# CONSIDERATIONS IN USING PROXY MEANS TESTS IN EASTERN CARIBBEAN STATES

unicef

UN WOMEN

# CONSIDERATIONS IN USING PROXY MEANS TESTS IN EASTERN CARIBBEAN STATES

**BY DEBBIE BUDLENDER**

**for the UN Women Multi-Country Office for the Caribbean and the UNICEF Office for the Eastern Caribbean Area**

St.Lucia, 2014

Thanks are due to the following who provided information and guidance: Christine Arab, Jawad Aslam, Augustus Cadette, Anne Case, Angus Deaton, Isiuwa Iyahen, Edwin St Catherine, Violet Warnery and Ingrid Woolard.

UN Women is the UN organisation dedicated to gender equality and the empowerment of women. A global champion for women and girls, UN Women was established to accelerate progress on meeting their needs worldwide.

UN Women supports UN Member States as they set global standards for achieving gender equality, and works with governments and civil society to design laws, policies, programmes and services needed to implement these standards. It stands behind women's equal participation in all aspects of life, focusing on five priority areas: increasing women's leadership and participation; ending violence against women; engaging women in all aspects of peace and security processes; enhancing women's economic empowerment; and making gender equality central to national development planning and budgeting. UN Women also coordinates and promotes the UN system's work in advancing gender equality.

# TABLE OF CONTENTS


EXECUTIVE SUMMARY 5

I. CONSIDERATIONS IN USING PROXY MEANS TESTS IN EASTERN CARIBBEAN STATES 9

- Introduction 10
- What is a Proxy Means Test? 11
- Accuracy & Reliability 11
- Expenditure Versus Income 12
- Individual Versus Household 12
- The Proxy Variables 13
- The Reliability of the Underlying Survey 14
- Currency of the Survey Data 15
- Exclusion and Inclusion Errors 16
- Assessing the Errors 19
- Choosing the Eligibility Cut-Off 19

II. MEASURING POVERY 21

- Poverty and Indigence Lines 22
- Unpaid Care Work 22
- Equivalence Scales 23
- Economies of Scale 23
- Adult Equivalents 25
- Testing the Effect of Different Adjustment Scales 27

III. ALTERNATIVES TO PROXY MEANS TESTS 29

- Multidimensional Poverty 30
- Community-based Targeting 31
- Comparing Alternative Targeting Systems 32

IV. IMPLEMENTING A PROXY MEANS TEST 34

Application-based or Survey of Potential Beneficiaries 35

Defining the Beneficiary Unit 36

Other Eligibility Criteria 36

Combining a Proxy Means Test with Other Targeting Approaches 37

Transparency 38

V. CONCLUSION AND RECOMMENDATIONS 40

VI. REFERENCES 42

# EXECUTIVE SUMMARY

## Introduction

The social safety net assessments conducted in Eastern Caribbean states in and around 2009 recommended that countries should develop a proxy means test (PMT) to replace the diverse approaches they were using to target poverty-related benefits. Proxy mean tests were argued to be an objective and transparent alternative mechanism, especially given the unreliability of income data.

This paper aims to assist countries in exploring the recommendations by explaining what PMTs are, highlighting their strengths and weaknesses and presenting opportunities and challenges.

## What is a Proxy Means Test?

Both 'standard' and proxy means tests aim to fulfil the same function—determining beneficiary eligibility. In contrast to standard means tests, which are based on household or individual income, PMTs are based on potential beneficiaries' non-income characteristics, which are combined in a formula to derive a proxy for income.

The characteristics (also referred to as 'variables') to be used as proxies in the PMT, and the weights to be attached to each of them, are derived through statistical analysis of an existing household survey dataset. In the Eastern Caribbean states, the analysis has commonly been based on the Surveys of Living Conditions, which were conducted during the 2000s.

## Challenges in Determining the PMT Equation

In almost all cases, PMT analysis is done on expenditure (how the individual or household chooses to use their means) rather than income (what their means enables them to do). What is not much discussed in the literature is the extent to which the expenditure data may be incorrect given the tedious, complicated and sensitive nature of the relevant survey questions.

Similar to many other poverty-targeting mechanisms, most PMTs aim to predict household rather than individual expenditure. This focus is based on an implicit assumption that the benefit of the available income, or expenditure, is spread evenly (or according to need) across all individuals in the household. This creates a challenge when a social protection measure is intended to assist particular individuals (such as the elderly or children), but social or family norms give other individuals more control over household expenditure decisions.

The independent variables chosen to model or predict expenditure (the dependent variable) need to reflect characteristics that are easily observable by an outsider, difficult to lie about and do not change rapidly. Variable choice is constrained by the questions asked in the survey on which the model is based. Typically, the variables chosen include those related to household assets, size of the household and demographics.

Incorporating variables that are related to the head of household is questionable, as differing definitions, variations in culture or even who answers the door to the interviewer may determine who is named as head. Further, the characteristics of the head of household may not adequately describe the characteristics of other household members.

PMTs may not be appropriate for programmes that

deal with emergency or crisis situations, because the tests aim to use variables that do not change rapidly. Further, the needs associated with emergency or crisis may not relate only to (or even primarily to) poverty.

The reliability of the equation on which the PMT is based depends in large part on the reliability of the underlying survey. One of the challenges in small island states such as those of the Eastern Caribbean is the small population size and the linked small sample size. There are often further challenges with the quality of some of the survey variables, including ostensibly objective characteristics such as education.

PMT equations are usually derived on the basis of household expenditure patterns that were formed while the household might have been receiving pre-existing social protection benefits. This can be a problem if the PMT is to be used to determine benefit eligibility because it could result in incorrectly specifying variables and their weights in ways that disadvantage households with the profile of existing beneficiaries.

The time lapse between collection of the underlying data and use of the PMT raises concerns as to whether the chosen characteristics and the selected weights still serve as good predictors of expenditure. The question of currency of survey data will continue to arise, as the PMT equations will need to be adjusted periodically to reflect changing conditions.

All surveys are based on a sample of households and are thus subject to a margin of error. In addition, the PMT equations give rise to exclusion ('undercoverage') and inclusion ('leakage') errors due to differences between the expenditure predicted by the equation and a household's actual expenditure. Exclusion errors (as usually defined), occur when people who should be within the government's desired target group are nonetheless excluded from receiving benefits because the PMT incorrectly predicts an expenditure that is higher than the programme cut-off. In contrast, inclusion errors occur when people who should not be within the target group are nonetheless included in a beneficiary group because the PMT incorrectly predicts an expenditure that is lower than the cut-off.

One of the classic papers on PMTs (Grosh & Baker, 1994) tested data from Jamaica, Bolivia and Peru using a poverty line set at around the 30th percentile of expenditure (a higher level than the poverty rates in Eastern Caribbean states). The paper found that their models (equations) correctly identified fewer than half of the households that should have been eligible (Grosh & Baker, 1994: 15). Similarly, Kidd & Wylde (2011: ii) found what they describe as "high in-built errors," which are especially severe when the target population for a benefit represents 20 percent of the population or less.

The choice of the cut-off point for targeting is a policy question, not a technical decision. Sabates-Wheeler et al (2014: 2) suggest that although policy makers may be inclined to focus on inclusion errors (because of the unnecessary costs incurred through such errors), exclusion errors should merit greater weight in discussion and programme design because of the "humanitarian" cost incurred when people who need assistance are excluded.

## Measuring Poverty

The most common approach used in the Eastern Caribbean states is to set an extreme poverty or indigent line at the level of expenditure needed to buy the basic minimum of calories (spread across the appropriate food groups) scientifically determined to be necessary. A household will only escape poverty at this level of expenditure if it meets two unrealistic assumptions. First, that it spends every cent in the most economical and judicious way, based on full knowledge of nutrition and prices in the market. Second, that it does not have other needs besides food.

The poverty line therefore adds to this basic food amount a further allowance for non-food items. The allowance is based on the average non-food expenditure of the bottom two quintiles of households.

Means tests—whether proxy or direct and whether based on income or expenditure—generally count only monetary amounts. These tests fail to consider the value of unpaid care work—unpaid work that is the work done mainly by women and includes caring for other household members (especially children),

cooking and other housework.

'Equivalence scales' is the generic term used to describe adjustments made in calculating poverty rates to account for differences in household size and composition. In the Eastern Caribbean, instead of deriving 'per capita' household expenditure by dividing total expenditure by the number of household members, a 'per (male) adult equivalent' expenditure is derived by dividing total expenditure by the sum of the individuals, with some individuals counting as only a fraction of a full adult male. Unusual when compared with current international practice, sex differentiated equivalence scales have been used up to the present in the Eastern Caribbean. This Eastern Caribbean practice is not in line with international best practice.

## Implementing a Proxy Means Test

Some countries aim to administer the PMT to the full population or to all households in poor areas, while others test only those who apply for benefits. The disadvantage of administering the PMT to the full population is the effort and expense involved. The disadvantage of the application-based approach is that poor, eligible households may not apply because they do not know about the benefit or because the time, effort, costs and other challenges associated with applying are too great.

In general, the PMT literature does not discuss whether the beneficiary unit assessed matches the country's legal obligations in terms of support. Further, any means test that is based on the household or family unit may not be appropriate for assistance that facilitates women and children escaping from domestic violence.

When used as part of a unified targeting system for different programmes, the eligibility thresholds on the PMT for the various social programmes may differ. In addition, there will often need to be further criteria used in determining eligibility for the various programmes, as the PMT is a proxy only for income. Other eligibility criteria might, for example, include presence of an elderly person, a child or a person with disabilities.

Consideration of other criteria alongside the PMT retains the PMT as the targeting mechanism with respect to poverty. It therefore does not address exclusion errors due to PMT equation mis-predictions that incorrectly exclude substantial numbers of those who should be eligible.

In some cases, countries do not disclose the variables and weights used in the PMT so as to reduce opportunities for manipulation. However, such non-disclosure reduces transparency and may result in reduced levels of public acceptance.

If the system is rights-based, an appeals mechanism would be a necessary component and would need to be able to override the PMT equation in cases where applicants are deemed to have been unfairly excluded.

## Recommendations

The following recommendations are put forward for those Eastern Caribbean countries that decide to use a PMT for targeting of benefits:

- Combining the PMT with other targeting methods can be considered, though doing so is unlikely to overcome all the challenges associated with PMTs;
- Approaches that take into account the income forgone by mothers and others with heavy care responsibilities should be considered;
- An alternative equivalence scale that does not differentiate on the basis of sex of household members should be used. The weighting for children should take into consideration expert opinion that this weight should be higher for middle- and higher-income countries than for poorer countries, and that childhood deprivation has long-lasting impacts on individuals, families and the country as a whole;
- A full costing of the roll-out of the PMT should be undertaken prior to its use, in order to ascertain the full financial, human resource and logistical implications; and

- Administrative justice requires that an appeals mechanism be in place for the targeting mechanism. To complement this mechanism, the responsible government agency could assess the cases of all rejected applicants on a periodic basis so as to avoid exclusion of vulnerable individuals who may not have the confidence or capacity to initiate an appeal. The appeals mechanism should consider the applicant's objective situation in order to determine whether the exclusion was fair.

# 1 CONSIDERATIONS IN USING PROXY MEANS TESTS IN EASTERN CARIBBEAN STATES

# CONSIDERATIONS IN USING PROXY MEANS TESTS IN EASTERN CARIBBEAN STATES

## Introduction

In and around 2009, most Eastern Caribbean states underwent social safety net assessments (SSNAs). These assessments made several recommendations for improving social safety net systems. A common recommendation to countries in the region was that they should replace the diverse approaches they were using to target poverty-related benefits with a proxy means test (PMT).

For example, the Saint Kitts and Nevis SSNA (Blank, 2010: 50) noted that existing systems required substantial staff time and effort and could be politically influenced. The SSNA argued that a unified system for identifying beneficiaries would increase efficiency, effectiveness and transparency of targeting. Given the unreliability of income data, the PMT was recommended as a possible option. An alternative recommendation, training social workers on the systematic identification of the poor, would be based on a one-off PMT-like analysis. This would identify the characteristics linked to poverty; a full PMT would require that regular household surveys be conducted to allow for updating the test.

Similarly, the Saint Lucia SSNA referred to the cost of administration of existing targeting systems, as well as the time, money and other costs imposed on beneficiaries given that different tests were applied for the various benefits. In addition, the tests were "not doing a very good job" of identifying appropriate beneficiaries (Blank, 2009: 46). A PMT was seen as an "objective and transparent" alternative mechanism (the approach was already being tested with the Koudemain programme using a PMT developed by the Statistics Department). As with Saint Kitts and Nevis, the unreliability of income data played a substantial role in proposing a PMT. The SSNA recommended subjecting the PMT to further tests and then using it for targeting by other programmes (which might use different cut-offs).

As of May 2014, countries were at different stages in taking the PMT recommendation forward.

This paper aims to assist this process by explaining what PMTs are, highlighting their strengths and weaknesses and by presenting opportunities and challenges. It does so by drawing on classic international literature on the topic (particularly that produced under the auspices of The World Bank), and by drawing on current practices in the Eastern Caribbean states.

The rest of this paper is divided into the following seven sections:

1. An explanation of what proxy means tests are;
2. An explanation and discussion of various aspects associated with the key issues of accuracy and reliability;
3. A discussion of the basic aspects of measuring poverty that are relevant for determination of targeting thresholds (cut-offs) for PMTs;
4. A discussion of the use of equivalence scales to adjust for household size and composition in measuring household poverty;
5. A discussion of some alternatives to PMTs;
6. A discussion of issues related PMT implementation; and
7. A brief conclusion.

The Executive Summary of this paper is available as a separate policy brief.

## What is a Proxy Means Test?

A PMT is used to determine eligibility for government benefit programmes. The test focuses on eligibility in terms of poverty.

Traditionally, means tests have used the level of income as the basis to determine whether an individual or household is eligible for poverty-related benefits. This is logical to the extent that poverty is seen as related to a lack of income or 'means'. It seems especially logical when the benefit takes the form of a cash grant.

PMTs aim to fulfil the same function as standard means tests—to determine beneficiary eligibility. In contrast to standard means tests, which are based on asking household or individuals about their income, PMTs are based on characteristics of the individual or household that are meant to predict their likely income. As the Saint Lucia SSNA explains: "A PMT uses data from household surveys to construct a scoring formula and a cut-off point for eligibility. Households that receive a score below the cut-off point are eligible for benefits; households that receive a score above the cut-off point are not eligible" (Blank, 2009: 46). The characteristics in the formula thus constitute a proxy for income.

IOS Partners/London School of Economics (2011) note that if the cut-offs are set at levels equivalent to the poverty and indigence lines for a country, then the PMT has intuitive appeal to policymakers (who are generally familiar with their country's poverty and indigence lines).

The main reasons usually given for why the PMT is considered to be more reliable than a test based on asking about income directly include:

- Individuals may be unwilling to state their income or may lie about their true income, especially if they know that access to a benefit depends on their responses;
- People in informal employment have little or no formal proof on income available, unlike those who receive a payslip as formal sector employees; and
- Income may fluctuate substantially over time in developing countries for people who are informally employed or reliant on sources such as remittances or income from seasonal industries. In contrast, the characteristics used for the PMT may be more stable.

## Accuracy and Reliability

The characteristics (also referred to as 'variables') to be used as proxies in the PMT, and the weights to be attached to each of them, are derived through statistical analysis of an existing household survey dataset. In the Eastern Caribbean states, the analysis has commonly been based on the Surveys of Living Conditions (SLC), which were conducted during the 2000s.

The analysis is used to determine which characteristics are statistically linked with poverty. Most analyses use either regression or principal component analysis. Statistical analysis also determines how strongly each of the characteristics is linked with expenditure (or,

more commonly, the log of expenditure). This enables derivation of the weight to be attached to that characteristic in the PMT. The analysis thus uses the chosen variables to produce an equation that predicts the likely expenditure of the household.

Several aspects of the statistical testing affect its reliability.

## Expenditure versus income

The PMT aims to provide a proxy for income, with income representing the ability of the individual or household to satisfy its needs. However, in almost all cases the analysis is done on expenditure (how the individual or household chooses to use their means) rather than income (what their means enables them to do). Woolard & Leibbrandt (1999) observe that conceptually, a measure based on expenditure identifies those who **fail** to meet a specified standard of living (a measure of what actually happens), whereas a measure based on income identifies those who are **unable** to meet a specified standard of living (a measure of what could potentially happen). Note that neither approach directly assesses the capacity or potential of a household or individual to earn income.

The literature justifies using expenditure because of the difficulties ascertaining income in a survey or a straightforward (non-proxy) means test. Using expenditure is also justified on the basis that it is likely to fluctuate over time less than income (a phenomenon known as 'smoothing'). What is not much discussed in the literature is the extent to which expenditure data may be incorrect. This is a serious concern where, as is generally the case, income data are collected by requiring respondents to record and/or remember every purchase they have made over a short period (perhaps through using a weekly diary as was done in the Eastern Caribbean Surveys of Living Conditions) as well as remember annual purchases.

A further concern is the extent to which individuals may omit to report some types of expenditure that may be considered less socially acceptable (such as expenditures on alcohol or tobacco). To the extent that these goods are consumed more by men than women, non-disclosure of these expenditures will increase men's apparent poverty relative to that of women.

Finally, focusing on expenditure is likely to reduce the extent of inequality between wealthier and poorer individuals and households because the wealthy tend to spend a smaller proportion of their income. The undercount of the means of the wealthy may not be serious for means testing, which focuses on the poorer end of distribution where there is likely to be little difference between income and expenditure.

## Individual versus household

Similar to many other poverty-targeting mechanisms, most PMTs aim to predict household rather than individual expenditure. This makes sense to the extent that much expenditure occurs on a household level. Indeed, in surveys, households are generally defined as a group of individuals that live together and pool income.

Use of the household implicitly assumes that the benefit of the available income, or expenditure, is spread evenly (or according to need) across all individuals in the household. This creates a challenge when a social protection benefit is intended to assist particular individuals (such as the elderly or children), but and social and family norms give other individuals more control over expenditure decisions. Nevertheless, IOS Partners/London School of Economics (2011: 1) state that the Belize PMT will measure the poverty level of "every household and individual" despite the test being done at the household level.

As Deaton & Muellbauer (1986: 742) state firmly, "there are cases in which [the assumption that everyone in the household has the same welfare level] would be clearly inappropriate, for example, in societies in which women and children are treated as the chattels of a dominant male." Citing evidence of widespread discrimination against older people in Africa, Kidd & Wylde (2011) suggest that a PMT should not be used for targeting of individual benefits such as old age, disability, child or unemployment benefits. Internationally, evidence suggests that monies placed in a woman's control are more likely to be used for the benefit of children and the family as a whole than

when placed in a man's control.

Because the PMT is based on household expenditure, the rest of this paper favours 'household' over 'individual or household'. However, it must not be forgotten that many benefits aim to target particular individuals rather than the household as a whole. Use of a PMT for such benefits may be a problem if there are inequalities within households that result in the individual who might have been targeted being assumed to benefit from a larger share than they do of household income.

## The proxy variables

The independent variables chosen to model or predict expenditure (the dependent variable) need to reflect characteristics that are easily observable by an outsider, difficult to lie about and do not change rapidly (if not, there is little advantage in using a PMT rather than a straightforward question about income). The choice of variables is constrained by the questions asked in the survey on which the model is based. Typically, the variables chosen include those related to assets of the household, the size of the household and demographics.

A relatively large number of variables may be included in the initial modelling. Final choice of variable is determined through statistical refinement of the model, which points to the variables that are the strongest determinants of predicted expenditure. In Saint Lucia, the SL-NET PMT that is currently being tested uses the following characteristics: (inverse of) household size, mean age of household members, proportion of household members with completed secondary school, number of bedrooms per capita, district, materials used for outer walls of dwelling, ownership of the dwelling, ownership of a washing machine, cable television connection, Internet connection, ownership of a refrigerator, ownership of a vehicle and ownership of a computer. The fact that the mean age of household members emerges as a significant predictor of expenditure (with a lower mean associated with lower expenditure), despite the use of adult equivalence scales (see discussion below), suggests strongly that poverty is more associated with children than with adults.

The characteristics are identified through statistical modelling rather than through 'common-sense'. Narayan & Yoshida's (2005) exploration of a PMT for Sri Lanka provides an example of modelling that resulted in identifying a counter-intuitive variable: living in the Western Province, the wealthiest province of the country, is strongly correlated with an increased risk of poverty. The counter-intuitive result is explained by the inclusion of another province factor that cancels out this effect. Similarly, in Belize, Stan Creek was ranked as the second richest rural district by the PMT (Wietzke, 2012). Also in Belize, ownership of a stove was found to be linked with poverty in urban areas (IOS Partners/ London School of Economics, 2011). While there might be a valid technical statistical explanation of such counter-intuitive results, this type of result is likely to encourage distrust of a technical instrument such as the PMT.

Though many PMTs include variables that relate to the head of household, Saint Lucia's PMT does not (in part because the data suggested that some households had more than one head while others had no head).

**Use of variables related to the head of household is questionable.** First, the definition (if it is provided) of 'head of household' often will not result in an objective and consistent choice among the members. For example, the manual for Grenada states that the head is defined on the basis of age or because they are the main source of the household's "economic supply" (Diaz 2014: 38) This does not identify a single person in cases where the oldest person in the household is not also the highest earner. In other surveys, there may be an assumption that members of the household will "know" who the head is. In such cases, variations in culture or even who answers the door to the interviewer may determine who is named as head. Characteristics based on head of household do not pass the test of being easily observable because the head itself is not easily observable.

Second, the characteristics of head of household may not adequately describe the characteristics of other household members. The Dominica PMT includes the head of household's occupation as one of the variables Fernandez (2008a). This assumes either that other members of the household are not employed or

that their work is less important income-wise than that of the head. Though it is arguable that the PMT modelling identifies significant characteristics and this is therefore not a question of assumptions, the lack of a clear definition of head of household could result in mis-targeting during implementation.

In Grenada (and perhaps elsewhere), the manual for the test repeatedly emphasizes the importance of addressing questions to the head of household or, in their absence, their spouse (Diaz 2014). In households with adult men and women, the head is likely to be a man. This instruction is given for general household questions as well as questions relating to children. The latter is surprising as it is generally recognized that women are likely to have better knowledge of children than men. Kidd & Wylde (2011) cite experience in Pakistan where women were observed to give different answers from men, with women's answers generally more accurate than men's. These considerations relate to the validity of the data used to derive the PMT formula.

**PMTs aim to use variables that do not change rapidly, which means that they may not be appropriate for programmes that deal with emergency or crisis situations.** For example, variables relevant to measuring short-term crisis need and medium-term poverty are likely to differ. Further, the needs associated with emergency or crisis may not relate only to—or even primarily to—poverty. This observation is especially relevant in the Eastern Caribbean states given the high levels of vulnerability to disasters.

Kidd & Wylde (2009) argue further that using ownership of assets in PMTs may discriminate against older applicants because households tend to accumulate assets over time; this does not necessarily indicate a decrease in their poverty level as measured by regular income.

## The reliability of the underlying survey

The reliability of the equation on which the PMT is based depends in large part on the reliability of the underlying survey. **One of the challenges in small island states such as those of the Eastern Caribbean is the small size of the population and the linked small size of the survey** (2,025 households in Belize; 2,688 individuals in Grenada; and a little over 1,200 households and 4,300 individuals in Saint Lucia). This limits the extent to which the data can be reliably disaggregated and the number of variables that can be used for the PMT. Nevertheless, Dominica has 27 variables (24 if one excludes dummies based on the same underlying variable); Saint Lucia has a more modest 13.

Cadette (2012: 10) notes further challenges with the quality of some of Saint Lucia's survey variables, including "considerable" missing values for education—ostensibly a question that should be easy to answer and not highly sensitive. This was problematic because one of the PMT variables is the proportion of household members who have completed secondary school. There are also substantial missing values for the earned income questions. While the latter could be seen as an underlying justification for using a PMT, the fact that so many income-related answers are missing (rather than possibly incorrect) raises concern as to the extent to which expenditure was fully recorded. This is especially so given that the number of questions and level of detail required for expenditure in the SLC was far greater than for income. In Saint Lucia, the main SLC interview alone lasted an average of two hours, in addition to which all adults and all employed children were required to complete expenditure diaries over a two-week period (Kairi Consultants, 2007b: 9).

In addition to missing data for survey participants, the PMT equation may result in further bias due to the absence of information for households in the targeted sample for whom questionnaires are not completed at all. In the Saint Lucia SLC of 2005/06, the response was 94 percent. The responses for those who did respond can be 'weighted up' to the full population, but errors will arise if those who respond share particular characteristics.

One aspect that does not seem to be considered in many PMT exercises, including those in the Eastern Caribbean, is the extent to which the current level of expenditure (as recorded in the survey used for modelling the PMT) reflects benefits that the household is already receiving from government. **If the PMT is to be used to determine eligibility for benefits—and in particular cash benefits—then it should reflect the expenditure that households would have had in the absence of existing benefits.** In practice, however, it seems that the PMT equations are derived on the basis of expenditure with existing benefits already in place.

This is an important consideration in the Eastern Caribbean islands where, for the most part, safety net reforms are attempting to streamline and regularize a set of existing benefits. If, as it seems is the case, no correction is made for these benefits (cash grants in particular), then the affected households will appear better-off than they actually are. If, as is likely, households already receiving benefits have particular characteristics, this will in turn affect the equation that underlies the PMT.

The Grenada exercise explicitly includes the household's access to pensions and public assistance as one of the indicators of well-being. This could result in these households being more likely to be excluded for whatever benefits are given based on the Grenada test derived from the analysis. In Saint Lucia, the SSNA recorded that 2,492 beneficiaries were receiving assistance from the Public Assistance Programme in April 2009 (Blank, 2009: 22). While this is a relatively small proportion of all households, the fact that beneficiaries will be concentrated among the poor could bias the PMT equation.

## Currency of the survey data

The analysis used in deriving the PMTs in the Eastern Caribbean has generally used surveys that were done some years previously. For example, the Saint Lucia PMT is based on data from the 2005/06 SLC, the Dominica PMT is based on data from 2002, and the Grenada test (which is not, in fact, a PMT, because it is not based on variables and weights derived from a regression on expenditure) is based on data collected in 2007/08. This raises concern as to whether the chosen characteristics, with the chosen weights, still serve as good predictors for expenditure up to a decade later. It also raises concerns going forward because for a PMT to remain effective in targeting, the variables and weights need to be regularly assessed and adjusted as necessary, which requires the regular production of new survey data.

The question is particularly relevant given that the SLCs were done prior to the major economic and global crisis that started in 2008. The question is also pertinent given that the variables include access to the Internet and cable TV and computer ownership — services for which costs and levels of access are likely to have changed rapidly in recent years. In Bangladesh, Sharif (2009) reports avoiding using variables such as the use of mobile phones because it was likely to have changed over the three years since the survey underlying the country's PMT was done.

Even a measure such as the percentage of household members who have completed secondary education is questionable when Saint Lucia's Labour Force Survey of 2012 suggested that only 40 percent of the population aged 15 years and above has attained this level of education, compared to 66 percent reported in the 2005/06 survey.

Wietzke's (2011) testing of the Belize PMT (the Beneficiary Targeting Index), found serious mismatches in data even over a relatively short period. The PMT was derived on the basis of the 2009 SLC. In 2011, the test was validated against the 2010 population census. Despite the relatively short lapse in time, Wietzke reports that the composition of the household changed between the time of the census and the validation exercise. The team was therefore able only to check the validity of person-specific variables relating to the head of household. The extent of the mismatch was 42 percent in respect of whether the household had a concrete roof, 41 percent in respect of whether there was a toilet, 64 percent in respect of whether the head of household had primary education or less and 66 percent in respect of whether the head had secondary or higher education, and 54 percent in respect of whether the head was employed. IOS Partners/London

School of Economics (2011) suggest that inaccuracy in reporting on employment of the head of household in rural areas may reflect high levels of self-employment.

Wietzke observes that while these results do not necessarily point to invalidity of census data, they raise concerns as to the reliability of the variables that were proposed for the PMT. He notes further that the index constructed was less reliable among the poorest households. Thus for an eligibility cut-off of 10 percent of households, 37 percent of households that would have been eligible on the basis of census data were not eligible on the basis of validation data. At a cut-off of 10 percent, the proportion in this situation was just over 25 percent of those that census data indicated should be eligible.

By dropping the variables on ownership of a computer and information on the education and employment status of the head of household, targeting accuracy was reduced for the poverty line and improved for the indigence line. However, even after improvement, the exclusion and inclusion rates (see below) were more than 50 percent. The accuracy of targeting also differed noticeably across rural and urban areas, but the sample was too small to allow for separate targeting equations for both areas.

**The question of currency of survey data will arise repeatedly if PMTs are implemented, as the PMT equations will need to be adjusted periodically to reflect changing conditions.** The Saint Kitts & Nevis SSNA (Blank, 2010) notes explicitly that ten-yearly surveys, as currently seems to be the practice, are not adequate if a PMT is to be used. In addition, if regular surveys are conducted and used for updating the PMT equation, the methodology and questionnaire would probably need to be more standardized across the years than currently seems to be the case.

## Exclusion and inclusion errors

All surveys are based on a sample of households, and are thus subject to a margin of error. Unfortunately, this particular error margin is rarely reported in the PMT literature. The PMT literature does, however, include extensive discussion of often substantial exclusion ('undercoverage') and inclusion ('leakage') errors.

Exclusion errors (as usually defined) occur when people who should be within the government's desired group are nonetheless excluded from receiving benefits because the PMT incorrectly predicts an expenditure for them that is higher than the cut-off. In contrast, inclusion errors occur when people who should **not** be within the target group are nonetheless included in a beneficiary group because the PMT incorrectly predicts an expenditure that is lower than the cut-off. Diaz et al (2014) use a more complicated definition for their work in Grenada. They define the inclusion error as relating to the extent to which transfers are larger than the poverty gap of particular families. Their use of this definition, based on a cash transfer, is interesting given that their test looks at other dimensions beyond income.

Exclusion and inclusion errors arise because the statistical analysis produces an equation that predicts expenditure, but the choice of variables can never include all factors that influence expenditure. There will therefore generally be a difference (error) between the expenditure predicted by the equation and the actual expenditure for a particular household. The statistical term r-squared describes the proportion of the variation in expenditure across households as explained by the chosen variables. In most cases in the literature, this is somewhere between 30 and 60 percent, leaving between 70 and 40 percent of the variation in expenditure unexplained by these proxy characteristics. Fernandez (2008a: 5) states that any level of r-squared above 0.50 is "acceptable" for this type of model (although he does not justify this value). IOS Partners/London School of Economics (2011) similarly states that "as a rule of thumb", one should aim for a level of 0.50 or higher, with the t-test applied at the 90 percent confidence level. As seen below, even r-squared values above 0.50 can result in substantial exclusion and inclusion errors.

In Saint Lucia, the PMT produced an r-square of 0.46, below the level deemed acceptable by Fernandez, and the following exclusion and inclusion rates:

- For the poverty line, set at around the 21st percentile to match the 21 percent poverty rate, an exclusion rate of 16 percent and an inclusion rate of 31 percent.

- For the indigence line, set at around the 1st percentile, an exclusion rate of 0 percent and an inclusion of 6 percent. While the error rates are small, only 15 households out of 1,200 were identified as indigent.

For Belize, the re-squared is 0.57, above the Fernandez level. However, the exclusion and inclusion rates, worryingly high and worse for urban than for rural areas, are as follows:

- For the poverty line, the exclusion rate is 24 percent and the inclusion rate 25 percent in rural areas, and 38 percent and 36 percent respectively in urban areas.
- For the indigence line, the exclusion rate is 39 percent and the inclusion rate 36 percent, even in rural areas (IOS Partners/London School of Economics, 2011).

For Dominica, the r-squared is reported as 0.57, above the Fernandez level. When the Dominica equation was tested on data from a different year, the exclusion rate was 39 percent and inclusion rate was 25 percent using a poverty line of EC$ 3,400. In Dominica, the PMT was also tested by applying it 'live' to 100 households in a poor urban area and 100 households in a poor rural area. Just over half (53) of the urban and 86 of the rural households were classified as poor by the PMT. It is not clear from the report whether this should be interpreted as equivalent to exclusion rates of 47 percent and 14 percent respectively.

In one of the classic papers on PMTs, Grosh & Baker (1994) argue that, rather than the r-squared, it is the exclusion and inclusion error rates that are important. They also consider the impact on poverty in terms of reducing poverty rates and poverty gaps, but that is beyond the scope of this paper as it brings in further questions that are not questioned here (such as the size and nature of benefits).

Grosh & Baker perform a series of tests on data from Bolivia, Jamaica and Peru using a poverty line set at around the 30th percentile of expenditure, which is at a higher level than poverty rates in Eastern Caribbean states. They conclude that **exclusion rates were "disappointingly high" (Grosh & Baker, 1994: 15), with the models (equations) correctly identifying fewer than half of the households that should have been eligible.** Inclusion errors increased as the poverty line was reduced, and vice versa. For Sri Lanka, Narayan and Yoshida (2005) find that increasing the poverty line from the 30th to the 40th percentile reduced the exclusion error from 43 to 22 percent, but increased the inclusion error from 35 to 44 percent. Wietzke observes that high levels of inaccuracy at lower cut-offs are of particular concern in Belize (and perhaps other Eastern Caribbean countries), in that governments are likely to make relatively small budgets available for social assistance programmes.

Table 1 shows further values for r-squared and exclusion and inclusion rates for poverty lines set at around 30 percent of households as quoted by Sharif (2009) from other studies. The rates are over-optimistic to the extent that programmes often target less than 30 percent of the population, which tends to increase error rates. Current programmes in the Eastern Caribbean states typically target an extremely small proportion of the population. For example, the SSNA for Saint Lucia recorded only 2,492 beneficiaries of the Public Assistance Programme in 2009, while the SSNA for Saint Kitts recorded only 641 beneficiaries for Assistance Pensions (Blank, 2010: 19).

Johannsen's (2006: 12) proposed measure for Peru includes some expenditure items (such as on clothing) alongside other non-expenditure variables; it is thus not a true PMT. It is nevertheless interesting for our purposes, as Johannsen finds that exclusion of the "powerful" monetary variables reduces the accuracy of the tool.

Kidd & Wylde (2011) assess the accuracy of PMT regression equations using databases from Bangladesh, Indonesia, Rwanda and Sri Lanka. For this investigation, they define the eligible population as the actual coverage of the programme in each country rather than on the basis of a poverty line. This approach results in inclusion and exclusion errors being the same in each case. They find what they describe as "high in-built errors," which are especially severe when targets represent 20 percent of the population or less (as is the case for many Eastern Caribbean programmes). They point out that in addition to inclusion and exclusion errors, survey sampling errors further diminished accuracy. They find that 7 to 11 percent of the target

## Table 1: Examples of Accuracy and Error Rates when Targeting 30 Percent of Households

| Country | R-squared | Exclusion (30%) | Inclusion |
|---|---|---|---|
| Sri Lanka | 0.56 | 43 | 35 |
| Pakistan | 0.53 | 43 | 35 |
| Armenia | 0.20 | | |
| Latin America | 0.30–0.40 | | |
| Egypt | 0.43 | 43 | 30 |
| Jamaica | | 69 | 44 |
| Bolivia (urban) | | 39 | 24 |
| Peru (urban) | | 54 | 35 |

**Sources: Grosh & Baker, 1994; Sharif, 2009**

population would be classified differently (as poor or non-poor), depending on whether the upper or lower limit of the sampling error on the regression coefficients was used.

Kidd & Wylde (2011) report r-squared values of 0.56 in Bangladesh, 0.37 in Indonesia, 0.44 in Rwanda, and 0.57 in Sri Lanka. Across these countries, exclusion and inclusion errors vary between 44 and 55 percent when 20 percent of the population is covered and between 57 and 71 percent when 10 percent is covered. These estimates thus confirm that higher error rates are likely to arise when the target population is relatively small. Sharif (2009: 3) also cites evidence from Pakistan of the difficulty of predicting expenditure accurately at the "left tail" of the expenditure distribution i.e. among the poorest households.

Ellis (2008: 2) takes issue with initiatives in Africa that aim to target the bottom "'non-viable' destitute" 10 percent of the population with cash grants. His observations are relevant to the extent that some of the Eastern Caribbean PMT initiatives aim to target the extreme poor or indigent with safety net measures rather than aiming to provide broader social protection. Ellis finds that in poor countries such as Ethiopia, Malawi and Zambia, the difference between the monthly expenditure of each of the deciles up to the sixth is only around two US dollars. Expressed differently, there are very small differences in well-being, life style and access to assets and income across the majority of the population.

In this situation, Ellis suggests that citizens understand, accept and perceive categorical targeting (where beneficiaries are selected on the basis of age or some other observable characteristic), as more fair than methods that differentiate based on the absence of able-bodied adults or other criteria. He notes that such categorical targeting also establishes the grant as a right. In contrast, where there is a PMT, applicants who have been rejected have very limited grounds to appeal against the rejection.

On the basis of their investigation, Kidd & Wylde (2011: III) observe that a PMT "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." They cite qualitative research in Mexico, Nicaragua and Peru in which community members said that whether or not one was deemed eligible for benefits that utilized a PMT was the result of luck or God's will, similar to the outcome of a lottery.

These worrying estimates probably understate the extent of the problem. In particular, where the estimates are derived—as is usual—from testing the predictive equation on the survey data, overfitting will bias the results in an optimistic direction because the testing uses the same data used for the prediction (Sharif, 2009). Use of one half of a sample to derive the equation and the other half of a sample to test it may avoid this, but sample sizes for Eastern Caribbean

countries are probably too small to allow this.

Splitting the sample into two will likely produce an optimistic picture in that the questionnaire, methodology and other aspects for the two half-samples will be the same. Kidd & Wylde (2011) perform a more robust test for Indonesia using a different dataset from that used for the original modelling. They find that both the r-squared and overall targeting errors are similar across the two surveys, but that 11 to 13 percent of the target population were treated differently. Unfortunately, tests using alternative data are not possible in the Eastern Caribbean states because of the absence of other data sets and, in particular, the absence of other data sets that include income or expenditure data.

Grosh & Baker (1994: 22) observe that results can be "considerably" improved if the equation is based on the poorest half of the population rather than the population as a whole. This too is probably not possible for small island states given the small overall sample size. In addition, restricting the equation to the poorest half could increase the bias that results from including expenditure from households that already receive benefits.

## Assessing the errors

Some argue that the size of the error rates overstates the problem because many of those who are incorrectly excluded are only just below the poverty line and many of those who are incorrectly included are only just above the poverty line. Including or excluding these households is thus not a serious error.

In engaging this argument, it is important to keep in mind that the poverty line is itself an artificial analytical construct. If the line is set at EC$ 2000, it is true that a household that had EC$ 1999 that then receives an extra EC$ 2 will not suddenly be in a noticeably different position. To that extent, the argument that exclusions and inclusions are less serious if they affect people near the cut-off is correct.

However, PMTs often incorrectly predict expenditure for a substantial number of households that are far from the cut-off. For Bangladesh, for example, Sharif (2009) reports that when the 20th percentile is used as the cut-off, 41 percent of those excluded were in the lowest 10 percent of actual expenditure. Similarly, one-third of households incorrectly predicted to be eligible were in the wealthiest 60 percent of the population. Though this level of accuracy may be better than existing programme targeting in Bangladesh, it is not sufficient to say that PMT targeting is adequate. In addition, it does not follow to say that PMT targeting is better than current methods of targeting used in the Eastern Caribbean states.

Sharif points out further that the estimates of accuracy reported in papers usually incorporate the implicit assumption that all potential beneficiaries know about the programme and are tested for eligibility, and that programme implementation is "perfect" Sharif (2009: 22). In making this assumption, papers underestimate the exclusion error. Kidd & Wylde (2011) report that while deskwork predicted exclusion errors of 42 percent and inclusion errors of 38 percent for Mongolia, during implementation the actual errors were 21 percent and 57 percent respectively.

The incorrect exclusions and inclusions result because the PMT equation reflects a prediction that is relatively accurate on average, but makes no claims to be accurate on an individual basis. This is especially concerning from a rights-based perspective in that it is difficult to justify the exclusion of a household that objectively qualifies and needs a benefit on the basis that the factors that statistically determine their expenditure as a specific household differ from those for some notionally average household.

## Choosing the eligibility cut-off

A decision maker who relies only on statistics might choose the point at which the sum of the exclusion and inclusion errors is smallest. However, Narayan and Yoshida (2005) observe that the choice of the cut-off point for targeting is a **policy** decision rather than a **technical** decision. If a government makes the policy decision to target households that are below a certain poverty level, it would constitute an enormous coincidence if this poverty level was the same as the point that minimized the sum of the exclusion and inclusion errors. If a government decides to restrict

the amount of money spent on a benefit, it is likely to favour a lower cut-off point, even if this has a higher exclusion error. Thus Sharif (2009) observes that PMTs are a good investment because, although costs are incurred for their administration, they reduce inclusion errors (and thus costs). If a government chooses a policy to ensure that those who are in need receive assistance, it is likely to favour a higher cut-off point—even if this means that some non-poor people are eligible for the benefit.

Sabates-Wheeler et al (2014: 2) suggest that policymakers might be inclined to focus on inclusion errors because of the unnecessary costs incurred through such errors. However, they cite Cornia and Stewart's observation that exclusion errors merit greater weighting because of the "humanitarian" cost incurred when people who need assistance are excluded and the objectives of the intervention thus undermined.

The choices are complicated by the fact that reducing inclusion errors could make more money available for benefits for those who are eligible if the benefit's total budget allocation held constant. The danger with this reasoning is that policymakers might instead decide to use the money saved for other purposes.

Grosh & Baker (1994: 9) observe that while those favouring universal subsidies traditionally favoured the minimization of exclusion errors, the "tight budgetary constraints" of the 1980s and 1990s saw a shift in emphasis towards targeting as a way of reducing expenditure. This shift resulted in a greater emphasis on reducing leakage. This bias might be even stronger in the wake of the ongoing global and financial crisis. However, these considerations must be balanced against greater recognition of the importance of a right to social protection.

Another complication in choosing a cut-off point is that there is usually a variance between the PMT equation's predicted percentage and the actual percentage of the population that is poor (in terms of the chosen poverty line). This statistical result is also reflected in the difference between the exclusion and inclusion rates at the poverty line point.

In general, where the percentage of the population targeted for eligibility is low, the percentage predicted by the PMT will be higher than the targeted percentage. For example, Grosh & Baker (1994: 16) find that the PMT equation will predict only 33 percent of the population to have expenditure below the actual maximum expenditure, according to the survey data, for the poorest 40 percent of the population. Conversely, a PMT equation will predict that 18 percent of the population has expenditure that is below the actual maximum expenditure, in the survey data, for the poorest 10 percent of the population. In Bangladesh, similarly, Sharif finds that the equation based on expenditure of the poorest 20 percent of households results in a model that predicts expenditure below the cut-off for only 13 percent of households (Sharif, 2009: 21). For Saint Lucia, 73 (83 percent) of the 88 households (unweighted) predicted to be indigent record expenditures above the indigence cut-off point. This phenomenon reflects the clustering of households at the lower end of the expenditure range, making it difficult to distinguish between poor households in terms of expenditure. IOS Partners/ London School of Economics (2011) note further that the use of ordinary least squares regression results in greater inaccuracy for the lowest (poorest) and higher (richest) ends of the expenditure distribution.

# 2 MEASURING POVERTY

# MEASURING POVERTY

## Poverty and Indigence lines

There are different ways of calculating a poverty line. Some favour a relative approach. A simplistic form of this approach categorizes the poorest 20 percent (or quintile) of the population as 'extreme poor' and the next poorest 20 percent as 'poor'. With this approach, a country can never reduce the poverty rate, as it is defined as the percentage of the population (or households) below the poverty line. A more sophisticated relative approach is to define poverty as having an income below some fraction of the median income in the country. The International Labour Organization uses this approach, with the fraction set at two-thirds, to determine the low pay rate for an individual.

Another common approach (one that seems to be used in the Eastern Caribbean states), is to set the extreme poverty (or indigent) line at the level of expenditure needed to buy the basic minimum of calories (spread across the appropriate food groups) scientifically determined to be necessary. The Caribbean Nutrition and Food Institute provided the oft-quoted amount of 2,400 kilocalories per day for an adult. No allowance is made for any expenditure other than on food.

A household will only escape poverty (or indigence) at this level of expenditure if it meets two unrealistic assumptions. First, it must spend every cent in the most economical and judicious way, based on full knowledge of nutrition and market prices. Second, it must not have other needs besides food. The poverty line adds to this basic food amount a further allowance for non-food items. The allowance is based on the average non-food expenditure of the bottom two quintiles of households. For Saint Lucia in 2005–2006, this yielded an extreme poverty (indigence) line that contained 1.2 percent of all households (1.5 percent of all people), and a poverty line that contained 21.4 percent of all people (28.8 percent of households).

The extent to which an indigence line based on food is meaningful might differ according to the level of development in a country. In South Africa, a country with a similar gross domestic product per capita to the Eastern Caribbean islands, Woolard & Leibbrandt (1999) find the relationship between caloric intake and other measures of poverty to be weak.

### Unpaid care work

Means tests, whether proxy or direct and whether based on income or expenditure, generally count only monetary amounts. Exceptions are often made by imputing amounts for owner-occupied housing and occasionally for self-consumed food production. However, means tests fail to consider the value of unpaid care work—unpaid work that is performed mainly by women and includes caring for other household members (especially children), cooking and other housework.

Folbre (2008) points out that if these costs were considered, analysts would not so readily assume that young children require less expenditure than adults as is frequently done (see discussion below). An approach that considers unpaid care work will also need to take into account the income foregone by mothers and others with heavy care responsibilities in terms of simultaneously managing income-earning and other work and by accepting lower-paid work because it more easily allows them to do unpaid caring. Folbre thus implicitly argues that one should consider the ability to earn income, rather than only whether a household or individual has enough income to satisfy its basic needs or incurs sufficient expenditure to satisfy its basic needs.

Folbre also critiques existing equivalence scales (see below), including the ones used in the United States, for omitting expenditure on college education as well as the cost of child care for younger children. More generally, she criticizes the fact that measures of need tend to focus much more on goods (such as food, housing and clothing) than on services, an approach she describes as "antiquated" (Folbre, 2008: 55).

## Equivalence Scales

'Equivalence scales' is the generic term used to describe adjustments made in calculating poverty rates in order to account for differences in the size and composition of households. Even in OECD countries, there is no single accepted scale. The two most common scales used are as follows (OECD, 2013):

- OECD or 'Oxford scale': Value of 1 to first household member, 0.7 for each additional adult and 0.5 for each child.
- OECD-modified scale: Value of 1 for first member, 0.5 for each additional adult and 0.3 for each child.

A third scale used by the OECD, the square root scale, divides household income by the square root of household size, irrespective of household members' age. This scale is used primarily to compare income inequality and poverty across countries, and so it is less relevant to the purposes of this paper.

The OECD notes that the difference between various measures increases with the size of the household. To the extent that poorer people tend to live in bigger households, the choice of equivalence scale affects the estimates for poor people most.

## Economies of scale

This discussion assumes that income and expenditure levels tell us how rich or poor a household is. However, complications arise in making this assessment for households with multiple persons. At one extreme, assume that the full benefit of the expenditure is available to each person in a household; a household with five members would be deemed five times as wealthy as a household of one member with the same level of total expenditure. At the other extreme, assume that the total expenditure is divided equally among all household members; in that household, each member only benefits from a fifth of the total amount.

The first extreme is clearly unrealistic. The second extreme is also unrealistic when considering what economists term 'economies of scale' and 'public goods'. Economies of scale refers to the principle that it typically costs less than double to do something for two people than to do it for one person. Public goods refers to instances where one person's consumption of a good or service does not result in there being less of that same good or service for another person to consume. The most common example given for these phenomena is housing, where a dwelling for two will not cost double the same quality of dwelling for one, and where the fact that one person lives in the dwelling does not prevent (a reasonable number of) other people living in the house. Food is a counter-example to a large extent, in that more than one person cannot eat the same portion of food. There are, however economies of scale in that larger amounts of food can often be bought more cheaply and there will be shared costs in preparing the food.

Food and housing are important examples because they generally account for a large part of expenditure. In Saint Lucia, the 2005/06 survey suggested that for households in the poorest quintile, 32 percent of expenditure went on food. Compare this to 19 percent in the wealthiest quintile (Kairi Consultants Limited, 2007: 48). In terms of housing, imputed and actual rentals accounted for an average of 14 percent of expenditure, of which more than 13 percent was imputed (i.e. the money is not spent, but the amount reflects what the household would need to pay if it did not own the household) (Kairi Consultants Limited, 2007a: 50).

Buhman et al (1998: 119) provide a categorization for what they describes as a "potpourri of equivalence scales" used in making adjustments for size of households. They divide the scales into four categories, two that they describe as "expert-based" and two as "survey-based." The expert-based scales include those that incorporate expert's views on how need will vary with family size (referred to as STAT) and those

that are based on the level used for poverty-related benefit schemes (PROG). The survey-based scales include those that are based on multivariate analysis of expenditure as is done in PMTs (CONS) and those that take into account respondents' assessment of the adequacy of their income (SUBJ). Deaton & Salman (undated: 49) similarly identify scales based on behavioural analysis using survey data (equivalent to CONS), on questions asking about subjective welfare (SUBJ), and on scales derived "in some reasonable, but essentially arbitrary way" (PROG and STAT).

The extent to which necessary expenditure should be adjusted for economies of scale related to the size of the household is referred to as elasticity. In their testing, Buhman et al use the following elasticity values—based on their reading of common practice—for the four approaches: 0.72 for STAT; 0.55 for PROG; 0.36 for CONS; and 0.25 for SUBJ.

Buhman et al use data from ten developed countries (Australia, Canada, Israel, Netherlands, Norway, Sweden, Switzerland, United Kingdom, United States and West Germany) for the period from 1979 through to 1983. The focus on developed countries differs from much of the other literature reviewed, which focuses on developing countries. It is germane to the aims of this paper, as the Eastern Caribbean islands probably lie somewhere between the developed countries and the poorer developing countries used in some of the literature.

Buhman et al focus on the size equivalence measure as they claim that this adjustment is the one most commonly used rather than, as seems to be the case for developing countries, adjustments for age. This is in line with Deaton & Zaidi's (undated) observation that economies of scale are likely to be more important for developed countries, where food constitutes a smaller proportion of expenditure of poor households. It is also in line with the OECD's (2013: 174) suggestion that equivalence scales will differ for low- and high-income households, as the proportion of income spent on housing tends to decrease as income rises and housing is one of the main sources of economies of scale (OECD, 2013: 174).

In exploring how poverty rates change as the definition of the poverty line changes, Buhman et al find that in most countries, poverty rates decline with increases in the adjustment for economies of scale (the 'elasticity of the equivalence'). This reflects the fact that poor people tend to live in larger households. However, the extent of the change differs widely, disallowing the derivation of any cross-country guideline.

For further analysis, they set the poverty line at half the median adjusted income (or well-being) measure. They then find that the composition of the population defined as poor changes when different equivalence scales are used. For single mothers with one child, the poverty rate drops by 40 percent or more in Germany, Israel, the Netherlands and the United Kingdom if one uses the PROG rather than the SUBJ approach. In Australia, older people account for 37 percent of the poor using SUBJ, 19 percent using PROG and 9 percent using STAT. Again, the results are too mixed to provide easy lessons.

Deaton & Zaidi favour the use of the expert approach (STAT), and advise strongly against the use of the survey-based approach (SUBJ). In respect of the latter, they comment that "it is hard to take these estimates seriously" (undated: 50). Folbre (2008: 40) points out that the results of behavioural approaches (equivalent to STAT) imply that children of wealthy families are "worth more" on the basis that their parents are able to spend more on them. The point could be broadened to a statement that behavioural approaches (of which PMT is one) reflect current practices rather than need. This is a problem to the extent that the household might 'choose' (in economists' terms) or 'be able' (in practical terms) to spend differently if it had more resources.

For the STAT approach, Deaton & Zaidi suggest a formula that includes both a household size adjustment and a single adjustment rate for children of all ages. They do not specify the exact values to be used for the two, but suggest that the child ratio could range from as low as 0.3 in the poorest countries to close to 1.0 for the United States and Western Europe. For the household size adjustment, they suggest a value of close to 1.0 for poorer countries and close to

0.75 for wealthier ones.

Deaton (1997: 243) notes that to the extent that the elderly tend to live in smaller households and children in larger households, the greater the adjustment for economies of scale, the less likely children are to be categorized as living in poor households. Conversely, no adjustment for economies of scale will make elderly people less likely to be found to live in poverty. There is also likely to be a gender effect to the extent that men are more likely than women to live alone. In the Saint Lucia SLC of 2005/06, for example, 67 percent of single-person households consisted of a man. Use of adjustments for economies of scale will therefore increase the relative extent to which men are judged to be living in poverty.

Folbre (2008: 74) argues against placing too much weight on economies of scale. She points out that at least some of the reduction in observed spending per capita should be interpreted as "belt-tightening" in order to meet the needs of a larger family rather than "increased efficiency" due to economies of scale. Deaton & Zaidi (undated) similarly suggest that even when using the expert approaches in poverty analysis, the simple per capita approach should be retained because of its simplicity, its widespread acceptance and because its strengths and weaknesses are well understood.

## Adult equivalents

The literature reviewed for this paper showed no evidence of any adjustments for economies of scale in the Eastern Caribbean. There are, however, often equivalence scale adjustments for children and sometimes for female adults. The 'per (male) adult equivalent' expenditure scale is derived by dividing total expenditure by the sum of the individuals, with some individuals counting as only a proportion of a full male adult (instead of deriving per capita household expenditure by dividing total expenditure by the number of household members). In essence, use of an adult equivalence scale rather than simple adjustments for household size reduces the relative chance of a household with children (and with female members where a gender adjustment is made) being considered poor when compared to a household consisting only of adults (males).

Some decades ago, it was relatively common to assume that women required somewhat less expenditure than men on the basis that they were smaller physically and less likely to do strenuous work, and therefore needed fewer calories. This led to the broader assumption that women generally cost less, as food constitutes a large proportion of expenditure (especially for the poor). This assumption is rarely used nowadays. It is seen as inequitable and also does not take into account that biologically (especially during pregnancy) women's needs may be equal to or even exceed those of men. Folbre (2008: 54) reports that the United States Department of Agriculture stopped weighting women less than men in their scales after threats that they would be prosecuted for sex discrimination.

Sex differentiated equivalence scales have been used up to the present in the Eastern Caribbean—a very unusual practice. For example, they were used in the SLCs and were incorporated in the derived variables in the Saint Lucia SLC data set (and probably in the datasets for other countries). This unusual Caribbean practice is reportedly based on an exercise done in Belize in the early 2000s (Kairi Consultants Limited, 2007b: 14-15).

Because these equivalence scales are embedded in the SLC data set, they are also used for SL-NET. What is especially unusual about the approach is that it treats the sex difference as if it starts at the age of one year—long before females tend to be smaller than males or engage in less strenuous work. Older adults are also weighted less than of an adult male aged 19 to 29 years. Overall, the equivalence rates range from 0.270 for all children under 1 year to 1 for and adult male 19 to 29 years, and back down to 0.618 for a woman aged 61 years and above.

Andaiye (2003: 85) notes that use of a sex differential will result in underestimation of poverty for female-headed households. This effect extends beyond female-headed households; the likelihood that a household will be classified as poor decreases as the female proportion of household membership increases.

It is not clear which method was used to develop the equivalence scale used in the Saint Lucia SLC. The Engel approach, used fairly commonly in the past, is based on the correct observation that the proportion of total expenditure allocated to food tends to decrease with increasing wealth. The approach then incorrectly uses this observation and the proportion of the household budget spent on food to determine the cost of an additional child in the household. The literature confirms that the Engel approach is theoretically unsound and "indefensible" (Deaton, 1997: 259; Brazilian Institute for Economics and Statistics et al, 1997). More generally, Deaton is clear that regression should not be used to derive equivalence scales (personal communication).

Experts are less dismissive of the theory of what is termed the Rothbarth approach, although they point out the difficulties of determining what constitutes the "adult goods" that are needed for the Rothbarth calculations and highlight the likelihood that the model omits important factors. They note that the Engel approach tends to give relatively high values for the adult equivalent (i.e. the proportion of an adult's expenditure that should be allowed for a child), while the Rothbarth approach tends to give lower values of around 0.25 (Deaton & Muellbauer, 1986: 720). They suggest that the appropriate value will lie somewhere between typical Engel and Rothbarth estimates, and that the equivalent will tend to be lower in developing countries. Folbre reports that in the United States, the recommended adult equivalent for children is 0.7 (Folbre, 2008: 52). This is higher than the level generally used in developing countries, despite the United States child estimate excluding a range of important items.

In his 1997 classic, Deaton (1997: 259) suggests that a weight of 0.4 for children under 5 years and 0.5 for children aged 5 to 14 might be appropriate, but acknowledges that these values are "arbitrary to some degree." In a later co-authored article (Deaton & Muellbauer, 1986: 722), child costs are estimated at between 0.3 and 0.4 for Indonesia and Sri Lanka, as compared to Engel estimates of around 0.82 if there is one child and 0.77 percent if there are two children. Deaton and colleagues argue further that the values will be higher for more developed countries where there are more substantial non-food expenditures incurred in respect of children. Overall, Deaton—an acknowledged global expert on these issues—is of the view that there are no exact methods for determining the size of either economies of scale or equivalence scales.

The Brazilian Institute of Economics and Statistics et al (1997) observe that whereas choice of equivalence scale may result in only small changes to aggregate measures (such as the proportion of the population that is poor), but that the use of different measures may change the demographic profile of households that are classified as poor. This, in turn, suggests that different equivalence scales will result in different households being classified as poor across the various scales.

Prior to the introduction of the Belize sex-age scales, the common approach in the Eastern Caribbean was based on calorie requirements computed by the Caribbean Food and Nutrition Institute. This scale equates a child under seven years of age to 0.2 of an adult; aged 7 to 12 to 0.3 of an adult; and aged 13 to 17 years to 0.5 of an adult (Saint Catherine, 2004; Tang, undated). These estimates are on the low side of those used internationally. Folbre's calculations, based on local estimates of actual costs, put the cost of even a preschool child at between 0.57 and 1.0 of an adult's (Folbre, 2008: 57).

Woolard & Leibbrandt (1999) question approaches that base the child adjustment on caloric requirements. They observe that the fact that a child of a specified age requires some percent the amount of calories required by an adult male does not mean that children's non-food expenditure decreases proportionately to an the non-food expenditure of an adult. Woolard & Leibbrandt follow the example of earlier South African research, using a child ratio of 0.5 alongside 0.9 for economies of scale.

Woolard & Leibbrandt's analysis suggests that choice of equivalence scales has very little impact on the rates of poverty or on geographical and racial patterns. Two caveats are in order here. First, the fact that patterns

on geography and race remain consistent reflects, at least in part, the strong differentials in these respects inherited from the Apartheid era. Second, while the rates might be similar using different methods, the individual households classified as poor and non-poor may differ. The latter is especially important when discussing targeting of benefits.

## Testing the effect of different adjustment scales

Kidd & Wylde (2011) test the effect of adult equivalence scales at the individual household level by comparing results without such scales and then using a scale in which children are given a weight of 0.5. They find that for Rwanda and Indonesia, 8 to 12 percent of households are rated differently (in terms of being poor or not) on the two scales; the difference affects more than 16 percent of households in Bangladesh. When these errors are combined with regression errors in Bangladesh—for a programme that covers 5 percent of the population—only 20 percent of the households ranked in the poorest 5 percent of reported actual expenditure are classified as poor using both the adjusted and unadjusted equivalence scales, 67 percent are excluded when using both scales, and 13 percent are included by one scale but not by the other. These disparities are important because they reflect the extent of cases that would receive a benefit if one approach was used but would not receive the benefit if the second approach was used.

This paper tests the impact of choice of equivalence scale in Saint Lucia using data from the SLC 2005/06. It compares results of three different approaches, namely the unusual Belize sex-plus-age adjustments used for the PMT, the simple child adjustments used prior to the introduction of the Belize approach and a simple per capita calculation with no adjustment for age. The same poverty threshold level at the individual (adult male) level is used for all three measures.

Use of the old scale results in a small decrease in the overall poverty rate—from 22 percent using the Belize scale to 21 percent with the old scale. In contrast, use of a simple per capita measure results in an increase of the overall poverty rate to 34 percent. Further analysis reveals that even the move to the old scale results in noticeable differences for particular family forms. For the purpose of the analysis, family form is defined on the basis of the presence in the household of a child or children under 18 years ('child'), men aged 18 or above ('man') and women aged 18 or above ('woman').

Table 2 shows poverty rates by family form and equivalence scale. Child-only households are omitted because the numbers are too small to produce reliable results (the same is probably true of the households that contain only adult men and children). In addition, the table does not distinguish between households with girl and boy children because the large number of different combinations of boys and girls would result in cell sizes too small for reliable analysis. If such analysis was possible, it would reveal the extent to which the Belize scale disadvantages households with girl children when assessing for poverty.

Table 2 shows a decrease in the poverty rate for all except the woman-only and man-only households when comparing the old approach with the Belize/SL-Net approach. This pattern is a result of the smaller child equivalent ratios in the old approach. These decreases outweigh the smaller increase from

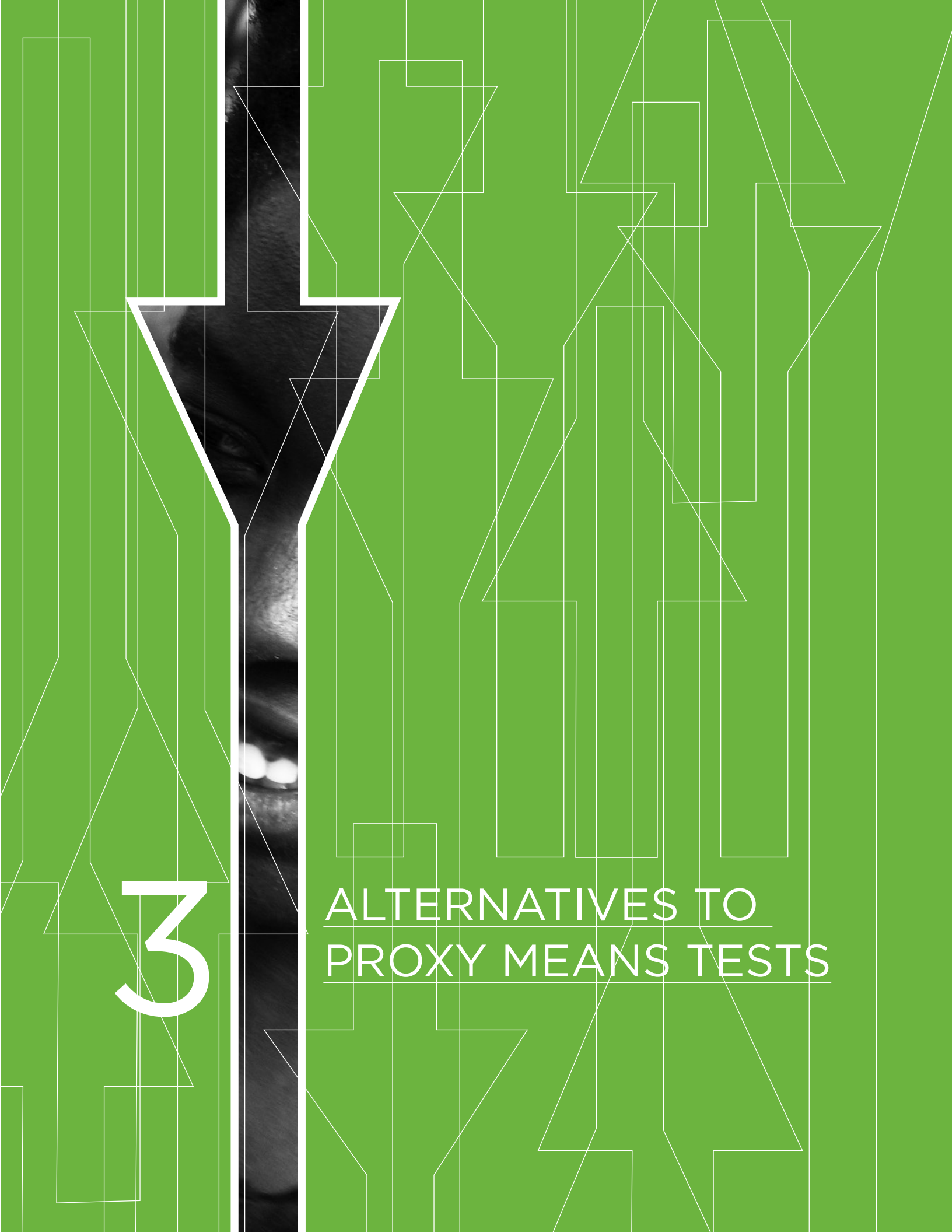
**Table 2: Poverty Rates by Family Form and Equivalence Scale, Saint Lucia 2005/06**

| | Child and Woman | Child and Man | Child, Woman and Man | Woman Only | Man Only | Woman and Man |
|---|---|---|---|---|---|---|
| All households | 6141 | 954 | 19771 | 4041 | 6428 | 8981 |
| Belize approach | 24% | 28% | 29% | 4% | 20% | 11% |
| Old approach | 17% | 17% | 24% | 15% | 24% | 18% |
| Per capita | 44% | 44% | 45% | 15% | 24% | 18% |

removing the sex differential. The per capita approach results in a substantial increase in the poverty rate for all household forms that include children. These households now have poverty rates three times as high as the woman-only households.

A similar comparison for households that consist of one adult woman and one or more children shows the poverty rate at 20 percent using the Belize approach, 12 percent using the old approach and 39 percent using an unadjusted per capita calculation. These calculations do not show which of the approaches is optimal, but provide strong evidence that choice of equivalence scale matters.

The relevance for the PMT is that changing the equivalence scale to align it with international good practice would necessitate a change in the weights used for the PMT variables, would likely result in a change in some of the predictor variables, and would result in a change in which households are identified as poor or non-poor.



# 3 ALTERNATIVES TO PROXY MEANS TESTS

# ALTERNATIVES TO PROXY MEANS TESTS

## Multidimensional poverty

Although PMTs are based on multiple indicators, their aim is to measure income poverty rather than multidimensional poverty. In contrast, the test derived for Grenada is described as a "multidimensional living conditions assessment" designed to target the "most deprived population" rather than a PMT (Diaz et al, 2014). Instead of trying to predict income or expenditure, the Grenada test focuses on, and attempts to maximize, the "kind of life of the household." Instead of using variables chosen on the basis of their statistical contribution to predicted expenditure, Diaz et al argue that policy makers should be given discretion to choose characteristics that are "desirable and socially valuable" and "normatively important [even] if they are not statistically predictive."

Diaz et al's approach is interesting given current interest in the Eastern Caribbean in multidimensional poverty measures. The interest is, however, not necessarily focused on the use of such a measure for targeting of particular households or individuals. The interest seems to be more generally whether it is possible to develop a regionally appropriate multidimensional instrument to measure overall levels of poverty. A multidimensional instrument may also give relatively accurate poverty rates for particular groupings (such as those defined geographically), because inclusion and exclusion errors cancel each other out. The instrument would nonetheless be inaccurate at the level of the individual household which is needed for the targeting of benefits.

IOS Partners/London School of Economics (2011: 3) discuss the advantages and disadvantages of wealth indices and multidimensional approaches when compared to a PMT. They observe that both approaches select variables as indicators of need or well-being in their own right rather than through a "detour via expenditure data." They observe that of these measures share a serious disadvantage; there is no intuitive definition of a poverty line based on assets, housing characteristics or other dimensions of deprivation. The levels are thus arbitrarily set, unlike where a PMT is used to target a poverty or indigence line. In a multidimensional poverty index such as the Alkire-Foster measure, the weights attached to variables are dimensions that are based on their perceived importance rather than the results of a statistical test. Further, the number of variables/dimensions in which a household must be deprived to pass the test is also arbitrarily set. Meth (2014) illustrates how the standard Alkire-Foster measure produces a multidimensional poverty rate for South Africa that is substantially lower than the standard expenditure-based poverty rate calculated by the official statistics agency.

Diaz et al (2014) test their index against survey data using the poverty line as a partial defining feature of eligible households. In addition, for a household to be eligible in terms of their testing of the multidimensional approach, they require that at least one household member qualifies in terms of programme eligibility rules (such as being disabled or chronically ill, pregnant or lactating, a child, or elderly). In doing so, they seem to confuse the purpose of a PMT—which is to identify those who are poor—with supplementary categorical characteristics that identify potential beneficiaries.

Diaz et al (2014) compare the results of their test against a standard PMT that is based on principal component analysis. The latter identifies a larger number of households as deprived (or poor) in that 89.2 percent of all households are identified as deprived on both measures, 9.3 percent are identified as deprived by the PMT but not on their multidimensional measure, and 1.5 percent are deprived on their measure but not on the PMT. They note that those households identified by the PMT (but not by their favoured multidimensional measure) as deprived tend to have higher deprivation in the child-related variables. This suggests that their measure is less child-friendly than the PMT.

Diaz et al (2014: 14) consciously chose only variables that were available in the SLC and the 2001 and 2011 censuses (and that might also be available in a future administrative register). The variables chosen to provide a dimension reflecting the importance of childhood in terms of the future of the country and the "intrinsic vulnerability" of children include the presence of toddlers, households in which children had died in recent years and households in which children were engaged in child labour. The variable on past mortality is questionable, given that it would be too late to assist the deceased child and there might not be other children in the household.

Diaz et al cite as a strength of their model that it does not apply differential weights to the various indicators. However, this is arguably a weakness—some indicators are far more important than others, even if one does not believe that the relative importance should be determined by regression or principal component analysis. The need for differential weighting becomes stronger when, as happens in their measure, the survey does not contain good questions to measure all aspects that they consider important contributors to multidimensional poverty.

The multidimensional approach is less scientific than the PMT, and the choice of indicators is likely to be fairly arbitrary (but at the same time constrained by the questions asked in a survey). The formula used is more complicated than would be the case in using a simple checklist of questions, but in essence the approach seems similar to having a simple checklist. One weakness relative to a PMT is that there does not seem to be the same concern with choosing easily verifiable variables.

## Community-based targeting

Using a real-life experiment rather than the desk-based research used in some other PMT investigations, Atlas et al (2010) provide a very thorough analysis of the relative strengths and weakness of PMT and community-based targeting in the Indonesia context. In the community-based approach, ordinary members of the village come together to rank each household so as to determine those who fall within the poorest percentiles. In a variation of this approach, to test elite capture, the ranking is done by community leaders rather than by ordinary members of the community. The comparison was done by dividing the 640 villages into three groups. The PMT was implemented in the first group; a community-based approach was used on the second group; and a hybrid approach was used in the third group. For the hybrid approach, a community ranking was used to identify poorer households, which were then subjected to the PMT.

Atlas et al find that community-based targeting performs worse than PMT when measured against per capita expenditure. However, community members show higher levels of satisfaction with the outcome of the community-based approach, there are fewer formal complaints and the ranking matches well with how individual community members rank themselves. The higher satisfaction made it easier for officials to disburse the grants. The community-based targeting also achieves better results than PMT in identifying the extreme poor with per capita consumption below one US dollar per day.

Atlas et al conclude that communities have a different concept of poverty than one based on income or expenditure alone. In particular, it seems that communities consider factors that relate to earning capacity rather than considering only the level of consumption (or of earning). Thus, for example, they favour widowed households over other households with the same expenditure level. The results also suggest that the community believes there are

economies of scale in that larger households were less likely to be deemed poor than smaller ones with the same per capita expenditure. In contrast, there is little evidence of a belief that children are cheaper than adults, as the community tended to rank households with children as poorer.

Their test produced no evidence of elite capture. In addition, gender composition of the community meetings resulting from holding them at different times of the day did not affect the targeting results. However, a challenge with the community-based approach was the effort and time required from community members and a substantial fall-off in accuracy of targeting after about 20 percent of households have been ranked.

## Comparing alternative targeting systems

Coady et al (2004: vii), after reviewing approximately 100 examples of targeted programmes in low- and middle-income countries, conclude that when it comes to targeting, there is a clear "central message," that "there are no magic bullets." Instead of calculating exclusion and inclusion errors, they try to estimate the extent to which targeting approaches have an outcome (or reach) that is different to that which would have been achieved by distributing benefits randomly across rich and poor. They find that about a quarter of the programmes reviewed are regressive, delivering more benefits to the rich than the poor. They find that Argentina's Trabajar public works programme performs best, transferring 80 percent of benefits to the poorest quintile.

Coady et al warn that readers should be cautious about their assessment as the performance measure "is a mishmash of various measures" (2004: 39). Their measure is also not very useful for review from a rights perspective, as it does not provide the number and proportion of poor people who are excluded, and thus denied their rights. Further, the variety of types of programmes included in their assessment (only 40 percent relate to cash transfers and only nine programmes, in five countries concentrated mainly in Latin America and the Caribbean, are PMTs) might decrease the direct relevance of some of the findings for exploring the PMT recommendation. Of interest, nevertheless, is that they state that 80 percent of variation in targeting performance resulted from differences within targeting methods (such as between one PMT and another PMT) rather than across methods.

Castañeda & Lindert (2005) present six country case studies whose targeting systems span verified means testing (VMT) in the United States, unverified means testing (UMT) in Brazil, and PMT in Chile, Colombia, Costa Rica and Mexico. The difference between the VMT and UMT is that in the former the reported income is checked against other sources by officials, while in the latter it is not. Castañeda & Lindert assess the success of the targeting methods in maximizing inclusion, minimizing exclusion, cost efficiency in terms of administration, and transparency. Their analysis utilizes data from the unified household information registries in these countries (which include all households that have been assessed, including those not deemed eligible), and programme-specific beneficiary lists that include only those deemed eligible.

Castañeda & Lindert find that Brazil's UMT is "reasonably accurate" (2005: 24), despite the reported income not being checked. PMTs are "nearly as accurate" as VMT, and "in some cases" (2005: 25) (but not all) more accurate than UMT. They suggest that one of the reasons that Brazil's UMT performs relatively well in targeting the poor is that it includes an element of geographical targeting.

Interpretation of their analysis on the accuracy of targeting is complicated by the fact that it is based primarily on registered households rather than on the population as a whole. The registers typically include only those households that have been assessed using the means test, including both those deemed eligible and those deemed ineligible. Coverage rates range from a very low 16 percent of the poorest quintile receiving Chile's old-age benefits to 60 percent for Mexico's Oportunidades programme. The low percentage for the old-age benefit is at least partly explained by the fact that not all poor households will include older people among their members. This example illustrates the weakness of judging targeting efficacy only by coverage of the poor when programmes have criteria

and purposes other than poverty alleviation.

Cost-wise, PMT is cheaper than VMT and has similar costs to UMT. However, Castañeda & Lindert note that this assessment does not take into account the sometimes heavy private costs that may be incurred by applicants in the various systems. The administrative burden of PMT is similar to that for UMT and less than for VMT.

Generalizing about the cost of a PMT is complicated by the fact that in some systems, the test is administered to the full population or to all residents in poor areas. Such full surveys obviously have associated costs, which can be substantial. In systems where the test is administered only to those who apply, the costs increase if those assessed as potentially eligible through an office-based interview are then subject to a home visit.

The Saint Lucia SSNA notes that internationally, most of the programmes that use a PMT still require an interviewer or social worker to visit the household to fill out the proxy-means testing form and to verify the conditions of the household as reported on the application form (Blank, 2009: 46). Castañeda & Lindert (2005) estimate that the cost of interviewing a household is 30 percent cheaper per interview when it is done through a survey rather than after an application, but the cost for the application-based interviews is reduced by the smaller number of interviews that need to be conducted.

# 4 IMPLEMENTING A PROXY MEANS TEST

# IMPLEMENTING A PROXY MEANS TEST

## Application-based or survey of potential beneficiaries

PMTs are administered to the full population, to all households in poor areas or to only those who apply. The disadvantage of administering the test to the full population is the considerable effort and expense involved. The disadvantage of the application-based approach is that poor eligible households may not apply because they do not know about the benefit or because the time, effort, costs and other challenges associated with applying are too great. Choosing this approach requires energetic promotion and social marketing of the benefits and details of how to apply, careful consideration of the costs that are likely to be incurred in doing so, and the challenges faced by poor households that want to apply. Coady et al (2004) note that publicizing the benefit and how to apply is an aspect that has often been neglected in favour of refinement of the formula or administrative aspects that focus on reducing inclusion error rather than reducing exclusion of those who might be eligible but do not apply.

Requiring applicants to provide proof of characteristics included in the test adds another potential barrier and burden. Coady et al (2004) describe how the substantial costs associated with applying for and collected benefits from Ecuador's Bono Solidario contribute to low take-up.

Chile, which spearheaded the use of PMTs in 1980 to determine eligibility for two large cash transfer programmes and water and housing subsidies, changed from a survey-based to an application-based approach in the 1990s. They did so on the basis of widespread knowledge of the benefits, a decrease in the poverty rate and the fact that poverty was not geographically concentrated, disallowing surveys based on poorer areas. Urban Mexico and the US also have application-based systems.

Designers of PMTs aim to select characteristics that are easily verifiable, objective, and difficult to lie about. Grosh & Baker (2014) note that family characteristics, such as the level of education, occupation and number of members, may be difficult to verify but are frequently assumed by programme managers to be reported accurately. In terms of assets, they note that households may remove small goods such as vehicles and radios when they are visited.

At least some of the variables used for a PMT will not be verifiable on the basis of an office visit. Many systems therefore provide that the interview is followed up by a home visit, either for all deemed eligible or for a sample. Grosh & Baker (1994: 30) suggest that where there are checks for only a sample, a penalty needs to be applied to those who are found to have given incorrect information. They draw an analogy in this respect with the way in which the US income tax system (and that of other countries) works.

The costs and burdens associated with a PMT are influenced by the length and nature of the questionnaire. Coady et al (2004) report that the form used in Chile is only two pages long and collects only the information required for the test. The same is true of the Saint Lucia PMT and Dominica's test. In contrast, the Mexico form is 20 pages long; the Grenada form is

six pages long.

Castañeda & Lindert (2005) discuss the challenges in ensuring that information in the registries is up-to-date and the discrepancy between the planned regularity of such updating and verification and what happens in practice. They observe further that most countries have not paid sufficient attention to ensuring the quality of the information in their registries.

## Defining the beneficiary unit

Castañeda & Lindert (2005) discuss some of the complications involved in linking individuals to households and/or families. They explain that in Colombia, Chile and Costa Rica the 'family unit' refers to a group of people within a household who are related by blood (similar to a nuclear family), whereas the 'household' may include more distant relations and people who are not related. In addition to the closeness of the relationship, the definition of the family unit also takes into account whether members are dependent on each other in terms of income. Further complicating the matter is that different programmes may use different beneficiary units, which will have different PMT scores. They provide the following example of a single household made up of four family units:

- Family 1: Father, Mother;
- Family 2: Married son, wife and their two children;
- Family 3: Unmarried daughter, her infant; and
- Family 4: Family friend living in house.

They note that eligibility criteria such as housing conditions, public services and durables in the home will be the same for all the families listed above, but criteria such as the education of family members, dependency ratio and crowding will differ. Castañeda & Lindert (2005) report that such complications are leading to Chile and Colombia to consider dropping the family-household distinction.

None of the literature reviewed discussed whether the beneficiary unit assessed matches the legal obligations in terms of support in the country's laws. In most countries, parents have a legal obligation to support their children (an obligation that might apply in the reverse). Similar obligations may not reside in other relationships, which raises questions regarding the validity of assessing or estimating income for a unit containing these other relationships.

None of the literature reviewed discussed the challenges that may arise when individuals join or leave the household (or want to do so). From a gender perspective, one needs to consider the case of a woman victim of domestic violence and how the definition of household, PMT approach, choice of person to whom the benefit is paid and the length of application process may facilitate or obstruct her choices regarding her situation and that of her children. Ultimately, any means test that is based on the household or family unit may not be appropriate for assistance that facilitates women and children escaping from domestic violence.

## Other eligibility criteria

In contrast to the PMTs, which focus on income poverty, the unified household targeting systems proposed in the Eastern Caribbean SSNAs are envisaged as serving as targeting mechanisms for multiple social programmes. If the PMT is used as a criterion in the unified targeting system, then the PMT eligibility thresholds may differ among social programmes. In addition, these other social programmes may need additional criteria to determine eligibility because the PMT is a proxy test for income and primarily applicable to poverty reduction programmes. In particular, it makes the most sense for cash grant programmes as the benefit is then directly compensating for the low income. In reality, many programmes do not have poverty reduction as their only aim, or maybe even their main aim. For this reason, among others, eligibility tests often include other elements alongside the PMT.

Grosh & Baker (1994: 26) provide examples of how Chile's PMT is used together with other eligibility criteria specific to particular benefits. In some cases (such as for the old age grant), the additional criteria include direct assessment of income. In such a case, the PMT serves as a first proxy test of income, with the direct test of income added for those who pass the first test.

IOS Partners/London School of Economics (2011) suggest that for Belize, the PMT could be supplemented by a simple multidimensional index that captures characteristics other than monetary poverty. This suggestion does not, however, cater for benefits that are targeted at particular characteristics (such as disability of a household member) rather than a cluster of characteristics.

Coady et al (2004) observe that 60 percent of the programmes they studied use more than one targeting method either in combination or sequence. For example, a child allowance (categorical targeting) may be means tested (individual assessment). (Coady et al seem to use 'individual assessment' to refer to assessment of individual households.) Across the sample of 122 programmes, a total of 253 targeting methods were used, illustrating how commonly multiple methods were used. Eligibility tests for cash transfers generally include some form of household assessment and often include other categorical criteria. Coady et al report that statistical analysis suggested that each additional method improved targeting performance (as measured by targeting of the poor) by about 15 percent. They note further that programmes targeting children generally performed better than those targeting the elderly in terms of reaching households in the poorest quintiles poverty-wise. This further illustrates the extent to which childhood is associated with poverty.

Geographical targeting was used in 52 of the 122 programmes examined by Coady et al, but might be less useful for small island states where there is limited variation in poverty among different areas. The 52 cases do not include those in which geographical variables are among those used in a PMT. SL-NET is one such case in that the formula predicts higher expenditure (and thus less likelihood of poverty) for households based in Castries (urban or rural), Gros Islet, Soufrière and Vieux Fort than for households in other geographical areas.

Conditions (or conditionalities) constitute a special category of eligibility test in that they require ongoing compliance in respect of particular behaviours rather than an assessment of a characteristic. Castañeda & Lindert (2005) suggest that conditions constitute a form of self-selection, but this assumes that individuals are always able to fulfil the conditions and thus can make the choice as to whether to do so.

## Combining a proxy means test with other targeting approaches

Supplementing the PMT with other criteria retains the PMT as the poverty targeting mechanism. This consideration does not address the weaknesses of a PMT and, in particular, does not address exclusion errors. This raises the question as to whether a PMT can be combined with other poverty targeting mechanisms, such as community targeting, multidimensional targeting, self-identification or direct questions about income. For example, if the PMT is flawed in targeting poverty, pairing it with another mechanism that correctly identifies households with children does not solve the weakness of the PMT. In contrast, it is possible that other methods to target poverty may be used to supplement or correct the problems in the PMT.

In ***community targeting,*** members or leaders of a local community identify households that they feel should be eligible for particular benefits. Advantages of the approach include the likelihood that local people have knowledge about households that is not easily available to outsiders and that the community may more readily accept locally made decisions as fair. Under this approach, the PMT could be applied only to those households that were first identified by the community. Disadvantages of this approach include the effort and expense required for the community process and the fact that the PMT will still incorrectly exclude some of those who are poor and should be eligible.

Alternatively, the community approach could be used after the PMT is applied, with the community (or community leaders) asked to identify households that were excluded but should nonetheless receive the benefit. It is unclear whether the community approach and PMT have, in practice, ever been used in this way. One can imagine a range of challenges in using it, such as favouritism or nepotism. It is also unclear whether the community approach has ever been used as the basis of a unified targeting mechanism rather than for

a particular programme or benefit. A general problem with community targeting is that there may be little consistency across communities, so that a household that is deemed eligible in one geographical area may be deemed ineligible in another.

In ***multidimensional targeting***, a checklist of characteristics is identified and then each household is scored against these characteristics. Given that the relevant characteristics are not derived from a regression and are not given weights according to the degree of their influence on income or expenditure, this approach could be considered a simpler—and less scientific—version of a PMT. Instead, the checklist of characteristics is based on assumptions as to what constitutes poverty. The approach has similar incorrect exclusionary characteristics as a PMT.

In ***self-identification***, the benefit is defined in a way that makes it unattractive to households that are not extremely poor, for example through payment of a very low wage on a public employment scheme. This approach does not seem appropriate for a targeting mechanism that is to be the basis of a unified system unless all benefits are of very limited value.

With ***direct questions about income***, instead of proxy questions the household or individual is asked about all income streams. Concerns about response accuracy (or lack thereof) to such questions were among the motivations for the PMT approach. Responses could be inaccurate because of the difficulty of determining a standard amount where income varies (such as informal employment). Responses could also be inaccurate because applicants do not declare all of their income in order to qualify as eligible. These are valid concerns, but the degree of possible inaccuracy needs to be compared to the exclusion and inclusion errors of the PMT. If questions about income are asked, at least some of dishonest applicants can be excluded by checking identification numbers against available databases such as government employment and tax records.

## Transparency

In some cases, countries do not disclose the variables and weights used in the PMT so as to reduce opportunities for manipulation. The World Bank further suggests that proxy variables should be changed regularly to prevent cheating. However, Kidd & Wylde point out that unless the situation being modelled has changed, this is likely to reduce the accuracy of the PMT as a predictor of poverty.

Castañeda & Lindert (2005) recognize transparency's importance because it increases the credibility of means tests and discourages fraud. Non-disclosure of variables and weights clearly reduces—or even removes—transparency. Coady et al (2004: 58) observe that while the formula-based PMT approach will be seen by some as "scientific," for others it will seem a "cold-hearted … mysterious black box" that serves to conceal what the social worker or government wants to do.

Coady et al (2004) report that in Armenia, the PMT formula is posted on the walls of welfare offices. However, qualitative research suggests that the scoring system, which was presented in mathematical notation, was not well understood despite extremely high literacy levels. In Chile, the formula was public until 1987. The formula was then changed and kept confidential, in part because it was deemed too difficult to explain and in part to discourage bribing of officials.

Coady et al suggest that programmes may publicly post a beneficiary list not only to promote transparency, but also to discourage the non-poor—who fear the stigma of being identified as a beneficiary—from applying. They suggest that stigma "is quite a blunt tool insofar as it can discourage participation among the poor and work against the promotion of dignity and self-worth as an outcome of development" (Coady et al, 2004: 76).

Provision for appeals is another aspect of transparency. Of the systems reviewed by Castañeda & Lindert (2005), only the United States had a well-established and effective formal appeals process. Other countries either did not have an appeals mechanism, had one that was not formalized or had one that was not followed in practice. In effect, where eligibility is determined on the basis of a PMT, the leeway for appeals is likely to be small, especially if the formula

used is not made public. However, if the system is rights-based, an appeals mechanism would be a necessary component, and would need to be able to override the PMT equation in cases where applicants are deemed to have been unfairly excluded.

# CONCLUSION AND RECOMMENDATONS

Castañeda & Lindert (2005: 1) claim that the objectives of targeting are consistent with "universal coverage of the poor," which is becoming a goal for social protection programmes in some developing countries. Their observation relates to targeting in general, rather than only to PMTs. The discussion above highlights the challenges of a PMT facilitating the achievement of this goal in that exclusion errors and other issues effectively undermine coverage of all poor households.

Sharif (2009) concludes his investigation by stating that a PMT may not be appropriate for targeting the extreme poor in Bangladesh, given the limited budget, the large errors when the cut-off is low, the ageing of available data available and the tendency of PMTs to exclude certain categories of households (such as those containing two old persons living with their grandchild). Though the Eastern Caribbean states are at a higher level of development than Bangladesh, some of these cautions might nevertheless be relevant, particularly the limitations of PMTs if the assistance programme's budget is restricted or the potential for systematic bias against certain household forms.

Most, if not all, PMTs aim to predict the expenditure of households rather than the expenditure of individuals. This focus is based on an implicit assumption that expenditures benefit all household members evenly (or according to need). This creates a challenge when a social protection benefit is intended to assist particular individuals (such as the elderly or children), but social and family norms give other individuals more control over expenditure decisions.

The reliability of the equation on which the PMT is based depends in large part on the reliability of the underlying survey. A further challenge to using a PMT for targeting is that it requires regularly conducting surveys in order to ensure that the PMT is up to date. This, in turn, will require substantial financial, human and other resources that add to the other costs associated with implementing a PMT. Wietzke (2011) proposes for Belize that an intercensal needs survey would be necessary at least every five years, if not more often. He estimates that the cost of such a survey would be between BZ$ 172,000 and BZ$ 720,000.

When PMT equations are tested, even on the same data set from which they are derived, the degree of inaccuracy tends to be particularly severe for indigence. Inclusion errors are especially severe at low cut-off points because of the clustering of incomes at the bottom end of the range. This reduces the effectiveness of PMT for benefits that aim to target a very small proportion of the population.

The poverty lines derived for the Surveys of Living Conditions conducted in Eastern Caribbean countries in the last decade use sex-plus-age adjusted equivalence scales. Differentiation by sex in equivalence scales has been criticized by leading practitioners, and the Eastern Caribbean has this differentiation starting at the surprisingly young age of one year. PMTs based on this approach will discriminate against women and girls by giving less weight to female than male poverty.

In terms of household forms, the literature does not describe any examples of PMTs that include variables that take into account the particular income and time poverty-related challenges faced by households where there is a single adult who is responsible not only for income-earning, but also for the provision of care for children or other household members.

A strong attraction of PMT is that it should produce the same result regardless of who administers it. One

disadvantage of this characteristic is its rigidity, in that it does not allow for the specific characteristics of a particular household to be taken into account. In addition, there is a "built-in error: the formula is designed to be right on average but will not correctly categorize every household" (Coady et al, 2004: 54). In statistical terms, there is a large standard error in the prediction for a particular household. The theoretical error is likely to be exacerbated in practice by incompletely reaching those who might be eligible.

Kidd & Wylde (2011) express concern as to whether those promoting the use of the PMT sufficiently understand what the approach can and cannot do. They suggest that policymakers should try to find ways of targeting that are easier to implement and less socially divisive. They note increasing interest for social protection that target resources to vulnerable groups such as the elderly, children, people with disability and the unemployed on a universal basis. They thus advocate for categorical targeting that does not include means tests.

It is disappointing that most of the literature is silent about rights. If accessing the programmes that will use the PMT is a right, then the fact that the formula has an inbuilt error is of serious concern because some citizens will be denied their rights on the basis of a formula that aims to reach the 'average' household or individual but does so with a degree of error. In addition, this paper has highlighted some instances where different approaches may introduce bias in favour for or against women, children and the elderly. Policymakers need to be aware of these potential biases so that they can make an informed decision as to which biases are desired and which not.

The following recommendations are put forward for those Eastern Caribbean countries that decide to use a PMT for targeting of benefits. These recommendations include some that should increase the child- and gender-sensitivity of the targeting:

- Combining the PMT with other targeting methods can be considered, though doing so is unlikely to overcome all the challenges associated with PMTs;
- Approaches that take into account the income forgone by mothers and others with heavy care responsibilities should be considered;
- An alternative equivalence scale that does not differentiate on the basis of sex of household members should be used. The weighting for children should take into consideration expert opinion that this weight should be higher for middle- and higher-income countries than for poorer countries, and that childhood deprivation has long-lasting impacts on individuals, families and the country as a whole;
- A full costing of the roll-out of the PMT should be undertaken prior to its use, in order to ascertain the full financial, human resource and logistical implications; and
- Administrative justice requires that an appeals mechanism be in place for the targeting mechanism. To complement this mechanism, the responsible government agency could assess the cases of all rejected applicants on a periodic basis so as to avoid exclusion of vulnerable individuals who may not have the confidence or capacity to initiate an appeal. The appeals mechanism should consider the applicant's objective situation in order to determine whether the exclusion was fair.

# REFERENCES


***Alatas V, Banerjee A, Hanna R, Olken BA & Tobias J.*** 2010. Targeting the Poor: Evidence from a field experiment in Indonesia. Working Paper 15980. National Bureau of Economic Research: Cambridge MA.

***Andaiye.*** 2003. "Smoke and Mirrors: The Illusion of CARICOM Women's Growing Economic Empowerment, Post-Beijing" in Nain GT & Bailey B (eds) Gender Equality in the Caribbean: Reality or Illusion. CARICOM Secretariat: Kingston: 73-107.

***Blank L.*** 2009. St Lucia Social Safety Net Assessment. Prepared for Government of St Lucia.

***Blank L.*** 2010. St Kitts & Nevis Social Safety Net Assessment. Prepared for Government of St Kitts & Nevis.

***Brazilian Institute for Geography and Statistics, Economic Commission for Latin America and the Caribbean and National Statistical Institute of Portugal.*** 1999. Equivalence Scales: A Brief Review of Concepts and Methods. Third Meeting of the Expert Group on Poverty Statistics (Rio Group), Lisbon, 22-24 November, 1999.

***Buhmann B, Rainwater L, Schmaus G & Smeeding TM.*** 1988. "Equivalence Scales, Well-being, Inequality, and Poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database" in Review of Income and Wealth 34(2): 115-142.

***Cadette, A.*** April 2012. Saint Lucia National Eligibility Test SL-NET. Ministry of Social Transformation, Local Government and Community Empowerment: St Lucia.

***Camacho A & Conover E.*** 2010. Manipulation of Social Program Eligibility: Detection, explanations and consequences for empirical research, downloaded www.aeaweb.org/aea/2011conference/program/retrieve.php?pdfid=266, 31 March 2014.

***Castañeda T & Lindert K.*** June 2005. Designing and Implementing Household Targeting Systems: Lessons from Latin American and the United States. Social Protection Discussion Paper no. 0526. World Bank: Washington DC.

***Coady D, Grosh M & Hoddinott J.*** 2004. Targeting of Transfers in Developing Countries: Review of Lessons and Experience. World Bank: Washington DC.

***Deaton A & Zaidi Salman.*** Undated. Guidelines for Constructing Consumption Aggregates for Welfare Analysis.

***Deaton A.*** 1997. The Analysis of Household Surveys: A microeconometric approach to development policy. Published for the World Bank. John Hopkins University Press: Baltimore and London.

***Deaton AS & Muellbauer J.*** 1986. On Measuring Child Costs: With Applications to Poor Countries. Journal of Political Economy 94(4): 720-744.

***Diaz BY with Narvaez L.*** November 2014. 'Codebook and Questionnaire Collection Manual: Grenadian Living Conditions Indicator Survey'. Support for Education, Empowerment and Development Program. Government of Grenada.

***Diaz Y, Espinoza FA & Markaki Y.*** 3 February 2014. Targeting Grenada's most deprived population: Towards a multidimensional living conditions assessment. Draft.

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

***Fernandez L.*** 2008a. Proxy Means Test Model to Select Poor Households for Social Programs in the Commonwealth of Dominica. Final Technical Note.

***Fernandez L.*** 2008b. Development of a Beneficiary Identification System and Beneficiary Registry in Dominica. [powerpoint]

***Folbre N.*** 2008. Valuing Children: Rethinking the economics of the family. Harvard University Press: Cambridge, Massachusetts and London.

***Grosh ME & Baker JL.*** July 1994. Proxy Means Tests for Targeting Social Programs: Simulations and speculation. Living Standards Measurement Study Working Paper No 118. World Bank: Washington DC.

***IOS Partners/ London School of Economics.*** May 2011. Report on the Design of the Beneficiary Targeting Index Module: Design and Implementation of the Single Information System of Beneficiaries for Belize.

***Johannsen J.*** October 2006. Operational Poverty Targeting in Peru – proxy means testing with non-income indicators. International Poverty Centre Working Paper No 30. Brasilia.

***Kairi Consultants Limited.*** 2007. Trade Adjustment and Poverty in St Lucia – 2005/06: Volume I: Main Report. Trinidad and Tobago.

***Kairi Consultants Limited.*** 2007b. Trade Adjustment and Poverty in St Lucia – 2005/06: Volume III: Quantitative assessment. Trinidad and Tobago.

***Kidd S & Wylde E.*** September 2011. Targeting the Poorest: An assessment of the proxy means test methodology. Australian Agency for International Development: Canberra.

***Meth C.*** 2014. "Ending can only be bad for SA's 'good story' on statistics". Business Day, 11 April 2014.

***Narayan A & Yoshida N.*** July 2005. Proxy Means Test for Targeting Welfare Benefits in Sri Lanka. Report No. SASPR-7. World Bank South-East Asia.

***Organization for Economic Cooperation and Development.*** 2013. OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth. OECD Publishing: Paris.

***Organization for Economic Cooperation and Development.*** Undated. What are equivalence scales? OECD Project on Income Distribution and Proverty, www.oecd.org/social/inequality.htm

***Sabates-Wheeler R, Hurrell A & Devereux S.*** February 2014. Targeting social transfer programmes: Comparing design and implementation errors across alternative mechanisms. WIDER Working Paper 2014/040. World Institute for Development Economics Research: Helsinki.

***Sharif IA***. August 2009. Building a Targeting System for Bangladesh based on Proxy Means Testing. SP Discussion Paper No 0914. World Bank.

***St Catherine E***. 2004. Some Quantitative Methods and Practices of Poverty Measurements and Poverty Statistics in the Eastern Caribbean. Presented at Workshop on Poverty Statistics in Latin America and the Caribbean, Rio de Janeiro, Brazil, 10-14 May 2004.

***St Lucia Survey of Living Conditions and Household Budgets*** 2005. Diary Schedule of Spenders 18 years of age and Over or Employed.

***Tang G.*** Poverty Measurement in the Caribbean: A Critique of Current Approaches from a Gender Perspective.

***Wietzke F-B.*** 2011. Report on the construction of the proxy means indicator, Belize MHDST. IOS Partners/ London School of Economics.

***Woolard I & Leibbrandt M.*** October 1999. Measuring Poverty in South Africa. Working paper no 99/33. Development Policy Research Unit, University of Cape Town: Cape Town.

UNICEF Office for the Eastern Caribbean Area
First Floor | UN House | Marine Gardens
Christ Church | Barbados
Tel: 246 467 6000
Fax: 246 436 3812
W: www.unicef.org/easterncaribbean
S.M.: www.facebook.com/UNICEFeasterncaribbean

UN Women Multi-Country Office - Caribbean
First Floor | UN House | Marine Gardens
Christ Church | Barbados
Tel: 246 467 6000
Fax: 246 437 6596
W: http://caribbean.unwomen.org/
S.M.: www.facebook.com/unwomencaribbean