WIDER Working Paper 2014/040

Targeting social transfer programmes

Comparing design and implementation errors across alternative mechanisms

Rachel Sabates-Wheeler, Alex Hurrell, and Stephen Devereux*

February 2014

Abstract: Identifying the poorest for selection into social transfer programmes is a major challenge facing programme implementers. An innovative cash transfer programme in northern Kenya trialled three targeting mechanisms to learn lessons about which approach is most effective at minimizing inclusion and exclusion errors. We conclude that community-based targeting is the most accurate of the three approaches, followed by categorical targeting by age and household dependency ratio. However, targeting performance is strongly affected by implementation capacity and modalities. Through a simulation exercise we show that a proxy means test would have performed better than single categorical indicators.

Keywords: targeting, social transfers, errors, Kenya
JEL classification: I38, I32, D60

Acknowledgements: The authors acknowledge the inputs of other members of the monitoring and evaluation team, notably Karen Tibbo, Fred Merttens, Patrick Ward, and Luca Pellerano.

---

*Rachel Sabates-Wheeler and Stephen Devereux are from the Centre for Social Protection at the Institute of Development Studies, Brighton, UK. Alex Hurrell is from Oxford Policy Management. Corresponding author r.sabates-wheeler@ids.ac.uk

This study has been prepared within the UNU-WIDER project 'ReCom–Foreign Aid: Research and Communication', directed by Tony Addison and Finn Tarp.

ISSN 1798-7237 ISBN 978-92-9230-761-5

Typescript prepared by Lesley Ellen for UNU-WIDER.

UNU-WIDER gratefully acknowledges specific programme contributions from the governments of Denmark (Ministry of Foreign Affairs, Danida) and Sweden (Swedish International Development Cooperation Agency—Sida) for ReCom. UNU-WIDER also gratefully acknowledges core financial support to its work programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy-making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

UNU-WIDER, Katajanokanlaituri 6 B, 00160 Helsinki, Finland, wider.unu.edu

1 Introduction

Targeting describes a range of mechanisms for identifying households or individuals who are defined as eligible for resource transfers, and simultaneously screening out those who are defined as ineligible. Achieving this simple objective is one of the most challenging aspects of implementing social transfer programmes, and typically requires trade-offs to be made between targeting accuracy and targeting costs (broadly defined). Accurate targeting is expensive and time-consuming—in emergency contexts, when time is of the essence, or in contexts where poverty is widespread, the budgetary costs of identifying and excluding the non-poor might exceed the savings. Conversely, crude targeting or no targeting (universal coverage) can be extremely wasteful of scarce resources.

These trade-offs inform not only the decision about **whether** to target, but also the choice of **how** to target; which targeting mechanism to apply. For instance, geographic targeting (blanket coverage of an area) is quick and cheap because no individual assessment is required, but it is usually inaccurate because it does not discriminate between local residents on characteristics of interest, notably their relative wealth. On the other hand, means testing is generally acknowledged as the most accurate mechanism for identifying the poor (if done rigorously), but it is costly to do well because it requires detailed and sensitive personal information about each potential beneficiary that must be elicited, verified, and regularly reassessed—many means tested programmes do 'retargeting' every year. Other targeting mechanisms, such as categorical approaches (including those focusing on age cohorts such as children or older persons), proxy means tests (using multiple indicators to identify the poor), and self-selection (for instance the labour requirement on public works programmes) are each associated with varying degrees of targeting accuracy or error and targeting cost (see Barrientos and Niño-Zarazúa 2011).

Social transfer programmes have spread rapidly throughout sub-Saharan Africa in recent years (Ellis et al. 2009; DFID 2011; Handa et al. 2011). Social transfers can be defined as regular non-contributory payments, in cash or in kind, provided by government or non-government organizations to individuals or households, with the objective of decreasing chronic or shock-induced poverty, addressing social risk, or reducing economic vulnerability (Samson et al. 2006: 2). Building on their origins in social safety nets (e.g. public works projects) and humanitarian relief interventions (emergency food aid), social transfers have recently become dominated by unconditional cash transfers to poor and vulnerable households (Hanlon et al. 2010). Because regular cash transfers are expensive, they are always targeted (sometimes to just a few thousand individuals, sometimes to several millions). But cash transfers are difficult to target well, not least because they are valuable and free, which creates incentives for applicants to misrepresent their true status in order to qualify and for officials to defraud the system.

The Hunger Safety Net Programme (HSNP) in northern Kenya is an example of an unconditional cash transfer programme. Launched in 2009, the HSNP aims to reduce extreme poverty by delivering regular cash transfers to some 300,000 poor and vulnerable individuals in the greater Mandera, Marsabit, Turkana, and Wajir districts. Targeting in this context presents considerable challenges, not just logistical, but also in terms of defining an appropriate and identifiable target population: appropriate in terms of being consistent with the programme's objective to reduce extreme poverty; and identifiable in terms of exhibiting specific observable and verifiable characteristics.

In contexts such as northern Kenya, targeting cannot rely on household income as the measure of relative living standards, since income flows are often irregular, hard to capture accurately, and verify, and do not capture self-production; a core livelihood strategy in these areas. An alternative measure of living standards, widely used for the analysis of poverty, is consumption expenditure, but this is generally not a feasible measure on which to target, since it requires capturing detailed information about all items consumed (food and non-food) over a given reference period for all households potentially eligible for the programme. Recording such detailed information from all households within a community being targeted requires so much time and resources that it becomes inefficient and impractical.

Alternative approaches involve identifying proxies for poverty or vulnerability, which are then used as measures for targeting. Such approaches usually rely on statistical analysis of the correlates of objective measures of poverty or vulnerability. This can either identify individual characteristics on which to target (e.g. orphan status, age, and disability, etc), or a set of characteristics which are jointly associated with poverty status (e.g. proxy means testing). However, such approaches rely on the availability and quality of suitable, accurate, and up-to-date data, and are, therefore, not always available or reliable. Another option is to abandon any explicit link between the targeting approach and objective measures of poverty or vulnerability, and instead target specific categories of individuals or households according to other criteria (e.g. logistical or geographical feasibility, social or political acceptability, and rights-based concerns). Finally, the types of households to be selected, or the actual households to be selected, could be chosen by communities themselves, according to more or less subjective or objective criteria, externally moderated or otherwise.

Due to the lack of detailed household-level data, the HSNP was not able to identify targeting measures that were explicitly associated with any objective measures of poverty. It was, therefore, decided that other mechanisms would be used as proxies for poverty targeting. During Phase 1 (2008-12) three targeting mechanisms were trialled: (1) community-based targeting; (2) households with high dependency ratios; and (3) older people (55 years or above). This experimental design provides a rare opportunity to compare outcomes across different targeting mechanisms within the same programme.

The focus of this paper is on targeting effectiveness, or accuracy.[1] This introduces another trade-off, between the two errors of targeting: **inclusion** (selection of ineligible beneficiaries) and **exclusion** (omission of eligible individuals or households). While policy makers tend to focus on inclusion errors, because they incur financial costs, Cornia and Stewart (1993) proposed that exclusion error should be weighted more highly than inclusion error because failure to reach people who need assistance is a humanitarian cost and compromises the achievement of programme objectives.

The question of choice between different mechanisms and their comparative effectiveness is a central concern of this paper. A comprehensive review of targeting outcomes on over 100 social transfer programmes (Coady et al. 2004) found that no single targeting mechanism performs best in all contexts—there was more variability in targeting performance within each mechanism than between them. The key determinant of targeting effectiveness is implementation capacity, or how rigorously the targeting is implemented. With these caveats, mechanisms can be ranked by their ability to identify the poor. The most accurate mechanism was found to be self-targeting (on public works programmes), followed by geographic targeting and means testing. Proxy means tests, community-based targeting and categorical targeting (of children and older people) achieved highly variable but generally poor results—sometimes they capture high proportions of

[1] Targeting mechanisms can also be assessed for their cost-effectiveness—how accurate they are in relation to their administrative cost—but this is beyond the scope of this paper.

poor people, but more often they do not. Finally, Coady et al. (2004) found that the use of multiple methods led to improved targeting accuracy over single methods.

Using new data from a randomized control trial evaluation of the HSNP, an innovative social transfer programme in Kenya that was specifically designed to test three different targeting mechanisms, we provide in this paper the first statistically robust assessment (to our knowledge) of the targeting effectiveness of different mechanisms within the same programme. Drawing on quantitative information from a baseline survey (conducted after the beneficiaries were identified but before they received any cash), we analyse the comparative performance of the three targeting mechanisms in terms of:

1. the ability of the mechanisms to identify the poorest and most vulnerable households in the population (coverage, inclusion, and exclusion by design);
2. whether the selected beneficiary households actually fulfil the eligibility criteria (implementation problems);
3. whether the selected households are actually poor; and
4. specification of a range of alternative targeting criteria as performance benchmarks for the three mechanisms actually used by the programme.

2 Targeting performance

Most social transfer programmes, including the HSNP, attempt to transfer resources to the poorest members of the population. So the measure of a targeting mechanism's effectiveness is how accurately it identifies poor people. According to Ravallion (2007: 7): 'A Type 1 [inclusion] error can be defined as incorrectly classifying a person as poor, while a Type 2 [exclusion] error is incorrectly classifying a person as not poor'. As noted above, targeted social transfer programmes are susceptible to two types of error:

- **inclusion error** can be quantified as the proportion of a programme's beneficiaries who receive transfers despite not being poor; and
- **exclusion error** can be quantified as the proportion of people in poverty who are omitted from a social transfer programme.[2]

Errors of inclusion and exclusion can arise at the design stage and/or during implementation.

Targeting errors by design: a common design challenge is a binding budget constraint which means that a programme cannot reach all poor households in the country, so either a quota is applied or the intervention is restricted to a geographical area, such as a district. (In effect, 'geographic targeting' becomes the first level of targeting.) This 'exclusion by design' is not, strictly speaking, a mistake, but planned 'under-coverage'. Similarly, geographic targeting will inevitably reach some non-poor beneficiaries, but this source of 'leakage by design' is intrinsic to the selection of this targeting mechanism, it is not a mistake made by programme administrators during implementation. Design errors will also arise where proxy measures are used to identify poor households, or where the targeting criteria have been selected with no explicit link to objective poverty measures. Such 'errors' can also be considered not as mistakes, for example in cases where it is known that a certain measure (e.g. old age) is not a perfect proxy for poverty, but is the best feasible option available (e.g. for logistical, social, or political concerns). So

[2] 'Inclusion error' is also known as leakage, vertical inefficiency, Type 1 error, or specificity. 'Exclusiion error' is also known as undercoverage, horizontal inefficiency, Type 2 error, or sensisitivity.

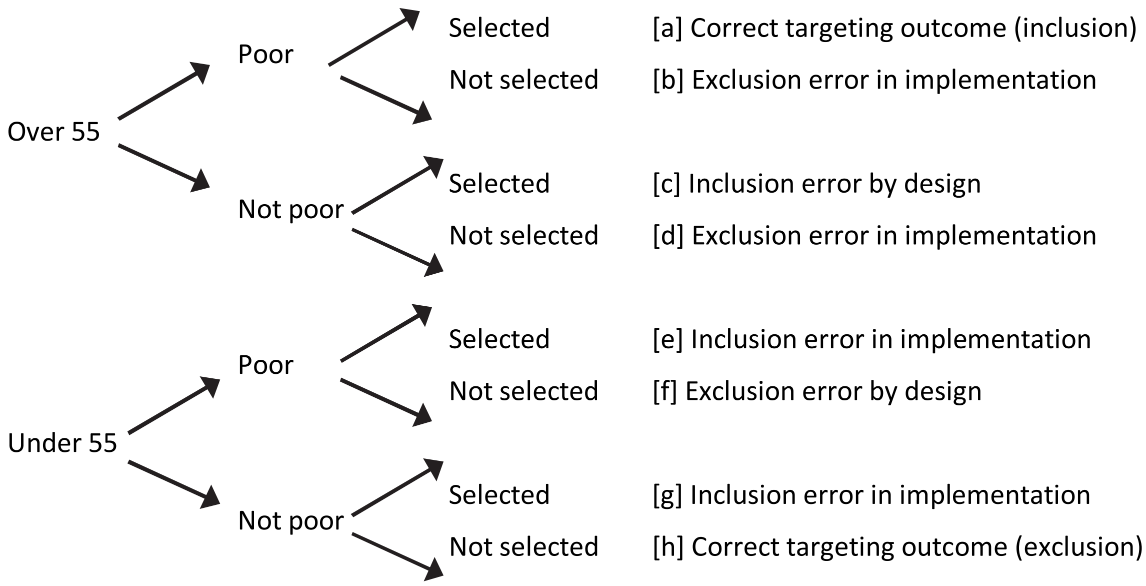
targeting may be done on the basis of old age even though it is known that some older persons will not be poor, because this is balanced by the fact that many older persons *are* poor and will, therefore, be reached by the programme. There are situations, however, where design errors should be viewed as a failure of the programme designers. For example, if the programme's stated objective was to target the poorest households, but its targeting criteria turned out to be only weakly associated with most objective measures of poverty—i.e. the programme decided to target older persons specifically *because* it thought they were poor, but it turned out that older persons tended to be richer. Similarly, targeting design failures can occur in the case of poorly designed proxy means testing which, either due to poor quality data or flawed statistical analysis, fails to accurately identify poor households.

Targeting errors in implementation: inclusion errors in implementation can occur if applicants misrepresent their true status in order to satisfy the eligibility criteria—say, by claiming to be poor when they are not—or if corrupt officials collude with ineligible applicants to register them. Exclusion errors in implementation can occur if eligible households miss the registration process (e.g. if pastoralists are away herding livestock) or if communities deliberately exclude marginalized members from a community-based targeting exercise. Both these possibilities are potential concerns for the HSNP.

For programmes such as HSNP, where targeting is based on characteristics that serve as proxies for poverty, such as age, assessing targeting effectiveness is complicated because two measures of targeting accuracy must be assessed: (1) the 'implementation accuracy' of the programme in reaching the eligible and excluding the non-eligible; and (2) the 'design accuracy' of the eligibility criteria as proxies for poverty. For instance, in some communities the HSNP targets all individuals over 55 years of age for a 'universal' social pension, so 'inclusion error in implementation' is the proportion of pension recipients who are under 55, while 'exclusion error in implementation' is all non-recipients over 55. But to assess the effectiveness of the social pension as a poverty targeting mechanism, we also need to know the percentage of people under 55 in HSNP communities who are poor (i.e. 'exclusion error by design') and the percentage of people over 55 who are non-poor (i.e. 'inclusion error by design').

All eight possible outcomes are depicted in Figure 1. Only two outcomes—inclusion of poor people over 55 [a] and exclusion of non-poor people under 55 [h]—are unambiguously correct targeting outcomes. If any poor or non-poor people over 55 are not selected into the programme, these are exclusion errors in implementation ([b] and [d]). If any poor or non-poor people under 55 are selected into the programme, these are inclusion errors in implementation ([e] and [g]). If any people over 55 who are selected into the programme are non-poor, this is inclusion error by design [c]. Finally, if people under 55 who are (correctly) not selected into the programme are poor, this is exclusion error by design (outcome [f]).

Figure 1: Targeting outcomes in a 'universal' social pension programme



Source: Authors' illustration.

So targeting performance must be assessed both in terms of **eligibility criteria** and in terms of the **poverty profile** of the population—these are not always the same. Both sources of inclusion and exclusion error in HSNP targeting design and implementation are reported in this paper.[3]

3 Targeting mechanisms

The range of possible targeting mechanisms can be grouped into six clusters. Each has its advantages and limitations, and evaluations of targeted social transfer programmes lead to the conclusion that there is no single optimal mechanism, either in theory or in practice. At the programme design stage, the choice of mechanism must weigh up different objectives—such as cost *versus* accuracy—while targeting performance in practice is highly variable, depending crucially on how well the targeting process is implemented (Coady et al. 2004).

1) **Means tests** assess the income and/or assets of individual applicants, and define anyone falling below a predetermined threshold as eligible. Means tests are expensive to apply, especially among poor populations with insecure livelihoods and unpredictable incomes. They are susceptible to under-reporting of incomes and assets by applicants, which requires verification and regular reassessments. They can create incentives to modify behaviour in order to become or remain eligible ('disincentive' and 'dependency' effects). Evidence from programmes in Albania (Alderman 2002), China (Ravallion 2007), and Krygyz Republic (Tesliuc 2004) finds that means tests are surprisingly inaccurate, with inclusion and exclusion errors both often exceeding 50 per cent.

[3] The decision trees for the different targeting mechanisms reviewed here, as well as the relevant statistics, are presented in Annex 3.

2) **Geographic targeting** concentrates resources within bounded areas (e.g. food insecure districts) and either defines all local residents as eligible or applies additional targeting mechanisms to narrow the eligible population. Location is a popular first-level targeting criterion for social transfers where budget constraints prohibit national coverage (e.g. on northern Kenya's HSNP). Universal coverage minimizes exclusion error, but only within the programme area. In Vietnam, programmes targeted at poor, remote communes have low inclusion error (8 per cent) but most poor people live in other rural areas or urban centres, so exclusion error (or under-coverage) at national level is very high (81 per cent) (Minot and Baulch 2002).

3) **Self-targeting** requires poor people to identify themselves, by lowering the value of transfers (e.g. giving food aid rather than cash) or raising the cost of accessing transfers (e.g. making beneficiaries queue). Public works programmes use both mechanisms to discourage non-poor people from applying: wages are set below local market rates, and a precondition for receiving payment is doing several hours of heavy manual labour. While inclusion errors on public works tend to be low, exclusion or under-coverage is usually high, because it is challenging to create enough workplaces, and because many poor people cannot work (HLPE 2012).

4) **Proxy means tests** choose a weighted combination of characteristics that are believed to be correlated with wellbeing or deprivation, and are more easily observed than income and assets (e.g. housing conditions). They are less expensive to administer and less prone to manipulation than means tests, but their performance depends critically on which proxies are applied and how they are weighted. Inclusion errors on proxy means test programmes in Brazil, Egypt, and Mongolia ranged from 16-57 per cent, while exclusion errors ranged from 19-45 per cent (Fiszbein and Schady 2009; Falkingham and Namazie 2002; Hodges et al. 2007; respectively).

5) **Categorical targeting** selects characteristics that are easy to observe and verify and are believed to be well correlated with poverty or vulnerability (e.g. disability). Popular 'vulnerable groups' include orphans, female-headed households, and older persons—the HSNP, for instance, targets older persons and households with high dependency ratios. Though relatively cheap and simple to administer, these crude 'proxy indicators' of poverty are often inaccurate. In Niger and Yemen, almost 80 per cent of older persons live in non-poor households and 80 per cent of poor households have no elderly members (Grosh and Leite 2009), so even a well-implemented social pension would generate high inclusion and exclusion errors by design.

6) **Community-based targeting** delegates beneficiary selection to community members. The main advantage is that local residents have personal knowledge of their neighbours, which project administrators do not. The main risk is that the process will be dominated by local elites who capture the benefits and exclude socially marginalized groups. A review of evidence found that: 'the targeting of poor communities and poor households within communities is markedly worse in more unequal communities (Mansuri and Rao 2004: 55). Because implementation modalities vary greatly, Coady et al. (2004) found that community-based targeting achieved the most variable results of all mechanisms.

We are limited in terms of which of these targeting mechanisms we are able to analyse. The programme we are evaluating does not use either means testing or self-targeting. We are not able to analyse the geographic aspect of HSNP targeting (i.e. the targeting impact of the design choice to implement the programme in four of the poorest districts in Kenya). Thus, for purposes of this paper and given the nature of the available data we examine three types of targeting mechanism: categorical, community-based targeting, and proxy means tests. Specifically, we are able to compare the actual outcomes of two categorical mechanisms (a social pension and a dependency ratio) and a community-based mechanism. We are also able to simulate outcomes using a range of categorical mechanisms as well as a proxy means test.

4 Targeting on the hunger safety net programme (HSNP)

The HSNP chose three targeting mechanisms and aimed to compare their performance. Two of the mechanisms are variants of categorical targeting (dependency ratio and older persons) and the third is community-based targeting. Clearly the three target populations corresponding to the three mechanisms are different as the eligibility criteria are different. In fact, the first step of our targeting analysis is to assess whether the programme reached the three intended target groups. To the extent that they did not, these are targeting errors in implementation. However, in the context of this programme the purpose of using the different mechanisms was to evaluate their comparative accuracy as proxies for poverty. The programme's objective was to target the extreme poor. So we are assessing how well each of the three mechanisms performed at actually selecting beneficiaries that are poor. Step two of our analysis is to assess whether households that are eligible to receive HSNP—because they are old, have high dependency ratio (DR) or through communities' selection—are in fact poor. Step three is to compare the performance of the three targeting approaches in terms of their 'pro-poorness'—which mechanism identifies the highest proportion of poor beneficiaries?

Targeting was implemented by sub-location, and only one targeting mechanism was used in each sub-location. For the areas covered by the evaluation, mechanisms were randomly allocated. This randomization underpins the evaluation methodology which is set out in Section 5 below. There was no retargeting during Phase 1, but individuals (social pensioners) and households (under dependency ratio and community-based targeting) leave the programme if they die or migrate out of the HSNP area.

1) **DR targeting:** a household's dependency ratio is defined by the HSNP as the number of individuals who are not working (under 18 or over 55 years of age, chronically ill, or disabled) as a proportion of total household size. Households with high dependency ratios living in designated sub-locations are eligible to receive assistance from the HSNP. The threshold for eligibility was set at 0.6 for Turkana and Marsabit, and 0.67 for Mandera and Wajir. (A DR of 0.67 can be interpreted as two 'dependents' for each 'worker'.) The logic underpinning DR targeting is that households with many dependents will be relatively or absolutely labour-constrained—some households have no working members at all—and that even if working individuals earn similar incomes, households with higher dependency ratios are poorer because they will have lower per capita incomes. But targeting based on household dependency ratios presents several challenges. Firstly, there are practical difficulties in establishing the correct age, degree of disability, and health status for every household member, which makes DR targeting administratively complex and time-consuming. Secondly, the assumption that a high DR signifies poverty might not hold in the pastoralist communities of northern Kenya, where wealthier households tend to be larger and children contribute to household income from a young age (e.g. by herding the household's livestock). This problem was recognized during the HSNP's inception phase, but a proposal to supplement DR targeting with a means test based on asset ownership was rejected.

2) **Categorical targeting by age (CTA):** all individuals aged 55 or above living in designated CTA sub-locations on the date of registration for HSNP were eligible to receive a non-contributory 'social pension'. Proof of age was established from the applicant's national identity card. If the applicant had no official documents, he or she was vetted by a committee representing the community. Social pensions are a popular social protection instrument with policy makers, but this approach to targeting on the HSNP encountered several problems. One supposed advantage of targeting older persons is that age is a single and easily verifiable characteristic—but few older persons in northern Kenya have birth certificates or accurate national identity cards. Another justification for social pensions is that providing support to older persons is easily understood and widely accepted—but in pastoralist cultures, where old age is often associated with power and wealth, the rationale for giving older persons (and no-one else) free cash was not at all obvious to community members. Finally, old age is supposedly associated with poverty, making this a robust proxy indicator—but cross-country evidence reveals that social pensions are relatively inaccurate in targeting the poor. In a review of 111 programmes, Coady et al. (2004) found that targeting older persons was the second worst mechanism in terms of reaching the poor.

3) **Community-based targeting (CBT):** in northern Kenya, CBT is the dominant form of targeting for programmes such as food aid, the rationale being that communities themselves are best placed to identify their poorest members. For the HSNP, communities selected households they considered most in need of assistance. Since the programme aims to address chronic poverty rather than acute need, each community was allocated a fixed quota that was set at 50 per cent of the expected number of households in the community. But since the community household population data was often inaccurate, and since chronic poverty levels may vary substantially between sub-locations, the standardized quota almost certainly overestimated the actual poverty rate (i.e. inclusion error) in some sub-locations and under-estimated it (i.e. exclusion error) in others. Lack of comparability across communities is a generic problem with CBT (Conning and Kevane 2002).

5 Evaluation design

The programme was implemented across the four former districts of Mandera, Marsabit, Turkana, and Wajir, and within these only in secure areas. The evaluation took place in 48 randomly selected sub-locations, out of the 356 sub-locations in the former districts of Mandera, Marsabit, Turkana, and Wajir that were assessed as secure and, therefore, where the HSNP could potentially operate.[4]

In each of the 48 sub-locations, beneficiaries were then selected for the programme according to the relevant mechanism. Once this was done, half the evaluation sub-locations were randomly assigned to be 'treatment' areas and received the programme payment immediately after the baseline survey had taken place in that sub-location. The other 24 sub-locations were assigned to be 'control' areas, where selected households would begin to receive transfers after two years. Because exactly the same targeting process took place at the same time in both treatment and control areas, the targeting analysis results relate to all 48 evaluation sub-locations.

[4] A 'sub-location' is an officially defined geographical unit corresponding to the lowest level of government administration (the sub-location chief).

The work presented here draws from both the quantitative survey and qualitative fieldwork. The quantitative survey comprises: (1) a household panel survey conducted on an annual basis (baseline, year 1 follow-up, year 2 follow-up), covering 5,108 randomly selected households in the 48 evaluation sub-locations, also sampled at random; and (2) quantitative community interviews conducted with a group of 10-12 community members on an annual basis (baseline, year 1 follow-up, year 2 follow-up) in the same 48 randomly sampled sub-locations. The qualitative fieldwork was dominated by community interviews with mixed-gender groups of community members (chief, elders, and others).

Because targeting was conducted in both treatment and control areas, households were sampled in the same way across both. Sixty-six beneficiary households were sampled from HSNP administrative records, using simple random sampling in each sub-location. Forty-four non-beneficiary households were sampled from household listings undertaken by the evaluation field teams in a sample of three randomly selected settlements within each sub-location. These settlements were stratified into three different types, and one settlement of each type was sampled.[5] Within settlements, all households were listed. During the listing, any beneficiary households were identified and then dropped from the sample frame

In this way a representative sample of beneficiary and non-beneficiary households was constructed. By comparing the characteristics of households selected by the HSNP targeting process (beneficiaries) with those not selected (non-beneficiaries), an assessment can be made of the degree to which HSNP was successful in targeting the poorest households in its operational areas. This comparison of beneficiary and non-beneficiary households is the basis of the targeting analysis.

Data analysis was undertaken using analytical weights that are the inverse of households' selection probabilities, taking as given that the sub-location was selected for inclusion in the study population. The estimates presented in this paper are, therefore, representative of the study population—that is, the population across the 48 sub-locations selected for inclusion in the study—rather than the entire population of the HSNP districts or even of the areas covered by the HSNP. Since the programme operated differently in some respects in the non-evaluation sub-locations, the findings also represent the programme as it operates in the evaluation sub-locations.

6 Findings

6.1 Characteristics of the sample: poverty, consumption, and demography

Table 1 shows the poverty rates in the HSNP districts according to the 2005-06 Kenya Integrated Household Budget Survey (KIHBS). These estimates are derived by estimating the proportion of households in the HSNP districts with total equivalized monthly expenditure below the official absolute, food, and hardcore poverty lines respectively. The estimates were calculated using the appropriate sample weights. On all measures it is clear that poverty rates in

[5] If there was no non-permanent settlement, a second permanent settlement was sampled. If there was no other permanent settlement (apart from main settlement), then a second non-permanent settlement was sampled. If there were neither enough permanent nor non-permanent settlements, then all remaining households were listed from the main settlement. Note that by definition there can only be one main settlement per sub-location. Large settlements (over approximately 300 households) were segmented into segments of approximately 100-150 households, and segments were then sampled using simple random sampling.

the programme areas are extremely high. In fact this was the basis for the decision to focus the programme on the HSNP districts, since these were identified by the KIHBS as being the four poorest districts in Kenya. In terms of targeting effectiveness, high absolute poverty rates are generally associated with low inclusion errors (since most households are poor), but high design exclusion errors (since the programme cannot afford to reach all poor households).

The household questionnaire collected information on each household's consumption and expenditure, which formed the basis for measuring consumption poverty.[6] Consumption is usually reported more reliably than income. Households were classified into five equal groups (quintiles) according to consumption expenditure levels such that quintile 1 corresponds to the poorest 20 per cent of households in the HSNP evaluation areas and quintile 5 to the least poor 20 per cent.

Table 1: Poverty rates in the HSNP districts according to KIHBS 2005-06

Poverty line	All HSNP districts
Absolute poverty line (percentage of households with total equivalized monthly expenditure below KES* 1,562 (rural) or KES 2,913 (urban)	85
Food poverty line (percentage of households with total equivalized monthly food expenditure below KES 988 (rural) or KES 1,474 (urban)	78
Hardcore poverty line (percentage of households with total equivalized monthly expenditure below KES 988 (rural) or KES 1,474 (urban)	64

Note: *Kenyan Shillings (KES)

Source: Authors' calculations using KIHBS 2005-06 data.

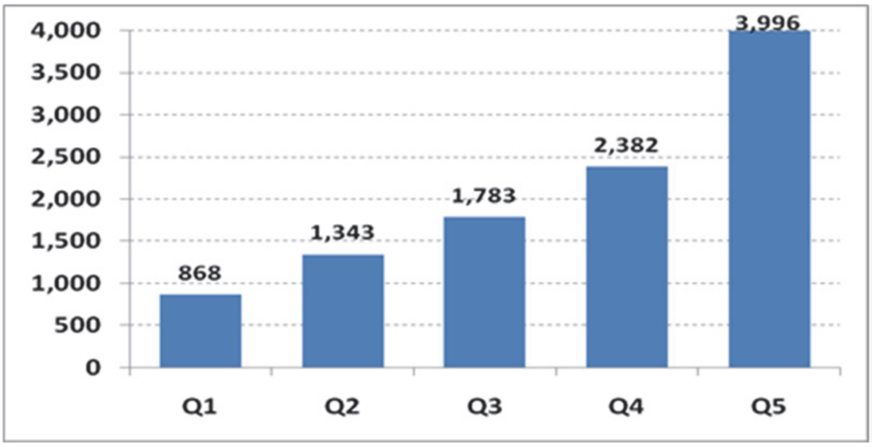
While the majority of households in our study population are poor in absolute terms, there is a substantial difference between the poorest and the least poor. The wealthiest quintile spends almost five times as much as the poorest per adult equivalent (KES 3,996 vs. KES 868), which indicates an appreciable degree of income inequality within the study population (see Figure 2). Given the resource constraints associated with social transfer provisioning in this area and these sharp inequalities, it is clear that targeting the poorest is still a fundamental concern, notwithstanding the fact that the majority of the population are poor in an absolute sense. For evaluating targeting in such situations we use a relative poverty line, described in the following section.

Although consumption expenditure is not a perfect proxy for household welfare, our analysis shows that it is highly correlated with many key dimensions of household well-being (see Annex 1, Table A1). On average, households in poorer quintiles spend a higher proportion of their total consumption budget on food, spend less on education and health services, own fewer

[6] This measure was standardized for the number of adult equivalents in each household and is used for comparing households on their level of consumption and poverty, as well as for defining each household's relative poverty status in the targeting analysis.

assets, have lower adult literacy and school enrolment rates, are more likely to have been ill or injured in the past three months, and have poorer quality housing. Furthermore, subjective poverty rates are significantly higher amongst households in poorer quintiles.

Figure 2: Mean monthly consumption expenditure per adult equivalent, by quintile in KES



Source: HSNP Monitoring and Evaluation (M&E) Baseline Evaluation Survey, Sep 2009-Oct 2010.

6.2 Programme coverage, targeting mechanisms, and poverty in the HSNP districts

Table 2 shows how programme coverage varies across the three targeting mechanisms. Overall coverage across the evaluation areas is 51 per cent, so just over half the households were selected for the programme. However, coverage varies substantially by targeting mechanism. Coverage in CTA areas is lowest, at 40 per cent, which is driven principally by the number of households containing at least one household member aged 55 or over. DR coverage is 66 per cent, which reflects the calibration of the DR eligibility cut-offs (between 0.6 and 0.7 depending on the district). The CBT coverage rate is determined by the 50 per cent quota set by programme administrators.

Table 2 also shows how consumption poverty and food security vary across the CBT, CTA and DR areas. Consumption poverty is defined using a 51 per cent relative poverty rate. A relative poverty line was chosen because it was not possible to apply inflation-adjusted KIHBS 2005-06 poverty lines to the baseline data, mainly because the instruments used to measure consumption expenditure were different, so applying poverty lines defined using KIHBS 2005-06 onto the HSNP baseline consumption expenditure data is not valid. The relative poverty line was calibrated at 51 per cent, in line with the HSNP coverage rate—given a 51 per cent coverage rate, it is hoped that those selected for the HSNP fall within the poorest 51 per cent. Households are defined as food insecure if they reported going entire days without eating during the worst recent period of food scarcity.

The variations in coverage rates by targeting mechanisms do not reflect variations in poverty and food security across the CBT, CTA, and DR areas. Poverty and food insecurity is lowest in CBT areas (42 per cent and 55 per cent respectively), but coverage in CBT areas is significantly higher than in CTA areas, which have greater levels of poverty and food insecurity. This finding is not surprising given that the CBT, CTA, and DR coverage levels were purposively set at different levels. However, combined with the fact that poverty and food insecurity also vary across CBT, CTA, and DR areas, it has significant implications for the targeting analysis.

Table 2: HSNP coverage, consumption poverty and food security by district (%)

	By targeting mechanism			All HSNP evaluation areas	
	CBT areas	CTA areas	DR areas	%	N
Coverage rate					
Proportion of households that are beneficiaries	47	40***	66***	51	5,108
Consumption poverty					
Proportion of households falling below 51% relative poverty line	42*	54	60	51	5,106
Food security					
Proportion of households identified as food insecure (went entire days without eating during worst period)	55	63	71*	63	5,106

Notes: (1) The 'N' column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes. (2) Asterisks (*) indicate that the targeting mechanism estimate is significantly different to the pooled mean across the other two mechanisms: *** = 99%; ** = 95%; * = 90%.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

The implication of variations across the targeting mechanisms in programme coverage (by design) and poverty rates (by chance, since the allocation of targeting mechanism across sub-locations was done randomly as part of the evaluation design) is that inclusion and exclusion errors—standard measures of targeting effectiveness—cannot be used for assessing the relative targeting effectiveness of the three mechanisms. Instead, ratio measures that compare the poverty rates among selected and non-selected households and to the overall poverty rate are used.

Table 3 shows comparative levels of poverty and food insecurity for beneficiary and non-beneficiary households. Beneficiary households are 30 per cent (13 percentage points) more likely to be among the *poorest* (bottom 51 per cent) as compared to non-beneficiary households (57 per cent vs. 44 per cent). In terms of food security, beneficiary households are only 16 per cent (9 percentage points) more likely to be food insecure compared to non-beneficiaries. Does this represent effective targeting? In order to understand how this compares to the targeting effectiveness of other cash transfer programmes around the world the Coady-Grosh-Hoddinott (CGH) index has been calculated and is also presented in Table 3. The CGH index is a measure of the effectiveness with which programmes are targeted. It is defined as the ratio of the value of transfers going to the poor to the (relative) size of the poor in the population.[7] This index is calculated for both the poverty measures used, giving values of 1.12 and 1.07 according to the consumption expenditure and food security measure respectively. This shows that poor households are 7-12 per cent more likely to have been selected for the programme under HSNP targeting than they would have been under random or universal targeting.

[7] So, for example, if the poorest 40 per cent of the population receive 40 per cent of the transfers by value, the ratio is 1. See Coady et al. (2004). Note that the CGH index takes into account resources transferred to the poor, rather than simply the proportion of households that are poor relative to the national poverty rate. This is consistent with our analysis, provided the value of the transfer is constant across households and there is not much variation in household size between rich and poor households. Since there are very few households receiving multiple benefits, and since household size is relatively similar across consumption quintiles, this approximation is valid and considerably simplifies the exposition of results.

Coady et al. (2004) presents empirical evidence on targeting efficiency and outcomes from 122 anti-poverty interventions in 48 countries. The median programme reviewed had an index of 1.25, implying that it transfers 25 per cent more resources to poor individuals than a universal programme. The ten best performing schemes, the majority of which are in the Americas, were shown to transfer two to four times more resources to the poor than would have occurred under a universal scheme. In other words, the targeting effectiveness of HSNP does not compare particularly well with other similar programmes in terms of targeting effectiveness at an aggregate level.

Table 3: Relative poverty rates and food security by beneficiary status

	All HSNP evaluation areas
Consumption poverty	
Proportion of households falling below 51% relative poverty line	
Beneficiary households (%)	57***
Non-beneficiary households (%)	44
Ratio of poverty rates: beneficiaries vs non-beneficiaries	1.30
CGH index: % of beneficiaries that are poor / poverty rate	1.12
Food security	
Proportion of households identified as food insecure (went entire days without eating during worst period)	
Beneficiary households	67**
Non-beneficiary households	58
Ratio of poverty rates: beneficiaries vs non-beneficiaries	1.16
CGH index: per cent of beneficiaries that are poor / poverty rate	1.07

Notes: Asterisks (*) indicate that the beneficiary household estimate is significantly different to the non-beneficiary household estimate: *** = 99per cent; ** = 95per cent; * = 90%.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

To better understand this concern, it is useful to decompose the targeting problem into issues of design and implementation (as discussed earlier). For the purposes of this paper, there are two ways of analysing targeting, based on the population to which we apply eligibility criteria.

1) The HSNP design applied only certain targeting mechanisms in specific sub-locations, and we can assess targeting for a given mechanism only for those sub-locations where it was implemented—in other words, for a restricted 'actual' sample. Furthermore, eligible households needed to be residents of those sub-locations. We restrict this analysis to those sub-locations where a specific mechanism was implemented and to households that are resident in those sub-locations.

2) We can also analyse targeting by applying eligibility criteria across the sample as a whole, irrespective of what particular mechanism was implemented in each sub-location—in other words, a hypothetical 'predicted' sample. Moreover, we can also specify alternative targeting mechanisms and simulate outcomes for comparative poverty-targeting purposes. For instance, we can simulate proxy means test targeting outcomes as well as other targeting criteria. This analysis can be undertaken across the sample as a whole.

The first measure of programme eligibility is useful for understanding both implementation and design precision within the existing programme and is the focus of the next section. The second measure provides an insight into the extent of coverage required should the programme expand to cover all eligible people within the sample locations. It is thus an informative measure for future programming scenarios and, therefore, forms the basis of the penultimate section of the paper.

6.3 Performance of targeting implementation: eligibility and selection

When evaluating programme implementation, key questions of interest are: what proportion of the households that meet the programme's eligibility criteria are benefiting from the programme? And, what proportion of beneficiary households are not in fact eligible, i.e. do not meet the programme's eligibility criteria? This assesses how well the programme has managed to identify and enrol its target group, and exclude those who are not part of the target group.

Table 4 provides answers to these questions, indicating coverage as well as inclusion and exclusion errors in implementation.[8] In terms of eligibility, we see that 54 per cent of households overall are eligible (defined as programme eligibility). This disaggregates across targeting mechanism as 47 per cent for CTA and 60 per cent for DR. A striking and encouraging finding is the very high eligibility rate among CTA beneficiaries, with 96 per cent of beneficiaries in CTA areas being CTA eligible. Implementation was less effective in DR areas with 70 per cent of beneficiaries in DR areas being DR eligible. Taken together, this indicates that the coincidence of beneficiaries and eligibility status was reasonably high overall, and the programme was successful in terms of enrolling the intended groups.

Table 4: Implementation errors, by targeting mechanism

	By targeting mechanism		All CTA and DR areas	
	CTA areas	DR areas	%	N
Eligibility rate: % of households that are eligible				
All households	47**	60	54	3,438
HSNP households	96***	70	79	2,047
Coverage rate: % of households covered by HSNP				
All households	40***	56	51	5,108
Eligible households	83	77	79	2,077
Inclusion errors:				
% of beneficiary households that do not meet eligibility criteria	4***	30	21	2,047
Exclusion errors:				
% of eligible households not covered by HSNP	17	23	21	2,077

Notes: (1) The 'N' column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes. (2) Asterisks (*) indicate that an estimate is significantly different to the estimate in the cell to its right: *** = 99%; ** = 95%; * = 90%.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

Finally, and in keeping with the strong coverage results, we see low inclusion and exclusion errors in implementation. CTA out-performs DR as a targeting mechanism for implementation

[8] We are unable to conduct this analysis for the CBT targeting mechanism as there were no pre-defined eligibility criteria for CBT. Instead it was left to the community to decide (within some parameters) who were eligible.

errors, with only 4 per cent inclusion error and 17 per cent exclusion error. In other words, the programme performs well in selecting those who are in fact eligible under the two targeting mechanisms. The explanation for the somewhat better targeting results for CTA is likely due to the fact that in the context of large and often very fluid household sizes, a CTA method will be more accurately implemented as the existence of older people in a household is less changeable over seasons and is more readily observable.

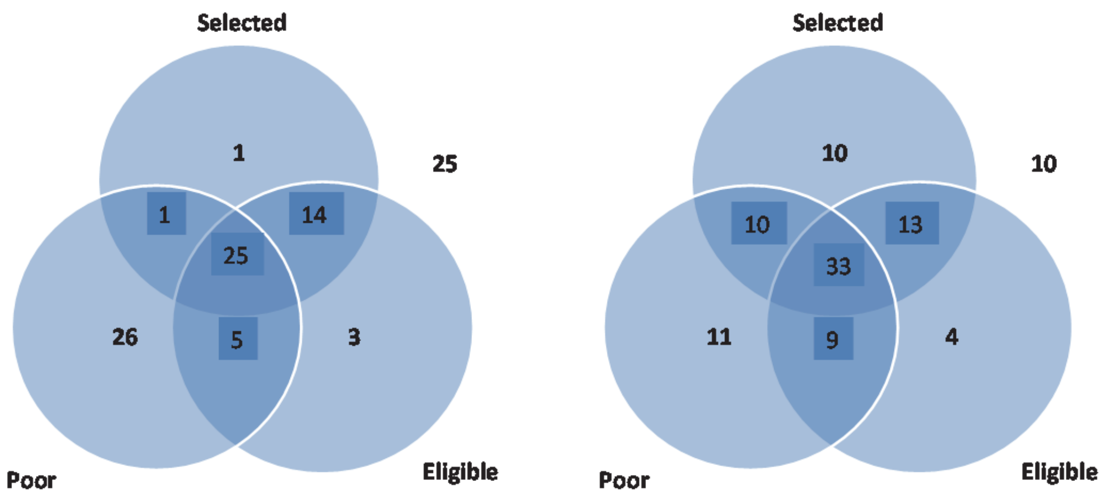
In the Venn diagrams below, the sum of all targeting errors is 100 per cent. For effective targeting, with well designed and implemented targeting criteria, the three circles should overlap as closely as possible, meaning that all 'poor' households are both 'eligible' and 'selected'.

Figure 3 shows that in terms of 'eligibility' and 'selection', 39 per cent of households in the sample were eligible for a social pension and were also selected for the programme, whereas 46 per cent of households were DR eligible and selected.

Figure 3: Targeting accuracy of social pension and dependency ratio targeting approaches.

(a) Percentage of households from total that are income poor, eligible, and/or selected for HSNP under social pension targeting

(b) Percentage of households from total that are income poor, eligible, and/or selected for HSNP under dependency ratio targeting



Note: See Annex 3 for the decision tree format for targeting accuracy of each mechanism.

Source: Authors' illustration.

6.4 Performance of targeting design: eligibility and poverty

Restricting the analysis to assess eligibility only for those sub-locations where the programme implemented specific targeting mechanisms and only for those households who are residents, it is possible to assess the characteristics of eligible households, compared to ineligibles, and in particular their poverty status. Are the eligible—whether they are selected or not—poor?

Disaggregating eligibility by specific targeting mechanism (see Table 5), we see some variation in the ability of the different mechanisms to identify the poor, with 58 per cent of eligible households in CTA areas being poor, which is significantly different to the 50 per cent of ineligibles that are poor in the same areas.[9] For DR areas, 68 per cent of households that are DR eligible are poor. Again, this is significantly different from the 48 per cent of ineligible households that are poor. However, in terms of food security, neither CTA nor DR targeting criteria pinpoint those households that are food insecure, an important finding given the context and objectives of the HSNP. There is a high degree of overlap between CTA and DR eligibility– unsurprisingly, households containing older persons tend to have high dependency ratios and vice versa. In CTA areas 70 per cent of eligible households would also have been eligible under DR targeting, while 48 per cent of eligible households in DR areas are also CTA eligible. This shows that due to the similar population profiles between CTA against DR not much is gained from the comparison, especially as a large proportion of the groups are not actually poor.

Table 5: Characteristics of eligible and ineligible households

	CTA areas		DR areas	
	Eligible households	Ineligible households	Eligible households	Ineligible households
Consumption poverty				
Proportion of households falling below 51% relative poverty line (%)	58**	50	68***	48
Food security				
Proportion of households identified as food insecure (went entire days without eating during worst period) (%)	64	62	72	69
Consumption expenditure				
Mean monthly consumption expenditure per adult equivalent (KES)	1,763***	2,152	1,600***	2,096
Household composition and eligibility overlap				
Proportion of households that (%):				
contain at least one member aged 55+ (CTA eligible)	100***	3	48***	24
are DR eligible	70***	51	100***	3
contain at least 1 orphan	18*	13	25***	15
contain at least 1 chronically ill member	8**	5	7*	4
contain at least 1 disabled member	16*	9	14***	6
contain at least 1 disabled or chronically ill member	22**	13	20***	10

Notes: Asterisks (*) indicate that the eligible household estimate is significantly different to the in eligible household estimate: *** = 99%; ** = 95%; * = 90%.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

Looking at the last four rows of the table, it is clear that, on average, both mechanisms tend to favour households containing orphans and with members that are chronically ill or disabled compared to ineligible households. This is evidenced by the significant differences between the eligible and ineligible columns. This indicates that the eligibility criteria for these mechanisms picks up more than just consumption poverty, but they are able to identify households along a range of other characteristics that might also proxy for different dimensions of poverty.

[9] Note that this is not a comparison of treatment and control, but rather of ineligible and eligible households; therefore, the differences we observe in Table 5 are not unexpected.

So, in contrast to the results in the previous section where we saw a high coincidence of beneficiaries and eligibility criteria from both mechanisms, here we find that eligibility criteria do not correspond as well to poverty status.

6.5 What factors determine selection/eligibility for CBT?

The descriptive statistics in the previous table reveal that 100 per cent of beneficiary CTA households contain at least one member over the age of 54. This suggests that targeting on age criteria has been successfully implemented. Some 69 per cent of households selected for DR targeting have dependency ratios above or equal to 0.6 (0.6 for Turkana and Marsabit; 0.67 for Mandera and Wajir). This is significantly more than for non-selected households.

It is much more difficult to understand the determinants of selection used with CBT, as no specific criteria were set out for identifying this target group. Criteria were 'suggested' as discussed above, but it is instructive, in retrospect, to analyse the key indicators that communities used to identify the poor. To do this, we use probit regressions to test for these determinants and we report the results below. The dependent variable equals 1 if the household was targeted for inclusion in the programme through CBT and 0 for households in CBT areas that were not selected. We specify independent variables that fall into a range of categories: household demographic categories, wealth (livestock, housing, and assets), food-aid receipt, and residency status. In addition, we control for household location by district, as well as running the regressions separately by district to check for consistency of targeting determinants across locations. We apply population weights.

The coefficients presented in Table 6 have been transformed into marginal effects: so, for example, the coefficient 0.029 in column one, associated with household size, means that every additional household member increases the likelihood that the household was selected into CBT by 2.9 percentage points (0.029 x 100). Dummy variables, such as whether the household head is female, are interpreted as 'switching the variable' from 0 to 1. The coefficient for 'fully settled' in column one means that a fully settled household is 19 percentage points more likely to be selected for inclusion in the programme under CBT than a partially settled household, after allowing for other characteristics of the households that are included in the model.

The first column shows the results from the whole sample. One potentially confounding factor in the overall regression results is that a particular characteristic may be more associated with CBT in one district and not another. Since these would tend to cancel each other out, our aggregated results would mask these changes. Accordingly, as a robustness check, we estimate the correlates of selection into CBT by district. These are presented in the last three columns.

The most striking result from the set of regressions above is that, by looking across variable significance between districts, we see that there is no general story to be told about CBT targeting. Clearly, different districts use different criteria. The only consistent results across the overall regression and three district regressions relate to the fully settled variable, with these households being 19.5 per cent more likely to be selected by CBT than partially mobile households. In Turkana, we see the expected signs and significance on chronic illness (that is, a household with at least one member who has a chronic illness is 17 per cent more likely to be selected for inclusion into CBT), the number of orphans in a household, the asset value (the higher the asset value the less likely the household is selected), and on whether the household has a toilet (this is a visible sign of wealth and the results show that households with a toilet in Turkana are 21 per cent less likely to be selected through CBT).

Table 6: Determinants of selection using CBT

VARIABLES	(Overall) CBT	(Turkana) CBT	(Marsabit) CBT	(Mandera) CBT	(Wajir) CBT
HH characteristics					
Has person over 54	-0.001	0.051	-0.008	0.066	-0.118*
	(0.038)	(0.038)	(0.052)	(0.083)	(0.066)
HH size	0.029*	0.017	-0.003	0.106***	-0.008
	(0.015)	(0.014)	(0.014)	(0.026)	(0.009)
Chronic illness	0.025	0.176*	-0.024	0.166*	0.039
	(0.058)	(0.097)	(0.086)	(0.095)	(0.078)
Disability	-0.042	-0.042	-0.009	0.001	-0.027
	(0.042)	(0.053)	(0.028)	(0.083)	(0.041)
Has orphan(s)	-0.053	-0.109**	-0.034	0.350***	-0.053
	(0.036)	(0.048)	(0.064)	(0.095)	(0.050)
Number of orphans	0.020	0.053***	-0.005	-0.096***	0.016
	(0.012)	(0.017)	(0.014)	(0.024)	(0.026)
Female head	0.025	-0.002	0.026	0.055	0.013
	(0.033)	(0.066)	(0.042)	(0.048)	(0.058)
%of 18 to 54 year olds	0.001*	0.002*	0.002	0.008***	-0.000
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Mobility status					
Fully settled	0.192**	0.313**	0.201**	0.505***	-0.145
	(0.094)	(0.159)	(0.095)	(0.099)	(0.115)
Fully mobile	0.208	-0.202	-0.163**	0.308***	0.038
	(0.199)	(0.229)	(0.082)	(0.100)	(0.197)
Wealth and assets					
Has livestock	-0.050	-0.007	-0.057	-0.007	0.150*
	(0.062)	(0.065)	(0.147)	(0.096)	(0.081)
Log (tropical livestock Unit)	-0.051	-0.012	-0.024	-0.327***	0.040
	(0.046)	(0.085)	(0.055)	(0.073)	(0.056)
Log (assets value)	-0.011	-0.051***	0.041***	-0.036***	-0.013
	(0.011)	(0.013)	(0.016)	(0.011)	(0.019)
Subjectively poor	-0.102	-0.313***	0.017	0.088	-0.029
	(0.082)	(0.115)	(0.072)	(0.081)	(0.073)
Poor (below 51% relative poverty line)	-0.095	-0.199**	-0.089	-0.049	0.026
	(0.063)	(0.078)	(0.064)	(0.099)	(0.057)
Household characteristics					
Has a toilet	0.119	-0.212***	0.089	0.314***	-0.013
	(0.113)	(0.078)	(0.099)	(0.095)	(0.066)
Has poor walls	0.174	0.070	0.003	-0.090	0.219**
	(0.118)	(0.154)	(0.098)	(0.108)	(0.111)
Food security and food aid					
Days without eating last hungry season	-0.031	0.038	0.013	-0.197***	-0.071
	(0.047)	(0.069)	(0.067)	(0.044)	(0.095)
Receiving food aid	0.004	0.104	0.025	-0.217*	0.173***
	(0.076)	(0.096)	(0.092)	(0.117)	(0.061)
On school feeding	-0.143*	-0.137	0.075	-0.158**	-0.101
	(0.077)	(0.112)	(0.064)	(0.062)	(0.074)
Observations	5,105	1,313	1,299	1,251	1,242

Notes: (1) The 'N' column denotes the overall sample size. Asterisks (*) indicate that the estimated regression coefficient is statistically significant: *** = 99%; ** = 95%; * = 90%.

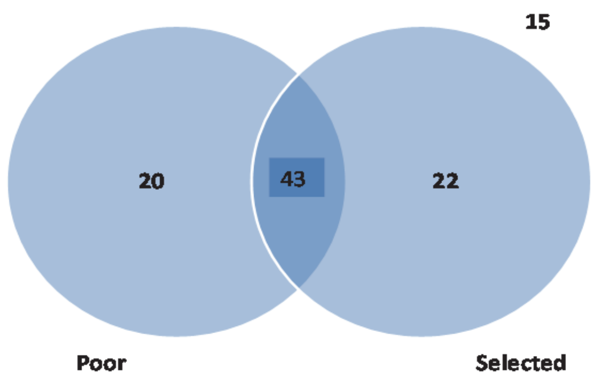
Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

Curiously, in Turkana (but in no other district) we see that if a household is poor (under the consumption 51 per cent coverage rate) or if a household perceives themselves as poor, they are less likely to be selected under CBT. The negative sign on this is worrying, as it indicates that the non-poor are more likely to be targeted under CBT, after adjusting for other factors. In Marsabit, we see an equally unanticipated result on asset values—the likelihood of being selected for CBT increases as asset levels increase. In Marsabit, households that are fully settled are significantly more likely to be selected under CBT than partially settled households and, conversely, households that are mobile are significantly less likely than partially settled households to be selected.

In Mandera we see that the food aid indicators are significant in explaining non-selection using CBT. This is not the case in other districts. So, for instance, if a household receives food aid or is part of a school feeding programme, they are less likely to be selected by CBT. As illustrated above, household size, orphans, and chronic illness are positive and significant.) In Mandera also, we see that livestock and asset ownership is a good predictor of selection through CBT, households with higher livestock and asset values being less likely to be selected for CBT. Strangely, households with more adult working age members are significantly more likely to be selected using CBT in Turkana and Mandera. The main expected predictors in Wajir are food aid receipt and the quality of the walls of a house—with both positively predicting selection through CBT.

The important finding here is that clearly the results are district-specific, indicating that CBT has not been implemented in a consistent manner across the different districts. To some extent, this is an expected feature of the CBT approach, since communities are free to come up with their own criteria. The other clear finding is that fully settled households are much more likely to be included in the programme under CBT, controlling for other factors, suggesting that the process is very liable to exclude semi- and fully-mobile households. Note, however, to the extent that fully settled households are more likely to be officially resident in the sub-location, this may in part reflect the residency requirement of the programme. Figure 4 shows that 43 per cent of households in the sample were selected for a social pension and were also poor.

Figure 4: Community-based targeting



Source: Authors' illustration.

6.6 Comparative effectiveness of targeting mechanisms: are selected households poor?

Inclusion error is the proportion of beneficiary households that are not living below the absolute poverty line, while exclusion error is proportion of households living below the absolute poverty line that are not covered by HSNP. Since the absolute poverty rate in the HSNP districts is very high (85 per cent), and much higher than programme coverage (51 per cent), it is unsurprising that we observe high exclusion errors (46 per cent). To reduce exclusion error would require an increase in programme resources to facilitate an expansion in programme coverage. Although inclusion errors are relatively low (11 per cent), the high poverty rate means that even random or universal targeting would only result in inclusion errors of 15 per cent. Therefore, in order to improve HSNP targeting effectiveness, efforts should be made to reduce inclusion errors.

Table 7 compares the proportion of households in poverty and food insecure by beneficiary status across the three different targeting mechanisms.

Table 7: Comparative targeting performance by mechanism

	Targeting mechanism		
	CBT areas	CTA areas	DR areas
Consumption poverty			
Proportion of households falling below 51% relative poverty line			
Beneficiary households (%)	51**	58**	63
Non-beneficiary households (%)	34	51	54
Food security			
Proportion of households identified as food insecure (went entire days without eating during worst period)			
Beneficiary households	65**	63	71
Non-beneficiary households	47	63	70

Notes: Asterisks (*) indicate that the beneficiary household estimate is significantly different to the non-beneficiary household estimate: *** = 99%; ** = 95%; * = 90%.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

As discussed above, because poverty and coverage rates vary across the CBT, CTA, and DR areas, a simple comparison of poverty rates among HSNP households cannot be used to compare the comparative targeting effectiveness of the three mechanisms. Instead, two more nuanced measures are used to compare targeting performance of the three mechanisms: the ratio of beneficiary and non-beneficiary poverty rates; and the CGH index (Coady et al. 2004). The ratio of beneficiary and non-beneficiary poverty rates gives an alternative measure of targeting effectiveness to the CGH index. For both measures, higher values indicate a better result in terms of targeting beneficiaries as compared to non-beneficiaries.

Table 8 shows that, on both measures, and using the two different poverty definitions (consumption poverty and food security), CBT comes out as performing best, followed by CTA and then DR (see the 'actual' columns). These results resonate with the assertions of the emerging literature on targeting which credits CBT for its ability to address information asymmetries affecting most other targeting methods. This is because community groups are perceived as having better information about the needs and poverty status of other community members, although risks around capture by elites also need to be recognised. This finding also reflects the fact that the DR and CTA targeting criteria were not chosen with any explicit link to objective poverty measures. Furthermore, as discussed above, DR was not implemented perfectly which could undermine its targeting performance.

In order to assess how CTA and DR would have compared if both had been implemented perfectly it is again necessary to use the poverty ratio and CGH measures. These are also presented in Table 8, in the 'predicted' columns. In terms of consumption poverty, the estimates show that DR would have performed almost as well as CBT if it had been implemented with 100 per cent accuracy. This implies that the implementation errors in DR targeting have drastically undermined the targeting effectiveness of this mechanism. Further analysis (not presented) reveals that this is driven by ineligible beneficiaries that are somewhat better off being covered by the programme in place of poorer eligible non-beneficiaries in the DR areas.

In contrast, even with 100 per cent implementation accuracy CTA targeting would not perform well from a consumption poverty targeting perspective. This is because in the HSNP districts old age does not appear to be strongly associated with poverty. However, since CTA targeting, unlike DR, was implemented effectively (96 per cent of beneficiaries fulfilled the eligibility criteria) the actual 'net' effectiveness of CTA and DR targeting was similar.

Table 8. Comparative targeting performance by mechanism: predicted versus actual

	CBT areas	CTA areas	DR areas		
	Actual	Predicted	Actual	Predicted	Actual
Consumption poverty					
Ratio of poverty rates: beneficiaries vs non-beneficiaries	1.50	1.17	1.15	1.42	1.16
CGH index: %of beneficiaries that are poor/ poverty rate	1.21	1.09	1.08	1.13	1.05
Food security					
Ratio of poverty rates: beneficiaries vs non-beneficiaries	1.37	1.04	1.00	1.05	1.01
CGH index: %of beneficiaries that are poor/ poverty rate	1.17	1.02	1.00	1.02	1.00

Notes: (1) Poverty rate and CGH indices are based on a 51% relative poverty line. (2) Predicted targeting performance is based on the poverty rates among eligible households, i.e. it is the targeting performance assuming 100% targeting accuracy whereby all eligible households are selected and all selected households are eligible.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

All mechanisms have different eligibility criteria. However, as described in the background to the HSNP, the purpose of the three mechanisms was to understand which of the three was a more accurate proxy for targeting the poor. We see that, while all three mechanisms are pro-poor, targeting using CBT was relatively more effective than either CTA or DR. The ambition of each is to target the poorest households. But CBT is highly subjective, as shown in the previous analysis. We see that there is a trade-off between how effectively the targeting mechanism is applied and poverty performance. That is, we find that CBT does best on poverty, but if DR had been applied as it was intended then DR would have performed almost as well as CBT. This has implications for the choice of targeting mechanism. Are there better ways to target such that a higher proportion of the poor are selected for the programme? In the next section we investigate this and show that it might be worth substituting targeting mechanisms that are difficult to implement for different mechanisms that give a better 'on the ground' outcome.

7 Alternative ways to target

It is possible to simulate the application of actual and hypothetical targeting mechanisms across the data set as a whole, disregarding the mechanisms that were implemented in practice in each area. This has the benefit of averaging out any differences between the areas in which the different mechanisms were implemented that may have occurred by chance. For simplicity, this analysis also removes the residency requirement imposed by the programme. Using this approach indicates that 59 per cent of households eligible for a CTA are among the poorest (under a 51 per cent relative poverty line), and 58 per cent of households eligible for a transfer under DR are amongst the poorest. This is similar for the lower relative poverty line, suggesting that differences between the two mechanisms are not intrinsic to the criteria used, but rather to their implementation and the populations in which they have been applied.

When we tabulate CTA-eligible households by DR-eligible households, we find that 29 per cent of the sample would qualify for both types of transfer based on the eligibility criteria alone (having at least one household member over 54 years of age and a dependency ratio above 0.6). Disaggregated further, we see that 70 per cent of those eligible for a CTA would also be eligible for a transfer under the DR criteria. There is, therefore, considerable overlap between the two categories of household.

Given the results above, two points stand out: (1) the targeting mechanisms, and their associated eligibility criteria, do not perform strongly on identifying the poor; and (2) implementation has been successful with respect to selecting the eligible households. Therefore, the obvious question becomes: is there an alternative targeting mechanism that would better identify the poor? In this section, we specify a range of alternative targeting criteria, as possible proxies for poverty or as alternative bases for cash transfers, and compare them against DR and CTA eligibility criteria for the population as a whole. Specifically, we simulate four alternative criteria:

1. a household is eligible if it contains at least one orphan;
2. a household is eligible if it contains at least one member who is chronically ill or disabled;
3. a household is eligible if it contains at least one child under the age of 6 years-old (this may be a criterion used under a child benefit-type programme); and
4. a household is eligible if it satisfies a threshold level under a proxy means test (PMT).

All the variables used to specify the PMT were relatively easy to collect and together are likely to predict the poverty status of a household. We used a total of 17 variables to construct the PMT measure. These are listed in Annex 2. Most of these variables reflect the community criteria used to establish eligibility during the CBT selection process.

This is an attempt to construct a PMT measure. Of course, we could change the variables and/or reduce the number of indicators included. These simulations and further analysis would refine the targeting criteria for future programming. We have set the PMT threshold for eligibility to match the programme's 51 per cent coverage rate; in other words, the bottom 51 per cent of households, ranked according to their PMT scores, are classified as eligible.

Table 9 provides the results of these different simulations. Looking along the top row, we see that, for our total sample, 59 per cent of CTA-eligible households and 58 per cent of DR-eligible households are poor. The rates are almost identical when we use orphans and illness/ disability as proxies for poverty.

We also run a simulation for households that have at least one child under six. In terms of a proxy for poverty we see that, comparatively, this mechanism performs the weakest, with only 48 per cent of eligible households being poor by our 51 per cent coverage standard. The demographic characteristics of households selected on this basis are also quite different. This does not appear to be a good proxy for targeting poor households, although it is recognized that there may be other reasons for targeting transfers at households with small children.

These simulation results suggest that a simple PMT approach would significantly out-perform the CTA and DR criteria, and the actual CBT performance, in identifying poorer households.

Table 9. Alternative targeting scenarios

Indicator	By eligibility status											
	CTA		DR		>= one orphan		>= one ill or disabled member		PMT-eligible		Child<=5	
	Elig	Inelig	Elig	Inelig	Elig	Inelig	Elig	Inelig	Elig	Inelig	Elig	Inelig
%of HHs that are amongst the *poorest* (bottom 51%)	59***	44	58***	40	58***	49	58***	49	74***	25	48**	56
Coady-Grolsch-Hoddinott targeting index score (based on 51% relative poverty line)	1.16		1.14		1.14		1.14		1.45		0.94	
%HHs that are subjectively poor	71***	64	69**	63	72***	65	72*	66	76***	56	67	69
Mean monthly consumption expenditure	1793***	2265	1814***	2423	1825***	2130	1929**	2102	1490***	2682	2126	1989
%of HHs in quintile 1 (poorest)	25***	17	24***	14	25***	19	21	20	35***	4	17***	24
%of HHs in quintile 2	21	19	23***	16	21	20	23*	19	27***	12	20	20
%of HHs in quintile 3	23**	18	21	19	23**	19	22	20	21	19	19	21
%of HHs in quintile 4	18*	21	19	21	17*	21	20	20	12***	28	22**	17
%of HHs in quintile 5	12***	25	12***	30	14***	21	14***	21	5***	36	21**	18
%of HHs contain >= 1 member aged 55+	100	0	50***	28	45**	39	59***	37	49***	32	27***	62
%of HHs that are DR eligible	71***	48	100	0	68***	55	71***	55	72***	42	61***	51
%of HHs with >= one orphan	21**	17	22***	14	100	0	23***	18	22***	15	14***	26
%of HHs with >= one chronically ill member	9***	5	8***	4	7	7	41***	0	5***	8	5***	9
%of HHs with >= one disabled member	16***	7	13***	7	13**	10	65***	0	13***	8	8***	14
%of HHs with >=1 disabled/ ill member	23***	11	20***	11	20**	15	100	0	17	15	13***	21

Notes: (1) The 'N' column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes. (2) Asterisks (*) indicate that an estimate is significantly different to the relevant comparator, as explained in Section 1 of the baseline report: *** = 99%; ** = 95%; * = 90%.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

The proportion of households that are eligible according to the PMT criteria is 74 per cent, considering the 51 per cent relative poverty line. Furthermore, the PMT measure is better able to identify the correct 'gradient' across consumption quintiles in terms of numbers and mean consumption. That is, we see that the percentage of PMT-eligible households declines as we move up the wealth quintiles, from 35 per cent in quintile 1 to just 5 per cent in quintile 5. This is what we would hope for from a targeting mechanism that aims to target the poorest households. However, this is a preliminary analysis and, like DR, PMT approaches can be difficult to implement effectively in practice. It is also an 'in-sample' prediction, and the same coefficients applied to another dataset (e.g. the actual information collected during the targeting process) would not be expected to have such a high predictive accuracy. So, rather than interpreting these results as recommending a PMT approach is best, it is suggested instead that the possible role of some type of explicitly poverty-focussed targeting approach for later phases should be further investigated, noting that this would require additional simulation work and should be complemented by qualitative analysis.

8 Conclusions

Even in very poor communities such as the arid and semi-arid lands of northern Kenya, income inequality is sufficiently high that targeting social transfers on the poorest individuals and households can be justified as a means of maximizing their poverty-reducing impacts. However, in contexts where the poverty headcount is very high, geographic targeting might be the simplest, cheapest, and most effective way of reaching large numbers of poor people. Given an 85 per cent poverty headcount, blanket coverage of the HSNP districts would incur an inclusion error of 15 per cent and an exclusion error of 0 per cent. The costs saved by excluding the non-poor in this context are limited, and might be exceeded by the costs of identifying them. On the other hand, if budget constraints and programme objectives necessitate targeting the 'poorest of the poor', many options exist. The HSNP deployed three targeting mechanisms in an effort to determine which mechanism was most effective at identifying and reaching absolutely poor and food insecure households in northern Kenya.

The analysis presented in this paper allows us to draw the following conclusions. First, in terms of overall effectiveness, HSNP targeting is pro-poor, but only mildly so. Beneficiary households are 30 per cent (13 percentage points) more likely to be among the poorest (bottom 51 per cent) as compared to non-beneficiaries (57 per cent vs. 44 per cent). In terms of food security, beneficiaries are only 16 per cent (9 percentage points) more likely to be food insecure compared to non-beneficiaries. Secondly, across the three targeting mechanisms trialled, community-based targeting (CBT) performed best: it was most effective at identifying the poorest and food insecure households. Importantly, given the increasing attention being paid to the social impacts of social transfers, CBT was also more likely to be perceived as a fair process by households and communities.

However, CBT also displayed some of the weaknesses that have been associated with community-based targeting in other contexts. Being based on relative rankings rather than absolute poverty measures, it is insensitive to variations in poverty levels across communities. In northern Kenya, poverty and food insecurity vary substantially across districts, but a quota was applied uniformly across all HSNP programme areas. Also, CBT depends on full participation of all community members and the avoidance of 'elite capture', which can distort targeting outcomes. In one HSNP district the evidence suggests that the poorest households were less likely to be selected, implying that the targeting process was indeed captured by local elites. This suggests a need to test the targeting effectiveness of 'moderated' CBT (where programme staff observe and guide communities to select households that conform to the target population) against ‘unmoderated’ CBT (where communities exercise their discretion over who should be selected for a programme).

This leads to a third broad conclusion. Variations in targeting performance reflect variations in the way each targeting mechanism was implemented in each locality. Implementation matters. As a rule, the more complicated the targeting criteria the worse the targeting performance. For instance, in HSNP areas the dependency ratio is a better proxy for consumption poverty than whether the household has an older person, so DR targeting should have performed better than CTA targeting. In fact, DR targeting was undermined by implementation errors, and it performed worst of all three mechanisms in terms of inclusion and exclusion errors. Conversely, targeting households with older persons is the simplest mechanism, so CTA targeting was implemented most effectively, but it did not perform well in terms of identifying the poorest households.

This paper also presented a simulation analysis that assessed programme coverage, the comparative characteristics of eligible and ineligible households, and targeting effectiveness under six alternative targeting options. The PMT approach significantly outperforms all other simulated targeting approaches and would also be expected to outperform the actual targeting performance of CBT (the best performing of the three HSNP mechanisms). Under PMT targeting, three times as many beneficiary households would be poor as non-beneficiaries (76 per cent and 26 per cent respectively). However, PMT approaches can be difficult to implement effectively in practice, and it is, therefore, informative to consider the degree to which implementation problems undermined the targeting effectiveness of DR. The simulation analysis also revealed that targeting households containing children would be the weakest mechanism—only 48 per cent of eligible households would be poor compared to 55 per cent of ineligibles—although it is recognized that there may be other reasons for targeting transfers at households with young children.

Several broader lessons for targeting social transfers in low-income countries can be drawn out of this experience in northern Kenya. Firstly, while targeting older persons or children might be justified on equity grounds—the right of every person to income security in old age, or the policy imperative to tackle child poverty—categorical targeting of older persons or children is not a robust proxy for poverty in many (perhaps most) contexts. Categorical targeting should be based on available survey data confirming that the indicator selected as a proxy for poverty is in fact strongly correlated with poverty in the specific local context.

Secondly, in low-income countries with weak administrative capacity, effective implementation of targeting mechanisms is likely to be challenging, and simpler mechanisms should therefore be preferred over complex approaches. Thirdly, given that there is invariably a trade-off between targeting complexity and targeting accuracy, some level of inclusion and exclusion errors must be expected and tolerated, especially if simpler (but cruder) mechanisms are preferred to ensure effective implementation. The key policy choice for programme designers is which targeting error to weight more highly—inclusion or exclusion.

References

Alderman, H. (2002). ‘Do Local Officials Know Something We Don't? Decentralization of Targeted Transfers in Albania’. *Journal of Public Economics*, 83(3): 375-404.

Barrientos, A. and M. Niño-Zarazúa (2011). ‘Social Transfers and Chronic Poverty: Objectives, Design, Reach and Impact’. Manchester: Chronic Poverty Research Centre.

Coady, D., M. Grosh, and J. Hoddinott (2004). 'Targeting Outcomes Redux'. *World Bank Research Observer*, 19 (1): 61-85.

Conning, J. and M. Kevane (2002). 'Community-based Targeting Mechanisms for Social Safety Nets: A Critical Review'. *World Development*, 30(3): 375-394.

Cornia, G. and F. Stewart (1993). 'Two Errors of Targeting'. *Journal of International Development*, 5 (5): 459-496.

Department for International Development (DFID) (2011). 'Cash Transfers evidence paper'. London: DFID.

Ellis, F., S. Devereux and P. White (2009). 'Social Protection in Africa'. Cheltenham: Edward Elgar.

Falkingham, J. and C. Namazie (2002). 'Measuring Health and Poverty: A Review of Approaches to Identifying the Poor'. London: DFID Health Systems Resource Centre.

Fiszbein, A., N. Schady et al. (2009). 'Conditional Cash Transfers: Reducing Present and Future Poverty'. Washington DC: World Bank.

Grosh, M. and P Leite (2009). 'Defining Eligibility for Social Pensions: A View from a Social Assistance Perspective', Chapter 12 in Holzmann, R., D. Robalino, and N. Takayama (editors) 'Closing the Coverage Gap: The Role of Social Pensions and Other Retirement Income Transfers'. Washington DC: World Bank.

Handa, S., S. Devereux, D. and Webb (editors) (2011). 'Social Protection for Africa's Children'. London: Routledge.

Hanlon, J., A. Barrientos, and D. Hulme (2010). 'Just Give Money to the Poor: The Development Revolution from the Global South'. Kumarian Press.

HLPE (2012). *Social Protection for Food Security*. A Report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security. Rome: CFS.

Hodges, A., A. Dufay, K. Dashdorj, K. Jong, and U. Budragchaa (2007). 'Child Benefits and Poverty Reduction: Evidence from Mongolia's Child Money Programme'. New York: UNICEF.

Mansuri, G. and V. Rao (2004). 'Community-Based and -Driven Development: A Critical Review'. Policy Research Working Paper 3209, Washington DC: World Bank.

Minot N. and B. Baulch (2002). 'The Spatial Distribution of Poverty in Vietnam and the Potential for Targeting'. MSSD Discussion Paper No. 42, Washington DC: IFPRI.

Ravallion, M. (2007). 'How Relevant Is Targeting to the Success of an Antipoverty Program?'. Policy Research Working Paper 4385, Washington DC: World Bank.

Samson, M., I. van Niekerk, and K. MacQuene (2006). 'Designing and Implementing Social Transfer Programmes'. Cape Town: Economic Policy Research Institute.

Tesliuc, E. (2004). 'Mitigating Social Risks in Kyrgyz Republic'. Washington DC: World Bank.

ANNEX 1

Table A1: Household welfare by consumption expenditure quintile	Consumption expenditure quintile					All HSNP evaluation areas	
	Q1 (poorest)	Q2	Q3	Q4	Q5	Estimate	N
Food, health, and education expenditure							
Mean share of food expenditure in total monthly household expenditure (KES)	83***	80***	77	77*	73***	78	5,105
Mean monthly household health expenditure (KES)	58***	72***	85***	138	277***	126	5,105
Mean monthly household education expenditure (KES)	47***	137**	198	293**	415***	218	5,105
Household assets and livestock ownership							
Mean value of all household assets owned by household (KES)	9095**	12478*	15230*	27226	66917**	26184	5,105
Mean value of productive assets owned by household (KES)	718***	1548***	2370	3011**	4042**	2337	5,106
Proportion of households owning livestock (%)	78	77**	75**	70	53***	70	5,106
Mean tropical livestock units owned currently (for households owning livestock) (TLUs)	1.0***	1.4*	2.0	2.2**	2.8***	1.8	3,778
Education and health status							
Proportion of adults aged 18+ that are literate (%)	14***	19	18**	27**	35***	22	12,611
Proportion of children aged 6-17 that are currently attending school (excluding *duksi* and *madrasah*) (%)	40***	49	51	62***	68***	53	10,540
Proportion of people ill/injured in the past 3 months (excl. chronic illness) (%)	34***	25	20	20	13***	23	28,065
Household dwelling characteristics							
Proportion of households with a sand/earth floor (%)	97**	94**	94***	85	70***	88	5,106
Proportion of households with walls made of natural materials (%)	98***	94**	93***	83*	66***	87	5,106
Subjective poverty							
Proportion of households reporting that they are 'struggling' (%)	68***	60	65***	57	39***	58	5,106
Proportion of households reporting that they are 'unable to meet household needs' (%)	20***	14***	8**	6***	3***	10	5,106

Notes: (1) The 'N' column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes. (2) Asterisks (*) indicate that a quintile estimate is significantly different to the pooled mean across the other four quintiles: *** = 99%; ** = 95%; * = 90%. (3) Consumption quintiles are defined according to the distribution of consumption expenditure over the study population such that each quintile contains 20% of the population.

Source: HSNP M&E Baseline Evaluation Survey, Sep 2009-Oct 2010.

ANNEX 2

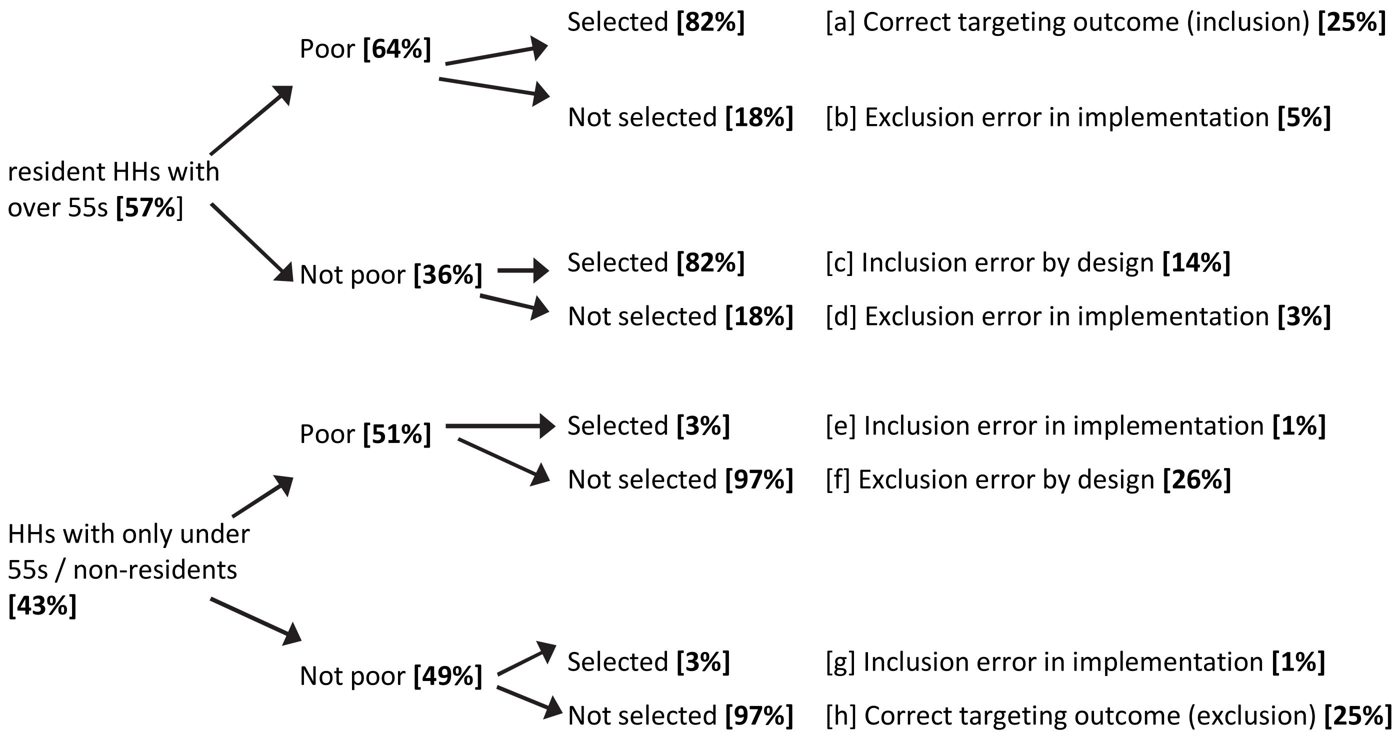
1. Whether a household receives food aid (Y/N)
2. Whether a household is part of a school feeding programme (Y/N)
3. Whether the household has a toilet in the home (Y/N)
4. The number of rooms in the house
5. An indicator of whether the walls of the house are poor quality
6. Whether the household has at least one disabled member (Y/N)
7. Whether the household has at least one chronically ill member (Y/N)
8. Whether the household owns livestock (Y/N)
9. Household size (number of members)
10. The age of the head of household
11. The number of orphans in the household
12. Whether the head is a female (Y/N)
13. Whether the head is a child (Y/N)
14. Whether the household has any members over 54 years-old
15. The DR of the household
16. The settlement/residency status of the households (fully settled, partially settled, fully mobile)
17. The district where the household is located

ANNEX 3: Decision trees

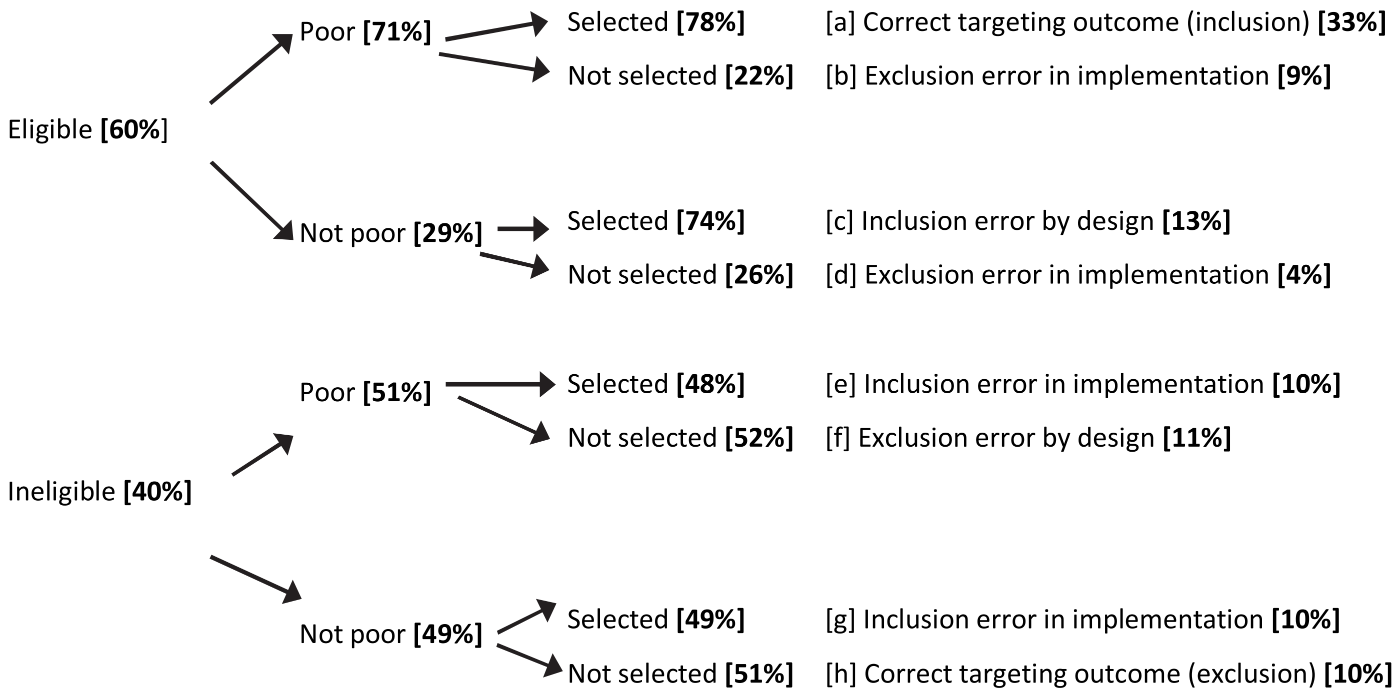
Stats for the targeting tree diagram:

Social Pension:

Poor = in bottom national decile (amongst poorest 10% of Kenyan households)



Dependency Ratio:



CBT:

Poor **[63%]** → Selected **[69%]** [i] Correct targeting outcome (inclusion) **[43%]**
→ Not selected **[31%]** [ii] Exclusion error in implementation **[20%]**

Not poor **[37%]** → Selected **[61%]** [iii] Inclusion error in implementation **[22%]**
→ Not selected **[39%]** [iv] Correct targeting outcome (exclusion) **[14%]**