CENTRE D'ETUDES
ET DE RECHERCHES
SUR LE DEVELOPPEMENT
INTERNATIONAL

SERIE ETUDES ET DOCUMENTS DU CERDI

Finding the Best Indicators to Identify the Poor

Adama Bah

Etudes et Documents n° 24

November, 2013

CERDI
65 BD. F. MITTERRAND
63000 CLERMONT FERRAND - FRANCE
TÉL. 04 73 17 74 00
FAX 04 73 17 74 28
www.cerdi.org

The author

Adama Bah

Clermont Université, Université d'Auvergne, CNRS, UMR 6587, CERDI, F-63009 Clermont Fd, France and The National Team for the Acceleration of Poverty Reduction (TNP2K), Indonesia Vice-President Office
Email: adamabah@gmail.com

This work was supported by the LABEX IDGM+ (ANR-10-LABX-14-01) within the program "Investissements d'Avenir" operated by the French National Research Agency (ANR)

La série des *Etudes et Documents* du CERDI est consultable sur le site :

http://www.cerdi.org/ed

Directeur de la publication : Patrick Plane

Directeur de la rédaction : Catherine Araujo Bonjean

Responsable d'édition : Annie Cohade

ISSN : 2114 - 7957

Acknowledgments

I would like to thank Joppe de Ree, Julia Tobias, Sam Bazzi and Sudarno Sumarto for providing helpful comments on previous drafts, and Maciej Czos for editing the paper. Any remaining errors are my sole responsibility. Support for this paper has been provided by the Australian Government through the Poverty Reduction Support Facility (PRSF). The findings, interpretations and conclusions expressed in this material do not necessarily reflect the views of the Government of Indonesia or the Government of Australia.

Abstract

Proxy-means testing (PMT) is a method used to assess household or individual welfare level based on a set of observable indicators. The accuracy, and therefore usefulness of PMT relies on the selection of indicators that produce accurate predictions of household welfare. In this paper I propose a method to identify indicators that are robustly and strongly correlated with household welfare, measured by per capita consumption. From an initial set of 340 candidate variables drawn from the Indonesian Family Life Survey, I identify the variables that contribute most significantly to model predictive performance and that are therefore desirable to be included in a PMT formula. These variables span the categories of household private asset holdings, access to basic domestic energy, education level, sanitation and housing. A comparison of the predictive performance of PMT formulas including 10, 20 and 30 of the best predictors of welfare shows that leads to recommending formulas with 20 predictors. Such parsimonious models have similar predictive performance as the PMT formulas currently used in Indonesia, although these latter are based on models of 32 variables on average.

Keywords: Proxy-Means Testing, Variable/Model Selection, Targeting, Poverty, Social Protection.
JEL codes: I38, C52.

"The essential ingredients [of specification searches] *are judgment and purpose, which jointly determine where in a data set one ought to be digging and also which stones are gems and which are rocks."*

E. E. Leamer (1978)

Introduction

Proxy-means tests (PMT) have been increasingly used to identify poor households in developing countries. PMT deals with the following problem, which is a typical situation in most developing countries. There is no official registry that contains accurate and up to date information on household and/or individual revenue. Self-reported income is therefore unverifiable, and it can be potentially time consuming and costly to identify the poor. Proxy-means testing (PMT) allows assessing household welfare based on observable indicators.

The implementation of PMT requires two distinct data sources. First, a household survey containing information on household expenditures (or income) and socioeconomic characteristics is required. It is used to estimate the correlation between household consumption, and a set of observable characteristics based on simple regression methods.[1] The set of weights (coefficients) derived from these consumption regressions provides a scoring formula, which is used to compute "PMT scores", or predicted consumption, in a targeting survey.[2] A targeting survey, or "census of the poor", is administered to all households or individuals considered as potentially eligible for social protection programs. It collects information on socioeconomic indicators that enter the PMT formula estimated using the household consumption survey.

The accuracy, and therefore the usefulness of PMT for targeting social programs, depends critically on the quality of the indicators available in the targeting survey: they need to be good predictors of welfare

[1] See Sumarto *et al.* (2007) for a discussion of the consumption regression approach used for PMT in comparison with alternative approaches to estimating household welfare with the objective to identify the poor.
[2] A similar approach is also used to develop "small area estimation poverty maps", following the method proposed by Elbers *et al.* (2003). The difference is that the second data source used for poverty maps is typically a population census.

or poverty. In addition, such predictors need to be easy to collect and verify,[3] as well as limited in number, given that targeting survey interviews should be short. In the context of the implementation of a PMT method for targeting social protection programs, the key issue is therefore to identify among a large set of candidate variables the predictors that are cost-effective to be included in the targeting survey because they are reliable in producing good predictions of household welfare.

It is not straightforward to identify good predictors of welfare. For a number K of candidate variables, the data set does not "admit a unique set of inferences" (Leamer 1985). Instead, the model space is "infinite dimensional" (Leamer 1983): there are 2^K possible models (with different sets of predictors) that can be estimated. When K is larger than 30, there are billions of possible models: it is impossible to find the single best one. Moreover, among all possible models, there are a large number of models that provide good predictions, which is acknowledged in the method developed in this chapter. Through random sampling of models from the entire model space, this method allows first assessing the level of predictive performance that can be expected from a good model, and then identifying which variables are more likely to produce such models.

The model random sampling method relates to the "sensitivity analyses" advocated by Leamer (1985) to address model uncertainty.[4] The robustness of a variable is assessed here by evaluating whether changes in the set of predictors with which this variable is combined lead to differences in the contribution of this variable to model predictive performance. In the literature, a number of analyses of the sensitivity of model parameters to changes in assumptions, in particular to changes in the set of predictors, have been conducted using the extreme-bounds tests proposed by Leamer[5] – see *e.g.* Levine and Renelt (1992).

[3]Since this data is collected for the purpose of potentially providing program benefits, respondents might be tempted to give answers that increase their chances of receiving these benefits. This is an even greater risk when the targeting system is being updated, as households might have made the link between being surveyed and receiving benefits. The indicators selected for the targeting survey should therefore be easily verifiable by the enumerators and the risk of measuring them with error should be low.

[4] Leamer discusses the need for sensitivity analyses to assess the robustness of the empirical relationship between a given variable or set of variables and the outcome of interest. The sensitivity analysis approach and debate have mostly been limited to the macroeconomic literature, in particular the cross-country growth literature.

[5] Leamer (2010) considers however that the extreme-bounds analysis proposed in Leamer (1983) has been "poorly understood and inappropriately applied."

However, this approach has been criticized, for instance by Granger and Uhlig (1990) and Sala-i-Martin (1997), for giving the same importance (weight) to all models. Indeed, some models are obviously not "likely to be the true model" (Sala-i-Martin 1997), in the sense that they have a weak performance. Therefore, the robustness of a variable should not be rejected based on such models. Instead it is recommended to "restrict attention to better fitting models" (Sala-i-Martin *et al.* 2004), which is the approach adopted in this chapter.

I implement the model random sampling method on a set of 340 indicators drawn from the 2007 Indonesian Family Life Survey. These indicators are subsequently ranked according to their probability of being included in good predictive models and to their contribution to the predictive performance of these models. I find as good predictors of welfare, and therefore useful in both a targeting survey and a PMT formula, variables that span the categories of private asset holdings, access to basic domestic energy, education level, sanitation and housing. The model random sampling method leaves the decision regarding the predictors to be used in both the targeting survey and the PMT formulas to the researcher and/or the policymaker. Yet, it provides them with useful information to make this decision in a transparent manner, such as the expected increase in predictive performance from the inclusion of a given variable.

The prediction accuracy that can be expected from PMT formulas that include the best predictors of welfare is discussed, to illustrate the use of the results in terms of predictor ranking. In particular, comparing the predictive performance of models with 10, 20 and 30 of the best variables, I find overall that good predictions are obtained with 20 predictors and therefore recommend using such parsimonious models. These results contribute to the literature on targeting social programs by further demonstrating that targeting using PMT has inherent relatively important errors, even when including a large number of good predictors. Therefore, it is recommended that PMT is used in combination with other targeting methods, in line with Coady *et al.* (2004), especially in rural areas. Moreover, these results suggest that a targeting system developed based on PMT should include a mechanism to address grievances regarding households that are wrongly excluded or included in the list of beneficiaries of a social program.

The remainder of the paper is organized as follows. Section 2 presents the Indonesian Family Life Survey (IFLS), the data source used for this analysis. Section 3 discusses existing models selection methods and describes the model sampling method proposed here. Section 4 presents the results in terms of the variables that are identified as good predictors of welfare. Section 5 discusses the predictive performance of PMT formulas including 10, 20 and 30 of these predictors. Section 6 concludes with a discussion of the implications of the findings in terms of the indicators - and models - that will contribute to improve targeting accuracy in Indonesia.

Section 2 – The Indonesian Family Life Survey

The Indonesian Family Life Survey (IFLS)[6] is a large-scale longitudinal survey which provides extensive information on households that are representative for about 83% of the Indonesian population living in 13 provinces in 1993.[7] This paper uses the cross-sectional data of the 2007 wave of the survey, the IFLS4, which has a sample size of 12,945 households.

Compared to the SUSENAS, national socioeconomic survey conducted annually by the Indonesian National Statistics Office (*Badan Pusat Statistik*, BPS) which covers more than 200,000 households, the IFLS contains more detailed information on households and individuals. This serves better the purpose of identifying good predictors of poverty and welfare. The IFLS is composed of 11 household books, of

[6] More information on the IFLS is available on the RAND Corporation website (http://www.rand.org/labor/FLS/IFLS.html).

[7] The IFLS is a longitudinal survey. The sampling scheme for the first round IFLS1, collected in 1993, has therefore determined the sampling in subsequent rounds – IFLS2 in 1997, IFLS3 in 2000 and IFLS4 in 2007 – which follow the original IFLS1 households and their splitoffs. The IFLS1 surveyed 7,224 households and more than 22,000 individuals. They were selected based on the sampling scheme of the 1993 Susenas, nationally representative socioeconomic survey of about 60,000 households, which stratified on provinces, then on urban-rural areas within provinces. The 13 provinces were selected not only to maximize the representation of the population, but also to capture the cultural and socioeconomic diversity of Indonesia in a cost-effective way, given the size and terrain of the country. Within the 13 provinces, 321 enumeration areas (EAs) were randomly selected, with an oversampling of urban EAs and EAs in smaller provinces to facilitate urban–rural and Javanese–non-Javanese comparisons.

which 4 are at the household level and 7 at the individual level. Book K provides information on household composition at the time of the survey and on the dynamics in the household demographics. Book 1 and Book 2 provide information on household expenditures and socioeconomic characteristics such as housing characteristics, household businesses (farm and nonfarm), private assets, and non-labor income. Book 3A and 3B are answered to by at least 1 household member aged 15 and above, and they collect data on individual characteristics, among which education level, health, community participation and employment. Book 5 is administered to children aged below 15; it provides information on school participation, health and labor participation. Books US1 and US2 collect data on physical health, including weight, height and other health-related measurements for all household members.

Using these 8 books,[8] I construct 340 variables which are *potentially good* predictors of welfare and/or poverty, at the household or individual level. Many of these candidate variables are discrete and measure the same phenomenon. For instance, a dummy is created for each type of wall or floor, in order to allow disentangling the specific types that matter the most for predicting welfare.[9] Overall, the candidate variables can be grouped for convenience into 14 categories which refer to different manifestations of poverty, based on the literature on poverty and welfare: demographics, demographic dynamics, education level, school participation, literacy, health status, nutrition, employment, housing, basic domestic energy services,[10] sanitation, private assets, business assets and community participation. The indicators grouped in the categories demographics, education level, school participation, literacy and employment are calculated separately for the household head and for the other household members, by gender and age group, where relevant and possible.

[8] The 3 household books that are not used in this paper are the tracking book (T), the one answered by ever-married women (4) and the one on cognitive assessment (EK).

[9] See the Appendix for the full list of variables. More details on their definition, as well as summary statistics are available upon request.

[10] Basic domestic energy is the basic energy or energy services needed to achieve standard daily living requirements. It includes lighting, cooking, heating or cooling, as well as energy to provide basic services linked to health, education and communications (*e.g.* drinking water).

The welfare indicator I use is the logarithm of real household per capita expenditures, computed as the sum of food and non-food expenditures (excluding durable goods) divided by household size and adjusted for price differences using provincial urban/rural poverty lines. I use the 2000 IFLS provincial urban/rural poverty lines from Strauss *et al.* (2004) which are based on the poverty lines for February 1999 calculated by Pradhan *et al.* (2001). These 2000 IFLS lines are inflated to reflect 2007 prices using the inflation rate of the official province urban/rural poverty lines between 2000 and 2007. Lastly, expecting that the characteristics of the poor are different across urban and rural areas, I estimate consumption regression models separately for each area.[11]

Section 3 – Variable and model selection using a random sampling method[12]

The accuracy of PMT scores is negatively affected by the use of weak predictors of welfare for their construction. When eligibility to social protection programs relies exclusively on PMT scores, this is problematic. The media in Indonesia, for example, regularly feature troubles related to the allocation of key social protection programs such as *Raskin* ("Rice for the Poor"), the subsidized rice program. The imperfect relationship between PMT scores and actual household welfare is one of the reasons for these allocation problems. In this section, I first discuss existing approaches to selecting indicators to be used in a PMT formula (PMTF) before presenting the model random sampling method as an alternative for identifying the best predictors of welfare.

[11] In urban areas, the sample size is of 6,984 households, and 5,961 households in rural areas. All regressions include cross-sectional sampling weights from Strauss *et al.* (2009).

[12] Variable selection and model selection are used interchangeably in this paper. It is considered that the appropriate model is identified through specification search, which consists in selecting the set of explanatory variables that is appropriate given the objective of the empirical researcher, following Leamer (1978).

3.1 Existing model selection methods

In the academic literature on PMT, as well as for the practical implementation of PMT around the world,[13] variables used for estimating PMTF are usually selected through manual or automated specification searches. In the manual approach, a set of initial candidate predictors is first selected, based on their anticipated correlation with welfare or poverty, on their availability in consumption surveys, and on the ease and accuracy with which they can be collected. Per capita expenditures are then regressed on this list of *candidate* predictors and their *a posteriori* significance level in the full regression is considered to select the *final* set of predictors (see *e.g.* Ahmed and Bouis 2002).

More common are automated selection procedures, such as stepwise, meant to reduce the number of predictors. Such procedures have been used in PMT simulations (*e.g.* Grosh and Baker 1995 and Sharif 2009), as well as recently for Indonesia's targeting schemes (World Bank 2012). They are easily implemented using standard econometric softwares, which is convenient when there is a large set of initial candidate predictors. However, "stepwise procedures are not intended to rank variables by their importance" and are "not able to select from a set of variables those that are most influential" as they tend to be unstable as to which variables are included in the final model (James and McCulloch 1990). The whimsical[14] nature of stepwise procedures seems to depend notably on the degree of correlation between the initial candidate predictors, as shown by Derksen and Keselman (1992). Furthermore, models identified by such automated procedures are more subject to chance features of the data and frequently fail to predict as accurately when applied to samples other than the estimation sample (Judd and McClelland 1989).

With traditional approaches to specification searches, whether manual or automated, the focus is on the models; the variables that should be included in the targeting survey are those that are selected in the final

---

[13] See e.g. Glewwe and Kanaan (1989), Grosh and Baker (1995), Ahmed and Bouis (2002), Narayan and Yoshida (2005) and Sharif (2009) for PMT simulations for various countries; Castaneda and Lindert (2005) for a review of the experience of Latin America countries.

[14] This term is used by Leamer (1983) to refer to the lack of robustness of econometric results to basic changes in specifications or functional forms.

models. However, these approaches do not "find" the best model according to some objective criterion such as the (adjusted) R2. Instead, such methods select one model among many good (or excellent) models and dismiss the rest. As a result, slight changes - in the sample, in the set of candidate variables or in the implementation procedure - may lead to models that include different sets of predictors and yet appear equally good in terms of fit or predictive performance.[15] In Indonesia, PMTFs have been developed for each of the 497 districts based a combination of manual and stepwise procedures. These PMTFs include different sets of predictors for different districts, which leads to an overall large number of indicators to be collected in a targeting survey.[16] These procedures are not appropriate to assess whether there is a smaller set of indicators that would provide similar predictions across all districts. They are therefore not ideal for identifying which of the 340 candidate variables work best in predicting welfare, and should therefore be included in a targeting survey, which is the purpose of this paper.

3.2 A model random sampling method for selecting PMT indicators

The imperfect world of specification searches is one in which one must work with models that are good at best (according to the adjusted R2, or any other criterion of predictive performance), rather than with a single best model among all possible models. In fact, there are usually a very large number of good models that can be constructed given an initial set of predictors. The first step in the proposed new approach is establishing what the general characteristics of "good models" are, before identifying the variables that are included in such models in a second step.

[15] This issue has been referred to as the model uncertainty issue in the literature (see *e.g.* Leamer 1978, Leamer 1985, Temple 2000): there are several models which are good, in the sense that they provide estimates which are validated according to the diagnostic test results, but these models yield different conclusions regarding not only the variables but also their parameters. For a further discussion of issues arising from traditional model selection approaches, see *e.g.* Raftery (1995).

[16] The Indonesian targeting survey, the Data Collection for Social Protection Programs (*Pendataan Program Perlindungan Sosial* – PPLS), which was administered to more than 25 million households in 2011 to establish the national targeting system, collects information on over 50 indicators. Note however that the procedure to develop the 2011 PPLS questionnaire was different from the procedure discussed here since the questionnaire was designed and implemented before the 497 PMTFs were developed.

As mentioned above, the primary interest of this paper is to evaluate the performance of each of the 340 preselected candidate variables in predicting welfare. Designing a targeting survey questionnaire requires the identification of a limited number of indicators which will provide good predictions of welfare. I therefore focus for convenience on linear predictive models with a fixed size of 10 variables.[17]

With 340 candidate predictors – which are identified prior to any model or variable selection – one can construct approximately $\frac{340!}{330!10!} = 2 \times 10^{19}$ different linear prediction models of 10 variables. The model space is far too large to evaluate all models. Instead 10-variable models are randomly selected from the entire model space (of models with 10 variables).

The two-step model random sampling method

The first step aims to gain insight on what a good model is. I sample at random from the space of 10-predictor models a large number, s_1, of models of the form:

$$\text{Ln } PCE = \alpha + \beta_k X_k + \varepsilon, \quad \text{where } k = 1,..,10 \qquad (1)$$

For each model sampled, the measure of predictive performance, the R-squared (R2)[18] is stored. The sampling distribution of R2 approximates the population (true) distribution of R2 when s_1 becomes large and therefore can be used to gauge the distribution of predictive performance across the entire model space.[19] From this sample of s_1 models, I divide the estimated R2 in 1000 quantiles, or "permilles".[20] This allows classifying subsequent models into different permilles. Models appearing in the top permille are those that produce a good fit and that are therefore considered as good for predictions.

[17] Sala-i-Martin *et al.* (2004), who also develop a method for ranking the predictors of economic growth by their "importance", argue that fixing model size is "easy to interpret, easy to specify and easy to check for robustness". Model size can be straightforwardly extended if needed, and can in practice depend, for instance, on the budget and time available for the survey, or on the need to identify indicators that fit on a 1-page questionnaire (or on a scorecard) for the targeting survey. Robustness checks are conducted for model sizes of 5, 20 and 30, see section 4.

[18] There are other possible measures for model predictive performance, such as the Akaike or Bayesian information criteria (AIC, BIC). However, as argued by Granger and Uhlig (1990), the R-squared is a "relevant statistic and some exact results are achievable using it." Furthermore, since models of equal size are compared, similar results would be obtained with other measures, including the adjusted R-squared.

[19] s_1 is chosen here to be 300,000 but it can be readily increased; results in terms of the R2 distribution converge already with s_1=60,000.

[20] Note that in theory this approach can be easily extended to estimate 10,000 or 100,000 quantiles. However, the sample size (of models) needs to be increased accordingly, knowing that this takes processing time.

The second step focuses on the characteristics of these "good" models and aims to identify which of the 340 candidate variables appear often in the best models, i.e., a model from the top 0.1%. After having estimated R2 permilles in the first step, I again proceed to randomly select models of the form (1) from the model space, this time with the objective to evaluate the performance of each single variable X_k, when combined with a random sample of 9 additional covariates. Intuitively, good variables are those that contribute to model predictive performance, regardless of the other controls they are paired with.

In this second step, I randomly sample s_2 models from the model space.[21] For each randomly selected model *i*, it is first evaluated within which permille its R2 falls ($\mathrm{R}^2_{\mathrm{q}-1} < \mathrm{R}^2_{\mathrm{i}} < \mathrm{R}^2_{\mathrm{q}}, q \in (1; 1000)$). Then, information is collected on the 10 variables that compose model *i*, their added value and coefficient sign. Based on all models including a given variable X_k, three indicators of performance are computed for each variable. The first indicator of performance is the probability of inclusion of variable X_k in a good model. This indicator measures the probability of inclusion for each predictor when a model from the top permille is randomly selected. Variables that occur frequently in top predictive models are good variables. (Note that when this probability is close to 1, the variable is necessary to obtain a good predictive power.)

$$P(X_k \text{ included} \mid \mathrm{R}^2_{\mathrm{q}-1} < \mathrm{R}^2_{\mathrm{i}} < \mathrm{R}^2_{\mathrm{q}}),\ q \in (1; 1000) \qquad (2)$$

Intuitively, (2) is the fraction of models of a certain predictive performance level that include the preselected variable X_k.

The second indicator of performance used to asses each candidate predictor is its *added value*, conditional on inclusion in a model of a certain predictive performance. The added value is defined as the difference between the R2 with and without the variable X_k included in the regression model. It measures how much each variable "adds" to the predictive power of the model. Variables that add a lot, and especially those that add a lot to good models, are good variables.

[21] s_2 is chosen here to be 1,000,000. On average, when estimating a sample s_2=1,000,000 random models, each candidate predictor is selected in about 3,000 models. Similar results in terms of variable ranking are obtained when estimating s_2=100,000 random models.

The third indicator of variable performance considered is the sign of the coefficients of variable X_k in all models in which it is included. Variables that are robustly related to poverty or consumption should have the same sign, regardless of the other covariates included. In other words, a good predictor of welfare not only has a high contribution to the R2 of good models but also a correlation with welfare that is constantly either positive or negative.

3.3 Advantages of the model random sampling method

The model random sampling method provides several interesting pieces of information useful to the policymaker interested in implementing an effective PMT, especially when the pieces of information are combined.

The starting point of this method is the estimation of the distribution of model performance across the model space. It gives a sense of the predictive performance of any model that can be expected given the data. Secondly, instead of delivering a single "best" model according to a certain criterion, it specifically acknowledges that it is practically impossible to find the single best model. Instead, the model random sampling method establishes a benchmark to make an assessment of the quality of models.

The method also allows assessing the relationship between the individual performance of a given variable and model performance. The final variable ranking appears robust to the number of initial predictors or their degree of correlation. This presents the advantage of allowing comparisons of good predictors of welfare across time, location and welfare/poverty definitions. It can therefore be expected to provide reliable predictions, especially outside of the estimation sample.

Section 4 - Results: the best predictors of welfare

In this section I first present the results in terms of the estimated R2 distribution across the model space, and then discuss the specific variables that are identified as good predictors of welfare, both separately for urban and rural areas. Note that the dependent variable used, (the logarithm of) per capita expenditures, has automatically a strong correlation with (the logarithm of) household size, which obtains an inclusion probability approaching 1 in the top urban and rural models. The results presented here therefore rely on the estimation of randomly selected models in which (the logarithm of) household size is automatically included - with a randomly selected subset of 9 other predictors.[22]

4.1 The empirical distribution of predictive performance over the model space

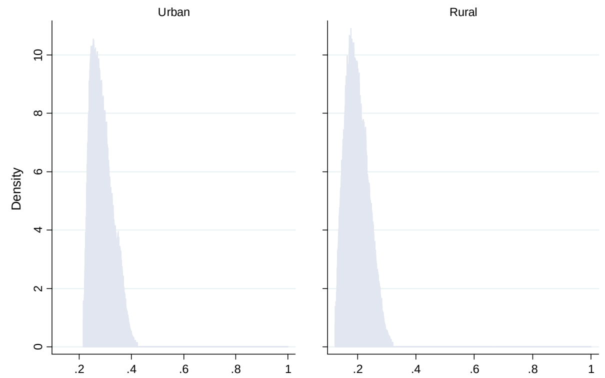
Figure 1 shows the R2 distribution over the space of 10-predictor models for urban and rural areas. The distribution of R2 across the model space shows why targeting using PMT has inherent errors: 99.9% of models (of 10 predictors) predict less than 42% of the variation in actual household per capita expenditures in urban areas, and less than 33% in rural areas.[23] Model performance varies widely between urban and rural areas. In urban areas, on average models of 10 predictors have an R2 of 0.29, whereas in rural areas the average R2 of 10-predictor models is about 0.20. It suggests that it might be easier to identify poor households in urban areas than in rural areas based on the prediction of their per capita expenditures using observable characteristics. This is possibly explained by households tending to share similar socioeconomic characteristics in rural areas more than in urban areas, which makes the distinction between poor and non-poor households based on observables more complex.[24]

[22] Robustness checks are conducted without the forcing of household size. The results for the ranking of the remaining predictors, available upon request, are the same.

[23] When applying the random sampling method to models of 30 predictors for instance, it is found that 99.9% of these models have an R2 lower than 0.50 in urban areas and 0.40 in rural areas.

[24] Another reason for the lower predictive performance of rural models might be the fact that the poorest households are located in rural areas, whereas PMT does not allow targeting accurately enough the poorest of the poor. This stems from Ordinary Least Squares (OLS)-based predictions being inaccurate at the bottom of the consumption distribution (Grosh and Baker 1995).

Figure 1 - R-squared distribution over the space of models of 10 predictors



Note: R2 distribution estimated from 100,000 models randomly selected from the model space.

4.2 The best predictors of household welfare

Tables 1a and 1b show the list of predictors that have a probability greater than 0.05 to be included in the top 0.1% models and that have a consistent sign probability, ranked in descending order of their inclusion probability, in urban and rural areas respectively.[25] Variable inclusion probabilities in the top 0.1% models are reported in column (1); the conditional probabilities to obtain a model which performance classifies it among the top 0.1% when a given variable is included - referred to henceforth as model probability - are reported in column (2). Variable added values, expressed as the average share of added value in model R2, and their coefficient sign probabilities are reported in columns (3) and (4) respectively. They are both conditional on each variable being included in a top 0.1% model. Tables 1a and 1b show that there is a high correlation between variable inclusion or model probability and their added values.

The best predictors of welfare appear to belong to different variable categories, spanning from education and employment to housing and asset ownership.[26] In urban areas, the first three variables appear significantly better than all other variables: they have a probability higher than 40% to be included in a top 0.1% model. In such models, these variables contribute on average to increasing the R2 by more than

[25] See Annex 1.1 for the full list of variables, and a similar assessment of their performance in the top 0.1% models.
[26] Variable categories that are not among the best predictors of welfare include variables that are less easy to collect and verify such as demographic dynamics, health status and nutrition.

12%. In rural areas, one variable, cooking with wood, is significantly better than all other variables: more than 80% of the top 0.1% models include this variable. Furthermore, cooking with wood contributes to nearly one-fifth of the R2 of these best models. This corresponds to an increase of about 6 percentage points in the R2 of the top 0.1% rural models. Interestingly, it is also the only variable among the best predictors that has a negative correlation with household welfare.[27]

Table 1a: Best predictors for the top 0.1% urban models

Variables	Inclusion prob. (1)	Model prob. (2)	Added value (3)	Sign prob. (4)
(D) Asset: Fridge	0.631	0.215	0.123	1
(Log) Avg years of schooling in the HH	0.470	0.160	0.121	1
(D) HHH max education level: university	0.432	0.147	0.120	1
Nb of rooms in the house	0.206	0.070	0.079	1
(D) HH cooks with gas	0.201	0.068	0.078	1
(D) HH cooks with wood	0.128	0.043	0.070	0
(Log) house floor size	0.121	0.041	0.064	1
(D) Max. education level in HH: university	0.103	0.035	0.047	1
(D) Floor type: ceramic	0.086	0.029	0.043	1
(D) Drinking water source: mineral water	0.076	0.026	0.037	1
(D) Toilet: own with septic tank	0.072	0.024	0.033	1
(D) Asset: Vehicle	0.064	0.022	0.035	1
(D) >=1 HHM aged >15 enrolled in school	0.064	0.022	0.030	1
(D) Non drinking water source inside the house	0.063	0.022	0.032	1
(D) Garbage disposed in trash can	0.057	0.019	0.031	1
(D) Drinking water source: well	0.053	0.018	0.029	0
(D) HHH max education level: primary school	0.052	0.018	0.028	0
(D) Asset: Receivables	0.051	0.017	0.028	1

Note: HH stands for household, HHH for household head and HHM for household member; >=1 stands for at least 1. (D) indicates dummy variables. In all random models, the dependent variable is household adjusted per capita expenditures, and household size is included. The sign probabilities (probability that the variable has a positive sign) and added values (average difference between the R-squared of the 10-variable model - including the variable of interest - and the R-squared of the 9-variable model - excluding the variable of interest - as a share of the average R-squared of the top 0.1% 10-predictor models) are conditional on the variable being included. The model probability is the probability that the model is among the top 0.1% models, conditional on including a given variable.

[27] In both urban and rural areas, most of the best predictors, such as fridge ownership for instance, seem to allow differentiating households that are rich, rather than the poor. This may also explain why PMT does not perform well in targeting accurately the poorest of the poor. Note that the candidate variables constructed from categorical variables such as the type of wall or of cooking fuel include dummies for all categories. There is no missing category (among those specified on the questionnaire).

Table 1b: Best predictors for the top 0.1% rural models

Variables	Inclusion prob. (1)	Model prob. (2)	Added value (3)	Sign prob. (4)
(D) HH cooks with wood	0.823	0.280	0.187	0
(D) Asset: TV	0.301	0.102	0.122	1
(D) Asset: Fridge	0.268	0.091	0.110	1
(D) Asset: HH Appliances	0.264	0.090	0.112	1
Nb of rooms in the house	0.189	0.064	0.104	1
(Log) house floor size	0.171	0.058	0.095	1
(Log) Avg years of schooling in the HH	0.136	0.046	0.086	1
(D) Asset: Vehicle	0.103	0.035	0.058	1
(D) Max. education level in HH >= senior sec.	0.092	0.031	0.059	1
(D) HH cooks with gas	0.085	0.029	0.073	1
(D) Non-farm business asset: 4-wheel vehicle	0.071	0.024	0.054	1
Nb HHM aged 15-64	0.069	0.024	0.052	1
(Log) Max. years of schooling in the HH	0.067	0.023	0.057	1
(D) Non drinking water source inside the house	0.066	0.022	0.043	1
(D) HHH max education level: university	0.060	0.020	0.049	1
(D) Toilet: own with septic tank	0.059	0.020	0.046	1
(D) HHM primary job status: gvt employee	0.058	0.020	0.051	1
(D) Drinking water source: mineral water	0.051	0.017	0.041	1
(Log) HH avg per capita annual earnings	0.051	0.017	0.033	1
(D) Max. education level in HH: university	0.050	0.017	0.048	1

Note: HH stands for household, HHH for household head and HHM for household member; >=1 stands for at least 1. (D) indicates dummy variables. In all random models, the dependent variable is household adjusted per capita expenditures, and household size is included. The sign probabilities (probability that the variable has a positive sign) and added values (average difference between the R-squared of the 10-variable model - including the variable of interest - and the R-squared of the 9-variable model - excluding the variable of interest - as a share of the average R-squared of the top 0.1% 10-predictor models) are conditional on the variable being included. The model probability is the probability that the model is among the top 0.1% models, conditional on including a given variable.

Similarly to the R2 distribution, there appears to be best predictors that are specific to urban areas, and others specific to rural areas, in addition to some variables that overlap between the two areas. This suggests that obtaining good predictions of household welfare in Indonesia requires estimating PMT formulas separately for urban and rural areas at least.[28] The predictors that appear only in the top 0.1% rural models are variables that allow distinguishing households whose living is not exclusively dependent on farming and agriculture and that have diversified income sources (*e.g.* ownership of non-farm business

[28] Implementing the model random sampling method at a more disaggregated geographic level would similarly allow identifying whether good predictors of welfare are specific to provinces or even districts, and therefore whether it is desirable to develop PMTFs for each of these areas.

assets, being a government employee, the number of working-aged household members). In addition, in rural areas ownership of private assets such as a TV or other appliances, which could be categorized as non-vital or “convenience” assets, appear to contribute largely to distinguish households by their welfare status.

In urban areas, are among the best predictors of welfare variables that are closely related to the location of the dwelling, such as the type of floor, garbage disposal in a trash can or drinking water from a well. These variables are likely to be correlated with dwellings located in slums, or disadvantaged neighborhoods within cities; they describe the inequality in access to basic services and sanitation, as well as the use of low quality construction material. Other predictors that are specific to urban areas relate to the education participation of household members aged 15 and above, as well as ownership of financial assets.

Robustness checks are conducted to assess whether the variable rankings obtained are sensitive to the number of predictors selected in the random models or to the number of models estimated in the first and second steps of the method. The results in terms of predictor ranking, available upon request, are not affected by such changes in the implementation of the method. The model random sampling method provides a ranking for all predictors in terms of their probability of being included in the top performing models and their performance in explaining the variation in consumption or poverty (see Appendix). The identification of good predictors of welfare is the first step in the estimation of a PMT formula (PMTF). The second step relates to how to combine them in a PMTF, with practical considerations such as the number of predictors to include. The next section discusses the predictive performance of PMTFs using the good predictors of welfare.

Section 5 – The performance of proxy-means test formulas using the best predictors

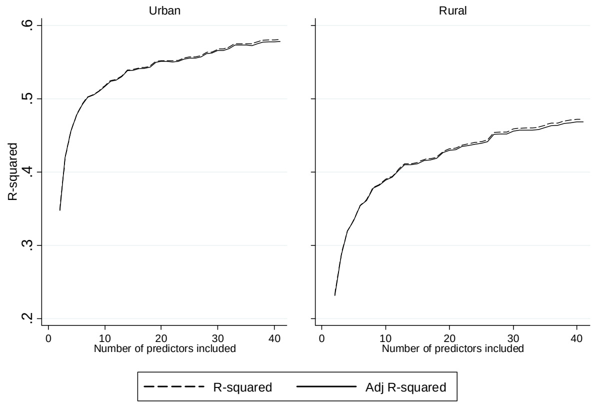
Identifying the number of variables to include in the final PMTF involves a trade-off between (i) the completeness of information, in order to explain as much of the variation in consumption or poverty as possible, and (ii) the restriction of the number of variables, to limit the costs of collecting accurate data. Furthermore, formulas with a high number of predictors are more subject to being not as valid within as outside the estimation sample, due to their increased variance of predictions or to an over-fitting problem. In this section, I compare PMTFs with 10, 20 and 30 of the best predictors of welfare, in terms of their fit, as well as prediction errors and targeting incidence, to illustrate that this tradeoff between completeness of information and model parsimony is not always in favor of the first.

5.1 Model fit

Figure 2 shows the observed increase in R2 and adjusted R2 obtained when increasing the number of predictors one-by-one, in descending order of their inclusion probability in the top 0.1% 10-predictor models. In urban areas the R2 reaches 0.50 with only 6 predictors. In rural areas, an R2 of 0.39 is obtained with 10 predictors, and increasing the number of predictors to 40 yields an R2 higher by less than 10 percentage points, at about 0.47. This shows that the marginal returns of adding more predictors decreases in both areas, especially after about 10 predictors.

The difference between the R2 and the adjusted R2 of models of all sizes is hardly distinguishable, including with 40 predictors. This suggests that even using the adjusted R2 as selection criterion leads to selecting models with a large number of variables. Based on Figure 2, models of 40 predictors (or more) should be selected in both areas as they have the highest adjusted R2.

Figure 2 - Evolution of the R-squared with increases in the number of predictors



5.2 Model prediction errors

In this section, I assess whether gains in terms of fit from increasing the number of predictors also leads to lower prediction errors, both in and out of sample. Prediction errors are measured by Type I errors, or undercoverage, and Type II errors, or leakage rates. Undercoverage refers to households that are in the target population based on their actual welfare, but are predicted to be above the eligibility cutoff. Leakage refers to households wrongly predicted to have a welfare level that is below the program eligibility cutoff whereas their actual welfare is above this cutoff. I use the predicted 30th percentile as eligibility cutoff, which amounts to comparing the actual and predicted expenditure deciles.[29] I also calculate the severe undercoverage and severe leakage rates. The former refers to households whose expenditures classify them in the first actual expenditures decile but that are predicted to be above the 30th percentile eligibility cutoff point; the latter refers to households that are predicted in the first expenditure decile whereas their actual expenditures classify them above the 30th percentile.

[29] Using the predicted 30th expenditure percentile as eligibility cutoff, which is the practice adopted in Indonesia, shifts the focus on household relative position within the distribution as predicted by the PMTF. This has the main advantage to allow offsetting the errors inherent to OLS predictions, which, by shrinking the consumption distribution, tend to predict higher expenditures for the poorest households. In addition, it allows better planning for programs, since it ensures that the number of households identified as eligible is closer to the expected coverage, in this case the poorest 30%.

In-sample prediction errors of models with 10, 20 and 30 predictors are shown in Panel A of Table 2. Undercoverage and severe undercoverage appear lower in urban areas, across all models. Models of 20 predictors are performing best in terms of undercoverage and severe undercoverage, in both areas. About 11% of the poorest 10% are predicted above the 30th percentile with a 20-predictor model in urban areas. Leakage and severe leakage rates appear marginally lower with 30-predictor models, especially in rural areas.

Table 2: Prediction accuracy of urban and rural models of different sizes - full and cross-validation samples.

	Urban			Rural		
Nb of Predictors	10	20	30	10	20	30
Panel A: Full Sample Results						
R-squared	0.518	0.552	0.568	0.39	0.432	0.459
Adjusted R-squared	0.517	0.551	0.566	0.389	0.43	0.456
Severe Undercoverage	15.6	11.1	16	28.4	25.5	26.8
Undercoverage	29	25	29.7	43	42	42.8
Leakage	48.4	45.5	45.3	33	32.1	29.5
Severe Leakage	29.4	29.8	27.5	19.3	17.8	16.6
Panel B: Half Specification Sample Results						
R-squared	0.515	0.549	0.567	0.39	0.433	0.462
Adjusted R-squared	0.514	0.546	0.563	0.388	0.429	0.456
Severe Undercoverage	16.9	13.2	18	28.1	24.9	26.3
Undercoverage	28.3	25.6	29.3	43.8	42.5	43.8
Leakage	48.9	46.5	45.8	33	31.6	29.3
Severe Leakage	31.2	31.5	28.9	20.8	18.4	17.2
Panel C: Half Test (out) Sample Results						
Predicted R-squared	0.513	0.546	0.525	0.385	0.437	0.475
Severe Undercoverage	16	12.9	15	29.1	26.7	26.5
Undercoverage	28.7	25.7	30.3	42.3	42.1	42.1
Leakage	47.3	45	45	33.1	32.8	29.9
Severe Leakage	28.1	28.7	26.6	18.9	17	15.6

Note: undercoverage refers to Type I error; severe undercoverage refers to the share of households in the poorest 10% that are predicted to be above the eligibility cutoff. Leakage refers to Type II error, and severe leakage refers to the share of households that are predicted in the first decile whereas their actual expenditures place them above the third decile. All error rates calculated by comparing the predicted with the actual expenditure deciles, with the predicted 30th percentile as eligibility cutoff. The figures for the specification and test samples are average values over 3 samples randomly drawn from the full sample.

In addition to full sample predictions, and in order to reduce the risk of selecting a model that overfits the data, cross-validation tests are conducted. Models are estimated using one half of the sample – estimation or specification sample - and the formulas generated are used to estimate predicted consumption in the other half of the sample – test sample.[30] This procedure allows mimicking the real-world situation where PMT models are estimated using a consumption survey and applied to calculate the predicted consumption of households surveyed in the targeting survey. The cross-validation procedure also provides a test of the stability of the models when estimated with a smaller sample. The prediction results for the specification and test samples are shown in Panels B and C of Table 2. All models produce similar results in terms of R2 and prediction error rates, both with a smaller estimation sample and out-of-sample. However, the performance of the urban model with 30 predictors can appear slightly less robust, since the predicted R2 out of sample is 4 percentage points lower than the R2 obtained with the full and half specification samples, and the difference in severe undercoverage and leakage rates is also slightly higher than for other models.

Table 3: Overall prediction errors at different cutoff points.

		Predicted 30th	Actual 30th	Predicted 40th
10-Var	Coverage	30	23	40
	Undercoverage	39	49	30
	Leakage	40	34	31
20-Var	Coverage	30	24	40
	Undercoverage	37	46	29
	Leakage	38	34	30
30-Var	Coverage	30	23	40
	Undercoverage	37	47	29
	Leakage	37	32	30

Note: the 30th and 40th predicted eligibility cutoffs compare households in the actual and predicted expenditure deciles, while the actual 30th percentile eligibility cutoff focuses on households whose predicted consumption is below the actual 30th consumption percentile.

Table 3 presents the national level targeting prediction results, based on the estimation of urban and rural models of different sizes. Undercoverage and leakage are based on two cutoffs, the actual and the

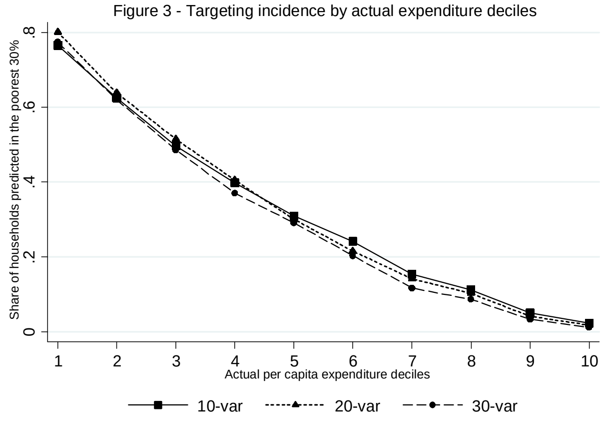
[30] The cross-validation tests are conducted on three different randomly drawn estimation and test samples in order to ensure the robustness of the test conclusions.

predicted 30th expenditure percentiles.[31] When using the actual 30th percentile as eligibility cutoff, the coverage is lower, about 23%, which leads to higher undercoverage and lower leakage compared to the predicted 30th percentile. Table 3 shows that combined errors rates from models of 20 and 30 predictors are very similar, and at the 40th percentile all 3 models produce similar error rates.

5.3 Targeting incidence

Lastly, I consider the combined targeting incidence and the distribution of predictions errors of urban and rural models of 10, 20 and 30 predictors. Targeting incidence refers to the share of households in each decile of the actual expenditure distribution that are predicted to be below the eligibility cutoff. The distribution of prediction errors focuses on how households mis-predicted to be eligible or non-eligible are distributed across actual consumption deciles. The idea is indeed that undercoverage is less of a serious problem if households falsely excluded are close to the cutoff as opposed to at the very bottom of the distribution; similarly, leakage is less grave if it includes households that are just above the cutoff compared to households that are at the highest end of the distribution.

Figure 3 - Targeting incidence by actual expenditure deciles



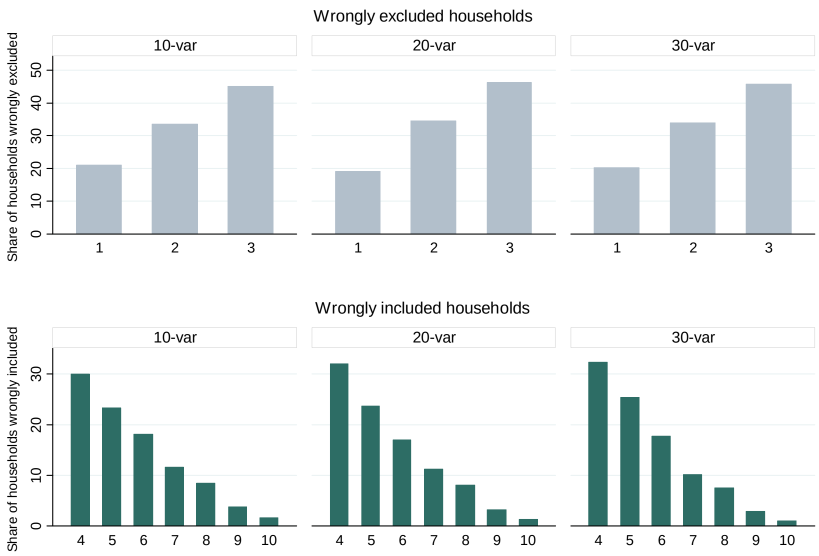
[31] The actual 30th expenditure percentile has been used as eligibility cutoff in other PMT simulation studies (*e.g.* Grosh and Baker 1995; Narayan and Yoshida 2005).

Even though deriving PMTFs induces rather significant errors in identifying the poor, targeting appears highly progressive, and households that are wrongly excluded or included largely tend to be classified in deciles that are relatively close to the eligibility cutoff, for all models. Figure 3[32] shows that with all models the share of households predicted below the 30th percentile decreases rapidly when going up in the deciles. The main gain in terms of increases in the share of poor households that are identified as beneficiaries appears when going from 10 predictors to 20 predictors, although it is a relatively small gain. The 30-predictor models perform similarly to the 10- and 20-predictor models overall, except for the shares of households from the higher deciles, which are slightly lower with the 30-predictor models. Figure 4 shows that with all models more than 40% of households that are wrongly predicted to be above the eligibility cutoff are actually in the third decile, and more than 50% of households that are wrongly predicted to be below the predicted 30th expenditure percentile are actually in the fourth and fifth deciles. Figure 4 also confirms that models of 20 predictors exclude slightly fewer households from the first decile, whereas models of 30 predictors perform slightly better in terms of inclusion errors, which concern more households from the fourth and fifth deciles.

Using 10, 20 or 30 predictors appears to produce relatively similar overall predictive performance. In other words, models of 30 predictors are not significantly better, especially in terms of prediction errors and targeting incidence. Based on the findings of this section, using models with 20 predictors therefore seems a good trade-off between prediction accuracy and model parsimony. This is in line with Grosh and Baker (1995), who highlight that increasing the number of predictors has diminishing returns in terms of the probability of collecting inaccurate information and thus the costs of verification. With 2 models, one for urban and one for rural areas, using 20 predictors implies to collect information on 28 indicators overall, which appears reasonable.

[32] When disaggregating the targeting incidence and the distribution of inclusion and exclusion errors from urban and rural models, a trend similar to that observed in Figures 3 and 4 is obtained, although, as discussed previously, urban models produce better predictions.

Figure 4 - Actual decile distribution of wrongly excluded and included households



Section 6 – Recommendations and concluding remarks

Leamer (2010) advises that "it would be much healthier for all of us [economists] is we could accept our fate, recognize that perfect knowledge will be forever beyond our reach and find happiness with what we have." In this chapter, it is acknowledged that it is impossible to find the best model to predict household welfare. Instead, I propose a new method for selecting the predictors to use in PMT formulas, based on the estimation of the distribution of predictive performance across the model space and on the assessment of the sensitivity of the contribution of each candidate variable to model performance.

I focus here on the recommendations based on the findings for Indonesia, where a PMT approach has been used for identifying beneficiaries of social protection programs since 2005. Most recently, in 2011, the newly established Unified Database for Social Protection Programs[33] (UDB) is also based on PMT.

[33] The Unified Database, which is managed by the National Team for the Acceleration of Poverty Reduction (Tim Nasional Percepatan Penanggulangan Kemiskinan – TNP2K) under the Office of Indonesia's Vice-President, is a national registry for identifying potential beneficiaries of social protection programs. Over 25 million households, or 96 million individuals, have been surveyed using the targeting survey PPLS11 and are registered in the Unified

As mentioned earlier, PMTF specific to each of the 497 Indonesian districts were developed using the national socioeconomic survey, SUSENAS, and applied to data collected using the targeting survey, 2011 Data Collection for Social Protection Programs (*Pendataan Program Perlindungan Social* - PPLS). The findings of this paper provide useful recommendations for the updating of the UDB, planned for 2014.

The correlates identified using the IFLS data confirm the multidimensional aspect of welfare and poverty: the predictors shown to be robust predictors for welfare belong to different variable categories, spanning from education and employment to housing and asset ownership. Most of these variables are included in both the SUSENAS and PPLS. The first recommendation is therefore to add the few good predictors that are not yet included in the Susenas and PPLS questionnaires, such as the number of rooms in the house, the garbage being disposed of in a trash can and the distinction between private and business assets. Moreover, the emphasis should be put on ensuring that the data collected is of good quality. This is particularly important for the variables with the highest added values in predicting households' socioeconomic status. For collecting data on household size and on other demographic indicators such as the number of members aged between 15 and 64, it is thus recommended to use as much as possible official documents when they are available to the respondents to complete the questionnaires. For other critical variables, in the education level, housing and access to energy services categories in particular, enumerators have to be carefully trained to collect quality information on these variables.

A combination of manual specification searches and stepwise procedure was used to select the indicators that entered the 497 PMTF developed for the UDB. The method proposed in this paper can be used as an alternative to identify – for each district or at a higher geographical level – good predictors to be used for the updating of the PMTF in 2014. The implementation procedure adopted here can be relatively easily automated using standard statistical softwares. Further, the robustness checks carried out show that the ranking of predictors is consistent, including with the estimation of a lower number of random models.

Database – it is the largest database of its kind in the world. More information about the Unified Database is available on http://bdt.tnp2k.go.id.

Regarding the predictive performance of the PMTFs, the results obtained in this paper are not directly comparable with the models that have been developed for ranking households in the UDB, since they were estimated using the SUSENAS instead of the IFLS, and since a model was developed for each district. However, the average performance and size of these district-specific PMTF can be compared for illustrative purposes with the ones developed here. On average, the district PMT formulas currently used in Indonesia have 32 variables and are based on consumption regression models that have an R-squared of 0.5. These models have predicted undercoverage and leakage rates at the predicted 40^{th} percentile cutoff of about 30%. The 10-, 20- and 30-predictor urban and rural models developed in this paper have been shown to lead to similar prediction error rates. According to World Bank (2012), there are significant gains in terms of targeting accuracy in estimating PMT formulas at greater levels of geographical disaggregation, with the greatest gain obtained when going from provincial- to district-level models. This suggests that using the model random sampling approach to identify the best predictors of welfare could potentially allow both using a lower number of predictors and improving the accuracy of the PMT-based welfare predictions used for targeting in Indonesia.

References

Ahmed A., and H. Bouis (2002): “Weighing What’s Practical: Proxy-Means Tests for Targeting Food Subsidies in Egypt”, *Food Policy*, 27: 519-540.

Castaneda T., and K. Lindert (2005): “Designing and Implementing Household Targeting Systems: Lessons from Latin America and the United States”, Social Protection Discussion Paper Series No. 0526”, Washington DC: The World Bank.

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

Derksen S., and H. J. Keselman (1992): “Backward, Forward and Stepwise Automated Subset Selection Algorithms: Frequency of Obtaining Authentic and Noise Variables”, *British Journal of Mathematical and Statistical Psychology*, 45(2): 265–282.

Elbers C., Lanjouw J. O. and P. Lanjouw (2003): "Micro-Level Estimation of Poverty and Inequality," *Econometrica*, 71(1): 355-364.

Glewwe P. and Kanaan, O. (1989): "Targeting assistance to the poor using household survey data", Policy Research Working Paper Series 225, Washington DC: The World Bank.

Granger C. W. J. and H. F. Uhlig (1990): "Reasonable Extreme-Bounds Analysis," *Journal of Econometrics*, 44(1-2): 159-70.

Grosh M. (1994): *Administering Targeted Social Programs in Latin America: From Platitudes to Practice*. Washington DC: The World Bank.

Grosh M., and J. Baker. (1995)" "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation", Working Paper No. 118, Living Standards Measurement Study, Washington DC: The World Bank.

James F. C., and C. E. McCulloch (1990): "Multivariate Analysis in Ecology and Systematics: Panacea or Pandora's Box?" *Annual Review of Ecology, Evolution, and Systematics* 21: 129-166.

Judd C. M., and G. H. McClelland (1989): *Data Analysis: A Model Comparison Approach*. San Diego, CA: Harcourt, Brace, Jovanovitch.

Leamer E. E. (1978): *Specification Searches: Ad Hoc Inference with Non-experimental Data*, New York: Wiley, 1978.

Leamer E. E. (1983): "Let's take the Con Out of Econometrics", *American Economic Review*, 73(3): 31-43.

Leamer E. E. (1985): "Sensitivity Analysis Would Help", *American Economic Review*, 75: 308-313.

Leamer E. E. (2010): "Tantalus on the Road to Asymptopia," *Journal of Economic Perspectives*, 24(2): 31-46.

Levine, R. and D. Renelt (1992). A Sensitivity Analysis of Cross-Country Growth Regressions, *American Economic Review*, 82(4): 942-963.

Narayan A., and N. Yoshida. (2005): "Proxy Means Test for Targeting Welfare Benefits in Sri Lanka", PREM Working Paper 33258, South Asia Region, Washington DC: The World Bank.

Pradhan M., Suryahadi A., Sumarto S., and L. Pritchett (2003): "Eating Like Which "Joneses?" An Iterative Solution to the Choice of a Poverty Line "Reference Group"". *Review of Income and Wealth* 47 (4): 473–487.

Raftery A. E. (1995): “Bayesian model Selection in Social Research,” *Sociological Methodology*, 25: 111-163.

Sala-i-Martin X. (1997): “I Just Ran Two Million Regressions”, *American Economic Review*, 87(2): 178-183.

Sala-i-Martin X., Doppelhofer G., and R. I. Miller (2004): “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach.” *American Economic Review*, 94(4): 813-835.

Sharif, I. (2009): “Building a Targeting System for Bangladesh based on Proxy Means Testing”, Social Protection Discussion Paper No. 0914, Washington DC: The World Bank.

Strauss J., Beegle K., Dwiyanto A., Herawati Y., Pattinasarany D., Satriawan E., Sikoki B., Sukamdi and F. Witoelar (2004): *Indonesian Living Standards Before and After the Financial Crisis: Evidence from the Indonesia Family Life Survey*, Santa Monica, CA: RAND Corporation, 2004.

Strauss J., Witoelar F., Sikoki B. and A.M. Wattie (2009): “The Fourth Wave of the Indonesia Family Life Survey (IFLS4): Overview and Field Report”, RAND Labor and Population, April 2009. WR-675/1-NIA/NICHD.

Sumarto S., Suryadarma D. and A. Suryahadi (2007): “Predicting Consumption Poverty using Non-Consumption Indicators: Experiments using Indonesian Data,” *Social Indicators Research*, 81(3): 543-578.

Temple J. (2000): “Growth Regressions and What the Textbooks Don’t Tell You”, *Bulletin of Economic Research* 52 (3): 181-205.

World Bank (2012): *Targeting poor and vulnerable households in Indonesia*. Public expenditure review (PER). Washington D.C. - The World Bank. http://documents.worldbank.org/curated/en/2012/01/15879773/targeting-poor-vulnerable-households-indonesia

Appendix: Ranking of all candidate predictors – by category

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Basic energy services								
(D) HH cooks with gas	**0.201**	**0.068**	**0.078**	**1.00**	0.085	0.029	0.073	1.00
(D) HH cooks with wood	**0.128**	**0.043**	**0.070**	**0.00**	**0.823**	**0.280**	**0.187**	**0.00**
(D) Drinking water source: mineral water	0.076	0.026	0.037	1.00	0.051	0.017	0.041	1.00
(D) Non drinking water source inside the house	0.063	0.022	0.032	1.00	0.066	0.022	0.043	1.00
(D) Drinking water source: well	0.053	0.018	0.029	0.00	0.033	0.011	0.025	0.00
(D) HH pays for drinking water	0.044	0.015	0.021	1.00	0.042	0.014	0.025	1.00
(D) HH cooks with kerosene	0.039	0.013	0.065	0.05	0.055	0.019	0.056	0.40
(D) Non drinking water source: well	0.037	0.013	0.025	0.00	0.016	0.006	0.008	0.06
(D) Drinking water source inside the house	0.035	0.012	0.016	1.00	0.034	0.011	0.034	1.00
(D) Non drinking water source: pipe	0.031	0.010	0.020	1.00	0.019	0.007	0.005	0.95
(D) Same drinking and non drinking water source	0.027	0.009	0.015	0.08	0.018	0.006	0.010	0.00
(D) Drinking water source: pump	0.026	0.009	0.010	0.00	0.015	0.005	0.004	0.94
(D) HH pays for drinking water delivered	0.021	0.007	0.006	1.00	0.022	0.008	0.005	0.96
(D) Drinking water source: improved (MDG)	0.020	0.007	0.006	0.10	0.023	0.008	0.003	0.96
(D) HH pays for non drinking water delivered	0.020	0.007	0.003	0.95	0.011	0.004	0.001	0.82
(D) Drinking water source: pipe	0.019	0.007	0.003	0.95	0.016	0.006	0.003	1.00
(D) electricity	0.019	0.007	0.003	1.00	0.031	0.010	0.021	1.00
(D) Drinking water source: river/creek	0.018	0.006	0.001	0.00	0.017	0.006	0.002	0.00
(D) Drinking water source: spring	0.017	0.006	0.000	0.88	0.018	0.006	0.001	0.16
(D) Non drinking water source: river/creek	0.016	0.006	0.001	0.00	0.014	0.005	0.004	0.00
(D) Drinking water source: collection bassin	0.016	0.006	0.001	0.00	0.013	0.004	0.002	0.00
(D) Non drinking water source: rain	0.016	0.006	0.000	0.13	0.016	0.006	0.002	1.00
(D) HH pays for non drinking water	0.015	0.005	0.005	1.00	0.021	0.007	0.005	0.95
(D) Non drinking water source: collection bassin	0.015	0.005	0.001	0.00	0.012	0.004	0.004	0.00
(D) Non drinking water source: pump	0.013	0.005	0.003	0.08	0.033	0.011	0.016	1.00
(D) HH cooks with electricity	0.013	0.005	0.001	0.92	0.013	0.005	0.001	0.14
(D) Non drinking water source: pond	0.013	0.005	0.000	0.23	0.016	0.006	0.001	0.00
(D) HH cooks with charcoal	0.012	0.004	0.000	0.33	0.016	0.006	0.002	0.00
(D) Drinking water source: pond	0.010	0.003	0.000	0.00	0.011	0.004	0.001	0.00
(D) Non drinking water source: spring	0.007	0.002	0.000	0.86	0.009	0.003	0.001	0.00
(D) Drinking water source: boiled or mineral	0.006	0.002	0.003	1.00	0.016	0.006	0.006	1.00
Private assets								
(D) Asset: Fridge	**0.631**	**0.215**	**0.123**	**1.00**	**0.268**	**0.091**	**0.110**	**1.00**
(D) Asset: Vehicle	0.064	0.022	0.035	1.00	**0.103**	**0.035**	**0.058**	**1.00**
(D) Asset: Receivables	0.051	0.017	0.028	1.00	0.036	0.012	0.038	1.00
(D) Asset: HH Appliances	0.043	0.015	0.032	1.00	**0.264**	**0.090**	**0.112**	**1.00**
(D) Asset: TV	0.042	0.014	0.025	1.00	**0.301**	**0.102**	**0.122**	**1.00**
(D) Asset: Jewelry	0.041	0.014	0.027	1.00	0.030	0.010	0.021	1.00

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(D) Asset: Second house	0.036	0.012	0.023	1.00	0.023	0.008	0.016	1.00
(D) Asset: Land	0.019	0.007	0.003	0.74	0.013	0.004	0.002	0.77
(D) Asset: Savings	0.019	0.007	0.002	0.74	0.018	0.006	0.002	0.58
(D) Asset: Land	0.018	0.006	0.009	1.00	0.019	0.007	0.019	1.00
(D) Other assets	0.014	0.005	0.004	1.00	0.017	0.006	0.008	1.00
(D) Asset: Livestock	0.010	0.003	0.002	0.40	0.020	0.007	0.010	1.00
(D) Asset: HH furniture	0.010	0.003	0.000	0.20	0.015	0.005	0.001	1.00
Education level								
(Log) Avg years of schooling in the HH	**0.470**	**0.160**	**0.121**	**1.00**	**0.136**	**0.046**	**0.086**	**1.00**
(D) HHH max education level: university	**0.432**	**0.147**	**0.120**	**1.00**	0.060	0.020	0.049	1.00
(Log) Max. years of schooling in the HH	0.110	0.038	0.073	0.93	0.067	0.023	0.057	1.00
(D) Max. education level in HH: university	**0.103**	**0.035**	**0.047**	**1.00**	0.050	0.017	0.048	1.00
(D) HHH max education level: primary school	0.052	0.018	0.028	0.00	0.020	0.007	0.013	0.00
(D) HHH max education level: senior sec. and above	0.047	0.016	0.040	0.83	0.020	0.007	0.013	0.95
(D) Max. education level in HH: senior sec. and above	0.044	0.015	0.021	1.00	**0.092**	**0.031**	**0.059**	**1.00**
(D) HHH max education level: senior secondary	0.038	0.013	0.050	0.89	0.030	0.010	0.015	1.00
(D) >=1 HHM aged >15 ever attended school	0.032	0.011	0.008	0.39	0.012	0.004	0.013	0.83
(D) >=1 HHM left school before 15	0.030	0.010	0.015	0.00	0.021	0.007	0.010	0.00
(D) Max. education level in HH: primary	0.028	0.009	0.022	0.00	0.044	0.015	0.042	0.00
(D) Max. education level in HH: senior secondary	0.028	0.009	0.007	0.44	0.030	0.010	0.020	0.84
(D) >=1 HHM left school before 12	0.027	0.009	0.008	0.04	0.010	0.003	0.014	0.00
(D) Max. education level in HH: junior secondary	0.022	0.008	0.010	0.00	0.021	0.007	0.003	0.23
(D) HHH has no education	0.021	0.007	0.011	0.24	0.023	0.008	0.018	0.04
(D) >=1 HHM left school after 18	0.021	0.007	0.008	1.00	0.024	0.008	0.026	1.00
(D) >=1 HHM aged <15 has ever attended school	0.019	0.007	0.001	0.74	0.019	0.007	0.002	0.10
(D) HHH max education level: junior secondary	0.013	0.005	0.004	0.15	0.016	0.006	0.001	0.12
(D) >=1 HHM aged <15 entered PS at 10 or above	0.013	0.005	0.000	1.00	0.012	0.004	0.000	0.00
(D) >=1 HHM aged <15 entered PS at 8 or above	0.012	0.004	0.000	0.33	0.011	0.004	0.001	0.00
(D) >=1 HHM aged <15 entered PS at 7 or below	0.010	0.003	0.001	0.50	0.016	0.006	0.001	0.24
Housing								
Nb of rooms in the house	**0.206**	**0.070**	**0.079**	**1.00**	**0.189**	**0.064**	**0.104**	**1.00**
(Log) house floor size	**0.121**	**0.041**	**0.064**	**1.00**	**0.171**	**0.058**	**0.095**	**1.00**
(D) Floor type: ceramic/marble/granite/stone	**0.086**	**0.029**	**0.043**	**1.00**	0.038	0.013	0.036	1.00
(D) Floor type: cement/brick	0.032	0.011	0.016	0.00	0.015	0.005	0.001	0.19
(D) Type of wall: bamboo/woven/mat	0.029	0.010	0.007	0.00	0.033	0.011	0.017	0.00
(D) House status: rented/contracted	0.025	0.008	0.006	1.00	0.022	0.008	0.010	1.00
(D) House building: single unit multiple levels	0.021	0.007	0.012	1.00	0.026	0.009	0.009	1.00
(D) Type of roof: metal plates	0.021	0.007	0.007	1.00	0.042	0.014	0.036	1.00
(D) House status: self-owned	0.021	0.007	0.003	0.90	0.016	0.006	0.001	0.35
(D) Type of roof: asbestos	0.020	0.007	0.002	0.05	0.013	0.005	0.001	1.00

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(D) Type of roof: bamboo/grass/foliage	0.020	0.007	0.000	0.15	0.009	0.003	0.001	0.33
(D) Type of roof: concrete	0.019	0.007	0.001	1.00	0.018	0.006	0.002	1.00
(D) Floor type: dirt	0.018	0.006	0.008	0.00	0.048	0.016	0.031	0.00
(D) Type of roof: tiles/shingles	0.018	0.006	0.004	0.06	0.042	0.014	0.044	0.00
(D) Floor type: lumber/board	0.018	0.006	0.001	0.78	0.017	0.006	0.004	0.78
(D) House status: occupied	0.017	0.006	0.009	0.00	0.013	0.004	0.003	0.00
(D) House building: duplex unit single level	0.017	0.006	0.000	0.06	0.012	0.004	0.000	0.08
(D) Type of wall: masonry	0.016	0.006	0.002	1.00	0.021	0.007	0.011	1.00
(D) House building: duplex unit multiple levels	0.016	0.006	0.001	1.00	0.015	0.005	0.000	0.31
(D) House building: multiple units & levels	0.016	0.006	0.001	0.25	0.014	0.005	0.000	0.27
(D) House building: multiple unit single level	0.016	0.006	0.000	0.38	0.013	0.004	0.002	1.00
(D) House building: single unit & level	0.014	0.005	0.008	0.00	0.014	0.005	0.004	0.00
(D) House built on stilt	0.014	0.005	0.001	1.00	0.013	0.005	0.001	0.43
(D) Floor type: tiles/terrazzo	0.013	0.005	0.005	0.00	0.015	0.005	0.002	0.13
(D) Type of wall: lumber/board/plywood	0.013	0.005	0.000	0.15	0.017	0.006	0.002	0.56
(D) Floor type: bamboo	0.012	0.004	0.000	0.00	0.013	0.005	0.001	0.00
(D) Type of roof: wood	0.011	0.004	0.002	1.00	0.010	0.003	0.003	1.00
Sanitation								
(D) Toilet: own with septic tank	0.072	0.024	0.033	1.00	0.059	0.020	0.046	1.00
(D) Garbage disposed in trash can, collected by sanitation service	0.057	0.019	0.031	1.00	0.015	0.005	0.006	1.00
(D) Toilet: Creek/river/ditch	0.041	0.014	0.017	0.00	0.032	0.011	0.020	0.00
(D) House has ventilation	0.029	0.010	0.012	1.00	0.021	0.007	0.005	1.00
(D) House yard well kept	0.027	0.009	0.012	1.00	0.013	0.004	0.006	1.00
(D) Piles of trash around House	0.027	0.009	0.003	0.00	0.013	0.005	0.003	0.00
(D) Sewage: Permanent pit	0.025	0.008	0.000	0.88	0.013	0.004	0.001	0.77
(D) House has a kitchen outside	0.021	0.007	0.003	0.00	0.013	0.005	0.003	0.00
(D) House in stagnant water	0.020	0.007	0.006	0.00	0.012	0.004	0.003	0.00
(D) Sewage: Disposed in yard/garden	0.020	0.007	0.001	0.00	0.014	0.005	0.003	0.00
(D) Toilet: animal stable	0.020	0.007	0.001	0.00	0.016	0.006	0.001	1.00
(D) Sewage: Sea, beach	0.020	0.007	0.001	0.15	0.013	0.005	0.000	0.93
(D) Garbage disposed into river/creek/sewer	0.019	0.007	0.003	0.00	0.014	0.005	0.004	0.00
(D) Toilet: pond/fishpond	0.018	0.006	0.000	1.00	0.008	0.003	0.000	0.00
(D) Sewage: Paddy field/other field	0.018	0.006	0.000	0.72	0.013	0.005	0.001	1.00
(D) Toilet: sea/lake	0.018	0.006	0.000	0.78	0.008	0.003	0.000	0.00
(D) Toilet: shared	0.017	0.006	0.005	0.12	0.014	0.005	0.001	0.07
(D) House w/ 1 room for cooking and sleeping	0.017	0.006	0.003	0.00	0.013	0.004	0.007	0.00
(D) Garbage disposed in pit	0.017	0.006	0.001	0.00	0.013	0.005	0.001	0.86
(D) Sewage: Disposed into river	0.016	0.006	0.002	0.00	0.018	0.006	0.003	0.05
(D) House is next/under a stable	0.016	0.006	0.002	0.13	0.014	0.005	0.000	0.67

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(D) Garbage disposed in sea, lake, beach	0.016	0.006	0.001	0.25	0.014	0.005	0.000	1.00
(D) Garbage disposed in paddy field/other field	0.015	0.005	0.000	0.00	0.015	0.005	0.001	0.00
(D) House has medium-sized yard	0.014	0.005	0.004	1.00	0.010	0.003	0.010	1.00
(D) Toilet: public	0.014	0.005	0.004	0.00	0.017	0.006	0.001	0.00
(D) Sewage: Drainage ditch (stagnant)	0.014	0.005	0.000	0.29	0.016	0.006	0.000	0.41
(D) Sewage: Drainage ditch (flowing)	0.012	0.004	0.003	1.00	0.021	0.007	0.007	1.00
(D) Human/animal waste near house	0.012	0.004	0.001	0.00	0.013	0.005	0.003	0.00
(D) Sewage: Hole (without permanent lining)	0.012	0.004	0.001	0.00	0.023	0.008	0.000	0.25
(D) Garbage diposed in forest, mountain	0.012	0.004	0.000	0.00	0.008	0.003	0.000	0.75
(D) Garbage disposed in yard and let decompose	0.011	0.004	0.003	0.00	0.017	0.006	0.001	0.00
(D) Toilet: own without septic tank	0.011	0.004	0.001	0.09	0.014	0.005	0.002	0.07
(D) Toilet: sewer	0.011	0.004	0.000	0.00	0.015	0.005	0.002	0.00
(D) Garbage burned	0.009	0.003	0.004	0.11	0.013	0.005	0.001	1.00
(D) Sewage: Pond/fishpond/lake/pool	0.009	0.003	0.000	0.00	0.013	0.005	0.000	0.07
(D) Toilet: yard/field	0.009	0.003	0.000	0.33	0.012	0.004	0.001	1.00
Demographics								
Nb HHM aged 15-64	0.037	0.013	0.021	0.97	0.069	0.024	0.052	1.00
Dependency ratio	0.030	0.010	0.016	0.03	0.042	0.014	0.040	0.00
(D) HH size = 1	0.029	0.010	0.024	1.00	0.020	0.007	0.008	0.95
(D) HHH aged <=30	0.027	0.009	0.001	0.27	0.019	0.007	0.002	0.15
(Log) Max. age in the HH	0.023	0.008	0.003	0.13	0.017	0.006	0.004	0.11
(D) >=1 separated HHM	0.023	0.008	0.002	0.00	0.009	0.003	0.000	1.00
(Log) HHH age	0.022	0.008	0.001	0.59	0.020	0.007	0.003	0.76
Child dependency ratio	0.021	0.007	0.009	0.10	0.029	0.010	0.025	0.00
(D) Household size <= 4	0.020	0.007	0.008	0.00	0.018	0.006	0.007	0.00
Nb HHM aged 16-18	0.020	0.007	0.008	1.00	0.012	0.004	0.003	0.83
(D) Household size >= 10	0.020	0.007	0.002	1.00	0.016	0.006	0.001	0.88
(Log) Avg age in the HH	0.020	0.007	0.002	0.40	0.015	0.005	0.003	0.69
(D) Household size >= 6	0.019	0.007	0.008	1.00	0.017	0.006	0.010	1.00
(D) HHH aged >=65	0.019	0.007	0.005	0.00	0.023	0.008	0.009	0.00
(D) >=1 married HHM	0.019	0.007	0.004	0.00	0.015	0.005	0.001	0.88
Adult (elderly) dependency ratio	0.018	0.006	0.008	0.00	0.020	0.007	0.014	0.05
Nb HHM aged > 64	0.017	0.006	0.005	0.06	0.022	0.008	0.007	0.04
Nb HHM aged 7-15	0.017	0.006	0.004	1.00	0.014	0.005	0.002	0.60
Nb HHM aged 5-12	0.017	0.006	0.003	0.76	0.015	0.005	0.003	0.19
(D) >=1 HHM divorced	0.017	0.006	0.000	0.00	0.014	0.005	0.001	0.00
Nb HHM aged <5	0.016	0.006	0.003	0.13	0.020	0.007	0.010	0.00
(D) Widowed HHH	0.015	0.005	0.003	0.00	0.011	0.004	0.005	0.00
Nb HHM aged < 15	0.015	0.005	0.003	0.73	0.020	0.007	0.011	0.14
(D) Household size >= 8	0.014	0.005	0.004	1.00	0.014	0.005	0.006	1.00

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nb HHM aged 19-24	0.014	0.005	0.003	0.86	0.015	0.005	0.005	1.00
(D) >=1 widowed HHM	0.014	0.005	0.002	0.07	0.021	0.007	0.003	0.05
(D) >=1 HHM has more than 1 wive	0.014	0.005	0.000	0.86	0.011	0.004	0.003	1.00
(D) Male HHH	0.013	0.005	0.001	0.92	0.015	0.005	0.002	1.00
Nb HHM aged 13-15	0.012	0.004	0.001	0.92	0.022	0.008	0.001	1.00
(D) >=1 single HHM	0.010	0.003	0.001	0.60	0.011	0.004	0.005	0.00
(D) Married HHH	0.009	0.003	0.001	0.44	0.022	0.008	0.002	0.96
(D) >=1 male HHM	0.008	0.003	0.002	0.00	0.016	0.006	0.001	0.59
(D) Household size <= 2	0.007	0.002	0.002	1.00	0.014	0.005	0.006	1.00
Education participation								
(D) >=1 HHM aged >15 enrolled in school	0.064	0.022	0.030	1.00	0.032	0.011	0.021	1.00
Nb children in university	0.045	0.015	0.028	1.00	0.031	0.010	0.014	1.00
Nb HHM aged >15 enrolled in school	0.045	0.015	0.027	1.00	0.039	0.013	0.020	1.00
Nb HHM enrolled in school	0.025	0.008	0.014	1.00	0.016	0.006	0.004	0.82
Share of children aged 0-18 in senior sec. school	0.015	0.005	0.008	0.93	0.018	0.006	0.015	1.00
Nb HHM aged <15 enrolled in school	0.015	0.005	0.003	0.93	0.016	0.006	0.001	0.29
Nb children in primary school	0.014	0.005	0.002	0.64	0.013	0.005	0.001	0.14
(D) >=1 HHM aged <15 enrolled in school	0.013	0.005	0.001	1.00	0.017	0.006	0.001	0.28
Nb children in junior sec. school	0.013	0.005	0.001	0.92	0.009	0.003	0.002	1.00
Share of children aged 0-18 in junior sec. school	0.013	0.005	0.001	0.85	0.015	0.005	0.001	0.81
Share of children aged 0-18 in primary school	0.013	0.005	0.000	0.62	0.019	0.007	0.001	0.30
(D) >=1 HHM enrolled in school	0.010	0.003	0.011	1.00	0.013	0.004	0.002	0.92
Nb out-of-school children aged 7-15	0.008	0.003	0.000	0.38	0.014	0.005	0.002	0.00
Nutrition								
Avg body mass index (BMI) in the HH	0.042	0.014	0.026	1.00	0.030	0.010	0.047	1.00
Nb overweight (BMI>25) HHM	0.035	0.012	0.015	1.00	0.047	0.016	0.032	1.00
Nb underweight (BMI<18.5) HHM aged >20	0.030	0.010	0.011	0.00	0.013	0.005	0.010	0.00
(D) >=1 underweight HHM - BMI<18.5	0.029	0.010	0.016	0.00	0.026	0.009	0.022	0.00
Nb obese (BMI>30) HHM	0.029	0.010	0.007	0.96	0.024	0.008	0.005	1.00
Nb overweight (BMI>25) HHM aged >20	0.027	0.009	0.013	1.00	0.034	0.011	0.032	1.00
(D) >=1 HHM aged 20+ is under-weight (BMI<18.5)	0.027	0.009	0.010	0.00	0.025	0.009	0.013	0.00
(D) >=1 Adult HHM aged 20+ is overweight	0.025	0.008	0.010	0.96	0.028	0.009	0.024	1.00
(D) >=1 overweight HHM - BMI>25	0.025	0.008	0.009	0.92	0.032	0.011	0.020	0.91
(D) >=1 stunted (low height for age) child aged < 5	0.020	0.007	0.004	0.00	0.019	0.007	0.012	0.00
(D) >=1 wasted (low weight for height) child aged < 5	0.018	0.006	0.001	0.00	0.010	0.003	0.000	0.00
Nb underweight (BMI<18.5) HHM	0.017	0.006	0.004	0.00	0.023	0.008	0.022	0.00
(D) >=1 obese HHM - BMI>30	0.016	0.006	0.007	1.00	0.023	0.008	0.005	0.96
Nb underweight children aged < 5	0.013	0.005	0.003	0.00	0.008	0.003	0.005	0.00
Nb wasted children aged < 5	0.012	0.004	0.001	0.00	0.011	0.004	0.000	0.09
(D) >=1 child aged <5 w/ low weight for age5	0.011	0.004	0.003	0.00	0.017	0.006	0.005	0.00

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nb stunted children aged < 5	0.010	0.003	0.004	0.00	0.022	0.008	0.012	0.00
Business assets								
(D) Non-farm business asset: 4-wheel vehicle	0.041	0.014	0.026	1.00	0.071	0.024	0.054	1.00
(D) Non-farm business asset: building	0.037	0.013	0.014	1.00	0.045	0.015	0.032	1.00
(D) Non-farm business asset: land	0.023	0.008	0.011	1.00	0.024	0.008	0.020	1.00
(D) Non-farm business is own entirely by HH	0.023	0.008	0.010	1.00	0.041	0.014	0.032	1.00
(D) HH owns non-farm business assets	0.020	0.007	0.011	1.00	0.046	0.016	0.035	1.00
(D) Farm business asset: heavy equipment	0.019	0.007	0.000	1.00	0.020	0.007	0.010	1.00
(D) Non-farm business asset: others	0.017	0.006	0.011	1.00	0.037	0.012	0.032	1.00
(D) Farm business asset: small tools	0.016	0.006	0.003	0.00	0.018	0.006	0.003	0.89
(D) Farm business asset: vehicles	0.016	0.006	0.000	0.19	0.018	0.006	0.007	1.00
(D) Farm business asset: hard stem plants	0.015	0.005	0.001	0.87	0.028	0.009	0.022	1.00
(D) HH owns farm business assets	0.015	0.005	0.001	0.40	0.021	0.007	0.007	0.91
(D) Farm business asset: house/building	0.015	0.005	0.000	0.67	0.014	0.005	0.000	0.87
(D) Farm business asset: livestock/poultry	0.013	0.005	0.000	0.00	0.012	0.004	0.003	1.00
(D) Non-farm business asset: other vehicle	0.010	0.003	0.005	1.00	0.012	0.004	0.004	1.00
(D) Farm business asset: land	0.010	0.003	0.002	0.90	0.030	0.010	0.017	0.97
(D) Farm business asset: tractor	0.008	0.003	0.000	1.00	0.011	0.004	0.005	1.00
Earnings								
HHH annual earnings, % HH total annual earnings	0.038	0.013	0.004	0.00	0.020	0.007	0.007	0.19
(Log) HHH annual earnings	0.026	0.009	0.002	0.96	0.039	0.013	0.029	1.00
(Log) HH Avg HH per capita earnings	0.022	0.008	0.002	0.64	0.033	0.011	0.025	1.00
(Log) HH avg per capita annual earnings	0.019	0.007	0.003	1.00	0.051	0.017	0.033	1.00
(Log) Total HH annual earnings	0.015	0.005	0.003	1.00	0.042	0.014	0.032	1.00
(Log) Sum of all HHM annual earnings	0.012	0.004	0.002	0.58	0.017	0.006	0.022	1.00
(Log) Max. per capita earnings in the HH	0.012	0.004	0.001	0.58	0.036	0.012	0.024	1.00
Employment								
(D) >=1 HHM with no occupation (previous week)	0.032	0.011	0.016	0.00	0.017	0.006	0.011	0.00
(D) HHM primary job status: gvt employee	0.031	0.010	0.020	1.00	0.058	0.020	0.051	1.00
Nb HHM unoccupied in the past week	0.028	0.009	0.011	0.00	0.018	0.006	0.004	0.11
(D) >=1 HHM worked in the past week	0.027	0.009	0.002	0.12	0.020	0.007	0.007	0.95
Max total # of hrs worked last wk in additional job	0.025	0.008	0.000	0.38	0.013	0.004	0.002	1.00
(D) HHM primary job status: private worker	0.023	0.008	0.009	0.00	0.018	0.006	0.002	0.89
Nb HHM who work	0.022	0.008	0.009	1.00	0.032	0.011	0.021	1.00
(D) >=1 HHM worked in the past year	0.022	0.008	0.005	0.05	0.013	0.005	0.008	0.79
(D) >=1 HHM work in any family-owned business	0.021	0.007	0.004	0.90	0.021	0.007	0.013	0.95
(Log) HH Avg # of annual worked weeks - all jobs	0.020	0.007	0.003	0.30	0.023	0.008	0.011	0.96
(D) >=1 HHM works in a family farm business	0.020	0.007	0.002	0.25	0.016	0.006	0.004	0.82
Max total # of weeks worked per year in add. job	0.020	0.007	0.000	0.80	0.016	0.006	0.003	1.00
(D) HHM second. job status: private worker	0.019	0.007	0.005	1.00	0.019	0.007	0.013	1.00

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(Log) HH Avg # of worked weeks (past yr) - all jobs	0.019	0.007	0.002	0.42	0.019	0.007	0.015	0.95
Max normal # of hrs worked last wk in additional job	0.019	0.007	0.000	0.11	0.012	0.004	0.001	0.67
(D) >=1 HHM work in a fam-owned n-farm business	0.017	0.006	0.009	1.00	0.040	0.014	0.031	1.00
(Log) Max. HH # of work hrs (past week) - main job	0.018	0.006	0.002	0.33	0.014	0.005	0.008	1.00
Avg total # of hrs worked last wk in additional job	0.018	0.006	0.000	0.50	0.017	0.006	0.003	1.00
(D) >=1 HHM aged >15 not current. in the labor force	0.016	0.006	0.006	0.00	0.008	0.003	0.002	0.00
(Log) Max. HH # of annual worked weeks - main job	0.016	0.006	0.003	0.25	0.013	0.005	0.004	0.57
(D) >=1 HHM employed	0.015	0.005	0.004	0.20	0.013	0.005	0.005	0.86
Avg normal # of hrs worked last wk in additional job	0.015	0.005	0.000	0.13	0.012	0.004	0.001	0.83
(D) HHM second. job status: self employed	0.015	0.005	0.000	0.80	0.011	0.004	0.000	0.91
(D) >=1 HHM aged >15 never in the labor force	0.014	0.005	0.003	0.07	0.017	0.006	0.005	0.00
(D) >=1 HHM did not work in the past year	0.014	0.005	0.002	0.07	0.020	0.007	0.012	0.00
(D) HHM primary job status: unpaid family worker	0.014	0.005	0.001	0.93	0.014	0.005	0.001	0.73
(D) HHH works	0.014	0.005	0.001	0.64	0.019	0.007	0.005	0.85
(D) HHM second. job status: unpaid family worker	0.014	0.005	0.000	1.00	0.022	0.008	0.006	1.00
(D) HHH has no occupation (previous week)	0.013	0.005	0.004	0.00	0.014	0.005	0.004	0.20
(Log) Max. HH # of weekly working hours - main job	0.013	0.005	0.003	0.15	0.012	0.004	0.005	0.92
(Log) HH Avg # of annual worked weeks - main job	0.013	0.005	0.002	0.85	0.027	0.009	0.014	0.96
(D) HHM second. job status: gvt employee	0.013	0.005	0.001	1.00	0.017	0.006	0.006	1.00
Nb children <15 working in the past week	0.013	0.005	0.000	1.00	0.011	0.004	0.000	0.73
(Log) Max. HH # of weekly working hours - all jobs	0.012	0.004	0.003	0.17	0.021	0.007	0.007	0.68
(Log) Max. HH # of annual worked weeks - all jobs	0.012	0.004	0.003	0.25	0.014	0.005	0.005	0.87
(D) >=1 unemployed HHM	0.012	0.004	0.001	0.08	0.012	0.004	0.000	0.08
Avg total # of weeks worked per year in additional job	0.012	0.004	0.000	1.00	0.017	0.006	0.003	0.94
(D) >=1 child <15 worked in the past week	0.012	0.004	0.000	1.00	0.008	0.003	0.000	1.00
(D) >=1 HHM has an additional job	0.012	0.004	0.000	0.58	0.014	0.005	0.002	0.93
(Log) Max. HH # of work hrs (past week) - all jobs	0.011	0.004	0.001	0.36	0.022	0.008	0.008	0.78
(Log) HH Avg # of work hrs (past week) - main job	0.010	0.003	0.003	0.50	0.028	0.009	0.013	1.00
(Log) HH Avg # of weekly working hours - main job	0.009	0.003	0.002	0.56	0.026	0.009	0.018	1.00
(D) HHM primary job status: self employed	0.008	0.003	0.004	0.88	0.012	0.004	0.003	0.75
Literacy								
(D) >=1 HHM aged >15 able to read in Indonesian	0.031	0.010	0.008	0.40	0.033	0.011	0.015	0.97
(D) >=1 HHM aged >15 able to read and write	0.026	0.009	0.008	0.52	0.026	0.009	0.019	0.93
(D) >=1 HHM aged >15 able to write	0.022	0.008	0.007	0.50	0.023	0.008	0.017	0.88
(D) >=1 HHM aged >15 able to write in Indonesian	0.021	0.007	0.007	0.43	0.039	0.013	0.018	0.98
(D) >=1 HHM aged >15 able to read	0.016	0.006	0.008	0.50	0.020	0.007	0.017	0.90
(D) >=1 HHM speaks only Indonesian in daily life	0.016	0.006	0.003	0.94	0.022	0.008	0.009	1.00
Demographic dynamics								
Nb HHM aged <15 who left since prev. survey	0.026	0.009	0.009	1.00	0.021	0.007	0.010	1.00
Nb former HHM living in the same district	0.025	0.008	0.005	1.00	0.015	0.005	0.001	0.94

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(D) >=1 HHM aged <15 left since prev. survey	0.023	0.008	0.007	1.00	0.025	0.009	0.009	1.00
(D) >=1 HHM moved in for work	0.023	0.008	0.003	1.00	0.013	0.005	0.001	1.00
Nb HHM who left for work	0.021	0.007	0.006	1.00	0.015	0.005	0.001	1.00
Nb new HHM aged 15-64 since prev. survey	0.021	0.007	0.001	0.62	0.016	0.006	0.004	1.00
Nb former HHM living in the same subdistrict	0.021	0.007	0.001	0.95	0.021	0.007	0.000	0.86
Nb HHM aged 15-64 who left since prev. survey	0.020	0.007	0.002	1.00	0.015	0.005	0.000	0.56
(D) >=1 former HHM living in the same district	0.019	0.007	0.004	1.00	0.017	0.006	0.002	1.00
Nb HHM who moved in for work	0.019	0.007	0.003	1.00	0.019	0.007	0.000	0.95
Nb HHM who left the HH since prev. survey	0.018	0.006	0.004	1.00	0.013	0.004	0.002	0.85
Nb former HHM living in the same province	0.018	0.006	0.000	0.78	0.023	0.008	0.001	1.00
(D) >=1 HHM left the HH for family obligations	0.018	0.006	0.000	0.11	0.013	0.004	0.001	0.00
(D) >=1 former HHM living in the same village	0.017	0.006	0.006	1.00	0.018	0.006	0.002	0.95
(D) >=1 HHM aged 15-64 left since prev. survey	0.017	0.006	0.003	1.00	0.011	0.004	0.000	0.27
Nb new HHM since prev. survey	0.017	0.006	0.001	0.59	0.011	0.004	0.001	0.64
Nb HHM who moved in for school (start or finish)	0.017	0.006	0.001	1.00	0.017	0.006	0.000	0.94
(D) >=1 former HHM living in the same province	0.017	0.006	0.000	0.82	0.018	0.006	0.000	1.00
(D) >=1 HHM left the HH since prev. survey	0.016	0.006	0.004	0.94	0.014	0.005	0.002	0.93
(D) >=1 new HHM aged <15 since prev. survey	0.016	0.006	0.004	0.06	0.020	0.007	0.012	0.00
(D) >=1 HHM died since prev. survey	0.016	0.006	0.001	0.13	0.009	0.003	0.001	0.00
(D) >=1 former HHM living in other prov Indonesia	0.016	0.006	0.001	0.38	0.013	0.005	0.001	0.64
Nb former HHM living in other prov in Indonesia	0.015	0.005	0.000	0.20	0.015	0.005	0.000	0.69
Nb HHM who died since previous survey	0.015	0.005	0.000	0.20	0.013	0.005	0.001	0.00
Nb HHM who left for family obligations	0.015	0.005	0.000	0.53	0.012	0.004	0.001	0.00
(D) >=1 new HHM since prev. survey	0.014	0.005	0.002	0.57	0.010	0.003	0.001	0.30
(D) >=1 former HHM living in the same subdistrict	0.014	0.005	0.001	0.93	0.018	0.006	0.000	0.79
(D) >=1 former HHM living abroad	0.014	0.005	0.000	0.64	0.014	0.005	0.002	1.00
(D) >=1 new HHM aged >64 since prev. survey	0.014	0.005	0.000	0.21	0.014	0.005	0.001	1.00
Nb HHM who left for school (start or finish)	0.014	0.005	0.000	0.29	0.013	0.005	0.001	0.93
Nb HHM who left to become independent	0.013	0.005	0.002	0.92	0.014	0.005	0.001	0.80
Nb HHM who moved in for family obligations	0.013	0.005	0.001	1.00	0.014	0.005	0.001	0.00
Nb former HHM living abroad	0.013	0.005	0.000	0.69	0.015	0.005	0.003	1.00
Nb new HHM aged >64 since prev. survey	0.013	0.005	0.000	0.38	0.011	0.004	0.001	1.00
(D) >=1 new HHM aged 15-64 since prev. survey	0.012	0.004	0.001	0.58	0.014	0.005	0.002	0.87
(D) >=1 HHM left for school (start or finish)	0.012	0.004	0.000	0.33	0.015	0.005	0.002	1.00
Nb former HHM living in the same village	0.011	0.004	0.005	1.00	0.011	0.004	0.001	0.91
Nb new HHM aged <15 since prev. survey	0.010	0.003	0.003	0.40	0.015	0.005	0.011	0.00
(D) >=1 HHM left to become independent	0.010	0.003	0.002	1.00	0.013	0.004	0.001	0.77
(D) >=1 HHM left for work	0.009	0.003	0.004	1.00	0.018	0.006	0.001	0.95
(D) >=1 HHM moved in for family obligations	0.008	0.003	0.000	0.38	0.011	0.004	0.002	0.00
(D) >=1 HHM moved in for school (start or finish)	0.005	0.002	0.001	1.00	0.012	0.004	0.000	1.00

Table A1 : Ranking of all candidate predictors - by category (continued)

	Urban				Rural			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health status								
(D) >=1 HHM aged >15 reporting to be swht healthy	0.023	0.008	0.003	0.00	0.023	0.008	0.001	0.67
(D) >=1 HHM aged <15 w/ nausea (past month)	0.023	0.008	0.000	0.74	0.017	0.006	0.001	1.00
(D) >=1 HHM aged >15 reporting to be unhealthy	0.021	0.007	0.000	1.00	0.015	0.005	0.000	0.50
(D) Adult HHM unable to stand up from sitting - floor	0.021	0.007	0.000	0.48	0.016	0.006	0.001	0.94
(D) >=1 HHM aged <15 reported to be s. unhealthy	0.020	0.007	0.001	0.00	0.015	0.005	0.000	0.81
(D) >=1 HHM aged <15 w/ missed activity day	0.019	0.007	0.000	0.42	0.011	0.004	0.000	0.45
(D) Adult HHM unable to walk for 5 km	0.018	0.006	0.001	0.50	0.016	0.006	0.001	0.76
(D) >=1 HHM aged <15 w/ cough (past month)	0.018	0.006	0.000	0.28	0.010	0.003	0.000	0.20
(D) >=1 HHM aged <15 w/ toothache (past month)	0.018	0.006	0.000	0.06	0.014	0.005	0.002	1.00
(D) >=1 HHM aged <15 reported to be very healthy	0.017	0.006	0.001	1.00	0.012	0.004	0.002	0.00
(D) >=1 HHM aged <15 w/ breathing pb (past month)	0.017	0.006	0.001	1.00	0.018	0.006	0.000	0.84
(D) >=1 HHM aged >15 in bed for >=1 day	0.017	0.006	0.000	1.00	0.017	0.006	0.004	1.00
(D) Adult HHM unable to stand up from sitting - chair	0.017	0.006	0.000	0.12	0.011	0.004	0.000	0.45
(D) >=1 HHM <15 w/ fever (past month)	0.016	0.006	0.000	0.25	0.015	0.005	0.001	0.25
(D) Adult HHM unable to go to the bathroom	0.016	0.006	0.000	0.44	0.021	0.007	0.001	0.00
(D) Adult HHM unable to draw water from a well	0.015	0.005	0.000	0.80	0.013	0.005	0.001	0.14
(D) Adult HHM unable to carry a heavy load for 20m	0.014	0.005	0.000	0.43	0.016	0.006	0.001	0.82
(D) >=1 HHM aged <15 w/ stomach ache (past mth)	0.014	0.005	0.000	0.57	0.008	0.003	0.002	1.00
(D) >=1 HHM aged <15 reported to be smwht healthy	0.013	0.005	0.002	0.15	0.015	0.005	0.006	0.06
(D) >=1 HHM aged <15 w/ diarrhea (past month)	0.013	0.005	0.001	0.00	0.010	0.003	0.000	0.90
(D) >=1 HHM aged <15 in bed for >=1 day	0.013	0.005	0.001	0.00	0.013	0.005	0.000	0.07
(D) Adult HHM unable to sweep the floor	0.013	0.005	0.001	0.15	0.013	0.005	0.001	0.07
(D) >=1 HHM aged >15 reporting to be very healthy	0.013	0.005	0.000	0.92	0.015	0.005	0.001	0.00
(D) Adult HHM unable to dress	0.012	0.004	0.000	0.50	0.012	0.004	0.000	0.33
(D) >=1 HHM aged<15 w/ headache (past month)	0.011	0.004	0.000	0.55	0.018	0.006	0.000	0.79
(D) Adult HHM unable to easily bow, squat, kneel	0.010	0.003	0.000	0.50	0.013	0.005	0.000	0.36
(D) >=1 HHM aged >15 reporting to be s. unhealthy	0.010	0.003	0.000	0.30	0.022	0.008	0.002	0.96
(D) >=1 HHM aged <15 w/ runny nose (past month)	0.009	0.003	0.001	0.11	0.015	0.005	0.001	0.31
(D) >=1 HHM aged <15 w/ eye infection (past month)	0.009	0.003	0.000	0.78	0.007	0.002	0.000	1.00
(D) >=1 HHM aged >15 w/ missed activity day	0.008	0.003	0.001	1.00	0.015	0.005	0.006	1.00
(D) >=1 HHM aged <15 reported to be unhealthy	0.008	0.003	0.000	0.25	0.008	0.003	0.000	0.75
(D) >=1 HHM aged <15 w/ skin infection (past mth)	0.006	0.002	0.000	0.17	0.017	0.006	0.001	1.00
Community participation								
Nb HHM who participated in *arisan* (past year)	0.026	0.009	0.006	1.00	0.031	0.010	0.025	1.00
(D) >=1 HHM has participated in *arisan* (past year)	0.014	0.005	0.008	1.00	0.035	0.012	0.027	1.00

Notes: Results based on a random sample s_2=1,000,000 models of 10 predictors. Columns (1) and (5) show for each variable the probability of being included in a top 0.1% model; columns (2) and (6) show the probabilities that models including each variable are in the top 0.1%; columns (3) and (7) show the added values, as a share of model R2; and columns (4) and (8) show the positive sign probabilities. HH stands for household, HHH for household head and HHM for household member; >=1 stands for at least 1. *Arisan* is a rotating savings group. The numbers in bold show the best predictors in each area.