Journal of International Development
J. Int. Dev. **27**, 1521–1545 (2015)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jid.3186

TARGETING SOCIAL TRANSFER PROGRAMMES: COMPARING DESIGN AND IMPLEMENTATION ERRORS ACROSS ALTERNATIVE MECHANISMS

RACHEL SABATES-WHEELER[1]*, ALEX HURRELL[2] and STEPHEN DEVEREUX[1]
[1]*Centre for Social Protection at the Institute of Development Studies, Brighton, UK*
[2]*Oxford Policy Management, Oxford, UK*

Abstract: An innovative cash transfer programme in northern Kenya is the first of its kind to trial three targeting mechanisms to learn about which approach is most effective at identifying the poorest households while minimising inclusion and exclusion errors. Analysing data collected through a randomised controlled trial, 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. © 2015 UNU-WIDER. *Journal of International Development* published by John Wiley & Sons, Ltd.

Keywords: targeting; social transfers; errors; kenya
JEL Classification : I38, I32, D60

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. In emergency contexts, or in localities where poverty is widespread, the time and budgetary costs of identifying and excluding the non-poor might make poverty targeting inappropriate (Ellis, 2012). Conversely, crude targeting or no targeting (universal coverage) can be extremely wasteful of scarce resources (Devereux *et al*., 2015).

*Correspondence to: Rachel Sabates-Wheeler, Centre for Social Protection at the Institute of Development Studies, Brighton, UK.
E-mail: r.sabates-wheeler@ids.ac.uk

These trade-offs inform not only the decision about *whether* to target but also the choice of *how* to target. 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 performed 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 (Coady *et al*., 2004; Barrientos & Niño-Zarazúa, 2011).

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 individuals in Mandera, Marsabit, Turkana and Wajir counties. 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, because 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 is consumption expenditure, but this is generally not feasible given the vast quantities of detailed information required from all households in a targeted community and the associated costs.

Alternative approaches involve identifying proxies for poverty, which are then used as measures for targeting. Such approaches usually rely on statistical analysis of the correlates of objective measures of poverty. Due to a 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 (2008–2012), three targeting mechanisms were trialled: (i) community-based targeting; (ii) households with high dependency ratios; and (iii) older people (55 years or above). This innovative experimental design provides a rare opportunity to compare outcomes across different targeting mechanisms within the same programme. More commonly, in the targeting literature, numerous assumptions need to be made to allow researchers to compare efficiency and effectiveness results of different targeting mechanisms across different programmes, contexts, populations and institutional set-ups. This leads to inaccuracies and bias in results interpretation.

Using baseline survey data from a randomised control trial evaluation of the HSNP, we provide, in this paper, the first statistically robust assessment (to our knowledge) of the targeting effectiveness of different mechanisms within the same programme, allowing comparisons for the purpose of more robust generalisation.[1] The next section provides a

[1]The baseline survey was conducted after the beneficiaries were identified but before they received any cash.

review of targeting effectiveness. Then follows a description of targeting within the HSNP and a review of the evaluation methodology and data.

We lay out the poverty context of northern Kenya in Section 5.1. The rest of Section 5 is structured around the comparative performance of the three targeting mechanisms in terms of the following: (5.2) the combined performance of the mechanisms in identifying the poorest households in the evaluation areas (coverage, overall and of poorest);[2] (5.3) whether the eligible households are actually selected for the programme (implementation problems); (5.4) whether the eligible are, in fact, the poorest households (inclusion and exclusion by design); (5.5) the factors that determine selection under community-based targeting (CBT; probit regression analysis); and (5.6) the comparative targeting performance of the three mechanisms (coverage of poorest households by mechanism). In Section 6, we simulate a range of alternative targeting criteria as performance benchmarks for the three mechanisms actually used by the programme. We end with a discussion of the results.

2 TARGETING PERFORMANCE

Most social transfer programmes, including the HSNP, attempt to transfer resources to the poorest members of a population. So the measure of a targeting mechanism's effectiveness is how accurately it identifies poor people. Targeted social transfer programmes are susceptible to two types of errors: (i) *inclusion error* can be quantified as the proportion of a programme's beneficiaries who receive transfers despite not being poor; and (ii) *exclusion error* can be quantified as the proportion of people in poverty who are omitted from a social transfer programme (Cornia & Stewart, 1993).[3]

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

2.1 Targeting Errors by Design

A common design challenge is a binding budget constraint that 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 rather 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. Design errors will also arise where proxy measures are used to identify poor households or where targeting criteria are selected with no explicit link to objectively quantifiable poverty measures, such as consumption, income or assets. For example, even though it is well known that old age is not a perfect proxy for poverty, targeting may be performed on the basis of old age for logistical, social or political

[2]The poorest households are defined as those in the bottom 51 per cent of the consumption expenditure distribution. This 51 per cent relative poverty line was set in line with programme coverage in the evaluation areas (51 per cent).

[3]'Inclusion error' is also known as leakage, vertical inefficiency, type-1 error or specificity. 'Exclusion error' is also known as under-coverage, horizontal inefficiency, type-2 error or sensitivity.

considerations. Even though some older persons will not be poor, this ‘inclusion by design’ is balanced by the fact that many older persons *are* poor and will, therefore, be reached by the programme.

2.2 Targeting Errors in Implementation

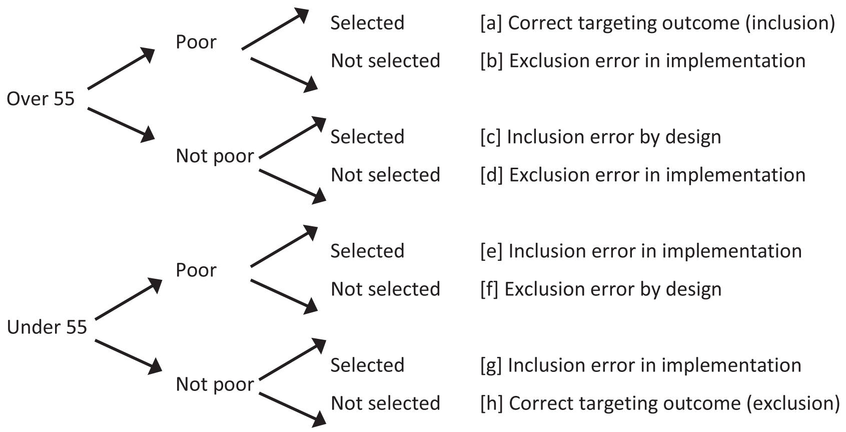
Inclusion errors in implementation can occur for numerous reasons. Behavioural incentives for applicants to misrepresent their true status in order to satisfy the eligibility criteria—say, by claiming to be poor when they are not—is one example. Castañeda (2005), Pritchett (2005) and others suggest targeting on ‘observable’ characteristics to minimise this problem. Poor design choices can also lead to errors in implementation, such as choosing inaccurate proxy variables (Hodges *et al*., 2007), or difficulties in verifying the actual values of proxies (Kidd & Wylde, 2011). 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 marginalised members from a CBT 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: (i) the ‘implementation accuracy’ of the programme in reaching the eligible and excluding the non-eligible; and (ii) 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 years while ‘exclusion error in implementation’ is all non-recipients over 55 years. 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 years in HSNP communities who are poor (i.e. ‘exclusion error by design’) and the percentage of people over 55 years 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 years [a] and exclusion of non-poor people under 55 years [h]—are unambiguously correct targeting outcomes. If any poor or non-poor people over 55 years are not selected into the programme, these are exclusion errors in implementation ([b] and [d]). If any poor or non-poor people under 55 years are selected into the programme, these are inclusion errors in implementation ([e] and [g]). If any people over 55 years who are selected into the programme are non-poor, this is inclusion error by design [c]. Finally, if people under 55 years who are (correctly) not selected into the programme are poor, this is exclusion error by design (outcome [f]). From the targeting outcomes, [e] (a poor person incorrectly included) and [d] (a non-poor person incorrectly excluded) are more preferred than the other four ‘incorrect’ outcomes.

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 errors in HSNP targeting design and implementation are reported in this paper.[4]

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



Source: Authors' illustration

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

3 TARGETING ON THE HUNGER SAFETY NET PROGRAMME

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 CBT. 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 the poorest households in evaluation areas.

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 the poorest households?

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 randomisation underpins the evaluation methodology, which is set out in the subsequent discussion. There was no retargeting during Phase 1, but individuals (social pensioners) and households (under dependency ratio and CBT) 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).

(2) *Categorical targeting by age (CTA)*: All individuals aged 55 years 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 policymakers, 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 poor individuals. 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. Because the programme aims to address extreme 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. The typical approach to CBT was for the community elders, facilitated by HSNP programme staff, to list all eligible beneficiaries during a baraza.[5] These households were ranked from most poor to least poor. Households were usually registered on paper and the data entered later into the management information system. The programme administration then printed the list of beneficiaries from the management information system for validation by the community, and then a proportional quota was applied to each sub-location's ranked list of households. Beneficiaries were officially enrolled into the programme once the community verification process was completed. But because the community household population data was often inaccurate and because extreme poverty levels may vary substantially between sub-locations, the standardised quota almost certainly overestimated the actual poverty rate (i.e. inclusion error) in some sub-locations and under-estimated it (i.e.

[5]A baraza is a community meeting. Barazas were convened at various stages of the targeting process under the different targeting mechanisms.

exclusion error) in others. Problems of consistency in target group identification across communities are a generic problem with CBT (Conning & Kevane, 2002; Mansuri & Rao, 2004).

Targeting efficiency can, of course, be compromised by various challenges that are specific to each approach. For example, categorical targeting based on age may exclude children without birth certificates or older persons who do not have identification documents and so cannot prove their age. Similarly, DR targeting may overlook a poor household that has sent some children to other households, to spread the burden and costs of raising them, thereby lowering the household's dependency ratio. Finally, community-based targeting is prone to bias against minorities (who are deliberately excluded), recent arrivals in the community (who might be unknown to community leaders) and households that are isolated (who have little voice in community matters). These considerations must be taken into account when evaluating the efficiency of any targeting mechanism.

4 EVALUATION DESIGN AND DATA

The programme was implemented across the counties 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 secure sub-locations where the HSNP could potentially operate.[6] The targeting methodology to be implemented in each of the 48 evaluation sub-locations was allocated at random.

The targeting analysis presented in this paper is based on the baseline survey data, which is representative of all households living in these evaluation sub-locations. In each of the sub-locations, beneficiaries were selected for the programme according to the relevant mechanism. Once this was carried out, half the evaluation sub-locations were randomly assigned to be 'treatment' areas and received their first 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 receiving transfers after 2 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.

The work presented here draws on (i) the baseline survey of a household panel survey conducted on an annual basis (baseline, year 1 follow-up and year 2 follow-up) covering 5108 randomly selected households in the 48 evaluation sub-locations, and (ii) quantitative community interviews conducted with a group of 10–12 community members on an annual basis (baseline, year 1 follow-up and year 2 follow-up) in the same 48 sub-locations.

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

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

each sub-location. These settlements were stratified into three different types, and one settlement of each type was sampled. 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 programme areas. This comparison of beneficiary and non-beneficiary households is the basis of the targeting analysis.

The household survey instrument was substantial in that it collected an extensive set of variables including consumption, assets, household structure, education, shocks, social networks, employment and production. The baseline survey was carried out between September 2009 and October 2010.

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. All standard errors have been calculated accounting for the clustered survey design.[7]

5 FINDINGS

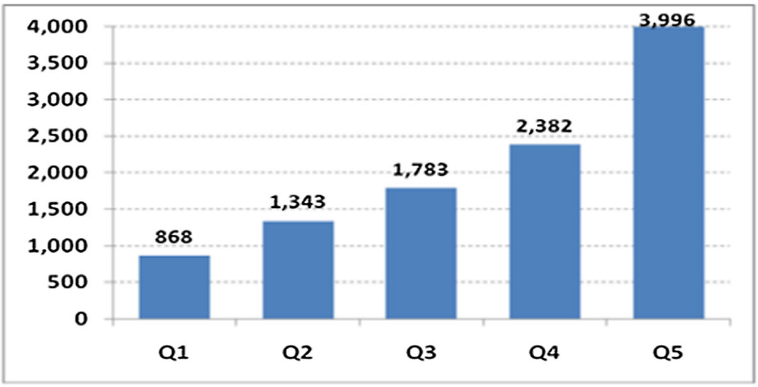
5.1 Characteristics of the Sample: Poverty, Consumption and Demography

While absolute poverty rates in the HSNP counties are extremely high according to national data, we find 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 (Kenyan Shilling (KES) 3996 vs. KES 868), which indicates an appreciable degree of income inequality within the study population (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. For evaluating targeting in such situations, we use a relative poverty line, described in the following section.

Our analysis shows that, while not a perfect proxy for welfare, consumption expenditure is highly correlated with many key dimensions of household well-being (see Appendix 1). 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 assets, have lower adult literacy and school enrolment rates, are more likely to have been ill or injured in the past 3 months, and have poorer quality housing. Furthermore, subjective poverty rates are significantly higher among households in poorer quintiles.

[7]For a more detailed explanation of the survey method, evaluation plan and analysis, please see Hurrell and Sabates-Wheeler, 2011.

DOI: 10.1002/jid



Source: HSNP Monitoring and Evaluation Baseline Evaluation Survey, Sep 2009-Oct 2010.

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

5.2 Programme Coverage, Targeting Mechanisms and Poverty in the HSNP districts

Table 1 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 years 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 1 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.[8] The variations in coverage rates by targeting mechanism do not reflect variations in poverty and food security across the CBT, CTA and DR areas. Poverty and food insecurity are lowest in CBT areas (42 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 intentionally 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.

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

[8]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.

Table 1. 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	5108
Consumption poverty					
Proportion of households falling below 51% relative poverty line	42*	54	60	51	5106
Food security					
Proportion of households identified as food insecure (went entire days without eating during worst period)	55	63	71*	63	5106

The 'N' column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes
Asterisks (*) indicate that the targeting mechanism estimate is significantly different to the pooled mean across the other two mechanisms:
***99%;
**95%;
*90%
HSNP, Hunger Safety Net Programme; CBT, community-based targeting; CTA, categorical targeting by age; DR, dependency ratio.
Source: HSNP M&E Baseline Evaluation Survey, September 2009 to October 2010.

Table 2 shows comparative levels of relative 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 with non-beneficiary households (57 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 with non-beneficiaries.

Does this represent effective targeting? In order to understand how this compares with the targeting effectiveness of other cash transfer programmes around the world, we calculated the Coady–Grosh–Hoddinott (CGH) index and this is also presented in Table 2. We chose the CGH index as (i) it has been applied widely, so we are able to compare our results here with other measures of targeting performance in a variety of programmes; (ii) it is easier to compare, across programmes, the normalised share of transfers going to the bottom X per cent than simply the share, and (iii) all targeting measurements display some weaknesses (see Ravallion, 2007 for a review of alternative measures) and often the choice comes down to a matter of preference.

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.[9] 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

[9]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).
[10]This analysis cannot be conducted for the CBT targeting mechanism as there were no pre-defined eligibility criteria for CBT.

  J. Int. Dev. **27**, 1521–1545 (2015) DOI: 10.1002/jid

Table 2. 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

Asterisks (*) indicate that the beneficiary household estimate is significantly different to the non-beneficiary household estimate:
***99%;
**95%;
*90%.
HSNP, Hunger Safety Net Programme; CGH, Coady–Grosh–Hoddinott index.
Source: HSNP M&E Baseline Evaluation Survey, September 2009 to October 2010.

have been selected for the programme under HSNP targeting than they would have been under random or universal targeting.

Coady *et al*. (2004) present empirical evidence on targeting efficiency and outcomes from 122 antipoverty 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 10 best performing schemes, the majority of which are in the USA, 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 well with other similar programmes in terms of targeting effectiveness at an aggregate level.

5.3 Performance of Targeting Implementation: Eligibility and Selection

To better understand the targeting performance of the programme, it is useful to decompose the targeting problem into issues of design and implementation (as discussed earlier). We turn first to targeting implementation errors—which relates to 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 3 illustrates coverage as well as inclusion and exclusion errors in implementation.[10] 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

[10]This analysis cannot be conducted for the CBT targeting mechanism as there were no pre-defined eligibility criteria for CBT.

Table 3. 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	3438
HSNP households	96***	70	79	2047
Coverage rate: % of households covered by HSNP				
All households	40***	56	51	5108
Eligible households	83	77	79	2077
Inclusion errors:				
% of beneficiary households that do not meet eligibility criteria	4***	30	21	2047
Exclusion errors:				
% of eligible households not covered by HSNP	17	23	21	2077

The'N' column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes.
Asterisks (*) indicate that an estimate is significantly different to the estimate in the cell to its right:
***99%;
**95%;
*90%.
HSNP, Hunger Safety Net Programme; CTA, categorical targeting by age; DR, dependency ratio.
Source: HSNP M&E Baseline Evaluation Survey, September 2009 to October 2010.

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.

In line with these strong coverage results, we also see low inclusion and exclusion errors in implementation. CTA outperforms DR as a targeting mechanism for implementation 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.

5.4 Performance of Targeting Design: Eligibility and Poverty

Next we assess the characteristics of eligible households, compared with ineligibles and, in particular, their relative poverty status.

Disaggregating eligibility by specific targeting mechanism (Table 4), we see some variation in the ability of the different mechanisms to identify the poorest households, with 58 per cent of eligible households in CTA areas being poor (defined on our relative 51 per cent measure), which is significantly different to the 50 per cent of ineligibles that are poor in the same areas. 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

Table 4. 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)	1763***	2152	1600***	2096
Household composition and eligibility overlap				
Proportion of households that (%):				
contain at least one member aged 55+ years (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

Asterisks (*) indicate that the eligible household estimate is significantly different to the ineligible household estimate:
***99%;
**95%;
*90%.
CTA, categorical targeting by age; DR, dependency ratio.
Source: HSNP M&E Baseline Evaluation Survey, September 2009 to October 2010.

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.

Looking at the last four rows of Table 4, it is clear that, on average, both mechanisms tend to favour households containing orphans and with chronically ill or disabled members compared with ineligible households. This is evidenced by the significant differences between the eligible and ineligible columns. This indicates that the eligibility criteria for these mechanisms pick up more than just consumption poverty, but they are able to identify households along a range of other characteristics.

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

5.5 What Factors Determine Selection/Eligibility for CBT?

The descriptive statistics in Table 4 reveal that 100 per cent of beneficiary CTA households contain at least one member over the age of 54 years. 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', such as households with no labour capacity or with very high dependency ratios, but it is instructive, in retrospect, to try to identify the key characteristics that communities used to identify their poorest members. To do this, we use probit regressions to analyse the factors associated with being selected under CBT. 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 survey weights.[11]

The coefficients presented in Table 5 have been transformed into marginal effects: so, for example, the coefficient 0.029 in the first column 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 (0.029×100). The coefficient for 'fully settled' in column 1 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 pooled sample. As a robustness check, we also estimate the correlates of selection into CBT by district (Table 5).

The most striking result from Table 5 is that there is no general story to be told about CBT. Clearly, different districts use different criteria. The only consistent results across all 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.

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 itself 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 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.

Overall, the key finding here is that 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, because communities are free to come up with their own criteria.

[11]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.

Table 5. Determinants of selection using CBT

Variables	(Overall) CBT	(Turkana) CBT	(Marsabit) CBT	(Mandera) CBT	(Wajir) CBT
Household (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	5105	1313	1299	1251	1242

The '*N*' column denotes the overall sample size. Asterisks (*) indicate that the estimated regression coefficient is statistically significant:
***99%;
**95%;
*90%.
CBT, community-based targeting.
Source: HSNP M&E Baseline Evaluation Survey, September 2009 to October 2010

5.6 Comparative Effectiveness of Targeting Mechanisms: Are the Poorest Households Selected?

We now move on to assess the comparative effectiveness of the three targeting mechanisms in selecting the poorest households in the evaluation areas. Due to varying poverty and coverage rates across the CBT, CTA and DR areas, as discussed earlier, a simple comparison of poverty rates among HSNP households cannot be used to compare targeting effectiveness of the three mechanisms. Instead, we use two other measures for this purpose: the ratio of beneficiary and non-beneficiary poverty rates; and the CGH index (Coady *et al*., 2004). The ratio of 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 with non-beneficiaries.

Table 6 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). 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 6 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.

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,

Table 6. 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

Poverty rate and CGH indices are based on a 51% relative poverty line.
Predicted targeting performance is based on the poverty rates among eligible households, that is, it is the targeting performance assuming 100% targeting accuracy whereby all eligible households are selected and all selected households are eligible.
HSNP, Hunger Safety Net Programme; CBT, community-based targeting; CTA, categorical targeting by age; DR, dependency ratio; CGH, Coady–Grosh–Hoddinott index.
Source: HSNP M&E Baseline Evaluation Survey, September 2009 to October 2010.

because 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.

6 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, the first two columns of Table 7 indicate 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 among 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.

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. Two points stand out here: (i) the targeting mechanisms, and their associated eligibility criteria, do not perform strongly on identifying the poor; and (ii) there is, therefore, considerable overlap between the two categories of households.

Therefore, the obvious question becomes: Is there an alternative targeting mechanism that would better identify the poorest households? In the subsequent discussion, we specify a range of alternative targeting criteria, as possible proxies for poverty, and compare them against DR and CTA eligibility criteria for the population as a whole. We use a proxy means test (PMT) methodology to simulate four alternative criteria, displayed in columns 3 to 6 in Table 7[12]:

(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 (this criterion may be used under a child benefit-type programme); and
(4) A household is eligible if it satisfies a threshold level under a PMT.[13]

[12]See Hurrell and Sabates-Wheeler (2011) for a more detailed explanation of this methodology.
[13]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 (predicted consumption expenditure). These are listed in Appendix 2. Most of these variables reflect the community criteria used to establish eligibility during the CBT selection process. The PMT threshold for eligibility was set 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.

DOI: 10.1002/jid

Table 7. Alternative targeting scenarios

	By eligibility status											
	1		2		3		4		5		6	
	CTA (if applied in all evaluation areas)		DR (if applied in all evaluation areas)		One or more orphans		One of more ill or disabled members		PMT-eligible		At least one child aged 5 or below	
Indicator	Elig	Inelig	Elig	Inelig	Elig	Inelig	Elig	Inelig	Elig	Inelig	Elig	Inelig
% of HHs that are among the *poorest* (bottom 51%)	59***	44	58***	40	58***	49	58***	49	74***	25	48**	56
Coady–Grosh–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+ years	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

The '*N*' column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes.
Asterisks (*) indicate that an estimate is significantly different to the relevant comparator, as explained in Section 1 of the baseline report:
***99%;
**95%;
*90%.
CTA, categorical targeting by age; DR, dependency ratio; PMT, proxy means test.
Source: HSNP M&E Baseline Evaluation Survey, September 2009 to October 2010.

Table 7 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 among the poorest households (bottom 51 per cent). 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 6 years. 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 among the poorest households (bottom 51 per cent). The demographic characteristics of households selected on this basis are also quite different. This does not appear to be a good proxy for targeting the poorest households, although we recognise that there may be other reasons for targeting transfers at households with small children.

Our simulation results suggest that a simple PMT approach would significantly outperform the CTA and DR criteria, and the actual CBT performance, in identifying poorer households. The proportion of the poorest households (bottom 51 per cent) that are eligible according to the PMT criteria is 74 per cent. 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 within sample prediction, and the same coefficients applied to another data set (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 that a PMT approach is best in all contexts, it is suggested this approach should be considered where the required data is available and there is sufficient capacity for effective implementation. The scope for combining an explicitly poverty-focused targeting approach (such as PMT) with community-based approaches should be considered.

7 CONCLUSIONS

The processes of selecting appropriate eligibility criteria and of identifying eligible people for selection into programmes present seemingly intractable challenges to social protection policymakers and administrators. Evidence from cross-country reviews suggests that trade-offs between targeting accuracy and targeting costs are inevitable and that choices need to be made about whether to invest more resources into improving targeting accuracy and whether to minimise inclusion error or exclusion error (Coady *et al.*, 2004; Devereux *et al.*, 2015). The findings presented in this paper reinforce Besley and Kanbur's (1990) theoretical proposition, that simpler approaches such as single proxy indicators are cheaper but less accurate than more complex approaches such as means tests. Our findings also confirm Coady *et al.*'s (2004) pragmatic proposition, that implementation is the single most important determinant of targeting success.

In contexts of widespread poverty or where the administrative and social costs of targeting are considered prohibitive, geographic targeting (blanket coverage of

DOI: 10.1002/jid

communities) is sometimes favoured as a way of minimising exclusion error with an acceptable level of inclusion error. However, even in very poor communities such as those in arid and semi-arid northern Kenya, income inequality is usually sufficiently high that targeting social transfers on the poorest individuals and households can be justified as a means of maximising their poverty-reducing impacts. The HSNP deployed three targeting mechanisms in an effort to determine which was most effective at identifying and reaching extremely poor and food-insecure households.

The analysis presented in this paper allows us to draw the following conclusions. Firstly, 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 with non-beneficiaries. In terms of food security, beneficiaries are only 16 per cent (9 percentage points) more likely to be food insecure compared with non-beneficiaries. Secondly, across the three targeting mechanisms trialled, 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 CBT 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 overall. 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 four 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 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.

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 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. A consideration for future research is the fact that while CTA and DR designs did not have quotas—selection of households is based on the criteria—in the areas where the CBT design was employed, a quota of 50 per cent was used. It would be useful to explore how the results may differ if no quota was imposed.

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 selected 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 always 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.

Reflecting on the implications of our findings, we conclude that designing a targeting strategy is not about choosing a targeting mechanism that will accurately identify the target group within a population with zero inclusion or exclusion errors. A strategic approach to targeting requires: first, identifying the target group that will maximise the achievement of the programme objectives; then selecting a mechanism for identifying members of the target group that is appropriate to the local context. Finally, sufficient resources must be dedicated to implement this targeting mechanism such that it achieves not 100 per cent accuracy, but a tolerable level of both inclusion and exclusion errors.

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. The authors are grateful for the insightful comments of the anonymous reviewers.

REFERENCES

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

Besley T, Kanbur R. 1990. The principles of targeting. In Policy, Research, and External Affairs Working Papers. World Bank: Washington DC.

Castañeda T. 2005. Targeting Social Spending to the Poor with Proxy–means Testing: Colombia's SISBEN System. Social Protection Unit, Human Development Network, World Bank: Washington DC.

Coady D, Grosh M, Hoddinott J. 2004. Targeting outcomes redux. *World Bank Research Observer* **19**(1): 61–85.

Conning J, Kevane M. 2002. Community-based targeting mechanisms for social safety nets: a critical review. *World Development* **30**(3): 375–394.

Cornia A, Stewart F. 1993. Two Errors of Targeting. *Journal of International Development* **5**(5): 459–496.

Devereux S, Masset E, Sabates-Wheeler R, Samson M, Rivas A, te Lintelo D. 2015. Evaluating the targeting effectiveness of social transfers: a literature review. IDS Working Paper and Centre for Social Protection Working Paper, 11. Institute of Development Studies: Brighton.

Ellis F. 2012. "We Are All Poor Here": economic differences, social divisiveness and targeting cash transfers in sub-Saharan Africa. *Journal of Development Studies* **48**(2): 201–214.

Hodges A, Dufay A-C, Dashdorj K, Jong K, Budragchaa U. 2007. Child Benefits and Poverty Reduction: Evidence from Mongolia's Child Money Programme. United Nations Children's Fund: New York.

Hurrell A, Sabates-Wheeler R. 2011. Hunger Safety-net Programme (Kenya) Targeting Effectiveness Evaluation Report. Oxford Policy Management (OPM): Oxford.

Kidd S, Wylde E. 2011. Targeting the Poorest: An assessment of the proxy means test methodology. Australian Agency for International Development (AusAID): Canberra.

Mansuri G, and Rao V. 2004. Community-based and -driven development: a critical review. *Policy Research Working Paper* 3209, Washington DC: World Bank.

Pritchett L. 2005. The Political Economy of Targeted Safety Nets. Social Protection Unit, Human Development Network, World Bank: Washington DC.

Ravallion M. 2007. "How relevant is targeting to the success of an antipoverty program?." *World Bank Policy Research Working Paper Series*, *Vol* (2007).

APPENDIX 1

HOUSEHOLD WELFARE BY CONSUMPTION EXPENDITURE QUINTILE

	Consumption expenditure quintile					All HSNP evaluation areas	
	Q1	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	5105
Mean monthly household health expenditure (KES)	58***	72***	85***	138	277***	126	5105
Mean monthly household education expenditure (KES)	47***	137**	198	293**	415***	218	5105
Household assets and livestock ownership							
Mean value of all household assets owned by household (KES)	9095**	12 478*	15 230*	27 226	66 917	26 184	5105
Mean value of productive assets owned by household (KES)	718***	1548***	2370	3011**	4042**	2337	5106
Proportion of households owning livestock (%)	78	77**	75**	70	53***	70	5106
Mean tropical livestock units owned currently (for households owning livestock) (TLUs)	1.0	1.4	2.0	2.2	2.8	1.8	3778
Education and health status							
Proportion of adults aged 18+ years that are literate (%)	14***	19	18**	27**	35***	22	12 611
Proportion of children aged 6–17 years 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	5106
Proportion of households with walls made of natural materials (%)	98***	94**	93***	83*	66***	87	5106
Subjective poverty							
Proportion of households reporting that they are ‘struggling’ (%)	68***	60	65***	57	39***	58	5106
Proportion of households reporting that they are ‘unable to meet household needs‘ (%)	20***	14***	8	6***	3***	10	5106

The ‘N’ column denotes the overall sample size. The sample sizes for the disaggregated estimates in other columns are based on smaller sample sizes.

Asterisks (*) indicate that a quintile estimate is significantly different to the pooled mean across the other four quintiles:

***99%;

**95%;

*90%.

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, September 2009 to October 2010.

APPENDIX 2

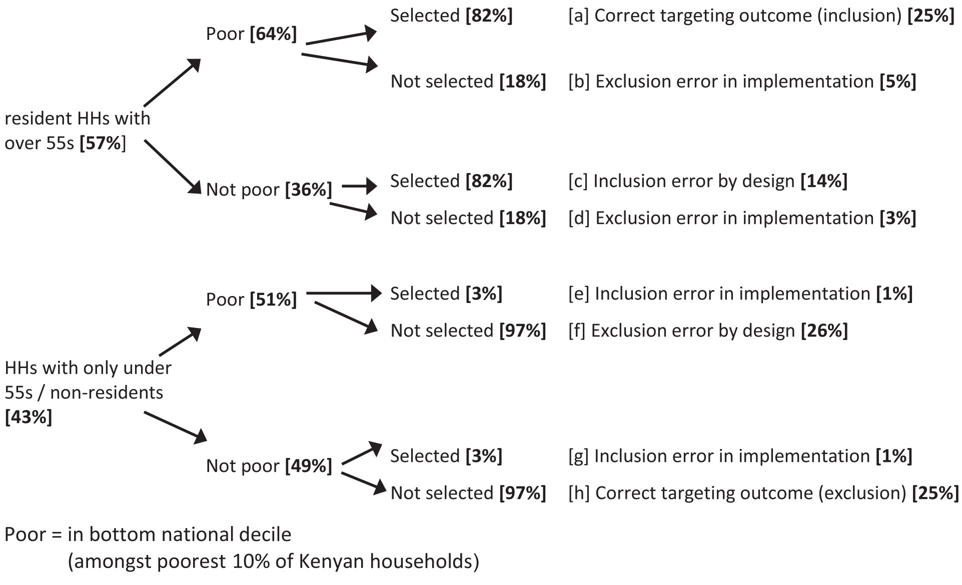
QUESTIONS USED TO CONSTRUCT THE PROXY MEANS TEST SCORE

(1) Whether the household receives food aid (Y/N)
(2) Whether the 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

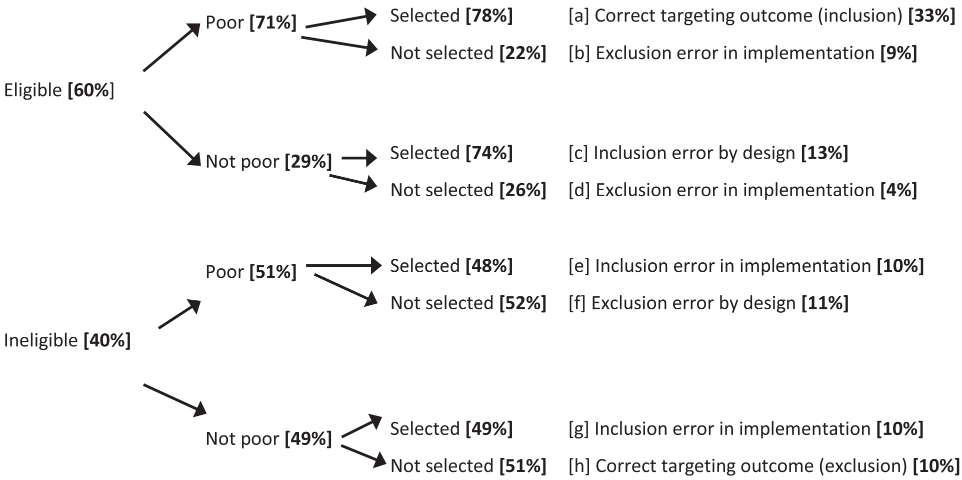
APPENDIX 3: DECISION TREES

Statistics for the targeting tree diagram:

Social Pension:



Dependency Ratio:



CBT: