UNITED NATIONS UNIVERSITY

UNU-WIDER

World Institute for Development Economics Research

Working Paper No. 2010/99

The Reliability of Small Area Estimation Prediction Methods to Track Poverty

Luc Christiaensen,[1] Peter Lanjouw,[2] Jill Luoto,[3] and David Stifel[4]

September 2010

Abstract

Tracking poverty is predicated on the availability of comparable consumption data and reliable price deflators. However, regular series of strictly comparable data are only rarely available. Poverty prediction methods that track consumption correlates as opposed to consumption itself have been developed to overcome such data gaps. These methods typically assume that the estimated relation between consumption and its predictors is stable over time—assumptions that usually cannot be tested directly. This study analyses the performance of poverty prediction models based on small area estimation (SAE) techniques. Predicted poverty estimates are compared to directly observed levels in a series of country settings that are widely divergent, but where data comparability over time is not judged to be a problem. Prediction models that employ either nonfood expenditures or a full set of assets as predictors, yield poverty estimates

…/.

Keywords: consumption prediction, price deflator, poverty dynamics, small area estimation

JEL classification: D12, D63, I32

[1] UNU-WIDER, Helsinki, email: luc@wider.unu.edu; [2] Development Economics Research Group of the World Bank, Washington, DC, email: planjouw@worldbank.org; [3] University of California Berkeley, Berkeley, CA, email: luoto@berkeley.edu; [4] Lafayette College, Easton, PA, email: stifeld@lafayette.edu):

This study has been prepared within the UNU-WIDER project Frontiers of Poverty Analysis, directed by Anthony Shorrocks.

UNU-WIDER gratefully acknowledges the financial contributions to its research programme by the governments of Denmark (Royal Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development—DFID).

ISSN 1798-7237 ISBN 978-92-9230-343-3

that match observed poverty fairly closely. This offers some support for using SAE techniques especially those based on models employing household assets, to approximate the evolution of poverty in settings where comparable consumption data are absent or settings where price deflators are of dubious validity. However, the findings also call for further validation especially in settings with rapid, transitory poverty deterioration, as in Russia during the 1998 financial crisis.

Acknowledgements

The paper has benefited from comments by participants in the UNU-WIDER 'Frontiers of Poverty Analysis' Conference held in Helsinki, 26-27 September 2008. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the executive directors of the World Bank or the governments they represent.

Acronyms and tables given at the back.

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Typescript prepared by T:mi LHR Editorial and Secretarial Assistance, for UNU-WIDER

1 Challenges in tracking poverty

Interest in understanding how poverty evolves over time is longstanding and has recently received additional impetus through the need to monitor progress towards halving poverty by 2015, the first Millennium Development Goal. Tracking poverty is predicated on the availability of poverty estimates that are comparable over time. Such measures are typically derived from survey-based household expenditure data. The simple act of constructing a survey-based consumption measure already poses considerable challenges (Deaton and Zaidi 2002); these only multiply when consumption expenditures and poverty estimates need to be compared over time.

First, consumption measures are often not available at regular intervals. For example, of the 48 countries in sub-Saharan Africa (SSA) included in the World Bank's PovcalNet[1] database, only 18 countries possess more than one national household consumption survey since 1995. Second, in those settings where multiple consumption measures are available, they are frequently not directly comparable. Even slight differences in questionnaire or survey design can yield quite different poverty estimates (Lanjouw and Lanjouw 2001; Gibson, Huang, and Rozelle 2003). Finally, the price deflators needed to capture real changes in command over goods and services, are also often missing or of dubious validity. More often than not, official consumption price indices (CPIs) deviate from price deflators calculated directly from the surveys, with little information available to adjudicate the choice.[2]

In response, poverty economists have been developing a series of different poverty prediction methods, exploiting the comparability of subsets of data within and across surveys.[3] The methods differ in the predictors and prediction techniques used, but they generally share the critical, and largely untested assumption that the estimated relation between the predictors and their welfare measure is stable over time. This cannot be taken for granted and has become an important stumbling block in furthering the use of poverty prediction techniques to overcome data constraints in tracking poverty over time.

The need for comparing and validating poverty prediction methods is perhaps best illustrated by the 'Great Poverty Debate' in India (Deaton and Kozel 2005). Following market liberalization in the early 1990s, the official poverty numbers for India showed a drop in poverty incidence from 36 per cent in round 50 of the national sample survey (NSS) (1993/4) to 26 per cent in round 55 (1999/2000), or a reduction in the number of poor people by about 60 million.

---

1 See www.research.worldbank.org/PovcalNet/jsp/CChoiceControl.jsp?WDI_Year=2007, accessed April 2010.

2 For instance, while households in Tanzania faced price increases of 93 per cent between 2000/01 and 2007 according to the national household budget surveys, the official CPI recorded only a price increase of 45 per cent (Hoogeveen and Ruhinduka 2009). Adjustment of the CPI, which was largely based on urban consumption baskets, also proved to go a long way in remedying the 1994-2003 growth-poverty paradox in Burkina Faso (Grimm and Günther 2006).

3 See for example, Ravallion (1996), Sahn and Stifel (2000), Kijima and Lanjouw (2003), Azzarri et al. (2006), Stifel and Christiaensen (2007), Tarozzi (2007), Mathiassen (2009), and Grosse, Klasen, and Spatz (2009).

However, these official numbers were received with scepticism. There was a widespread view that the underlying data were not comparable because reporting periods for various consumption items had changed between the two survey rounds. There were also lingering doubts about the accuracy of the price indices used to update the poverty lines (Deaton 2008). A number of different poverty calculations were proposed, each of them predicated on assumptions that were difficult to test. In contrast to the official estimates, one widely circulated alternative estimate puts the actual decline in poverty at only 2.8 percentage points, in effect implying an increase in the absolute number of poor people by about five million (Sen and Himanshu 2004). This particular estimate drew on alternative, abbreviated consumption data from the employment module of the NSS survey.

In an attempt to restore comparability across the Indian surveys via prediction methods, Deaton (2003) exploits the fact that the section of the consumption module that pertained to '30-day' expenditures, had not changed between rounds. He estimates the probability of a household in the 55th round being poor as a function of its per capita 30-day expenditures in that round and the relation observed between 30-day (log) per capita expenditures and total (log) per capita expenditures during the 50th round. The reliability of these new poverty estimates, suggesting a decline of 7 percentage points, depended on the validity of the assumption that there had been no change in the Engel curve relating 30-day type expenditures to total expenditures over time. This assumption rules out substitution effects following relative price shifts or changes in tastes between included and excluded expenditure subcomponents. Sen and Himanshu (2004) examine these assumptions in detail, and showed them to be far from innocuous.

Kijima and Lanjouw (2003) consider an alternative poverty prediction method. They also use a subset of explanatory variables that were strictly comparable between survey rounds, but confine their attention to variables, such as household demographics and stocks of assets, that came from outside the consumption module. The poverty estimates based on these predictions indicated a much less rapid decline in poverty during the 1990s than the official numbers and provided a qualitatively similar assessment of poverty decline as the Sen and Himanshu (2004) estimates. In this approach, the underlying relationship between consumption and its correlates is assumed to remain stable over time, thereby ruling out any possible changes in the 'returns' to factors such as education and labour. This too is a controversial assumption, especially in fast growing economies such as India.

Going one step further, Tarozzi (2007) uses both the 30-day consumption items and non-consumption variables such as educational status and land as predictors. He tests the validity of the stable parameter assumption on the 30-day consumption items and on the non-consumption variables, working with the much smaller NSS rounds that are fielded during the intervals between the large, 'quinquennial' rounds that underpin official poverty estimates. Tarozzi finds indirect support for the assumption of parameter stability. In his datasets the large reduction in poverty implied by the official figures received some empirical validation. However, his analysis also remains controversial because the year-to-year poverty changes implied by his calculations were difficult to accept. Concerns were expressed as to how well the 'thin' rounds were suited to this kind of analysis. Despite, or perhaps because of, all these efforts, the poverty trend in India during the 1990s remains a subject of intense debate.

In the absence of regularly fielded rounds of the same consumption surveys, researchers have also exploited the comparability and availability of data across time from alternative data sources. For example, Kenya had not conducted a national household budget survey since 1997, but conducted three demographic and health surveys (DHS) in 1993, 1998 and 2003. Stifel and Christiaensen (2007) estimate the relationship between assets and consumption in the 1997 national household budget survey and subsequently apply these estimates to comparable asset data in each of the DHS to predict household consumption and poverty in rural and urban areas. This yielded useful insights into the dynamics of poverty in Kenya between 1993 and 2003. To mitigate potential bias from the parameter stability assumption, they excluded household assets whose returns were considered more prone to change over time, such as labour and education variables, and included factors that affected the returns to assets over time such as rainfall and nutritional status. Even though the predictions of poverty from this study looked plausible when compared with trends in other indicators of wellbeing, it remains that the underlying assumptions for these predictions could not be verified.[4]

Clearly, empirical validation of the different model specifications and their underlying parameter stability assumptions is necessary before poverty prediction techniques can be routinely used. Such empirical verification is not straightforward because it requires, at a minimum, settings in which comparable consumption data *are not* missing. In such settings the exercise can be performed *as though* the data were missing, and afterwards predicted poverty can be checked against the 'truth'.

This paper makes a first contribution to filling this important void. It compares the poverty measures obtained directly from the data in a series of settings that have comparable expenditure data across time, with those obtained through application of an adapted version of the small area estimation (SAE) technique (Elbers, Lanjouw and Lanjouw 2003). Models based both on consumption subcomponents and on different combinations of non-consumption assets are explored. This provides a test of the predictive power of the most commonly used poverty prediction models, including the validity of the parameter stability assumption.

In particular, the study uses repeated cross-sections with highly comparable survey and questionnaire design from Vietnam (Vietnam living standards surveys (VLSS) of 1992/3 and 1997/8) and Russia (Russian longitudinal monitoring surveys (RLMS) of 1994, 1998, and 2003) as well as 2000-04 annual household panel data from Inner Mongolia and Gansu, two rural provinces in western China. Together they cover a range of different settings (low and middle income, rural/urban), spanning periods of deep structural change, accompanied by quite divergent evolutions of poverty. This puts the prediction techniques to a demanding test. The paper also presents another application for Kenya. Rather than a validation exercise, the purpose of this final application is to illustrate how prediction methods can help to confront comparability issues arising from problematic temporal cost-of-living indices.

In the three countries examined here, the poverty prediction methodology performs rather differently. In Vietnam, during a period of dramatic overall poverty decline and

4 Grosse, Klasen and Spatz (2009) provide a similar application to Bolivia, this time adjusting the parameters of the different regions based on the parameter changes observed over time in the urban subsample for which consumption data were available in all periods.

deep structural change, the poverty prediction method works quite well with models using certain expenditure components (non-rice food spending, nonfood spending), but equally well with comprehensive models specified on the basis of non-consumption assets. In this setting the underlying stability assumptions of the poverty prediction method appear to hold. In Russia, however, during a period of dramatic poverty increase between 1994 and 1998, only models based on food consumption are successful. No model is consistent in working well over the longer period between 1994 and 2003 in Russia, or in tracking poverty change between 1998 and 2003. In China, models based on non-expenditure assets work well, but only in the province of Inner Mongolia.

The findings combine to suggest that poverty prediction methods *can* work, even in settings where underlying structural changes are deep and poverty change is marked. However, the methods do not work everywhere. A preliminary meta-analysis, examining the country- and other context-specific circumstances that are correlated with the success of poverty prediction methods, provides some useful pointers, notably the importance of high explanatory power of the consumption model (as reflected in high R-squares). More experience is needed in order to get a better sense of when, and how, poverty prediction methodologies can be best applied.

The paper proceeds as follows. Section 2 briefly reviews the SAE methodology and reviews theoretical considerations in choosing consumption predictors (i.e., the consumption subcomponents and the non-consumption asset variables). Section 3 describes the data in more detail. The predictive power of the different prediction models with consumption subcomponents and different combinations of non-consumption assets across the different settings is assessed and the application to Kenya is presented in section 4. Section 5 concludes.

2 Methodological considerations in tracking poverty by poverty predictors

2.1 The adapted small area estimation technique

Following Kijima and Lanjouw (2003) and Stifel and Christiaensen (2007) the adapted version of the small area estimation (SAE) methodology developed by Elbers, Lanjouw and Lanjouw (2003) is used to impute a definition of consumption from one household survey into the other. The core intuition behind the adapted SAE methodology is to predict per capita consumption at the level of the household in survey round two using the available information on these households in round two (e.g., consumption subcomponents and/or non-consumption assets) as well as the parameter estimates derived from a model of consumption estimated from round one (including those concerning the distribution of the error term). By restricting the explanatory variables to those that are comparable across surveys, the method ensures an identical definition of consumption (welfare) across surveys, assuming that the relationship between consumption and its correlates remains stable over time. If non-consumption assets are used, it also circumvents the need for price deflators.

More formally,[5] let $W(c_t)$ represent the value at time t of the welfare measure (for example, poverty or inequality) based on the distribution of household-level per capita

5 For a detailed exposition, see Stifel and Christiaensen (2007).

consumption c at time t. Consider a log linear approximation to household consumption c_t:

$$\ln c_t = x_t'\beta_t + u_t \qquad (1)$$

where x_t are the p poverty predictors, such as the consumption subcomponents and/or the non-consumption assets, β_t is a vector of p parameters, and u_t is a heteroscedastic error term. Since only x_{t+k} is observed, not c_{t+k}, with k=1,…, T, household disturbances u_{t+k} of $\ln c_{t+k} = x_{t+k}'\beta_{t+k} + u_{t+k}$ are always unknown (only the distribution of u_t is known), and the expected value of W is taken given x_{t+k} and the model parameters of (1), i.e., $\mu_{t+k}^{s} = E[W^{s}(x_{t+k}, \beta_{t+k}, u_{t+k})]$ as opposed to $W(c_{t+k})$. The superscript‘s’ indicates that the expectation is conditional on a *sample* of the households in the same geographical area in period $t+k$ rather than a *census* of the households (in period t) as in poverty mapping.

Consistent estimates of u_{t+k} and β_{t+k} are obtained by taking a draw r from the estimated distributions of u_t and β_t respectively, which are obtained in estimating Equation (1). This yields $\hat{\mu}_{rt+k}^{s} = E[W^{s}(x_{t+k}, \hat{\beta}_{rt+k}, \hat{u}_{rt+k})]$. In doing so, the methodology imposes the assumption that the distributions of β_t remain constant over time, i.e., that the distributions of β_{t+k} and β_t are the same. Further, although the distribution of u_t is updated with x_{t+k} to estimate u_{t+k}, the relationship determining the heteroscedastic nature of the data-generating process is also assumed to be constant. Finally, given that the expectation is generally analytically intractable, an estimate of the expected value of $W(c_{t+k.})$ is obtained through simulating the process described above for different draws r, yielding $\tilde{\mu}_{t+k}^{s}$.

In pursuit of precise and consistent estimates of W_{t+k} in the absence of observations on the true c_{t+k}, it is important to understand which factors affect the difference between the estimator $\tilde{\mu}_{t+k}^{s}$ of the expected value of $W(c_{t+k})$ and the actual level of welfare for that geographical area at $t+k$. Four error components are distinguished: (1) the idiosyncratic $(W(c_{t+k}) - \mu_{t+k})$, (2) the sampling $(\mu_{t+k} - \mu_{t+k}^{s})$, (3) the model $(\mu_{t+k}^{s} - \hat{\mu}_{rt+k}^{s})$, and (4) the computational $(\hat{\mu}_{rt+k}^{s} - \tilde{\mu}_{t+k}^{s})$ error.

The *idiosyncratic* error component $(W(c_{t+k}) - \mu_{t+k})$ results because actual c_{t+k} is not known/used, but rather stochastic c_{t+k},, whereby the stochastic nature of c_{t+k} is known/assumed through the distributional features of u_{t+k}, i.e., $E[W(x_{t+k}, \beta_{t+k}, u_{t+k})]$ is calculated as opposed to $W(c_{t+k})$. This component depends on the explanatory power of the poverty predictors[6] and the sensitivity of the welfare measure to the stochastic nature of c_{t+1},[7] but becomes important only if the target population is small (Alderman et al. 2002). The interest here is in tracking welfare/poverty measures over time for major groups or areas for which representative data have been collected. Given that these populations (e.g., rural/urban, province) are usually rather large, the idiosyncratic error component tends to be small.[8]

Sampling error $(\mu_{t+k} - \mu^{s}_{t+k})$ arises because the consumption model is imputed into a sample, rather than a census. It depends on the sampling design, the sampling size and the population variance of the consumption measure. Error calculations on the poverty estimates should take the sampling design into account. The *computational* error $(\hat{\mu}^{s}_{rt+k} - \tilde{\mu}^{s}_{t+k})$ can be set arbitrarity small by selecting a sufficiently large number of simulations (Elbers, Lanjouw and Lanjouw 2002, 2003).

The *model* error component $(\mu^{s}_{t+k} - \hat{\mu}^{s}_{rt+k})$ follows from the fact that the parameters β_{t+k} as well as those describing the distribution of $u_{t+k,}$ are estimated, i.e., $E[W(x_{t+k}, \hat{\beta}_t, \hat{u}_{t+k})]$ is used as opposed to $E[W(x_{t+k}, \beta_{t+k}, u_{t+k})]$. The magnitude of the model error component depends in general on (1) the sensitivity of the welfare indicator to errors in estimated consumption, (2) the extent to which the measurement of the x variables in the target population deviates from the population of origin, (3) the precision of the coefficient estimates, which determines the distribution of $\hat{\beta}_{t+k}$, and the predictive power of the model, which affects the distributional parameters of $\hat{u}_{t+k}$, (4) the validity of the assumption that $\hat{\beta}_{t+k}$ does not vary across the consumption distribution when using OLS or GLS in estimating the basic consumption model, and (5) the validity of the assumption that the estimated distributions of $\hat{\beta}_t$ and the parameters used to estimate $\hat{u}_{t+k}$ are stationary.[9] It is the latter assumption that has received most attention

6 While consumption is clearly measured with error in practice, error free consumption measures are assumed here in the application. See Chesher and Schluter (2002) for rules to approximate the effect of measurement error in estimating welfare measures.

7 Ravallion (1988) and Lanjouw and Lanjouw (2001) show, for example, that when consumption becomes stochastic and equal to its true value plus an independently and normally distributed error term, as it has been modelled here, then the observed density function has the same shape as that of the true consumption, but with fatter tails. If the poverty line lies to the left of the mode of the distribution, the expected poverty headcount based on the stochastic consumption measure will be larger than that based on the true consumption measure, and vice versa. This pattern is quite general: for most measures of poverty, (including the widely used Foster-Greer-Thorbecke class of poverty measures), expected poverty will increase.

8 Elbers, Lanjouw and Lanjouw (2002) suggest that this error becomes negligible when population size approaches 10,000. Note that the population of concern is not the *sample*, which may be much smaller, but the actual true underlying *population*, which will concern provinces or regions and will thus be much larger.

9 A concern has also been expressed that SAE prediction techniques may be noisy (though still consistent) in the presence of small-area heterogeneity in the conditional distribution of consumption (Tarozzi and Deaton 2009). Unlike in poverty mapping, where the consumption model parameters are identified from a larger geographic region to impute welfare measures into its subregions, this concern

in the literature. Uncertainty around this assumption cannot be straightforwardly captured in standard error calculations on predicted welfare estimates.

2.2 Minimizing prediction error through astute predictors and estimation techniques

As has been illustrated in the 'Great Indian Poverty Debate' and shown empirically by Mathiassen (2009), a focus on minimizing model error is needed to minimize prediction error.[10] Model error from differences in the measurement of the poverty predictors (model error (2)) can be minimized by selecting surveys that maintain their survey and questionnaire design over time. The demographic and health surveys provide one good example, exploited earlier by Stifel and Christiaensen (2007).

In examining the predictive power of poverty predictors (model error (3)), two considerations are important: (1) the sensitivity of the poverty predictor to both upward and downward changes in income among those who are poor and those who are vulnerable to becoming poor, and (2) the likely stability of the relationship between the predictor and consumption over time.

Given Engel's Law, the income elasticity of different consumption subcategories likely differs depending on the level of income. Harrower and Hoddinott (2005) illustrate, for example, that food expenditures among poor rural villages in northern Mali were reasonably well smoothed in the face of income shocks, while nonfood expenditures were not. Following Bennett's Law, further differences in income sensitivity between staple and non-staple food expenditures are expected. The income elasticity of staple crop expenditures is likely highest among the poorest and changes in staple crop expenditures may thus be better at predicting improvements in distribution sensitive poverty measures than in predicting improvements in poverty headcounts. Among richer and urban households, non-staple food expenditures (such as eating out) are likely less robust against income declines than expenditures on staple foods.

Non-staple food expenditures may also be more sensitive to downward income shocks than nonfood expenditures. For example, depending on the depth of financial markets, reducing the service stream from existing possessions or housing may take more time to show up in expenditure numbers. Increases in income, on the other hand, may translate quicker into purchases of non-staple foods and durables alike (Elbers and Pouw 2009).

There is *a priori* no ground to assume that the Engel curve is more stable for certain consumption subcomponents. It is subject to changes in relative prices, changing tastes, and other demand shocks in all cases. The predictive power of both food and nonfood subcomponents of consumption expenditures will be considered below. These will be further divided—where the data permit—into staple and non-staple foods, as well as

---

is not applicable here. The area from which the consumption model is estimated, is the same as the area to which future poverty is imputed.

10 Predicting poverty for rural and urban areas in one region in Mozambique, based on a small set of poverty predictors from the labour survey and an estimated poverty prediction model from an earlier household budget survey, Mathiassen (2009) finds that about 80 per cent of the variance in the prediction error of the future poverty headcount could be attributed to model error. About 20 per cent was due to sampling error and only 1 per cent due to the idiosyncratic error.

frequent and infrequent nonfood expenditures (typically collected using 30-day and 1-year recall, respectively).

In addition to consumption subcomponents, which may be more closely correlated with overall consumption, but which are also more time consuming to collect and still require price deflators, five broad classes of non-consumption asset data are considered. These comprise: (1) geographic indicators such as rural/urban and regional location (proxying a household's agro-ecological, economic and institutional assets); (2) household demographic information and (3) educational and employment information such as sector of work by the household head (proxying the quantity and quality of their labour assets); (4) variables on the quality of housing such as presence or absence of electric lighting, permanent roofing material, and private water tap; and (5) ownership of consumer durables such as a bicycle, colour television, electric fan, etc. (proxying a household's physical assets).

Filmer and Scott (2008) find that asset indices, which are usually composed of housing quality indicators (asset class 4) and consumer durables (asset class 5), are better correlated with per capita consumption measures when consumption is measured with less error, when its transitory component is lower, and when its share of nonfood expenditures is higher (i.e., expenditures with a public-good component). The inclusion of variables that are more directly correlated with transitory income shocks such as rainfall, nutritional and health status, or situation specific variables such as arrears in pensions in Russia, or even measures of subjective wellbeing, could help capture better the transitory component in consumption.[11] Here, the study deliberately focuses on a sparser core set of assets that is commonly used in explaining variation in household consumption levels and usually readily available in the questionnaires. Including more time variant variables from outside, the questionnaire would increase the data compilation efforts in practice, but may contribute to greater precision and less bias.

To better capture economies of scale associated with the consumption of goods that have a public-good flavour and to increase the predictive power of the consumption model household, demographic information (asset class 2) can be incorporated. Yet *a priori*, one might expect that assets such as labour related variables, as well as education (asset class 2 and 3), would be more prone to parameter instability following structural or policy-induced, economic transformation.

Against this backdrop, poverty predictions derived from consumption models using different asset class combinations will be compared. Each time a careful selection of indicators from the different asset classes will be made using stepwise regression and other procedures so as to maximize the explanatory power of the consumption model and minimize model error, while at the same time preserving some parsimony in the final specification.

Turning to the estimation technique, the simulation based SAE technique deployed here has some advantages over the more standard consumption prediction techniques in the literature (see, for example, Azzarri et al. 2006) as well as the projection based estimators applied by Tarozzi and Deaton (2009) and Mathiassen (2009). First, unlike

[11] As they often represent important reasons for changing returns in assets, they could further mitigate the likelihood of violating the parameter stability assumption.

Azzarri et al., the SAE methodology provides consistent estimates of both the mean *and* the variance of consumption, and thus also a consistent estimate of the welfare measure in the future.[12]

Second, while the SAE method does impose some structure on the distribution of the idiosyncratic error term in the consumption model, the heteroscedasticity model applied within this approach permits a partial update of the distribution of the error term over time, reducing prediction error due to assumed stationarity of the error term.[13] Third, the technique is convenient to implement given the freely and readily downloadable PovMap2 software.[14]

2.3 Assessing the performance of SAE poverty predictions

To assess the performance of the SAE poverty prediction technique (including the empirical validity of the parameter stability assumptions), only surveys for which the survey design is fully comparable across rounds are examined. In particular, the SAE methodology is applied to strictly comparable explanatory variables of the two surveys, which are drawn either from the consumption modules of the two surveys, or from the other 'non-consumption or asset' sections of the two surveys. Poverty is then predicted in the latter round of each pair of surveys and the predicted poverty results derived from each 'class' of models are compared with the 'true' poverty numbers observed in the (second round) survey.

If the predicted poverty rates closely match the observed rates in a wide variety of settings, then this provides support to the contention that parameter instability need not be a pressing concern. Such a finding would suggest that SAE techniques and data on consumption subcomponents and/or household assets can indeed be used to

---

12 Azzarri et al. (2006) use only $x'_{t+k}\hat{\beta}_t$ to predict $\ln c_{t+k}$. Yet consistent estimation of $\hat{\beta}_t$ is not sufficient for the estimation of $W(c_{t+k})$ which is a function of c_{t+k}, and not a function of the distribution of conditional expectation $x'_{t+k}\beta_t$. For this reason, once $\hat{\beta}_t$ has been estimated within the SAE approach, an error term $\hat{u}_{t+k}$ is randomly drawn and added to $x'_{t+k}\hat{\beta}_t$ to recreate the conditional distribution of $\ln c_{t+k}$. Otherwise, the variance of the distribution of $\ln c_{t+k}$ is biased, resulting in a biased estimate of $W(c_{t+k})$.

13 Only the estimated parameters of the consumption variance equation are assumed to be stable over time, while the consumption variance predictors are allowed to change.

14 www.iresearch.worldbank.org/PovMap/PovMap2/PovMap2Main.asp. The wider applicability of the SAE method across different welfare measures is potentially traded off against greater accuracy of the estimated parameters in reflecting the effects of changes in the poverty predictors on poverty among the poor and the vulnerable in the projection methods. The latter identify the parameters on the poverty predictors from their effect on poverty status, as opposed to consumption in the SAE methodology, thereby implicitly holding the relationship between consumption and its predictors constant across the distribution of consumption. Relaxing this assumption through quintile regression estimation was not found to make a difference in producing small area estimations (Elbers, Lanjouw and Lanjouw 2002). When forecasting, it may matter as certain assets (e.g., luxury cars) may be very influential in determining the level of consumption, even though their change over time may not be very relevant to predict changes in poverty. Yet given the ancillary interest in this paper in examining the forecasting precision of operationally readily available tools, such as PovMap2, flexibility of the parameters along the distribution of consumption is not accommodated here, even though this could be readily pursued.

approximate the evolution of poverty within a country when comparable consumption data are absent. If, on the other hand, the results cannot capture the observed changes in poverty, caution in using such techniques would be warranted.

To judge the success of the prediction models, it is examined whether the predicted poverty estimate for the second survey round lies within the 95 per cent confidence interval around the observed poverty rate for that year. To be sure, standard errors can also be estimated around the predicted poverty rates. An alternative procedure would thus be to test whether the predicted poverty rate in the second round is statistically distinguishable from the observed poverty rate for that year. Such a test would be weaker than the tests performed here, as it would accommodate a degree of uncertainty around the predicted poverty rate, and thus reduce the likelihood of rejecting equality of the predicted and observed poverty rate. In other words, more imprecise poverty predictions would then (perversely) lead to a higher degree of validation of the prediction method. Moreover, given that the calculation of standard errors on the predicted poverty rates would not capture uncertainty associated with the underlying assumption of parameter stability, it is more appealing to apply this more stringent test of significance.

Finally, a meta-analysis is undertaken on the prediction results from the different surveys in order to explore whether there are any empirical regularities in explaining the performance of the prediction methodology. To do so, the prediction results from the different settings are pooled and related to a series of variables affecting prediction performance such as the predictive power of the consumption model (R-squared and number of sample observations), the characteristics of the poverty spells (direction of poverty change, time gap), the type of poverty measure and its level, as well as the set of poverty predictors used (different consumption subcomponents versus different non-consumption asset bundles).

3 Three plus one settings with multiple surveys of comparable design

The performance of the adapted SAE techniques is assessed in three different settings, each with highly comparable data, and each following or encompassing periods of substantial structural change. The case of Kenya where the technique helped in adjudicating the deflator choice is further considered as an additional application.

Following the introduction of the Doi Moi reforms in 1987, Vietnam experienced strong, broad-based economic growth throughout the 1990s resulting in an estimated drop in poverty from 60.6 per cent in 1992/3 to 37 per cent in 1997/8.[15] These estimates are based on the well respected and highly comparable Vietnam living standards surveys (VLSS) of 1992/3 and 1997/8 and are widely judged as reflecting the true course of events (see Agarwal, Dollar and Glewwe 2004). Both surveys are representative at the national and regional levels. The 1997/8 survey contains panel information on approximately 4,300 of the original 4,800 households interviewed in 1992/3, but has a total sample size of 6,002 households due to an expanded budget and sample design.

15 Poverty numbers in 1998 prices.

The second setting concerns Russia between 1994 and 2003, encompassing the height of the 1998 financial crisis. Poverty during this period is tracked based on the Russian longitudinal monitoring surveys (RLMS). These are nationally-representative surveys that track a panel of about 4,380 dwellings. They are not representative at the regional levels. Data from rounds 5, 8 and 12, corresponding to years 1994, 1998, and 2003, respectively, are used to test the performance of the SAE methodology. The surveys have similar consumption modules between rounds as well as many other similar survey components.[16] Nonetheless, there are concerns that the data are quite noisy, especially around Russia's financial crisis in 1998, which was accompanied by a sharp devaluation (Luttmer 2001; Wall and Johnston 2008).

The surveys document a sharp *increase* in recorded (consumption) poverty from 11.4 per cent in 1994 to 33.8 per cent in 1998, after which poverty is estimated to have fallen to 11.1 per cent by the twelfth round in 2003.[17] This sharp increase (and decline) in poverty, accentuated by lingering concerns about increased measurement error during the crisis period, provides an opportunity to explore whether the SAE prediction technology is equally adept at capturing a dramatic decline in welfare.

Thirdly, 2000-04 panel data from 800 and 700 rural households in Inner Mongolia and Gansu, two provinces in western China are used. Following the introduction of the household responsibility system in 1978, provinces in China experienced dramatic change, first along the coast, but later also in its hinterland. The surveys concerned here document this change as background to a World Bank supported poor-area development programme in both provinces. Consumption data were collected using the diary method, and the same questionnaire was administered following the same survey procedures throughout the panel. Poverty incidence (using the national poverty line) declined spectacularly from 19 per cent in 2000 to 6.2 per cent in Inner Mongolia and halved from 24.3 to 11.8 per cent in Gansu.

Finally, the methodology is applied in Kenya. The two latest household expenditure surveys in Kenya are the 1997 welfare monitoring survey (WMS) and the 2005/6 Kenya integrated household budget survey (KIHBS). These surveys were implemented during different periods of the year and more detailed consumption data were collected during the KIHBS—raising some questions regarding the comparability of the data.[18] However, it is the choice of the appropriate deflator that was generally considered to pose the greatest challenge to tracking the evolution of poverty during this period in Kenya. The official CPI almost doubled between 1997 and 2005-06, while the deflator based on recalculations of the rural and urban poverty lines suggested a much lower price increase (6 per cent in rural and 27 per cent in urban areas). Puzzlingly, changes in poverty lines and the CPI had largely mirrored each other in the surveys prior to 1997.

---

16 See www.cpc.unc.edu/rlms for more complete information about the survey and sampling design.

17 Poverty figures are authors' calculations based on RLMS expenditure data.

18 The 1997 WMS survey was carried out during three months (February -May 1997), while the data collection for the 2005 KIHBS spanned May 2005 till May 2006 during which field work was organized in 17 three-week cycles with all 69 districts covered in each cycle. The consumption data collection during the KIHBS was also more detailed. During the WMS, consumption data were collected for broad (aggregated) categories: 79 food (7 day recall) and 48 nonfood items compared with the use of more detailed categories during the KIHBS: 140 food items (7 day recall) and 184 nonfood items (1 month recall).

Against this background, application of the SAE methodology is intended here to serve as a check on the currently widely reported poverty numbers that are based on not fully comparable expenditure measures and on survey-based price deflators (World Bank 2008).

In each of these surveys, the consumption model is estimated separately for the geographic area concerned (national, rural, urban, province) using the initial survey year as base year. Poverty is subsequently predicted into the second survey, and the poverty rates are compared with the observed ones. The performance of the technique is tested using both the poverty headcount and the more distribution-sensitive poverty gap. The procedure is repeated using different combinations of expenditure subcomponents and different combinations of assets. Among the many indicators for each of the asset categories, those that maximized the R-squared were selected.

This way, the SAE poverty prediction technique is explored in a range of settings including quite different degrees of poverty change (increase, stagnation, decline), different time intervals (1-10 years), different levels of poverty as well as a wide variety of geographic environments (rural versus urban, lowlands versus highlands) each subjected to different shocks and structural change. Classified by the direction of poverty change and the geographic area, a total of 25 different settings are distinguished. Table 1 provides a summary (including the level of the poverty headcount in the base year and the observed poverty change). Application of different combinations of poverty predictors in each of these settings further provides the beginning of a database to conduct multi-variate meta-analysis on the importance of the different factors affecting the performance of the SAE poverty prediction methodology.

4 Performance of SAE poverty predictions using expenditures and asset models

The VLSS data allow testing the performance of consumption models using both expenditure subcomponents and household assets to predict poverty (Tables 2 and 3). The following subcategories of total expenditures were considered to estimate the consumption models: expenditures on food excluding rice (model 1); expenditures on all food (model 2); '30-day' nonfood expenditures, which include those expenditures that are asked with a 30-day recall period (model 3);[19] 'one-year' nonfood expenditures with a one year recall period (model 4); total nonfood expenditures (which is simply the sum of the previous two subgroups) (model 5).

In addition, the performance of increasingly elaborate non-consumption asset models is examined, all including geographic indicators. They are augmented either with demographic indicators (model 6); with demographic and educational variables (model 7); with demographic and educational variables as well as housing and consumer durables (model 8); or with housing and consumer durables only (model 9) to mitigate potential bias from changing returns to labour and education.

Growth in Vietnam was not only fast during the period of the survey, but also broad based with poverty falling dramatically, 23.2 percentage points nationwide, and across

19 This is similar to the specification by Deaton in his analysis of India's poverty numbers across NSS rounds (see Deaton 2003).

the board, despite a wide divergence in initial poverty incidence across provinces (from 35.3 per cent in the South Eastern province to 80 per cent in the northern Uplands in 1992/3). Poverty rates in the second period were predicted best by non-rice expenditure, annual nonfood expenditure, total nonfood expenditure, and the full asset models (columns 1, 4, 5, 8 and 9, respectively). In these models, the (absolute) difference between the predicted and observed poverty headcount was less than 3.4 percentage points on average despite declines in the observed poverty headcount measures between 14 and 35 percentage points (Table 2) and the poverty headcount point estimates fell well within the confidence intervals around the observed poverty rates in all but one or two of the ten regions for which poverty was predicted. Considering the poverty gap as poverty measure (Table 3), the picture does not change appreciably.

The SAE procedure appears to do a remarkably good job in tracking the poverty decline in Vietnam, despite a period of dramatic economic transformation. Interestingly, in terms of prediction performance, there is no clear basis for preferring models based on consumption subcomponents versus models based on non-consumption assets and household characteristics. It is noteworthy, however, that excluding rice consumption from the food component improves performance of the prediction model, while especially annual nonfood expenditures drive the performance of the nonfood expenditures models. The former observation is consistent with the lower income elasticity of demand for staple food than for non-staple food. Similarly, items recorded with a one-year recall period often contain more bulky and more expensive goods, with a higher income elasticity. When considering prediction models based on non-consumption assets and household characteristics, the message is to specify as rich a model as possible, with the only possible qualification that characteristics (such as education and demographics) that might be expected to experience changing returns could be omitted at relatively low cost.

From the Russian sample, the encouraging assessment of the SAE prediction approach, based on the Vietnam experience, is significantly tempered. Table 4 shows that only in the case of significant poverty increase (between 1994 and 1998), and only with a model based on food consumption, and only in the case of the headcount poverty rate, does the SAE approach appear to work. There is little systematic evidence that other model specifications work well, even in the context of relatively static poverty (between 1994 and 2003) or of dramatic poverty decline (1998 to 2003). The better performance of the food expenditure models during the first period is consistent with the dramatic decline in food eaten outside the home reported by Wall and Johnston (2008). Not only are non-staple food expenditures likely more sensitive to income declines (and increases), in middle-income countries, they also make up a larger share of overall food expenditure, rendering the food expenditure variable more sensitive to changes in income than in poorer countries. Working with the poverty gap does not alter the overall assessment (Table 5).

Nonetheless, the findings are not altogether surprising in light of the extensive literature studying the evolution of poverty during 1994-2003 based on the RLMS. Unable to truly disentangle measurement error from true transitory shocks, this literature emphasizes the relatively low levels of chronic poverty throughout this time period, highlighting the transitory nature of poverty during the period around Russia’s 1998 financial crisis (Luttmer 2001; Stillman 2001). Focusing on asset-based welfare measures instead, Wall and Johnston (2008) provide further indirect evidence that the consumption data for this period appear particularly noisy. The relatively low R-square

(0.35-0.40) of the different asset-based consumption models estimated here to predict poverty bears this out.

The relatively weak performance of the models in the Russian context, serves as an important reminder that good predictive power is not automatic, with a low R-square of the consumption model a sign counselling caution. Inclusion of more time variant variables could help mitigate such concerns, provided that their effects can be identified from the cross-sectoral variations, likely one of the reasons why the food expenditure variables were better at predicting the increase in poverty between 1994-98 (starting from a low level of poverty) than the subsequent decline between 1998-2003 (starting from a much higher level of poverty).

The findings from rural western China, where poverty also declined substantially, are somewhat more encouraging and more in line with the Vietnam experience, but only in the case of one province: Inner Mongolia, and not at all in the case of Gansu. In Inner Mongolia, models based on expenditure subcomponents do not do well, unlike Vietnam, but the full asset model as well as the asset model that omits demographic and educational characteristics, do well in capturing the dramatic decline in poverty from 19 per cent to roughly 6 per cent in a period of just five years. This assessment is slightly tempered in the case of the poverty gap —with the 2004 prediction based on models 5 and 6 at 1.5 per cent—just slightly outside the confidence interval of 0.4-1.2 on observed poverty for that year. Interestingly, in the case of the poverty gap, a model based on nonfood expenditures does succeed in tracking poverty decline in Inner Mongolia between 2000 and 2004.

In Gansu, however, the SAE approach does not perform well. Poverty is estimated at a much higher rate for the year 2004 than is observed for that year. Although, again, a model based on non-consumption assets and household characteristics does best, the predictions lie far above observed poverty in this province (Tables 6 and 7). The poor performance of the full food expenditure model in these rather poor settings where rice still makes up a substantial share of overall food expenditures is consistent with the earlier results from Vietnam and Russia. The results further suggest that welfare improvements in poorer settings are likely to translate quicker in the purchase of durable assets (and thus nonfood expenditures) than the purchases of non-food staples, which may be more sensitive to income declines, especially at higher income levels.

As a final application, poverty prediction results for Kenya are presented in Tables 8 (poverty incidence) and 9 (poverty gap). Consumption subcomponent models are not pursued here, given some changes in the consumption questionnaire design across the 1997 and 2005/6 surveys. Another key feature of the sample is the dramatic decline in poverty in Nairobi compared with only a slight decrease or stagnation in rural and other urban areas. Nonetheless, as was seen in Vietnam and in Inner Mongolia, the asset model performs well in this setting, again with the general prescription that the model should include housing quality characteristics and ownership of consumer durables (models 3 and 4).

The observed poverty numbers were obtained using a deflator derived from the survey, as opposed to the official CPI. The rather good performance of the (full) asset model in predicting the observed changes in poverty based on the survey deflator provides some support to the use of these survey-based deflators in analysing poverty in Kenya, and underscores the potential of asset-based poverty prediction models in adjudicating such choices.

To further explore the empirical regularities that are emerging from the case by case review of the three (plus one) country cases, a meta-analysis is performed on the poverty prediction results. In particular, a sample of 304 observations was obtained from the different models in the 3+1 country case studies—152 observations for the headcounts and 152 for the poverty gaps. The (absolute value of the) deviation between observed and estimated poverty levels in the second period (expressed in percentage terms)[20] divided by the observed poverty measure in that period, is taken as dependent variable. Using ordinary least squares with error terms corrected for heteroscedasticity at the country level, this dependent variable is subsequently regressed on the characteristics of the consumption model (R-squared and sample size), spell characteristics (the direction of poverty change and the length of the spell interval), the nature of the poverty prediction model (nonfood expenditures and full asset model with the remaining models as default), the poverty measure used and its initial level, and three country indicators (with China the omitted category).

Table 10 indicates that the single most powerful predictor of the success in tracking poverty via the SAE poverty prediction approach is the explanatory power of the underlying consumption model—a 10 percentage point increase in the R-squared is associated with a reduction in the difference between the observed and predicted future poverty by 12.5 per cent. This finding is not unexpected, and strongly confirms that the appeal of applying such techniques hinges on the kind of variables that are available to include in the model specification as well as the strength of their association with consumption.

As was seen during the country by country review, the nonfood models often reduce the prediction error compared with the other models. Use of the full asset model (compared with partial asset models or the food models), also further increases accuracy, but in the full sample of countries explored here, this effect is not statistically significant, because of the poor performance of the asset model in the Russian context. Re-estimation without Russia suggests that both the full asset and the nonfood expenditure models reduce the difference between predicted and observed poverty by a similar amount (i.e., between 25 and 30 per cent). Other possible correlates considered in this model are not statistically significant. It is likely that with further expansion of the database underpinning this meta-analysis, more nuances would emerge, in terms of our understanding of where SAE methods can be best applied.

5 Concluding remarks

The absence of comparable consumption data and price deflators at regular intervals has instigated the development of alternative methods to study the evolution of poverty over time. In essence these methods track a series of consumption correlates, instead of consumption itself. The correlates are mapped into consumption using an empirically

[20] To be precise, $(W^s(c_{t+k}) - \tilde{\mu}^s_{t+k}) / W^s(c_{t+k})$ is taken as dependent variable.

calibrated relationship between the two. Success of this approach hinges critically on the assumed stability of this relationship over time. But such an assumption is difficult to verify, and has so far gone untested. Until the performance of these models in predicting changes in poverty is scrutinized with actual data, one must be very circumspect with the application of such techniques in practice.

This paper provides the first step at filling this void, drawing on data from three surveys with highly comparable expenditure data, further complemented with a case study from Kenya, thus covering a wide range of different settings, periods of great structural change, and quite divergent poverty trajectories. An adapted version of the SAE technique described in Elbers, Lanjouw and Lanjouw (2002, 2003) is implemented. Consumption prediction models using consumption subcomponents and different combinations of non-consumption assets are tested, in effect using specifications that are plausible given the kind of data that are commonly available in most living standards surveys, light welfare monitoring surveys, and even in the many regularly conducted demographic and health surveys.

The success of the SAE approach varied across settings. In Vietnam, a variety of models, comprising both consumption subcomponents (non-rice food, nonfood) as well as non-consumption assets and household characteristics (particularly the full asset model) worked very well. This success is striking in light of the very deep structural transformation that Vietnam was going through between 1992/3 and 1997/8—transformations that would lead one to expect that parameter stability would fail to hold. In Russia, on the other hand, the SAE approach was much less successful. Part of the explanation for the poor performance in Russia may lie in the fact that the period under consideration was marked by a very sizeable economic shock leading to a sharp rise in poverty between 1994 and 1998 and then a similarly sized decline. The asset-based model specification, in particular, was found to not work well in this setting—suggesting that household decisions to acquire or draw down assets may be only loosely associated with macro events that are sudden and judged to be transitory. Other explanations may lie with the reported concerns in the literature about measurement error in the Russian consumption data (reflected also in the relatively low explanatory power of the Russian consumption models). In China, the prediction models worked well in one province, Inner Mongolia, but not at all in Gansu. In the Inner Mongolia setting, it is the asset models in particular that worked well. Different poverty measures did not affect performance of the prediction models.

Meta-analysis, bringing together the performance of all models in all of the settings considered here, confirms that the key determinant of success in producing reliable estimates based on these prediction methods is explanatory power in the basic consumption model. Controlling for this criterion, the only additionally significant determinant of accuracy in the prediction of poverty is the choice of the poverty predictors, with both nonfood expenditures and full asset models reducing prediction error compared to the other models. While expenditure subcomponent models may be appealing, they are not likely to be available in most settings where there are concerns about data non-comparability. More practical are models based on non-consumption assets, as such information is likely to be available, and comparable, even across otherwise non-comparable data. Given that the consumption subcomponent models still require appropriate price deflators, which can be hard to come by, the full asset models may also be more convenient from this perspective. The Kenya illustration further reveals that the SAE method outlined in this paper can be helpful in adjudicating between alternative (and conflicting) price deflators.

In conclusion, while collecting comparable consumption data and constructing reliable deflators remain the preferred options, these results provide cautious optimism that in the absence of (comparable) consumption and deflator data, poverty can still be tracked by tracking poverty predictors. They also indicate that further validation of the parameter stability assumption in more settings, for shorter and longer time periods, and especially in settings of rapid poverty deterioration is needed before the methodology can be promoted as a reliable proxy (or potentially even as cheaper alternative) on a larger scale. The additional explanatory power of more time variant variables, such as rainfall data in agriculture dependent settings, but also health variables and subjective poverty indicators deserves particular attention in this regard.

Works cited

Agarwal, N., D. Dollar, and P. Glewwe (eds) (2004). *Economic Growth, Poverty and Household Welfare in Vietnam*. Washington, DC: World Bank.

Alderman, H., M. Babita, G. Demombynes, N. Makhatha, and B. Özler (2002). ‘How Low Can You Go? Combining Census and Survey Data for Mapping Poverty in South Africa’. *Journal of African Economies*, 11 (2): 169-200.

Azzarri, C., G. Carletto, B. Davis, and A. Zezza (2006). -Monitoring Poverty Without Consumption Data: An Application using the Albania Panel Survey’. *Eastern European Economics*, 44 (1): 59-82.

Chesher, A., and C. Schluter (2002). ‘Welfare Measurement and Measurement Error’. *Review of Economic Studies*, 69 (2): 357-78.

Deaton, A. (2003). ‘Adjusted Indian Poverty Estimates for 1999-2000’. *Economic and Political Weekly*, 25 January: 322-6.

Deaton, A. (2008) ‘Price Trends in India and Their Implications for Measuring Poverty’. *Economic and Political Weekly*, 7 September: 3729-48.

Deaton, A., and Z. Salman (2002). ‘Guidelines for Constructing Consumption Aggregates for Welfare Analysis’. WB Living Standards Measurement Study Working Paper 135, Washington, DC: World Bank.

Deaton, A., and V. Kozel (2005). ‘Data and Dogma: The Great Indian Poverty Debate’. *The World Bank Research Observer*, 20 (2): 177-99.

Elbers, C., J. O. Lanjouw, and P. Lanjouw (2002). ‘Micro-Level Estimation of Welfare’. WB Policy Research Working Paper 2911. Washington, DC: World Bank.

Elbers, C., J. O. Lanjouw, and P. Lanjouw (2003). ‘Micro-Level Estimation of Poverty and Inequality’. *Econometrica*, 71 (1): 355-64.

Elbers, C., and N. Pouw (2009). ‘Modelling Sequencing Patterns in Asset Acquisition: the Case of Smallholder Farmers in Three Rural Districts of Uganda’. Amsterdam: Institute of International Development. Mimeo.

Filmer, D., and S. Kinnon (2008). ‘Assessing Asset Indices’. WB Policy Research Working Paper 4605. Washington, DC: World Bank.

Gibson, J., H. Jikun, and R. Scott (2003). 'Improving Estimates of Inequality and Poverty from Urban China's Household Income and Expenditure Survey'. *Review of Income and Wealth*, 49 (1): 53-68.

Grimm, M., and I. Günther (2006). 'Growth and Poverty in Burkina Faso – A Reassessment of the Paradox'. *Journal of African Economies*, 16 (1): 70-101.

Grosse, M., S. Klasen, and J. Spatz (2009). 'Matching Household Surveys with DHS Data to Create Nationally Representative Time Series of Poverty: An Application to Bolivia'. Courant Research Centre Discussion Paper 21. Göttingen: Georg-August-Universität.

Harrower, S., and J. Hoddinott (2005). 'Consumption Smoothing in the Zone Lacustre, Mali'. *Journal of African Economies*, 14 (4): 489-519.

Hoogeveen, J., and R. Ruhinduka (2009). 'Lost in Transition? Income Poverty Reduction in Tanzania since 2001'. Background paper to the Tanzanian Population and Human Development Report 2009. ?? check names

Kijima, Y., and P. Lanjouw (2003). 'Poverty in India during the 1990s: A Regional Perspective'. WB Policy Research Department Working Paper 3141. Washington, DC: World Bank.

Lanjouw, J. O., and P. Lanjouw (2001) 'How to compare Apples and Oranges: Poverty Measurement Based on Different Definitions of Consumption'. *Review of Income and Wealth*, 47 (1): 25-42.

Luttmer, E. (2001). 'Measuring Poverty Dynamics and Inequality in Transition Economies – Disentangling Real Events from Noisy Data'. WB World Bank Policy Research Working Paper 2549. Washington, DC: World Bank.

Mathiassen, A. (2009). 'A Model Based Approach for Predicting Annual Poverty Rates Without Expenditure Data'. *Journal of Economic Inequality*, 7 (2): 117-35.

Ravallion, M. (1988) 'Expected Poverty under Risk-Induced Welfare Variability'. *Economic Journal*, 98 (393): 1171-82.

Ravallion, M. (1996). 'How Well Can Method Substitute for Data? Five Experiments in Poverty Analysis'. *World Bank Research Observer*, 11 (2): 199-221.

Sahn, D., and D. Stifel (2000). 'Poverty Comparisons over Time and across Countries in Africa'. *World Development*, 28 (12): 2123-55.

Sen, A., and Himanshu (2004). 'Poverty and Inequality in India-I'. *Economic and Political Weekly,* 8 September: 4247-63.

Stifel, D., and L. Christiaensen (2007). 'Tracking Poverty over Time in the Absence of Comparable Consumption Data'. *World Bank Economic Review,* 21 (2): 317-41.

Stillman, S. (2001) 'The Response of Consumption in Russian Households to Economic Shocks'. WDI Working Papers Series 412. Ann Arbor: William Davidson Institute at the University of Michigan Stephen M. Ross Business School.

Tarozzi, A. (2007) 'Calculating Comparable Statistics from Incomparable Surveys, With an Application to Poverty in India'. *Journal of Business and Economic Statistics*, 25 (3): 314-36.

Tarozzi, A., and A. Deaton (2009). 'Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas'. *Review of Economics and Statistics,* 91 (4): 773-92.

Wall, M., and D. Johnston (2008). 'Counting Heads or Counting Televisions: Can Asset-Based Measures of Welfare Assist Policy-makers in Russia?'. *Journal of Human Development and Capabilities*, 9 (1): 131-47.

World Bank (2008). *Kenya Poverty and Inequality Assessment: Volume I: Synthesis Report*. Report No. 44190-Ke. Poverty Reduction and Economic Management Unit, Africa Region. Washington, DC: World Bank.

Acronyms

CPI	consumption price indices
DHS	demographic and health surveys
KIHBS	Kenya integrated household budget survey
NSS	national sample survey
RLMS	Russian longitudinal monitoring surveys
SAE	small area estimation technique
SSA	sub-Saharan Africa
VLSS	Vietnam living standards survey
WMS	welfare monitoring survey

Table 1
SAE poverty prediction technique tested in a multitude of settings

Poverty headcount (level (%), %point change)	Rural		Urban		Province		National		Total # obs.
Increase	RU94-98	(13.1%, 21.7%)	RU94-98	(10.6%, 22.7%)			RU94-98	(11.4%, 22.4%)	4
	RU94-03	(13.1%, 4.3%)							
Stagnation or modest change	KE97-05	(52.8%, -3.1%)	RU94-03	(10.6%, -2.5%)			RU94-03	(11.4%, -0.3%)	4
(-4%point change, +4% change)			KE97-05						
			Other urban	(43.2%, -0.5%)					
Decrease	VN92-97	(68.5%, -23.6%)	VN92-97	(28.6%, -18.6%)	VN92-97		VN92-97	(60.6%, -23.2%)	17
	RU98-03	(34.8%, -17.4%)	RU98-03	(33.3%, -25.2%)	Northern Uplands	(80.0%, -21.4%)	RU98-03	(33.8%, -22.7%)	
	GS00-04	(24.3%, -12.5%)	KE97-05		Red River Delta	(64.0%, -35.3%)	KE97-05	(50.8%, -4.2%)	
	IM00-04	(19.05%, -12.8%)	Nairobi	(40.0%, -19.4%)	North Central	(76.6%, -28.5%)			
					Central Coast	(53.2%, -18.0%)			
					Central Highlands	(72.9%, -20.5%)			
					South East	(35.3%, -27.7%)			
					Mekong River	(51.0%, -14.1%)			
Total # observations	7		6		7		5		25

Note: VN92-97 = Vietnam 1992/93-1997/8; RU94-98= Russia 1994-98; RU98-03=Russia 1998-03; RU94-03= Russia 1994-2003; GS00-04=China Gansu 2000 –2004; IM00-04=China Inner Mongolia 2000-2004; KE97-05=Kenya 1997-2005/6.

Table 2
Nonfood expenditures and the more complete asset models predict change in poverty headcount best in Vietnam 1992/3-1997/8

Poverty headcount (%) (standard error)	Observed levels		SAE predicted poverty levels in 1997/8								
Included in the model	1992/3	1997/8	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expenditure subcomponents											
Food: Nonrice			x	x	-	-	-	-	-	-	-
Food: Rice			-	x	-	-	-	-	-	-	-
Nonfood: 30-day			-	-	x	-	x	-	-	-	-
Nonfood: Annual			-	-	-	x	x	-	-	-	-
Non-consumption assets											
Geographic			-	-	-	-	-	x	x	x	x
Demographics			-	-	-	-	-	x	x	x	-
Education/profession			-	-	-	-	-	-	x	x	-
Housing quality			-	-	-	-	-	-	-	x	x
Consumer durables			-	-	-	-	-	-	-	x	x
National	**60.6**	**37.4**	39.3***	47.2	41.9	35.8***	33.3	54.6	55.6	38.2***	36.7***
	1.9	*1.6*									
Rural	**68.5**	**44.9**	46.6***	59.7	51.3	47.3***	41.0***	63.5	64.8	48.5***	44.2***
	1.7	*2.0*									
Urban	**28.6**	**9.0**	11.5***	14.5	13.0	8.6***	9.2***	21.7	22.8	11.8***	9.7***
	4.1	*1.5*									
Northern Uplands	**80.0**	**58.6**	67.6***	72.7	65.3***	58.4***	53.3***	76.0	78.4	62.3***	57.0***
	3.8	*5.6*									
Red River Delta	**64.0**	**28.7**	29.3***	45.4	41.1	34.6***	25.7***	57.2	57.4	32.5***	32.5***
	4.6	*3.4*									
North Central	**76.6**	**48.1**	55.1***	63.5	67.4	56.2***	51.3***	73.1	72.8	48.1***	47.9***
	4.1	*5.2*									
Central Coast	**53.2**	**35.2**	34.4***	47.3	39.9***	32.5***	35.7***	47.0	49.3	34.0***	31.9***
	6.0	*5.5*									
Central Highlands	**72.9**	**52.4**	54.5***	64.3***	64.3***	45.7***	47.4***	66.2***	64.0***	51.5***	49.2***
	13.9	*9.7*									
South East	**35.3**	**7.6**	11.3	14.9	12.8	10.2***	8.4***	27.3	28.6	12.3	16.8
	6.2	*1.5*									
Mekong River	**51.0**	**36.9**	38.6***	47.2	34.6***	33.8***	32.9***	42.7***	43.6	34.2***	30.7
	6.2	*3.0*									
No .of times difference NOT statistically different			9	1	4	10	9	2	1	9	8
Average absolute difference			3.1	11.8	7.8	3.4	3.0	17.1	17.9	2.4	3.0
# observed poverty ≥ predicted poverty			1	0	1	6	6	0	0	4	7
# observed poverty ≥ predicted poverty			9	10	9	4	4	10	10	6	3

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5% level from the observed poverty estimates, i.e., they fall within the 95% confidence interval around the observed poverty rates.

Table 3
Nonfood expenditures and the more complete asset models predict change in poverty gap best in Vietnam 1992/3-1997/8

Poverty gap (standard error)	Observed levels		SAE predicted poverty levels in 1997/8								
Included in model	1992/3	1997/8	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expenditure subcomponent											
Food: Nonrice			x	x	-	-	-	-	-	-	-
Food: Rice			-	x	-	-	-	-	-	-	-
Nonfood: 30-day			-	-	x	-	x	-	-	-	-
Nonfood: Annual			-	-	-	x	x	-	-	-	-
Non-consumption assets											
Geographic			-	-	-	-	-	x	x	x	x
Demographics			-	-	-	-	-	x	x	x	-
Education/profession			-	-	-	-	-	-	x	x	-
Housing quality			-	-	-	-	-	-	-	x	x
Consumer durables			-	-	-	-	-	-	-	x	x
National	**19.0**	**9.5**	9.9***	12.4	11.6	8.6***	8.3***	16.5	17.2	10.5***	9.4***
	0.9	*0.7*									
Rural	**22.0**	**11.6**	11.7***	16.4	14.8	11.8***	9.8***	19.1	19.7	13.7	11.5***
	1.0	*0.9*									
Urban	**7.3**	**1.7**	2.3***	3.1	3.1	1.8***	2.1***	5.4	5.6	2.6	1.9***
	1.2	*0.3*									
Northern Uplands	**26.7**	**16.8**	21.3***	23.0	20.0***	16.1***	14.7***	23.6	25.5	18.0***	15.9***
	2.6	*2.3*									
Red River Delta	**18.9**	**5.7**	6.1***	11.3	10.5	7.3***	5.2***	15.4	15.7	7.3***	7.3***
	1.9	*1.0*									
North Central	**25.3**	**11.8**	14.1***	17.5	20.3	14.3***	12.2***	22.7	23.0	12.9***	12.3***
	2.7	*1.9*									
Central Coast	**17.7**	**10.6**	9.0***	13.6***	12.4***	8.1***	9.7***	15.2***	15.7***	10.7***	9.5***
	3.2	*3.1*									
Central Highlands	**27.5**	**19.1**	15.3***	19.3***	22.2***	14.6***	16.3***	25.3***	22.9***	21.8***	17.1***
	8.5	*5.9*									
South East	**9.8**	**1.3**	2.3	3.0	5.8	2.1	1.8***	7.4	7.5	2.9	3.8
	2.0	*0.3*									
Mekong River	**15.0**	**8.1**	9.3***	11.8	9.5***	8.9***	8.2***	11.7	11.8	9.3***	7.5***
	1.5	*0.9*									
No. of times difference NOT statistically different			9	2	4	9	10	2	2	7	9
Average absolute difference			1.6	3.5	3.4	1.5	1.17	6.6	6.8	1.4	1.0
# observed poverty ≥ predicted poverty			2	0	0	4	6	0	0	0	6
# observed poverty < predicted poverty			8	10	10	6	4	10	10	10	4

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5% level from the observed poverty estimates, i.e., they fall within the 95% confidence interval around the observed poverty rates.

Table 4
Models have low predictive power in Russia: headcount

Poverty headcount (%) (standard error)	Observed levels		SAE predicted poverty levels in period 2					
Included in the model	Period 1	Period 2	(1)	(2)	(3)	(4)	(5)	(6)
Expenditure subcomponents								
Food expenditures			x	-	-	-	-	-
Nonfood expenditures			-	x	-	-	-	-
Non-consumption assets								
Geographic			-	-	x	x	x	x
Demographic			-	-	x	x	x	-
Education/profession			-	-	-	x	x	-
Housing quality			-	-	-	-	x	x
Consumer durables			-	-	-	-	x	x
Region	1994	1998						
National	**11.4**	**33.8**	35.0 ***	24.7	11.7	11.7	14.1	12.7
	0.6	*1.1*						
Rural	**13.1**	**34.8**	34.4 ***	22.7	14.0	15.9	22.4	18.2
	1.3	*2.0*						
Urban	**10.6**	**33.3**	34.2 ***	26.9	15.2	17.8	18.8	17.4
	0.7	*1.3*						
	1994	2003						
National	**11.4**	**11.1**	22.0	11.3 ***	9.8	8.2	8.5	8.4
	0.6	*0.6*						
Rural	**13.1**	**17.4**	22.6	11.8	11.2	11.3	9.9	13.1 ***
	1.3	*1.5*						
Urban	**10.6**	**8.1**	19.0	10.9	12.1	11.2	9.2 ***	11.2
	0.7	*0.6*						
	1998	2003						
National	**33.8**	**11.1**	24.4	15.6	30.1	28.1	26	31.7
	1.1	*0.6*						
Rural	**34.8**	**17.4**	29.8	21.6	32.0	30.4	30.6	38.4
	2	*1.5*						
Urban	**33.3**	**8.1**	18.5	14.8	27.7	29.2	27.3	30.0
	1.3	*0.6*						
No. of times difference NOT statistically different			3	1	0	0	1	1
Average absolute difference			7.3	5.7	14.0	13.3	11.7	14.1
# observed poverty ≥ predicted poverty			1	4	5	5	5	5
# observed poverty < predicted poverty			8	5	4	4	4	4

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5% level from the observed poverty estimates, i.e., they fall within the 95% confidence interval around the observed poverty rates.

Table 5
Models have low predictive power in Russia: poverty gap

Poverty gap (standard error)	Observed levels		SAE predicted poverty levels in period 2					
Included in the model	Period 1	Period 2	(1)	(2)	(3)	(4)	(5)	(6)
Expenditure subcomponents								
Food expenditures			x	-	-	-	-	-
Nonfood expenditures			-	x	-	-	-	-
Non-consumption assets								
Geographic			-	-	x	x	x	x
Demographic			-	-	x	x	x	-
Education/profession			-	-	-	x	x	-
Housing quality			-	-	-	-	x	x
Consumer durables			-	-	-	-	x	x
Region	1994	1998						
National	**3.8**	**12.9**	15.6	8.9	3.3	3.3	4.1	4.1
	0.2	*0.5*						
Rural	**4.1**	**13.2**	15.3	7.6	3.9	4.9	7.8	6.2
	0.5	*0.9*						
Urban	**3.7**	**12.7**	14.4	10.1	4.3	5.5	6.0	5.5
	0.3	*0.6*						
	1994	2003						
National	**3.8**	**3.6**	8.5	3.7 ***	2.8	2.1	2.3	2.6
	0.2	*0.2*						
Rural	**4.1**	**6.0**	8.3	3.5	3.1	3.2	2.8	4.3
	0.5	*0.6*						
Urban	**3.7**	**2.4**	6.3	3.4	3.4	3.2	2.6 ***	3.3
	0.3	*0.2*						
	1998	2003						
National	**12.9**	**3.6**	9.1	5.0	10.3	9.5	8.7	12.5
	0.5	*0.2*						
Rural	**13.2**	**6.0**	11.5	7.5	10.6	10.2	10.2	16.3
	0.9	0.6						
Urban	**12.7**	**2.4**	6.0	4.5	8.8	6.9	6.9	8.5
	0.6	*0.2*						
No. of times difference NOT statistically different			0	1	0	0	1	0
Average absolute difference			3.6	2.3	5.5	5.0	4.4	5.8
# observed poverty ≥ predicted poverty			0	4	5	5	5	5
# observed poverty < predicted poverty			9	5	4	4	4	4

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5% level from the observed poverty estimates, i.e., they fall within the 95% confidence interval around the observed poverty rates.

Table 6
More complete asset models (with and without demographic/education variables) perform well in Inner Mongolia in predicting poverty headcount, though not in Gansu

Poverty headcount (%) (standard error)	Observed levels		SAE predicted poverty levels in 2004					
Included in the model	2000	2004	(1)	(2)	(3)	(4)	(5)	(6)
Expenditure subcomponents								
Food expenditures			x	-	-	-	-	-
Nonfood expenditures			-	x	-	-	-	-
Non-consumption assets								
Geographic			-	-	x	x	x	x
Demographic			-	-	x	x	x	-
Education/profession			-	-	-	x	x	-
Housing quality			-	-	-	-	x	x
Consumer durables & agricultural assets			-	-	-	-	x	x
Inner Mongolia	**19.0**	**6.2**	10.8	9.8	18.5	18.2	7.3***	7.8***
	1.5	*0.9*						
Gansu	**24.3**	**11.8**	21.9	28.4	23.8	25.6	18.9	20.9
	1.8	*1.4*						
No. of times difference NOT statistically different			*0*	*0*	*0*	*0*	*1*	*1*
Average absolute difference			7.3	10.1	12.2	12.9	4.1	5.3
# observed poverty ≥ predicted poverty			0	0	0	0	0	0
# observed poverty < predicted poverty			2	2	2	2	2	2

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5 per cent level from the observed poverty estimates, i.e., they fall within the 95 per cent confidence interval around the observed poverty

Table 7
Nonfood expenditures and more complete asset models (with and without demographic/education variables) perform similarly well in Inner Mongolia in predicting poverty gap, though not in Gansu.

Poverty gap (standard error)	Observed levels		SAE predicted poverty levels in 2004					
Included in the model	2000	2004	(1)	(2)	(3)	(4)	(5)	(6)
Expenditure subcomponents								
Food expenditures			x	-	-	-	-	-
Nonfood expenditures			-	x	-	-	-	-
Non-consumption assets								
Geographic			-	-	x	x	x	x
Demographic			-	-	x	x	x	
Education/profession			-	-	-	x	x	
Housing quality			-	-	-	-	x	x
Consumer durables & agricultural assets			-	-	-	-	x	x
Inner Mongolia	**3.9**	**0.8**	2.0	1.2 ***	3.9	3.7	1.5	1.5
	0.4	*0.2*						
Gansu	**4.9**	**1.8**	3.6	6.1	5.0	5.6	4.2	4.5
	0.5	*0.3*						
No. of times difference NOT statistically different			*0*	*1*	*0*	*0*	*0*	*0*
Average absolute difference			1.5	2.4	3.1	3.3	1.5	1.7
# observed poverty ≥ predicted poverty			0	0	0	0	0	0
# observed poverty < predicted poverty			2	2	2	2	2	2

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5% level from the observed poverty estimates, i.e., they fall within the 95% confidence interval around the observed poverty rates

.

Table 8
Full asset model predicts headcount changes best in Kenya

Poverty headcount (standard error)	Observed levels		SAE predicted poverty levels in 2005/6			
Included in model	1997	2005/6	(1)	(2)	(3)	(4)
Non-consumption assets						
Geographic			x	x	x	x
Demographics			x	x	x	-
Education/profession			-	x	x	-
Housing quality			-	-	x	x
Consumer durables			-	-	x	x
National	**50.8**	**46.6**	45.6 ***	45.1	43.1	45.5 ***
	1.1	*0.6*				
Rural	**52.8**	**49.7**	50.5 ***	45.9	48.7 ***	44.9
	2.0	*0.7*				
Other urban	**43.2**	**42.7**	41.2 ***	46.7	45.5 ***	40.4 ***
	2.6	*1.6*				
Nairobi	**40.0**	**20.6**	34.0	28.6	24.8 ***	20.1 ***
	4.5	*2.5*				
No. of times difference NOT statistically different			*3*	*0*	*3*	*3*
Average absolute difference			4.2	4.4	2.9	2.2
# observed poverty ≥ predicted poverty			2	2	2	4
# observed poverty < predicted poverty			2	2	2	0

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5% level from the observed poverty estimates, i.e., they fall within the 95% confidence interval around the observed poverty rates.

Table 9
Full asset model predicts changes in poverty gap best in Kenya

Poverty gap (standard error)	Observed levels		SAE predicted poverty levels in 2005/6			
Included in model	1997	2005/6	(1)	(2)	(3)	(4)
Non-consumption assets						
Geographic			x	x	x	x
Demographics			x	x	x	-
Education/profession			-	x	x	-
Housing quality			-	-	x	x
Consumer durables			-	-	x	x
National	**16.2**	**16.6**	15.0	14.2	13.0	14.1
	0.5	*0.3*				
Rural	**22.3**	**17.8**	15.7	15.5	19.1	25.6
	2.3	*0.3*				
Other urban	**14.5**	**14.9**	13.4	16.1**	15.2***	12.5
	1.3	*0.7*				
Nairobi	**11.4**	**6.2**	10.9	8.0**	6.4***	5.2***
	2.2	*0.9*				
No of times difference NOT statistically different			*0*	*2*	*2*	*1*
Average absolute difference			2.5	1.9	1.4	3.4
# observed poverty ≥ predicted poverty			3	2	1	3
# observed poverty < predicted poverty			1	2	3	1

Note: *** denotes that the predicted poverty point estimates are not statistically different at the 5% level from the observed poverty estimates, i.e., they fall within the 95% confidence interval around the observed poverty rates.

Table 10
The better the predictive power of the underlying consumption model the more accurate the poverty prediction

Absolute value of the difference between predicted and observed poverty in period 2 divided by observed poverty in period 2	Model (1) OLS with robust s.e. at country level	
	Coef.	p-value
Characteristics prediction model		
R-squared of consumption model	-125.455	0.05
# of observations	0.003	0.20
Spell characteristics		
≥ 4% increase in poverty incidence	-50.76	0.48
≥ 4% decrease in poverty incidence	66.37	0.37
Years between surveys	-9.39	0.40
Poverty prediction model		
Full asset model	-19.83	0.16
Nonfood expenditure model	-35.83	0.04
Geographic area		
Rural sample	5.77	0.75
Urban	20.89	0.28
Province	31.35	0.21
Poverty measure		
Headcount (1=yes; 0=povgap)	-25.69	0.31
Headcount level	-1.44	0.14
Poverty Gap level	-4.82	0.17
Country dummy (China omitted)		
Vietnam	1.93	0.97
Russia	41.59	0.45
Kenya	-9.42	0.89
Constant	173.68	0.204
# observations	304	
R^2/Pseudo R^2	0.48	