DISCUSSION PAPER SERIES

No. 9740

TARGETING ULTRA-POOR HOUSEHOLDS IN HONDURAS AND PERU

Dean S. Karlan and Bram Thuysbaert

DEVELOPMENT ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9740.php**

ISSN 0265-8003

TARGETING ULTRA-POOR HOUSEHOLDS IN HONDURAS AND PERU

Dean S. Karlan, Yale University, Innovations for Poverty Action, MIT Jameel Poverty Action Lab, and CEPR
Bram Thuysbaert, Yale University, Ghent University and Innovations for Poverty Action

Discussion Paper No. 9740
November 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9740

November 2013

ABSTRACT

Targeting ultra-poor households in Honduras and Peru*

For policy purposes, it is important to understand the relative efficacy of various methods to target the poor. Recently, participatory methods have received particular attention. We ex-amine the effectiveness of a hybrid two-step process that combines a participatory wealth ranking and a verification household survey, relative to two proxy means tests (the Progress out of Poverty Index and a housing index), in Honduras and Peru. The methods we examine perform similarly to one another by various metrics. They all target most accurately in the cases of the poorest and the wealthiest households but perform with mixed results among households in the middle of the distribution. Ultimately, given similar performance, the analysis suggests that costs should be the driving consideration in choosing across methods.

JEL Classification: C81, O12 and O20
Keywords: participatory wealth rankings, poverty targeting and proxy means tests

Dean S. Karlan
Yale University
P.O. Box 208269
New Haven, CT 06520-8269
USA

Email: dean.karlan@yale.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=164317

Bram Thuysbaert
Yale University
P.O. Box 208269
New Haven, CT 06520-8269
USA

Email: bram.thuysbaert@yale.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=176830

*We thank Ford Foundation for funding; PLAN International, PLAN Honduras, ODEF Social, PLAN-Peru and Asociación Arariwa for collaboration on the underlying implementation; and, CGAP for collaboration on the overall initiative. Authors retained full intellectual freedom throughout the study. We thank Jennifer Severski, Adam Kemmis Betty, Maria Dolores Sanchez Liste, Juan Manuel Hernández-Agramonte, Edgar Salgado, Yann Guy, Andrew Hillis, Maria Dieci and Nathanael Goldberg, for project management, field support and data analysis support. All errors and opinions are those of the authors.

Submitted 07 November 2013

1 Introduction

Effectively identifying the appropriate recipients for aid programs is critical in order to maximize social impact with scarce resources. The erroneous inclusion of a household that is not part of the target population generally means resources wasted. Yet effective targeting is not costless. In theory, the economics are straightforward: screen such that the marginal cost of screening out the marginal ineligible participant is equal to the wasted resources transferred as a result of mistargeting.

Targeting poor households is difficult because the criteria for eligibility may be hard both to define and to verify. As there is no single defining characteristic of poverty, criteria for eligibility tend to be multidimensional and subject to much debate. Poverty lines based on per capita income or expenditure are often used, but it is also well recognized that they have limitations and represent a simplification of what it means to be poor (Ravallion 1998; Bebbington 1999; Alwang, Siegel, and Jorgensen 2001).

Once the relevant criteria are defined, verifying that certain households meet those criteria poses its own challenges. Measuring income for poor families, for example, is notoriously challenging: it derives mostly from informal sources and is often in kind rather than monetary (Deaton 1997). Verification may be further complicated when respondents, wishing to participate in the program, perceive an incentive to misreport information related to their eligibility. Such challenges may create a trade-off between accuracy and cost in identifying eligible households.

The challenges of verification require cost-effective solutions to targeting particular households. Three methods have been broadly proposed as solutions (Coady, Grosh, and Hoddinott 2002). First is geographic targeting, which uses national or regional poverty maps to select eligible households by region. While typically less precise relative to other methods, geographic targeting may suffice as an inexpensive and quick method in certain circumstances. Even when other selection methods are used, geographical targeting is often applied as a first filter. A second method is a proxy means test (PMT), in which field workers collect demographic, asset or housing information that can be used to approximate a household's poverty status. Compared to measurements of income or consumption, the inputs required for a PMT are both quicker to collect and easier to verify. However, for any PMT there is a substantial risk of targeting error. Moreover, PMTs typically lack transparency, potentially leading to accusations of favoritism or incompetence, which could undermine the legitimacy of the program. A third method is selection by village members themselves. The criteria for selection can range from nomination by local leaders to ranking through a Participatory Wealth Ranking (PWR). A PWR invites

village members to rank members of their community according to poverty levels. The poorest members, typically, are then eligible for the program. The increasing popularity of PWRs reflects a broader trend towards Participatory Rural Appraisals (PRAs) to collect information and design aid programs (Chambers 1994). Such participatory processes have the advantage of transparency and the incorporation of local knowledge, which is likely to be more precise than a PMT. On the other hand, there are a number of reasons why a PWR may not work well in practice. Local elites may manipulate the participatory process in order to include themselves, their family or members of a particular group. Moreover, local definitions of poverty may differ from the criteria of the program implementer.

The relative costs and merits of each method hold important implications for organizations seeking to target the poor effectively. PWRs in particular have received academic attention. Quantitative evaluations of PWR results using household surveys generally show that PWRs effectively identify households that are poor according to traditional measures of wealth. Adams et al. (1997) use expenditure, income and asset holdings to validate a PWR in rural Bangladesh, and find significant differences between wealth groups across all traditional socio-economic variables. Ojiako et al. (2009) and Temu and Due (2000) similarly find that PWRs in Nigeria and Tanzania, respectively, successfully identify poorer households. Van Campenhout (2006) studies a wealth ranking in rural Tanzania and finds that the wealth categories reflect asset holdings and schooling levels.[1]

While the results of PWRs are intrinsically interesting in terms of what we can learn about perceptions of poverty, their performance relative to other targeting methods is also directly relevant for policy. The relatively few studies that have made such comparisons have yielded mixed results. In a study in rural South Africa, Hargreaves et al. (2007) compare the results of a PWR with two survey-based methodologies that employ principal component analyses (PCA) to construct wealth indices. The PWR results are only weakly correlated with the survey-based tools, implying that one or the other (or both) is incorrect; however, in the absence of a credible benchmark it is impossible to determine which is more effective. Banerjee et al. (2007) evaluate a two-step targeting process used by Bandhan in India to establish eligibility for the same Graduation Program that is discussed in this paper. Using detailed household survey data to analyze the process, they find that the PWR approach compares favorably to the census-based methods used by the Indian government, although it is important to note that the latter fared

[1] Data were not available to compare the results with household income or expenditure.

particularly poorly. Most recently, Alatas et al. (2010) conducted a field experiment in Indonesia to compare a PMT and PWR, finding that while the PMT was more accurate when poverty was defined in terms of consumption, the villagers themselves were more satisfied with the results of the PWR (perhaps because this directly incorporates local conceptions of poverty, or because villagers are more satisfied if part of the process, regardless of the outcome).

We examine the effectiveness of a three-step "Targeting the Ultra-Poor" ("TUP") process relative to other methods in two different contexts, Honduras and Peru. The TUP method combines geographical targeting, PWRs and PMTs; we concentrate especially on steps two and three in this paper. The process was used to determine eligibility for the CGAP-Ford Foundation Graduation Program. In Gracias, Honduras, the program was implemented in 2008 by Plan International Honduras and ODEF Social. In Cusco, Peru, the program was implemented in 2010 by Plan International Peru and Asociación Arariwa. The programs aim to tackle extreme poverty by combining an asset transfer (livestock, e.g.) with training, cash transfers, and health services.

In both countries, the first step used geographical targeting. The intervention area was determined by the local organizations' area of operations and reference to regional poverty maps; villages were then selected using a simple scorecard. This paper does not evaluate the accuracy of this first step of identifying the broad geographical areas. The second and final steps, the focus of our analysis, was a PWR in the villages to determine the poorest households, and then a verification survey by the NGO that confirmed program eligibility and basic economic status questions. Our analysis uses data from a detailed household survey that was administered after the targeting process. The survey included the selected (i.e., identified as ultra-poor) households, as well as a random sample of non-selected ("excluded") households in the same village. In Honduras, 423 selected and 637 excluded households were surveyed in 15 randomly selected villages; In Peru, 470 selected and 537 excluded households were surveyed in 21 randomly selected villages.

Section 2 of the paper describes the program and the targeting process in both sites. The sample and data used for the subsequent analysis are described in Section 3. Section 4 compares the selected and excluded households at each stage in the targeting process across a range of variables in each country. Section 5 compares the accuracy of the TUP targeting process to two simple PMTs – a housing index and the Progress out of Poverty Index (PPI) – using per capita consumption as a benchmark of true poverty. Section 6 compares the performance of the TUP targeting process, the housing index and the PPI using other metrics of poverty such as assets and vulnerability. Section 7, examines in a regression framework characte-

ristics that predict a household's ranking in the PWR in a regression framework. Finally, section 8 summarizes our results and discusses the policy implications.

2 The targeting process

2.1 The CGAP-Ford Foundation Graduation Program in Honduras and Peru

The Graduation Program aims to help the poorest families "graduate" from extreme poverty within a 24-month period. The program incorporates five core elements: targeting of the poorest families, direct consumption support in the short term, access to savings services, one-to-one skills training and the transfer of a productive asset. The Graduation Program is being replicated in 10 pilot sites in Asia, Africa and Latin America (Syed and Montesquiou 2011). In two of these sites, Honduras and Peru, we built in additional data collection from the entire village, not just the identified participants, in order to analyze the targeting process.

The Honduras program is operated by Plan International Honduras and a local microfinance institution, ODEF Social, in the northern districts of Lempira department. Lempira was chosen as the intervention zone because of the high incidence of extreme poverty and because Plan already had projects in the area. The Peru program is operated by Plan International Peru and a Cusco-based microfinance institution, Asociación Arariwa, in two southern provinces of the Cusco department, Canas and Acomayo. As in Honduras, the intervention zones were selected according to poverty indicators and the existing presence of the local organizations.

In both sites, the program implementers first selected the poorest 80 villages in the intervention zone.[2] In Honduras, the Plan team already had considerable experience working in the area and selected the villages based on their prior observations and perceptions. In Peru, where prior knowledge of the project zone was more limited, the Plan team selected villages based on a simple scorecard that assessed access to basic services like roads, electricity, water, education and healthcare.

Within selected villages, households were selected using a two-step process. The first step was a PWR to which all village members were invited, which is described in Section 2.2. Households selected in this step progressed to the second step, a short verification survey applied by Plan field workers that was used to confirm eligibility according to criteria described in Section

2 In Peru, an additional six villages were later added, bringing the total to 86.

2.3. For program impact evaluation, the villages were then randomly assigned to treatment and control groups; within treatment villages selected households were also randomly assigned to treatment and control groups.

2.2 The participatory wealth ranking

All village members were invited to attend the PWR, which followed the geographical selection of eligible villages. In Honduras, the invitation was sent through local schools. The students informed their parents of the meeting that would take place in the school the following day. In Peru, Plan field workers went to each village a month prior to the meeting to set a date and invite participants. In both countries, field workers stressed the importance of a high level of participation and the attendance of women as well as men. In Peru, field workers would go ahead with meeting if more than 50% of households were represented, which, in most of the villages of the zone, is the established threshold at which communal decisions can be taken; in Honduras, the threshold was lower. Monitoring visits conducted by our research assistant at five of the meetings in Peru suggest that male participants typically outnumbered female participants by about three to one, although in other cases only women participated due to a misunderstanding of the meeting's aims. Comments from participants indicated that those who lived furthest away, typically the poorest households, were underrepresented at the meetings.

Each village PWR meeting was run by three field workers in a common area of the village, and lasted between two and three hours. In Honduras the meetings were conducted in Spanish, while in Peru most meetings were conducted in a mix of Quechua and Spanish, with some field workers speaking more in Quechua and others more in Spanish. The monitoring visits in Peru suggested that the level of participation, particularly among women, was higher when the meeting was predominately conducted in Quechua. After a brief introduction to the implementing organizations and the targeting process, participants prepared a sketch map of the village.[3] The map included landmarks such as roads, rivers, the school, village hall and different neighborhoods. In Honduras the facilitator attempted to engage all members, but in practice only two or three people tended to participate in the production of the map, while in Peru a group of four people was selected to work on the map whilst the other village members began

[3] In Honduras, no mention was made of the Graduation Program in order not to bias participant behavior in the PWR. In Peru, participants were given information only about the training elements of the project. The cash consumption stipend and asset transfer were not mentioned.

the wealth ranking. In parallel, other village members assisted a field worker in preparing index cards with the names of all the household heads in the village.

In Honduras, the ranking process began with a comparison between two families. For the first two households, the field worker read out the name of the household heads and then asked the village members if the two households lived in the same conditions. If so, the two index cards were placed in the same pile. If one lived better than the other, they were placed in separate piles. The field worker then picked up a third index card and asked if this household lived in similar, better, or worse conditions than the first two households. The process continued until all of the households were classified in piles. The number of categories varied between villages, depending on the responses of the village members. In homogenous villages, it was possible that the majority of households were categorized in the same group, although in such situations the field worker tried to encourage participants to identify subtle differences. The criteria used to distinguish between categories were implicit, rather than formally defined.

In Peru, on the other hand, the ranking process began with the definition of the wealth categories. The field worker proposed to the village members that in every village there are "families that have the most," "families that have neither a lot nor little," "families that have little," and "families that have the least." In many cases, the initial reaction of the meeting participants was to argue that everyone in the village is equally poor. However, with the use of examples, the field worker was able to demonstrate to the participants that although all may be poor, some are poorer than others.

Next, the village defined four wealth categories in terms of land, animal ownership and the house characteristics. For example, "families that have the most" might be defined as those that have more than 50 sheep, eight cows or 10 llamas; more than three *masas* of land (3,150m^2); and a house with four or more rooms. The "families that have the least" might have fewer than 14 sheep, two cows or three llamas; 0.5 *masas* (275m^2) of land; and a one room house. Anecdotal evidence suggests that this rather abstract exercise was difficult for many participants, and significant guidance was required from the case worker to produce a logical classification.

Index cards with the name of the household heads were then read out in random order. The location of the household was drawn on the village map and the participants decided in which category the household belonged. The index card was then placed in a cardboard box corresponding to that category. This process created a number of challenges. First, it was a time consuming process, particularly in large villages, and participants evidently tired towards the

end. Second, there was no established process for handling disputed cases, where some village members felt that the household should be in one category and others felt otherwise. Given the time constraint, the field worker needed to make a quick decision, and would typically go with the option that was being voiced most loudly, or appoint the village president to act as arbitrator. In general, there was little reference to the objective criteria established by the participants and it was unclear whether the classification of households reflected these criteria or not.

With the ranking complete, the next step at both sites was to determine the PWR categories that would be eligible for inclusion into the program. In practice, the norm was to select the poorest two categories from each village, which was normally over half of the households in the village. In poorer villages, the three poorest categories were selected.[4]

2.3 The verification step by NGO field workers

The next step of the TUP targeting process was a verification survey for households selected in the PWR. This was conducted by the NGO field workers at each house or in a community meeting. During the survey process, case workers would sometimes encounter additional households that had either not been ranked in the PWR or who claimed they had been ranked incorrectly, and would include these households in the survey. Not all households selected in the PWR were surveyed: some had migrated from the community, others were not at home, and still others did not meet the inclusion criteria defined for the project.

The first aim of the verification survey was to verify the suitability of the household for the project. We label the criteria used for this purpose the "programmatic" criteria. In Honduras the criteria applied were: (1) the household includes a child under the age of 18 to meet Plan's mission of helping children and (2) the household has lived in the village for at least three years. In Peru the criteria applied were: (1) the head of the household or their spouse is younger than 60, and would therefore be capable of managing an enterprise for several years to come; (2) the household includes a child under the age of 18 and (3) the household head doesn't live outside the community for more than six months of the year.

The second aim of the verification survey was to confirm that the household was indeed poor, in order to correct for errors or manipulation during the PWR: we label these "poverty"

[4] In a few villages in Honduras, four categories were selected.

criteria. In Honduras, this took the form of two criteria: (1) the household has a monthly per capita income of 600 Lempira or less, the monthly cost of a basic food basket and (2) the household meets at least two of the following three criteria: (a) having one manzana or less of land under cultivation[5]; (b) having minors in the household who work in income-generating or productive activities; (c) not currently participating in a development program. In Peru, four criteria were used, including a PMT: (1) neither the household head nor the spouse have a formal profession or occupation; (2) the household head does not own a second home outside of the community; (3) the household does not currently borrow money from formal sources[6]; (4) the household has a PPI score of 30 or less. The PPI was chosen as a PMT method because it was well-tested in Peru and simple to apply and calculate. With a PPI score of 30 or more, there is a 50% probability that the household is not below the national poverty line (Schreiner, 2009).

3 Sample and data

After the targeting steps were complete, we randomly selected 15 villages (out of 40 treatment villages) from Honduras, and 21 villages (out of 40 treatment villages) from Peru to be included in the targeting analysis study. Within each of these communities, an extensive socio-economic survey was administered to the selected households as part of an impact evaluation study. In addition, for the purpose of this targeting analysis, we also surveyed a random sample of the excluded households. Since a higher proportion of households selected for the program were sampled than those excluded for the program, we use sampling weights throughout the analysis to make the sample representative at the village level. In the 15 selected communities in Honduras, a total of 1,060 households were surveyed – 423 selected households and 637 non-selected households, whereas in the 21 villages in Peru 470 selected and 537 non-selected households were surveyed, for a total sample of 1,007 households.

Two filters were applied to define the sample frame for analysis. First, as mentioned above, several "programmatic" criteria were applied in the verification step (e.g. presence of a child under 18 in the household). These programmatic criteria reflect the priors of the implementing organizations about which types of households are suitable for the intervention and do not

[5] One manzana equals about 1.7 acres or 0.7 hectare.

[6] Households that had a loan with the microfinance institution Caja Nuestra Gente, which had recently entered into an agreement with the government to provide credit to beneficiaries of the conditional cash transfer scheme, Programa JUNTOS, were not excluded.

necessarily relate to poverty. To focus on how well the targeting methods select the poor, we therefore remove from the sample all households that do not satisfy the programmatic criteria.[7]

Second, some households identified as poor in the PWR were not surveyed later in the verification step. In some cases, there was no respondent available when the NGO field workers returned to the community. In other cases, the household could not be located in the community. For these households, we do not know whether, had they been surveyed, they would have been selected or not. Since we are interested in understanding how each of the steps contributed to the final selection, we discard from the sample the households that were identified as poor in the PWR but were not verified by the NGO in the verification survey. After applying these two filters, we are left with 897 households in the analysis sample in Honduras and 717 households in Peru (Table 1 for Honduras and Table 2 for Peru).

Of the 897 households in the Honduras sample, 702 were categorized as ultra-poor at the PWR stage. This corresponds to 62% of households (using sampling weights). Of these, relatively few failed to pass the verification step by the NGO (67 households or 17%).[8] Overall, the two-step TUP targeting process selected 52% of households for the program (conditional on meeting the programmatic criteria). In Peru, 64% of households were identified as ultra-poor at the PWR step, and of these, only 14% were excluded at the verification step. In all, 59% of households in Peru were finally identified as ultra-poor and selected for inclusion into the program. The final selection rates in Honduras and Peru – 52% and 59% respectively – may seem high, but note that the program villages were purposefully selected because of their high incidence of extreme poverty.

4 Selection at each stage of the TUP process

We begin our analysis by examining differences between how households are categorized at each step of the TUP targeting process. Table 3 (Honduras) and Table 4 (Peru) show the means of several welfare indicators for each group at a given step. We compare the groups in terms of demographics and education (Panel A), household assets (Panel B), productive assets

[7] Since we do not have data from the verification survey for the households excluded in the PWR step, we check programmatic eligibility using our household survey data.

[8] From Table 1, (62%-52%) / 62% = 17%.

and income (Panel C), and consumption, poverty, and vulnerability (Panel D). The number of households in each step is displayed in the last row of the table. The total sample sizes in the "PWR step" and "Final selection" are the same, while the sample size in the "Verification step" is the same as the number of households selected in the PWR.

4.1 Participatory Wealth Ranking

In both countries, households selected by the PWR are consistently poorer across a range of welfare indicators than households excluded in the PWR. In columns 1 – 3 of Tables 3 and 4, we see that all statistically significant differences between those excluded and selected by the PWR (N=564 for Peru, N=702 for Honduras) reflect sorting into groups by higher and lower welfare, particularly in Honduras where all indicators except whether or not the household is female-headed and sheep and goat ownership show a statistically significant difference. In both countries, indicators of education, household assets, productive assets, and consumption, poverty, and vulnerability reflect consistent sorting into groups by welfare status. To get a sense of the magnitudes of the differences in terms of the underlying variance, we standardize the difference by the overall standard deviation of the variable of interest (column 4). The largest standardized differences occur in Honduras for the housing index (0.96), the household asset index (0.85) and the years of education of the household head (0.51) and in Peru for the number of cattle (0.89), the household asset index (0.65) and the cultivated land area (0.55).

More statistically significant differences between selected and excluded households emerge in Honduras than in Peru. Notably, the PWR in Peru did not sort households by weekly income per capita or food security, although the lack of a significant difference for the former may stem from the challenge of measuring income. Reported income for both groups is much lower than consumption measures, reflecting this challenge.

4.2 Verification

Selection at the verification step (Columns 5-8 of Tables 3 and 4) shows fewer statistically significant differences relative to the PWR. Note however that the number of excluded households in this step is small, reducing our power to detect statistically significant differences. In Honduras, differences emerge in boys' enrollment, assets (the asset index, latrine ownership, the housing index, and land ownership) and food security, each reflecting sorting consistent with poverty status. In Peru, we find statistically significant differences in the expected direction only for girls' school enrollment and the PPI. The latter is not unexpected given that the PPI score was itself one of the criteria used in the verification step. Excluded households

have fewer sheep and goats than the selected households. In Peru, sheep and goats are livestock typical of poorer households, with the richer households holding cattle, lamas and alpacas.

Given the smaller (in standardized units) and fewer differences within the verification step relative to the PWR step, it appears that the verification step mostly served to identify and correctly exclude a few wealthier households, while the PWR effectively sorted households broadly into poor and wealthy categories.

4.3 Final selection

Columns 9-12 of Tables 3 and 4 show the results of the complete TUP targeting process, comparing households selected in both the PWR and verification steps with those excluded in either one. In both countries, the significant differences in the final selection echo those in the PWR. Taken as a whole, the selection process effectively targeted poorer households according to a wide range of indicators. As can be seen in columns 13-14, the PWR was responsible for the majority of the differences between the two finally selected and excluded groups in both Honduras and Peru.[9]

5 TUP targeting in comparison to other targeting methods

5.1 Consumption per capita as a benchmark

The mean comparisons examined above indicate that the TUP selection process broadly sorted households by welfare status. Such sorting is an important, but not sufficient indication of how well the process identifies poor households: the process may have erred more towards false positives, or false negatives, which may have important welfare consequences, and the process may have worked well but not as well as other methods (or worked equally as well, but cost more). We use a well-established measure of poverty, consumption per capita, as a benchmark to compare the TUP targeting process with random selection and two PMTs: the PPI and a housing index.

[9] The numbers in columns 13-14 indicate the relative contribution of the PWR and verification steps. Let M be the mean of the variable of interest, M_1 its mean conditional on being selected in the PWR step and M_2 the mean conditional on being finally selected, the contribution of the PWR is then defined as $(M-M_1)/(M-M_2)$ and the contribution of the verification step as $(M_1-M_2)/(M-M_2)$.

We choose consumption per capita as our benchmark assuming that it represents the best available proxy for wellbeing. Caution is needed however, for several reasons. First, the measurement of consumption for poor households is inherently difficult. Respondents do not always remember their expenditures accurately, do not tend to measure the consumption of their own produce in standardized units, and may perceive an incentive to inflate or deflate their reported expenditure. Second, survey-based measurement of consumption usually refers to a short time period; for some households, those time periods may not be representative of their typical consumption habits. Third, households have different consumption preferences: a household that chooses to spend as little as possible and save for the future may appear poorer than it actually is when consumption is used as the benchmark. Fourth and finally, consumption does not capture other dimensions of poverty, such as vulnerability to shocks or social and political inclusion. In sum, while we consider consumption to be the best benchmark available to compare targeting tools, we interpret the results with caution.

5.2 Other targeting methods

We compare the TUP selection process with two proxy means tests: the PPI and a housing index. The PPI is a poverty scorecard that estimates the likelihood that a household is poor based on ten questions related to demographics, education, housing and assets. The information takes about five minutes to collect and many of the answers are readily verifiable if the questionnaire is performed in the home. Answers to each question correspond to a certain number of points; the sum of these points yields a score out of 100 for the overall survey. Each score is then associated via a scorecard with a probability that the household falls below the poverty line. The scorecard is calibrated using data from the relevant country's national household survey. The choice and weighting of indicators is based on their correlation with poverty, the ease of collecting and verifying the information, and the liability of the indicator to change over time as poverty status changes (Schreiner 2010; Schreiner 2009)

We constructed a housing index using principal components analysis (PCA), a statistical technique often used in the creation of socio-economic status indices from household survey data. Five variables were entered into the PCA: the total number of rooms in the house and dummy variables indicating whether the house has a cement floor, a cement wall, a latrine and electricity access. From this set of correlated variables, PCA creates uncorrelated components that explain the variance in the data and thus provide synthesized information on the underlying concept—in this case housing quality. The components are ordered so that the first component explains the largest amount of variation in the data (Vyas & Kumaranayake, 2006). This

first principal component is then used as a relative index of housing quality, which we use as a proxy for overall wellbeing.

5.3 Results

Our approach evaluates how well the various targeting methods categorize households at various deciles in the distribution of consumption per capita. We compare the TUP targeting method against a naively random selection, the PPI, and the housing index. For the PPI and housing index, we choose the poverty line – i.e. the cut-off value to be categorized as poor – in each country so that the X% of households gets selected where X is the percentage of households selected by the actual TUP targeting process (52% for Honduras and 59% for Peru). For each decile of the consumption distribution, we then calculate the fraction of households in that decile that would be selected by a given targeting method. A perfect targeting tool would select all households in percentiles of consumption per capita less than X% and not select any households in percentiles above X%. A naively random sampling would achieve a rate of X% for each decile.

Table 5 (Honduras) and Table 6 (Peru) show the results of this comparison. Columns 1, 2, 3, 5, and 8 show the fraction of households within a particular decile of consumption per capita that would be selected by the given targeting tool. As column 2 shows, perfect targeting would select 100% of respondents in all percentiles before the cutoff point for selection (52% in Honduras and 59% in Peru), and 0% afterwards. The other columns present the p-values from tests of equality of proportions selected between two of the targeting tools. With some exceptions— particularly in Honduras — both tables show few consistent differences in performance among the targeting tools and between each targeting tool and a mere random selection.

In Honduras, there is some evidence both for differences between each targeting tool and a random sample and among targeting tools. Each targeting tool significantly outperforms the random sample in at least three deciles, concentrated in one or both tails; the PPI performs the best, with favorable significant differences relative to random sampling in six of the 10 deciles (mostly in the tails: deciles 1, 3, 4, 8, 9, 10). Among targeting tools, no tool significantly outperforms the other in more than two deciles. Here too, significant differences generally emerge in the tails—all four differences are in the bottom or top two deciles. Of the three tools, the TUP process fares the worst, with PPI and housing performing at a similar level.

In Peru, the only test that shows more than one difference from mere random selection is the PPI, but of the four significant differences, only two correspond to favorable performance (households in the 1st and 10th deciles are more and less likely to be selected, while households

in the 4th and 5th decile are less likely to be selected). Of the thirty tests (3 methods x 10 deciles) performed that compare targeting tools to each other, only one shows a significant difference.

Figures 1 and 2 visually depict the performance of each targeting tool using consumption per capita as a benchmark.[10] An ideal targeting tool would have a straight line from the point (0,1) to (X,1), followed by a straight line from (X,0) to (0,1) (where X corresponds to the percentage of households finally selected in each country). Even though each targeting tool displays a downward sloping trend, none comes close to the ideal values. In both graphs, there is evidence that the tools perform relatively well at the tails, with the sharpest changes in slope occurring at the very left and right ends of the graphs. The shape implies that the targeting tools are better at identifying the very poorest and the very richest than they are at correctly categorizing individuals in the middle. In Honduras, the tools perform particularly well in the tails, as both Table 6 and Figure 1 demonstrate.

In both graphs, no tool consistently outperforms the others. The PPI in Peru, for instance, outperforms the TUP process and housing index in including the very poorest, but in the rest of the distribution has mixed performance. The housing index in Honduras also starts out well beginning in the left tail and through the ninth decile rivals the other tools, but then rises for the last decile, including a higher proportion of wealthy households.

Taken together, the two figures paint a mixed picture: while the targeting tools each perform slightly better than random selection—particularly in the tails and in Honduras—their relative performance shows few consistent patterns.

6 Other definitions of poverty

For the reasons outlined in section 5.1, we present our results using consumption per capita as "the" measure of poverty with some caution. Moreover, even if consumption per capita is the best proxy for poverty, given measurement error, other measures may shed light on the relative performance of targeting tools. Tables 7 and 8 thus extend the comparison of TUP with other targeting tools by examining each tool's performance using additional benchmarks for poverty. Each cell shows among households selected by the targeting method in the column heading the percentage of households that rank in the bottom X% (52% in Honduras and 59% in

[10] The graphs depict fractions of non-parametrically estimated density functions. Because of the non-parametric smoothing, they do no perfectly map into the fractions in Tables 5 and 6.

Peru) according to the poverty metric on the left. As before, we compare the targeting tools to random selection, which would select X% of households in the bottom X% of the poverty distribution for each metric. A perfect targeting tool would only select those bottom X%; hence, the number selected for each metric is 100%.

In the first row of the tables we show how each tool fares on average for consumption per capita. The results mirror the trends from Tables 5 and 6: among random selection, TUP, the housing index, and PPI, there are no significant differences in Peru, while all targeting tools perform better than random selection in Honduras, but no significant differences exist among them.

Clearer differences emerge among the tools when the asset index is used as a benchmark. In Honduras, each tool outperforms the random sample on average (all three tests have p-values <0.01), and TUP and PPI both fare better than the housing index (66% vs. 58%, p-value<0.01 and 67% vs. 58%, p-value<0.01, respectively). The same trend exists in Peru (71% vs. 64%, p-value=0.02 and 72% vs. 64%, p-value=0.02, respectively), except that the housing index underperforms relative to random selection.

A similar trend exists among the tools and in each country when using the total value of animals and total cultivated land as benchmarks. In both countries, the relative ranking for each measure for the three tools is the same. For Honduras and Peru, respectively, the percent correctly identified as poor, using total value of animals, is TUP (66% and 70%) > PPI (62% and 61%) > Housing (56% and 58%) > random (52% and 59%). Similarly, for total cultivated land as the "true" measure of poverty, the success rates for Honduras and Peru, respectively, are TUP (61% and 66%) > PPI (58% and 63%) > Housing (53% and 59%) > random (52% and 59%). These results suggest that for assets, total value of livestock, and total value of cultivated land, TUP generally outperforms random selection and the housing index, and is slightly better than the PPI.

Few consistent differences among the tools appear when vulnerability to reductions in food consumption[11] and years of education of the household head are each used as the true measure of poverty. In Honduras, each of the tools is significantly better than random selection for both metrics and the TUP process and the PPI generally outperform the housing index. In Peru,

[11] As a measure of vulnerability to reductions in food consumption, we use an index based on questions in the survey about whether adults in the household reduced or skipped meals, did not eat for an entire day or whether meals were reduced for children in the past 12 months, and the frequency of these events.

no tool does better than random selection for the two metrics, except for the housing index when the education metric is used.

Overall, the comparison unveils two insights into the TUP selection process. First, when judged using six different poverty metrics, the TUP process typically performs better than random selection. Second, the TUP process, compared to PPI and the Housing index, leads to selecting households with fewer assets, less land, and less valuable livestock.

Figures 3 and 4 capture TUP's performance visually. The graphs plot the percentage of households selected by the TUP targeting process against the rank of those households according to three metrics of poverty: the asset index, consumption, and vulnerability to reductions in food consumption. TUP performs best according to the asset index in Peru and Honduras, showing a consistent negative slope from including poor households on the left to excluding rich households on the right. A similar, but weaker, trend exists for consumption. The TUP process shows an inverse relationship between vulnerability and selection in Peru and a direct, but weak relationship in Honduras.

7 Understanding PWR rankings

TUP's favorable performance along some poverty indicators, but not others, makes poignant the question of what observable information predicts how village members categorize households. Tables 9 and 10 analyze these criteria by regressing a household's group number (ranking) in the PWR on a host of covariates. Column 1 shows the results of this regression, while column 2 displays results from the same regression with different outcome variables selection by the complete TUP process as the outcome variable.

With the six poverty metrics and other covariates included in one regression, the PWR process shows trends consistent with earlier results: households with a lower score on the asset index in Honduras and Peru and households with less livestock and land in Peru are more likely to be ranked as poor, while there is no difference in the rankings among households according to their vulnerability to reductions in food consumption. In both countries, the education level of the household head is a statistically significant explanatory variable. Several other covariates are statistically significant, especially in Peru, indicating that, conditional on the poverty metrics, villagers take into account other characteristics when ranking households. In Honduras household size and having savings make a household less likely to be categorized as poor, whereas the household head being a widower has the opposite effect. In Peru, household size, and the household head being over 60 years old make a household less like likely to be

categorized as poor in the PWR, while households with heads under 30 are more likely to be categorized as poor.

Interestingly, in Peru, the PWR ranking favors households that have received support from NGOs or from the government in the past year. Assuming these NGOs are effectively targeting poor households, this result may not come as a surprise. On the other hand, perhaps households that receive support from NGOs are thereafter branded as poor and hence more likely to be ranked as such by other village members, irrespective of their wealth. Or some households may simply be more successful in garnering support from outside organizations, which could filter through in the PWR process. In line with the latter interpretation, we find in Peru that households that have higher degrees of social integration are more likely to be ranked as poor. As measures of social integration, the Peru regressions include a community participation index[12], the number of potential lenders someone could rely on in times of need and the number of times someone attended community meetings.[13] Both the community participation index and the number of potential lenders in the community are positively correlated with the likelihood of being categorized as poor. Though not definitive evidence, this is suggestive of elite capture.

8 Policy implications and conclusion

Data from household surveys of selected and excluded households in the targeting process for the Graduation Program support the effectiveness of PWRs in sorting households by poverty status. The subsequent step in the TUP targeting process—the verification step—produced fewer and smaller differences between selected and excluded households, but seems to have filtered some wealthy households out of the group selected in the PWR step. Looking beyond average differences to benchmark the TUP method against the distribution of consumption per capita shows that the TUP process fares poorly, but is mostly indistinguishable from two alternative targeting methods based on proxy means testing. Differences that do emerge surface within the tails of the consumption distribution.

This raises the question of why, given that households in these close-knit communities are likely to know each other very well, the PWR is not more accurate. One explanation finds sup-

[12] The community participation index is an index of participation in a range of community-based organizations.

[13] Unfortunately, these variables are not available for the Honduras sample.

port in our analysis that uses alternative poverty metrics: when compared against asset, land, and livestock ownership, both the TUP process and PPI method outperform random selection and the housing index. This demonstrates that local definitions of poverty incorporate variables other than consumption. Results from a multivariate regression controlling for consumption also reinforce this point, as asset ownership, livestock ownership, and education predict the rank in the PWR in both countries, after controlling for consumption. Moreover, conditional on poverty metrics, the PWR ranking is correlated with household head traits, prior support from NGOs and government and community participation. Noticeably absent from the list of predictors of the PWR ranking is vulnerability to reductions in food consumption.

Ultimately, the decision on what mechanism to use to target should be driven by a cost-benefit analysis. Table 11 provides an analysis of the estimated cost of each of the methods, in both countries. The TUP targeting process (both the participatory wealth ranking and the verification step) costs about US$7 per selected household, whereas the PPI or housing index would cost about US$5.5 per selected household. Thus the approaches have quite similar costs, but the TUP targeting process is the most expense. That cost is divided almost equally, half coming from the PWR and half from the verification step.[14] Thus if the verification step was deemed unnecessary (and our evidence suggests it contributed little to the poverty targeting), the PWR method would be substantially cheaper than the PPI or housing index methods which require household visits.[15]

Of course, the relative costs of the different targeting methods are a function of context-specific parameters such as the average number of households in a community and the percentage of targeted households. In Figures 5 and 6, we plot the (hypothetical) targeting cost per selected household as a function of these parameters for the Peruvian case – fixing the other parameters at their observed values. As seen in Figure 5, the PWR is substantially cheaper than the PPI/housing index independent of the percentage of households that is targeted (for an observed average village size of slightly under 100 households). Figure 6 shows however that

[14] The PPI/housing index and the verification step both make use of short household surveys. The cost of the verification survey is lower than the PPI/housing index survey because the verification survey includes only households selected in the PWR step.

[15] Our analysis focused on the poverty dimension of the targeting process and showed that the verification step contributed little to the poverty targeting. As mentioned above, the verification step also served a second objective, to verify a household's programmatic eligibility. If the verification step were deemed unnecessary, the screening of the programmatic criteria could be made part of the PWR process.

this relationship does not hold for small villages (less than 60 households) where the PPI/housing index is cheaper than the PWR. As village size increases, the PWR becomes cheaper relative to the PPI/housing index because the marginal cost of ranking one more household in the PWR is close to zero.

On the benefit side, the benefits of one approach versus another depend critically on the social welfare function one is maximizing, and implicitly from that, what the lost benefits are from resources "wasted" on delivering services to untargeted individuals. For instance, someone who is barely above the bar (thus not "ultra-poor" in a binary sense) still would benefit from the program and serve the greater social purpose of the program, just not as much as the person categorized as ultra-poor. In future work, we will examine the impact of the overall TUP program, and as part of that analysis, we will examine heterogeneity in impact with respect to baseline wealth.

References

Alatas, V., & al, e. (2010). Targeting the Poor: Evidence from a Field Experiment in Indonesia. *NBER working paper 15980* .

Banerjee, A., & al, e. (2007). Targeting Efficiency: How well can we identify the poor? *Institute for Financial Management and Research Centre for Micro Finance Working Paper Series No. 21* .

Banerjee, A., Duflo, E., Chattopadhyay, R., & Shapiro, J. (2007). Targeting efficiency: How well can we identify the poor? *Institute for Financial Management and Research Centre for Micro Finance, Working Paper Series No. 21* .

Chambers, R. (1994). Participatory Rural Appraisal (PRA): Analysis of Experience. *World Development , 22* (9), 1253-1268.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Baltimore: John Hopkins.

Perova, E., & Vakis, R. (2009). *Welfare impacts of the "Juntos" Program in Peru: Evidence from a non-experimental evaluation.* Washington, D.C.: World Bank.

Rabbani, Prakash, & Sulaiman. (2006). Impact Assessment of CFPR/TUP: A Descriptive Analysis Based on 2002-2005 Panel Data. *CFPR/TUP Working Paper Series No. 12* .

Ravallion, M. (1998). *Poverty Lines in Theory and Practice.* Washington, D.C.: World Bank.

Schreiner, M. (2010). *Progress out of Poverty Index. A Simple Poverty Scorecard for Honduras.* Grameen Foundation.

Schreiner, M. (2009). *Progress out of Poverty Index: A Simple Poverty Scorecard for Peru.* Grameen Foundation.

Skoufias, E., & Coady, D. (2002). Are the welfare lessons from imperfect targeting important? *FCND Discussion Paper No. 125* .

Skoufias, E., Davis, B., & de la Vega, S. (2001). Targeting the Poor in Mexico: An Evaluation of the Selection of Poor Households into PROGRESA. *World Development , 29* (10), 1769-1984.

Syed, M., & Montesquiou, A. (2011). *Reaching the Poorest: Lessons from the Graduation Model.* CGAP.

Thorp, R., & Paredes, M. (2010). *Ethnicity and the Persistence of Inequality.* Bastingstoke, England: Palgrave Macmillan.

Vyas, S., & Kumaranayake, L. (2006). Constructing socie-economic status indices: how to use principal components analysis. *Health Policay Plan , 21* (6), 459-468.

Table 1: Description of sample - Honduras

		Number of households	% of sample (using sampling weights)
1	Total households in sample	1060	100%
2	Excluded for not having child under 18 in house	85	11%
3	Excluded for having lived in the village for 3 years or less	20	2%
4	Households identified as poor in the PWR step and not interviewed in verification step	60	10%
5	Total households in analysis sample	897	77%
	PWR step		
6	Total households in analysis sample	897	100%
7	Considered non-poor in the PWR	195	38%
8	Considered poor in the PWR	702	62%
	Verification step		
9	Considered poor in the PWR	702	62%
10	Excluded for having monthly income per capita of more than L. 600	24	3%
11	Not meeting 2 or more of the criteria below (12-14)	28	4%
12	Doesn't meet "7000 m2 or less land under cultivation"	71	8%
13	Doesn't meet "minors working in income-generating or other productive activities"	372	34%
14	Doesn’t meet "not participating in a development program"	7	1%
15	Finally selected	635	52%

Table 2: Description of sample - Peru

		Number of households	% of sample (using sampling weights)
1	Total households in sample	1007	100%
2	Excluded b/c household head spends more than 6 months a year outside village	16	2%
3	Excluded b/c household head and spouse over 60	111	20%
4	Excluded for not having child under 18 in house	214	38%
5	Households identified as poor in the PWR step and not interviewed in verification step	39	5%
6	Total households in analysis sample	717	51%
	PWR step		
7	Total households in analysis sample	717	100%
8	Considered non-poor in the PWR	154	36%
9	Considered poor in the PWR	563	64%
	Verification step		
10	Considered poor in the PWR	563	73%
11	Excluded for having a second home	5	1%
12	Excluded for having a formal job	1	0%
13	Excluded for having a formal loan	9	2%
14	Excluded for PPI score > 30	16	3%
15	Finally selected	536	59%

Table 3: Selection at each stage of TUP targeting process - Honduras

	PWR step				Verification step				Final selection				Contributions to final selection	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Excluded	Selected	p-value (1)=(2)	Stand. diff. (1)-(2)	Excluded	Selected	p-value (5)=(6)	Stand. diff. (5)-(6)	Excluded	Selected	p-value (9)=(10)	Stand. diff. (9)-(10)	PWR step	Verif. Step
Panel A: Demographics and education														
Female-headed household	0.11	0.14		-0.10	0.15	0.14		0.03	0.12	0.14		-0.07	114%	-14%
	(0.31)	(0.35)			(0.36)	(0.35)			(0.32)	(0.35)				
Years of education of household head	1.28	0.84	***	0.51	0.92	0.83		0.11	1.20	0.83	***	0.44	91%	9%
	(1.02)	(0.66)			(0.78)	(0.64)			(0.98)	(0.64)				
School enrollment (boys: 12-17 years)	0.54	0.39	***	0.30	0.61	0.36	**	0.51	0.55	0.36	***	0.38	70%	30%
	(0.50)	(0.50)			(0.49)	(0.48)			(0.50)	(0.48)				
School enrollment (girls: 12-17 years)	0.59	0.44	**	0.30	0.44	0.43		0.02	0.55	0.43	**	0.24	97%	3%
	(0.49)	(0.49)			(0.50)	(0.50)			(0.50)	(0.50)				
Panel B: Household assets														
Household has radio	0.82	0.71	***	0.25	0.74	0.70		0.10	0.80	0.70	***	0.23	85%	15%
	(0.39)	(0.45)			(0.44)	(0.46)			(0.40)	(0.46)				
Household has bicylce	0.26	0.14	***	0.30	0.12	0.15		-0.06	0.23	0.15	***	0.21	109%	-9%
	(0.44)	(0.35)			(0.33)	(0.35)			(0.42)	(0.35)				
Asset index	1.90	-0.15	***	0.85	0.27	-0.23	**	0.21	1.56	-0.23	***	0.74	90%	10%
	(2.78)	(1.75)			(1.77)	(1.74)			(2.69)	(1.74)				
Latrine with water or septic tank	0.73	0.41	***	0.57	0.57	0.38	**	0.35	0.70	0.38	***	0.56	79%	21%
	(0.55)	(0.56)			(0.59)	(0.55)			(0.56)	(0.55)				
Housing index	1.12	-0.30	***	0.96	-0.01	-0.36	**	0.24	0.88	-0.36	***	0.83	90%	10%
	(1.56)	(1.14)			(1.20)	(1.11)			(1.56)	(1.11)				
Panel C: Productive assets and income														
Total cultivated land (m2)	9632	5521	***	0.40	6892	5244	*	0.16	9040	5244	***	0.37	85%	15%
	(12527)	(8347)			(6680)	(8622)			(11570)	(8622)				
Number of cattle	0.50	0.08	***	0.37	0.06	0.08		-0.02	0.40	0.08	***	0.28	102%	-2%
	(1.61)	(0.69)			(0.32)	(0.74)			(1.44)	(0.74)				
Number of sheep/goats	0.02	0.04		-0.09	0.08	0.03		0.20	0.03	0.03		0.00	-1650%	1750%
	(0.14)	(0.29)			(0.36)	(0.27)			(0.21)	(0.27)				
Total weekly income per capita (Lempiras)	74.98	50.59	***	0.30	46.74	51.35		-0.06	68.89	51.35	**	0.22	109%	-9%
	(98.65)	(66.09)			(55.79)	(67.96)			(91.83)	(67.96)				
Household has business	0.19	0.10	***	0.26	0.10	0.10		0.00	0.17	0.10	**	0.20	99%	1%
	(0.39)	(0.30)			(0.30)	(0.30)			(0.37)	(0.30)				
Panel D: Consumption, poverty and vulnerability														
Total weekly consumption per adult equivalent (Lempiras)	166.91	128.49	***	0.32	148.03	124.53		0.20	162.81	124.53	***	0.32	79%	21%
	(120.23)	(118.01)			(230.51)	(77.11)			(151.17)	(77.11)				
Household below $1.25 poverty line	0.33	0.47	***	-0.29	0.45	0.48		-0.06	0.36	0.48	***	-0.25	91%	9%
	(0.47)	(0.50)			(0.50)	(0.50)			(0.48)	(0.50)				
Total ppi score	36.37	32.08	***	0.40	33.64	31.77		0.17	35.78	31.77	***	0.37	84%	16%
	(11.67)	(9.86)			(10.30)	(9.75)			(11.44)	(9.75)				
Food security index	0.52	-0.02	***	0.35	0.39	-0.09	**	0.31	0.49	-0.09	***	0.38	73%	27%
	(1.29)	(1.62)			(1.34)	(1.66)			(1.30)	(1.66)				
Number of observations	195	702	.		67	635	.		262	635				

Statistical signficance denoted * = 10%, ** = 5%, *** = 1%. The p-values are based on robust estimates of the standard errors. Columns (4), (8) and (12) show the standardized difference between the means of the excluded and selected groups, i.e. the difference divided by the overall standard deviation of the variable. Columns (13) and (14) indicate the relative contribution of the PWR and verification steps. Let M¬ be the mean of the variable of interest, M¬1 its mean conditional on being selected in the PWR step and M2 the mean conditional on being finally selected, the contribution of the PWR is then defined as (M-M1)/(M-M2) and the contribution of the verification step as (M1-M2)/(M-M2).

Table 4: Selection at each stage of TUP targeting process - Peru

	PWR step				Verification step				Final selection				Contributions to final selection	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Excluded	Selected	p-value (1)=(2)	Stand. diff. (1)-(2)	Excluded	Selected	p-value (5)=(6)	Stand. diff. (5)-(6)	Excluded	Selected	p-value (9)=(10)	Stand. diff. (9)-(10)	PWR step	Verif. step
Panel A: Demographics and education														
Female headed household	0.06	0.13	*	-0.22	0.14	0.13		0.03	0.07	0.13	*	-0.18	103%	-3%
	(0.24)	(0.33)			(0.35)	(0.33)			(0.26)	(0.33)				
Years of education of household head	6.33	5.21	***	0.31	6.42	5.10		0.37	6.35	5.10	***	0.35	79%	21%
	(3.44)	(3.63)			(4.26)	(3.55)			(3.55)	(3.55)				
School enrollment (boys: 12-17 years)	0.98	0.95	*	0.17	0.94	0.95		-0.04	0.97	0.95		0.14	107%	-7%
	(0.14)	(0.14)			(0.24)	(0.22)			(0.16)	(0.22)				
School enrollment (girls: 12-17 years)	0.93	0.89		0.14	1.00	0.88	*	0.40	0.94	0.88		0.19	67%	33%
	(0.26)	(0.26)			-	(0.32)			(0.25)	(0.32)				
Panel B: Household assets														
Household has radio	0.98	0.95	*	0.13	0.91	0.95		-0.22	0.97	0.95		0.07	164%	-64%
	(0.15)	(0.22)			(0.29)	(0.21)			(0.18)	(0.21)				
Household has bicylce	0.68	0.48	***	0.41	0.42	0.48		-0.12	0.65	0.48	***	0.34	107%	-7%
	(0.47)	(0.50)			(0.50)	(0.50)			(0.48)	(0.50)				
Asset index	1.33	-0.15	***	0.65	0.64	-0.22		0.38	1.24	-0.22	***	0.64	88%	12%
	(1.73)	(2.37)			(2.54)	(2.35)			(1.87)	(2.35)				
Latrine with water or septic tank	0.22	0.25		-0.07	0.18	0.26		-0.19	0.22	0.26		-0.10	94%	6%
	(0.42)	(0.43)			(0.38)	(0.44)			(0.41)	(0.44)				
Housing index	0.40	0.00	***	0.34	0.11	-0.01		0.10	0.37	-0.01	***	0.32	94%	6%
	(1.00)	(1.22)			(0.98)	(1.24)			(1.00)	(1.24)				
Panel C: Productive assets and income														
Total cultivated land (m2)	8144	4134	***	0.55	4091	4138		-0.01	7626	4138	***	0.48	100%	0%
	(8730)	(5793)			(3335)	(5966)			(8348)	(5966)				
Number of cattle	4.59	2.39	***	0.89	2.92	2.34		0.23	4.38	2.34	***	0.82	94%	6%
	(2.85)	(1.81)			(2.14)	(1.77)			(2.82)	(1.77)				
Number of sheep/goats	13.31	11.30		0.13	6.15	11.75	***	-0.36	12.42	11.75		0.04	264%	-164%
	(17.23)	(14.37)			(10.95)	(14.55)			(16.74)	(14.55)				
Total weekly income per capita (Lempiras)	21.56	21.33		0.01	22.25	21.25		0.03	21.65	21.25		0.01	51%	49%
	(27.16)	(30.84)			(23.07)	(31.44)			(26.65)	(31.44)				
Household has business	0.28	0.14	***	0.37	0.23	0.13		0.26	0.28	0.13	***	0.38	86%	14%
	(0.45)	(0.34)			(0.42)	(0.34)			(0.45)	(0.34)				
Panel D: Consumption, poverty and vulnerability														
Total weekly consumption per adult equivalent (Soles)	43.87	36.67	**	0.31	31.40	37.10		-0.24	42.38	37.10	**	0.23	120%	-20%
	(25.63)	(21.56)			(18.20)	(21.76)			(25.17)	(21.76)				
Household below $1.25 poverty line	0.13	0.26	***	-0.32	0.35	0.25		0.25	0.16	0.25	*	-0.23	120%	-20%
	(0.34)	(0.44)			(0.48)	(0.43)			(0.36)	(0.43)				
Total ppi score	17.69	16.15	**	0.21	19.46	15.85	*	0.48	17.92	15.85	***	0.27	65%	35%
	(7.13)	(7.67)			(9.48)	(7.43)			(7.48)	(7.43)				
Food security index	-0.12	0.07		-0.12	0.01	0.07		-0.04	-0.11	0.07		-0.11	93%	7%
	(1.68)	(1.57)			(1.46)	(1.57)			(1.65)	(1.57)				
Number of observations	154	563			27	536			181	536				

Statistical signficance denoted * = 10%, ** = 5%, *** = 1%. The p-values are based on robust estimates of the standard errors. Columns (4), (8) and (12) show the standardized difference between the means of the excluded and selected groups, i.e. the difference divided by the overall standard deviation of the variable. Columns (13) and (14) indicate the relative contribution of the PWR and verification steps. Let M¬ be the mean of the variable of interest, M¬1 its mean conditional on being selected in the PWR step and M2 the mean conditional on being finally selected, the contribution of the PWR is then defined as (M-M1)/(M-M2) and the contribution of the verification step as (M1-M2)/(M-M2).

Table 5: Mistargeting of different selection methods by consumption decile - Honduras

Decile of per capita consumption distribution	(1) Random selection	(2) Perfect targeting	(3) Complete TUP targeting	(4) p-value (1)=(3)	(5) Housing Index targeting	(6) p-value (1)=(5)	(7) p-value (3)=(5)	(8) PPI targeting	(9) p-value (1)=(8)	(10) p-value (3)=(8)	(11) p-value (5)=(8)	(12) Number of obs.
1st decile	52%	100%	54%	0.64	69%	0.00	0.04	72%	0.00	0.02	0.59	95
2nd decile	52%	100%	57%	0.38	67%	0.01	0.17	54%	0.69	0.66	0.09	91
3rd decile	52%	100%	61%	0.12	64%	0.04	0.75	63%	0.05	0.86	0.92	95
4th decile	52%	100%	56%	0.53	55%	0.56	0.95	68%	0.00	0.16	0.08	88
5th decile	52%	100%	65%	0.03	55%	0.55	0.16	54%	0.63	0.14	0.92	96
6th decile	52%	20%	59%	0.20	54%	0.67	0.48	54%	0.62	0.53	0.93	93
7th decile	52%	0%	46%	0.35	52%	0.88	0.36	53%	0.87	0.43	0.99	84
8th decile	52%	0%	47%	0.42	43%	0.15	0.59	40%	0.06	0.41	0.73	82
9th decile	52%	0%	39%	0.04	27%	0.00	0.02	29%	0.00	0.14	0.83	77
10th decile	52%	0%	30%	0.00	29%	0.00	0.79	28%	0.00	0.64	0.82	70

This table compares the complete TUP selection process to alternative targeting procedures: a housing index and the PPI score. Households were ranked according to per capita consumption and for each decile of the consumption distribution, we calculate the % of households that were (or would have been) selected by each of the methods. Since the complete selection process identified the poorest 52% and not a complete ranking, the housing index and PPI methods mimick that selection, i.e. a household is selected if it ranks among the 52% lowest scoring households.

Table 6: Mistargeting of different selection methods by consumption decile - Peru

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Decile of per capita consumption distribution	Random selection	Perfect targeting	Complete TUP targeting	p-value (1)=(3)	Housing Index targeting	p-value (1)=(5)	p-value (3)=(5)	PPI targeting	p-value (1)=(8)	p-value (3)=(8)	p-value (5)=(8)	Number of obs.
1st decile	59%	100%	70%	0.37	72%	0.27	0.77	77%	0.10	0.29	0.31	53
2nd decile	59%	100%	62%	0.64	65%	0.33	0.77	68%	0.14	0.52	0.71	68
3rd decile	59%	100%	70%	0.22	55%	0.62	0.03	62%	0.71	0.17	0.26	80
4th decile	59%	100%	56%	0.71	58%	0.93	0.82	47%	0.08	0.22	0.19	66
5th decile	59%	100%	51%	0.38	60%	0.94	0.54	46%	0.10	0.49	0.28	67
6th decile	59%	100%	57%	0.85	62%	0.77	0.73	58%	0.91	0.95	0.83	63
7th decile	59%	40%	70%	0.11	60%	0.88	0.30	62%	0.62	0.37	0.81	76
8th decile	59%	0%	48%	0.20	54%	0.60	0.54	66%	0.36	0.21	0.37	64
9th decile	59%	0%	64%	0.47	63%	0.51	0.90	54%	0.44	0.33	0.22	72
10th decile	59%	0%	40%	0.02	41%	0.04	0.87	41%	0.07	0.91	0.99	57

This table compares the complete TUP selection process to alternative targeting procedures: a housing index and the PPI score. Households were ranked according to per capita consumption and for each decile of the consumption distribution, we calculate the % of households that were (or would have been) selected by each of the methods. Since the complete selection process identified the poorest 52% and not a complete ranking, the housing index and PPI methods mimick that selection, i.e. a household is selected if it ranks among the 52% lowest scoring households.

Table 7: Poverty rates among targeted households using different poverty metrics - Honduras

Poverty metric	(1) Random selection	(2) Perfect targeting	(3) Complete TUP targeting	(4) p-value (1)=(3)	(5) Housing Index targeting	(6) p-value (1)=(5)	(7) p-value (3)=(5)	(8) PPI targeting	(9) p-value (1)=(8)	(10) p-value (3)=(8)	(11) p-value (5)=(8)
Total weekly consumption per adult equivalent	52%	100%	58%	0.00	62%	0.00	0.15	61%	0.00	0.14	0.86
Asset index	52%	100%	66%	0.00	58%	0.01	0.00	67%	0.00	0.64	0.00
Vulnerability to reductions in food consumption	52%	100%	59%	0.00	56%	0.06	0.25	60%	0.00	0.65	0.09
Total value of livestock owned	52%	100%	66%	0.00	56%	0.08	0.00	62%	0.00	0.07	0.01
Total cultivated land	52%	100%	61%	0.00	53%	0.64	0.00	58%	0.01	0.20	0.01
Education level of household head	52%	100%	59%	0.00	54%	0.36	0.03	58%	0.01	0.66	0.05

This table shows poverty rates among targeted households for each of the targeting methods (column) using different poverty metrics (row). Since the complete selection process identified the poorest 52%, a household is considered poor according to each of the metrics if it ranks among the 52% lowest scoring households on the metric.

Table 8: Poverty rates among targeted households using different poverty metrics - Peru

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Poverty metric	Random selection	Perfect targeting	Complete TUP targeting	p-value (1)=(3)	Housing Index targeting	p-value (1)=(5)	p-value (3)=(5)	PPI targeting	p-value (1)=(8)	p-value (3)=(8)	p-value (5)=(8)
Weekly consumption per adult equivalent	59%	100%	61%	0.42	61%	0.55	0.99	62%	0.39	0.81	0.82
Asset index	59%	100%	71%	0.00	64%	0.19	0.02	72%	0.00	0.71	0.02
Vulnerability to reductions in food consumption	59%	100%	58%	0.71	58%	0.81	0.92	57%	0.60	0.77	0.72
Total value of livestock owned	59%	100%	70%	0.00	58%	0.75	0.00	61%	0.65	0.00	0.40
Total cultivated land	59%	100%	66%	0.00	59%	0.92	0.01	63%	0.28	0.22	0.21
Education level of household head	59%	100%	63%	0.19	67%	0.01	0.09	59%	0.93	0.25	0.01

This table shows poverty rates among targeted households for each of the targeting methods (column) using different poverty metrics (row). Since the complete selection process identified the poorest 59%, a household is considered poor according to each of the metrics if it ranks among the 59% lowest scoring households on the metric.

Table 9: Regressions - Honduras

	(1)	(2)
	Group number in PWR (higher = poorer)	Complete TUP targeting (0/1)
Poverty metrics		
Total weekly consumption per adult equivalent (L 100)	-0.04	-0.03*
	(0.03)	(0.01)
Asset index	-0.16***	-0.05***
	(0.02)	(0.01)
Vulnerability to reductions in food consumption	0.00	0.02*
	(0.02)	(0.01)
Total value of livestock owned (L 10000)	-0.04	0.00
	(0.05)	(0.03)
Total cultivated land (ha)	-0.05	-0.03*
	(0.04)	(0.02)
Education level of household head	-0.12***	-0.06**
	(0.04)	(0.02)
Other covariates		
Household size	-0.04***	-0.01*
	(0.02)	(0.01)
Household head under 30	0.05	0.08
	(0.10)	(0.05)
Household head over 60	-0.12	-0.08
	(0.11)	(0.06)
Household head is widow(er)	0.23*	0.10
	(0.13)	(0.06)
Household received transfer from another hh last year	-0.06	-0.05
	(0.08)	(0.04)
Household received support from government last year	0.09	0.01
	(0.08)	(0.04)
Household received support from NGO last year	0.15	-0.01
	(0.10)	(0.05)
Household holds savings	-0.21**	-0.08
	(0.10)	(0.05)
Household took loan in past 12 months	0.04	0.03
	(0.07)	(0.04)
Constant	4.47***	0.96***
	(0.17)	(0.08)
Number of observations	834	834

OLS estimates with robust standard errors. Statistical signficance denoted * = 10%, ** = 5%, *** = 1%.

Table 10: Regressions - Peru

	(1) Group number in PWR (higher = poorer)	(2) Complete TUP targeting (0/1)
Poverty metrics		
Total weekly consumption per adult equivalent (10 S)	-0.04*	-0.03***
	(0.02)	(0.01)
Asset index	-0.06***	-0.03***
	(0.02)	(0.01)
Vulnerability to reductions in food consumption	-0.02	-0.01
	(0.02)	(0.01)
Total value of livestock owned (1000 S)	-0.10***	-0.03***
	(0.02)	(0.01)
Total cultivated land (ha)	-0.11***	-0.04
	(0.04)	(0.03)
Education level of household head	-0.03***	-0.02***
	(0.01)	(0.01)
Other covariates		
Household size	-0.03*	-0.02**
	(0.02)	(0.01)
Household head under 30	0.40***	0.11**
	(0.10)	(0.05)
Household head over 60	-0.32*	-0.22***
	(0.19)	(0.08)
Household head is widow(er)	0.08	-0.01
	(0.11)	(0.06)
Household received transfer from another hh last year	0.03	0.03
	(0.09)	(0.05)
Household received support from government last year	0.26***	0.12**
	(0.09)	(0.05)
Household received support from NGO last year	0.17**	0.03
	(0.08)	(0.04)
Household holds savings	0.16*	0.09**
	(0.08)	(0.04)
Household took loan in past 12 months	0.04	-0.02
	(0.09)	(0.04)
Household suffered an income shock in the past 12 months	0.08	0.15***
	(0.07)	(0.04)
Communal participation index	0.04**	0.03***
	(0.02)	(0.01)
Number of potential lenders in the community (for S/. 50)	0.02*	0.01**
	(0.01)	(0.01)
Number of times attended communal meetings in past 12 months	-0.01	0.00
	(0.00)	(0.00)
Self-reported economic status (1-10)	-0.10***	-0.02
	(0.03)	(0.01)
Constant	3.74***	0.96***
	(0.29)	(0.16)
Number of observations	635	635

OLS estimates with robust standard errors. Statistical signficance denoted * = 10%, ** = 5%, *** = 1%.

Table 11: Estimated costs of targeting for Peru and Honduras

	Peru			Honduras		
	Total	Per hh screened	Per hh selected	Total	Per hh screened	Per hh selected
Two step targeting	$33,127	$3.85	**$7.4**	$26,372	$3.30	**$6.34**
PWR	$16,340	$1.90	**$3.7**	$13,333	$1.67	**$3.21**
Verification	$16,787	$1.95	**$3.8**	$13,038	$1.63	**$3.13**
PPI/Housing index	$26,230	$3.05	**$5.9**	$20,967	$2.62	**$5.04**

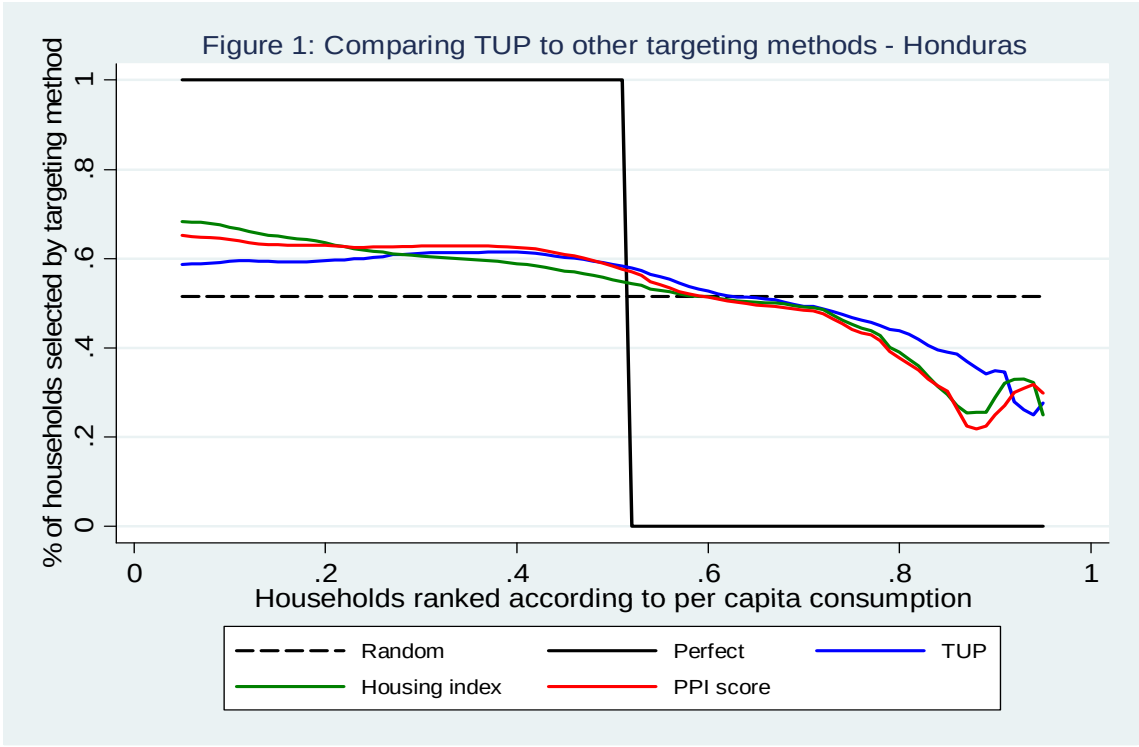


Figure 1: Comparing TUP to other targeting methods - Honduras
% of households selected by targeting method
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
Households ranked according to per capita consumption
Random
Perfect
TUP
Housing index
PPI score

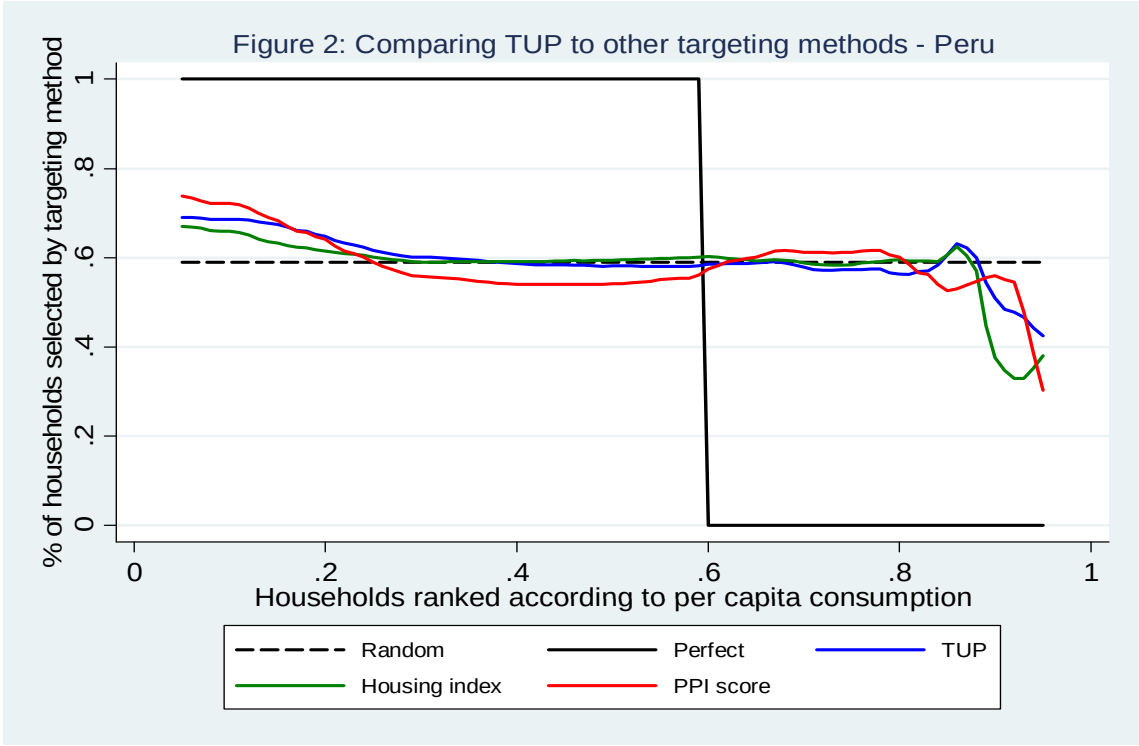
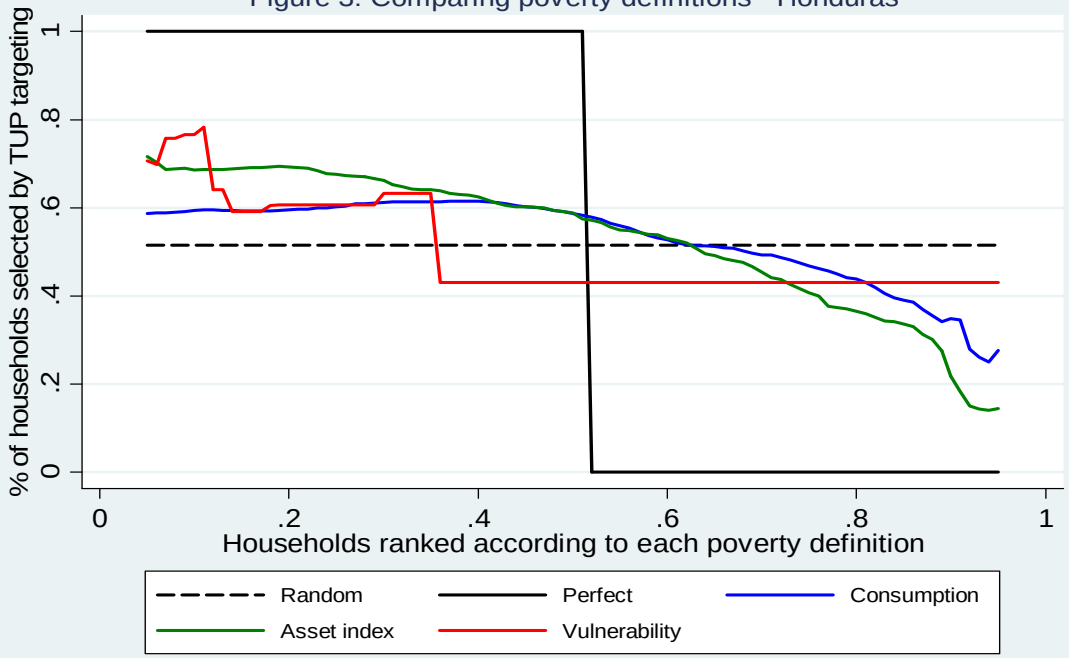


Figure 2: Comparing TUP to other targeting methods - Peru
% of households selected by targeting method
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
Households ranked according to per capita consumption
Random
Perfect
TUP
Housing index
PPI score

Figure 3: Comparing poverty definitions - Honduras



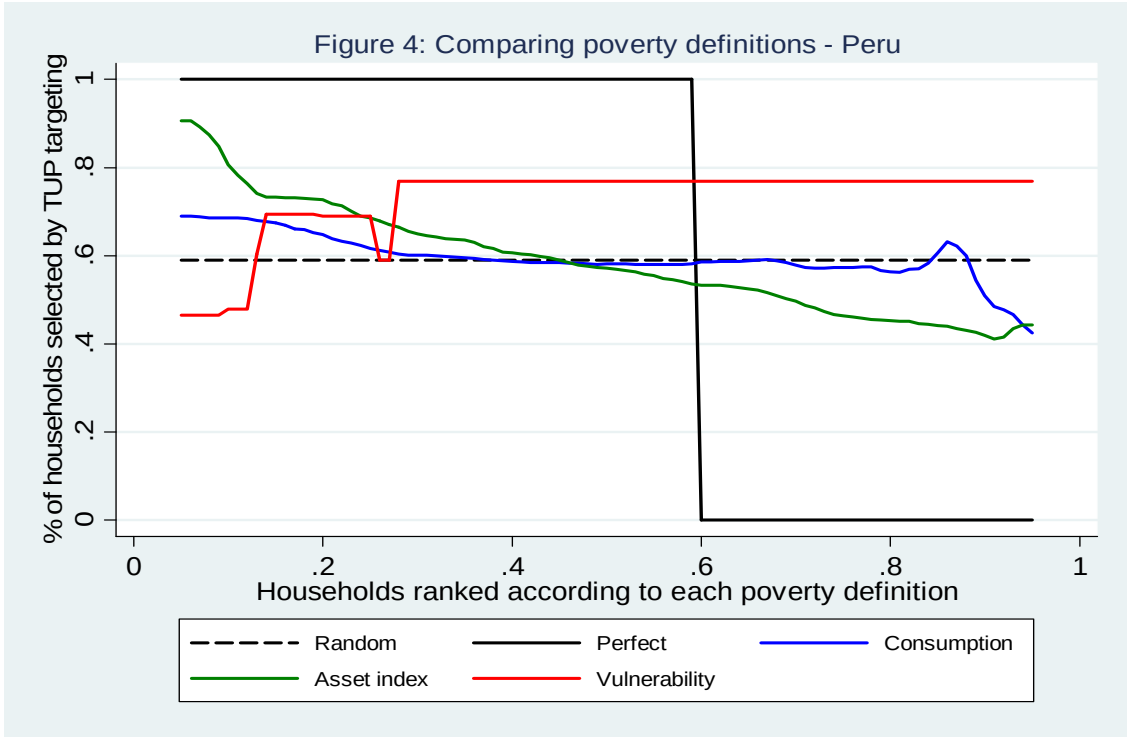


Figure 4: Comparing poverty definitions - Peru
% of households selected by TUP targeting
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
Households ranked according to each poverty definition
Random
Perfect
Consumption
Asset index
Vulnerability

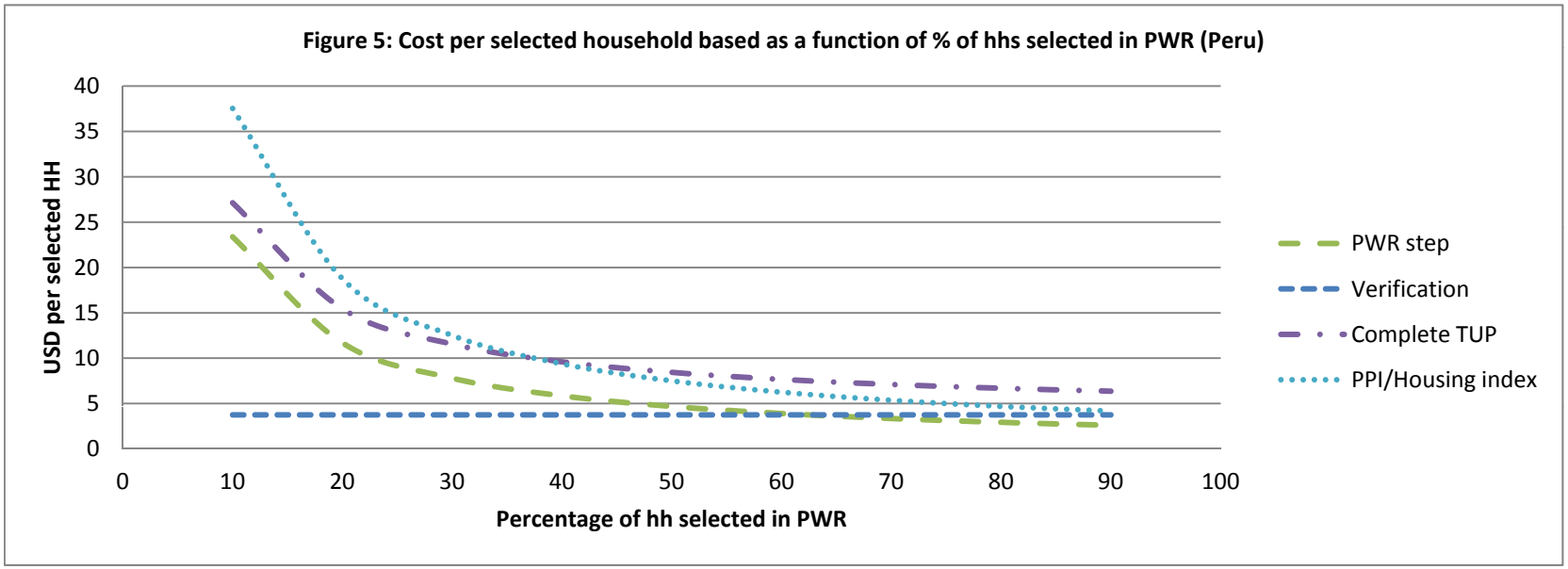


Figure 5: Cost per selected household based as a function of % of hhs selected in PWR (Peru)
USD per selected HH
40
35
30
25
20
15
10
5
0
0
10
20
30
40
50
60
70
80
90
100
Percentage of hh selected in PWR
PWR step
Verification
Complete TUP
PPI/Housing index

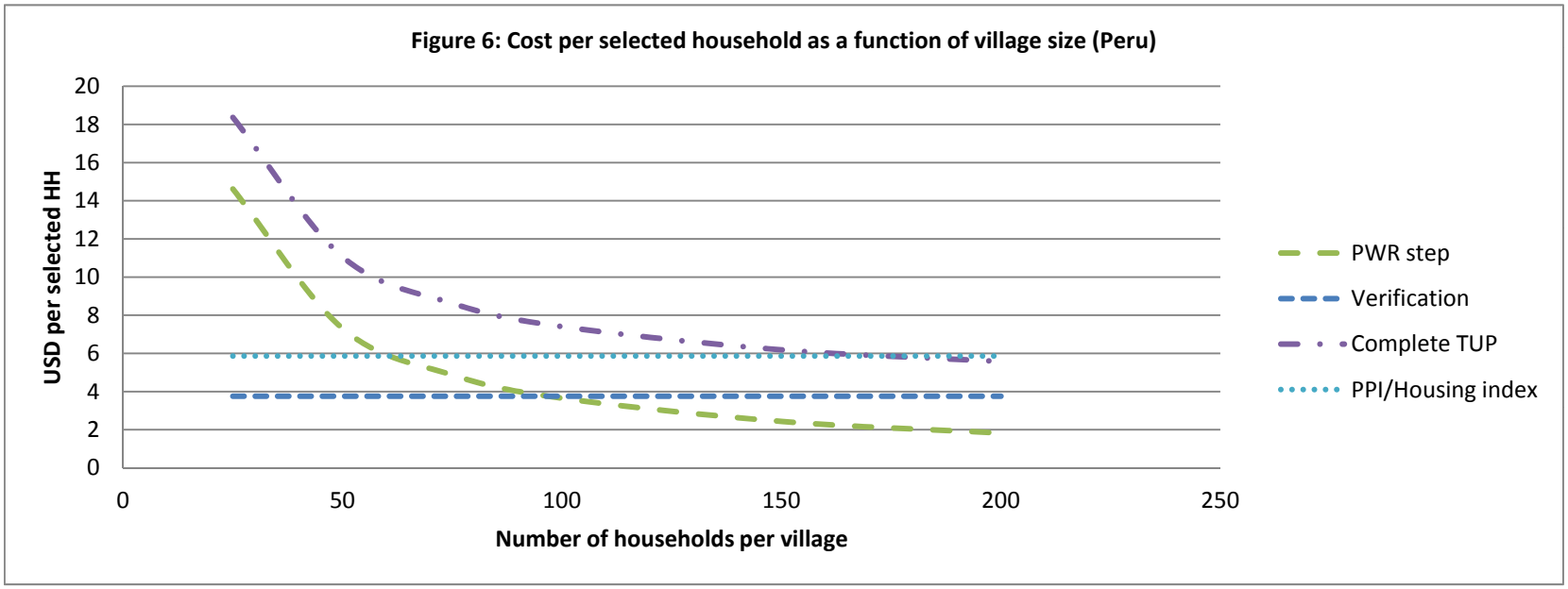


Figure 6: Cost per selected household as a function of village size (Peru)
USD per selected HH
20
18
16
14
12
10
8
6
4
2
0
0
50
100
150
200
250
Number of households per village
PWR step
Verification
Complete TUP
PPI/Housing index