

SOUTH ASIA REGION

PREM WORKING PAPER SERIES



Proxy Means Test for Targeting Welfare Benefits in Sri Lanka

Ambar Narayan and Nobuo Yoshida

July 2005

Report No. SASPR-7



A WORLD BANK DOCUMENT

The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent.

Public Disclosure Authorized

Public Disclosure Authorized

Public Disclosure Authorized

Public Disclosure Authorized

About the SASPR Working Paper

The purpose of the SASPR Working Paper Series is to provide a quick outlet for sharing more broadly research/analysis of issues related to development in South Asia. Although the primary source of such research/analysis is SASPR staff, other contributors are most welcome to use this outlet for rapid publication of their research that is relevant to South Asia's development. The papers are informal in nature and basically represent views/analysis of the concerned author(s). All papers submitted for publication are sent for an outside review to assure quality. I provide only a very light editorial touch. For enquiries about submission of papers for publication in the series or for copies of published papers, please contact Naomi Dass (telephone number 202-458-0335).

Sadiq Ahmed
Sector Director
South Asia Poverty Reduction and Economic Management
World Bank, Washington D.C.

Preface and Acknowledgements

This paper was prepared (during FYs 2003-2004) as a part of the task “Technical Assistance for Welfare Reform for Sri Lanka” that the World Bank has been engaged in since 2003. The task is led by Tara Vishwanath (SASPR), with Ambar Narayan, Princess Ventura, Nobuo Yoshida (SASPR), Francisco Ayala, Yoko Kijima, Hernando Quintero (consultants), and S. Shivakumaran (Welfare Benefits Board) as team members. One of the recommendations for targeting formula made in this paper has been evaluated using the data collected during the targeting pilot conducted in June-August, 2003 by the Welfare Benefits Board (Government of Sri Lanka). This formula has also been accepted as the method for targeting Samurdhi transfers in the North and East of Sri Lanka, for which implementation efforts are ongoing as of June 2005. The results of this paper and the complementary analysis of pilot data were presented at the policy workshop organized by Welfare Benefits Board (Colombo, November, 2003). The paper has also provided significant inputs into the Poverty and Social Impact Analysis (PSIA) on Welfare Reforms in Sri Lanka, which has just been completed.

The authors are grateful for comments and suggestions from a number of individuals during the course of the work for this paper. In particular, acknowledgements are due to Tara Vishwanath, Margaret Grosh (HDNSP), Francisco Ayala, and S. Shivakumaran. The analysis also benefited from discussions with members of the Steering Committee (Government of Sri Lanka) set up to guide the technical work for Welfare Reform, and discussions with a broad range of participants at the policy workshop organized by the Welfare Benefits Board. The authors bear all responsibility for remaining errors and omissions.

Table of Contents

Introduction	1
I. Proxy Means Test: Rationale and Methodology.....	2
II. Developing a Proxy Mean Testing Formula (PMTF) for Sri Lanka	3
III. Deriving a Model: Results from OLS Regressions and Simulations.....	7
IV. Targeting Errors and Choice of Cutoff Points	13
V. Comparisons with Alternative Methods of Estimation.....	18
VI. Simulations to Measure Welfare Improvements	20
Conclusion.....	25
References	28
Figures	29
Appendix	i

Introduction

The Welfare system of Sri Lanka during recent years – of which the Samurdhi program is the most significant component – has suffered from serious design flaws that have led to considerable mis-targeting. A recent evaluation of Samurdhi suggests that its targeted foodstamp program, which constitutes 80 percent of the total program budget, misses about 40 percent of households ranked in the poorest quintile, while almost 44 percent of the total budget is spent on households from the top 3 quintiles. Qualitative results suggest that political factors, including party affiliation or voting preferences influence allocation of Samurdhi grants. Large-scale leakage of benefits has led to the program covering as much as half of the population – far above the poverty rate, estimated at less than 25 percent – with the result that the benefits are spread too thinly and the small size of transfers has little impact on poverty.¹

Responding to the need for reform, the new Welfare Benefit Act was enacted by the Parliament in July 2002 to rationalize the legal and institutional framework of all social welfare programs, and to improve the targeting performance of the Samurdhi foodstamp program in particular. On the governance side, the Act mandates an independent Welfare Benefits Board and lower level Selection Committees to set eligibility criteria, validate entry and exit into the program, and redress appeals. On the technical aspect of identifying beneficiaries, the Act envisages setting objective criteria for selection of beneficiaries, with a longer-term objective of integrating the criteria to cover *all* welfare programs in the country. If implemented effectively, these reforms are likely to enhance objectivity and transparency, thereby minimizing the scope for political interference in the selection process and increasing targeting efficiency.²

A key element of this plan for implementation involves developing objective measures, based on easily observable and verifiable indicators, to identify potential beneficiaries of the program. In this context, the *objective of this note* is to analyze viable options for such an objective measure, namely a *means-testing formula*, using the household data from the Sri Lanka Integrated Survey. The results of this analysis will be presented to the Steering Committee appointed by the Government of Sri Lanka for implementing the Welfare Reform, to help decide on a targeting formula to be tested through a pilot targeting exercise that is planned in the near future.³

Section I of this note will briefly discuss the rationale and methodology for proxy means tests, placing them in the context of some of the relevant economic literature on the topic. Section II will describe the data and methodology for developing the targeting formula for the Sri Lankan case. Section III will present results and identify the models considerable best suited for the purpose. Section IV will analyze the targeting outcomes derived from simulations with the models and relate these to possible choices for eligibility thresholds. Section V will compare the selected models with some alternative methods for estimating the model. Section VI will derive and compare welfare gains for a few examples of payment schedules used to distribute benefits to the eligible groups identified by the formulas, and Section VII will conclude by summarizing some of the important conclusions from the exercise.

¹ See Glinskaya (2000) for a detailed evaluation of the Samurdhi program.

² To oversee implementation of reforms in accordance with the Act, a Steering Committee has been set up constituting senior officials from key departments.

³ A simultaneous exercise conducted in Sri Lanka by a team comprising of local statisticians will develop alternate versions means-testing formulae using a different data source, the Consumer Finance Socio-Economic Survey. Along with the options presented here, these formulae are also candidates for testing through the pilot. The pilot exercise will test and validate the formulae, application form(s) to collect the information, and institutional and database capacity is planned for June 03, covering around 50,000 households spread over 4 representative geographical regions within the country.

I. Proxy Means Test: Rationale and Methodology⁴

Targeting benefits to the poor first requires a precise definition of the target group. Once the target group is established, a methodology must be found for identifying individuals or households that are in that group and for excluding those who are not. For instance, if the poor are identified as a target group for a program, one must be able to make a precise judgment about the level of welfare or the means of the recipient.

In principle, conducting a *means test* that correctly measures the earnings of a household is the best way to determine eligibility when the poor are the target group, as is the case with Samurdhi. In practice, however, such straightforward means tests suffer from several problems. First, applicants have an incentive to understate their welfare level, and verifying that information is difficult in developing countries where reliable records typically do not exist. Second, income is also considered an imperfect measure of welfare in developing countries, since it is unlikely to measure accurately imputed value of own-produced goods, gifts and transfers, or owner-occupied housing. Incomes of the poor in developing countries are also often subject to high volatility due to factors ranging from seasonality of agriculture and sporadic nature of employment in the informal sector. Since adjustments for such volatility are hard to make in practice, actual welfare from income measures are likely to be highly distorted. In the light of these difficulties, rigorous means tests are largely reserved for industrialized economies where a well-educated labor force is concentrated in jobs in which cash is paid regularly and payments are reported to tax or welfare authorities. Where means-testing is used in developing countries, it is greatly simplified, at a considerable cost to accuracy.⁵

Given the administrative difficulties associated with sophisticated means tests and the inaccuracy of simple means tests, the idea of using proxy means tests that avoid the problems involved in relying on reported income is appealing. Proxy means test involves using information on household or individual characteristics correlated with welfare levels in a formal algorithm to proxy household income or welfare. These instruments are selected based on their ability to predict welfare as measured by, for example, consumption expenditure of households. The obvious advantage of proxy means testing is that good predictors of welfare – like demographic data, characteristics of dwelling units and ownership of durable assets – are likely easier to collect and verify than are direct measures like consumption or income. The efficacy of proxy means testing is indicated by a recent comparative study of targeting in Latin America (Grosh, 1994), which has found that, among all targeting mechanisms, proxy means tests tend to produce the best incidence outcomes in developing countries.

Academic evidence and practical experience with Proxy Means Tests

A number of simulations in academic papers by various authors show how proxy means test could work, and the welfare gains likely produced by implementing such a targeting system. Haddad, Sullivan and Kennedy (1991) used household survey data from Ghana, the Philippines, Mexico and Brazil to show that some variables that would be very simple to collect could serve as good proxies for the measures of caloric adequacy that are usually used as the standard measures of food and nutrition security, which are harder to collect as they rely on the memory of individuals and on the anthropometric indicators of pre-school children. Glewwe and Kanaan

⁴ This section draws extensively from Grosh and Baker (1995) and Grosh (1994).

⁵ Simple means tests are performed as part of the food stamp programs in Jamaica (prior to 2002), Honduras, Sri Lanka and Zambia. In Jamaica and Sri Lanka, this evaluation has been largely subjective and does not contain any systematic examination or weighting of certain factors. Evaluations reveal that the two programs delivered only 56 and 57 percent of its benefits respectively to those in the poorest 40 percent of the population.

(1989) have used regression analysis on a data from Cote d'Ivoire to predict welfare levels based on several combinations of variables that are fairly easy to measure. The paper demonstrated that simple regression predictions could improve targeting markedly over untargeted transfers.⁶ In a recent study, Grosh and Glinskaya (1997) used regression analysis with data from Armenia to show how the targeting outcomes of a current cash transfer program can be improved by using a suitable proxy mean test formula.

Grosh and Baker (1995) carries out simulations on Living Standards Measurement Survey data sets from Jamaica, Bolivia and Peru to explore what kind of information can best be used in a proxy means test and how accurate such tests might be expected to be. Their results show that more information is generally better than less for a targeting formula, though there are diminishing returns. The proxy systems all have significant undercoverage, but they cut down leakage so much that the impact on poverty is better with imperfect targeting than with none.

While academic exercises have been useful in developing such a proxy mean test system, more insights on the implementation of such programs can be gained by looking at actual experiences on the ground, in Chile where it has existed since 1980, and more recent programs in Costa Rica, Colombia and Jamaica. The Ficha CAS in Chile uses a form filled out by a social worker that collects information on household characteristics such as location, housing quality, household composition and education and the work done by the household members. Scores are then assigned using a complicated algorithm and then used to determine eligibility for two large cash transfer programs and for water and housing subsidies, and if so, the level of subsidy.

II. Developing a Proxy Mean Testing Formula (PMTF) for Sri Lanka

Data Selected for the Analysis

The data used for this exercise is the Sri Lanka Integrated Survey (SLIS), conducted by the World Bank in collaboration with local institutions in 1999-2000. This is a multitopic household survey in the style of an LSMS, with modules on consumption, income, employment, health, nutrition, fertility, education, and living conditions. It also includes information on benefits received from existing welfare programs, including Samurdhi, and a detailed community module. The SLIS covered around 7500 households drawn from all regions of Sri Lanka, including the areas under conflict during those times (the North-East) and was designed to be representative at the national and provincial levels.⁷ However, the results from the North-East sample were found by many to be inconsistent with the conventional wisdom about the region. In the absence of any other study for the region for the last ten years that can serve as a sound basis for comparison, it is impossible to determine one way or other the veracity of the SLIS North-East findings. But given the controversy surrounding data from this region, and the fact that fieldwork was disrupted by the prevailing conflict conditions, it was considered best to conduct the exercise *excluding* the North-East sample of the SLIS.

The exclusion of the North-East left us with around 5600 households for the rest of Sri Lanka. Taking into account the sample weights, the North-East sample amounted to about 12 percent of the total sample, which is consistent with this region's estimated share in the country's population. Using the sampling weights, the residual sample by design is representative for the

⁶ Glewwe (1990) took the same basic approach of predicting welfare. Instead of using regressions, he solved a poverty minimization problem to derive weights for each household variable. While theoretically more appropriate, the poverty minimization technique is much more difficult to compute, and produces results not dissimilar from those based on regression analysis.

⁷ The sampling strategy of the SLIS was developed by the Department of Census and Statistics in Sri Lanka.

entire country excluding the North-East. Exclusion of the North-East, while unfortunate, should be seen in the context that no survey has managed to cover the region during the past decade of conflict – in other words, the problem is certainly not unique to the SLIS.

On all other accounts, the SLIS appears to be well suited to the purpose of developing the PMTF for Sri Lanka. It has rich and detailed information on *most correlates of poverty* (more so than any other household survey in Sri Lanka), along with information on the *benefits received from the existing Samurdhi program*, which makes it amenable for comparisons between the existing program and the proposed formulas on targeting efficiency and welfare implications. The SLIS is also the *most recent source of representative household data* available for Sri Lanka. On the downside, it is not clear yet whether the SLIS will be sustained in future – a decision that will hinge on whether it becomes a part of the Government of Sri Lanka's poverty and social monitoring system. If it is not repeated, future updating of the formula using new information will become a problem. All these factors (along with the results presented here) must be taken into account before deciding in favor of adopting either version of the PMTF – the one developed here using SLIS, or the one being developed meanwhile in Sri Lanka using CFSES (1996-97).

Selecting an indicator for actual household welfare

We choose per capita household consumption expenditure (monthly) as the welfare measure that would be proxied by a set of easily observable indicators. This includes all expenditures on non-durables, the imputed value of non-durables received as gifts or produced in the household, and the imputed value of owner-occupied housing; it excludes expenditures on durable goods and assets.⁸

In development literature, consumption expenditure is generally considered a more accurate measure of welfare than income for several reasons. First, because consumption expenditures tend to be less variable than income over seasons, it is more likely to indicate the household's "true" economic status, as a result of households with sporadic incomes smoothing their consumption patterns over time. Second, in practice, consumption is generally measured with far greater accuracy than income in a household survey, primarily because households' sources of income may include home-based production, own farms and businesses. Calculating the flow of *net* incomes from these sources turn out to be a big problem since the flow of costs and returns from these activities are often inaccurately reported by households.

Predicting welfare: the choice of Ordinary Least Squares (OLS)

To predict welfare, the consumption variable is regressed, using OLS method, on different sets of explanatory variables. The case for using OLS as the model for predicting welfare is driven primarily by convenience and ease of interpretation. The first problem with using an OLS model is that many of the explanatory variables are likely to be endogenous to (and thus not independent explanators of) household welfare. This problem is however of less concern to us, since our objective is solely to *identify* the poor and not to explain the *reasons* for their poverty. Second, Grosh and Baker (1995) points out that strictly speaking, OLS is inappropriate for predicting poverty since the technique minimizes the squared errors between the "true" and the predicted levels of welfare, which is a different theoretical problem from that of minimization of poverty. That said, OLS is considered convenient and useful by these authors when a large numbers of

⁸ Although the consumption expenditure aggregate should include imputed value of flow of services from durable goods, in practice this is very hard to calculate. Thus the convention followed here, as in the case of data from many other countries, is to exclude all expenditures on durable goods. In this analysis, the actual variable used as the indicator for welfare is the *natural log* of per capita consumption expenditure (monthly), which is found to work well in the regressions.

predictor variables, including continuous variables, are available.⁹ Moreover, using OLS has the advantage of being able to intuitively interpret the coefficients of the predictors on welfare – a feature that is likely to appeal to a policymaker and more amenable to achieving political consensus in the country.

Predicting welfare: the choice of variables

Selection of variables to predict welfare as measured by per capita consumption should take into account two separate criteria: *correlation* between the welfare measure and the predictor, which will determine accuracy of the prediction, and *verifiability* of the predictor, which will determine the accuracy of information used to impute welfare. The types of predictors used for this exercise, discussed below, were arrived at after judging all possible predictors on the basis of these two criteria.

Location variables are obviously the most easily verifiable, and the same is true for *characteristics of the community*, when it is defined in simple terms like the presence of a bank or administrative offices. *Housing quality* may also be easily verified by a social worker visiting the home. *Household characteristics*, such as the number of members and dependents, and age, education and occupation of the household head, are less easy to verify. However, it is generally felt that these information, firstly, are not overly difficult to verify, and secondly, that households are less likely to misrepresent such information. Using program officers who live in the same community as the applicant households to collect the information – as is envisaged for Sri Lanka – also makes it more likely that such information will be reported correctly.

Ownership of durable goods or farm equipment is verifiable by inspection – however they can be misrepresented by the household removing the goods from the home during an expected visit by the social worker, which is easier to do with small or mobile items than for items such as stoves or refrigerators. The general presumption in the literature is also that people are more willing to lie about ownership of such items than they are about household characteristics. However, these variables tend to have high predictive power for welfare, and therefore including them can reduce mis-targeting substantially.

Ownership of productive assets is again not easy to verify. The presence of livestock is verifiable to some extent. As for land ownership, while it may not be measured perfectly, one can reasonably expect that program officers who belong to the community will have local knowledge about whether a household owns a large amount of land or not, which will deter misrepresentation. The fact that these variables are likely to have high correlations with poverty in rural areas makes a strong case for including them as predictors of welfare.

Very briefly, the steps in the procedure for arriving at the PMTF run as follows. The original set of variables belonging to the six broad categories is identified based on the two criteria mentioned above. Dichotomous variables are then created for some of the continuous variables in order to identify those characteristics that discriminate between poor and rich households. The set of selected predictors are then introduced in a weighted OLS regression of (log of) per capita monthly consumption expenditure.¹⁰ Different subsets of variables are checked for possible multicollinearity, and a few variables are adjusted or dropped as necessary to reduce such

⁹ An algorithm that does solve the problem of minimizing poverty is found in Ravallion and Chao (1989), and could be a better tool for designing a transfer scheme than the OLS model. However this algorithm is very difficult to use when a large number of predictive variables are available, and is difficult to interpret for policymakers. See Grosh and Baker, Annex I for a fuller discussion.

¹⁰ In order for the results to be representative for the entire population (excluding the North-East), all OLS regressions are weighted, where the weight of each observation is equal to the product of the household specific sampling weight and the size of the household.

problems. A stepwise regression is then used with the remaining set of variables because it is designed to eliminate from the regression variables that are not statistically significant and do not increase the model's overall explanatory power. From this process, different models (described in detail later) evolve based on the subset of variables entering into the regression.

Determining Eligibility

Each model predicts a certain level of welfare, as measured by (log of) per capita monthly consumption expenditure. These predicted welfare levels are used to assign individuals to eligible or ineligible groups, based on an eligibility cutoff point. While an obvious choice for this cutoff point is the poverty line applicable to this data for Sri Lanka, this is not an option since no poverty line has been calculated using SLIS data till date.¹¹ Instead, we define the eligibility cutoff point by the welfare level of a certain percentile of the individual welfare distribution, using “true” welfare. Since it is not known yet what percent of the population the program would target when it is implemented, we consider a range of cutoff points, defined by specific percentiles of actual/true per capita consumption expenditures (e.g. 25th, 30th, 40th).¹²

The selection of the cutoff point is essentially a policy, and not a technical decision. By simulating a wide range of scenarios corresponding to different cutoff points for each model, we seek to achieve two objectives. Firstly, the exercise will show the sensitivity of the model and its attendant errors in targeting to changes in cutoff points. Second, the simulations will help the government make a policy decision on what the cutoff point should be, taking into account the tradeoffs inherent in choosing a relatively higher cutoff vis-à-vis a low one.

As a reference, it is useful to note as a useful reference point that the head-count poverty rate for Sri Lanka was estimated at 25 percent, using HIES data in 1995-96. One must however keep in mind that there is no way to compare the 25th percentile from SLIS data to the poverty line computed from HIES, and the former can, at best serve as an *indicative* poverty line for the SLIS.

Evaluating the targeting formulae

As with all regression analyses, different specifications of the model and different samples of the population yield different results and it is not always easy to say which specification is superior. However, a variety of tests can be conducted, which, taken together, can be used to select one model over another. We use two types of criteria to evaluate alternate options for the PMTF. The first criterion is the regression's R^2 , which is the proportion of the variation in consumption that is explained by the regression model. Higher the R^2 , the better are a particular set of variables in predicting welfare.

The second criterion involves looking at measures that indicate the ability of various models to identify the poor properly. No matter what model is used, given that it can predict welfare only with some imperfection, it is likely that some truly eligible people will be left out, while others who are not eligible will benefit. Following Grosh and Baker (1995) and related literature for other countries, we evaluate targeting accuracy of alternate models using Type I and II errors, from which rates of undercoverage and leakage are derived, and incidence of benefits across income/consumption groups. Individuals are categorized in four groups according to whether their *true* and *predicted* (by the regression model) welfare levels fall above or below the defined

¹¹ The SLIS consumption module is somewhat different (in terms of itemization of food products) from that in the HIES (1985/86 and 1990/91), so that the consumption baskets are not perfectly comparable across these surveys. Since existing poverty lines for Sri Lanka have been based on the baskets developed from these HIES, these should not be used for the SLIS sample.

¹² An eligibility cutoff point of percentile X is defined as the per capita actual monthly consumption expenditure of a household *i*, such that X percent of the total population in the sample have per capita monthly expenditures less than that of household *i*

eligibility cutoff point. Those whose true welfare falls below the eligibility threshold constitute the “target” group, while those with predicted welfare below the eligibility threshold constitute the “eligible” group. Individuals whose true and predicted welfare measures put them on the same side of the cutoff line are targeting “successes”.

A person who is incorrectly excluded by the formula is a case of *Type I error* (see Table 1). Conversely, a person incorrectly identified as being eligible constitutes a case of *Type II error* (Table 1). *Undercoverage* is calculated by dividing the number of cases of Type I error by the total number of individuals who should get benefits $[e_1/n_1]$. *Leakage* is calculated by dividing the number in the Type II error category by the number of persons served by the program $[e_2/m_1]$. Undercoverage reduces the impact of the program on the welfare level of the intended beneficiaries, but carries no budgetary cost. Leakage, on the other hand, has no effect on the welfare impact of the program on the intended beneficiaries, but increases program costs.

Table 1: Illustration of Type I and II errors

	<i>Target group</i>	<i>Non-target group</i>	<i>Total</i>
<i>Eligible: predicted by PMTF</i>	Targeting Success (s_1)	Type II error (e_2)	m_1
<i>Ineligible: predicted by PMTF</i>	Type I error (e_1)	Targeting Success (s_2)	m_2
<i>Total</i>	n_1	n_2	n

While it would be preferable to have low levels of leakage *and* undercoverage, in reality one may face tradeoffs between these two objectives. In general, the higher the priority assigned to raising the welfare of the poor, the more important it is to eliminate undercoverage. Conversely, if saving program costs is a higher priority, it is more important to minimize leakage. Lowering leakage, besides being cost-efficient, can also be welfare increasing in the presence of a budget constraint – lower the leakage of benefits to ineligible individuals, higher would be the amount available for transfers to those who are eligible.

The last criterion to evaluate targeting efficiency is by looking at how a specific PMTF allocates potential beneficiaries across the expenditure distribution. It is preferred that a model has good *incidence*, i.e. most of the identified beneficiaries belong to the bottom of the consumption (income) distribution, and relatively few, if any, from the top of the distribution.

III. Deriving a Model: Results from OLS Regressions and Simulations

Simulations with the basic set of models

Table 1 below summarizes some key results – R^2 , undercoverage rate and leakage rate – for a set of models. All models are OLS regressions of (log of) per capita monthly consumption measured in Sri Lankan rupees on a set of predictors.¹³ For all models, stepwise regressions are used to eliminate insignificant variables, and retain only those whose statistical significance is above a prescribed limit (equal to 80 percent, unless otherwise specified).

Model 1: Contains the full set of predictors. These include selected variables for location (province, rural/urban); community characteristics (presence of a bank or divisional headquarters in the community); household assets (consumer non-durables, farm equipment); household’s ownership of land and livestock; characteristics of household head (age, education, main activity, marital status); household demographics (household size, number of

¹³ The consumption measure used here *includes* the transfers received by the household from the government, in various forms that include Samurdhi foodstamps as well as benefits from other programs. For more discussion on the reasons for choosing this measure of consumption see Section II, Appendix.

dependents, whether children attend school); housing characteristics (owned housing or not, type of floor, wall and latrine, number of rooms) and value of Samurdhi foodstamp received.

Model 2: *dropping* variable for value of Samurdhi foodstamp received from Model 1

Model 3: *dropping* variable for value of Samurdhi foodstamp received and province dummies from Model 1

Model 4: *dropping* variable for value of Samurdhi foodstamp received, province dummies and land ownership from Model 1

Model 5: *retaining* only community characteristics, housing characteristics and location (rural/urban) from Model 3

Model 6: *dropping* community characteristics from Model 3

Model 7: *retaining* only those variables from Model 3 whose coefficients have very high statistical significance (99 percent level and above)

While a large number of models with different combinations of predictors have been tried out, the above are enough to present a logical story of the results. Model 1 represents the entire set of variables that provide the “best fit” for per capita consumption with this data (apparent from the R-square, which is the highest for this model), as long as the list of variables are limited to those that can be measured and verified with some degree of precision. Adding more variables to this set – it has been observed – leads to almost no improvement to the R-square, as well as undercoverage and leakage rates.

Table 2: Results from different models

Models	Undercoverage Rate for different cutoff percentiles				Leakage Rate for different cutoff percentiles				R-square
	25	30	35	40	25	30	35	40	
1	0.47	0.41	0.35	0.29	0.37	0.35	0.32	0.29	0.59
2	0.51	0.42	0.37	0.30	0.39	0.35	0.33	0.29	0.58
3	0.52	0.43	0.37	0.28	0.39	0.35	0.33	0.30	0.56
4	0.53	0.45	0.37	0.28	0.40	0.36	0.34	0.31	0.56
5	0.72	0.61	0.46	0.34	0.43	0.41	0.37	0.35	0.42
6	0.52	0.43	0.37	0.30	0.39	0.35	0.33	0.31	0.56
7*	0.52	0.43	0.37	0.28	0.39	0.36	0.33	0.31	0.56

Note: (1) Undercoverage and leakage rates are calculated for the “poverty line” being equal to the eligibility cutoff threshold in every case¹⁴

(2) The 25th, 30th, 35th, and 40th percentiles of actual consumption amount to Rs. 1129, 1201, 1270 and 1347 monthly per capita (at 2000 prices) respectively

Although Model 1 achieves the best results, it includes a problematic variable – membership in existing Samurdhi program. As and when the program is reformed, this variable will change in character, and therefore cannot be used any time *after* the first time. The PMTF, by all considerations, should continue to be used for entry into and exit from the program for at least some length of time to ensure consistency over time. Which is why, in spite of the fact that this variable is reasonably correlated with poverty and adds to the explanatory power of the OLS regression, in our view it *should not* be included in the selected PMTF.

Thus Model 2, which is Model 1 minus the Samurdhi membership variable, is the one to focus on, and offers the best prediction and lowest error rates under the circumstances. With the cutoff point set at 25 percent of the population, Model 2 has an undercoverage rate of 51 percent and a leakage rate of 39 percent; the corresponding rates fall to 42 percent and 35 percent with a cutoff point equal to the consumption of the 30th percentile of the population.

For various reasons, using weights for the province a household belongs to may be problematic for targeting benefits. Thus we also consider options that *omit the province location variables*,

¹⁴ For instance, when the poverty line is the 25th percentile of actual per capita consumption, or Rs. 1129, the eligibility threshold is also a predicted per capita consumption of Rs. 1129.

first case of that being Model 3 – identical to Model 2 except for the fact that the province dummies are now dropped. The undercoverage rate is 1 percentage point higher in Model 3 as compared to Model 2 for cutoff points equal to the 25th and 30th percentiles; the R-squared falls by 0.02 in Model 3 from that in Model 2.

Models 4 and 5 illustrate two cases for what happens as the set of predictors are reduced further. Model 4 omits the land ownership variables from Model 3, and Model 5 excludes, in addition to land ownership, all demographic characteristics of the household, information on household head, and that on household durables and other assets. The “sacrifice” in fit and error rates is apparent comparing the R-squared, undercoverage and leakage rates with those of Model 3. In case of Model 4, the undercoverage and leakage rates increase by 1 to 2 percent depending on the cutoff point; for Model 5 the undercoverage rates are much higher (almost 20 percent for the lower cutoff points), and so are the leakage rates (almost 4 percent for the lower cutoff points).¹⁵

Model 6, similarly, shows what happens if the two community level variables used in Model 3 are omitted. While the R-squared and error rates are similar to those in Model 3, both undercoverage and leakage rates in Model 6 are higher (by 2 and 1 percent respectively) when the cutoff point is the per capita consumption of the 40th percentile of the population.

Finally, Model 7 takes Model 3 as a starting point, and omits all but the most significant variables (with significance level of 99 percent or higher). This restricted set of variables is seen to work almost as well as Model 3 – R-square and undercoverage rates are identical, and leakage is just 1 percentage point higher for Model 7 for the 30th and 40th percentile cutoff points.

Comparison with PMTF results from other countries

The results in general compare well with those from similar exercises conducted for other countries. For a poverty line and eligibility cutoff equal to the 30th percentile of actual per capita consumption, Models 2, 3 and 7 yield undercoverage rates of 42-43 percent and a leakage rate of 35 percent. For the same poverty line and cutoff in percentile terms, a similar exercise using Jamaica data for 1989 yields undercoverage and leakage rates of around 41 and 34 percent respectively; the corresponding rates are 39 percent and 24 percent for urban Bolivia, and 54 and 35 percent for urban Peru (1990 data for both cases) (Grosh and Baker). Using Jamaica data for 2000, the corresponding rates are 69 and 44 percent.

Choice of Models

In view of these results, we make a few observations and recommendations. *Model 2* is the best overall performer. However, we find a few problems with this model:

First, Model 2 includes province variables, which *may* be problematic to include in a formula politically. Further, the lack of data for the North-East makes determining the province level weight for this region highly problematic, which in turn adds to the arguments against using such weights.

Second, it turns out that the coefficients on the province dummies for North-Central and Uva (for regression results, refer to Table A-2, Appendix) lead to *positive* weights for location in these provinces, with Western Province as the reference group. This is counter-intuitive, since the latter is known as by far the wealthiest province in Sri Lanka (and the average consumption expenditures from the SLIS itself confirms that), while North-Central and Uva are relatively poor by any measure. While such counter-intuitive results are entirely possible

¹⁵ While all the cases are not presented in Table 2, omitting other combinations of variables – for example retaining household assets and omitting housing information – exhibit similar drops in R-square and increase in error rates, as compared to Model 3 (these results are available upon request).

in a regression set up where the coefficient measures the “province effect” after controlling for other effects, it makes using these coefficients as weights all the more difficult.¹⁶

The third problem is related to the point above. It turns out that while Model 2 has the most accurate predictions overall, because of the strongly positive weight on North-Central, the undercoverage rate is much higher for that province (62 percent) compared to the rest of the country (see Table 3). In contrast, undercoverage is much below the country average for Central and Southern provinces. On the other hand, Model 7 has undercoverage rates that are far more uniform across provinces, although as seen above, its error rates for the entire country are slightly higher than those for Model 2. Such unevenness in identification of the poor across provinces, in our view, is problematic.

Given these difficulties, we recommend Model 7 – which is the restricted version (retaining only the highly significant coefficients) of Model 3 – as the preferred option. Both Models 7 and 3 avoid the problems associated with assigning weights to location in provinces, and also achieve far more uniform levels of undercoverage across provinces. Model 7 is preferable to Model 3, because it achieves almost identical results with fewer variables, reducing the time and cost for collecting information.

<i>Province name</i>	Undercoverage		leakage	
	<i>Model 2</i>	<i>Model 7</i>	<i>Model 2</i>	<i>Model 7</i>
Western	0.61	0.51	0.31	0.31
Central	0.25	0.46	0.37	0.31
Southern	0.34	0.38	0.32	0.33
North-Western	0.53	0.49	0.42	0.40
North-Central	0.62	0.35	0.37	0.49
Uva	0.39	0.30	0.28	0.38
Sabaragamuwa	0.40	0.44	0.40	0.38
<i>All Provinces</i>	<i>0.42</i>	<i>0.43</i>	<i>0.35</i>	<i>0.36</i>

Next, it will be useful to explore the tradeoffs in overall error rates and accuracy in predictions between Models 2 and 7. Figure 1 shows that for the lower two cutoff points (25th and 30th percentiles of actual per capita consumption), undercoverage rates are about 1 percentage point higher for Model 2, while for the 40th percentile cutoff point Model 7 has lower undercoverage. Figure 2 shows that leakage rates are lower for Model 7 for the 25th percentile cutoff point, and lower for Model 2 for the 35th and 40th percentile. Finally, as listed in Table 2, R-squared for Model 7 is 2 percentage points lower than that for Model 2.

Thus on balance, while Model 2 appears to perform marginally better than Model 7, the differences are quite small and sometimes even favor Model 7. So the recommendation in favor of Model 7 over Model 2 does not seem to lead to sizeable “sacrifices” in terms of accuracy of predictions. The advantage of Model 7 vis-à-vis Model 2 on the other hand lies in the smaller number of variables used in the former, at the cost of little higher undercoverage and/or leakage rates for most likely cutoff points.

Finally, Model 2 should be not ruled out as an option. If achieving maximum accuracy in the aggregate – rather than avoiding large differences in errors across provinces – is the prime concern, and assigning weights to provinces (even if some are counter-intuitive) is not problematic, Model 2 should be considered since it does yield the best predictions. Ultimately, this would be a question for policymakers to decide on. That said, in our opinion, Model 7 offers the best combination of results, and would be our focus for the remaining analysis.

¹⁶ In this case, it turns out that the possession of certain consumer durables (fan, electric cooker) are highly negatively correlated with location in North-Central and Uva provinces. The coefficient on the dummies for these 2 provinces in that case capture the location effect after controlling for the effect of these assets, and these location effects turn out to be counter-intuitive.

Exploring error rates in rural¹⁷ and urban areas separately

Looking at the overall undercoverage and leakage rates, as we have done so far, does not say anything about how well the formulae are able to predict separately for rural and urban regions. This is considered important, since if it is the case that one or both models have disproportionate error rates in either rural or urban region, one may need to consider new options that reduce errors for the relevant region.

The first row of tables 4 and 5 list the undercoverage and leakage rates with Model 7 for different cutoff points for rural and urban regions. Undercoverage is found to be considerably higher for urban region than for rural (Table 4). For instance, when the cutoff point is set at the 30th percentile, the rural and urban undercoverage rates are 41 and 71 percent respectively. The gap between rural and urban areas is much smaller for leakage rates (Table 5), e.g. 4 percentage points for the 30th percentile cutoff.

Table 4: Undercoverage rates in Rural/Urban areas

<i>cutoff</i> <i>ptiles</i> →	<i>Total</i>			<i>Rural</i>			<i>Urban</i>		
	25	30	40	25	30	40	25	30	40
model7*	0.53	0.43	0.29	0.50	0.41	0.26	0.78	0.71	0.53
model8*	0.53	0.43	0.28	0.51	0.42	0.27	0.67	0.53	0.35
model9	0.51	0.42	0.28	0.50	0.41	0.27	0.58	0.46	0.44
Samurdhi	NA	NA	0.42	NA	NA	0.40	NA	NA	0.62

Table 5: Leakage rates in Rural/Urban areas

<i>cutoff</i> <i>ptiles</i> →	<i>Total</i>			<i>Rural</i>			<i>Urban</i>		
	25	30	40	25	30	40	25	30	40
model7*	0.39	0.35	0.31	0.39	0.35	0.30	0.4	0.39	0.33
model8*	0.39	0.35	0.31	0.39	0.35	0.30	0.47	0.38	0.36
model9	0.38	0.35	0.30	0.38	0.35	0.30	0.31	0.26	0.30
Samurdhi	NA	NA	0.43	NA	NA	0.42	NA	NA	0.57

Notes: "NA" refers to "Not Applicable" since the current Samurdhi covers 40 % of population.

However, the problem of undercoverage in urban areas is less important than it appears. The urban sector comprises only 14 percent of the total sample, and has far lower incidence of poverty – only 18 percent of the urban population, for instance, fall in the bottom 30 percent of the population in terms of per capita consumption. The low share of population in the urban sector and their relatively better economic status combine for a very low share of the urban sector in the total number of poor in the country. Again, if the 30th percentile is taken as the poverty line, only 8 percent of those below that line live in urban areas. This implies that even though undercoverage in urban areas using Models 2 and 7 appear high in percentage terms, these errors amount to relatively small numbers of poor urban people who are actually left out of the program.

That said, it is well worth looking into whether these errors in urban areas can be minimized, and at what cost to the rural and overall error rates. This is done using Models 8 and 9. **Model 8** is identical in all respects to Model 7, with one exception – the *dummy for urban region is dropped* in Model 8. The idea behind leaving out the urban dummy is simply to see what happens when the one variable that is most responsible for introducing an “anti-urban bias” into the models (given the positive and significant coefficient of the urban dummy in all models) is omitted.

A comparison between Models 7 and 8 is instructive (Tables 4 and 5). Although overall undercoverage rates are similar for the two models, Model 7 has lower undercoverage for rural areas, while Model 8 has a similar advantage in urban areas. There is almost nothing to choose between the two models for leakage rates – the only difference appears to be in urban areas for some cutoff points, where leakages are marginally higher for Model 8.

¹⁷ Rural areas include the plantation sector in all models. Introducing a variable for the plantation sector, separate from rural and urban areas, lead to no change in the model’s fit or error rates.

Finally, **Model 9** is the case where regressions are performed separately for rural and urban areas – starting with the same set of variables as in Model 7 and retaining the significant ones (at 80 percent level) for each regression. This model is in some sense, the ideal in that it allows the best model for each sector to be estimated separately – allowing for structural differences – and would naturally be expected to minimize the error rates for each sector. In spite of this, a comparison of Model 9 with the others does not reveal a significant advantage from using the former. Models 7 and 8 yield overall undercoverage rates that are within 2 percentage points of those from Model 9. Model 9 does yield lower undercoverage for urban areas, especially in comparison to Model 7; the gap is much smaller when the comparison is instead with Model 8. On leakage, again overall rates are quite similar between Models 7, 8 and 9 (a difference of 1 percentage point or less); urban leakage rates are lower with Model 9 for some cutoff points.

In our opinion, the marginal gains from using the setup of Model 9 does not justify the considerable operational complications involved in using separate formulas for urban and rural areas. As Table 3 shows, the primary gain from using Model 9 would be in reducing urban undercoverage. However, if this objective is important, Model 8 should be considered superior, yielding results that are close to those from Model 9, without the need to consider separate formulas depending on where a household is located.

Recommendations for formulas for proxy-means testing

The above exercise concludes with our set of recommended options for the PMTF having expanded to *two* – *Models 7 and 8*. As discussed, these options represent various policy choices and constraints, and thus the exercise of selecting on formula from this set is best left to the government. Here, we summarize the discussion so far, with the goal of informing this decision-making in the best way possible. Even though we favor Model 7 or 8, Model 2 is included in the list of options below, to complete a list of reasonable alternatives for decision-makers to consider.

- Model 2 is the most comprehensive model – incorporating province dummies and variables from all categories mentioned above.
 - Yields the best fit and the lowest error rates on the aggregate
 - Province weights: (a) may be hard politically to incorporate in a formula (b) some weights are not intuitive, which reduces their acceptability
 - Because of the weights, rates of undercoverage vary widely – some provinces covered far better than others
- Model 7 omits province location variables, and restricts the set to variables that are highly significant (99 percent level and above)
 - Fit and error rates are close, but not identical to those for Model 2
 - Avoids the problems in Model 2 due to the use of province weights
 - Undercoverage rates are more uniform across provinces than for Model 2 – which is desirable in our opinion
- Model 8 is identical to Model 7, with urban location variables omitted. Therefore all the pros and cons of Model 7 vis-à-vis Model 2 apply.
 - Yields overall error rates very similar to those for Model 7
 - Reduces urban undercoverage as compared to Model 7, at the cost of increasing rural undercoverage
 - Could be selected over Model 7 *in case* reducing undercoverage in urban areas specifically is a high priority

The formulae derived from Models 7 and 8 are presented in detail in Table 6 below, based on the regression results listed in Table A-2, Appendix. The scores are arrived at by multiplying the

coefficients from the particular models by 100, and rounding to the nearest integer. By construction, these models are such that every variable entering into the regressions is significant at 99 percent or above. The calculation to arrive at a score for a household uses the intercept/constant as the starting point. For any dummy variable, if the condition is true for the household, then the weight for the variable is multiplied by 1 and added to the constant. For continuous variables, the measure of the variable for the household is multiplied by the weight and the result added to the constant. Both the formulas are more likely to assign benefits to larger households; households where all children do not go to school; households with few durable goods and amenities, little land and livestock and poor housing; households with older heads; and where the head is a female widow, has lower levels of education and does not work as a salaried employee.

IV. Targeting Errors and Choice of Cutoff Points

This section will examine the targeting errors associated with the models recommended in the above section more closely. This will include exploring what these errors that are estimations based on a sample are likely to translate to in the population, and analyzing the incidence of targeting and the distribution of errors in targeting. Not only will this help understand how the targeting benefits will be distributed among different economic strata of the population, but also suggest ways in which each type of error can be reduced. The most important choice in this regard will be the choice of cutoff points, and the analysis here will show that as cutoff points are raised or lowered, for a given target group and a formula, there is a clear tradeoff between undercoverage and leakage errors. Given this, the choice of a cutoff point will hinge on the precise nature of policy priorities, and political and budgetary constraints.

Confidence intervals for targeting errors

The targeting errors of the models described above (Table 6) are given by Tables 4 and 5. While considering these errors, it is important to note that since these are estimates based on a sample, they are characterized by standard errors, which in turn determine the confidence intervals of these estimates. Tables A-3 and A-4 (Appendix) lists the 95 percent confidence intervals for the estimated undercoverage and leakage rates for each of the two models (Models 7 and 8). For instance, with a cutoff equal to the 30th percentile of actual per capita consumption, and a target group defined by the population under a poverty line equal to the 30th percentile, targeting using Model 7 will result – with a 95 percent probability – in an undercoverage rate between 40 and 47 percent (where the point estimate is given by 43 percent).

The confidence intervals are important for policymakers to consider, since they provide a reasonable indication of the range within which targeting errors can be expected to fall. In that sense, they indicate the accuracy of the targeting errors estimated from a sample survey. In general, the 95 percent confidence intervals for our models form a small range around each point estimate for undercoverage/leakage rate, as a direct result of the relatively small standard errors of all the point estimates. This indicates, firstly, the accuracy of the point estimates, and secondly, a high likelihood that the targeting errors a policymaker would be concerned about would be quite close to the figures estimated in Tables 4 and 5.

Comparing predictions of models with existing Samurdhi coverage

It is useful to compare the simulated results with the models described so far with the targeting of the existing Samurdhi system (refer to Tables 4 and 5). The Samurdhi foodstamps go to about 40 percent of the SLIS sample (excluding the North-East), which implies that the relevant comparison to make is between the undercoverage and leakage of the Samurdhi targeting system and that of our models taking the cutoff point as the per capita monthly consumption of the 40th percentile of the sample.

Table 6: Weights on each variable for the selected models

<i>Variables</i>	<i>Dummy</i>	<i>Model7</i>	<i>Model8</i>
Location			
Rural/Estate	*	-10	
Community characteristics			
Public/Private bank in community	*	7	8
Divisional Secretariat in community	*	8	9
Household assets			
Car/van	*	40	40
Cooker (kerosene/gas/electric)	*	15	17
Bicycle/Tricycle	*	4	4
Fan	*	11	11
Refrigerator	*	11	12
Motorcycle/Scooter	*	9	8
Radio/CD/Cassette player	*	4	4
Sewing Machine	*	7	7
Tractor	*	15	15
TV/Video player	*	7	8
Land and livestock			
Cultivable land owned by household : 1<Acres<=2	*	8	7
2<Acres<=4	*	8	8
Acres>4	*	17	16
Livestock (any)	*	8	8
Household head			
<u>Not</u> a female who is widowed/separated/divorced	*	6	5
Age: 70-79	*	-6	-5
80 and above	*	-13	-13
Education: Passed OL or Grade 11	*	7	7
Passed AL/GAQ/GSQ	*	10	10
Has Degree/PG/Diploma	*	17	16
Work: Salaried employment or in business	*	5	5
Household demographics			
Household size: 3-4 members	*	-22	-23
5-6 members	*	-39	-39
7-8 members	*	-51	-52
>8 members	*	-59	-59
<u>All</u> children age 5-16 attend school	*	7	6
Housing characteristics			
Dwelling owned by hhold	*	4	4
Fuel for cooking: Gas/electricity	*	12	13
Toilet: Private and flush type	*	16	16
No. of Rooms (excl. kitchen/bath) <i>per</i> hhold member		17	16
Walls: <u>Not</u> cabook/mud/plank/cadjan	*	6	6
Constant		715	707
<i>Notes:</i>			
1) All scores are derived from regressions of (log of) per capita consumption expenditure on a set of variables			
2) The score for each variable is its coefficient in the regression, <i>multiplied</i> by 100, and <i>rounded</i> to the nearest integer			
3) The aggregate score for each household is calculated as <i>constant +/- the weight on each variable</i>			
• For each dummy variable (indicated by *), multiply the score by 1 <i>if true</i> for household, by 0 <i>if not true</i>			
• For each continuous variable, multiply the score by the value of the variable for the household			
4) Regressions include only variables with significance level of 99 percent and above			

As Tables 4 and 5 show, the errors are significantly lower for both the models considered, when compared to existing Samurdhi targeting, for the relevant cutoff point. Overall undercoverage is 13-14 percentage points lower for our models, compared to the Samurdhi undercoverage of 42 percent. Similar gaps exist for the rural areas taken separately, where the Samurdhi undercoverage rate is around 40 percent. In urban areas, the Samurdhi undercoverage rate of 62 percent is higher 9 percent or more than that of any of our models. In terms of overall leakage, the Samurdhi rate of 43 percent is at least 12 percent higher than that of the two models, including a gap of 12 percentage points for leakage in rural areas and between 21 and 24 points in urban areas. Thus by all considerations – undercoverage, leakage and the same in urban and rural areas considered separately – the models just described perform much better in targeting the poor than the existing Samurdhi targeting system.

Incidence of targeting and distribution of errors

One concern common to both the models described above is the relatively high rate of undercoverage. However, in judging the characteristics of a model, it is also important to look at *whom the model targets*, and conversely, *who is missed*. While it is always unsatisfactory to fail to cover those who fall below the poverty line, the error is less grave if the people who are excluded fall only just below the poverty line rather than at the very bottom of the welfare distribution. Similarly, while it is always undesirable to cover those who are outside the target group, the error is more palatable if people who are incorrectly included fall just above the poverty line rather than at the top of the distribution.

Cutoff at 30th percentile of per capita consumption¹⁸: Table 7 and Figure 3 show the incidence – how the targeted population is distributed among various groups using Models 7 and 8, when the cutoff threshold is set at the consumption of the 30th percentile. Both models show highly progressive targeting – less than 2 percent of the richest quintile are identified as beneficiaries, compared to two-thirds of those in the bottom quintile. For both models, coverage declines from the first to the second to the third decile (e.g. from 73 to 52 to 44 percent with Model 7) – which is desirable. Figure 3 shows this progressive pattern of targeting for either model.¹⁹

Table 7: Coverage by deciles if cutoff=30th percentile

<i>decile</i>	<i>Model7</i>	<i>Model8</i>
1	0.73	0.74
2	0.52	0.53
3	0.44	0.44
4	0.29	0.29
5	0.25	0.24
6	0.18	0.18
7	0.11	0.11
8	0.06	0.06
9	0.02	0.03
10	0.00	0.01
Total	0.26	0.26

The question of who are missed out is addressed by Figure 4, which shows the distribution of the cases of Type I error or undercoverage among the bottom three deciles. The largest proportion of those who are missed by either model belong to the group closest to the poverty line (the 3rd decile), followed by the second and the lowest deciles in that order. In case of Model 7, 43 percent of the poor individuals

¹⁸ The 30th percentile of actual per capita consumption expenditure (monthly) is defined by a cutoff point of 709 using either Model 7 or 8.

¹⁹ Note that both models target *less than 30 percent* of the population on the aggregate (last row of Table 5), even though the cutoff is set at the 30th percentile. This is because the 30th percentile in terms of actual consumption is not equal to the 30th percentile in terms of *predicted* consumption. For example, Model 7 predicts consumption such that only 26 percent of the population has predicted consumption less than the cutoff point (which is the *true* consumption of the 30th percentile of the population).

missed by the formula belonged to the 3rd decile, 37 percent to the second, and 21 percent to the first.

The second question to consider is the converse of the above – namely, who are the undeserving recipients of the program as identified by each model. Clearly, the errors of leakage are less damaging, higher the proportion of cases that are located just above the poverty line. Figure 5 shows that the highest proportion of Type II errors occur among individuals just above the poverty line (4th decile), and the proportion declines monotonically with higher deciles. As many as 58 percent of the undeserving recipients identified by Models 7 or 8 belong to the two deciles just above the designated poverty line.

Thus overall it appears that in spite of significant errors in identifying the poor, these formulas do induce targeting that is highly progressive. Moreover, those who deserve benefits but are missed out, as well as those who are identified incorrectly as beneficiaries overwhelmingly tend to belong to groups that are relatively close to the poverty line, which has obvious desirable welfare implications. Finally, the two models considered are very similar in terms of overall incidence and distribution of Type I and Type II errors.

Cutoff at the 40th percentile of per capita consumption²⁰: As seen before (Tables 2, 4 and 5), changing cutoff points has significant effect on undercoverage and leakage rates. In order to see how it affects incidence of targeting and distribution of errors, we consider a case with the cutoff threshold for targeting set at the per capita consumption of the 40th percentile of the distribution, with the target group also expanded to include the bottom 40 percent of the distribution.

Table 8: Coverage by consumption deciles if cutoff=40th percentile

<i>decile</i>	<i>Model7</i>	<i>Model8</i>	<i>Samurdhi</i>
1	0.89	0.91	0.69
2	0.75	0.75	0.60
3	0.68	0.69	0.54
4	0.54	0.55	0.49
5	0.45	0.47	0.47
6	0.35	0.34	0.43
7	0.25	0.24	0.36
8	0.15	0.16	0.25
9	0.05	0.05	0.18
10	0.01	0.01	0.06
Total	0.41	0.42	0.41

As shown in Table 8, targeting is again highly progressive with either model. Moreover, with the higher cutoff point, even lower proportions of people in the lowest deciles are missed (compare with Table 5). On the other hand, the benefits now “leak” to a higher proportion of the top deciles (about 7 percent of the top 3 deciles with Model 7, compared to about 3 percent earlier).

Table 8 and Figure 6 also contrast the incidence of targeting using these models, with the existing Samurdhi targeting (the comparison is valid since Samurdhi reaches about 40 percent of the population). While Samurdhi targeting is progressive, the just-derived models perform far better in that respect. For instance, 12 percent of the

population in the highest quintile receives Samurdhi benefits, while the corresponding figure for Model 7 or 8 is only 3 percent.

Just as before, we look at the distribution of cases of Type I and Type II errors among the different consumption groups. The overall patterns are similar to that for the previous case – the highest proportion of Type I as well as Type II errors occur among the groups closest to the poverty line, and the proportions decline as one moves further away from the poverty line (Figures 7 and 8).

²⁰ The 40th percentile of actual per capita consumption expenditure (monthly) amounts to a cutoff point of 721, using either Model 7 or 8.

Choice of cutoff points

From the above, it is important to note two things that happen as the cutoff point and the target group are expanded from the 30th percentile to the 40th percentile. First, the share of the lowest deciles in total number of cases of undercoverage/Type I error falls (e.g. for Model 7, the share of the bottom two deciles drops from 57 percent to 31 percent). Second, the share of the highest deciles in total number of cases of leakage/Type II error increases (again for Model 7, the share of the top three deciles combined rises from 10 percent to 17 percent). These are consistent with what was seen above, that incidence of targeting improves among the poorest groups, but seemingly at the cost of incorrectly including some more people from the richest groups into the target population.

Table 9: Different cut-off levels for a given target group

Cutoff level	Undercoverage rate	Leakage rate	Coverage
30 pctile of actual ^a	0.43	0.35	0.26
30 pctile of predicted ^b	0.37	0.37	0.30
40 pctile of actual ^c	0.22	0.44	0.41

Note: Model 7 is used to obtain the predicted value of per capita expenditure. The *target group* is the bottom 30 percent defined by actual per capita consumption expenditure

The analysis so far also implies that for a given target group, raising the cutoff point will increase coverage. Note that all calculations of undercoverage and leakage till now have assumed that the poverty line that defines the target population is the *same* as the cutoff point adopted for defining eligibility.

This need not always be the case. For instance, while the policy *objective* may be to target the poorest 30 percent of the population, to achieve this one has the option of setting the cutoff point used by the targeting formula to determine eligibility at various levels (e.g. at the 25th, 30th or 35th percentile of consumption).

Table 9 shows the error rates when different cutoff points are chosen, for a given target group – the people below the poverty line equal to the 30th percentile of actual per capita consumption. Undercoverage and leakage depend on the Type I and Type II errors in targeting *this group*, for different cutoff points (in terms of the predictions from Model 7) to select beneficiaries.

When the cutoff is set equal to the 40th percentile of actual per capita consumption, Model 7's predictions lead to undercoverage of 22 percent and leakage of 44 percent for the target group, with an overall coverage of 41 percent of the entire population. This should be contrasted with the results from the 1st row, which represents the standard case of the cutoff for eligibility being equal to the poverty line that defines the target group. As expected, raising the eligibility cutoff from the 30th to the 40th percentile (from Rs. 1201 to Rs. 1347 per capita) reduces undercoverage significantly among the target group (by 21 percentage points), but with higher leakage (by 9 percentage points). Overall coverage also increases by 15 percentage points. Table 9 also shows that if instead the cutoff point is raised to Rs. 1237, one achieves undercoverage and leakage rates (equal to 37 percent for both) somewhere between the above two cases. The intuitive appeal of this cutoff lies in the fact that since it is equal to the 30th percentile of *predicted* per capita consumption, exactly 30 percent of the population is identified as beneficiaries using this cutoff.

From this discussion, the following points emerge as key factors in deciding on cutoff points.

- First, it is important to define the *target group*, which will depend on the “poverty line” used to determine who belongs to this group.

^a An individual is eligible for benefit if his/her predicted per capita expenditure is below 30th percentile of actual expenditure (Rs. 1201).

^b Eligible for benefit if predicted per capita expenditure is below 30th percentile of *predicted* expenditure (Rs. 1237).

^c Eligible for benefit if predicted per capita expenditure is below 40th percentile of actual expenditure (Rs. 1347).

- The criteria used to define this line may be *absolute* (e.g. money amount required to attain a minimum level of basic needs) or *relative* (e.g. the 25th or 30th percentile of per capita consumption expenditure).
- In general, as the target population expands (poverty line is set higher), targeting errors, especially undercoverage tend to get reduced. This is because the targeting model's predictions are better on the average for identifying a larger group.
- Most importantly, defining the target group would depend on policy priorities, political realities and budgetary constraints
- Second, one must define the *cutoff point* for eligibility (defined in terms of the “scores” or predictions from the model).
 - If undercoverage is a prime concern, raising the cutoff point is likely to increase coverage for a *given* target group, but at the cost of higher leakage.
 - Moreover, if there is a hard budget constraint, the benefits available for the targeted population will reduce as leakage becomes higher. These competing factors need to be taken into account in deciding the optimal cutoff point from the point of view of undercoverage and leakage rates.

V. Comparisons with Alternative Methods of Estimation

In this section we will compare the results from the models derived so far with those derived using different methodologies. This will help understand whether any of the alternate methodologies yield significantly different results, as well as indicate the pros and cons of using alternate methodologies.

Estimations using only a poorer segment of the population

The first alternate methodology consists of using the poorer segment of the population to derive a PMTF. Grosh and Baker presents a case where they use only the poorest half of the population as the basis for building the targeting models, and show that such an approach leads to significantly lower undercoverage. This is because it puts more emphasis on accurately predicting the welfare of those near the bottom of the distribution, where the improvements are most relevant to the goal of poverty reduction. They also claim that using only the bottom half of the welfare distribution may be more realistic than using the whole population, since for most social welfare programs, members of the upper class are unlikely to bother making the necessary effort to claim eligibility.

However, in the current context, it seems problematic to mimic their approach completely, that of building the model on the data from the poorest half of the population, and simulate the error rates using the *same* part of the population (assuming that the richer half of the population would be completely left out of the targeting mechanism). While the first part of the exercise is possible, the simulations in our case must be conducted using the *entire* population to get an idea of what the error rates are likely to be in practice. This is because when the program is implemented, it will be impossible to determine who *truly* belongs to the poorest half of the population. This wouldn't be a problem if *ex ante* one could be sure that the higher income groups would be unlikely to apply for the benefits. But experience in Sri Lanka suggests otherwise, especially in view of the fact that 12 percent of the population in the *highest* quintile are currently receiving Samurdhi benefits.

In view of this difficulty, we try out two approaches that are similar to Grosh and Baker's approach. The first approach (**Model 10**) adopts a two-stage selection method. The first stage consists of a regression on the entire sample using the variables from Model 7, from which households whose predicted per capita expenditure is less than that of the 70th percentile of actual

per capita expenditure are selected. At the second stage, consumption is again regressed on the variables from Model 7 for the households *selected from the first stage*. The beneficiaries selected are those whose predicted consumption from the second stage regression is below the defined cutoff point. Undercoverage and leakage rates are computed taking into account the *entire* sample.²¹ The first stage regression – a departure from Grosh and Baker’s exercise – identifies in practice the poorer section of the population, on whom the second-stage model can be applied. Implementing this model would involve using two formulae – the first to eliminate the richest 30 percent of the population, and the second to select beneficiaries from the set surviving the first-round elimination.

Table 10: Comparing Model 7 with Models 10 and 11

Undercoverage rate				
Cut-off pctlile	25	30	35	40
model7	0.52	0.43	0.37	0.28
model10	0.55	0.46	0.40	0.31
model11	0.50	0.38	0.29	0.20
Leakage rate				
model7	0.39	0.36	0.33	0.31
model10	0.39	0.37	0.33	0.32
model11	0.40	0.38	0.37	0.35

The second approach, adopted in **Model 11**, involves regressing the variables from Model 7 on the bottom 80 percent of the population, with the estimated coefficients used to predict consumption expenditures of the *entire sample*. This approach may yield lower errors in targeting because, as mentioned above, it better captures the characteristics of that section of the population whose welfare is more relevant to the program. At the same time, it does not involve identifying the poorer section of the population *before* the model can be applied, and is in that sense possible to implement using information collected

from application forms for the program.

The error rates from Models 10 and 11 are shown in Table 10. First, it is apparent that Model 10 does no better than our benchmark, Model 7, and in fact does worse for all cutoff points in terms of undercoverage. Model 11, on the other hand, yields lower undercoverage rates and higher leakage rates than does Model 7, with the gaps in leakage somewhat smaller than those for undercoverage. Thus on the whole, Model 11 appears to perform marginally better than our best model so far, namely Model 7.²²

Having said that, in our opinion, the improvement in predictions using Model 11 is not large enough to justify its selection over Model 7. Firstly, the improvement is not unambiguous, since lower undercoverage has to be weighed against the higher leakage rates from using Model 11. Secondly, even the improvement in undercoverage is not very high (between 2 and 5 percentage points) for the cutoff points likely to be the most relevant (25th and 30th percentiles). Thirdly, the method of estimating the poverty predictors using only a certain part of the sample, depending on the poverty ranking of households, appears to be rather arbitrary. There is no reason, *prima facie*, for example to estimate the model for the bottom 80 percent (and not, say, the bottom 60 percent) of the population. For these reasons, Model 7 is easier to understand and explain – and given that

²¹ Stepwise regressions are used for both stages, where variables with p-values of less than 0.2 are retained. Note that Type I error for the whole model is the sum of such errors from the two stages – the poor who are excluded by the first stage regression, and then those who are excluded by the second stage regression. The source of leakage can however only be the second stage regression, since the first stage regression does not identify anyone as a beneficiary, but just “selects” qualified households for the second stage.

²² Both Models 10 and 11 were run for a number of possible “splits” of the sample. The cases presented here – namely the 30-70 split of the sample for the two stages of Model 10, and Model 11 being estimated on the poorest 80 percent of the sample – yielded the best results overall among numerous possibilities.

the advantages of Model 11 are marginal, Model 7 still appears to be the best choice in our opinion.²³

Testing sensitivity to “overfitting”

In the existing literature, a methodology sometimes adopted consists of using *half the sample* to run the regressions to predict welfare, and testing the predictions from this model by calculating undercoverage and leakage rates on the *other half of the sample*. The utility of this lies in reducing the likelihood of “overfitting” the sample.²⁴ By separating the estimation and the testing of the model between two non-overlapping parts of the sample, the model will likely be subjected to a harder test – by minimizing the bias in favor of the model that may occur when the predictions from the model are used on the same observations that were used to derive the coefficients. This is important to consider, so as to mimic as far as possible the real-world situation where our models will be applied to impute the welfare of households who will not be the same set of households for whom the formula is estimated. We conduct this exercise as a test for the sensitivity of our existing models in two ways: firstly, to see whether the coefficients of the model using this method are significantly different from those derived using the whole sample; secondly, to see how the targeting errors are when the new method, involving a “harder test” of accuracy in targeting, is used.

This exercise is conducted for the two sets of variables used in Models 7 and 8, and the two new models are called **7a** and **8a** respectively. Firstly, the coefficients from Model 7a turn out to be reasonably close to those from the original Model 7, and similarly those from Model 8a to the original Model 8. This suggests that the original models are quite robust to adjustments for overfitting. Secondly, the error rates using the new method are close to those of the original models for various cutoff points and poverty lines, with the former being usually higher by 2-3 percentage points.²⁵

Thus subjecting our OLS-based methodology to a harder test does not lead to significant increases in targeting errors; moreover, the coefficients or weights of the variables after adjusting for overfitting are similar to those of the original models. These results essentially validate our models 7 and 8, the methodology underlying them, as well as the results from our simulations of targeting errors with these models.

VI. Simulations to Measure Welfare Improvements

The exercise so far has examined the targeting formulas exclusively from the point of view of targeting efficiency. A related, and equally important question is the subject of this section, namely, what kind of welfare improvements can be expected with the application of these formulas in transferring benefits to eligible beneficiaries? In addressing this question, it is important to bear in mind that the exercise here is meant to be merely indicative, firstly, because it will be conducted under rather restrictive simplifying assumptions, and secondly, because it will be limited to showcasing a few cases which could serve as useful examples as abstractions of real world situations.

²³ Although this entire exercise is conducted in the form of comparisons between Model 7 and adaptations of methodology using the same set of variables, *similar results and conclusions hold* if Model 8 was selected instead, and the same changes in methodology were implemented using its variables.

²⁴ See Grosh and Glinskaya (1997), Hentschel et al (1998) for applications of this method

²⁵ Table A-5, Appendix lists some of the results for Models 7a and 8a, comparing these to results from Models 7 and 8 (for an eligibility cutoff point at the 40th percentile of actual per capita consumption).

Definitions and assumptions

To conduct this exercise, a measure of welfare has to be first defined. In this exercise, welfare will be measured by the Foster-Greer-Thorbecke (FGT) index; for various values of its parameter α : $\alpha=0$ reduces the FGT index to the well-known headcount ratio, which merely measures the proportion of poor individuals in the population; $\alpha=1$ reduces it to the “poverty gap” measure, which measures the average distance of those below the poverty line from the poverty line; $\alpha=2$ reduces it to the squared poverty gap or severity of poverty measure. Only joint consideration of these three indices can give an adequate description of poverty that satisfy two famous axioms of poverty measurement.²⁶

(a) even if the number of the poor is the same, if there is a welfare reduction in one poor household, a measure of poverty should detect an increase of poverty (captured by an increase in the poverty gap index); (b) even if the average welfare of the poor is the same, if there is a transfer from one poor household to another poor household, relatively better off, a measure of poverty should detect an increase of poverty (captured by an increase in the severity of poverty).

The simulations conducted here to calculate the loss/gain in welfare, measured by the FGT index, will attempt to measure the impact of not merely the *introduction* of a transfer payment system, but the *replacement* of the existing system of payments with a new one. Most of the existing literature for different countries, on the other hand, has considered the former, and in that sense our task is somewhat more complicated.

To make such calculations, we make some simplifying assumptions. Firstly, we assume that *all existing* Samurdhi foodstamp benefits to households would have been used by the household towards their monthly consumption of non-durables, and that *in the absence of* existing Samurdhi foodstamp benefits, the welfare (as measured by consumption) of the recipient households would be reduced exactly by the amount of benefits received. Secondly, we assume that *all additional* welfare benefits households would receive will be used for consumption; this implies that when new benefits are introduced, their welfare would increase by the amounts they receive.²⁷

Such assumptions are rather restrictive for two sets of reasons. First of all, in practical terms, the Samurdhi transfers received by households in the existing system are not used completely for immediate consumption – a varying amount is held back as compulsory deductions for reasons that include forced saving and insurance premiums, and some foodstamps may be used later by the household. The data does not allow us to measure the value of foodstamp that was *actually* available to the household after deduction, and used for the household’s consumption last month. Secondly, these assumptions do not allow for substitution in the household’s consumption basket on receiving or losing welfare benefits. In other words, it does not account for the possibility that a household would be substituting some of its food purchases by the goods he can buy with the foodstamps, which would imply that its monthly consumption on non-durables would not fall by an amount equal to the benefits lost if foodstamps were to be lost or reduced. Conversely, when a household starts receiving fresh benefits, because of similar substitution effects, the entire amount of benefits received may not be translated into consumption as our second assumption above requires it to do.

²⁶ A. K. Sen (1976); for the definition of FGT index and a detailed discussion, see Section II (Appendix)

²⁷ The second assumption is standard for most simulations conducted with targeting formulas in various countries (see Grosh and Baker (1995)). The first assumption is an additional one in our case, necessitated by the nature of our exercise.

Methodology for calculating welfare loss/gain

To measure welfare loss/gain, total welfare is first calculated on the current data – using FGT indices for the three values of α , taking the 25th percentile of actual per capita consumption as the poverty line – as the baseline case with the existing Samurdhi payments. Under the above assumptions, the “new” scenario is then constructed by first *deducting* the existing foodstamp payments from the consumption expenditures of households, and then by *assigning* payments to the consumption expenditures of households deemed eligible by Model 7, such that all new payments add up to the budget constraint (equal to the total existing Samurdhi foodstamp payments in the sample).²⁸

The target group is defined by the poverty line, in this case the 25th percentile of per capita actual consumption (amounts to a monthly per capita consumption of Rs. 1129), which is used for welfare calculations using FGT indices. The eligible group comprises of households whose predicted consumption (or “scores” from Table 5), using Model 7, is less than the selected cutoff point. The selected cutoff point for this exercise, *unless otherwise specified*, is also equal to the 25th percentile of actual per capita consumption expenditure.

The net gain/loss in welfare is calculated by taking the *difference* between the value of each welfare index in the new scenario and that for the baseline scenario (as a percentage of the latter), for each of the different payment schemes considered. The net change in welfare will be determined by a number of factors, including how well the new formula targets vis-à-vis the existing Samurdhi system, the coverage that will depend on the cutoff point and the budget, and the *method of distributing the available budget* for payments among the eligible group. The last-named is the focus of this analysis, to see what kinds of payment methods result in the highest gains in welfare, depending on what measure of welfare is adopted. In doing so, there will be no attempt to analyze all possible options for payment mechanisms to find the most optimal scheme. Rather, the focus will be on providing examples that indicate various possibilities and likely tradeoffs between various payment schemes.

Table 11 lists the welfare levels, using the approach described above, for seven alternate payment schemes, as well as the current scenario under the existing Samurdhi system. The numbers in parentheses indicate the percentage change in welfare for every scenario from the current scenario, as a percentage of the latter. The poverty line for these calculations, taken at the 25th percentile of actual per capita consumption, should be regarded as merely illustrative; qualitatively similar results hold if somewhat higher or lower poverty lines are used for the calculations.

The payment schemes are chosen to illustrate a broad range of options. The first three are the simplest to implement, since they involve fixed payments – per household, household member or dependent – to each eligible household.²⁹ Coverage using either of these options is about 20 percent of the population. Option 4 is similar, with the only difference that Rs. 200 per capita is paid to every household, starting from the poorest and moving upwards the consumption scale, till the budget is exhausted – this makes for a coverage of about 17 percent of the population. Option 5 involves a progressive payment scheme, whereby the benefits paid in per capita terms is the highest (Rs. 300) for households in the bottom decile, and declines for the second and third deciles. Option 6 involves paying a fixed amount of Rs. 500 per household, coupled with a per dependent component of Rs. 213, for all households below the cutoff point. Finally, option 7 involves the most customized payment schedule, whereby the poorest household receives enough

²⁸ The total Samurdhi foodstamp budget in the sample, expanded by the sample weights, amounts to about Rs. 6.8 billion, which is not far from official figures of foodstamp budget in 1999 (equal to Rs. 8.2 billion).

²⁹ *Dependent* refers to any household member of age 16 or below, or above age 60.

to achieve the consumption of the next poorest, these two poorest households receive enough to attain the consumption of the next poorest, and so on till the budget is exhausted.

It is important to note that option 7 has the most egalitarian or redistributive impact, by ensuring that the marginal rupee always goes to the neediest. This is confirmed by the figures that show that poverty gap and squared poverty gap both improve the most (11 and 17 percent respectively) under this scheme. However, this scheme is almost impossible to implement in practice, since it involves different payments for *every* household. That said, it serves as a useful benchmark for other, more practical schemes to be judged against.

Table 11: Comparison between different transfer schemes

Transfer Option	Description	α^a			Avg. transfer per Beneficiary ^e (Rs.)	Coverage (% of population)
		0	1	2		
Option 1	Fixed amount <i>per household</i> for all households below cutoff ^b	0.247 (-1.2)	0.053 (-11.7)	0.018 (-10.0)	176.46	19.57
Option 2	Fixed amount <i>per capita</i> for all households below cutoff ^b	0.245 (-2.0)	0.053 (-11.7)	0.018 (-10.0)	176.46	19.57
Option 3	Fixed amt. <i>per dependent</i> for all households below cutoff ^b	0.244 (-2.4)	0.053 (-11.7)	0.018 (-10.0)	176.46	19.57
Option 4	Rs. 200 <i>per capita</i> for all households till budget runs out ^c	0.245 (-2.0)	0.053 (-11.7)	0.018 (-10.0)	200.00	17.24
Option 5	Progressive: households in bottom decile get Rs. 300 per capita, the second decile Rs. 200; the third decile Rs. 100 ^c	0.246 (-1.6)	0.052 (-13.3)	0.017 (-15.0)	165.45	20.87
Option 6	Fixed amt. of Rs. 500 <i>per hhold</i> + Rs. 213 <i>per dependent</i> for all hholds below cutoff ^b	0.246 (-1.6)	0.053 (-11.7)	0.018 (-10.0)	176.46	19.57
Option 7	Marginal rupee to individuals in the neediest household ^d	0.248 (-0.8)	0.052 (-13.3)	0.017 (-15.0)	146.50	23.35
Current		0.25	0.06	0.02	84.97	40.64

Notes: (1) % changes in the welfare level for every scenario from that under the current system of Samurdhi (as a % of the latter) are in parentheses. (2) FGT measures are used to measure welfare. (3) All transfer options are based on per capita consumption expenditure predicted in [Model 7](#). (4) The fixed budget for program benefits is equal to the total amount of Samurdhi foodstamps received by households in sample. (5) The poverty line for the welfare calculation is set at the 25th percentile of actual per capita consumption. (6) All transfer amounts are *monthly* figures.

^a: α represents the sensitivity to the income distribution among the poor. For $\alpha=0$, the FGT measure is the Headcount ratio; for $\alpha=1$, the Poverty Gap; for $\alpha=2$, the squared Poverty Gap/severity of poverty.

^b: Cutoff point for eligibility = 25th percentile in actual per capita monthly consumption expenditure; *dependents* refer to all members of the household of age 16 or below or above age 60.

^c: Households receive benefits per capita, starting with the poorest, until the fixed budget is exhausted.

^d: The level of transfer is determined sequentially: the poorest household is given enough money per capita to reach the level of the second poorest. Then these two are given enough money to reach the level of the third poorest. This process is repeated until the budget is exhausted. Thus marginal rupee of benefit always goes to the neediest.

^e: Average transfer per beneficiary is calculated by dividing the total transfer budget by *all members* of the households receiving transfers (even if households receive different amount per member, as in options 3, 5, 6 or 7)

Welfare gains with simulations

The first thing to note from Table 11 is that *all* transfer options listed lead to only marginal gains (around 2 percent) for α -value of 0. This is because the total amount available for transfers, even when coverage is only around 20 percent of the population, is not enough to make a sizeable dent in the headcount poverty rate, which simply counts the number of people below the poverty line. This is all the more true since a large proportion of the eligible population belong to the lowest consumption deciles, who would need much higher transfers to move out of poverty.

On the other hand, all the options listed do lead to significant improvements when more distribution-sensitive welfare measures are chosen, namely when α is equal to 1 or 2. These welfare measures (α -value of 1 or 2), which are able to capture improvements in the average welfare of the poor and distribution among the poor, thus appear more sensitive to the changes we introduce. In other words, the scenarios we develop with the new targeting formula lead to greater gains in reducing the depth and severity of poverty than in reducing the number of poor in the population.

Finally, it is important to bear in mind that these simulations do not represent an *addition* in benefits on the aggregate, but rather reflect the net welfare gains from better targeting and redistribution among the poor of the *same aggregate level* of payments. As a result, the welfare gains would naturally be less than what would be the case if a targeted transfer scheme were *introduced* for the country.

Comparison between payment options

Option 5 – involving a progressive scale by which households in each decile receive a different amount in per capita terms – leads to the largest improvements in welfare for α -values of 1 and 2 (by 10 and 16 percent respectively).³⁰ In fact this option achieves a level of welfare equivalent to that with option 7, which represents the most elaborate scheme designed to maximize poverty impact. Options 1, 2, 3, 4 and 6 lead to very similar results in terms of improving welfare for α -values of 1 and 2, with gains of around 9 percent and 12-13 percent respectively over those from the existing scenario. Measured by headcount, options 2 to 4 lead to gains of around 2 percent, while the rest lead to somewhat smaller gains compared to the existing welfare level.

Thus on current evidence there is very little to choose between the different payment schemes considered here. The reasons are mainly to do with the fact that the *average* transfer per person in the eligible group, and the coverage as a proportion of the population are similar for all the options considered, and especially so for options 1 through 6. Although the differences are small, option 3 appears to achieve the best combination of results when all three values of α are considered. Option 3 is also attractive since it is easy to implement and makes simple intuitive sense, which may lead to greater political and public acceptance. If improving headcount is not a priority, option 5 is the best since it leads to the highest welfare gains when any redistribution-sensitive measure is adopted.

It should also be noted that options 2 to 4, which are based on payments to households fixed in *per capita* or *per dependent* terms, perform better in terms of reducing poverty headcount than option 1 which pays a flat amount to all eligible households. This is because payments based on number of members or dependents in a household are likely to be more progressive, since poorer households are often larger, with greater number of dependents.

To summarize, while the cases are considered here are by no means exhaustive, they do provide some useful indications on the comparison of welfare gains, depending on the payment scheme that is adopted. Adopting a progressive scale, like in option 5, is likely to lead to the best improvements in terms of average welfare of the poor or depth of poverty, and the distribution of welfare among the poor. On the other hand, to balance these objectives with that of achieving a modest decline in the number of poor people, *a scheme like option 3 may be the best option* – which is also easy to implement and intuitively appealing.

It is also important to bear in mind that the choice of the payment scheme is not independent of other policy decisions, namely the choice of target group and that of the cutoff point. In fact, all these factors need to be considered together to arrive at a combination of choices that is best

³⁰ Very similar results would hold if option 4 were to instead consist of a progressive scale of payments *per dependent* to households, with payments declining from the 1st to the 2nd to the 3rd decile.

suited to the policy objectives and political compulsions of the country. The next sub-section provides an example of how these different choices are linked in determining the welfare gains from a system.

Changing the cutoff point for a given payment scheme

Table 12 lists the results from changing the eligibility cutoff point, for a given payment scheme (Option 2) and a fixed poverty line (25th percentile of actual consumption) that defines the target group. As the cutoff point for eligibility is increased from the 25th to 30th percentile and beyond, the payment per dependent (and average payment per capita) declines since the budget is fixed at the same level as before. Coverage among the population naturally improves.

The welfare calculations show that at least for this payment scheme and the selected poverty line, there is no gain in welfare from increasing the eligibility cutoff. To see why this may happen, note that raising the cutoff can improve welfare by reducing undercoverage among the target population, and reduce welfare by increasing leakage to the non-target population, which reduces the benefits available to the target households for a given budget. For the case considered here, the second effect appears to outweigh the first, with the result that as the cutoff is raised, there is an overall reduction in welfare for any value of α . This happens in spite of improvements in coverage, among the entire population as well as among the poor.³¹

Table 12: Welfare gains from Option 3 for different eligibility cutoffs

Cutoff percentile	α			Average transfer per Individual (Rs.)	Coverage (% of population)
	0	1	2		
25 th	0.244 (-2.35)	0.053 (-8.53)	0.018 (-11.84)	176.46	19.57
30 th	0.244 (-2.25)	0.054 (-7.40)	0.018 (-10.54)	131.15	26.33
35 th	0.246 (-1.68)	0.054 (-6.29)	0.019 (-8.83)	103.90	33.24
40 th	0.247 (-1.37)	0.055 (-4.82)	0.019 (-6.17)	83.71	41.25

Notes: (1) % changes in the welfare level for every scenario from that under the current system of Samurdhi (as a % of the latter) are in parentheses. (2) FGT measures are used to measure welfare. (3) All transfer options are based on per capita consumption expenditure predicted in Model 7. (4) The fixed budget for program benefits is equal to the total rupee amount of Samurdhi foodstamps received by households in sample. (5) The poverty line for the welfare calculation is set at the 25th percentile of actual per capita consumption. (6) All transfer amounts are *monthly*.

Conclusion

In conclusion, it will be useful to summarize some of the key findings in this report, at the same time highlighting how these findings can inform key policy decisions of the government about the targeting system.

Based on OLS regressions of (log of) per capita consumption expenditures on a set of probable predictors of welfare, we recommend a set of models that can be used to derive the scores for a formula for targeting. Because of unreliability of data from the North-East, we exclude this region from our regressions and simulations. Model 7, one of our recommended models, includes as its set of predictors selected variables of the following types: location (rural/urban/estate), community characteristics, household assets, land and livestock ownership, household demographics including household size, age, education, occupation and marital status of the household head, and housing conditions. An alternate model suggested is Model 8, which is

³¹ The fact raising the cutoff results in improved coverage among the poor, for a given poverty line/target group, is illustrated by Table 9 above, albeit for a different poverty line equal to the 30th percentile.

identical to Model 7 with the exception that the urban/rural location variable is omitted from the former.

Targeting error rates, taking the country as a whole, are similar for the two models, and these rates compare well with those seen with similar formulas developed for other countries. They also represent significant improvements over the targeting efficiency of the existing Samurdhi system. Finally, the incidence of targeting by these formulas are shown to have desirable properties in general – they are highly progressive and a large majority of targeting errors occur among households that are relatively close to the poverty line on either side.

In addition, Model 2, which adds province level dummy variables to the basic setup of Model 7 is a notable case, in that it results in the most accurate predictions and lowest errors in general. However, using province-level weights can be problematic in practice. Firstly because the excluded region, North-East lacks such weight at the moment. Secondly, some of the weights run counter to the relative economic ranking and conventional wisdom about the provinces, because of multi-collinearity in the regression with variables like certain household assets, which may reduce the acceptability of these weights. Given these difficulties, and the fact that Model 2 leads to very marginal improvements in targeting over the other models, we focus on Models 7 and 8 as possible choices. The choice between these two would depend on whether reducing undercoverage in urban areas, at the cost of increasing undercoverage in rural areas, is a priority.

The selection of the target group, which involves selecting a “poverty line” that defines which households would be targeted, is another important policy decision for the government. Larger the target group, the more accurate are the predictions of our models – an advantage that has to be weighed against the fact that for a given budget, expanding the target group would reduce the impact of program on those who are the most needy.

For a given target group, the choice of the eligibility cutoff point – in terms of the *predicted* consumption or the score derived from the models – is crucial in determining coverage and targeting errors. Higher the cutoff point, lower is the likelihood of undercoverage of the target group, but at the cost of higher leakage to those outside the target group. On balance, what leads to higher or lower welfare levels is uncertain, and would depend on the payment system adopted, the target group chosen, and the cutoff points considered. Moreover, the decision on cutoff points would also depend on – beyond simulated (and imperfect) calculations of welfare gains/losses – on the policy priorities, budgetary and political constraints faced by the government.

Subsequently, a few simulations with different payment schemes suggest that on the whole, replacing the existing system with targeting using one of our formulas would lead to welfare gains. While this analysis is not exhaustive and is conducted under restrictive assumptions on household behavior, it suggests that the gains can be significant in terms of reducing the depth and severity of poverty. All the schemes considered yield very similar results, but based on small differences that do exist, a payment scheme that uses a progressive scale to provide benefits – benefits per household member are the highest for the first decile, second highest for the 2nd decile and so on – appears to result in the highest reduction in depth and severity of poverty. A simpler scheme, one that pays every household a fixed amount per dependent (child and elderly) in the household, yields a balanced combination of reduction in depth, severity and incidence of poverty.

Notably, all these choices – of the target group, cutoff point and the payment system – are not independent of each other, and must be considered together in the context of the policy objectives and priorities. The objective of this report has been to present a set of alternatives with their associated tradeoffs to aid the government in narrowing the set of choices to align with their policy priorities, and budgetary and other constraints. As these choices and policy priorities

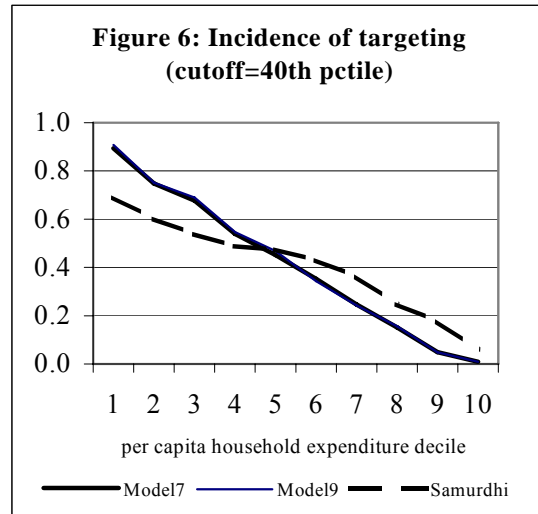
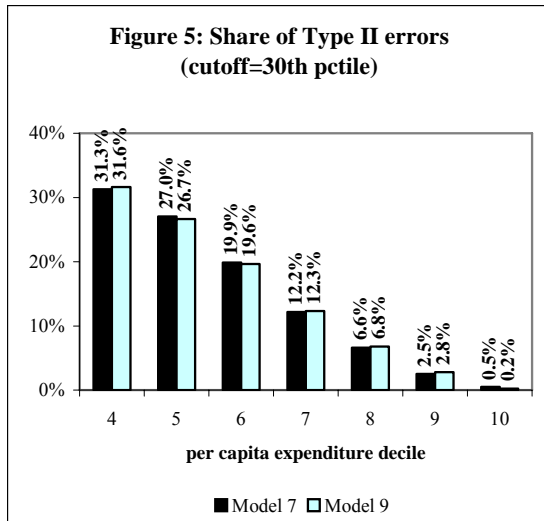
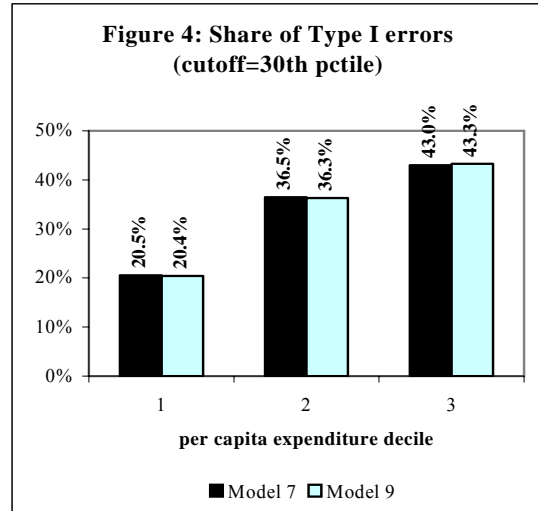
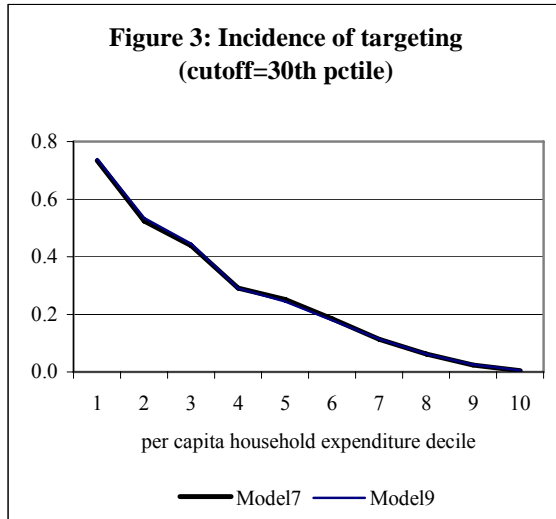
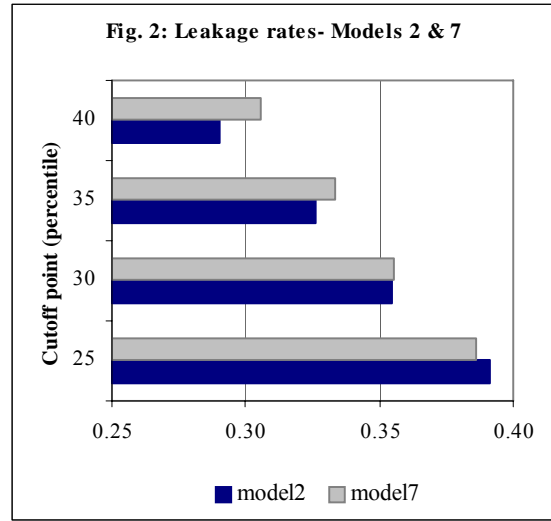
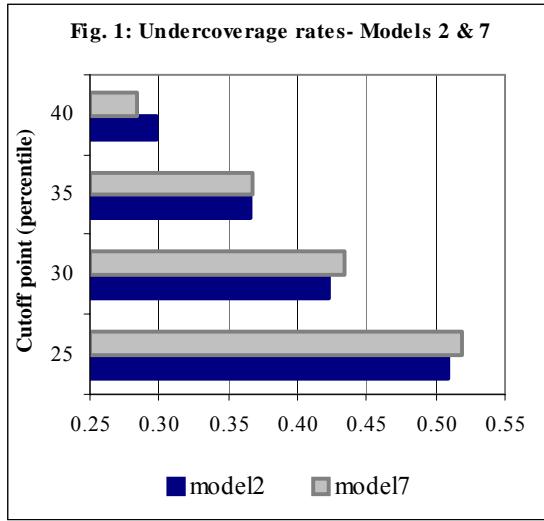
emerge, more exhaustive analysis, especially in developing suitable payment schemes must be undertaken.

Finally, the selection of the targeting formula itself should be considered an evolving exercise, which may benefit from fine-tuning based on local knowledge and the knowledge gained from the pilot exercise. Moreover, as noted before, the important question of whether the formulas developed here would be suitable for targeting or not should be considered in the context of future survey plans. To enable future updating of the models developed here, future rounds of SLIS, or a similar survey that collects the kind of information used in these models, would be necessary. Therefore, decisions on adopting these SLIS-based formulas should be taken keeping in view Sri Lanka's plans for survey-based poverty monitoring in the medium term.

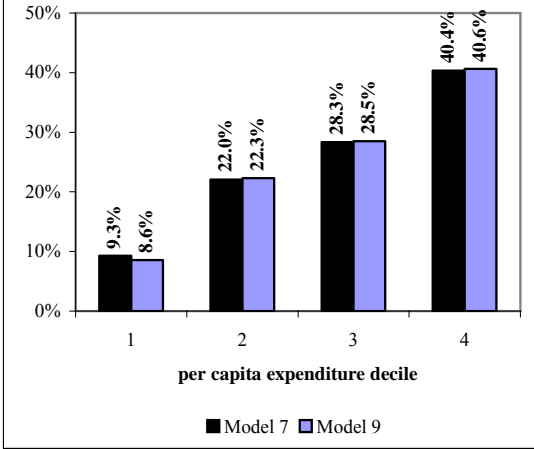
References

- Glewwe, P. 1990. "Efficient Allocation of Transfers to the Poor." Living Standards Measurement Study Working Paper No. 70. World Bank, Washington, D.C.
- Glewwe, P. and O. Kanaan. 1989. "Targeting Assistance to the Poor: A Multivariate Approach Using Household Survey Data." Policy, Planning and Research Working Paper No. 225, World Bank, Washington, D.C.
- Glinskaya, E. 2000. "An Empirical Evaluation of Samurdhi Program". Background paper to Sri Lanka Poverty Assessment. SASPR, World Bank, Washington, D.C.
- Grosh, M. 1994. "Administering Targeted Social Programs in Latin America : from Platitudes to Practice". Washington, D.C. : The World Bank, c1994. 174 p.
- Grosh, M., and J. Baker. 1995. "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation." LSMS Working Paper No.118, World Bank, Washington, D.C.
- Grosh, M and E. Glinskaya. 1997. "Proxy Means Testing and Social Assistance in Armenia." Draft. Development Economics Research Group, World Bank.
- Haddad, L., J. Sullivan, and E. Kennedy. 1991. "Identification and Evaluation of Alternative Indicators of Food and Nutrition Security: Some Conceptual Issues and an Analysis of Extant Data." Draft. International Food Policy Research Institute, Washington, D.C.
- Hentschel, J., J. Lanjouw, P. Lanjouw and J. Poggi. 1998. "Combining Census and Survey Data to Study Spatial Dimensions of Poverty: a Case Study of Ecuador". Policy Research Working Paper No. 1928, World Bank.
- Ravallion, M. and K. Chao. 1989. "Targeted Policies for Poverty Alleviation under Imperfect Information: Algorithms and Applications." *The Journal of Policy Modeling*, 11(2): 213-224.
- Sen, A. K. 1976. "Poverty: an Ordinal Approach to Measurement". *Econometrica*, 48: p. 437-446

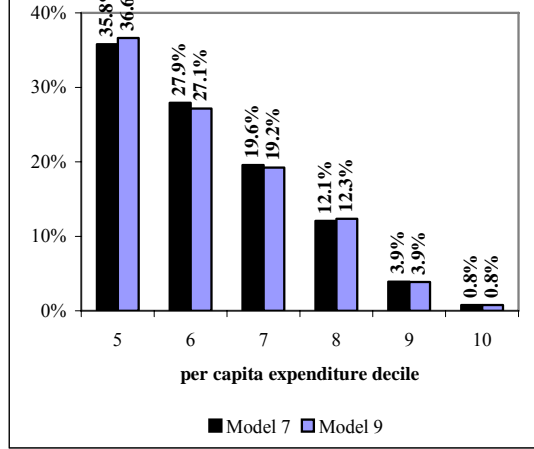
Figures



**Figure 7: Share of Type I errors
(cutoff=40th pctile)**



**Figure 8: Share of Type II errors
(cutoff=40th pctile)**



Appendix

I. Additional Tables

Table A-1: Description of all models

<i>Model</i>	<i>region</i>	<i>samurdhi</i>	<i>asset</i>	<i>community</i>	<i>head</i>	<i>member</i>	<i>ownership</i>	<i>housing</i>	<i>Selection by significance</i>	<i>Others</i>
Model 1	Y	Y	Y	Y	Y	Y	Y	Y	Medium ¹	
Model 2	Y	D	Y	Y	Y	Y	Y	Y	Medium	
Model 3	Drop province dummies	D	Y	Y	Y	Y	Y	Y	Medium	
Model 4	Drop province dummies	D	Y	Y	Y	Y	Drop landown dummies	Y	Medium	
Model 5	Drop province dummies	D	D	Y	D	D	D	Y	Medium	
Model 6	Drop province dummies	D	Y	D	Y	Y	Y	Y	Medium	
Model 7	Drop province dummies	D	Y	Y	Y	Y	Y	Y	High ²	
Model 8	D	D	Y	Y	Y	Y	Y	Y	High	
Model 9	Drop non-urban dummy	D	Y	Y	Y	Y	Y	Y	Medium	Regress OLS for urban and rural separately.
Model 10	Drop non-urban dummy	D	Y	Y	Y	Y	Y	Y	Medium	Two Stage ³
Model 11	Drop non-urban dummy	D	Y	Y	Y	Y	Y	Y	Medium	Regress OLS for the bottom 80 % of expenditure groups ⁴

Notes: * “Y” refers to “variables in a category are included before stepwise regressions. “D” refers to “all variables in a category are dropped before stepwise regressions.” Furthermore, we select variables using stepwise procedure: In the first step, we execute an OLS regression with all variables of categories included, and then keep variables with a certain level of significance for the next step. We continue this procedure until all estimated coefficients of variables have a certain level of significance (the p-value is either lower than 0.2 or 0.01, see below).

1. After stepwise regressions, only reasonably significant variables (their p-value<0.2) are selected.
2. After stepwise regressions, only highly significant variables (their p-value<0.01) are selected.
3. At the first stage regression, we regress model 7 and keep sample if the predicted per capita expenditure is smaller than 70 percentile of actual per capita expenditure. At the second stage, regress model 7 for the remaining sample. Using the estimated coefficients, compute the predicted per capita expenditure for the regression sample.
4. Regress model 7 using sample of the bottom 80 percent of per capita expenditure. Then using the estimated coefficients, compute the predicted per capita expenditure for the whole sample.

Table A-2: Regression results from OLS estimations
(Dependent variable: Log of actual per capita monthly consumption expenditure of household)

Variable	Description of variable	Model 2	Model 7	Model 8
prov2	1 if province=Central 0= otherwise	-0.180 (10.47)**		
prov3	1 if province=Southern 0= otherwise	-0.052 (3.05)**		
prov5	1 if province=North Western 0= otherwise	-0.029 (1.63)		
prov6	1 if province=North Central 0= otherwise	0.155 (6.93)**		
prov7	1 if province=Uva 0= otherwise	0.029 (1.37)		
prov8	1 if province= Sabaragamuwa 0= otherwise	-0.078 (4.06)**		
non_urban	1= lives in Rural/Estate 0= lives in urban area	-0.071 (4.49)**	-0.098 (6.11)**	
car_van	1= hhold has Car/van 0= otherwise	0.404 (16.80)**	0.402 (16.33)**	0.403 (16.32)**
cooker	1= hhold has Cooker (kerosene /gas/ electric); 0= otherwise	0.137 (7.11)**	0.147 (7.58)**	0.166 (8.68)**
cycle	1= hhold has Bicycle/Tricycle 0= otherwise		0.040 (3.68)**	0.036 (3.27)**
fan	1= hhold has Fan 0= otherwise	0.092 (5.83)**	0.108 (6.77)**	0.114 (7.15)**
fridge	1= hhold has Refrigerator 0= otherwise	0.112 (6.05)**	0.112 (5.92)**	0.118 (6.20)**
gen_pump	1= hhold has Pump/generator 0= otherwise	0.134 (2.40)*		
m_cycle	1= hhold has Motorcycle/Scooter 0= otherwise	0.074 (4.56)**	0.089 (5.39)**	0.083 (5.01)**
radio	1= hhold has Radio/CD player/Cassette player; 0= otherwise	0.050 (3.75)**	0.044 (3.21)**	0.043 (3.12)**
sew_mach	1= hhold has Sewing Machine 0= otherwise	0.078 (6.48)**	0.073 (5.93)**	0.073 (5.95)**
tractor	1= hhold has Tractor 0= otherwise	0.109 (2.91)**	0.149 (3.87)**	0.149 (3.88)**
tv_vcr	1= hhold has TV/VCR 0= otherwise	0.075 (5.88)**	0.072 (5.68)**	0.075 (5.87)**
bank_com	1= Public/Private bank in community 0= otherwise	0.056 (4.27)**	0.073 (5.59)**	0.083 (6.46)**
ds_com	1= Divisional Secretariat in community; 0= otherwise	0.112 (5.27)**	0.083 (3.88)**	0.091 (4.24)**
wid_f	0= Head is female and widowed /separated /divorced; 1= otherwise	0.055 (3.51)**	0.056 (3.49)**	0.053 (3.25)**
ageHcat4	1= hhold head Age: 70-79 0= otherwise	-0.041 (2.03)*	-0.055 (2.69)**	-0.054 (2.61)**
ageHcat5	1= hhold head Age: 80 + 0= otherwise	-0.116 (3.45)**	-0.134 (3.95)**	-0.133 (3.89)**
edulevH4	1= hhold head Passed OL or Grade 11 0= otherwise	0.065 (4.22)**	0.065 (4.16)**	0.065 (4.16)**
edulevH5	1= hhold head Passed AL/GAQ/GSQ 0= otherwise	0.102 (4.41)**	0.103 (4.32)**	0.103 (4.33)**
edulevH6	1= hhold head Has Degree /PG /Diploma; 0= otherwise	0.190 (4.41)**	0.169 (3.86)**	0.163 (3.71)**

activH34	1= hhold head in Salaried employment or business; 0= otherwise	0.043 (3.41)**	0.049 (3.80)**	0.050 (3.89)**
landown2	1=Cultivable land owned by household: 1<Acres<=2; 0= otherwise	0.053 (2.52)*	0.075 (3.48)**	0.069 (3.19)**
landown3	1=2<Acres<=4 0= otherwise		0.084 (3.31)**	0.079 (3.09)**
landown4	1=Acres>4 0= otherwise	0.158 (3.63)**	0.166 (3.74)**	0.159 (3.58)**
lstk	1=hhold has Livestock (any) 0= no livestock	0.071 (3.76)**	0.084 (4.39)**	0.082 (4.24)**
dysize2	1= hhold size: 3-4 members 0= otherwise	-0.209 (7.25)**	-0.220 (7.49)**	-0.227 (7.71)**
dysize3	1= hhold size 5-6 members 0= otherwise	-0.362 (11.86)**	-0.387 (12.67)**	-0.393 (12.84)**
dysize4	1= hhold size 7-8 members 0= otherwise	-0.483 (14.03)**	-0.512 (15.30)**	-0.516 (15.36)**
dysize5	1=hhold size 8 + members 0= otherwise	-0.537 (13.04)**	-0.587 (15.02)**	-0.586 (14.94)**
rsch5_16	1=All children in hhold of age 5-16 attend school; 0=otherwise	0.064 (3.63)**	0.065 (3.67)**	0.062 (3.45)**
ndep	# of dependents age<=16 or age>60	-0.010 (2.06)*		
dwellten1	1=Dwelling <u>owned</u> by hhold 0= not owned by hhold	0.024 (2.00)*	0.038 (3.09)**	0.035 (2.82)**
flrtyp3567	1= type of floor: cement terrazo tiles brick; 0=other	0.029 (1.97)*		
fuel1	1=Fuel for cooking: Gas/electricity 0= other	0.131 (5.86)**	0.122 (5.37)**	0.126 (5.53)**
latrtyp1	1=Toilet: Private and flush type 0= other	0.164 (10.01)**	0.163 (9.93)**	0.162 (9.80)**
rmsmem	No. of Rooms (excl. kitchen/bath) <i>per</i> hhold member	0.160 (11.21)**	0.165 (11.50)**	0.159 (11.05)**
walltyp137	0=Walls: cabook/mud/plank/cadjan 1= other	0.071 (5.32)**	0.062 (4.77)**	0.063 (4.80)**
Constant		7.162 (172.46)**	7.145 (175.16)**	7.071 (181.02)**
#Observations		5257	5257	5257
R-squared		0.58	0.56	0.56

Table A-3: 95% Confidence Intervals for undercoverage rates

<i>Cutoff percentile</i>	<i>Model 7</i>	<i>Model 8</i>
25	[0.48 ,0.57] (0.022)	[0.48 ,0.57] (0.021)
30	[0.40 ,0.47] (0.019)	[0.39 ,0.46] (0.019)
35	[0.33 ,0.40] (0.018)	[0.33 ,0.40] (0.018)
40	[0.25 ,0.32] (0.016)	[0.25 ,0.31] (0.017)

Notes: All calculations for poverty line same as the eligibility cutoff; confidence intervals are in square brackets; standard errors are in parentheses

Table A-4: 95% Confidence Intervals for leakage rates

<i>Cutoff percentile</i>	<i>Model 7</i>	<i>Model 8</i>
25	[0.34 ,0.44] (0.027)	[0.34 ,0.44] (0.027)
30	[0.31 ,0.40] (0.022)	[0.31 ,0.39] (0.022)
35	[0.29 ,0.37] (0.020)	[0.30 ,0.38] (0.020)
40	[0.27 ,0.34] (0.018)	[0.27 ,0.34] (0.018)

Notes: All calculations for poverty line same as the eligibility cutoff; confidence intervals are in square brackets; standard errors are in parentheses

Table A-5: Models 7a, 8a and Samurdhi after addressing "overfitting"*

	Split (1)		Split (2)	
	Undercover	Leakage	Undercover	Leakage
Model 7a	0.33 (0.28)	0.32 (0.31)	0.28 (0.28)	0.31 (0.31)
Model 8a	0.32 (0.28)	0.32 (0.31)	0.28 (0.28)	0.32 (0.31)
Samurdhi*	(0.42)	(0.43)	(0.42)	(0.43)

Notes: 1) To split the sample in half, households were sorted first by province, then by sector, and then by per capita consumption. In split (1), the even numbered observations were used for estimating predicted consumption level, and error rates were computed using the set of odd numbered observations. Split (2) reverses the roles of the odd and even numbered observations.
2) The numbers in parentheses are the error rates using the corresponding models *without* addressing overfitting
3) * Samurdhi error rates are computed for the entire sample
4) All error rates are calculated for a target group of bottom 40 percent and a cutoff of 40th percentile of actual per capita consumption

II. Notes to the main text

Clarifying the consumption measure

The consumption measure used for the exercise *includes* the transfers received by the household from the government, in various forms that include Samurdhi foodstamps as well as benefits from other programs. Since the objective of the Welfare reform is to *replace* the existing system of targeting (particularly, Samurdhi), it can be argued that what should matter for gauging welfare of a household is consumption *net* of Samurdhi foodstamps. However, such a net measure of consumption, in our judgement, should not be used in this case for a number of reasons.

Firstly, the Samurdhi foodstamps variable is reported in the data with some error, which primarily stems from the fact that a part of the transfers actually do not enter into the household's consumption, but is deposited in the Samurdhi Banks as "forced savings" of the household, or deducted in some cases as premiums for social insurance. This would not matter, if it was known for sure how much is held back from the households as savings. However, this is difficult to determine in practice, particularly since households – in response to the question posed in SLIS – have sometimes reported amounts received *net* of these deductions, and sometimes in *gross* terms, depending on their own interpretation of the question.

Secondly, even if the actual value of foodstamps *received* by the household last month were known, there is no way to determine how many of those were actually used to buy foodstuff last month, and the value of foodstuff bought last month (since the price of food from the Samurdhi cooperative store where these stamps are accepted may be different from the market price). Moreover – a more subtle point – it is not known to what extent the Samurdhi foodstamp would have *substituted* for purchases that would be made even without the foodstamps. All these problems make it very unlikely that a naive calculation of consumption net of the value of Samurdhi foodstamps received last month will yield a true measure of the relative welfare status of households *in the absence of Samurdhi*, which is the objective of such an exercise.

Thirdly, the Welfare reform is eventually meant to cover all welfare programs for targeting purposes. Thus deducting *only* Samurdhi foodstamps from the consumption of households to determine their relative welfare status will not be enough, and consistency will demand that welfare of households be calculated net of *all* welfare transfers. However, detailed information

on the other programs – particularly which benefits were received in cash and entered the consumption bundle of last month – are not available from the SLIS.

For all the above reasons that would introduce various kinds of measurement error, it was considered best to measure welfare by consumption – without netting out any transfers received. However, a sensitivity test was conducted by comparing the results from our “best” model with a similar model where the dependent variable was per capita consumption net of the reported Samurdhi foodstamp receipts. The results show an extremely high degree of overlap between the predictions of the two models under the two different measures of consumption – the models “agree” on more than 90 percent of the predictions. These results imply that how consumption is defined as a dependent variable in these models actually matters little – and using either definition one arrives at PMTFs whose predictions are quite similar in terms of identifying the poor.

*Defining the Foster-Greer-Thorbecke (FGT) measure for welfare*³²

The formula for the FGT index is given by:
$$P_{\alpha} = \frac{1}{n} \sum_{i=1}^q \left(\frac{Z - y_i}{Z} \right)^{\alpha}$$

Where Z = poverty line; y_i = income of the i th person; q = the number of poor; n = total population; α = parameter that determines sensitivity to distribution of welfare among the poor. The n people in the population are ranked by welfare from poorest to richest: $i = (1, 2... q..n)$.

When $\alpha = 0$, the FGT measure collapses to the Headcount ratio or the percentage of the population that is below the poverty line. This measure can give estimates of how many of the poor should be served by poverty programs, but is insensitive to differences in the depth of poverty. Suppose the poverty line is Rs. 100, there are ten people in the economy and two are poor. The Headcount index will give the same result ($P_0 = 0.2$) if there are two poor people with welfare (measured by, say, consumption expenditure) of Rs. 95 as it would with two incomes of Rs. 5; yet clearly, in the latter case poverty is more severe.

When $\alpha = 1$, the FGT index becomes the Poverty Gap, a measure of the depth of poverty. This measures the total consumption shortfall as a percentage of the poverty line. Thus, in the case of the two poor people with consumption expenditure of Rs. 95, $P_1 = 0.01$. With two poor people earning Rs. 5, P_1 would be 0.19. This implies that P_1 will increase *even if* the number of the poor is the same, if there is a welfare reduction in one poor household.

The drawback to the Poverty Gap measure is that it will estimate the poverty to be the same when one poor person has an income of Rs. 90 and the other an income of Rs. 10 as it would when both have an income of Rs. 50. Yet many would argue that the former situation represents lower welfare if the *distribution* of welfare among the poor is a concern. This is overcome for $\alpha > 1$. For instance, when $\alpha=2$, the first case gives $P_2 = 0.082$ and the second gives 0.025. In other words, P_2 will register an increase *even if* the average welfare of the poor is the same, if there is a transfer from one poor household to another poor household that is relatively better off. The drawback to using $\alpha = 2$ is that the measure is hard to interpret.

³² Based largely on Grosh (1994), Box 3.1, p. 25