WPS5683

Policy Research Working Paper 5683

Small Area Estimation-Based Prediction Methods to Track Poverty

Validation and Applications

Luc Christiaensen
Peter Lanjouw
Jill Luoto
David Stifel

The World Bank
Development Research Group
Poverty and Inequality Team
June 2011

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. Price deflators are also often missing or disputed. In response, poverty prediction methods that track consumption correlates as opposed to consumption itself have been developed. These methods typically assume that the estimated relation between consumption and its predictors is stable over time—an assumption that cannot usually be tested directly. This study analyzes the performance of poverty prediction models based on small area estimation techniques. Predicted poverty estimates are compared with directly observed levels in two country settings where data comparability over time is not a problem. Prediction models that employ either non-staple food or non-food expenditures or a full set of assets as predictors are found to yield poverty estimates that match observed poverty well. This offers some support to the use of such methods to approximate the evolution of poverty. Two further country examples illustrate how an application of the method employing models based on household assets can help to adjudicate between alternative price deflators.

This paper is a product of the Poverty and Inequality Team, Development Research Group, and relates to the Equity and Development research program. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at lchristiaensen@worldbank.org.

Small Area Estimation-Based Prediction Methods to Track Poverty: Validation and Applications[1]

Luc Christiaensen, Peter Lanjouw, Jill Luoto, and David Stifel

Key words: consumption prediction, price deflator, poverty dynamics, small area estimation

JEL codes: D12, D63, I32

[1] Luc Christiaensen, Senior Economist, World Bank, corresponding author, lchristiaensen@worldbank.org, Peter Lanjouw, Research Manager of the Poverty Group in the Development Economics Research Group of the World Bank, planjouw@worldbank.org, Jill Luoto, Associate Economist, Rand Corporation jluoto@rand.org, David Stifel, Associate Professor, Lafayette College, Pennsylvania, stifeld@lafayette.edu.
The paper has benefited from comments by two anonymous referees and 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.

1 Challenges in Tracking Poverty

Interest in understanding how poverty evolves over time is longstanding. It has received additional impetus through the call 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[2] 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; Beegle, de Weerdt, Friedman and Gibson, 2010). 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

[2] Read on April 2010 http://iresearch.worldbank.org/PovcalNet/jsp/CChoiceControl.jsp?WDI_Year=2007

calculated directly from the surveys, with little information available to adjudicate the choice.[3]

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.[4] 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 a 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 percent in Round 50 of the National Sample Survey (NSS) (1993/1994) to 26 percent in Round 55 (1999/2000), or about a reduction in the number of poor people by 60 million.

However, these official numbers were received with skepticism. 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). Different poverty calculations were proposed, each of them

[3] For instance, while households in Tanzania faced price increases of 93 percent between 2000/01 and 2007 according to the national Household Budget Surveys, the official CPI recorded only a price increase of 45 percent (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).

[4] 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).

predicated on assumptions that were difficult to test. In contrast to the official estimates, one widely circulated alternative put 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 5 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) exploited the fact that the section of the consumption module that pertained to "30-day" expenditures, had not changed between rounds. He estimated 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 sub-components. Sen and Himanshu (2004) examined these assumptions in detail, and showed them to be far from innocuous.

Kijima and Lanjouw (2003) considered an alternative poverty prediction method. Drawing on a small-area estimation (SAE) approach introduced by Elbers, Lanjouw and Lanjouw (2002, 2003) for the purpose of developing “poverty maps”, they also used a subset of explanatory variables that were strictly comparable between survey rounds, but confined 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, ruling out possible changes in the "returns" to factors such as education and labor.[5] This too is a controversial assumption, especially in fast growing economies such as India.

Going one step further, Tarozzi (2007) used both the 30-day consumption items and non-consumption variables such as educational status and land as predictors. He tested the validity of the stable parameter assumption on the 30-day consumption items and on the non-consumption variables, using the much smaller NSS rounds that are fielded during the intervals between the large, "quinquennial" rounds that underpin the official poverty estimates. Tarozzi found 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 remained disputed because the year-to-year poverty changes implied by his calculations were difficult to accept. Concerns were expressed as to how well suited the "thin" rounds were 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 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

[5] Strictly speaking the approach requires that the relationship between consumption and its predictors is stable at the overall, equation, level. Changes in returns to different factors can be accommodated as long as these are offsetting.

Surveys (DHS) in 1993, 1998 and 2003. Stifel and Christiaensen (2007) estimated the relationship between assets and consumption in the 1997 national household budget survey and subsequently applied these estimates to comparable asset data in each of the DHS surveys 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, the authors excluded household assets whose returns were considered more prone to change over time, such as labor 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.[6]

Clearly, empirical validation of the different model specifications and their underlying parameter stability assumptions is necessary before the kind of poverty prediction techniques described above can be routinely used. Such empirical verification 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 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 the SAE-based approach also employed by Kijima and Lanjouw (2003), Christiaensen and Stifel (2007)

[6] 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 was available in all periods.

and Grosse, Klasen and Spatz (2009). 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 this study, we use repeated cross-sections with highly comparable survey and questionnaire design from Vietnam (Vietnam Living Standards Surveys (VLSS) of 1992/1993 and 1997/1998) and rural household panel data (2000-2004) from two western provinces in China—Gansu and Inner Mongolia—to assess the validity of the SAE-based poverty prediction methodology. These two settings span periods of deep structural change, accompanied by marked reductions in poverty. At first glance one would not expect model stability in such settings, and so these applications put the prediction techniques to a demanding test. In addition, the paper presents two applications, for Russia and Kenya respectively. Rather than providing a validation exercise, the two applications demonstrate how the poverty prediction methods considered here can help to confront comparability issues arising from problematic temporal cost-of-living indices.

The results are quite encouraging. In Vietnam the poverty prediction method works quite well both with models using certain expenditure components (particularly non-rice food spending and non-food spending) and with comprehensive models specified on the basis of non-consumption assets. Similarly, in rural Gansu and Inner Mongolia, models based on non-expenditure assets work fairly consistently, while models using certain expenditure subcomponents also work satisfactorily in some instances. The broad conclusion is that the general approach appears to work, and its underlying assumptions of stable parameters seem to hold in these two “test” settings of

rapid poverty reduction and structural change, particularly when models using non-consumption assets are employed.

Application of the method further lends some credence to a recent literature that argues that official inflation adjustments proposed for Russia during a period of deep crisis in the late 1990s, overstated cost-of-living increases and attendant increases in poverty. The illustration for Kenya indicates that cost-of-living deflators for the early 2000s derived directly from household survey data appear more credible than official CPI deflators, resulting in an assessment that poverty declined moderately during the period between 1997 and 2005/6.

We end with a preliminary meta-analysis in which we examine whether there are country or other context-specific circumstances that are correlated with the success of poverty prediction methods. The analysis provides some useful pointers, highlighting in particular the importance of high explanatory power of the consumption model (as reflected in high R-squareds), which turns out to be a defining characteristic of both the comprehensive asset model as well as the non-staple food and nonfood expenditure models.

The paper proceeds in section 2 with a brief review of the SAE methodology and the theoretical considerations in choosing consumption predictors (i.e. the consumption subcomponents and the non-consumption asset variables). Section 3 describes the case study data in more detail and examines the validity of the approach in Vietnam and China. Application of the method to assess alternative positions regarding the rate of change of prices over time in the two additional case study settings of Russia and Kenya

is presented in section 4. The insights emerging from the meta-analysis are highlighted in Section 5. Section 6 offers concluding remarks.

2 Methodological Considerations in Tracking Poverty via Poverty Predictors

The Adapted Small Area Estimation Technique

Following Kijima and Lanjouw (2003) and Stifel and Christiaensen (2007) we employ an adapted version of the small area estimation (SAE) methodology originally proposed by Elbers, Lanjouw and Lanjouw (2002, 2003) to impute a definition of consumption from one household survey into the other. Consider two household surveys covering a given population in two separate time periods, designated round 1 and round 2. The core intuition behind the adapted SAE methodology is to predict per capita consumption at the level of the household in a survey fielded at time $t+1$ using the available information on these households at time $t+1$ (e.g. consumption subcomponents and/or non-consumption assets) as well as the parameter estimates derived from a model of consumption estimated using a survey fielded at time t (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, but assumes 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, 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

consumption c at time t.[7] Consider a log linear approximation to household consumption c_t:

$$\ln c_t = x_t' \beta_t + u_t \tag{1}$$

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 heteroskedastic error term. Since only x_{t+1} is observed, not c_{t+1}, household disturbances u_{t+1} of $\ln c_{t+1} = x_{t+1}' \beta_{t+1} + u_{t+1}$ are always unknown (only the distribution of u_t is known), and the expected value of W is taken, given x_{t+1} and the model parameters of (1), i.e. $\mu_{t+1}^s = E[W^s(x_{t+1}, \beta_{t+1}, u_{t+1})]$ as opposed to $W(c_{t+1})$. The superscript 's' indicates that the expectation is conditional on a *sample* of the households in the same geographical area in period $t+1$ rather than a *census* of the households (in period t) as in the ELL (2002, 2003) poverty mapping application.

Consistent estimates of u_{t+1} and β_{t+1} 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 $\hat{\mu}_{rt+1}^s = E[W^s(x_{t+1}, \hat{\beta}_{rt+1}, \hat{u}_{rt+1})]$. 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+1} and β_t are the same. Further, although the distribution of u_t is updated with x_{t+1} to estimate u_{t+1}, the relationship determining the heteroskedastic 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

[7] For a detailed exposition see Elbers, Lanjouw and Lanjouw (2002, 2003) and also Stifel and Christiaensen (2007).

$W(c_{t+1})$ is obtained through simulating the process described above for different draws r, yielding $\tilde{\mu}^{s}_{t+1}$.

In pursuit of precise and consistent estimates of W_{t+1} in the absence of observations on the true c_{t+1} it is important to understand which factors affect the difference between the estimator $\tilde{\mu}^{s}_{t+1}$ of the expected value of $W(c_{t+1})$ and the actual level of the welfare measure $W(c_{t+1})$ at t+1. ELL (2002, 2003) point to three error components in the SAE approach to developing poverty maps. First, idiosyncratic error $(W(c_{t+1}) - \mu_{t+1})$, captures the fact that the welfare indicator deviates from its expected value as a result of the realizations of the unobserved component of consumption. As noted in ELL (2002, 2003), this error component declines as the size of the "target" population over which $W(c_{t+1})$ is being estimated increases. In the present application our "target" population generally comprises the total national population, or very large sub-groups of the population (such as state- or province-totals). As a result, the idiosyncratic error is unlikely to contribute meaningfully to the overall error.

Second, model error $(\mu_{t+1} - \hat{\mu}_{rt+1})$ arises from the fact that $\tilde{\mu}_{t+1}$ is a function of consistent parameter estimates that themselves have been estimated with some error. This error component does not diminish with the size of "target population" and can only be reduced via careful specification of the prediction models. Third, computational error $(\hat{\mu}_{rt+1} - \tilde{\mu}_{t+1})$ arises from the simulation-based method of computation. This error component can be set arbitrarily small by increasing the number of draws. In the present context of applying the SAE approach to imputations from one survey to another survey,

one must in addition allow for the sampling design and structure of the survey into which consumption is being predicted. Hence, an additional sampling component $(\mu_{t+1} - \overset{s}{\mu}_{t+1})$ must be added to the three error components arising from the standard SAE poverty mapping application. The size of the sampling error will depend on the design and sample size of the household surveys being analyzed.

It is important to note that alongside the error components described above, prediction error will also depend on underlying assumptions, notably that the relationship between consumption and covariates remains stable over time, and that covariates are defined identically across surveys.

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), one clear task facing the analyst is to select model specifications that keep model error to a minimum.[8] In addition, success of the procedure will depend on selection of 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). Two further 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.

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

Our focus is on two distinct, consumption and non-consumption classes of models. We experiment within these classes with alternative specifications. Another option would be to follow Tarozzi (2007) and mix both consumption and non-consumption covariates in our specifications. Possibly these would lead to models with greater explanatory power and predictive success. However, as we are interested in later sections of this paper to exploit the fact that non-consumption models dispense with a need to deflate expenditures for price variation, we have chosen in this paper to keep the classes of models separate, and to retain some parsimony in the number of models to estimate and compare.

Considering models based on consumption first, we examine the predictive power of both food and non-food sub-components of consumption expenditures. These will be further divided—where the data permit—into staple and non-staple foods, as well as frequent and infrequent non-food expenditures (typically collected using 30-day and 1-year recall respectively). 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 non-food 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 still high 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 non-food 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 more quickly into purchases of non-staple foods and durables alike (Elbers and Pouw, 2009).

Turning to the non-consumption models, 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 labor 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, color 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 consumption measures expressed in per capita terms, the less measurement error there is in the consumption measure, the lower is the transitory component, and the higher is the share of non-food expenditures (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 well-being, could help capture better the transitory component in consumption.[9] In this paper, we focus deliberately on the relatively sparse set of assets that is readily available in most conventional surveys and that has been commonly found to explain well variation in household consumption levels. Including more time variant variables from outside the questionnaire would possibly contribute to greater precision and less bias but would likely increase the data compilation efforts in practice.

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

Accordingly, poverty predictions derived from consumption models using different asset class combinations will be compared. From within each of the asset classes, care will be taken to specify models that achieve satisfactory explanatory power but that also minimize model error. A balance is sought whereby additional regressors that add to explanatory power can be added only if parameter estimates are precisely estimated. Concerns about overfitting also prompt a preference for relatively parsimonious specifications. As a basic starting point, it is generally appealing to check the performance of a particular specification on a random sub-sample of the survey data

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

set that was used to calibrate the consumption model. Only specifications that are able to predict actual observed consumption levels well should be used to impute consumption into later rounds of survey data.

The simulation based SAE technique deployed here has some attractive features that are woth noting.[10] First, the approach of estimating consumption rather than estimating final welfare indicators directly allows the estimation of any number of poverty or inequality indicators with the same simulation.[11] Second, it 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] Third, while the SAE method does impose some structure on the distribution of the idiosyncratic error term in the consumption model, the heteroskedasticity 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] Fourth, the technique is convenient to implement given the freely and readily downloadable PovMap2 software.[14]

---

[10] To be sure, the small area estimation approach is just one of a variety of valid approaches that can be employed (see for example, Azzarri et al. (2006), Tarozzi and Deaton (2009), and Matthiasen (2009)).

[11] This is in contrast to approaches, such as Deaton and Tarozzi (2009) or Tarozzi (2007), which require the estimation of different first stage models for different welfare indices.

[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 lnc_{t+k}. Otherwise, the variance of the distribution of $lnc_{t+k,}$ 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] http://iresearch.worldbank.org/PovMap/PovMap2/PovMap2Main.asp.

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), we assess whether the predicted poverty rates closely match the observed rates in a wide variety of settings. If so, parameter instability may not be a pressing concern, suggesting that SAE techniques and data on consumption subcomponents and/or household assets can indeed be used to 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.

In particular, to judge the success of the prediction models, we focus on whether the imputation-based poverty estimate for the second survey round lies within the 95% confidence interval around the "true" (directly-estimated) poverty rate for that year. To be sure, standard errors can also be estimated around the imputation-based poverty rates. An alternative procedure would thus be to test whether the imputation-based poverty rate in the second round is statistically distinguishable from the directly estimated poverty rate for that year. Such a test would be less conservative than the one we apply: it would reduce the likelihood of rejecting equality of the predicted and observed poverty rate. Moreover, it is important to acknowledge that the calculation of standard errors on the imputation-based poverty rates would not capture the uncertainty associated with the underlying assumption of parameter stability.

To shed preliminary light on how characteristics of the consumption models and settings affect the predictive power of the SAE prediction techniques, a meta-analysis of the prediction results from these different surveys and models is further pursued in

section 5. In particular, the SAE prediction techniques are validated and applied in four country settings—validated in Vietnam and rural China and applied in Russia and Kenya. Separate consumption models are estimated for different geographic areas within each country (national, rural, urban, province). The settings further differ in the evolution of poverty observed during the spell (decline, stagnation, increase) and sample size. Classified by the direction of poverty change and the geographic area, a total of 25 different settings can be considered. Table 1 provides a summary (including the level of the poverty headcount in the base year and the observed poverty change).

3 Testing Poverty Predictions Using Two Surveys of Comparable Design

We start with an assessment of the performance of the adapted SAE techniques in Vietnam and China, two different settings with highly comparable data and each encompassing periods of substantial structural change. We then move to the case of Russia and Kenya to demonstrate that the poverty prediction technique based on model specifications that eschew regressors requiring adjustment for cost of living variations can help to adjudicate questions about the appropriate choice of price deflators.

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% in 1992/1993 to 37% in 1997/1998. These estimates are based on the well respected and highly comparable Vietnam Living Standards Surveys (VLSS) of 1992/1993 and 1997/1998 and are widely judged as presenting an accurate picture of the evolution of living standards (Agarwal, Dollar and Glewwe, 2004). Both surveys are representative at the national and regional levels. The 1997/1998 survey contains panel

information on approximately 4300 of the original 4800 households interviewed in 1992/1993, but has a total sample size of 6002 households due to an expanded budget and sample design.

The second setting is China, where we draw on 2000-2004 panel data from 800 and 700 rural households from the two western provinces of Inner Mongolia and Gansu, respectively. Following 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 examined here document this change as background to a World Bank-supported poor area development program 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 sharply from 19 percent in 2000 to 6.2 percent in Inner Mongolia and halved from 24.3 to 11.8 percent in Gansu.

Performance of SAE poverty predictions using expenditures and asset models

The VLSS data allow testing the performance of consumption models using both expenditure sub-components and household assets (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)[15]; "one-year" nonfood expenditures with a one year recall

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

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 of which include geographic indicators. The models 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 labor 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 the board, despite a wide divergence in initial poverty incidence across provinces (from 35.3 percent in the South Eastern province to 80 percent in the northern Uplands in 1992/3). Poverty rates in the second period were predicted best by non-rice expenditure, annual non-food expenditure, total non-food 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 head count 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). The poverty headcount point estimates fell well within the confidence intervals around the directly observed poverty rates in the great majority of predictions for the ten regions considered. The picture does not change appreciably for the poverty gap as poverty measure (Table 3).

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 sub-components 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 non-food expenditures drive the performance of the non-food 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 that might be expected to experience changing returns (such as education and demographics) can be omitted at relatively low cost. As expected, the consumption subcomponent and full asset models display higher predictive power in the underlying consumption models (R-squareds on average exceeding 0.62 and 0.58 for the consumption subcomponent and full asset models respectively; R-squareds for most of the other asset models varying between 0.25 and 0.45).

The findings from rural western China, where poverty also declined substantially, are similarly encouraging and in line with the Vietnam experience. In Inner Mongolia, models based on expenditure sub-components do not do well, unlike in Vietnam, but the full asset model as well as the asset model that omits demographic, educational and agricultural asset characteristics, do well in capturing the dramatic decline in poverty

from 19 percent to roughly 6 percent in a period of just 5 years (Table 4). 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 percent – just slightly outside the confidence interval of 0.4-1.2 on observed poverty for that year (Table 5). Interestingly, in the case of the poverty gap, a model based on non-food expenditures does succeed in tracking poverty decline in Inner Mongolia between 2000 and 2004.

In Gansu, the SAE approach performs well in tracking the halving of poverty from 24 percent to 12 percent with models based on overall food expenditures and with non-staple food expenditures, as well as with a non-consumption model based on the full array of assets. The poverty gap results are also best with these three sets of models. As in Vietnam, the performance of the different models is again broadly consistent with their predictive performance in the underlying consumption models. Comparing the R-squareds of the different first stage consumption models in Inner Mongolia, they are highest for the full asset (0.41) and non-food models (0.73), but below 0.31 for the other models. A similar trend is observed in Gansu (R-squareds equal to 0.59 for the full asset model, but only 0.38 for the food staple model, which tends to have a lower R-squareds in all settings).

Overall, results for both Vietnam and rural China are quite encouraging. Against a background of deep structural change, the poverty prediction method appears to work consistently well with comprehensive models specified on the basis non-consumption assets, and in nearly as many cases with models using certain expenditure components (non-staple food spending, non-food spending). From this analysis it seems that the

underlying stability assumptions of the poverty prediction method appear to hold, particularly when models using non-consumption assets are employed.

4 Gauging Cost-of-Living Adjustments in Russia and Kenya

Following an assessment of the validity of our poverty prediction method in Vietnam and rural China, we now turn to two applications that consider the all too common situation where temporal poverty comparisons need to be made in the face of uncertainty about the appropriate price deflator. Our first example concerns Russia between 1994 and 2003, straddling 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 4380 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 analyzed. 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).

Conventional analysis of these surveys based on official price deflators documents a sharp *increase* in recorded (consumption) poverty from 11.4% in 1994 to 33.8% in 1998, after which poverty is estimated to have fallen back to 11.1% by 2003.[17] While this trend has in general been accepted in the literature on Russian poverty, recent papers have raised some doubts. Stillman and Thomas (2008) find almost no effect of the 1998

[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.

crisis in Russia on nutritional status. This could be understood if Russian households were able to fully smooth food consumption during this period of dramatic expenditure shortfalls. However, it would also be consistent with an overstatement of the increase in poverty in 1998. This latter observation could be obtained if inflation between 1994 and 1998 had somehow been overstated. Indeed, Gibson, Stillman and Trinh (2008) find evidence of a substantial overstatement in the CPI for urban Russia. To the extent that this latter finding holds for Russia more generally, and that the price deflators that accompany the RLMS data track the Russian CPI, it is possible that poverty in 1998, estimated from RLMS data, is also overstated. We apply our poverty prediction method to the RLMS data to probe these alternative narratives.

In our second examination of poverty trends in the face of uncertain price-deflators we consider the case of Kenya. Two recent 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 was 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 2005-6, while the deflator based on recalculations of the rural and urban poverty lines suggested a much lower price increase (6 percent in

[18] The 1997 WMS survey was carried out during 3 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 was collected for broad (aggregated) categories: 79 food (7 day recall) and 48 non-food items compared with the use of more detailed categories during the KIHBS: 140 food items (7 day recall) and 184 non-food items (1 month recall).

rural and 27 percent in urban areas). Puzzlingly, changes in survey-based poverty lines and the CPI had largely mirrored each other in the surveys prior to 1997. Application of the SAE methodology is attempted here to check on the poverty numbers for Kenya that are based on survey-based price deflators (World Bank, 2008).

Poverty Predictions in the Face of Uncertain Inflation Adjustments

According to the SAE prediction approach, poverty in Russia rose between 1994 and 1998, but by much less than official statistics would suggest. Table 6 shows that while official statistics indicate that the incidence of poverty rose from 11.4 percent in 1994 to 33.8 percent in 1998, the poverty prediction method based on a comprehensive model of non-consumption assets suggests that the headcount rose only from 11.4 to 14.1 percent during this five year period. Interestingly, a model specification that includes an indicator capturing households' subjective assessment of their wellbeing supports the finding of a relatively modest increase in poverty (Table 6, model 5).[19]

The broad finding of a more muted increase in poverty (relative to what official statistics imply) is also evident separately for rural and urban areas, although in absolute terms the SAE approach based on the comprehensive assets specification does suggest that poverty increased appreciably in rural areas, from 13.1 percent to 22.4 percent. When the SAE approach is applied to track poverty between 1994 and 2003, the approach indicates that poverty in 2003 was only marginally lower than in 1994, while the official statistics suggest that poverty in 2003 remained roughly the same. The qualitivate

[19] Gibson, Stillman and Trinh (2008) also find that subjective wellbeing indicated a much milder decline in living standards between 1994 and 2001 than official estimates had suggested.

conclusion of a muted increase in poverty between 1994 and 1998 is also reached with the poverty gap as welfare indicator (Table 7).

While the findings reported above offer some support to an emerging view in the literature that official assessments as to the decline in Russian living standards during the financial crisis in the late 1990s may be overstated, it is also important to acknowledge that they cannot settle the debate about the welfare impacts of the Russian financial crisis. The nature of the financial crisis in Russia, a macro-shock that the Russian populace was unlikely to have been able to anticipate, could well have resulted in major income shortfalls that seriously compromised living standards. In such circumstances, households are probably unable (and unwilling) to quickly run down their asset holdings, and so a poverty prediction model based on asset holdings would be unlikely to track directly the immediate income cuts associated with the crisis. A richer prediction model, better able to capture short term behavioral responses, would be desirable in this context. The inclusion of subjective well-being indicators represents one step in that direction, and it is of some comfort that results with this model support the asset-only model. However, further analysis and investigation along these lines is warranted.

Poverty prediction results for Kenya are presented in Tables 8 (poverty incidence) and 9 (poverty gap). A key feature of this example 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 China, 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 analyzing poverty in Kenya, and underscores the potential of asset based poverty prediction models in adjudicating such choices.

5 Preliminary Meta Analysis

The SAE poverty prediction technique has been explored here 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. Application of different combinations of poverty predictors in each of these settings 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.

In particular, the (absolute value of the) deviation between observed and estimated poverty levels in the second period divided by the observed poverty measure in that period[20], is taken as dependent variable. Using Ordinary Least Squares with error terms corrected for heteroskedasticity 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), the nature of the

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

poverty prediction model (full asset model, full asset model without education, staple and non-staple food expenditure model and non-food expenditures models, with the remaining asset models as default), the poverty measure used (headcount or poverty gap) and its initial level, and country indicators (China the omitted category). As our analysis of the Russia case suggests that the official poverty estimates for 1998 may be overstated, and given that but we do not have an alternative "correct" inflation that would allow us to compare our poverty predictions against adjusted official poverty estimates, the Russian estimates are not included in our meta analysis.[21] The resulting sample has 204 observations—102 observations for the headcounts and 102 for the poverty gaps.

At first sight, the R-squared of the underlying consumption model and the choice of poverty predictors appear not to affect the poverty predicting power of the SAE approach (Table 10, col 1). Yet, as seen before, both are likely highly correlated, introducing multicollinearity and less precise estimates. When the indicators of the type of poverty prediction models used are excluded (Table 10, col 2), the explanatory power of the underlying consumption model emerges as a a powerful predictor of the success in tracking poverty via the SAE poverty prediction approach—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 13 percent. This finding is not unexpected. It 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.

[21] Nonetheless, inclusion of the findings for Russia does not change the core insights emerging from the meta-analysis.

However, not all models are equally adept in predicting consumption, with the non-staple food expenditure model tending to yield the highest R-squared for the first stage consumption regressions, closely followed by the non-food and full asset models (Table 10, column 3), yielding R-squareds that were 0.36, 0.29 and 0.22 higher than those of the other asset models. The R-squareds from the staple food expenditure and full asset models without education was not different from the other more parsimonious asset models. The consumption models also tended to work slightly better in urban areas compared with the national and provincial consumption models.

Returning to the correlates of the poverty prediction error (Table 2, col 2) in the country studies examined here, the SAE approach worked better in urban settings and the prediction error tended to be smaller when the underlying sample was larger, when predicting poverty headcounts, and in settings with higher initial poverty levels. The two latter findings, however, partly reflect the mechanics of the arithmetic, with similar absolute changes from higher levels yielding smaller relative (percent) changes. Given a further expansion of the database underpinning this meta-analysis, more nuances in our understanding of where SAE methods can best be applied are likely to emerge.

6 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 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 rarely been tested. Until the performance of these models in predicting changes in poverty is scrutinized with actual data, one must remain cautious with the application of such techniques in practice.

This paper provides a first step at filling this void, drawing on data from two surveys with highly comparable expenditure data, further complemented with two case study applications, 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.

Our poverty predictions were found to be broadly successful in two country settings that offer an opportunity to compare predicted poverty against observed poverty. In Vietnam, a variety of models, comprising both consumption sub-components (non-rice food, non-food) 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 went through between 1992/3 and 1997/8 — transformations that would lead one to expect that parameter stability would fail to hold. In the rural Chinese provinces of Inner Mongolia and Gansu, the prediction

method worked similarly well, particularly with the full asset model. Different poverty measures did not affect performance of the prediction models.

We next applied the poverty prediction methodology to two case studies, Russia and Kenya, where there are grounds for questioning official cost of living adjustments that underlie temporal poverty comparisons. Our SAE approach of predicting poverty on the basis of an asset model dispenses with the need to introduce inflation adjustments. Applying this method to the Russian data for 1994, 1998 and 2003 indicates that, indeed, there may be some basis to recent suggestions in the literature that the official inflation indices overstate the rise in prices between 1994 and 1998, the height of the financial crisis. In Kenya, official inflation indices have been found to differ markedly from those calculated directly from household survey data. The SAE approach was found to lend support to poverty trends estimated on the basis of the inflation indices derived from the survey data as opposed to the official series.

Finally, a preliminary meta analysis of our results confirms that the key determinant of success in producing reliable estimates based on these prediction methods is explanatory power in the basic consumption model. In this, it is especially the non-staple food and non-food expenditures and full asset models that perform well. While expenditure sub-component 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. Consumption sub-component models also still require appropriate price deflators, which can be hard to come by.

Together, these results combine to provide cautious optimism that poverty can be tracked in the absence of (comparable) consumption data by tracking poverty predictors. However, the results also clearly call for further validation of the parameter stability assumption in more settings, for shorter and longer time periods, and especially in settings of rapid poverty deterioration. In such endeavours, the additional explanatory power of more time variant variables, such as rainfall data in agricultural dependent settings, but also health variables and subjective poverty indicators deserves particular attention.

Works Cited

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

Alderman, Harold, et al., 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 Carlo, Gero Carletto, Benjamin Davis, and Alberto, Zezza, 2006, Monitoring Poverty Without Consumption Data: An Application using the Albania Panel Survey, *Eastern European Economics*, 44-1: 59-82.

Beegle, Kathleen, Joachim De Weerdt, Jed Friedman and John Gibson, 2010, Methods of Household Consumption Measurement through Surveys: Experimental Results from Tanzania, *Policy Research Working Paper Series, 5101*, World Bank: Washington D.C..

Deaton, Angus, 2003, Adjusted Indian Poverty Estimates for 1999-2000, *Economic and Political Weekly,* January 25: 322-326.

Deaton, Angus, 2008, Price Trends in India and Their Implications for Measuring Poverty, *Economic and Political Weekly,* September 7: 3729-3748.

Deaton, Angus, and Salman, Zaidi, 2002, Guidelines for Constructing Consumption Aggregates for Welfare Analysis, *Living Standards Measurement Study Working Paper 135,* World Bank: Washington D.C.

Deaton, Angus, and Valerie, Kozel, 2005, Data and Dogma: The Great Indian Poverty Debate, *The World Bank Research Observer,* 20-2: 177-199.

Elbers, Chris, Lanjouw, Jean.O. and Peter, Lanjouw,2002, Micro-Level Estimation of Welfare, *Policy Research Working Paper 2911*, World Bank: Washington D.C.

Elbers, Chris, Lanjouw, Jean O., and Peter, Lanjouw, 2003, Micro-Level Estimation of Poverty and Inequality, *Econometrica*, 71-1: 355-364.

Elbers, Chris, and Pouw, Nicky, 2009, Modelling Sequencing Patterns in Asset Acquisition: the Case of Smallholder Farmers in Three Rural Districts of Uganda, mimeo, Amsterdam Institute of International Development.

Filmer, Deon, and Kinnon, Scott, 2008, Assessing Asset Indices, *World Bank Policy Research Working Paper 4605*, World Bank: Washington D.C.

Gibson, John, Jikun, Huang, and Scott, Rozelle, 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.

Gibson, John, Stillman, Steven, and Trinh, Le, 2008 CPI Bias and Real Living Standards in Russia During the Transition, *Journl of Development Economics*, 87-1: 140-160.

Grimm, Michael, and Isabel, Günther, 2006, Growth and Poverty in Burkina Faso – A Reassessment of the Paradox, *Journal of African Economies*, 16-1: 70-101.

Grosse, Melanie, Stephan, Klasen, and Julius, 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*, Georg-August-Universität Göttingen, Göttingen, Germany.

Harrower, Sarah, and John Hoddinott, 2005, Consumption Smoothing in the Zone Lacustre, Mali, *Journal of African Economies*, 14-4: 489-519.

Hoogeveen, Johannes, and Remidius, Ruhinduka, 2009, *Lost in Transition? Income Poverty Reduction in Tanzania since 2001*, background paper to the Tanzanian Population and Human Development Report 2009.

Kijima, Yoko, and Peter, Lanjouw, 2003, Poverty in India during the 1990s: A Regional Perspective, *Policy Research Department Working Paper 3141*, World Bank: Washington D.C.

Lanjouw, Jean O., and Peter, 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, Erzo, 2001, Measuring Poverty Dynamics and Inequality in Transition Economies – Disentangling Real Events from Noisy Data, *World Bank Policy Research Working Paper 2549*, World Bank: Washington D.C..

Mathiassen, Astrid, 2009, A Model Based Approach for Predicting Annual Poverty Rates Without Expenditure Data, *Journal of Economic Inequality*, 7-2: 117-135.

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

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

Sen, Abhijit, and Himanshu, 2004, Poverty and Inequality in India-I, *Economic and Political Weekly*, September 8: 4247-4263.

Stifel, David, and Luc, Christiaensen, 2007, Tracking Poverty over Time in the Absence of Comparable Consumption Data, *World Bank Economic Review*, 21-2: 317-341.

Stillman, Steven, and Duncan, Thomas, 2008, Nutritional Status During an Economic Crisis: Evidence from Russia, *Economic Journal*, 118-531: 1385-1417.

Tarozzi, Alessandro, 2007, Calculating Comparable Statistics From Incomparable Surveys, With an Application to Poverty in India, *Journal of Business and Economic Statistics*, 25-3: 314-336.

Tarozzi, Alessandro, and Angus, Deaton, 2009, Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas, *Review of Economics and Statistics*, 91-4: 773-792.

Wall, Martin, and Deborah, 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: 131-147.

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

Table 1: The multitude of settings in which the SAE poverty prediction technique is tested (Vietnam/China) or applied (Russia/Kenya)

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

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: Non-food expenditures and the more complete asset models predict future 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

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

Table 3: Non-food expenditures and the more complete asset models predict future 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

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

Table 4: China 2000-2004: complete asset models perform well for poverty headcount across provinces; the food, and especially the non-staple food, models, only in Gansu.

Poverty headcount (%) *(standard error)*	Observed levels		SAE predicted poverty levels in period 2											
Included in the model	2000	2004	(1)		(2)	(3)		(4)	(5)	(6)	(7)		(8)	
Expenditure subcomponents														
Food expenditures[1)] (total)			x		-	-		-	-	-	-		-	
Food (staples)			-		x	-		-	-	-	-		-	
Food (non-staples)			-		-	x		-	-	-	-		-	
Nonfood expenditures			-		-	-		x	-	-	-		-	
Non-consumption assets														
Geographic			-		-	-		-	x	x	x		x	
Demographic			-		-	-		-	x	x	x		-	
Education/Profession			-		-	-		-	-	x	x		-	
Agricultural assets (incl. Land)			-		-	-		-	-	-	x		-	
Housing quality			-		-	-		-	-	-	x		x	
Consumer durables			-		-	-		-	-	-	x		x	
Inner Mongolia	**19.0**	**6.2**	8.6		12.6	2.7		9.8	19.3	18.8	8.3	***	7.8	***
	1.5	*0.9*												
Gansu	**24.3**	**11.8**	13.3	***	29.0	12.9	***	32.4	21.1	21.5	13.5	***	15.5	
	1.8	*1.4*												
No. of times difference NOT statistically different			*1*		*0*	*1*		*0*	*0*	*0*	*2*		*1*	
average absolute difference			*2.0*		*11.8*	*2.3*		*12.1*	*11.2*	*11.2*	*1.9*		*2.7*	
# observed poverty ≥ predicted poverty			*0*		*0*	*1*		*0*	*0*	*0*	*0*		*0*	
# observed poverty < predicted poverty			*2*		*2*	*1*		*2*	*2*	*2*	*2*		*2*	

*** denotes that the predicted poverty point estimates are not statistically different at the 5 percent level from the observed poverty estimates, i.e. they fall within the 95 percent confidence interval around the observed poverty rates. [1)]Staples include grains, beans and potatoes; Non-staple foods include meat, fruit and vegetables and other foods

Table 5: China: 2000-2004: complete asset model performs well for poverty gap across provinces; the food, and especially the non-staple food, models, only in Gansu; the non-food model only in Inner Mongolia.

Poverty gap *(standard error)*	Observed levels		SAE predicted poverty levels in period 2											
Included in the model	2000	2004	(1)		(2)	(3)		(4)		(5)	(6)	(7)		(8)
Expenditure subcomponents														
Food expenditures [1)](total)			x		-	-		-		-	-	-		-
Food (staples)			-		x	-		-		-	-	-		-
Food (non-staples)			-		-	x		-		-	-	-		-
Nonfood expenditures			-		-	-		x		-	-	-		-
Non-consumption assets														
Geographic			-		-	-		-		x	x	x		x
Demographic			-		-	-		-		x	x	x		-
Education/Profession			-		-	-		-		-	x	x		-
Agricultural assets (incl. land)			-		-	-		-		-	-	x		-
Housing quality			-		-	-		-		-	-	x		x
Consumer durables			-		-	-		-		-	-	x		x
Inner Mongolia	**3.9**	**0.8**	1.5		1.6	0.3		1.2	***	4.3	4.0	1.5		1.5
	0.4	*0.2*												
Gansu	**4.9**	**1.8**	1.6	***	4.8	1.8	***	6.4		4.0	4.1	2.1	***	2.6
	0.5	*0.3*												
No. of times difference NOT statistically different			*1*		*0*	*1*		*1*		*0*	*0*	*1*		*0*
average absolute difference			*0.45*		*1.9*	*0.25*		*2.5*		*2.9*	*2.8*	*0.5*		*0.8*
# observed poverty ≥ predicted poverty			*1*		*0*	*1*		*0*		*0*	*0*	*0*		*0*
# observed poverty < predicted poverty			*1*		*2*	*1*		*2*		*2*	*2*	*2*		*2*

*** denotes that the predicted poverty point estimates are not statistically different at the 5 percent level from the observed poverty estimates, i.e. they fall within the 95 percent confidence interval around the observed poverty rates. [1)]Staples include grains, beans and potatoes; Non-staple foods include meat, fruit and vegetables and other foods

Table 6: Official estimates of the rise in the Russian poverty headcount during the 1998 crisis may be overstated

Poverty headcount (%)	Observed levels						
(standard error)							
Included in the model	Period 1	Period 2	(1)	(2)	(3)	(4)	(5)
Non-consumption assets							
Geographic			x	x	x	x	x
Demographic			x	x	x	-	x
Education/Profession			-	x	x	-	-
Housing quality			-	-	x	x	-
Consumer durables			-	-	x	x	-
Subjective perception of quality of life			-	-	-	-	x
Region	1994	1998					
National	**11.4**	**33.8**	11.7	11.7	14.1	12.7	13.2
	0.6	*1.1*					
Rural	**13.1**	**34.8**	14.0	15.9	22.4	18.2	16.9
	1.3	*2.0*					
Urban	**10.6**	**33.3**	15.2	17.8	18.8	17.4	11.5
	0.7	*1.3*					
	1994	*2003*					
National	**11.4**	**11.1**	9.8	8.2	8.5	8.4	9.2
	0.6	*0.6*					
Rural	**13.1**	**17.4**	11.2	11.3	9.9	13.1	12.4
	1.3	*1.5*					
Urban	**10.6**	**8.1**	12.1	11.2	9.2	11.2	7.4
	0.7	*0.6*					
No. of times difference NOT statistically different			0	0	1	0	1
average absolute difference			12.4	12.8	13.8	13.6	11.5
# observed poverty ≥ predicted poverty			5	5	5	5	6
# observed poverty < predicted poverty			1	1	1	1	0

*** denotes that the predicted poverty point estimates are not statistically different at the 5 percent level from the observed poverty estimates, i.e. they fall within the 95 percent confidence interval around the observed poverty rates. [1)] Staple foods: bread and potatoes; non-staple foods: meat, fruit, fat, eggs and "eat-out"

Table 7: Official estimates of the rise in the Russian poverty gap during the 1998 crisis may be overstated

Poverty gap *(standard error)*	Observed levels						
Included in the model	Period 1	Period 2	(1)	(2)	(3)	(4)	(5)
Non-consumption assets							
Geographic			x	x	x	x	x
Demographic			x	x	x	-	x
Education/Profession			-	x	x	-	-
Housing quality			-	-	x	x	-
Consumer durables			-	-	x	x	-
Subjective perception of quality of life			-	-	-	-	x
Region	1994	1998					
National	**3.8**	**12.9**	3.3	3.3	4.1	4.1	4.1
	0.2	*0.5*					
Rural	**4.1**	**13.2**	3.9	4.9	7.8	6.2	5.5
	0.5	*0.9*					
Urban	**3.7**	**12.7**	4.3	5.5	6	5.5	3.3
	0.3	*0.6*					
	1994	2003					
National	**3.8**	**3.6**	2.8	2.1	2.3	2.6	2.8
	0.2	*0.2*					
Rural	**4.1**	**6.0**	3.1	3.2	2.8	4.3	3.9
	0.5	*0.6*					
Urban	**3.7**	**2.4**	3.4	3.2	2.6	3.3	2.3
	0.3	*0.2*					
No. of times difference NOT statistically different			0	0	1	0	1
average absolute difference			3.5	3.8	4.3	4.4	3.6
# observed poverty ≥ predicted poverty			5	5	6	5	5
# observed poverty < predicted poverty			1	1	01	1	1

*** denotes that the predicted poverty point estimates are not statistically different at the 5 percent level from the observed poverty estimates, i.e. they fall within the 95 percent confidence interval around the observed poverty rates. [1)] Staple foods: bread and potatoes; non-staple foods: meat, fruit, fat, eggs and "eat-out"

Table 8: Full asset model predicts future headcount 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	

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

Table 9: Full asset model predicts future 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	

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

Table 10: The better is the predictive power of the underlying consumption model, the more accurate is the poverty prediction.

OLS with robust s.e. at country level	Relative deviation[1)] from observed poverty in period 2		Relative deviation[1)] from observed poverty in period 2, w/o asset model		R-squared of first stage consumption model	
	Coef.	p-value	Coef.	p-value	Coef.	p-value
	(1)		(2)		(3)	
Characteristics prediction model						
R-squared of consumption model	-71.50	0.30	-139.79	0.05	-	-
# of observations	-0.01	0.00	-0.01	0.00	0.00	0.19
Spell characteristics						
≥ 4% decrease in poverty incidence	9.06	0.51	22.46	0.11	-	-
Poverty prediction model						
Full asset model	-55.17	0.21			0.22	0.03
Full asset model w/o education	-61.75	0.15			0.06	0.13
Staple food expenditure model	20.36	0.30			0.07	0.13
Non-staple food expenditure model	-53.79	0.15			0.36	0.02
Non-food expenditure model	-52.88	0.27			0.29	0.00
Geographic area						
Rural sample	8.41	0.02	6.14	0.31	-0.09	0.12
Urban	-54.97	0.00	-54.04	0.00	0.03	0.00
Poverty measure						
Headcount (1=yes; 0=povgap)	-35.29	0.04	-35.49	0.05	-	-
Headcount level	-2.10	0.00	-2.13	0.00	-	-
Poverty Gap level	-7.34	0.00	-7.41	0.00	-	-
Country dummy (China omitted)						
Vietnam	68.06	0.01	69.06	0.01	0.03	0.33
Kenya	44.02	0.06	53.43	0.03	0.06	0.03
Constant	207.03	0.01	195.85	0.00	0.33	0.02
# observations	204		204		204	
Adj-R^2	0.43		0.35		0.73	

1) Absolute value of the difference between predicted and observed poverty in period 2, divided by observed poverty in period 2