



Targeting the Poorest:

An assessment of the proxy means test methodology

September 2011





Acknowledgements

This paper was researched and written by Stephen Kidd and Emily Wylde with the able assistance of Zoltan Tiba, Daniel Stein and Oliver Vanden Eynde. We would like to thank the anonymous reviewers who provided food for thought and many helpful suggestions. We are also grateful to Nobuo Yoshida, Matt Wai-Poi and Amri Ilmma of the World Bank for sharing information, data and STATA code for Bangladesh and Indonesia.

The proxy means test is a targeting methodology with many advocates and detractors but there has been little research on its effectiveness. While the views expressed in this publication are the researchers, the Australian Agency for International Development (AusAID) supports evidence-based debates. We hope this paper will encourage others to investigate further the strengths and weaknesses of the proxy means test methodology.

© Commonwealth of Australia 2011

This work is copyright. You may download, display, print and reproduce this material in unaltered form only (retaining this notice) for your personal, non-commercial use or use within your organisation. Apart from any use as permitted under the *Copyright Act 1968*, all other rights are reserved. Requests and inquiries concerning reproduction and rights should be addressed to Commonwealth Copyright Administration, Attorney General's Department, Robert Garran Offices, National Circuit, Barton, ACT 2600 or posted at <http://www.ag.gov.au/cca>.

Published by the Australian Agency for International Development (AusAID), Canberra, September 2011.

This document is online at www.ausaid.gov.au/publications

Disclaimer The views expressed in this publication are those of the authors and not necessarily those of AusAID.

For further information, contact:

AusAID
GPO Box 887
Canberra ACT 2601

Phone (02) 6206 4000
Facsimile (02) 6206 4880
Internet www.ausaid.gov.au

Cover image: An Oportunidades staff member administers a proxy means test to a potential program beneficiary
Photo: National Coordination of Oportunidades, Archives SEDESOL, Mexico.

Executive Summary

The purpose of this study is to help AusAID staff, their counterparts in partner governments and others working in the field of social protection to better understand the strengths and weaknesses of a targeting methodology known as the proxy means test (PMT). As social protection practitioners search for effective ways to target poor people in developing countries, proxy means testing has become increasingly popular. The methodology estimates household income by associating indicators or ‘proxies’ with household expenditure or consumption. This study assesses its accuracy, objectivity, transparency and ease of implementation.

Despite the substantial literature on the PMT there is a dearth of comprehensive analysis. This means that developing country governments—and those advising on social protection—do not have an adequate basis upon which to consider the merits of proxy means testing or assess the methodology against alternatives. ‘Targeting the Poorest’ gives policymakers and those working in social protection information to help them decide whether to use the PMT. It attempts to explain the methodology’s considerable inaccuracy at low levels of coverage and sets out other challenges to its use. The theoretical basis of the methodology and its implementation are examined as are its associated costs.

Proxy means testing uses multivariate regression to correlate certain proxies, such as assets and household characteristics, with poverty and income. This study assesses regression accuracy in Bangladesh, Indonesia, Rwanda and Sri Lanka and finds that the PMT has high in-built errors, especially at relatively low levels of coverage (20% of the population and below). Exclusion and inclusion errors vary between 44% and 55% when 20% of the population is covered and between 57% and 71% when 10% is covered.

Part of the reason for this is the imperfect correlation between multiple proxies and household consumption. Additionally, the PMT methodology is based on national household survey data that represent ‘reality’ at one point in time and are inherently inaccurate to varying degrees. Other issues are sampling errors in household surveys and assumptions made in applying the PMT, which increase the arbitrary nature of the methodology yet affect whether individual households receive social protection benefits.

Implementing proxy means testing presents a number of challenges. Enumerators are not always objective when conducting surveys and do not always have time to verify proxies within households. Some proxies can also be difficult to verify—such as level of education, age and household assets—and interviewees can influence survey results, with children and men not as reliable as women.

Another challenge of proxy means testing relates to crises and shocks faced by households, including minor ones that are part of every day life. As a result, households that fall into poverty but do not suffer a related change in the household characteristics and assets used as proxies cannot receive social protection benefits.



The PMT is expensive to administer and has associated social and political costs. There is evidence that it can generate social conflict and stigmatise beneficiaries. Politically the methodology—as with all other forms of poverty targeting—is less likely to be popular because it excludes the middle class and those who are better-off.

This study’s findings show that the PMT is inherently inaccurate, especially at low levels of coverage, and it relatively arbitrarily selects beneficiaries. It therefore functions more like a simple rationing mechanism that selects some poor and non-poor but excludes large numbers of eligible poor from receiving benefits and support.

‘Targeting the Poorest’ does not provide an in-depth assessment of other targeting methods or formally compare them with the PMT methodology. It suggests, however, that other methods used to develop social protection schemes—which do not directly target poor people—may be better at including intended beneficiaries and avoiding the pitfalls of the PMT.

Contents

Executive Summary	ii
1. Introduction	1
1.1. Study methodology	3
1.2. Structure	3
2. The targeting effectiveness of proxy means testing in practice	4
3. Errors embedded within the regression methodology	6
3.1. Household surveys: a robust tool for targeting?	6
3.2. Designing the proxy means test: errors in modelling poverty	8
3.2.1. Accuracy of proxy means test regression analysis	9
3.2.2. Understanding errors inherent in proxy means test regressions	11
3.2.3. Impact of methodological choices	14
3.2.4. Sampling errors: impact on proxy means testing	20
4. Implementing proxy means testing: issues with practice	21
4.1. Finding the beneficiaries	21
4.2. Capacity of enumerators	22
4.3. Verifying indicators	24
4.4. Identity of the interviewee	24
4.5. Community verification	25
4.6. Political interference	25
4.7. Recertification and poverty dynamics	25
4.8. Responsiveness to the household life cycle	26
4.9. Appeals	27
4.10. Addressing intra-household poverty	28

5. Social and political costs of the proxy means test methodology	29
5.1. Social costs	29
5.2. Political costs	30
6. Conclusion	31
Annex 1: Methodology for devising proxy means test scores	33
Annex 2: Variables used in proxy means tests for each study country	34
Annex 3: Explaining higher errors found in other analysis of proxy means tests	36
Annex 4: Impacts of equivalence scales on the selection of beneficiaries	39
References	40

Figures

Figure 1:	Exclusion and inclusion errors with the PMT at different cut-off levels	9
Figure 2:	Exclusion and inclusion errors with the PMT at different cut-off levels: Bangladesh and Rwanda	10
Figure 3:	Scatter plots from Indonesia and Bangladesh indicating predicted versus actual poverty rankings using a target cut off of 20%	11
Figure 4:	Distribution of actual versus predicted values of proxy means test scores	12
Figure 5:	Actual versus predicted values of proxy means test scores for Bangladesh, Indonesia, Rwanda and Sri Lanka	13
Figure 6:	Per capita consumption versus proxy means test scores by household, Bangladesh	14
Figure 7:	Percentage of households in targeted populations treated differently depending on the use of equivalence scales at 10% and 20% cut-off levels	15
Figure 8:	Proportion of target households correctly identified in Bangladesh using two equivalence scale scenarios (per capita and 0.5) at three cut-off levels (5%, 10%, 20%)	16
Figure 9:	Percentage of households in targeted populations treated differently depending on the treatment of missing variables at 10% and 20% cut-off levels, Sri Lanka	18
Figure 10:	Percentage of households in targeted populations treated differently depending on the variables used in the proxy means test regression, Indonesia	19
Figure 11:	Percentage of households in targeted populations treated differently when the lower or upper bound of the confidence interval on regression coefficients is used	20
Figure A:	Cumulative distribution of actual consumption and proxy means test score, Sri Lanka	38
Figure B:	Percentage of households in targeted populations treated differently depending on the use of equivalence scales	39

1. Introduction

In recent years there has been significant and increasingly polarising debate over what targeting methodology should be used for social protection programs, including universal methods that target everyone in a demographic category or methods that target just the poor. Within this debate there has been little attention paid to whether targeting methodologies deliver what they set out to achieve.

Social protection analysts often accept that simple methods, such as universal targeting, are relatively effective largely because their targeting criteria (such as age) can be easily determined. Methods targeting the poor are more complex, especially in developing countries with large informal economies which make it difficult to accurately assess incomes given that, for example, most people do not report how much they earn.

The PMT methodology, originally developed in Chile and designed to address this problem, is briefly described in Box 1.

Box 1: The proxy means test methodology – a brief description

The PMT is based on national household surveys. Given that household income in developing countries is often difficult and expensive to measure accurately, the methodology relies on household assets and other indicators—or proxies—to estimate household welfare.

To work, the proxies used need to be easy to measure. They include demographic characteristics (such as age of household members and size of household), human capital characteristics (such as education of household head and enrolment of children in school), physical housing characteristics (such as type of roof or floor), durable goods (such as refrigerators, televisions or cars) and productive assets (such as land or animals). Regressions are run to find the proxies that most correlate with welfare. While individual proxies may be weakly correlated with welfare, multiple proxies show stronger correlations.

The PMT uses a set of proxies (usually between 10 and 30) that best explain welfare. Each proxy is given a weight based on its estimated impact on household expenditure.

Enumerators visit households to see if they have the proxies being used in the PMT. Then, using the agreed weights, a score is calculated for each household. Households that score below the cut-off point are eligible for the social protection program being considered.

In the past decade proxy means testing has spread across the world and is widely regarded to accurately identify the poor.¹ The World Bank (2009a:7) recently stated that the PMT has been:

... proven to work particularly well in countries with high levels of informality and where personal and household income is difficult to verify with any degree of precision.

A number of reports comment on the efficacy of implementation, suggesting the methodology is objective² and transparent³ and that proxies can be easily measured and verified during household visits.⁴ Overall, most studies portray the PMT as a scientific and technocratic solution to poverty targeting.

This study investigates these claims, examining the methodology's accuracy and the robustness of its implementation. It also looks at the social and political costs associated with proxy means testing.

This study focuses only on the PMT methodology. It does not examine in detail other targeting methodologies or formally compare them to proxy means testing. It does not suggest ways to improve the PMT but rather assesses its value as a targeting methodology to enable policymakers and others to make informed judgments on the appropriateness of its use.

Some issues faced by the PMT are typical of other poverty targeting methods and some apply to all methodologies. To paint a complete picture of the PMT this study describes some issues the methodology has in common with other approaches, as well as some specific to the PMT alone.

A key aspect of this study is how accurate the methodology is in measuring exclusion and inclusion errors.⁵ An exclusion error is defined as the proportion of those eligible for a social protection program but who are excluded from it as a result of inaccurate targeting. An inclusion error is defined as the proportion of those selected for a program who are not eligible for it. For the purpose of this study the eligible population corresponds to the coverage of the program. So if a program has 10% coverage, it is assumed that those eligible form the poorest 10% of the population.⁶ The measure of inclusion and exclusion errors used in this study will result in the two errors being the same.⁷

1 See, for example, Castañeda and Lindert (2005:40); Grosh et al. (2008:101); Ahmed and Bouis (2002:68); Johannsen (2006:3); and Hanlon et al. (2010:108).

2 Skoufias et al. (1999:1); Narayan and Yoshida (2005:1); Grosh et al. (2008:100); World Bank (2009c); and Chamberlin et al. (2009).

3 Castañeda and Lindert (2005:5); World Bank (2009a;2009b;2009c).

4 Maluccio (2008), Coady et al (2002:29), Coady et al. (2004).

5 Across the literature these errors can be defined differently. For example, inclusion errors can mean: a) program beneficiaries not eligible (however eligibility is defined); b) those not eligible for the program but who are beneficiaries; or, c) non-poor who are program beneficiaries. Exclusion errors can mean: a) intended beneficiaries of the program unable to access it; or b) poor unable to access it. In the literature it is often unclear which definition is being used.

6 Determining inclusion and exclusion errors with reference to a poverty line and rate that may be higher than the program's coverage rate does not enable a robust assessment of targeting accuracy. Such an assessment would make it difficult to differentiate between errors due to the methodology with the impact of a limited budget.

7 For example, in a program targeted at the poorest 10%, if 40% of those eligible are excluded then the other 40% included must not have been eligible (that is, they are not among the 10% poorest), making both exclusion and inclusion errors 40%.

1.1. Study methodology

The study's methodology adopted two approaches. The first was an extensive review of proxy means testing literature. A number of qualitative studies provided insights, especially with implementation challenges.

The second was an assessment of effectiveness. This tested the methodology using household survey data from Bangladesh, Indonesia, Rwanda and Sri Lanka. In three of these countries a proxy means testing formula was already developed. This study applied these formulae to their respective household survey data. In the fourth country—Rwanda—the PMT formula was developed for this study by replicating the standard practice used elsewhere.

By combining the quantitative and qualitative information gleaned through the two approaches the study paints a detailed picture of the PMT methodology.

1.2. Structure

Section 2 examines quantitative and qualitative evidence on the accuracy of the PMT methodology as used by social protection programs across the world, and the implications of this on communities.

Section 3 examines the causes of the inaccuracies associated with proxy means testing, assessing whether a national household survey is an appropriate basis for a targeting mechanism and analysing some issues inherent in the PMT's regression methodology. Section 4 examines the methodology's relatively arbitrary nature, presenting the results of the literature review and describing how these affect accuracy. Section 5 examines the social and political costs of the methodology and Section 6 concludes the paper.

2. The targeting effectiveness of proxy means testing in practice

There is little robust evidence in the literature on the PMT's effectiveness after implementation. Most assessments focus on inclusion errors and paint a sketchy picture.⁸ However, one study of Mexico's *Oportunidades* (formerly known as *Progresa*), a program that reaches around 20% of the population, estimated an exclusion error of 70% and an inclusion error of 36% (Veras et al. 2007:2). Another study of a program in Armenia using the PMT saw the better-off benefiting more than the poor.⁹

Another way to assess targeting accuracy is to examine the perception of communities in programs where the PMT has been applied. Studies have shown that communities do not understand why the methodology excludes large numbers of poor and eligible people. The view of one doctor in a community in Mexico typifies this:

Frankly, I don't know how they got the data for *Progresa* because there are families here in this community who are poor, poor. There are large families that do not have support from *Progresa* and we have proof. I have been here eight years and know the entire community inside-out ... and I've found that there are many poor people who do not have *Progresa* and we do not know why they have been left outside the program. (Adato et al. 2000:27)

This sentiment is repeated in other studies of communities that participated in proxy means testing. Qualitative research in Mexico, Nicaragua and Peru indicates that some community members ascribe the omission of poor households to luck or God's will, describing the PMT methodology as similar to a lottery.¹⁰ In Mexico's *Progresa* program one facilitator remarked:

What I understand is that when people ask me why they are not in the list, despite being surveyed, I tell them that I don't know the precise reason but that probably it was a lottery and because it was a lottery, no-one knows who needs it most. (Adato et al. 2000)

Community members—both chosen and rejected—find the omissions of eligible households difficult. In Nicaragua, for instance, one non-beneficiary's mother described a conversation with her son:

He ... well he gets sad because he says, "Look mom, if we were beneficiaries ... if you were a beneficiary you could perhaps buy me some shoes", because right now he has no shoes ... So I tell him: "If I were a beneficiary of course you would have your shoes and your clothes already, but you know the reality, not all the women here are beneficiaries, but maybe we become beneficiaries soon, let's not lose our faith". (Adato and Roopnaraine 2004:78)

8 A well-known assessment of targeting accuracy is the Coady et al. (2004) study examining 85 programs. However, the ranking of programs by targeting accuracy is somewhat misleading because it conflates different maximum scores and does not account for program coverage or exclusion errors.

9 Coady et al. (2004).

10 Adato (2000); Adato et al. (2000); Adato and Roopnaraine (2004); Huber et al. (2008:49).

In some communities in Nicaragua and Peru, beneficiary families have held collections to buy school supplies for non-beneficiary families since they are so poor.¹¹ One *Progresa* beneficiary stated:

I feel bad because at times someone I know says to me: “Ay sister! I haven’t got anything to give to my children; and then I think and I say that those of us who are part of *Progresa*, on the day that they pay us, why don’t we cooperate between all of us, some with soup, others with other things and we make a fund and we give it to those who aren’t in *Progresa*”.

(Adato 2000)

The few results available on the PMT’s targeting accuracy depict a problematic picture of its accuracy and effectiveness. Sections 3 and 4 examine the reasons for this.

¹¹ Adato and Roopnaraine (2004:78f); Huber et al. (2008:46).

3. Errors embedded within the regression methodology

The PMT uses information from household surveys to undertake regressions that correlate household assets and household expenditure (or consumption). Assessing the methodology therefore requires examining whether the household survey is a robust tool for targeting and the accuracy of the regressions.

3.1. Household surveys: a robust tool for targeting?

A household survey is designed to measure household consumption, expenditure, income and other factors to assess national wellbeing and poverty levels. It is less suited to form the basis of a targeting mechanism, and to do so would need to provide a better representation of reality. The less accurate it is the weaker the methodology's foundation becomes—inaccuracies and misrepresentations in a household survey interact with errors arising from the regressions between the proxies and household consumption.

Two main types of errors exist in household surveys: non-sampling and sampling.¹²

Non-sampling errors are incorporated into household surveys in many ways. For example, these surveys:

- assume households are static entities, when in reality they can be dynamic and changeable—even during a survey period—making it challenging at times to define exactly who is a household member¹³
- can miss certain types of households (in particular the very rich and very poor)¹⁴ and certain types of household members (such as those residing temporarily like migrants or children in boarding schools)
- do not always accurately represent certain types of households, such as homes to people with disability
- include inaccurate information supplied by respondents¹⁵
- adopt different approaches to survey design, which can influence the accuracy of reporting on consumption.

Inaccuracies are inherent in household surveys due to sampling error; for example, samples are relatively small, often between 2000 and 10 000 households. While all results have an inherent margin of error and ideally should be presented with caution, this is rarely done (Haughton and Khandker 2009:13).

12 Sampling errors are due to observing a sample of the population rather than the whole population. Non-sampling errors are due to all other types of errors.

13 See, for example, Beaman and Dillon (2010) for an analysis of the impact of household definitions on welfare measures, and Udry (1996) on the operational difficulties of implementing the 'common pot' definition.

14 Narayan and Yoshida (2005) demonstrate how two household surveys conducted in Sri Lanka are not comparable.

15 Consumption is often understated and a 5% understatement can translate into a 10% overstatement of the poverty head count. The way questions are asked also matters. In a 1993–94 household survey in Cambodia per capita consumption was measured as 2262 riels. In a 1999 survey it was only 1799 riels, despite robust economic growth, according to Haughton and Khandker (2009:22–24;83).

The quality of results from household surveys depends not only on sampling and non-sampling errors, but also on the methodological choices made when analysing data. Key decisions involve: what to include in the measure of household consumption¹⁶; adjustments made for the consumption of children compared to adults; and if and how to take account of economies of scale in the household. Other issues relate to the level of detail in the questionnaire, the accuracy of reporting and the extent of missing variables, all of which require analysts to make assumptions.¹⁷

Methodological choices and analyst assumptions also influence results. One example relates to the choice to use equivalence scales. Household surveys are designed to estimate consumption of entire households. To determine individual welfare, household consumption is divided by the number of people in the household.¹⁸ Because not everyone in a household consumes the same (young children, for example, consume less than working adults) different members can be given different weightings, called equivalence scales. Using such scales can lead to different results, as illustrated in Box 2. Another example illustrating this is the World Bank's (2002:129) analysis of the 1998–99 Pakistan Integrated Household Survey which estimated a poverty rate of 26.9% using an equivalence scale of 0.8 compared to a poverty rate of 32.6% without an equivalence scale.¹⁹

Box 2: Equivalence scales: impact on household survey results

Imagine a household of two adults with four children aged 0 to 15 years of age, with a household consumption of \$1200 a year in a country where the per capita poverty line is \$250 a year. If no equivalence scales are used, the consumption of each individual in the household would be \$200 a year and the household would be below the poverty line. If, however, an equivalence scale of 1 were applied for adults and 0.5 for children 0 to 15 years of age, then the consumption of each individual would be \$300, making the household apparently 50% better off and no longer poor.

As well as influencing poverty estimates, equivalence scales can change how types of households are ranked on relative poverty. Lanjouw et al. (1998), for example, assessed poverty rates in seven countries in Eastern Europe and the former Soviet Union: without equivalence scales the elderly were found to be better off than average while households with more than two children were poorer than average. A small change in equivalence scales reversed these rankings.

16 Durable goods, such as cars, bicycles and televisions, can be problematic. Haughton and Khandker (2009:25) point out that consumption should build in depreciation and only include the amount of such goods consumed during the survey year, although this is difficult to calculate and so is often ignored. Researchers also rarely include the value of publicly provided goods and services—Haughton and Khandker (2009:17).

17 Deaton and Zaidi (2002) discuss the challenges in analysing household consumption.

18 In reality, household surveys cannot accurately assess individual households because the intra-household distribution of wealth and income is unknown. It is more accurate to say such surveys can calculate the per capita poverty levels of households within which specific categories of people live.

19 Using an updated dataset for Pakistan (the 2007–08 Pakistan Social and Living Standards Measurement Survey and Household Integrated Economic Survey) Kidd, Wylde and Tiba (2010) calculated a poverty rate of 25.1% without equivalence scales, 17% with an equivalence scale of 0.8 for children, and 7% with an equivalence scale of 0.5 for children.

There is no correct answer on equivalence scale choice.²⁰ Best practice is to conduct thorough sensitivity analyses on equivalence scales so analysts are aware of the effects of their assumptions and can ensure data interpretations are robust. Much of the poverty analysis feeding into proxy means testing lacks this analysis.

A further concern with household surveys as a basis for the PMT is the use of consumption or expenditure as the basis for determining household wellbeing. This is not problematic when the household survey is measuring poverty, as it is usually accepted that consumption and expenditure are better reported than income in developing countries and may be a more accurate wellbeing measure. Means tests, on the other hand, should measure a household's capacity to care for itself, as determined by income rather than consumption. Yet a PMT's accuracy depends on the correlation between the proxies and consumption or expenditure, and the inaccuracies this leads to are impossible to measure.

This discussion does not undermine household surveys as useful instruments for measuring welfare, but they are only approximations of reality. As Haughton and Khandker (2009:34) point out:

There is no ideal measure of well-being; all measures of poverty are imperfect. This is not an argument for avoiding measures of poverty, but rather for approaching all measures of poverty with a degree of caution.

According to the literature, it is rare that the accuracy of the household survey upon which a PMT is based is assessed. Consequently, a key concern is whether PMTs are being promoted with sufficient caution, or whether their accuracy is in danger of being oversold.²¹

3.2. Designing the proxy means test: errors in modelling poverty

For proxy means testing to be accurate, the regressions it uses to determine proxies and their respective weightings need to be accurate. To test this, this study undertook simulations on household surveys using data from Bangladesh, Indonesia, Sri Lanka and Rwanda. The first three countries were chosen because the PMT methodology had already been developed and was accessible. Rwanda enabled the study to examine a low-income African country, and although a PMT formula or analysis did not exist simulations were created based on the approach used in the three other countries.

Within the household survey data there are hundreds of potential proxies. Through regression analysis the PMT methodology aims to identify the proxies most correlated with poverty that can easily and accurately be assessed when visiting households during the targeting survey (Annex 1 details the methodology for calculating PMT scores). The selected proxies tend to be broadly similar across tests, as indicated in Annex 2, which sets out the proxies used in the four study countries.

20 Deaton (1997:241–268) discusses the theoretical and empirical issues involved. Similar challenges arise in the choice of how to deal with economies of scale in household survey analysis; for example, it is cheaper per capita to house a couple than to house two individuals separately. Haughton and Khandker discuss this more fully (2009:27ff;89ff) and Anand and Morduch (1996) discuss the significant role of economies of scale in measuring consumption in Bangladesh.

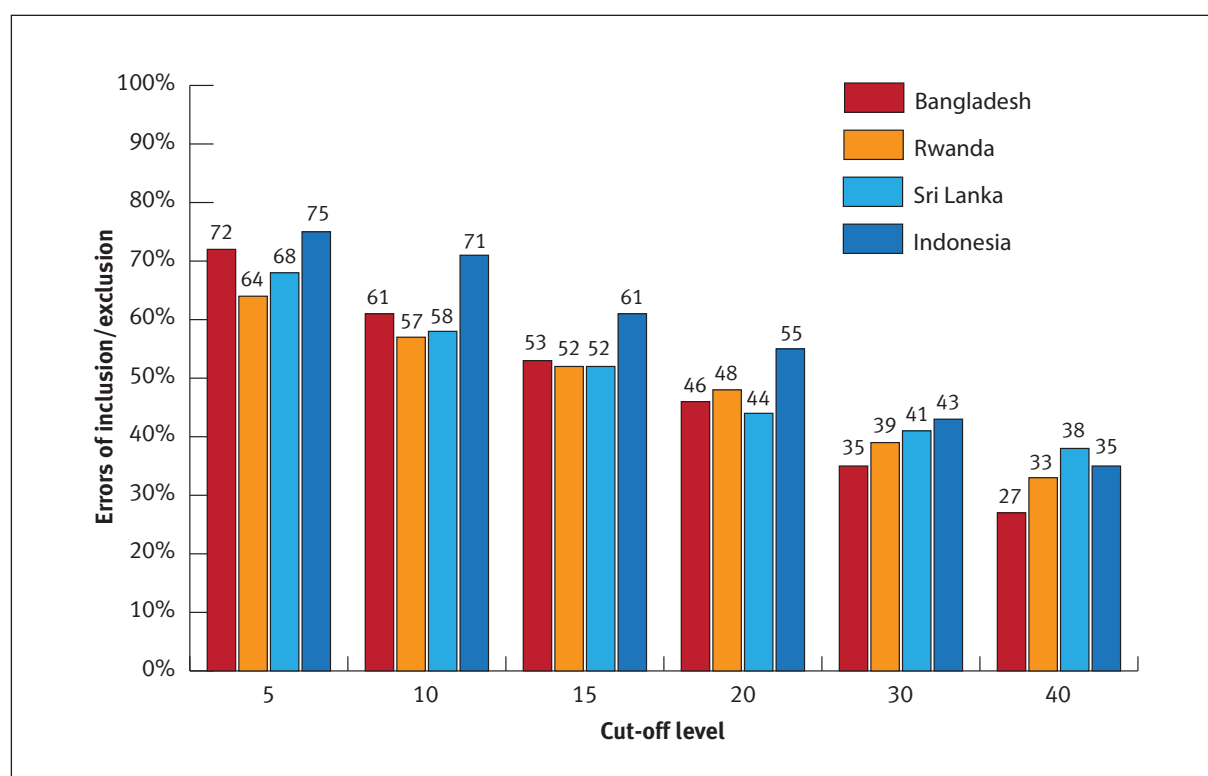
21 For example, the World Bank (2009a:37) notes that in Bosnia two household surveys produced different outcomes in PMTs. Neither was 'correct' since both included some degree of inaccuracy.

This section presents the simulation results, discusses the accuracy of the PMT regression analysis and errors inherent in the regressions, and examines the impact of assumptions and sampling errors.

3.2.1. Accuracy of proxy means test regression analysis

The study assessed PMT accuracy against household survey data at different cut-off levels. In the literature PMTs are commonly assessed against relatively high population coverage of 20% to 40%, but in developing countries (with limited program resources) assessing lower levels of population coverage (below 20%) is more realistic. Assessment results are set out in Figure 1.

Figure 1: Exclusion and inclusion errors with the PMT at different cut-off levels

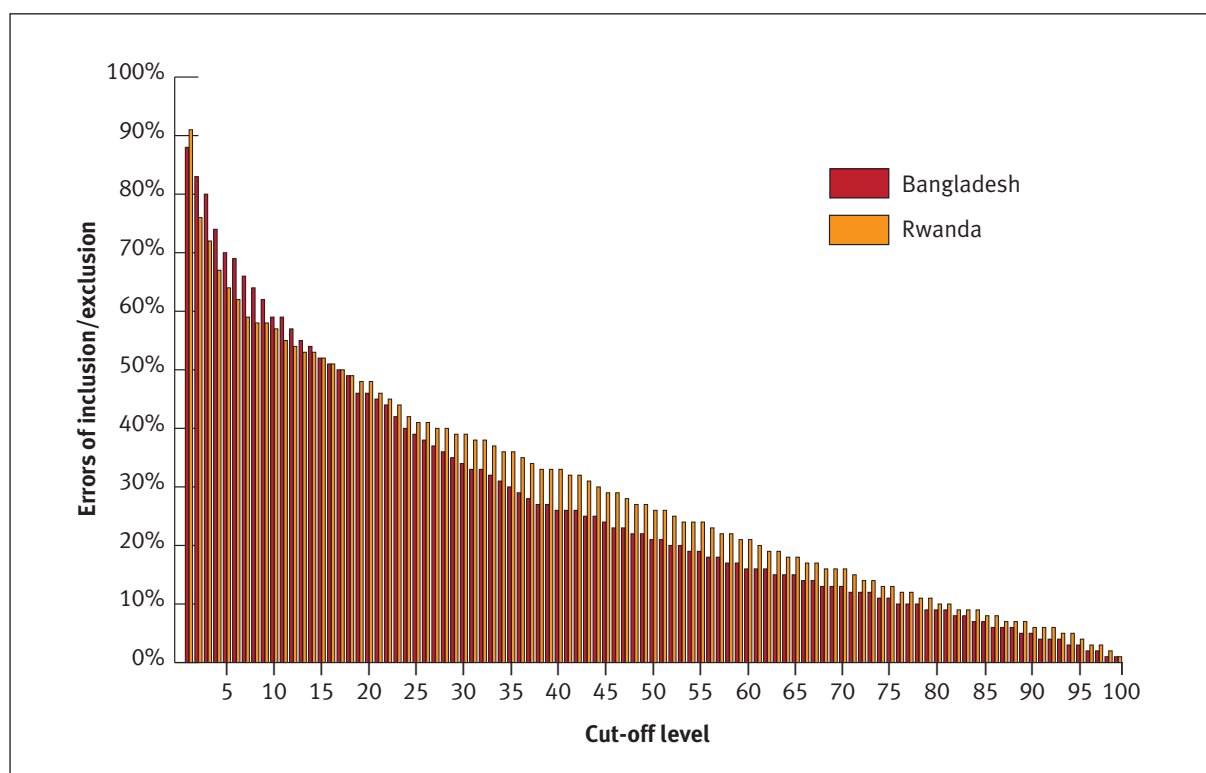


Targeting errors increase the lower the cut-off level (that is, the smaller the size of the program). In Bangladesh, for example, errors of inclusion and exclusion increase from 46% to 61% when program size is halved (that is, by targeting the poorest 10% rather than 20%). In Indonesia the increase is from 55% to 71%. Even at higher levels of coverage, the theoretical targeting errors are large. Figure 1 indicates errors of between 35% (Bangladesh) and 43% (Indonesia) at 30% coverage.²²

²² The researchers also analysed theoretical errors in Pakistan and found similar results. At 10% coverage, the exclusion error would be 67%, at 20% coverage it would be 50%, and at 30% coverage it would be 42% (Kidd, Wylde and Tiba 2010). A recent analysis of Nepal's PMT indicated similar errors (Kidd et al. 2011).

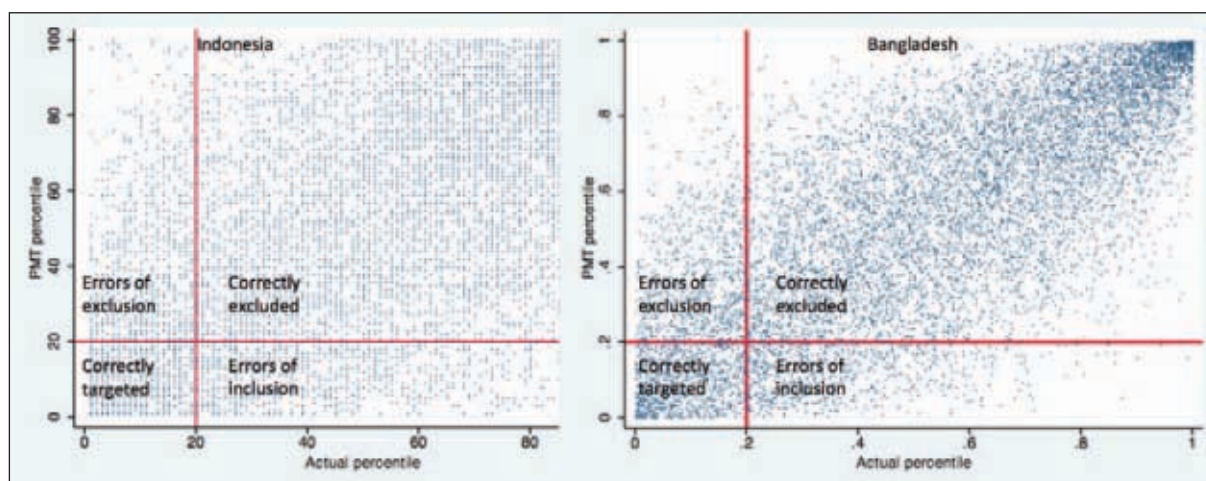
Taking this analysis a step further, Figure 2 shows the distribution of exclusion and inclusion errors that would be found at cut-off levels between 0% to 100% of the population in Bangladesh and Rwanda. At coverage below 10% theoretical errors are high and achieving inclusion and exclusion errors below 30% would require targeting at least the bottom 35% to 50% of the population. The pattern is similar in Indonesia and Sri Lanka.

Figure 2: Exclusion and inclusion errors with the PMT at different cut-off levels: Bangladesh and Rwanda



Another way of viewing performance is to examine scatter plots comparing household ranking in the poverty distribution to their ranking using the PMT score. As Figure 3 indicates for Indonesia and Bangladesh, there is a broad scatter of households across the graph indicating poor targeting. The large number of households incorrectly excluded and included at 20% coverage can be observed in the top left and bottom right-hand quadrants.

Figure 3: Scatter plots from Indonesia and Bangladesh indicating predicted versus actual poverty rankings using a target cut off of 20%



A key finding is that the magnitudes of errors are consistent across all four countries, reflecting problems with the PMT approach rather than with its application in a given country.²³ The errors found in this study are broadly consistent with those found by other analysts. For example, in Sri Lanka, Narayan and Yoshida (2005:17) found an exclusion error of 43% at 30% coverage, and in Bosnia and Herzegovina the World Bank (2009a:36) found an exclusion error of 45% at 20% coverage. Divergences in reported findings from this pattern result from differences in the methodology used (Annex 3 explains in detail).

3.2.2. Understanding errors inherent in proxy means test regressions

A number of reasons explain inaccuracies in the PMT targeting. Underlying them is that regressions rarely explain more than half of the variation in consumption between households (Coady et al. 2004:54).

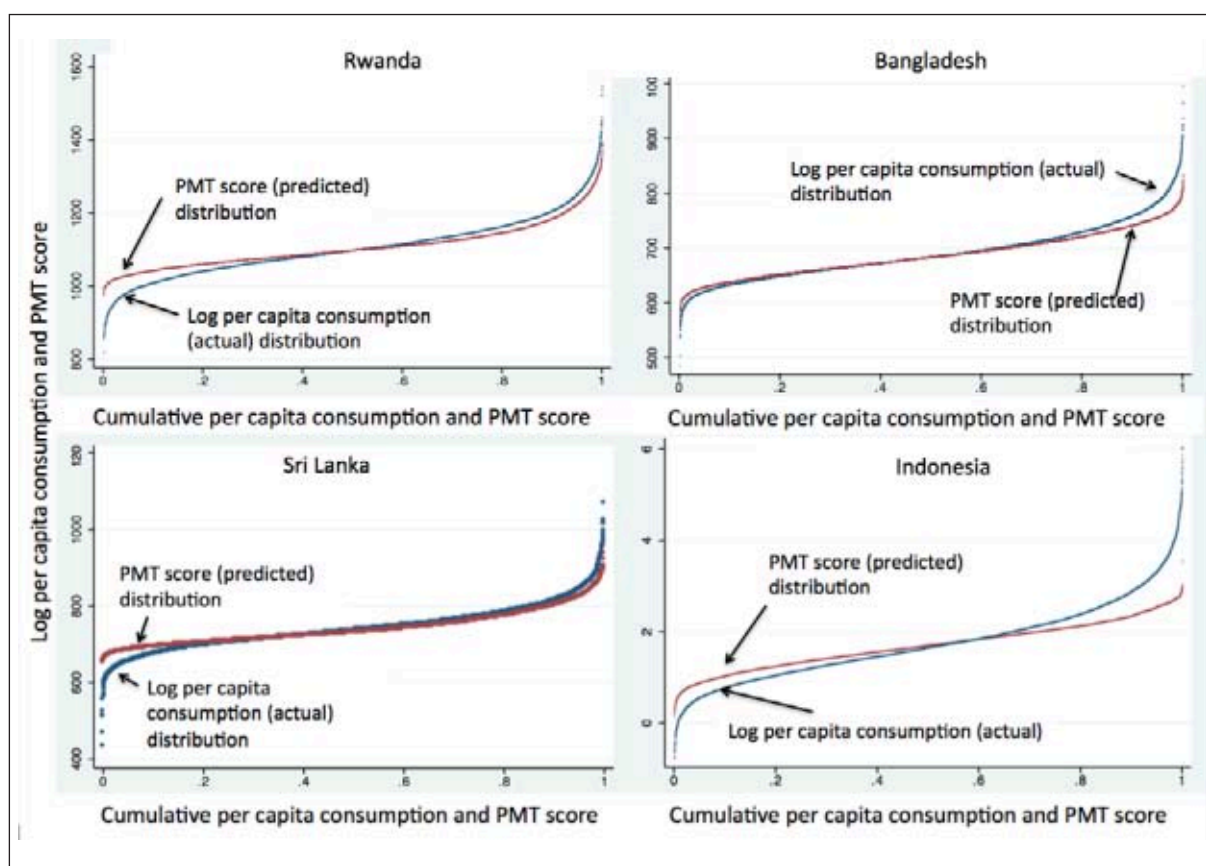
The R-squared can be used to measure regression accuracy, with a perfect fit producing a score of 1. With the Bangladesh PMT it is 0.56, Rwanda 0.44, Indonesia 0.37 and Sri Lanka 0.57.²⁴ These fairly typical results are found elsewhere in the world where R-squared values vary between 0.3 and 0.6.²⁵ The question is whether this is a strong enough basis upon which to build a case for using the methodology.

²³ This is not to say targeting would not improve if more accurate regression models were used, since the ordinary least squares regressions employed in PMTs are somewhat simplistic. The IRIS Center (2005; 2008) has shown that methods such as two-stage, quantile, linear probability, and variance ratio can result in higher accuracy, particularly at the poorer end of distribution. This study does not assess the PMT against more accurate models, but encourages a richer discussion of implementation on the ground.

²⁴ The R-squared provides a measure of the explanatory power of the model. In technical terms it is the fraction of the variation in the dependent variable explained by the independent variables. The regression results this study obtained for Bangladesh are consistent with those found by Sharif (2009) as the same variables and same dataset were used.

²⁵ In its simulation of a PMT for Bosnia the World Bank (2009a) found an R-squared result of 0.488 for its model, and for comparison provides R-squared results for other countries: Armenia, 0.2; Georgia, 0.62; Russia, 0.58; Sri Lanka 0.59. Ahmed and Bouis (2002) calculated an R-squared result for their work in Egypt of 0.43 and stated that R-squared values in Latin America ranged from 0.3 to 0.4.

Figure 4: Distribution of actual versus predicted values of proxy means test scores

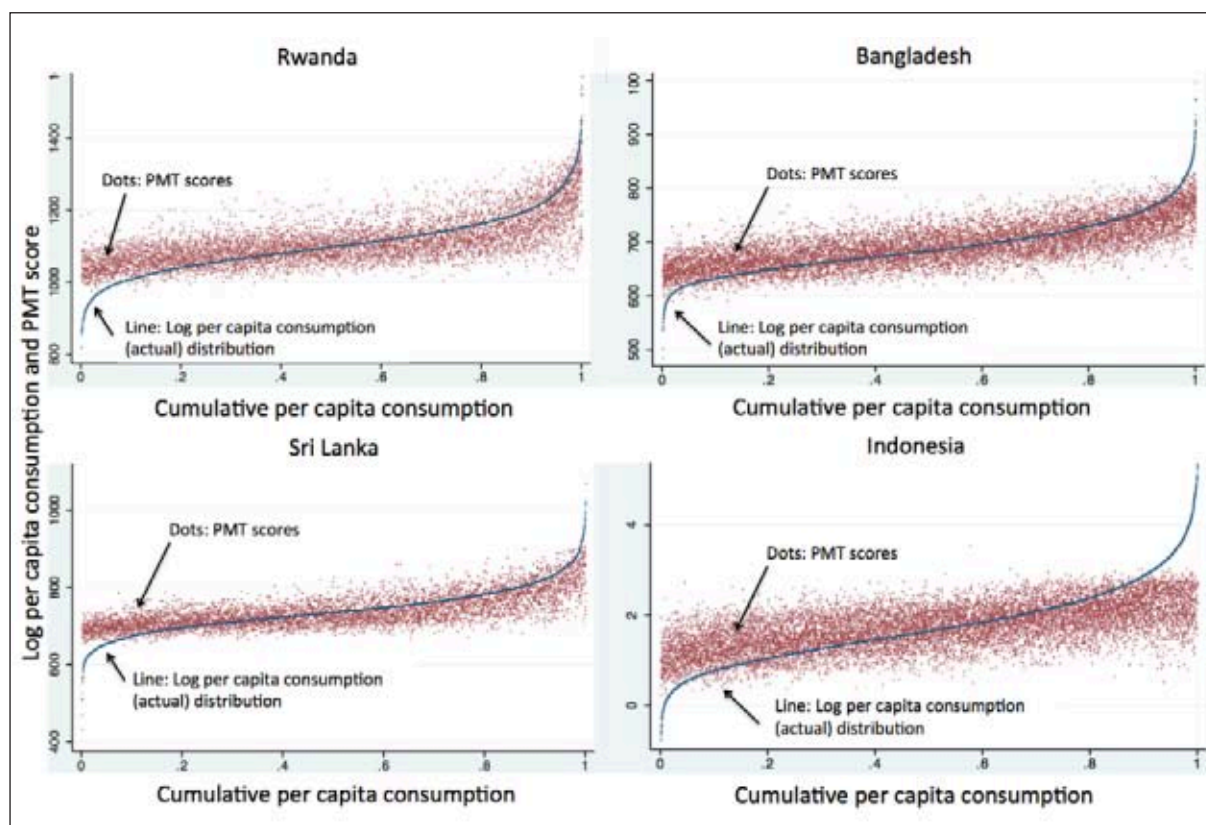


One concern with PMT regressions is the challenge of undertaking modelling among the poorest. Figure 4 shows the cumulative distribution of predicted PMT scores (red lines) across the cumulative distribution of actual consumption (blue lines) from poorest to richest. In Rwanda and Sri Lanka the actual and predicted distributions diverge at the left (poorer) tail—which they also do, to a lesser extent, in Bangladesh. In Indonesia they diverge throughout the distribution. This is partly because the methodology is inherently less capable of reflecting the same degree of variation simply because it includes only a few variables of those available in the full consumption measure. The PMT scores are therefore flatter in shape—there is little to distinguish the households at the left of the distribution from each other.

The second aspect of poor performance is the extent to which the PMT score for individual households reflects actual ranking in the consumption distribution. Figure 5 shows that PMT scores do not perform well in distinguishing between poor households. So the PMT performs the weakest at the point where it would be expected to find the best correlation between assets and consumption.²⁶

²⁶ IRIS (2005) found similar performance of traditional PMT models (ordinary least squares) at identifying the very poor, which they refer to as 'poverty accuracy'.

Figure 5: Actual versus predicted values of proxy means test scores for Bangladesh, Indonesia, Rwanda and Sri Lanka



Another issue is whether some assets are measured accurately. With the PMT, choices are made on how to treat particular assets because asking households for greater detail may prove difficult. For example, a survey may require a ‘yes’ or ‘no’ answer on whether households possess animals, but it may not ask about the number of animals. Huber et al. (2008:46) note this issue among community members in Peru’s Juntos program. This lack of detail influences the accuracy of PMT results if, for example, a household with one cow is treated the same as a household with 100 cows.²⁷

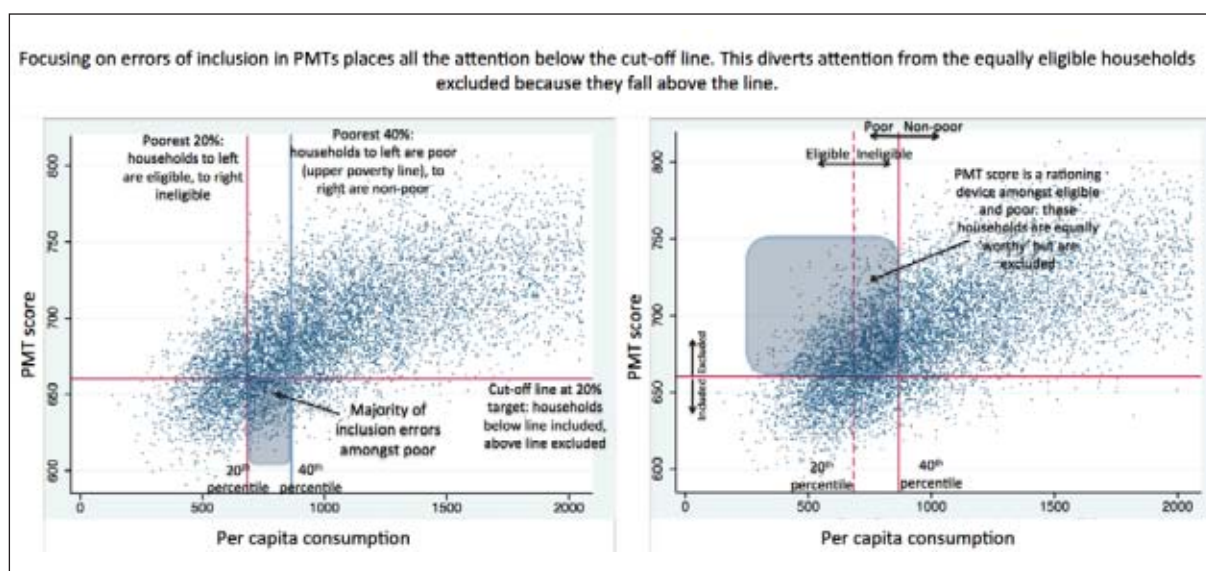
The literature on PMTs often focuses on inclusion errors, possibly because minimising ‘leakage’ to non-eligible households is considered to be of greater importance. It may also be because the methodology performs poorly in terms of exclusion errors. Where a cut-off line is below the poverty line most theoretical errors of inclusion are from households that, while not in the poorest target population, are nevertheless still poor. The location of inclusion errors for Bangladesh households is demonstrated in the left-hand graph of Figure 6, with most errors falling between the 20th percentile of consumption and the 40th (the upper poverty line). Focusing below the horizontal cut-off line might persuade policymakers

²⁷ It is likely to be related initially to problems in the original PMT analysis and design, when analysts decide how to interpret data to specify formulas and determine which questions will be asked in the PMT ‘scorecard’ surveys during implementation.

that the PMT is accurate enough to serve as a basis for targeting, and any incorrectly selected beneficiaries are still poor households, albeit not extreme poor.

This all obscures the fact that the majority of the target population and those in the poorest 40% remain above the cut-off line, as indicated in Figure 6 (right-hand graph), and would be excluded from program benefits. In this example a PMT targeting the poorest 20% will reach 46% of the targeted poorest quintile and 42% of the 'poor'. Although the PMT seems to identify the poorest, in this case it functions as a de facto rationing device, arbitrarily selecting between poor and extreme poor households.

Figure 6: Per capita consumption versus proxy means test scores by household, Bangladesh



3.2.3. Impact of methodological choices

The methodological choices made in applying the PMT have an impact on targeting performance. This study examines three key choices:

1. selection of equivalence scale
2. ways in which missing variables are treated in the household survey (for example, when data is absent due to non-responses, incorrectly completed questionnaires or errors in data entry)
3. decisions over which variables to include.

To examine each choice, alternatives were modelled using cut-off values (targeting 5%, 10%, 20% etc.). Overall, the choice made did not change inclusion and exclusion errors at the aggregate level, but the selection of individual households changed markedly. When testing the choices at a given cut-off value some households were:

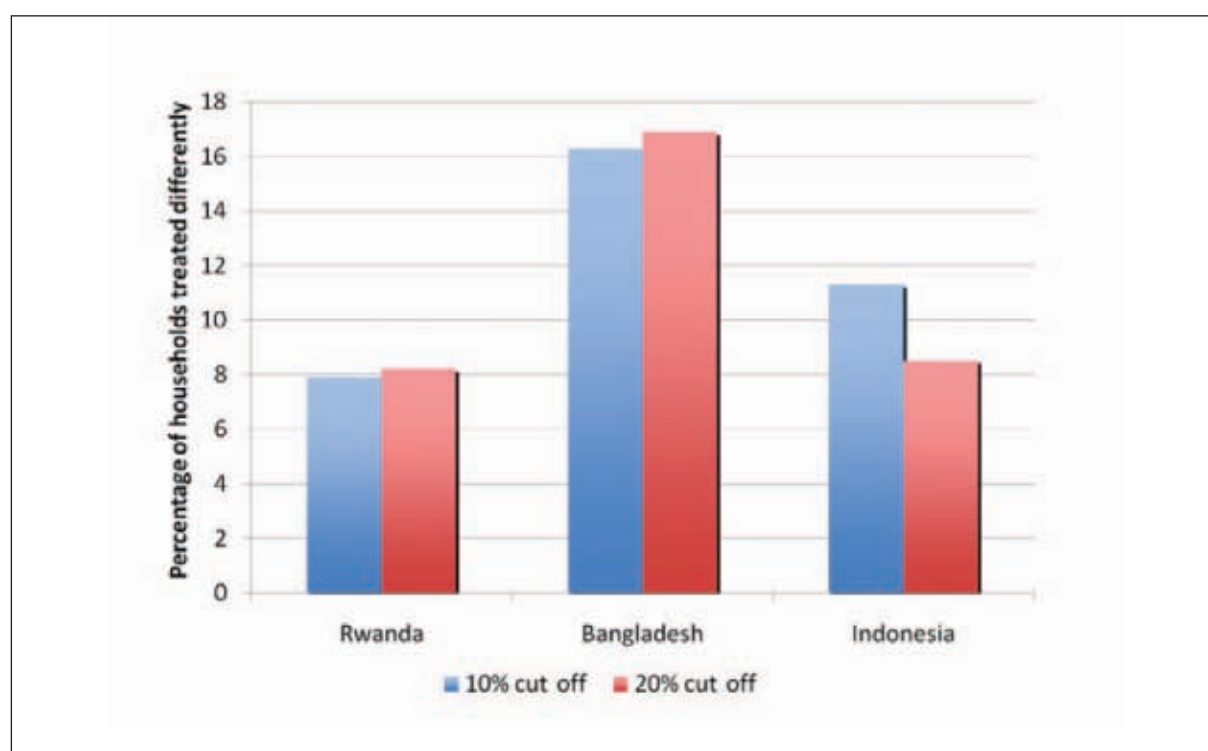
- included in both scenarios (correctly or incorrectly, but consistently)
- excluded in both scenarios (correctly or incorrectly, but consistently)
- treated inconsistently by being included in one scenario but excluded in the other.

This inconsistency is due to methodological choices that help determine PMT scoring formulas. Households score differently—and are ranked differently—depending on the formula used. This then affects eligibility for social protection programs.

As outlined in Section 3.1, the choice of equivalence scale can have a major affect on the way household-level poverty is conceived and quantified. To assess the degree of impact the study used two scenarios: one without equivalence scales and one using an equivalence scale of 1 for adults and 0.5 for children.²⁸

Figure 7 shows these scenarios at 10% and 20% cut-off levels for Bangladesh, Indonesia and Rwanda. The number of households in the target populations treated differently is examined. More comprehensive results against a range of cut-off levels are in Annex 4.

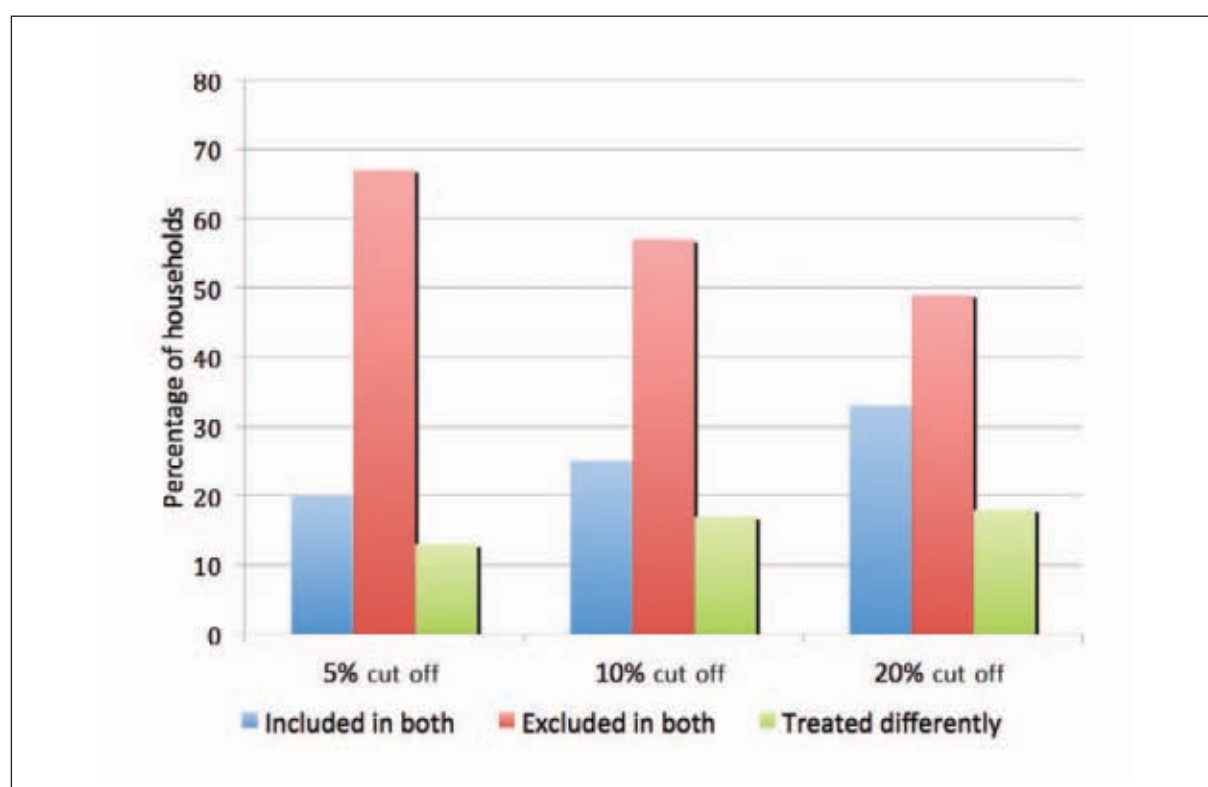
Figure 7: Percentage of households in targeted populations treated differently depending on the use of equivalence scales at 10% and 20% cut-off levels



²⁸ 0.5 for children means children are assumed to require half the consumption of an adult. This does not imply any particular merit for an equivalence scale of 0.5 for children and is used for illustrative purposes.

The impacts on the target group were relatively high. In Rwanda and Indonesia between 8% and 12% of households were treated differently, while in Bangladesh the number reached more than 16%. When combined with normal regression errors only a small proportion of the target population is identified correctly in both scenarios. This is illustrated in Figure 8 using results from Bangladesh, which shows the proportion of target households at coverage levels of 5%, 10% and 20% that are identified correctly using the two equivalence scales. At 5% coverage, only 20% of households in the poorest 5% of actual (per capita) consumption are included under both per capita and 0.5 equivalence scales, whereas 67% are excluded under both scenarios and 13% are included in one but not in the other.

Figure 8: Proportion of target households correctly identified in Bangladesh using two equivalence scale scenarios (per capita and 0.5) at three cut-off levels (5%, 10%, 20%)



Another critical choice is how to deal with missing variables in household survey data. Ideally strong survey design, implementation and data entry practices minimise the number of non-responses recorded in the dataset, but in practice the number can be significant. A decision must therefore be made to impute values to non-responses or treat them as missing, impacting on the estimated values of the coefficients in the regression and hence on the weights used in the PMT score.

Missing variables were problematic with the Sri Lanka data due to the number of missing values for key variables in the PMT regression. This was largely related to asset holdings, where the data should have indicated if a household owned the asset, but in practice for many households the dataset does not have any information. To demonstrate this, the number of observations that ignored missing values is 5619,

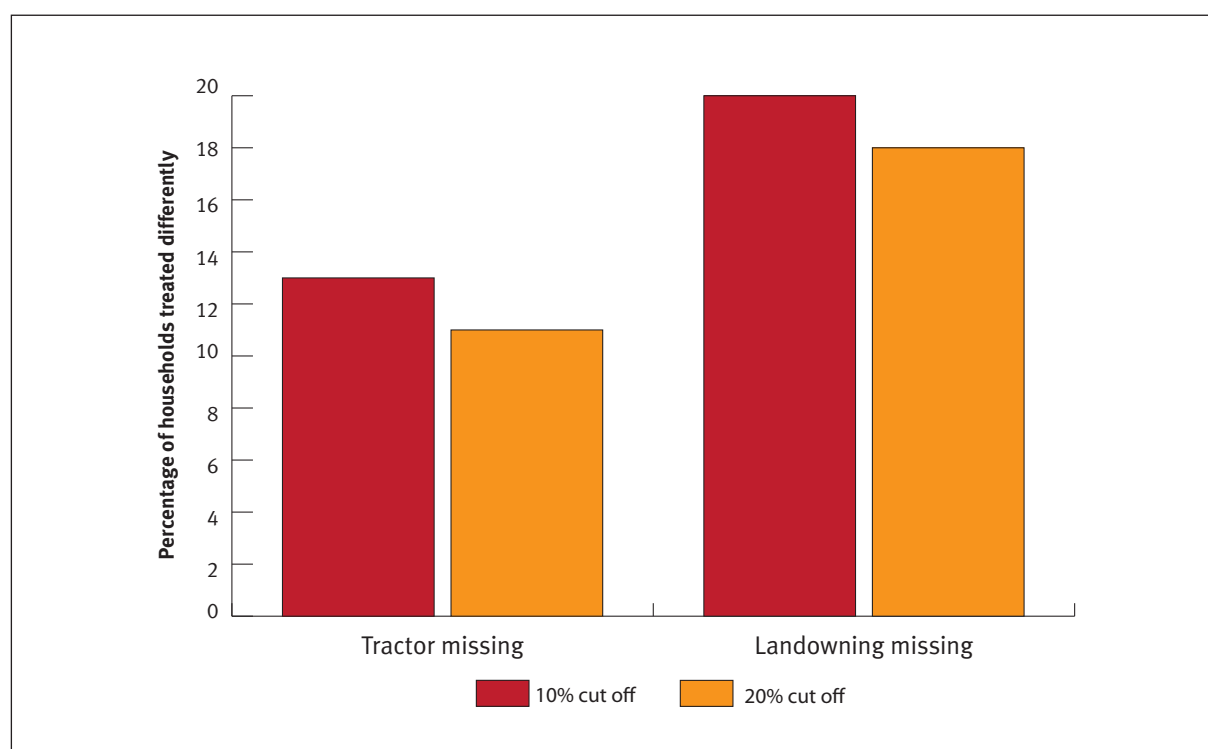
whereas the number of observations included if all missing variables are treated as ‘missing’ is 344.²⁹ In other words, removing the observations with missing variables means that only 344 households have information on all variables used in the model.

Instead of dropping observations with missing values, analysts can choose to impute a value, often the mean or median. In the case of the PMT regression variables, however, a zero value is more appropriate since the missing value is more likely to signify that the household did not own an asset (for example, in the case of landholding, tractor ownership and durable goods). In the context of dummy variables (where a household does or does not possess an attribute), the median or mean value is not that meaningful and would overstate ownership of the assets involved (since the means and medians will always be larger than zero).

This study ran a sensitivity analysis that compared the affects on households of treating all values as zero with treating all values as zero and tractor ownership and landholding size as missing, respectively. There is a relatively large impact on how individual households are affected even though aggregate targeting errors were fairly consistent across scenarios. As Figure 9 illustrates the treatment of landownership is significant, since this variable has a large weight in the original model. For example, at a cut-off value of 10%, 20% of households in the poorest decile of the target population are treated differently, as are 18% of the target population using a cut-off value of 20%. When the tractor variable is treated as missing, impacts are also large, with only 523 observations in the model.

²⁹ This compares with the regression used by Narayan and Yoshida (2005) which had 5257 observations. Although the current study’s results with 5619 observations are similar, there are expected differences in formula weights relating to assumptions about how to treat missing variables and which households to include or drop from the model.

Figure 9: Percentage of households in targeted populations treated differently depending on the treatment of missing variables at 10% and 20% cut-off levels, Sri Lanka



A further methodological choice made is which variables to include in the PMT regression. This issue arose in the Indonesia example for which this study used another dataset—the Indonesia Family Life Survey 4 (IFLS4) from 2007–08—that has a different questionnaire and sampling design from that used for targeting purposes in practice (the Susenas dataset). The IFLS4 represents about 80% of the population nationally but contains a wider variety of household asset variables than the Susenas does. The Susenas dataset, however, has the advantage of being representative down to the district level and it covers the entire country.³⁰

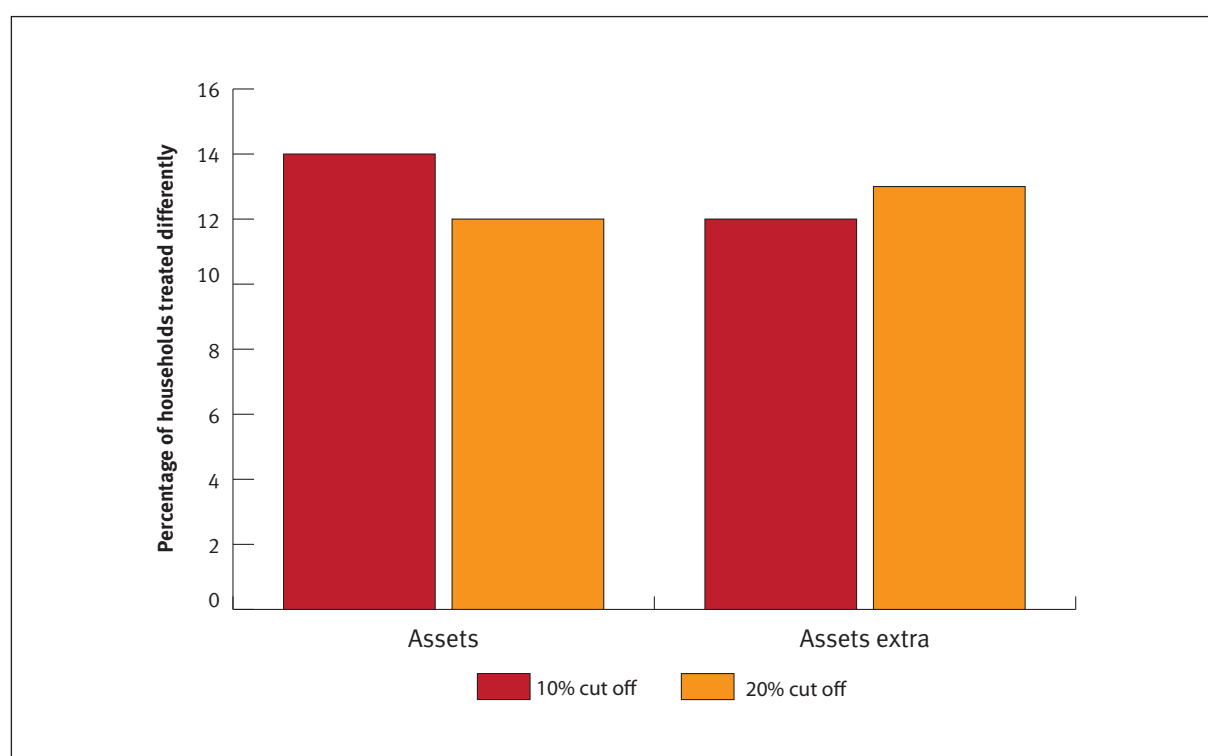
The study’s base model is the one coming closest to that used by the PMT program employing the Susenas dataset.³¹ This was then compared to the base model with two further variations. The first adds household variables commonly used in other countries (for example, assets such as a refrigerator, television and jewellery) and an educational variable. The second variation adds variables on household living conditions (such as if there is stagnant water around the house or if the kitchen is outside) and whether the household has suffered a shock (for example, loss of household member or injuries).

³⁰ Atypically the PMT formulas used by Indonesia’s cash transfer program are different for each district.

³¹ However, the set of community variables (the existence of health and educational facilities) is slightly smaller as there is no direct equivalent in the IFLS4 dataset.

The overall targeting errors are similar to the base model, which is not surprising since the comparison models have similar goodness of fit (as measured by the R-squared values).³² However, as Figure 10 indicates for Indonesia, the impact at the household level is significant, with 11% to 13% of the target population being treated differently.

Figure 10: Percentage of households in targeted populations treated differently depending on the variables used in the proxy means test regression, Indonesia



These three scenarios show how PMT results are affected not only by household-level data but also by methodological choices. One choice is not better than another, although the decision in the original analysis in Bangladesh, Sri Lanka and Indonesia not to compare the affects of using equivalence scales was arguably problematic. The key point is that crucial choices, even seemingly minor ones, need to be given adequate consideration because their impact on results can be high and the implications should be presented to policymakers.

³² The study's original model has an R-squared of 0.374, the 'asset model' has an R-squared of 0.403, and the asset extra model has an R-squared of 0.405.

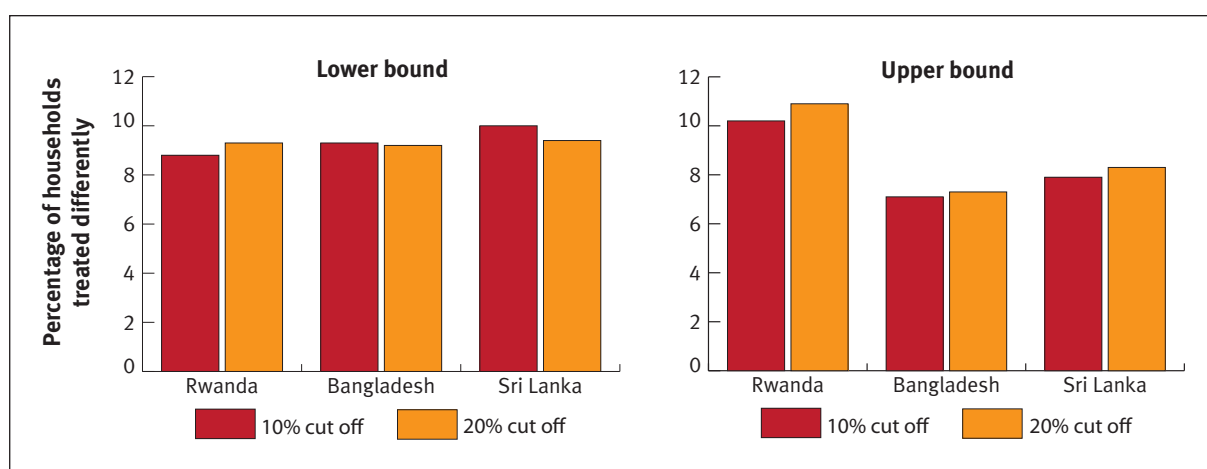
3.2.4. Sampling errors: impact on proxy means testing

Sampling errors raise further implications for PMT accuracy. As discussed in Section 3.1, surveys can only provide estimates of values based on a sample. Their precision can be calculated so that each coefficient in the regression model (that is, the weights in the PMT formula) can be understood as falling within a range.

Two scenarios, one using values at the lower bound of the 95% confidence interval and one at the upper bound, show the impact on households when different weights within a range are used. In other words, there is 95% confidence that the values for the coefficients lie between the lower and upper bound.

Although these scenarios fall within a narrow confidence interval, the impact on individual households is significant, as shown in Figure 11 for Rwanda, Bangladesh and Sri Lanka. The scenarios model the impact on the target population at coverage levels of 10% and 20%. They show that between 7% and 11% of the target populations were treated differently depending on whether the lower or upper bound of the sampling error was considered.

Figure 11: Percentage of households in targeted populations treated differently when the lower or upper bound of the confidence interval on regression coefficients is used



4. Implementing proxy means testing: issues with practice

Once the PMT scoring system is developed, implementing it presents a number of challenges, and its theoretical targeting accuracy should not be substituted for actual accuracy. Some PMT assessments assume ‘perfect implementation’ and therefore produce simulated results that perform better than what is achieved in reality.³³

As with all poverty targeting, implementing proxy means testing can result in some level of error. For example, in Mongolia theoretical exclusion errors were 42% and inclusion errors were 38%. Following implementation, the real errors were 21% and 57% respectively.³⁴

Few studies evaluate how well the PMT is implemented, although one qualitative study (GHK Consulting Limited 2009) examined the pilot phase of a PMT in Pakistan’s Benazir Income Support Programme. This study, as well as research conducted in Mexico, Peru and Nicaragua, and Castañeda and Lindert’s (2005) overview of implementation in Latin America, document a range of implementation challenges.

This section examines how errors can be introduced at each implementation stage and shows that, as with other poverty targeting methods, proxy means testing is difficult to implement well.

4.1. Finding the beneficiaries

To minimise PMT’s exclusion errors all potential beneficiaries need to be surveyed, but this is not easy, especially among the poor.³⁵ Proxy means testing employs two methods to identify beneficiary households. The first—a census method—aims to visit every household. In the second—an on-demand survey—applicants visit centres or offices to register, with enumerators often subsequently visiting houses to verify information.

Census method

Although the census method aims to visit every household, this does not guarantee complete coverage for reasons that vary from country to country. For example, people may not be at home when enumerators arrive and enumerators may not be able or willing to return.³⁶ In Pakistan enumerators often found maps confusing and could not find some households and did not enter some areas for security reasons

33 Examples include Araujo and Carraro (undated); Narayan and Yoshida (2005:13); Jalan and Murgai (2006); World Bank (2005:17f,71); and World Bank (2009a), although the latter acknowledges its assessment assumed perfect implementation.

34 Hodges et al. (2007). The reduction in the exclusion error resulted from a significant (but unintended) expansion in coverage due to weak program implementation and not from an improvement in targeting.

35 Orozco and Hubert (2005:8) note the poor are often the most difficult to register. Coady et al. (2002) argue that insufficient attention is given to poor people not registering for programs and there is a need to increase outreach programs.

36 Adato et al. (2000); Adato and Roopnaraine (2004:16); Huber et al. (2008:46f). In Pakistan enumerators ran out of forms and never returned (GHK 2009).

or because households were too far away (GHK 2009). In Colombia, municipalities prioritised easier-to-reach areas to reduce costs (Castañeda and Lindert 2005:11,13). In other cases households refuse to participate: in Mexico some people hid from census enumerators, concerned about giving authorities personal information (Adato et al. 2000); in Peru some evangelical communities boycotted the census, provoked by preachers that the coding would label them as ‘the Anti-Christ’ (Huber et al. 2008:48).

On-demand survey method

On-demand surveys are often used to reduce costs but can result in higher exclusion errors (Castañeda and Lindert 2005:4). In urban Mexico, 51% of eligible urban households did not register for *Progresa* (Coady and Parker 2005).

The success of on-demand surveys depends on the quality of outreach programs—people need to know that they need to register. In Mexico outreach program quality carried out by municipalities varied.³⁷ Around 25% of eligible households did not hear about the on-demand survey and up to a further 14% did not know where to register.³⁸ In Pakistan the quality of outreach also varied and lack of community awareness affected enumeration quality.³⁹

Another issue with on-demand surveys is the need for household members to get to registration centres. In some cases centres are far away or expensive to travel to. In urban Mexico households with a car had a higher chance of being accepted on to the program, presumably because car owners could more easily travel to registration centres (Coady and Parker 2005:26).

As with other targeting methods another challenge relates to households obtaining the correct documents to be allowed into the program.⁴⁰ This may have contributed to the regressive nature of the PMT in Armenia (Coady et al. 2004). Language can be a barrier for Indigenous people—some households in Mexico were excluded from *Progresa* because they did not speak Spanish (Coady and Parker 2005:26; Adato et al. 2000).

4.2. Capacity of enumerators

Training enumerators to be good at what they do and to work consistently is important to proxy means testing. Capacity varies, however, and can be an area of weakness.

Castañeda and Lindert (2005:13) note that ‘the quality of human resources can significantly affect the interview process.’ In Pakistan, for example, training the organisations sub-contracted to conduct

37 Coady and Parker (2005:5); Orozco and Hubert (2005:3). Peru’s Juntos program did not organise an information campaign and, as a result, many families were unaware of its purpose (Huber et al. 2008:48).

38 Coady and Parker (2005:22) have demonstrated how higher quality communication campaigns are associated with greater participation in urban Mexico.

39 In Pakistan an effective information campaign was considered ‘essential’ for the Benazir Income Support Programme to succeed, but the national radio and television awareness campaign was weak and the evaluation team did not know it existed (GHK 2009: 87f).

40 Coady et al. (2004:70); Huber et al. (2008:49); GHK (2009); Hannigan (2010). In Mongolia, 26% of poor households that did not apply for access to the child money program cited lack of documents as the reason. Obtaining these documents can impose fees, fines and opportunity costs (Hodges et al. 2007:31).

surveys lasted one day and was not considered consistent or adequate (GHK 2009:55,65f,80). This likely had implications for quality at other levels because the program used a system of cascading training. Training the large number of enumerators needed to conduct a survey can also be challenging, with one Pakistan survey method requiring 148 000 staff to be implemented nationally.

PMTs can also be administratively demanding. While some countries have stronger administrative capacity, engaging sufficient personnel with the right skills can still be challenging. Although willing to provide training, Chile, for example, could not source enough qualified enumerators to run its PMT (Grosh and Baker 1995).

Another issue is enumerator honesty. In Mongolia one study euphemistically noted ‘large anecdotal evidence that documents the subjectivity of the process’ (World Bank 2006:3). Bribery concerns in connection with proxy means testing have been noted in Chile despite its civil service’s good reputation (Grosh and Baker 1995). In Pakistan local enumerators surveyed friends and relatives, a conflict of interest (GHK 2009:117,123).

Better supervision helps improve enumerator performance (Castañeda and Lindert 2005:6), although supervisors can face excessive demands, especially when budgets are constrained. In Pakistan supervisors were expected to quality control every questionnaire at the end of each day. This meant reviewing up to 300 questionnaires from up to 10 enumerators daily (GHK 2009:54f,140,171).

Enumerator work load and the physical strain of interview conditions can be demanding and incentive payments can drive enumerators to complete as many surveys as possible in a day, potentially compromising quality. In Costa Rica and Pakistan enumerators struggled to visit households due to the heat and in the latter some had to be replaced (GHK 2009:71,112,17; Viquez 2005:22). Despite this enumerators in Pakistan conducted more than double the expected number of interviews (GHK 2009:61).

Enumerators who are not adequately trained or who are rushed to meet targets are less able to deal with issues that arise during surveys. In Pakistan some enumerators found it difficult to identify the household; while officially defined as a group of people sharing a cooking pot, some enumerators conflated it with a married couple.⁴¹ Other challenges were knowing how to deal with households with migrant workers, seasonal workers, the status of livestock held but not owned, and the position of servants in better-off households.

In Costa Rica, Mexico and Peru enumerators changed survey results or used their own judgment to make decisions on eligibility when it was clear that the PMT result was incorrect (Orozco and Hubert 2005:24; Viquez 2005:19; Huber et al. 2008:46f). While this may have been for the right reasons, it has the potential for corrupt practice.

41 GHK (2009). A similar situation was also noted in Nicaragua (Adato and Roopnaraine 2004:69).

4.3. Verifying indicators

Part of PMT's attraction is the assumption that proxies can be easily observed and verified. In theory this reduces the chance that households will supply false information. In reality observing and verifying proxies, including age, education and occupation, is not always easy. Often people do not have birth certificates, they can easily under-report their education and provide false occupational information. Productive assets, such as livestock, land size and land ownership, can be difficult to verify and durable goods, if relatively small, can be removed from the house during inspections (Narayan and Yoshida 2005). In Mexico verifying whether a car parked outside a person's house actually belonged to them was not always possible, with some claiming they did not own the car but rather were looking after it (Adato et al. 2000).⁴² In Palestine the Ministry of Social Affairs estimates that half of households gave false answers to its initial PMT survey.

Visiting houses to verify proxies is critical to proxy means testing success although this is not automatic (Viquez 2005:22). In Armenia the PMT was regressive and one reason may have been the absence of household visits (Coady et al. 2004). In Pakistan enumerators only entered 31% of houses (GHK 2009:115), and in both Pakistan and Peru there is evidence that some people were interviewed in the street without household verification (Huber et al. 2008:45; GHK 2009).

Most PMT household surveys take about 20 minutes with longer interviews significantly increasing implementation costs and shorter ones compromising verification quality. In most cases in Pakistan enumerators did not verify proxies through observation, probe answers or cross-check information. Respondents sometimes provided information on other households and data was collected by text or phone (GHK 2009).

Assuming that proxies are easy to verify can be a significant weakness and lead to insufficient rigour during household visits. In urban Mexico households answer questions on proxies well in advance of a home visit and therefore have time to change matters before the enumerator arrives (Orozco and Hubert 2005:20). That households may manipulate proxies led Chile to stop making proxies or weights public because of concerns with fraud (Grosh and Baker 1995). The World Bank (2009a) proposes that proxy variables should be changed regularly to prevent people from gaming the system. This could undermine PMT's accuracy because the best explanatory variables are unlikely to change over time (Coady et al. 2004).

4.4. Identity of the interviewee

Interviewee identity can influence the reliability of PMT survey answers. Normally the household head should be interviewed, but if not home when the enumerator visits then others, including children, are allowed to provide survey information, even though children are less likely than their parents to be fully aware of their household situation. In Mexico respondents only have to be over 15 years of age to be interviewed; in Peru children as young as 12 have been interviewed; and in Cambodia schoolchildren

⁴² Coady and Parker (2005:33) suggest that reporting false information could be a problem in Mexico. In rural Mexico households that exaggerate their wealth for fear of being stigmatised may also exaggerate their poverty once they see the rewards of joining the program (Adato 2000).

have completed questionnaires.⁴³ Gender is also an important factor: in Pakistan it was noted that women gave different answers to men that were often more accurate (GHK 2009).

4.5. Community verification

Community verification is often proposed as a way to address inclusion and exclusion errors and expose households that manipulate information. Beneficiary lists are publicised so communities can challenge beneficiary selection, assuming that the community knows who is poor or better-off and that those who falsely indicate they are poor will be stigmatised.

However, there is little evidence that this works, partly because community meetings sometimes do not take place. Community verification through meetings under *Progresa* in Mexico was rare (Adato 2000)⁴⁴ and the same applies to Nicaragua and Peru (Adato and Roopnaraine 2004:17,21).⁴⁵ Publicly questioning beneficiary choice can be divisive (Grosh et al. 2008:118) and making community verification work in urban areas would present particular challenges in the absence of tight-knit communities.

4.6. Political interference

PMT can be subject to political interference. In Colombia Castañeda and Lindert (2005:11) note that politics influenced program area selection. In Mexico Adato et al. (2000:9) suggest some areas were not visited because enumerators had their own political agenda. In Pakistan's Benazir Income Support Programme there were examples of politics influencing the communities or households in the PMT survey (GHK 2009:15). Inclusion of appropriate safeguards would reduce errors relating to such interference.

4.7. Recertification and poverty dynamics

Although household welfare and income can change rapidly the PMT methodology captures households at one point in time and assumes their incomes are static. The reality for many countries is that many of the poor are clustered just above or below the poverty line and people move in and out of poverty throughout their lives. For example, in Indonesia 58% of households were poor at least once during 1998–99 although the poverty rate was only 37%, and 38% of households that were poor in 2004 were not poor in 2003 (World Bank 2006:159). In Pakistan, 67% of households were vulnerable to poverty from 2001 to 2004, compared to a national poverty rate of around 27% (World Bank (2009c:77). And in urban Mexico only 7% of extreme poor households in 2002 still had the same status in 2007 (Rascon and Rubalcava 2008).

⁴³ Adato et al. (2000); Huber et al. (2008:45); Fiszbein and Schady (2009:71).

⁴⁴ Several reports (Orozco and Hubert 2005:3,8; Grosh et al. 2008:118; Hanlon et al. 2010:109) indicate that although community meetings take place in Mexico households are rarely removed. Skoufias et al. (1999:iv) note only 0.1% of beneficiary selections were disputed.

⁴⁵ In Peru's Juntos program community meetings were led by program officials who indicated who should be removed from the program (Huber et al. 2008:46f).

In countries like these following initial targeting the poverty status of beneficiary and non-beneficiary households is likely to change relatively quickly. Some beneficiaries move out of poverty and some non-beneficiaries become poor. Even the possession of some assets or proxies can change dramatically over a few short years. As a result targeting accuracy can degrade, potentially rapidly.

Re-certification and keeping up-to-date registration databases is therefore critical for program accuracy. While this is implicitly acknowledged—for example in Latin America household re-certification reportedly takes place every two to three years⁴⁶—actual frequency is uneven. In Mexico re-certification should occur every six years in rural areas and every three years in urban districts (Castañeda and Lindert 2005), although six years after initial targeting households still not been recertified (Orozco and Hubert 2005:8). Jamaica's PATH program initially planned for re-certification every two years. This was then changed to every four but does not appear to have happened (World Bank 2009b).

But even with sufficient re-certification the PMT is not a good targeting mechanism in major external crises, such as floods, droughts or earthquakes, or in the large or small crises that hit individual households as part of day-to-day life, such as if a breadwinner becomes ill or unemployed causing a significant fall in income (Coady et al. 2004; Johannsen et al. 2010). With major external crises, the logistical requirements for mobilising a PMT are such that it cannot be implemented quickly. With individual household crises, there may be no immediate change in assets or other proxy indicators, especially those with the greatest explanatory value (such as number of household members, age and education).

4.8. Responsiveness to the household life cycle

The PMT methodology is potentially weak in responding to the household life cycle for various reasons including shifts in the value of assets, the stage at which assets are accumulated and their depreciation.

As households mature, they are likely to accumulate assets without necessarily experiencing an increase in income, particularly in informal economies. Younger households may have similar incomes as mature households but fewer assets (that take time to accumulate). In addition, while assets held by more mature households may be relatively old—a 15-year-old refrigerator for instance—the PMT often does not account for depreciation. For example, in Indonesia certain proxies discriminated against households headed by those over 40 years of age. Possessing assets such as tin roofs and cement floors reduces the chances of older people being identified as poor even when their income is the same as younger people who have not accumulated such assets (Hannigan 2010).

Inheritance can also be an issue, including of assets not worth a great deal, as expressed by this poor Jamaican:

Not because you have furniture mean you have money. Like my mother dead left plenty of things for me but me can't afford food and the fridge need to fix and me can't fix it.
(Mathematica 2007)

⁴⁶ Grosh et al. (2008:114f); Rawlings and Rubio (2003); Coady et al. (2002).

The elderly are particularly susceptible to discrimination by proxy means testing since, while they may have experienced a significant reduction in income as they become less able to work, the assets they have accumulated during their working lives may count against them. Even education may discriminate against older people since their qualifications may be of little use in entering the labour market as they grow older. For this reason Jamaica's PATH program has adjusted weightings for older people living alone.

4.9. Appeals

All social protection programs should have an appeals system, especially ones using the PMT, where up to 50% or more of eligible poor may be excluded. However appeals are not a regular feature of the methodology, meaning a large number of people excluded from programs have no recourse. Chile, Costa Rica and Brazil do not have a formal appeals process, and while Mexico's *Oportunidades* has one around service delivery and benefits (the Citizen Attention Program) it does not deal with complaints relating to household selection or program eligibility.⁴⁷ Kenya's Orphans and Vulnerable Children Cash Transfer program has a grievance system but it has not been activated because authorities fear too many appeals (Ward et al. 2010:14).

While an appeals system should be an integral part of any social protection program, designing one that fairly, objectively and transparently addresses eligibility is a major challenge. One difficulty with proxy means testing is determining the basis upon which appeals can be made. Basing it on income poverty might make it difficult to implement if half or more of the eligible poor are excluded: if all of them appealed the methodology would become ineffective. However if appeals are based on scores against proxies there would be no means of redress for large numbers of otherwise eligible households.

Another difficulty is designing an appeals system that is supported by a transparent targeting system. Some recommend transparency on a human rights basis (Sepulveda 2009) or to promote the system's credibility (Castañeda and Lindert 2005:9), but in reality the 'black box' nature of the PMT methodology does not easily lend itself to transparency. Many find the methodology difficult to understand—as seen in Mexico and Indonesia⁴⁸—and some countries withhold information on proxies and weightings to control fraud (Castañeda and Lindert 2005:32). Chile, for example, used to make information on proxies publicly available but stopped because of concerns that people may manipulate them (Grosh and Baker 1995). In Armenia posters with proxies and weightings were put up in public places but people were confused by the information (Coady et al. 2002:30).

⁴⁷ Castaneda and Lindert (2005:60); A. Hijuelos, Advisor to The National Coordinator, *Oportunidades* (personal communication, June 1, 2011).

⁴⁸ Adato (2000); Adato et al. (2000); Orozco and Hubert (2005:3); Hannigan (2010).

4.10. Addressing intra-household poverty

Because household surveys cannot provide information on individual poverty levels proxy means testing cannot address variations in individual wellbeing within households and, despite suggestions to the contrary, the PMT cannot identify poor individuals.⁴⁹ The best the methodology can do is assume that everyone in a household will receive the same share of a social protection program benefit. But evidence across Africa indicates this is not the case: older people are increasingly excluded by households that prioritise children and working adults.⁵⁰ Since proxy means testing cannot identify and target those who are poorer or more deserving within households, it should not be used to target individual benefits such as old age pensions, child grants, disability or unemployment benefits.⁵¹

⁴⁹ See, for example, Grosh and Leite (2009:165). The PMT cannot be used to identify families either but some programs that use the methodology, such as in Pakistan and Costa Rica, target the family not the household (Viquez 2005:12f; GHK 2009:11,123).

⁵⁰ Van der Geerst (2002); Miguel (2005); Aboderin (2006); Kidd and Whitehouse (2009).

⁵¹ Chile recognised this in its 2008 reforms to its pension system and moved away from using the PMT to target pensioners to a system focused on individual income: See: <http://www.pensionreforms.com/Preview.aspx?274>.

5. Social and political costs of the proxy means test methodology

The costs of proxy means testing—as with other forms of targeting—can go beyond the financial and include social and political costs.⁵² Although it is difficult to quantitatively measure these costs they can be significant and should be included when considering targeting options (Adato and Roopnaraine 2004:79).

5.1. Social costs

There is growing evidence that targeting, including proxy means testing, can impact negatively on community cohesion.⁵³ A number of studies document how the methodology can create divisions in communities between beneficiaries and non-beneficiaries. Due to the PMT's targeting inaccuracies both groups can be relatively similar in composition, including poor and non-poor households. Discord can be exacerbated by opportunities for beneficiaries to engage in group activities and building new social networks that exclude non-beneficiaries.⁵⁴

Studies in Mexico and Nicaragua point to feelings of despair, resentment, anger and jealousy among non-beneficiaries⁵⁵, even leading to conflict. In some communities in Mexico non-beneficiaries threw rubbish onto streets while beneficiaries were cleaning them; in others, non-beneficiaries knocked down recently mended fences (Adato 2000:32). Non-beneficiaries have withdrawn their labour from voluntary community activities, as in Indonesia where 'huge exclusion errors are causing tension and unrest.'⁵⁶ The initial distribution of Indonesia's *Program Keluarga Harapan* provoked stone throwing and in one community a building was burnt down and program facilitators saw their work as dangerous (Hannigan 2010). In Lebanon proxy means testing led to riots in some refugee camps.⁵⁷

The selection of some beneficiaries over others is often resisted, particularly in communities where the separation between poverty levels is slight or imperceptible. One community member in Mexico expressed a typical concern: 'You ask us to tell you how we can know who is poor and who is not and we are telling you that here we are all poor.'⁵⁸ Similar concerns have been reported in communities in

52 Sen (1995) discusses this in relation to targeting in general. Also Coady et al. (2004:7ff); Kidd et al. (2011).

53 Adato (2000); Adato et al. (2000); Gonzalez de la Rocha (2003); Adato and Roopnaraine (2004); Gallardo (2008); Huber et al. (2009); Hannigan (2010).

54 Adato (2000); Adato et al. (2000); Adato and Roopnaraine (2004).

55 Adato (2000); Adato et al. (2000); Adato and Roopnaraine (2004).

56 Hannigan (2010).

57 Georgia Rowe, personal communication.

58 Adato (2000:23).

Nicaragua, Peru and Indonesia.⁵⁹ ‘We are all poor’ is an expression of egalitarian values and a reaction to the divisions that targeting creates.⁶⁰

Social costs can take other forms and stigma can occur in any non-universal program (Grosh et al. 2008:104), as in Mexico when some refused to admit how poor they were when initially surveyed.⁶¹ One facilitator in Mexico stated that:

Poor people can’t eat like other people. So, due to shame they said that they were eating well. There is the shame of eating beans and chillies and, at times, the families don’t have food. It is for this reason that many feel hurt and don’t say anything or think it is better to hide and not respond. So, many families who are really in need, many poor families, still don’t have *Progresas*. (Adato et al. 2000)

5.2. Political costs

One way to evaluate PMT’s political costs is to compare the methodology with more inclusive targeting mechanisms and the amount of funding each receives.

Conventional economic theory argues that programs excluding the middle-class are less likely to receive significant public funding.⁶² Comparing the costs of programs using proxy means testing with programs such as universal old age pensions bears this out. The latter tend to have significantly larger budgets.⁶³ For example, while many of the largest programs using proxy means testing have budgets of no more than 0.4% of a country’s gross domestic product, many universal pension programs command budgets in excess of 1%.⁶⁴ The consequence of this for poor households is that they may well receive smaller benefits from a tightly targeted PMT program compared to a universal program with broader political support.⁶⁵

59 Adato (2000); Adato et al. (2000); Adato and Roopnaraine (2000); Huber et al. (2008:49); Hanlon et al. (2010:110); Hannigan (2010); Kidd (1999).

60 Cf. Kidd (1999); Ellis (2008).

61 Also Fiszbein and Schady (2009:78). Reverse stigmatisation can also occur. In Mexico and Nicaragua poor people unfairly excluded felt stigmatised because they could not participate in many of the program’s community activities and because they and their children were markedly poorer than the beneficiaries (Adato 2000:26; Adato and Roopnaraine 2004).

62 Sen (1995:14); Moene and Wallerstein (2001:22); Mkandawire (2005:13); Pritchett (2005); Fiszbein and Schady (2009:59f); Kidd et al. (2011).

63 Kidd et al. (2011).

64 Mexico’s *Oportunidades* is probably the world’s largest program using proxy means testing.

65 World Bank (1990); Moene and Wallerstein (2001:22); Kidd et al. (2011).

6. Conclusion

Within the social protection field proxy means testing has become an increasingly popular way to target poor people in developing countries and is widely regarded as effective. This methodology estimates household income by associating indicators or ‘proxies’ with household expenditure or consumption. This study examined PMT’s accuracy, objectivity, transparency and ease of implementation, filling a gap in the literature and providing evidence for policymakers and practitioners on its usefulness in particular settings.

Using data from Bangladesh, Indonesia, Rwanda and Sri Lanka, the study looked at the regression accuracy of the PMT methodology at coverage levels between 5% and 40% of the population. It found that exclusion and inclusion errors vary between 44% and 55% respectively when 20% of the population is covered and between 57% and 71% respectively when 10% is covered. At these coverage levels—which are realistic for many programs in developing countries with limited financial resources—eligible households have no better than a one-in-two chance of being selected, and in some cases even view proxy means testing as a lottery.

Household selection is also influenced by methodological choices that may be sound for a household survey analysis, where broad trends and patterns in poverty can be used to inform policy choices, but in the context of a PMT these choices selectively determine which households are poor. Common methodological choices for which there is no right or wrong approach include equivalence scales, treatment of missing variables and deciding which variables to include in the PMT regression. Relatively small differences in any of these areas can lead to significant differences between how households are identified. When different equivalence scales are used, for example, up to 16% of Bangladeshi households may find themselves identified as poor, or not; and in the case of missing variables up to 20% of Sri Lankan households will be treated as poor, or not. In Indonesia the study found that up to 13% of households will be treated as poor or not, depending on which available variable is used as proxies to determine poverty status.

These findings highlight the significant discretion that exists in making methodological choices. In the absence of sensitivity analyses that present the implications of different choices, there is an impression that inaccuracies in the PMT regression build cumulatively upon inaccuracies in the household survey analysis. This combination of theoretical errors means a majority of eligible poor households may be permanently excluded from social grant benefits as a result of PMT scoring, especially in environments where there is no or infrequent re-certification or a sound appeals processes.

Implementation adds a further set of challenges that can lead to additional exclusion and inclusion errors. While it is difficult to know the size of the errors introduced by implementation—they vary from context to context as the Mongolia example shows—they can be significant. Errors can be introduced through implementation practices, from locating potential beneficiaries to verifying the information households provide. PMTs are also administratively demanding, and even countries with stronger capacity can still face difficulties engaging sufficient personnel with the skills needed to apply the

methodology. Other factors affecting implementation include enumerator honesty, political interference and difficulties observing and verifying proxies, including age, education and occupation.

What is clear is that ‘perfect implementation’ should not be assumed—as occurs in some studies—and that the PMT’s theoretical targeting accuracy should not be substituted for actual accuracy. As with any targeting mechanism, it is possible to reduce implementation errors but this can increase financial costs.

Although this study’s findings show that the errors inherent in the PMT methodology are significant, there is so far no evidence that other forms of poverty targeting perform any better in developing countries with large informal sectors, weak administrative capacity and low fiscal space.

Other approaches to social protection propose different ways of targeting that may be easier to implement, socially less divisive and politically popular. For example, there are growing calls for developing countries to build social protection systems that direct resources to vulnerable groups—such as older people, children, persons with disability and the unemployed—on a universal basis.⁶⁶ Sound evidence exists that targeting vulnerable groups on a universal basis may generate a higher level of political support than an approach focused on developing large-scale ‘safety nets’ that may exclude a majority of the poor. While such universal targeting approaches have been adopted by many developed countries and an increasing number of developing countries, there are important trade-offs relative to other forms of targeting that need to be considered, relating to financial, social and political costs, and coverage of the poor.⁶⁷

A key concern of this study is whether PMT is being promoted with a sufficiently realistic appreciation of what it can achieve. This study aims to contribute to a better understanding of proxy means testing’s advantages, disadvantages and trade-offs in comparison to other targeting methods, and help practitioners and policymakers make better decisions on whether to adopt the PMT methodology.

66 Townsend (2007); Behrendt et al. (2009); Ellis (2010); Kidd and Calder (2011).

67 Farrington and Slater 2009.

Annex 1: Methodology for devising proxy means test scores

The PMT scoring system is developed by analysing household survey data using these general steps:

1. Consumption data is analysed to estimate total household consumption. This involves making technical decisions regarding adjustments for seasonality, constructing localised price indicators and deciding whether to include different kinds of consumption items (for example, durables and one-off expenditures such as weddings). For present purposes it was therefore important to use 'official' consumption aggregates.
2. Household consumption is adjusted for the number of people in the household. In many instances per capita consumption is used, however this assumes both adults and children consume the same amounts. Another approach is to use an equivalence scale to calculate per-adult equivalent consumption, assuming, for example, that children have a consumption weight of 0.5 of each adult.
3. A regression model of the correlates of poverty is estimated. Often, the World Bank uses a stepwise regression technique, whereby the statistical software iterates through possible combinations of variables, keeping those that are significant and dropping those that are not. This is, however, a highly technocratic approach, which prioritises a good 'fit' of the model over selection based on a more nuanced understanding of the underlying poverty situation.
4. The PMT regression then provides the PMT scoring system with the estimated coefficients serving as the weights.
5. PMT scores are calculated for each household based on these weights. The scores are essentially predicted values of consumption.

Targeting efficiency and poverty impact of the PMT scoring system can then be evaluated, comparing how well the predicted values in the PMT score compare with actual consumption at different program coverage levels. For example, with a targeted coverage of 20%, targeting efficiency measures the extent to which the PMT score accurately predicts if a household's consumption is in the bottom two deciles.

Annex 2: Variables used in proxy means tests for each study country

Variable	Rwanda	Bangladesh	Sri Lanka	Indonesia
Human capital variables				
Education of household head		X	X	
Highest level of education in household	X			X
Female literacy	X			
Number of children in school			X	X
Demographic characteristics				
Household size	X	X	X	X
Number of children	X	X		X
Gender/marital status of head (for example, widow)		X	X	X
Age of household head		X	X	X
Dependency ratio				X
Household assets				
Own home			X	X
Type of wall construction	X	X	X	X
Type of roofing material	X	X		
Type of latrine	X	X	X	
Number of rooms per capita			X	
Type of cooking fuel	X			X
Radio, television and other forms of electronic or communication devices	X	X	X	
Bicycle, car, motorcycle or other means of owned transport	X	X	X	
Furniture	X			
Access to electricity		X		
Cooker, heater, fan, air conditioning		X	X	X

Variable	Rwanda	Bangladesh	Sri Lanka	Indonesia
Productive assets				
Landholding size	X	X	X	
Livestock (for example, cows, goats and sheep)	X	X	X	
Use of fertiliser	X			
Livelihood options				
Agricultural or nonfarm wage labour	X	X		
Non-farm independent business	X			
Agricultural production of cash or staple crops	X			
Receipt of foreign remittances		X		
Sector of work (for example, informal, industry or agriculture)			X	X
Geographic dummies	X	X	X	
Community variables				
Presence of midwife				X
Population density				X
Asphalt road				X
Bank in community			X	
Divisional secretariat in community			X	

Annex 3: Explaining higher errors found in other analysis of proxy means tests

Errors of inclusion and exclusion are by definition symmetrical whenever the program size is equal to the target population. In the type of simulation undertaken to test targeting efficiency of a program aiming to reach a given percentage of the population, this study would expect the errors to be the same. However, in other assessments of PMT scoring across a range of countries errors of inclusion and exclusion are often different. Table A, for example, provides results from Pakistan, where at 10% coverage errors of exclusion (under coverage) are given as 88.2% and errors of inclusion (leakage) are given as 50.9% (World Bank 2009c: 87). A similar scenario appears in work undertaken in Sri Lanka (Narayan and Yoshida 2005).

Table A: Targeting performance of the Pakistan proxy means test formula by percentage

	%	%	%	%
Target group (poorest percentage)	10.0	20.0	25.0	30.0
Coverage	2.4	13.0	19.0	25.7
Under coverage	88.2	61.0	52.2	42.8
Leakage	50.9	40.0	37.0	33.2

These counter-intuitive results are not explained in the text in the Pakistan example. However the analysis for Bangladesh (which also shows differing inclusion and exclusion errors) provides some clues as to what is happening.⁶⁸ Rather than using cut-off points based on where households fall in the distribution of the PMT score (covering all households with the lowest percentage of PMT scores), the authors instead use cut-off points based on the bottom percentage of *actual* household consumption (covering all households that fall in the bottom percentage of the consumption distribution). In other words, they use cut offs of *actual* rather than *predicted* values.

This leads to additional errors which are referred to as ‘coverage’ in the studies mentioned above and explains why errors of inclusion and exclusion do not add up. In the Pakistan example in Table A, when targeting the poorest 10% the program (as simulated in the analysis) would only reach 2.4% of the population. Similarly, targeting the poorest 20% would reach 13%, and at 30% it would reach 25.7%.

This approach to establishing cut-off values is problematic for two reasons:

1. It misunderstands the point of simulating targeting efficiency in the household survey data. The point of the PMT methodology is that, in the real world, there is not full information on a household’s consumption, only proxy indicators. So a program with a budget large enough to cover the poorest

⁶⁸ That this methodology must be pieced together from different assessments highlights some of the issues around transparency and clarity in the way proxy means testing is presented and assessed.

20% would calculate PMT scores for all households in the program area and would then be able to include the 20% of households with the lowest scores. Simulations based on targeting a given percentage of the population in terms of actual consumption is, therefore, not that relevant to the way in which programs are implemented.

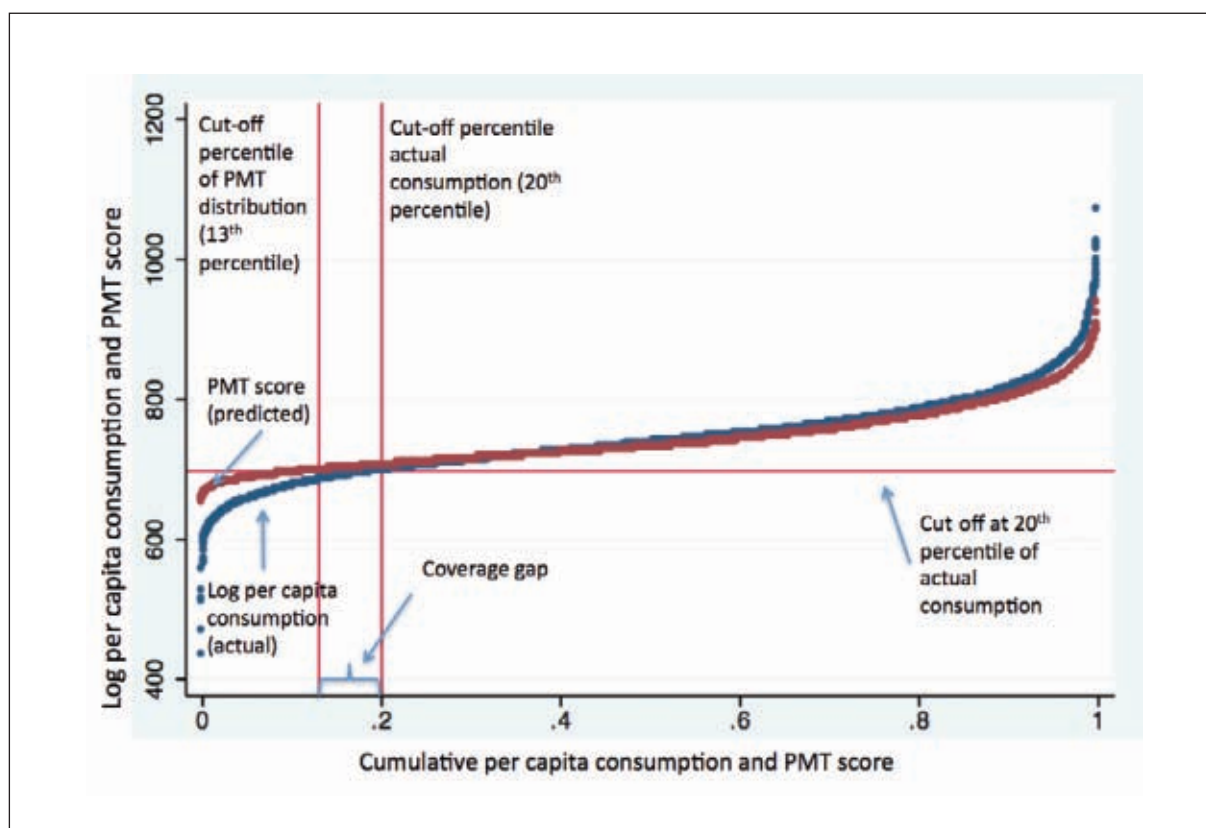
2. It muddles the analysis of inclusion and exclusion errors by introducing an additional coverage error. As a result it is difficult to understand how much exclusion error is related to the poor predictive power of the PMT formula and how much is related to coverage that is lower than targeted.⁶⁹

With Pakistan (Table A), the under coverage (exclusion) errors of 88.2% do not reflect the errors if the program covered 10% of the population, but rather if it covered only a quarter of that (2.4%). This tells what the exclusion errors would be if programs only spent a quarter of their budgets, which is not helpful for ex ante analysis, since the ‘under-spending’ relates only to theoretical rather than actual information.

The reasons for the large differences between targeting the poorest percentage of the population based on PMT (predicted) scores or consumption (actual) are important to understanding how well the PMT methodology identifies the poorest. Again with Pakistan—and also in this study’s simulations across four countries—the gap between target coverage and the ‘coverage’ error grows smaller as the size of the target population increases. This is because the PMT scores perform worse at the bottom tail of the distribution, as shown in Figure A using data from Sri Lanka. At around 35% the two distributions cross, at which point using the cut-off values from actual consumption leads to errors related to the opposite effect of a larger program size than targeted.

⁶⁹ This practice overstates exclusion errors for a given target population and undermines the arguments for using the PMT. Exclusion errors would be lower if the cut off were calculated based on the PMT score.

Figure A: Cumulative distribution of actual consumption and proxy means test score, Sri Lanka



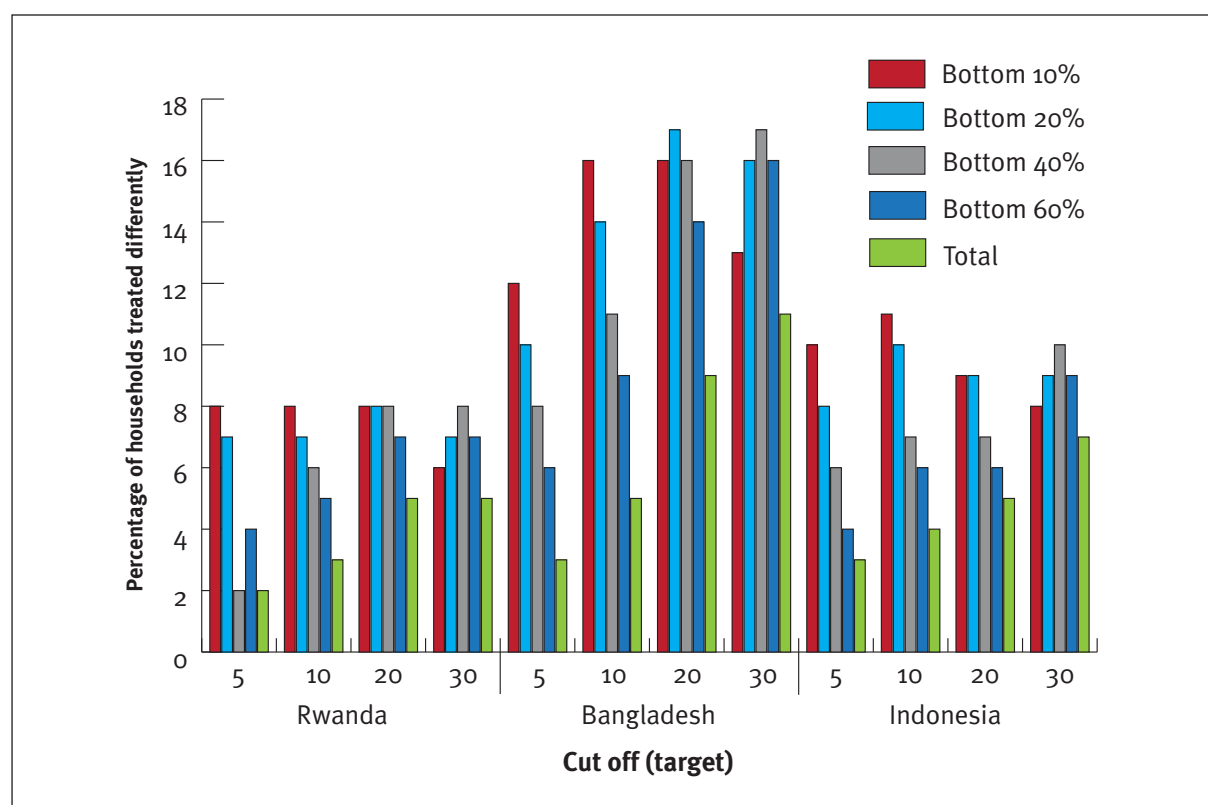
The ‘coverage’ error issue brings to the fore some of the problems inherent in identifying the poorest of the poor using the PMT. The PMT score (predicted value) will necessarily have a lower variation than actual consumption, since it is based only on information available in the model.

Annex 4: Impacts of equivalence scales on the selection of beneficiaries

The impact of two equivalence scales was tested on the selection of household beneficiaries—one using no scaling (per capita) and one using a scaling factor of 0.5 for children. Figure B shows the key results of the scenario modelling for a range of cut offs and coverage levels. In general, the impacts appeared greatest for those living in households near the cut-off value, meaning it is important to disaggregate the impacts for different groups of the population as done below. For example, at a cut-off value of 5% households in the bottom 10% receive the most impact, while at a cut off of 30% households in the 20th to 60th percentiles receive the most impact.

Total impacts are fairly small, ranging from around 2% to 6% in Rwanda and from around 3% to 11% in Indonesia. However, the impacts on the poor—the target group—are much higher: around 8% in Rwanda for cut offs of 5% to 30% and around 8% to 10% in Indonesia for cut offs of the same levels. In Bangladesh the impacts were much higher, ranging from above 10% of the target group at a cut off of 5% to 14% to 16% for cut offs between 10% and 30%.

Figure B: Percentage of households in targeted populations treated differently depending on the use of equivalence scales



References

- Aboderin, I 2006, *Intergenerational Support and Old Age in Africa*, New Brunswick, NJ, and London: Transaction Publishers.
- Adato, M 2000, *El Impacto de Progreso Sobre las Relaciones Sociales en la Comunidad*, unpublished paper by International Food Policy Research Institute (IFPRI).
- Adato, M, Coady, D & Ruel, M 2000, *Evaluación de Operaciones de Progreso desde la Perspectiva de los Beneficiarios, las Promotoras, Directores de Escuela y Personal de Salud*, unpublished paper by IFPRI.
- Adato, M & Roopnaraine, T 2004, *A Social Analysis of the Red de Protección Social (RPS) in Nicaragua*, final report, IFPRI.
- Ahmed, AU & Bouis, HE 2002, *Weighing What's Practical: Proxy Means Tests for Targeting Food Subsidies in Egypt*. Food Consumption and Nutrition Division, IFPRI, Washington DC, United States.
- Alatas, V, Banerjee, A, Hanna, R, Olken, BA & Tobias, J 2010, *How to Target the Poor: Evidence from a Field Experiment in Indonesia*. NBER working paper no. 15980; May.
- Anand, S & Morduch, J 1997, 'Poverty and the Population Problem,' in Massimo Livi-Bacci & Gustavo de Santis (eds), *Population and Poverty in Developing Countries*, Oxford University (Clarendon) Press, Oxford, England.
- Araujo, MC & Carraro, L undated, *A Proxy-Means Test Exercise for the Selection of Beneficiaries of Poverty Targeted Programs in Mongolia*, unpublished document.
- Baulch, B & Hoddinott J 2000, *Economic Mobility and Poverty Dynamics in Developing Countries*, Frank Cass Publishers, London, England.
- Beaman, L & A. Dillon, A 2010, *Do Household Definitions Matter in Survey Design? Results from a Randomized Survey Experiment in Mali*. International Food Policy Research Institute, discussion paper 01009. IFPRI, Poverty, Health and Nutrition Division.
- Behrendt, C, Cichon, M, Hagemejer, K, Kidd, S, Krech, R & Townsend, P 2009, 'Rethinking the Role of Social Security in Development', P. Townsend (ed), *Building Decent Societies: Rethinking the Role of Social Security in Development*. International Labour Office, Geneva.
- Castañeda, T & Lindert, K 2005, *Designing and Implementing Household Targeting Systems: Lessons from Latin American and The United States*, Social Protection Unit, Human Development Network, The World Bank.
- Chamberlin, C, Abbott, D, Bruni, L, Rigolini, J, Rokotunida I & Waqavonovono E 2009, *Social Assistance in Fiji: Context and Framework for a Technical Assistance Program*, unpublished manuscript of Ministry of Women, Social Welfare and Poverty Monitoring, and World Bank.

Coady, D, Grosh, M & Hoddinott, J 2002, *Targeting of Transfers in Developing Countries: Review of Experience and Lessons*, Social Safety Net primer series, Social Protection Anchor.

Coady, D, Grosh, M & Hoddinott, J 2004, *Targeting of Transfers in Developing Countries*, The International Bank for Reconstruction and Development/The World Bank, Washington DC, United States.

Coady, D & Parker, S 2005, *Program Participation under Means-Testing and Self-Selection Targeting Methods*, FCND discussion paper 191, Food Consumption and Nutrition Division, International Food Policy Research Institute, Washington DC, United States.

Deaton, A 1997, *The Analysis of Household Surveys: A Micro-econometric Approach to Development Policy*, World Bank, Washington DC, United States.

Deaton, A & Zaidi, S 2002, *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*, Living Standards Measurement Study working paper 135, World Bank, Washington DC, United States.

Devereux, S 2009, *Social Protection and the Global crisis. Wahenga Regional Hunger and Vulnerability Programme*, viewed June 2010, http://www.wahenga.net/sites/default/files/Social_protection_and_the_global_crisis.pdf.

Ellis, F 2008, *'We Are All Poor Here': Economic Difference, Social Divisiveness, and Targeting Cash Transfers in Sub-Saharan Africa*, paper prepared for the conference entitled Social Protection for the Poorest in Africa: *Learning from Experience*, held 8 to 10 September 2008, Uganda.

Ellis, F 2010, *Social Protection in Africa: Where Are we, and How Did We Get Here?* Frontiers of Social Protection brief, no: 7, September 2010, viewed September 2010, http://www.wahenga.net/sites/default/files/briefs/7.%20FoSP%20Brief%207%20-%20SP%20in%20Africa%20-%20where%20are%20we,%20and%20how%20did%20we%20get%20here_web.pdf.

Fiszbein, A & Schady, N 2009, *Conditional Cash Transfers: Reducing Present and Future Poverty*. World Bank, Washington DC, United States.

Gallardo, E 2008, *Política Social y Vida Comunitaria. Efectos Del Programa Oportunidades en dos Comunidades Mazatecas de Oaxaca*, Masters Thesis at El Colegio de La Frontera Norte, Tijuana, México.

GHK 2009, *Pakistan—Process Evaluation of Benazir Income Support Programme (BISP) Scorecard-Based Poverty Targeting under the Test Phase*, GHK Consulting Limited, London, England.

Gonzalez de la Rocha, M 2003, *México: Oportunidades y Capital Social. Seminary-Capital Social y Programas de Superación de la Pobreza: Lineamientos para la Acción*, Comisión Económica para América Latina y el Caribe, Naciones Unidas.

Grosh, M & Baker, J 1995, *Proxy Means Tests for Targeting Social Programs: Simulations and Speculation*, LSMS working paper no.118, World Bank, Washington DC, United States.

Grosh, M, del Ninno, C, Tesliuc, E & Ouerghi, A 2008, *For Protection & Promotion: The Design and Implementation of Effective Safety Nets*, Washington DC, The International Bank for Reconstruction and Development/The World Bank.

Grosh, M & Leite, P 2009, 'Defining Eligibility for Social Pensions: A View from a Social Assistance Perspective', *Closing the Coverage Gap: Role of Social Pensions and Other Retirement Income Transfers*, World Bank Publications, Washington DC, United States.

Hanlon, J, Barrientos, A & Hulme, D 2010, *Just Give Money To The Poor: The Development Revolution from the Global South*, Kumarian Press, Sterling, United States.

Hannigan, L 2010, *Social Assistance in Indonesia—Understanding the Gaps*, draft version of paper presented at the Annual Indonesia Update Conference, 24 to 25 September 2010, Australian National University, Canberra, Australia.

Haughton, J, & Khandker, SR 2009, *Handbook on Poverty and Inequality*, The International Bank for Reconstruction and Development/The World Bank, Washington DC, United States.

Hodges, A, Dashdorj, K, Yun Jong, K, Dufay, A-C, Budragchaa, U & Mungun, T 2007, *Child Benefits and Poverty Reduction: Evidence from Mongolia's Child Money Programme*, working paper, Division of Policy and Planning, United Nations Children's Fund (UNICEF) May 2007.

Hill, P 1986, *Development Economics on Trial: The Anthropological Case for the Prosecution*. Cambridge University Press, Cambridge, England.

Huber, L, Zarate, P, Durand, A, Madalengoitia, O & Morel, J 2008, *Programme Juntos: Certezas y Malentendidos en Torno a las Transferencias Condicionadas—Estudio de Caso de Seis Distritos Rurales del Peru*, United Nations Populations Fund, Instituto de Estudios Peruanos and UNICEF, Lima, Peru.

IRIS Center 2005, *Note on Assessment and Improvement of Tool Accuracy*, Developing Poverty Assessment Tools Project.

IRIS Center 2008, *Accuracy Results for 20 Poverty Assessment Tool Countries*, Developing Poverty Assessment Tools Project.

Jalan, J & Murgai, R 2006, *An Effective 'Targeting Shortcut?': An Assessment of the 2002 Below-Poverty Line Census Method*, Draft, India Safety Nets Review.

Johannsen, J 2006, *Operational Poverty Targeting in Peru—Proxy Means Testing with Non-Income Indicators*, working paper, International Poverty Centre, Brasilia, Brazil.

Johannsen, J, Tejerina, L & Glassman, A 2010, *Conditional Cash Transfers in Latin America: Problems and Opportunities*, unpublished manuscript of the Inter-American Development Bank.

Kidd, SW 1999, *Love and Hate among the People without Things: the Social and Economic Relations of the Enxet people of Paraguay*, doctoral dissertation, University of St. Andrews, St. Andrews, Scotland.

Kidd, SW & Whitehouse, E 2009, 'Pensions and Old Age Poverty,' in *Closing the Coverage Gap: the Role of Social Pensions and Other Retirement Income Transfers*, World Bank, Washington DC, United States.

Kidd, SW, Calder, R & Wylde, E 2011, *Assessing Targeting Options for Nepal's Social Grants: What Does the Evidence Tell Us?* Briefing paper published by United Kingdom's Department for International Development, Nepal and Development Pathways.

Kidd, SW & Calder, R 2011, *Conditional Cash Transfer Programmes: Their Relevance for Nepal*. Briefing paper published by United Kingdom's Department for International Development, Nepal and Development Pathways.

Kidd, SW, Wylde, E & Tiba, Z 2010, *Building a Social Protection Floor in Pakistan*. Unpublished report for UNICEF Pakistan.

Lanjouw, P, Milanovic, B & Paternostro, S 1998, *Poverty and Economic Transition: How do Changes in Economies of Scale Affect Poverty Rates for Different Households?*, policy research working paper 2009, World Bank, Washington DC, United States.

Maluccio, J 2008, *Household Targeting in Practice: The Nicaraguan Red de Protección Social*, Middlebury College Economics discussion paper no. 08-02, Department of Economics, Middlebury College, Vermont, United States.

Mathematica 2007, *Evaluation of Jamaica's PATH program: Final Report*, Mathematica Policy Research Inc., Washington, DC, United States.

Miguel, E 2005, 'Poverty and Witch Killing', *Review of Economic Studies*, vol. 72 (4):1153-72.

Ministry of Social Affairs/Palestine undated, *Social Safety Net Reform Project, the First Conditional Cash Transfer project in the MENA Region of the World Bank*, PowerPoint presentation.

Mkandawire, T 2005, *Targeting and Universalism in Poverty Reduction*, social policy and development program paper, no. 23, December 2005, United Nations Research Institute for Social Development.

Moene, KO & Wallerstein, M 2001, 'Targeting and the Political Support for Welfare Spending', *Economics of Governance*, 2 (1): 3-24.

Narayan, A, & Yoshida, N 2005, *Proxy Means Test for Targeting Welfare Benefits in Sri Lanka*, working paper 33258, South Asia Region, World Bank, Washington DC, United States.

Orozco, M, & Hubert, C 2005, *La Focalización de Desarrollo Humano Oportunidades de Mexico*, Serie de Documentos de Discusión sobre la Protección Social No. 0531, Unidad de la Protección Social, Red de Desarrollo Humano, El Banco Mundial, Washington DC, United States.

Pritchett, L 2005, *A Lecture on the Political Economy of Targeted Safety Nets*, social protection discussion paper series, no. 0501, Social Protection Advisory Service, The World Bank, Washington DC, United States.

Rascon, E, & Rubalcava L 2008, *Diagnóstico General de las Barreras a la Acumulación del Capital Humano en Zonas Urbanas en México (Perfil de la Pobreza Urbana)*, Consultant Report. Inter-American Development Bank, Washington DC, United States.

Rawlings, L & Rubio, G 2003, *Evaluation of the Impact of Conditional Cash Transfer Programs—Lessons from Latin America*, policy research working paper 3119, World Bank, Washington, DC, United States.

Rawls, J 1971, *A Theory of Justice*, Cambridge, Massachusetts, Harvard University Press.

Sen, A 1995, 'The Political Economy of Targeting,' in van de Wall, D, and Nead, K (eds.) *Public Spending and the Poor: Theory and Evidence*. World Bank, The John Hopkins University Press, Baltimore, United States.

Sepulveda, M 2009, *Promotion and Protection of all Human Rights, Civil, Political, Economic, Social and Cultural Rights, Including the Right to Development*, report to the United Nations General Assembly, prepared by the Independent Expert on Human Rights and Extreme Poverty, Human Rights Council, Eleventh Session, agenda item 3, 27 March 2009.

Sharif, IA 2009, *Building a Targeting System for Bangladesh Based on Proxy Means Testing*, social policy discussion paper no. 0914, World Bank, Washington DC, United States.

Skoufias, E, Davis B & de la Vega, S 1999, *Targeting the Poor in Mexico: An Evaluation of the Selection of Households into Progresa*, International Food Policy Research Institute, Washington DC, United States.

Slater, R & Farrington, J 2009, *Cash Transfers: Targeting*. Overseas Development Institute, project briefing, no: 27, November. Overseas Development Institute, London.

Townsend, P 2007, *The Right to Social Security and National Development: Lessons from OECD Experience for Low-Income Countries*. Issues in social protection, discussion paper 18. Social Security Department, International Labour Office.

Udry, C 1996, 'Gender, Agricultural Production, and the Theory of the Household', *Journal of Political Economy*, vol. 104, no. 5, pp. 1010–1046.

Van der Geerst, S 2002, 'Respect and Reciprocity: Care of Elderly People in Rural Ghana', *Journal of Cross-Cultural Gerontology*, 17(1): 3–31.

Veras, F, Peres R & Guerreiro, R 2007, *Evaluating the Impact of Brazil's Bolsa Família: Cash Transfer Programmes in Comparative Perspective*, IPC evaluation note no. 1, International Poverty Centre, Brasilia, Brazil.

Viquez, R 2005, *Sistema de Identificación de la Población Objetivo: SIPO en Costa Rica*. Serie de Documentos de Discusión sobre la Protección Social, no. 0531, Unidad de la Protección Social, Red de Desarrollo Humano, El Banco Mundial, Washington DC, United States.

Ward, P, Hurrell, A, Visram, A, Riemenschneider, N, Pellerano, L, O'Brien, C, MacAuslan, I & Willis, J 2010, *Cash Transfer Programme For Orphans And Vulnerable Children (CT-OVC), Kenya Operational And Impact Evaluation, 2007–2009*, Final Report, report prepared by Oxford Policy Management.

World Bank 1990, *World Development Report: Poverty*, A World Bank Publication, Oxford University Press.

--- 2001, *Project Appraisal Document on a Proposed Loan in the Amount of US\$40 Million to Jamaica for a Social Safety Net Project*, report no: 22359 JM, 9 August 2001. Latin America and the Caribbean Region, The World Bank, Washington DC, United States.

--- 2002, *Pakistan Poverty Assessment: Poverty in Pakistan: Vulnerabilities, Social Gaps, and Rural Dynamics*, report no. 24296-PAK, October 28, 2002. The World Bank, Washington DC, United States.

-
- 2005, *Egypt—Toward a More Effective Social Policy: Subsidies and Social Safety Net*, Social and Economic Development Group—Middle East and North Africa Region, The World Bank, Washington DC, United States.
- 2006, *Making the New Indonesia Work for the Poor*, The World Bank, Washington DC, United States.
- 2009a, *Social Transfers in Bosnia and Herzegovina: Moving Towards a More Sustainable and Better Targeted Safety Net*, The World Bank, Washington DC, United States.
- 2009b, *Implementation Completion and Results Report on a Loan in the amount of US\$40 Million to the Government of Jamaica for a Social Safety Net Project*, report no: ICR00001226, 17 September 2009. Latin America and the Caribbean Regional Office, The World Bank, Washington DC, United States.
- 2009c, *Project Appraisal Document on a Proposed Credit to the Islamic Republic of Pakistan for a Social Safety Net Technical Assistance Project*, report no: 47288-PK, 8 May 2009.