The Reliability of Small Area Estimation Prediction Methods to Track Poverty

1 September, 2008

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

Abstract:

In response to the challenge of obtaining repeated and comparable consumption measures over time necessary to track poverty, poverty practitioners have developed a series of poverty prediction methods. Using small area estimation techniques (SAE) this study analyzes the empirical validity of the assumptions underlying these methods by comparing the predicted poverty measures with the poverty rates observed in repeated national household surveys from Vietnam, Russia and Kenya. The assumption that the Engel curve is stable over time when subcomponents of the household consumption measure are used to predict poverty is contradicted by the results which point to substitution between the included and excluded sub-components of expenditures. However, the assumption of stable returns to assets, when household assets are used to predict poverty, is not rejected. These findings provide cautious optimism for using SAE techniques applied to household assets to approximate the evolution of poverty. In case of a major economy wide "shock", as in Russia during the period under study, the methodology is however less appropriate.

[1] Paper submitted for presentation at the WIDER Conference on Frontiers of Poverty Analysis in Helsinki, 26-27 September, 2008. Luc Christiaensen and Peter Lanjouw are at the World Bank. Jill Luoto is at the University of California Berkeley and David Stifel is at Lafayette College, Pennsylvania.

1. Introduction

Monitoring progress towards achieving the Millennium Development Goal of halving poverty by 2015 requires comparable poverty estimates over time. Poverty measures are typically derived from survey based household expenditure estimates. Constructing such estimates is challenging in practice (Deaton and Zaidi, 2002) and these challenges only increase when comparing expenditures and poverty estimates over time. Not only are consumption measures often not available at regular intervals, but even when they are available, they are often not comparable. Even slight differences in questionnaire or survey design can lead to quite different poverty estimates (Lanjouw and Lanjouw, 2001). In addition, consumption price deflators are often calculated using urban consumption baskets only, which can be quite different from rural consumption patterns, rendering them inappropriate for tracking national poverty.

This has prompted poverty economists to develop a series of different poverty prediction methods (Ravallion, 1996; Sahn and Stifel, 2000; Kijima and Lanjouw, 2003; Azzari et al., 2006; Luoto, 2006; Stifel and Christiaensen, 2007). The methods differ in the data sources and prediction techniques used, as well as their underlying assumptions, none of which are uncontroversial. The recent "Great Poverty Debate" in India (Deaton and Kozel, 2005) vividly illustrates the challenges involved in developing reliable poverty prediction methods .. The official poverty numbers for India showed a decline in poverty from 36 percent in Round 50 of the National Sample Survey (NSS) (1993/1994) to 26 percent in Round 55 (1999/2000). However, these numbers were largely considered not to be comparable since the reporting periods were changed for various

consumption items between surveys, and the accuracy of the price indices used to update the poverty lines over time and across space has also come into question.

In an attempt to restore comparability across the Indian surveys, Deaton (2001) used the fact that the section of the consumption module on "30-day" expenditures did not change between the 50th and 55th rounds to estimate the probability of a household in the 55th Round being poor as a function of its per capita 30-day expenditures and the relationship observed between 30-day per capita expenditures and total per capita expenditures from the 50th Round (with all expenditure variables in log form).[2] The reliability of these poverty estimates depends on the validity of the assumption that there is no change in the Engel curve relating 30-day expenditures to total expenditures over time. It rules out any substitution effects following relative price shifts between the included and excluded expenditure sub-components.

Kijima and Lanjouw (2003, 2005) on the other hand use a subset of explanatory variables outside the consumption module that are strictly comparable between survey rounds to predict consumption figures. The poverty estimates based on these predictions show a less rapid decline in poverty during the 1990s than the official numbers. In this approach, the underlying relationship between consumption and its correlates is assumed to remain stable over time, thereby ignoring any potential changes in the "returns" to factors such as education and labor. This may be too stringent an assumption, especially in fast growing economies such as India during the period under study (Deaton and Kozel, 2005). So far, no consensus has emerged regarding the true poverty trend in India during the 1990s.

[2] One argument Deaton and Kozel cite in favor of Deaton's methodology is the fact that these "30-day" expenditures comprise at least 20% of household total expenditures on average and can explain around 80% of the variation in total expenditures (Deaton and Kozel 2005).

Stifel and Christiaensen (2007) build on the theoretical insights emerging from the Great Indian Poverty Debate to predict poverty over time in Kenya, which had not conducted a national household budget survey since 1997. They first estimate a model of consumption that is a function of household assets and other basic indicators derived from a Household Budget Survey, and then impute this model into a series of Demographic and Health Surveys that contain comparable asset data. To mitigate potential bias from the parameter stability assumption, they exclude household assets whose returns are more prone to change over time such as labor and education variables. They further control for potential factors that affect the returns to assets over time through inclusion of time variant variables such as rainfall and nutritional status. While triangulation with the evolution of GDP, other non-monetary indicators of welfare, and the results from other sample surveys suggests that their poverty predictions are plausible, in the absence of another comparable household budget survey, they were unable to test the empirical validity of their stability assumption.

Both Kijima and Lanjouw (2003) and Stifel and Christiaensen (2007) used an adapted version of the small area estimation (SAE) methodology developed by Elbers, Lanjouw and Lanjouw (2003) to impute a definition of consumption from one household survey into the other. This technique is superior to the more standard consumption prediction techniques in the literature (see for example Azzarri et al., 2006), as it provides consistent estimates of both the mean *and* the variance of consumption, and thus also a consistent estimate of the change in poverty over time[3].

The lack of empirical verification of the validity of the underlying assumptions in these different poverty prediction methodologies has been an important void in the

[3] See Stifel and Christiaensen (2007) for a more elaborate discussion of this point.

poverty literature and an important obstacle in furthering the use of alternative poverty prediction techniques to track poverty over time. Such empirical verification has not been easy in practice as it requires, at a minimum, settings in which comparable consumption data *are not* missing, so that the exercise can be performed *as though* they were, after which the predicted results can be checked against the "truth". Using repeated cross sections and/or panel data of nationally representative household living standard measurement type surveys that are comparable over time in terms of questionnaire and survey design from three widely divergent settings, this paper makes a first contribution to filling this void by comparing the poverty predictions obtained through SAE techniques applied to consumption subcomponents and non-consumption assets with the poverty measures obtained directly from the data.

First, the study empirically explores the validity of the poverty prediction methodology in Vietnam using the Living Standards Surveys (VLSS) of 1992/1993 and 1997/1998. They span a period of rapid poverty reduction following the Doi Moi reforms introduced in1987. This provides a context in which the stability assumptions that underpin the application of the methodology are put to a demanding test. Second, we test the methodology using the 1994, 1998, and 2003 Russian Longitudinal Monitoring Surveys (RLMS) which encompass the 1998 financial crisis and its subsequent period of recovery. Finally, using the 1997 and 2005 household expenditure surveys, the approach is also tested in Kenya, a low income setting, characterized by economic stagnation during much of the 1990s and a couple of years of more robust growth more recently.

The results from Vietnam and Russia suggest a substitution between the included and excluded sub-components of expenditures, contradicting the assumption of a stable

Engel curve, and the use of consumption subcomponents as information base to predict poverty over time. However, the assumption of stable returns to assets is rejected in neither Vietnam nor Kenya. These findings provide cautious optimism for using SAE techniques applied to household assets to approximate the evolution of poverty. In case of a major economy wide "shock", as in Russia during the period under study, the methodology is likely less appropriate

The paper proceeds as follows. Section 2 briefly reviews the SAE methodology and discusses the theoretical considerations in chosing the consumption predictors (i.e. the consumption subcomponents and the non-consumption asset variables).. Section 3 describes the data, presents the poverty prediction results and discusses their performance in predicting poverty for Vietnam, Russia and Kenya respectively. Section 4 concludes.

2. Methodological considerations

The Adapted Small Area Estimation Technique

The core intuition behind the 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 assets) as well as the parameter estimates (including those concerning the distribution of the error term) derived from a model of consumption estimated from round one. By restricting the explanatory variables to those that are comparable across surveys, the method ensures an identical definition of consumption (welfare) across surveys,

circumventing the need for price deflators, but assuming that the relationship between consumption and its correlates remains stable over time.

More formally[4], let H represent the poverty headcount, based on the distribution of household-level per capita consumption, y_h. Using data at t, model the log of consumption y_{ht} for household *h* at t as:

(1) $$\ln y_{ht} = \mathbf{x}_{ht}\beta + \mu_{ht},$$

where $\mathbf{x}_{ht}\beta$ is a vector of k parameters and μ_{ht} is a disturbance term that satisfies $E[\mu_{ht} | x_{ht}] = 0$. The vector of consistent estimators $\hat{\beta}$ from equation (1) obtained using the survey data at t is then used to predict consumption levels at t+1, generating a distribution of predicted values for $\hat{y}_{ht+1}$.

The conditional distribution of the national and subnational poverty headcounts, H, at t+1 are obtained based on the generated distribution of predicted values for $\hat{y}_{ht+1}$. A separate consumption model (1) is estimated for each subnational level (r). In particular, because the household-level disturbances at t+1 are unknown, the expected value of H is estimated using x_{ht+1} and the model of consumption in (1) as:

(2) $$\lambda_r^s = E[H \mid X_r^s, \xi_r],$$

where ξ_r is the vector of model parameters, including those that describe the distribution of the disturbances, and the superscript 's' indicates that the expectation is conditional on the sample of households at t+1 from region *r* rather than a census of households (Kijima and Lanjouw 2003).

[4] For a more detailed discussion of the application of the SAE technique to predict poverty over time, see Kijima and Lanjouw (2003) and Stifel and Christiaensen (2007).

We replace the unknown vector ξ_r with the consistent estimators $\hat{\xi}_r$ estimated from the survey data at t to construct the estimator for λ_r^s, $\hat{\lambda}_r^s$. One hundred simulated draws are performed to derive the estimator $\hat{\lambda}_r^s$ in each model. See Kijima and Lanjouw (2003) for more detail on the prediction error associated with the estimator $\hat{\lambda}_r^s$ of the expected value of H for a given region.

To test empirical validity of the parameter stability assumptions, the adaptation of the SAE statistical methodology is applied to the 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. The predicted poverty results derived from each "class" of models are then compared with the “true” poverty numbers observed in the survey.

If the predicted poverty rates closely match the observed ones in many different settings, then this serves as evidence in support of the parameter stability assumption and the use of SAE techniques and data on the consumption subcomponents and/or household assets to estimate the evolution of poverty within a country when comparable consumption data are absent. If, on the other hand, our results cannot capture the observed changes in poverty, it provides evidence that the use of such techniques may be ill-advised and any results from such an exercise should be viewed with caution.

To assess the performance of the SAE technique applied to the different subsets of poverty predictors, the poverty headcount and the poverty gap will be used as poverty measures. Each time, the sign, (statistical) significance and size of the difference between the predicted and the observed poverty will be assessed. The first two moments of the distribution of the deviation between the predicted and observed poverty rates

across the different subnational levels/regions will furthermore be examined to assess consistency in the performance of the different consumption models.

Consumption Models Using Expenditure Sub-Components

The assumption of a stable Engel relationship between the sub-components of consumption expenditures and total expenditures is tested in a variety of ways by disaggregating the consumption aggregate into different subgroups of expenditures and relating each subgroup to total consumption. At a minimum we distinguish between food and non-food expenditures, and their combination, with further disaggregations also examined depending on the data set. Each of the subgroups may further exclude certain items that are asked about in differing manners between surveys and therefore arguably may capture different information.[5]

In effect, this exercise serves as an indirect test of the assumption of a stable Engel relationship. This is because the resulting poverty estimates can be compared across the different models to check for evidence of substitution patterns between the various sub-components of expenditures. To see this, consider the following example. If the estimated poverty headcount derived from a model relating just nonfood expenditures to total expenditures is 20%, while it is 30% when a model includes both food and nonfood expenditures and 40% when a model includes only food expenditures, this suggests the presence of substitution patterns between food and nonfood expenditures.

[5] More information on the included and excluded expenditure sub-components in the empirical applications are available from the author upon request.

To be more precise, if predicted poverty in the nonfood model ($\hat{H}_{nf}$) is 20%, while it is 30% in the model using both food and nonfood expenditures ($\hat{H}_{f+nf}$), and 40% in the food model ($\hat{H}_{f}$), we can write the following:

$$\hat{H}_{nf} < \hat{H}_{f+nf} < \hat{H}_{f} \Rightarrow \hat{Y}_{nf} > \hat{Y}_{f+nf} > \hat{Y}_{f}$$

where $\hat{Y}_{xx}$ is the corresponding estimate at any given percentile in the distribution of total consumption based on a model using only nonfood expenditures (*nf*), food and nonfood expenditures combined (*f+nf*), or only food expenditures (*f*), respectively.[6] But this implies the following,

$$\hat{Y}_{nf} > \hat{Y}_{f+nf} > \hat{Y}_{f} \Rightarrow \hat{\beta}_{nf} X_{nf} + \hat{\varepsilon}_{nf} > \hat{\beta}_{f+nf}(X_{f} + X_{nf}) + \hat{\varepsilon}_{f+nf} > \hat{\beta}_{f} X_{f} + \hat{\varepsilon}_{f}$$

Now, because in expectation all disturbance terms are mean zero, we can write the first statement below, while the second two are given by definition[7]:

$$E[\hat{\beta}_{nf} X_{nf}] > E[\hat{\beta}_{f+nf}(X_{f} + X_{nf})] > E[\hat{\beta}_{f} X_{f}]$$

$$X_{nf} < X_{f} + X_{nf}$$

$$X_{f} < X_{f} + X_{nf}$$

Together, these inequalities imply that we must have:

$$E[\hat{\beta}_{nf}(X_{f} + X_{nf})] > E[\hat{\beta}_{nf} X_{nf}] > E[\hat{\beta}_{f+nf}(X_{f} + X_{nf})] > E[\hat{\beta}_{f} X_{f}],$$

which implies that

$$E[\hat{\beta}_{nf}] > E[\hat{\beta}_{f+nf}], E[\hat{\beta}_{f+nf}] <> E[\hat{\beta}_{f}].$$

[6] This assumes the distributions of predicted expenditures are the same across specifications.
[7] This assumes each category of consumption is strictly positive.

But this necessarily implies some degree of substitution between food and nonfood expenditures in this example, which undercuts the assumption of a stable Engel curve relating these items. In the results presented below, a similar logic is used to comment on the evidence of a stable Engel relationship between the included and excluded sub-components of consumption across a variety of models, in addition to the more direct comparisons based on sign, significance and size of the difference between predicted and observed poverty change.

"Non-Expenditure" Models Using Household Assets

The focus is on five broad classes of household asset related information that are typically readily available within the questionnaires and commonly useful in explaining variation in household consumption levels: 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 household durables such as a bicycle, color television, electric fan, etc. (proxying a household's physical assets). [8]

[8] The inclusion of time variant variables such as rainfall, relative price changes, nutritional and health status that affect the returns to assets as pursued by Stifel and Christiaensen (2007) could further improve the consumption models and poverty predictions by reducing the likelihood of violating the parameter stability assumption. Here, the study deliberately focuses on a sparser core set of assets that is usually readily available in the questionnaires. Including time variant variables from outside the questionnaire often substantially increases the data compilation efforts in practice, and it may not be necessary.

Variables from these five general classes that are strictly comparable across surveys are first identified. A subset of variables from these different classes is then selected using stepwise and other procedures so as to maximize the explanatory power of the consumption model and minimize model error (Stifel and Christiaensen, 2007). A priori, assets such as labor and education are likely more prone to parameter instability following economic or policy change. To test the poverty predicting performance and the parameter stability assumption for different asset classes, poverty predictions derived from consumption models using different asset class combinations will be compared.

3. The Performance of SAE Poverty Predictions Using Consumption Sub-Components and Household Assets Across Different Settings

Vietnam

Vietnam is a country that experienced widespread economic growth throughout the 1990s and this growth was accompanied by a large reduction in poverty. More importantly for our purposes, the Vietnam Living Standards Surveys (VLSS) of 1992/1993 and 1997/1998 are very well respected data sources that cover this period and document this reduction in poverty. The surveys estimate a drop in national poverty from 58% in 1992/1993 to 37% in 1997/1998, and these figures are widely regarded as reflecting the true course of events in Vietnam during this time.[9] Table 1 presents the official poverty estimates for Vietnam from both surveys, including at the regional level.

[9] When the 1992 consumption data and poverty line are adjusted forward to 1998 prices to be comparable with VLSS 1997/1998, the poverty headcount in 1992/1993 increases slightly to 61%. Table 1 presents the poverty numbers for 1992-93 cited in 1998 prices.

The VLSS 1992/1993 covers 4800 households and is representative at the national and regional levels. It was conducted between September 1992 and October 1993 and contains data on the schooling, health, employment, migration, housing, and fertility of household members, as well as household consumption and ownership of a variety of household durables. The second VLSS round was collected between December 1997 and December 1998, and contains vastly similar information as the earlier survey, with many components of the two surveys' questionnaires identical. 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. It is representative at the national and regional levels.

The VLSS data allow testing the performance of consumption models using both expenditure sub-components and household assets to predict poverty. Table 1 presents the results when sub-components of expenditures are used. The following subcategories of total expenditures were considered to estimate the consumption models: expenditures on food excluding rice; expenditures on all food; "30-day" nonfood expenditures, which include those expenditures that are asked with a 30-day recall period[10]; "one-year" nonfood expenditures with a one year recall period; total nonfood expenditures (which is simply the sum of the previous two subgroups); and total expenditures (a combination of all previous subgroups of expenditures).

[10] This is the specification most similar to that used by Deaton in his analysis of India's poverty numbers across NSS rounds (see Deaton 2003??).

Table 1
Test of Assumption of Stable Parameter Estimates: Expenditure Models, Vietnam 1992-93 to 1997-98

Included in Models:		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Food Expenditures: Nonrice			X	X				X	
Food Expenditures: Rice				X				X	
Nonfood Expenditures: 30-day goods					X		X	X	
Nonfood Expenditures: Annual goods						X	X	X	
Region		1992-93 VLSS							1997-98 VLSS
National	**P0**	60.6	39.3	47.2	41.9	35.8	33.3	42.6	37.4
		(1.9)	(1.2)**	(1.1)	(1.5)**	(1.5)**	(1.2)**	(0.7)	(1.6)
	P1	19.0	9.9	12.4	11.6	8.6	8.3	11.0	9.5
		(0.9)	(0.4)**	(0.5)	(0.6)**	(0.5)**	(0.4)**	(0.3)**	(0.7)
Rural	**P0**	68.5	46.6	59.7	51.3	47.3	41.0	53.0	44.9
		(1.7)	(1.4)**	(1.1)	(1.6)**	(2.0)**	(1.6)**	(0.8)	(2.0)
	P1	22.0	11.7	16.4	14.8	11.8	9.8	13.8	11.6
		(1.0)	(0.5)**	(0.5)	(0.7)	(0.7)**	(0.5)**	(0.3)**	(0.9)
Urban	**P0**	28.6	11.5	14.5	13.0	8.6	9.2	10.1	9.0
		(4.1)	(1.1)**	(1.1)	(1.7)**	(1.1)**	(1.0)**	(0.6)**	(1.5)
	P1	7.3	2.3	3.1	3.1	1.8	2.1	2.1	1.7
		(1.2)	(0.3)**	(0.3)	(0.5)**	(0.3)**	(0.3)**	(0.2)**	(0.3)
Northern Uplands	**P0**	80.0	67.6	72.7	65.3	58.4	53.3	65.3	58.6
		(3.8)	(3.3)**	(2.2)**	(4.1)**	(4.2)**	(3.5)**	(1.6)**	(5.6)
	P1	26.7	21.3	23.0	20.0	16.1	14.7	19.5	16.8
		(2.6)	(1.8)**	(1.3)**	(2.1)**	(1.8)**	(1.5)**	(0.8)**	(2.3)
Red River Delta	**P0**	64.0	29.3	45.4	41.1	34.6	25.7	36.6	28.7
		(4.6)	(2.1)**	(2.6)	(3.2)**	(3.5)**	(2.5)**	(1.5)**	(3.4)
	P1	18.9	6.1	11.3	10.5	7.3	5.2	8.1	5.7
		(1.9)	(0.6)**	(0.9)	(1.2)	(1.0)**	(0.6)**	(0.4)**	(1.0)
North Central	**P0**	76.6	55.1	63.5	67.4	56.2	51.3	56.7	48.1
		(4.1)	(2.6)**	(2.3)	(3.0)	(4.1)**	(2.8)**	(1.6)**	(5.2)
	P1	25.3	14.1	17.5	20.3	14.3	12.2	15.2	11.8
		(2.7)	(0.9)**	(0.9)**	(1.5)	(1.7)**	(0.9)**	(0.6)**	(1.9)
Central Coast	**P0**	53.2	34.4	47.3	39.9	32.5	35.7	40.5	35.2
		(6.0)	(2.9)**	(3.7)**	(4.0)**	(3.0)**	(2.9)**	(1.9)**	(5.5)
	P1	17.7	9.0	13.6	12.4	8.1	9.7	11.1	10.6
		(3.2)	(1.0)**	(1.4)**	(1.6)**	(0.9)**	(0.9)**	(0.6)**	(3.1)
Central Highlands	**P0**	72.9	54.5	64.3	64.3	45.7	47.4	58.1	52.4
		(13.9)	(5.3)**	(5.2)**	(10.2)**	(6.0)**	(7.3)**	(3.9)**	(9.7)
	P1	27.5	15.3	19.3	22.2	14.6	16.3	18.4	19.1
		(8.5)	(1.9)**	(2.0)**	(5.3)**	(2.2)**	(2.6)**	(1.5)**	(5.9)
South East	**P0**	35.3	11.3	14.9	12.8	10.2	8.4	9.4	7.6
		(6.2)	(2.7)**	(2.4)**	(3.9)**	(2.7)**	(1.0)**	(2.4)**	(1.5)
	P1	9.8	2.3	3.0	5.8	2.1	1.8	2.0	1.3
		(2.0)	(0.7)**	(0.6)**	(2.2)**	(0.6)**	(0.3)**	(0.8)**	(0.3)
Mekong River	**P0**	51.0	38.6	47.2	34.6	33.8	32.9	40.7	36.9
		(6.2)	(1.9)**	(2.0)**	(3.1)**	(2.9)**	(2.5)**	(1.5)**	(3.0)
	P1	15.0	9.3	11.8	9.5	8.9	8.2	9.5	8.1
		(1.5)	(0.7)**	(0.8)	(1.2)**	(1.0)**	(0.8)**	(0.4)**	(0.9)

Notes: Column (1) contains official poverty numbers from VLSS 1992-93. Column (8) contains official poverty numbers from VLSS 1997-98. Standard errors in parentheses. Columns (2) through (7) contain poverty estimates using differing categories of explanatory variables as explained in text. The X's at the top of a column indicate if a category of explanatory variables is included in the models whose results are presented. (**): Signifies that poverty estimate is not statistically different at 5% levels from official poverty statistic for 1997-98 as listed in column (8).

Included in the model		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Food Expenditures: Nonrice			x	X				X	
Food Expenditures: Rice				X				X	
Nonfood Expenditures: 30-day goods					x		X	X	
Nonfood Expenditures: Annual goods						x	x	X	
Validation criteria									
% of times predicted poverty NOT statistically different from official poverty statistic (at 5% level)	P_0		100	50	90	100	100	80	
	P_1		100	50	70	100	100	100	
Average (absolute) difference between predicted and official poverty statistic across regions	P_0		3.1	11.8	7.7	3.4	3.0	5.4	
	P_1		1.6	3.5	3.4	1.5	1.1	1.6	
Standard deviation of (absolute) difference between predicted and official poverty statistic across regions	P_0		2.8	3.6	5.2	2.7	1.9	2.6	
	P_1		1.5	2.0	2.1	1.4	0.9	1.0	

Across the columns of this table, there is evidence of substitution patterns between the various sub-components of expenditures. Generally, the models that are restricted to include only food expenditures in columns 2 and 3 under-predict welfare (and therefore overestimate poverty), indicating their limitation in capturing the large declines in poverty between surveys. This is consistent with the notion that expenditures on food constitute the bulk of any "basic needs" of a household's welfare, so that total expenditures on food might be expected to be less sensitive to changes in overall welfare. Meanwhile, the models that include only the matching sub-components of non-food expenditures (columns 4, 5 and 6) largely capture the considerable changes in welfare across surveys. At the national level, each of the models that include matching "30-day" nonfood expenditures (column 4), "one-year" nonfood expenditures (column 5), and all nonfood expenditures combined (column 6), has a resulting poverty estimate that is not

statistically different from the "true" estimate for 1997-1998 (listed in column 8). This would be encouraging news in isolation.

However, column 7 combines all matching food and nonfood expenditures to predict welfare, and its resulting poverty estimates are consistently higher than those based on the models using nonfood expenditures alone (columns 4, 5 and 6), and consistently lower than those based on the models using only expenditures on food (columns 2 and 3). Such a scenario matches our example outlined earlier in section 3, and because this implies that $\hat{\beta}_{nf} >> \hat{\beta}_{f+nf} > \hat{\beta}_{f}$, it suggests evidence of substitution taking place between food and nonfood expenditures between surveys. This, in turn, violates the assumption of a stable Engel curve over time, and therefore could draw into question the viability of a methodology that relies on this assumption.

A list of the non-consumption related asset variables comparably defined between surveys for Vietnam and their associated summary statistics is contained in Table 2. The results indicate that between the two surveys many Vietnamese households improved their housing structures and acquired durables such as electric fans. This is evidence of an improvement in welfare that we expect will contribute to lower poverty predictions.[11]

[11] Since the cross-price elasticity of food with household durables and housing materials is expected to be small, it is unlikely that many households acquire these items at the expense of food and other necessary consumption and therefore our estimates will not overestimate poverty reduction.

Table 2
Vietnam: Summary Statistics of Comparably Defined "Non-Consumption" Variables from VLSS 1992-93 and VLSS 1997-98

	1992-93	1997-98		1992-93	1997-98
Geographic Variables:	**Mean**	**Mean**	***Housing Quality Indicators:***	**Mean**	**Mean**
Urban household (0/1)	0.20	0.22	Permanent Wall Material (0/1)	0.54	0.65
Regional Indicators:			Permanent Flooring Material (0/1)	0.43	0.63
Northern Uplands (0/1)	0.16	0.18	Permanent Roofing Material (0/1)	0.67	0.81
Red River Delta (0/1)	0.22	0.20	Household owns second home (0/1)	0.01	0.03
North Central (0/1)	0.13	0.14	Private Drinking Tap (0/1)	0.08	0.14
Central Coast (0/1)	0.12	0.11	Permanent Drinking Tap (0/1)	0.15	0.29
Central Highlands (0/1)	0.03	0.04	Private Laundry Tap (0/1)	0.00	0.00
South East (0/1)	0.13	0.13	Permanent Laundry Tap (0/1)	0.01	0.02
Mekong River (0/1)	0.22	0.21	Garbage Collection Service (0/1)	0.09	0.17
			No Toilet in Home (0/1)	0.27	0.04
Household Demographic Variables:			Permanent Toilet in Home (0/1)	0.19	0.27
Household Size	4.96	4.69	Private Toilet in Home (0/1)	0.61	0.70
Age of household head	45.34	47.79	Electric Lighting in Home (0/1)	0.49	0.78
Number of Females	2.57	2.42	Electric Fuel Source in Home (0/1)	0.11	0.23
Spouse present in home (0/1)	0.81	0.81	Permanent Housing Structure (0/1)	0.16	0.16
Female head of household (0/1)	0.27	0.26	Private Bath in Home (0/1)	0.05	0.06
Number of Adults 15+	3.06	3.16	Semipermanent Home (0/1)	0.47	0.59
Number of Children 0-15	1.87	1.51	Private House (0/1)	0.93	0.94
Number of Children 0-5	0.60	0.36	Number of rooms in home	2.05	2.28
Number of Children 5-10	0.65	0.54			
Number of Children 10-15	0.62	0.61	***Household Durables Ownership:***		
Number of Elderly 65+	0.27	0.33	Color Television (0/1)	0.09	0.38
Dependency Ratio	0.92	0.79	Black & White Television (0/1)	0.13	0.18
Number of Female Children	0.92	0.74	Stereo/Radio or Cassette Player (0/1)	0.26	0.46
Number of Female Adults	1.65	1.68	Camera (0/1)	0.01	0.03
			Refrigertor (0/1)	0.04	0.09
Household Educational and Professional Variables:			Air Conditioner (0/1)	0.00	0.01
Household head has any education (0/1)	0.88	0.91	Washing Machine (0/1)	0.00	0.02
Household head education in levels	1.86	1.81	Electric Fan (0/1)	0.31	0.68
Average education of female adults in levels	1.64	1.63	Car (0/1)	0.00	0.00
Average education of male adults in levels	1.90	1.87	Bicycle (0/1)	0.65	0.73
Proportion of children aged 5-15 in school	0.77	0.64	Motorbike (0/1)	0.11	0.20
Household head is a professional (0/1)	0.20	0.12	Boat (0/1)	0.05	0.04
Household operates a farm (0/1)	0.70	0.62	Sewing Machine (0/1)	0.15	0.16
Household head works in agriculture (0/1)	0.64	0.56	Bed (0/1)	0.93	0.95
Household head job category	0.97	1.01	Clock (0/1)	0.33	0.70
At least one hh member works in agriculture (0/1)	0.78	0.71	# of Tools (hammer, axe, etc.)	7.98	8.07
Self Employed Non-Farm Business in HH	0.45	0.44			

Source: Author's calculations from VLSS 1992-93 and 1997-98.

Table 3 contains the resulting poverty estimates from this exercise using the non-expenditure variables from the Vietnamese data. The empirical models of consumption used to derive these estimates are contained in the Appendix.[12] As before columns (1) and (8) contain the official poverty statistics from the two VLSS rounds for 1992/1993 and 1997/1998, respectively. Column (2) contains estimated poverty numbers based on a model of consumption that utilizes only information from the first two classes of variables, geographic and household demographic information. When broken down by

[12] The Appendix only includes the estimated models of consumption that include regressors from all five classes of explanatory variables for each geographic region with Vietnam. Estimated models that exclude certain categories of variables are available from the authors upon request.

region, these models are tailor-fitted to each region and therefore differ across regions, but each derives its explanatory variables from the first two classes of variables. In column (2) the estimated poverty numbers for 1997/1998 are lower than the official numbers for 1992/1993, but they are statistically different and higher than the official numbers for 1997/1998, both at the national level and across most of the regions. It is concluded that more information is necessary to adequately capture variation in consumption levels over time. Column (3) of Table 4 allows for the inclusion of variables describing the education and employment of household members, along with the geographic and demographic sets of variables. Again, we see that such a model does not capture the large drop in poverty between surveys for Vietnam.

When we further add housing characteristics to the model in column (4), our predicted poverty numbers for 1997/1998 significantly decrease, although not enough to be statistically indistinguishable from the official numbers for 1997/1998 for every region. The final addition of the set of variables describing ownership of household assets to the models presented in column (5) results in predicted poverty rates for 1997/1998 that largely replicate the official numbers for Vietnam. In column (5), none of the imputed poverty estimates is statistically different from the official numbers for 1997/1998 at the 5% level when broken down by region or rural/urban location for either the poverty headcount (P0) or the poverty gap index (P1). This is suggestive evidence that the assumption of stable point estimates over time is plausible within this class of models that draws explanatory variables from the "non-consumption" sections of surveys. Since the magnitude of poverty reduction varied widely across regions within Vietnam

between surveys, the ability of our models to capture this variation when all classes of explanatory variables are included lends support to the methodology.[13]

The precipitous drop in predicted poverty rates in columns (4) and (5) relative to the models in columns (2) and (3) suggests that the inclusion of information on housing characteristics and household asset ownership is necessary to adequately capture changes in welfare over time using this methodology. This agrees with our intuition that an improvement in welfare should correspond to improvements in housing quality and asset ownership. More importantly, it also supports the notion that the underlying relationship between consumption and these explanatory variables remains stable over time, even in the presence of rapid economic growth and accompanying changes.

[13] It is important to note that the regional poverty estimates are each based upon a model of consumption that best fits the data from that region in 1992/1993 and therefore the estimated models differ across regions when imputed into the 1997/1998 data. We do not restrict our regional estimates to share a single specification since it is likely that geographic variation leads to a different relationship between consumption and its correlates.

Table 3

Test of Assumption of Stable Parameter Estimates: Non-Expenditure Models, Vietnam 1992-93 to 1997-98

Included in Models:		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geographic Indicators			X	X	X	X	X	X	
Household Demographics			X	X	X	X			
Education/Profession Variables				X	X	X			
Housing Quality Variables					X	X	X	X	
Asset Ownership Variables						X	X	X	
Region		1992-93 VLSS							1997-98 VLSS
National	P0	60.6	54.6	55.6	46.6	38.2**	36.7**	37.3**	37.4
		(1.9)	(1.4)	(1.2)	(1.4)	(1.3)	(1.4)	(1.4)	(1.6)
	P1	19.0	16.5	17.2	13.4	10.5**	9.4**	9.6**	9.5
		(0.9)	(0.7)	(0.6)	(0.6)	(0.5)	(0.5)	(0.6)	(0.7)
Rural	P0	68.5	63.5	64.8	56.3	48.5**	44.2**	44.0**	44.9
		(1.7)	(2.0)	(1.7)	(1.5)	(1.7)	(2.1)	(2.1)	(2.0)
	P1	22.0	19.1	19.7	16.8	13.7**	11.5**	11.3**	11.6
		(1.0)	(1.0)	(0.9)	(0.7)	(0.7)	(0.8)	(0.8)	(0.9)
Urban	P0	28.6	21.7	22.8	18.3	11.8**	9.7**	11.4**	9.0
		(4.1)	(2.6)	(2.3)	(1.8)	(1.3)	(1.5)	(1.7)	(1.5)
	P1	7.3	5.4	5.6	4.5	2.6**	1.9**	2.3**	1.7
		(1.2)	(0.9)	(0.8)	(0.6)	(0.4)	(0.4)	(0.4)	(0.3)
Northern Uplands	P0	80.0	76.0	78.4	74.5**	62.3**	57.0**	57.1**	58.6
		(3.8)	(3.4)	(2.5)	(2.9)	(3.0)	(3.6)	(3.6)	(5.6)
	P1	26.7	23.6	25.5	24.3**	18.0**	15.9**	16.2**	16.8
		(2.6)	(2.3)	(1.8)	(1.6)	(1.4)	(1.5)	(1.4)	(2.3)
Red River Delta	P0	64.0	57.2	57.4	49.6	32.5**	32.5**	39.8**	28.7
		(4.6)	(3.3)	(3.1)	(3.1)	(2.7)	(2.4)	(2.7)	(3.4)
	P1	18.9	15.4	15.7	12.3	7.3**	7.3**	8.9**	5.7
		(1.9)	(1.5)	(1.4)	(1.3)	(0.7)	(0.7)	(0.9)	(1.0)
North Central	P0	76.6	73.1	72.8	60.9**	48.1**	47.9**	49.7**	48.1
		(4.1)	(3.8)	(3.5)	(4.0)	(4.0)	(4.7)	(4.1)	(5.2)
	P1	25.3	22.7	23.0	17.0**	12.9**	12.3**	12.1**	11.8
		(2.7)	(2.1)	(2.1)	(1.7)	(1.4)	(1.5)	(1.4)	(1.9)
Central Coast	P0	53.2	47.0**	49.3**	36.0**	34.0**	31.9**	26.1**	35.2
		(6.0)	(5.1)	(4.0)	(3.8)	(2.7)	(3.0)	(3.1)	(5.5)
	P1	17.7	15.2**	15.7**	10.9**	10.7**	9.5**	7.2**	10.6
		(3.2)	(2.4)	(1.8)	(1.5)	(1.1)	(1.1)	(1.2)	(3.1)
Central Highlands	P0	72.9	66.2**	64.0**	54.2**	51.5**	49.2**	44.1**	52.4
		(13.9)	(13.0)	(6.2)	(4.5)	(3.1)	(6.0)	(7.7)	(9.7)
	P1	27.5	25.3**	22.9**	18.9**	21.8**	17.1**	12.9**	19.1
		(8.5)	(8.1)	(3.4)	(2.0)	(1.6)	(2.3)	(2.9)	(5.9)
South East	P0	35.3	27.3	28.6	21.8	12.3**	16.8	17.5	7.6
		(6.2)	(3.2)	(2.9)	(1.8)	(1.7)	(2.0)	(2.4)	(1.5)
	P1	9.8	7.4	7.5	5.6	2.9**	3.8	4.0	1.3
		(2.0)	(1.3)	(1.1)	(0.7)	(0.5)	(0.6)	(0.8)	(0.3)
Mekong River	P0	51.0	42.7**	43.6**	36.9**	34.2**	30.7**	28.4**	36.9
		(6.2)	(3.2)	(2.8)	(2.9)	(2.6)	(2.8)	(2.7)	(3.0)
	P1	15.0	11.7**	11.8**	9.7**	9.3**	7.5**	6.8**	8.1
		(1.5)	(1.3)	(1.3)	(1.0)	(1.0)	(0.9)	(0.9)	(0.9)

Notes: Column (1) contains official poverty numbers from VLSS 1992-93. Column (8) contains official poverty numbers from VLSS 1997-98. Standard errors in parentheses. Columns (2) through (7) contain poverty estimates using differing categories of explanatory variables as explained in text. The X's at the top of a column indicate if a category of explanatory variables is included in the models whose results are presented. Column (7) includes a variation on the definition of variables describing houshold asset ownership as explained in text. (**): Signifies that poverty estimate is not statistically different at 5% level from official poverty statistic for 1997-98 as listed in column (8).

Included in the model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geographic Indicators		X	X	X	X	X	X	
Household Demographics		X	X	X	X			
Education/Profession Variables			X	X	X			

Housing Quality Variables						X	X	X	X	
Asset Ownership Variables							X	X	X	
Validation criteria			1992-93 VLSS							1997-98 VLSS
	% of times predicted poverty NOT statistically different from official poverty statistic (at 5% level)	P_0		30	30	50	100	90	90	
		P_1		30	30	50	100	90	90	
	Average (absolute) difference between predicted and official poverty statistic across regions	P_0		17.1	17.9	9.6	2.4	3.0	5.3	
		P_1		6.6	6.8	3.8	1.4	1.0	1.9	
	Standard deviation of (absolute) difference between predicted and official poverty statistic across regions	P_0		6.6	6.5	6.9	1.6	2.9	4.4	
		P_1		6.6	6.5	6.9	1.6	2.9	4.4	

Given the a priori theoretical concerns about the stability of the returns to labor and education, especially in rapidly changing economic and political environments, it is useful to test how the model performs when it includes only information on housing quality and durables ownership, along with basic geographic indicators performs (column (6)). While model (5) is slightly more precise in predicting poverty as demonstrated by the first two moments of the (absolute) difference between the predicted and observed poverty rates, this is completely driven by the poor performance of model (6) in the South East region (South East). It is the only region for which model (6) estimates statistically different poverty numbers than those observed in 1997/1998. When the poverty predictions for the South East region are dropped from the moment calculations, the performance of model (5) and (6) is in essence indistinguishable[14].

To assess how sensitive the results are to the precise types of assets included in the imputed models, the ownership of durables is redefined by classifying the number of

[14] Excluding the South East region, the average (absolute) difference between the predicted and the official poverty numbers in 1997/98 are 2.2 and 2.3 for model (5) and (6) respectively for the poverty headcount and 1.3 and 0.8 for the poverty gap. The standard deviation are 1.5 and 2.0 for model (5) and model (6) for the poverty headcount, and 0.7 in both cases for the poverty gap.

medium or high-valued durables a household has in either survey, rather than imputing the point estimate associated with ownership of a color television in particular, for example. This is done in an effort to avoid "over-fitting" our models to be relevant only for Vietnam. Classifying the various assets into medium or high-value categories is done in Table 4, which also contains summary statistics for ownership of these assets from the two surveys. Column (7) of Table 3 presents results from models that include these altered definitions of asset ownership, as well as geographic and housing quality indicators. At the national level, they result in poverty predictions that are particularly close to the true numbers for 1997/1998, suggesting the precise types of assets a household owns is not important to our results. When these models are run on individual regions of the country, again with the exception of the South East, none of the predicted poverty numbers are statistically distinguishable from the official numbers for 1997/1998. However, many of the point estimates are quite a bit lower than the official numbers, suggesting *too much* poverty reduction between surveys when run at the regional level. The inclusion of the individual assets is preferred.

Table 4
Household Asset Ownership

	1992-93	1997-98
Medium Assets:	**VLSS**	**VLSS**
Overall Average Count	1.78	3.16
Color Television	0.09	0.38
Black/White Television	0.13	0.18
Stereo Equipment/ Radio or Cassette Player	0.26	0.46
Camera	0.01	0.03
Electric Fan	0.31	0.68
Wall or Table Clock	0.33	0.70
Bicycle	0.65	0.73
Big Assets:		
Overall Average Count	0.20	0.37
Refrigerator or Freezer	0.04	0.09
Air Conditioner	0.00	0.01
Washing Machine	0.00	0.02
Car	0.00	0.00
Motorbike	0.11	0.20
Boat	0.05	0.04

Finally, while it is encouraging to see that the predicted poverty numbers closely reflect the official numbers for Vietnam when appropriate models of consumption are utilized, these results could be due simply to a "convenient" placing of the poverty line. To test for this, the entire distributions of the predicted household consumption figures for 1997/1998, based on a model of consumption derived from VLSS 1992/1993, are compared with the "official" household consumption levels from the VLSS 1997/1998.[15]

Figure 1 plots the distribution of predicted consumption levels for each household in the 1997/1998 VLSS at the national level with the associated 95% confidence bounds alongside the actual consumption figures for every household and their associated 95%

[15] Although by definition the "official" consumption levels for each household are also an approximation, we refer to the consumption levels estimated using this variant of the SAE methodology to be the "predicted" consumption figures, and the final consumption aggregates as contained in the VLSS data to represent the "official", "true", or "actual" consumption figures.

confidence bounds.[16] The vertical line in the figure represents the official, national poverty line (expressed in log terms). The "true" and predicted consumption figures for 1997/1998 largely overlap at all point in their distributions. Only at the very top of the distribution are there households with actual and predicted consumption figures whose confidence bounds nowhere overlap. The imputed model would slightly overestimate poverty (i.e., underestimate the reduction in poverty between surveys) if the poverty line were drawn among the top 10% of households in the distribution of welfare levels. This constitutes a greater than two standard deviation increase over the official poverty line. The imputed welfare estimates never differ by a statistically significant margin from the official welfare levels when considering rural households alone as demonstrated in Figure 2.[17] In sum, that the imputed poverty estimates largely replicate the official numbers cannot be attributed to a "lucky" placing of the poverty line.

Evidence from Vietnam seems to support the assumption of stable parameter estimates over time in the use of SAE methodology to update poverty estimates in a country, so long as the explanatory variables are drawn from sections of the surveys other than the consumption modules. The findings further suggest that there may not a lot to be gained from including labor and education variables in addition to geographic indicators and private asset variables such as housing quality and durables. This is encouraging given possible lingering doubts ex ante about the stability of the returns to these assets in rapidly changing environments. The results using subcomponents of

[16] The distribution of predicted consumption figures drawn in Figure 1 is that which results from the model presented in Table A1 in the Appendix, with corresponding poverty estimate in column (5) of Table 3.
[17] A similar figure for only urban households cannot be created since the survey design is such that some strata include only a single primary sampling unit, which prohibits the creation of a 95% confidence bound for actual welfare levels.

consumption are less promising, as there is evidence against the assumption of a stable Engel curve relating the included and excluded sub-components.

Russia

The Russian Longitudinal Monitoring Surveys (RLMS) are nationally-representative surveys that comprise a panel of dwellings and have similar consumption modules between rounds as well as many other similar survey components. Data from rounds 5, 8 and 12, corresponding to years 1994, 1998, and 2003, respectively, are used to test the assumption of stable point estimates using the SAE methodology.[18] In Russia, unlike Vietnam, the surveys document a sharp *increase* in recorded poverty from 15.0% in 1994 to 38.6% in 1998, after which poverty is estimated to have fallen to 11.6% by the twelfth round in 2003.[19] The sharp increase in poverty observed in 1998 was a result of the Russian Financial Crisis, and we test whether the SAE methodology can capture this "shock" in poverty rates over time.

Table 5 presents results using the matching sub-components of consumption expenditures as explanatory variables to test the assumption of a stable Engel curve in Russia between survey rounds.[20] While the models capture the general directions of the large swings in poverty rates associated with the Russian Financial Crisis of 1998, they are not systematically convincing across all specifications and data rounds, and in

[18] Information on the precise models of consumption utilized for this exercise on the RLMS data available from the author upon request.
[19] Poverty figures are authors' calculations based on RLMS income data. Corresponding poverty headcounts using consumption data from rounds 5, 8 and 12 are 11.4%, 33.8%, and 11.1%, respectively.
[20] Data limitations prevent us from replicating the exact sub-categories of consumption expenditures as we did with the Vietnam data. We disaggregate the Russian expenditures data into only matching food and nonfood expenditures.

aggregate they suggest that the assumption of a stable Engel curve between rounds is likely inappropriate.

Table 5: Russia – Expenditure Sub-Components Models Rounds 5,8 and 12

National, Poverty Headcount Consumption Poverty	(1)	(2)	(3)		(4)		(5)	(6)	(7)
Food expenditures			X				X		
Nonfood expenditures					X		X		
	1994	1998						1998	2003
Baseline	5	8						8	12
1994	11.4 (0.6)	-	35.0 (1.2)	**	26.7 (1.4)		38.4 (0.9)	33.8 (1.1)	
1994	11.4 (0.6)	-	21.8 (1.1)		11.3 (1.0)	**	20.6 (0.9)		11.1 (0.6)
1998		33.8 (1.1)	25.8 (1.4)		14.4 (0.9)		15.0 (0.7)		11.1 (0.6)

Notes: Standard errors are in parentheses. Columns (1) and (2) contain official consumption poverty numbers from the baseline rounds 5 and 8 respectively. Columns (6) and (7) contain the official poverty numbers that are imputed (either round 8 or round 12). Columns (3) through (5) contain poverty estimates using differing subcomponents of expenditures. The X's at the top of a column indicate if a category of explanatory variables is included in the models whose results are present. (**): Signifies that the poverty estimate is not statistically different at the 5% level from the official poverty statistic for 2005 as listed in column (7)

When a model of consumption is estimated in 1994 based on matching sub-components of expenditures between rounds 5 and 8 (corresponding to years 1994 and 1998), Table 5 correctly estimates a steep rise in predicted poverty. The estimate based on the model using only matching food expenditures does not differ from the "true" poverty estimate for 1998. However, the model that includes only nonfood expenditures underestimates poverty by a statistically significant amount, while the model that includes both food and nonfood expenditures overestimates poverty by a statistically significant amount. The pattern of these results is similar to those found in Vietnam, in that the models that combine food and nonfood expenditures generally overestimate poverty. However, here the "truth" against which our results are being measured is a significant increase in poverty between rounds, and the model based only on food expenditures outperforms the model based only on nonfood expenditures.

When a 1994 model of consumption is imputed into the 2003 data (corresponding to rounds 5 and 12), bypassing the "crisis" year of 1998, the resulting poverty estimates overestimate poverty in the models that include food expenditures. Now it is the model that includes only nonfood expenditures that predicts a poverty rate that does not differ significantly from the "true" rate for 2003. How informative this is remains unclear, as the models that are restricted to include only the matching nonfood expenditures between any two rounds have much smaller sample sizes. This is because many households are recorded as having zero expenditures on the matching nonfood items in the Russian data, and are subsequently dropped from the analysis when converted into log form.

When a 1998 model of consumption is imputed into the 2003 data, we see a similar pattern to that found in Vietnam where the model that includes only food expenditures overestimates poverty, while the models that include nonfood expenditures correctly show a large decrease in predicted poverty rates, although their predictions do not fall within the 95% confidence bound for the "true" poverty rate in 2003. In sum, the pattern of these results can be summarized in the following way:

$$5 \to 8 : \hat{H}_{nf} < \hat{H}_{f} \leq \hat{H}_{f+nf}$$
$$5 \to 12 : \hat{H}_{nf} < \hat{H}_{f+nf} \leq \hat{H}_{f}$$
$$8 \to 12 : \hat{H}_{nf} \leq \hat{H}_{f+nf} < \hat{H}_{f}$$

The pattern of results observed when a round 8 model is imputed into round 12 most closely resembles that from Vietnam. This is also the only span where we are once again trying to predict a large decline in poverty rates. Overall, using the reasoning outlined in section 2, there is no convincing evidence in support of the assumption of a stable Engel curve over time.

Table 6 presents corresponding results for the "non-consumption" models based on matching variables from all five different categories outside the consumption modules of the Russian surveys. As poverty statistics in Russia are calculated based on both income and consumption expenditures the lower panel of Table 6 also presents results using household income as the dependent variables in the imputation exercise. One difficulty specific to the Russian setting that may constrain the ability of these "non-consumption" models to capture a large increase in poverty for 1998 is that during the crisis households may have purchased durables to be able to retain their savings in the face of runaway inflation. The 1998 survey includes an additional set of questions asking precisely whether households undertook such behavior during the crisis, and over 12% of households admitted to these types of purchases.

Table 6: National and regional "Non-consumption" income and consumption models 1994, 1998, 2003

Poverty Headcount	(1)	(2)	(3)		(4)	(5)
	1994	1998	Non-consumption assets		1998	2003
Round	5	8			8	12
Consumption						
Region						
National	11.4		12.2		33.8	
	(0.6)	-	(0.8)		(1.1)	
National		33.8	26.9			11.1
		(1.1)	(0.9)			(0.6)
National	11.4		7.9			11.1
	(0.6)	-	(1.1)	**		(0.6)
	9.3		8.8			2.9
Region 1: Moscow/St. Petersburg	(1.5)		(2.3)	**		(0.8)
	10.1		2.5			6.6
Region 2: North/Northwestern	(2.4)		(1.6)	**		(2.0)
	9.8		5.4			9.8
Region 3: Central/Central Black Earth	(1.2)		(1.2)	**		(1.3)
	13.8		10.6			16.6
Region 4: Volga/Viask/Volga Basin	(1.6)		(1.9)	**		(1.7)
	9.9		12.2			10.5
Region 5: North Caucuses	(1.5)		(2.2)	**		(2.1)
	12.1		4.8			15.1
Region 6: Urals	(1.8)		(1.4)			(1.7)
	12.9		5.6			8.7
Region 7: Western Siberia	(2.4)		(2.9)	**		(1.1)
	12.5		7.4			12.9
Region 8: Eastern Siberia/Far East	(2.2)		(1.9)	**		(1.9)

Poverty Headcount	(1)	(2)	(3)		(4)	(5)
	1994	1998	Non-consumption assets		1998	2003
Round	5	8			8	12
Income						
Region						
National	15.0		17.9		38.6	
	(0.7)		(1.3)		(1.1)	
National		38.6	34.5			11.6
		(1.1)	(1.0)			(0.7)
National	15.0		10.7			11.6
	(0.7)		(1.0)	**		(0.7)
	8.2		3.3			3.4
Region 1: Moscow/St. Petersburg	(1.6)		(1.0)	**		(1.0)
	14.4		7.9			10.7
Region 2: North/Northwestern	(2.7)		(2.6)	**		(2.4)
	14.0		10.3			10.1
Region 3: Central/Central Black Earth	(1.5)		(1.8)	**		(1.4)
	13.8		15.8			12.0
Region 4: Volga/Viask/Volga Basin	(1.7)		(2.2)	**		(1.5)
	17.3		22.9			17.7
Region 5: North Caucuses	(2.1)		(2.8)	**		(2.4)
	13.1		7.9			12.0
Region 6: Urals	(1.6)		(1.5)	**		(1.7)
	23.9		10.4			10.0
Region 7: Western Siberia	(3.4)		(2.5)	**		(1.8)
	19.7		8.1			16.9
Region 8: Eastern Siberia/Far East	(2.6)		(1.7)			(2.2)

Notes: Standard errors are in parentheses. Columns (1) and (2) contain official consumption poverty numbers from the baseline rounds 5 and 8 respectively. Columns (4) and (5) contain the official poverty numbers that are imputed (either round 8 or round 12). Column (3) contains the imputed poverty estimates using non-consumption variables. The regional poverty rates are only demonstrative as the RLMS data are not designed to be representative at the regional level. (**): Signifies that the poverty estimate is not statistically different at the 5% level from the official poverty statistic for 2005 as listed in column (7)

The top panel of Table 6 presents results with consumption expenditures as the outcome of interest. Again, the methodology largely fails to capture the large swings in poverty rates across these tumultuous years in the Russian economy. However, in both the consumption and income panel more promising results are garnered when the "shock" to poverty rates of 1998 is bypassed, and a model of income or consumption based on round 5 data from 1994 is imputed into the round 12 data of 2003. Despite nine years separating these two rounds of surveys, the methodology is able to produce an estimated income poverty headcount for 2003 of 10.7% that is not statistically different from the "official" headcount of 12.1% for the income specification, and a predicted poverty

headcount of 7.9% which cannot be rejected to be equal to the "true" consumption poverty headcount of 11.1%.

However, because the "true" national poverty rates from 1994 and 2003 are nearly equal, it is unclear how informative this evidence is. Thus, we compare the "true" and "estimated" poverty rates for eight geographic regions covered by the RLMS data between 1994 and 2003 using regional models of consumption estimated for 1994.[21] The different regions of Russia had greater variation in their evolution of poverty during this time than did the national aggregate. However, given that the RLMS data is not designed to be representative at the regional level, this exercise and the "true" regional poverty rates are only demonstrative. Across a period of nine years, seven of the eight geographic regions have estimated income and consumption poverty headcounts that do not differ significantly from their "true" poverty headcounts, which supports the much better performance of the non-consumption models for the 1994-2003 period..

In sum, the evidence from the Russian panel confirms the precarious nature of the stable Engel curve assumption already indicated in the Vietnam data and highlights that, even when using household assets, the SAE methodology may not do a good job in all circumstances, especially not when there are big swings in poverty rates that are largely the byproduct of a major "shock" to the economy (in this case, the Russian Financial Crisis of 1998).

[21] We do not perform a similar analysis at the regional level with the 1998 data, as our national-level results do not encourage further scrutiny at the regional level with the 1998 data.

Kenya

The two latest household expenditure surveys in Kenya are the 1997 Welfare Monitoring Survey (WMS) and the 2005 Kenya Integrated Household Budget Survey (KIHBS). While these surveys were implemented during different periods of the year and more detailed consumption data was collected during the KIHBS, a lot of care was taken in the questionnaire design and the poverty analysis to enable comparability over time.[22] During this period, Kenya observed several years of recession, followed by an economic recovery that started in 2003, driven by horticulture, electricity and communications in the primary, secondary and tertiary sectors respectively (World Bank, 2008). This combination of recession and recovery resulted in real annual per capita GDP growth of only about 0.2 percent during 1997-2006, a limited decline in the national poverty headcount (from 50.8 percent in 1997 to 46.6 percent in 2005), and some deterioration of the living conditions of those who remained poor with the national poverty gap increasing from 16.2 to 16.6 respectively.

Poverty declined most rapidly in Nairobi (almost by 50 percent), while it continued to hover around 43 percent in the other urban areas. Rural areas experienced only a marginal decline of about 3 percentage points. These trends are reflected in the evolution of household assets (Table A11) with improvements especially in secondary education, better housing quality and sanitation, and more household durables (sofas, TVs). Progress in these variables is especially pronounced in Nairobi.

[22] 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).

The focus is on testing whether the "stable returns to assets" assumption holds despite substantial swings in the economy during the period under study. The same five categories of assets are considered as in the Vietnam data (geographic indicators, household demographics, education, housing quality variables, ownership of durables). Those assets that were comparable across the surveys were identified and poverty estimates in 2005 were obtained by applying the SAE technique to different combinations of subsets of the five categories, each time selecting those variables that maximized the model's predictive power of consumption in 1997 through stepwise regression.

Column (2) in Table 7 presents the results of the sparsest consumption model, based only on the geographic indicators and the household demographics. While this performs well at the national level, across the different models it displays on average the highest (absolute) deviation from the "true" poverty numbers. The fact that none of its predicted poverty rates is statistically significantly different from the observed poverty rates is misleading in this case, as the predicted poverty numbers all display high standard errors, indicating imprecise estimation. This is not surprising given that the consumption model is largely driven by the geographic indicators, only complemented with the household demographics. Deviations from the observed poverty numbers are especially large in Nairobi, Nyanza and the Western Province. Augmenting model (2) with the education variables (as in column (3)) improves the performance slightly.

Adding variables capturing housing quality (column (4)) and ownership of durables (column (5)) considerably improves the poverty predicting performance of the model, as in Vietnam. In model (4) the rural, other urban, and Nairobi poverty headcounts are predicted at 50.4, 44.1, and 20.6 percent respectively, compared with

observed rates of 49.7, 42.7 and 20.6 percent. For 9 of the 10 poverty headcount predictions is the predicted not statistically different from the observed one. The results for model (5) are similar.

Given its a priori theoretical supremacy, the performance of a model limited to the geographic indicators, housing quality and durable ownership variables (i.e. without the labor and education variables) is also examined (column 6). Echoing the findings from the Vietnam sample, model (6) performs slightly better at the national level than (4) and (5), but displays a slightly higher average (absolute) deviation from the observed poverty rates across the regions. In sum, the results from Kenya very much support the earlier findings from Vietnam (and Russia for 1994-2003) that application of the SAE technique provides a good alternative to predict poverty when applied to a broad set of asset indicators. The models perform best when geographic indicators, housing quality and durable ownership are included. The predictive power improves only slightly when labor and education are further included.

Table 7: Test of Assumption of Stable Parameter Estimates: Non-Expenditure Models, Kenya 1997 to 2005

Included in Models		(1)	(2)		(3)		(4)		(5)		(6)		(7)
___ Geographic Indicators			x		x		x		x		x		
___ Household Demographics			x		x		x		x				
___ Education					x		x		x				
___ Housing Quality							x		x		x		
Asset Ownership									x		x		
		1997											2005
Region		WMS											KIHBS
National	P_0	**50.8**	46.8	**	46.3	**	45.5	**	44.4	**	46.8	**	**46.6**
		1.1	*1.6*		*1.3*		*1.4*		*1.2*		*1.2*		*0.6*
	P_1	**16.2**	15.5	**	14.8		14.1		13.6		14.7		**16.6**
		0.5	*0.8*		*0.7*		*0.6*		*0.5*		*0.6*		*0.3*
Rural	P_0	**52.8**	52.4	**	47.9	**	50.4	**	50.2	**	45.5		**49.7**
		2.0	*1.5*		*3.3*		*2.7*		*1.9*		*1.1*		*0.7*
	P_1	**22.3**	16.6	**	16.3	**	17.8	**	19.9	**	26.1		**17.8**
		2.3	*0.7*		*2.1*		*3.0*		*2.1*		*0.8*		*0.3*
Other Urban	P_0	**43.2**	41.2	**	43.4	**	44.1	**	45.8	**	40.6	**	**42.7**
		2.6	*3.7*		*2.6*		*2.2*		*2.8*		*3.0*		*1.6*
	P_1	**14.5**	13.4	**	15.0	**	15.4	**	16.1	**	13.2	**	**14.9**
		1.3	*1.6*		*1.2*		*1.2*		*1.3*		*1.3*		*0.7*
Nairobi	P_0	**40.0**	36.1	**	32.5		20.6	**	26.8	**	19.5	**	**20.6**
		4.5	*10.4*		*5.0*		*4.5*		*2.6*		*3.1*		*2.5*
	P_1	**11.4**	11.8	**	9.3	**	5.5	**	7.4	**	5.3	**	**6.2**
		2.2	*4.3*		*2.0*		*1.8*		*1.2*		*1.2*		*0.9*
Central	P_0	**39.0**	36.7	**	34.4	**	32.8	**	28.2	**	32.3	**	**30.4**
		3.2	*4.0*		*3.2*		*3.4*		*3.1*		*2.9*		*1.7*
	P_1	**10.5**	9.7	**	9.1	**	8.9	**	6.9		8.8	**	**9.6**
		1.1	*1.5*		*1.1*		*1.1*		*1.0*		*1.1*		*0.8*
Coast	P_0	**60.1**	52.6	**	53.7	**	54.8	**	57.1	**	62.5	**	**58.7**
		3.0	*5.2*		*4.1*		*3.8*		*3.0*		*3.2*		*2.1*
	P_1	**20.4**	18.5	**	18.6	**	19.3	**	20.1	**	22.4	**	**21.0**
		1.7	*2.6*		*2.2*		*2.1*		*1.5*		*2.0*		*1.0*
Eastern	P_0	**58.6**	54.9	**	54.9	**	55.1	**	52.9	**	53.1	**	**50.4**
		3.7	*3.4*		*3.7*		*2.9*		*3.2*		*2.8*		*1.6*
	P_1	**17.1**	15.8	**	15.7	**	15.4	**	15.0	**	15.5	**	**17.5**
		1.6	*1.5*		*1.7*		*1.3*		*1.3*		*1.3*		*0.8*

Included in Models		(1)	(2)		(3)		(4)		(5)		(6)		(7)
___ Geographic Indicators			x		x		x		x		x		
___ Household Demographics			x		x		x		x				
___ Education					x		x		x				
___ Housing Quality							x		x		x		
Asset Ownership									x		x		
		1997											2005
Region		WMS											KIHBS
Nyanza	P_0	**63.1**	56.7		54.7		52.4	**	49.8	**	55.2		**46.4**
		4.1	*3.8*		*3.9*		*3.9*		*3.7*		*3.7*		*1.6*
	P_1	**22.0**	19.0	**	19.0	**	16.9	**	15.1	**	18.4	**	**16.2**
		2.4	*1.9*		*2.2*		*1.7*		*1.7*		*1.7*		*0.7*
Rift Valley	P_0	**50.6**	49.1	**	46.4	**	45.2	**	42.9		46.2	**	**48.4**
		2.1	*2.7*		*2.6*		*2.2*		*1.9*		*2.3*		*1.2*
	P_1	**15.8**	15.8	**	14.4		14.2		13.0		14.2		**17.3**
		1.0	*1.3*		*1.2*		*0.9*		*0.8*		*1.0*		*0.6*
Western	P_0	**58.5**	59.4	**	59.6		59.4		59.4	**	57.9	**	**53.1**
		2.7	*3.2*		*2.6*		*2.5*		*3.1*		*3.3*		*2.0*
	P_1	**19.8**	21.2	**	20.5	**	20.9	**	19.7	**	19.2	**	**18.6**
		1.3	*1.6*		*1.4*		*1.2*		*1.5*		*1.6*		*0.9*
Validation criteria													
% of times predicted poverty NOT statistically different from official poverty statistic (at 5% level)	P_0		100		70		90		90		80		
	P_1		90		80		80		70		70		
Average difference between predicted and official poverty statistic across regions	P_0		5.4		4.5		3.0		3.4		3.2		0.0
	P_1		2.1		1.9		1.4		2.0		2.3		0.0
Standard deviation of difference between predicted and official poverty statistic across regions	P_0		4.7		3.6		2.2		2.0		2.4		0.0
	P_1		1.5		1.0		1.0		1.1		2.2		0.0

Notes: Column (1) contains official poverty numbers from WMS 1997. Column (8) contains official poverty numbers from KIHBS 2005. Standard errors are in italics. Columns (2) through (6) contain poverty estimates using differing categories of explanatory variables as explained in the text. The X's at the top of a column indicate if a category of explanatory variables is included in the models whose results are present. (**): Signifies that the poverty estimate is not statistically different at the 5% level from the official poverty statistic for 2005 as listed in column (7)

4. Conclusions

This study has empirically examined the primary assumption underlying the adaptation of the SAE methodology in producing reliable poverty estimates over time. Within the class of "consumption" models that utilize sub-components of the consumption aggregate in building a model of consumption to estimate poverty, there was evidence of substitution patterns between the included and excluded sub-components of expenditures. This sheds doubt on the assumption of a stable Engel curve over time.

Within the class of "non-consumption" models examined here, there was no conclusive evidence against the assumption of stable parameter estimates over time, especially not where limited or rapid declines in poverty were observed. These asset models worked best when geographic indicators, housing quality and durable ownership variables were also included. The additional inclusion of labor and education variables did not deteriorate the predictive power, providing some evidence that the returns to these assets are likely also relatively stable. The improvement of the predictions from including the labor and education variables were however marginal, indicating that the results do not hinge on the inclusion of these types of variables. Those who might still want to exclude them a priori on theoretical grounds would thus not stand to lose a lot of predictive power.

Overall, the empirical findings in this study provide cautious optimism for the validity of the assumption of stable returns to assets and the ability of this adaptation of the SAE methodology to correctly approximate the evolution of poverty within a country. Nonetheless, the methodology should not be used indiscriminately, especially not when an economy has undergone a major “shock”, as illustrated by the results from Russia.

Yet outside of this possibility, the methodology holds promise for generating estimates of welfare where they otherwise do not exist.

Works Cited

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.

Deaton, Angus. 2001. "Adjusted Indian Poverty Estimates for 1999-2000." *Economic and Political Weekly*. January 25, 2003. pp. 322-326.

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. 2003. "Micro-Level Estimation of Poverty and Inequality." *Econometrica*, 71(1): 355-364. January.

Kijima, Yoko, and Peter Lanjouw. 2005. "Economic Diversification and Poverty in Rural India." *Indian Journal of Labour Economics*, 8(2): 2005.

Kijima, Yoko, and Peter Lanjouw. 2003. "Poverty in India during the 1990s: A Regional Perspective." Policy Research Department Working Paper No. WPS3141, The World Bank.

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), March.

Luoto, Jill. 2006. " Re-Estimating Uganda's Poverty Numbers: Assessing the Robustness of the Observed Increase in Poverty." University of California, Berkeley. *Mimeo*.

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 Counties in Africa." *World Development*, 28(12): 2123-55.

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.

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.

Appendix

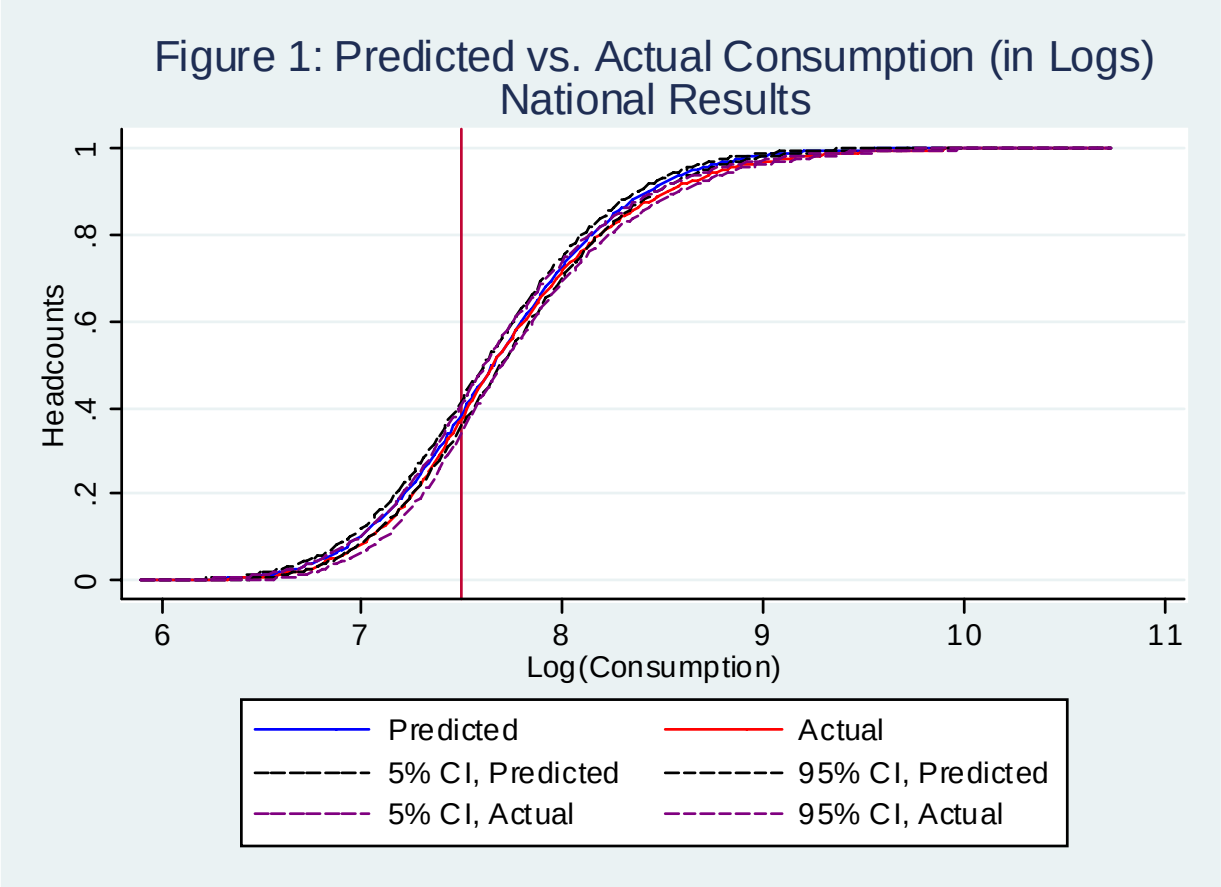
Figure 1: Predicted vs. Actual Consumption (in Logs)
National Results
Headcounts
0
.2
.4
.6
.8
1
6
7
8
9
10
11
Log(Consumption)
Predicted
Actual
5% CI, Predicted
95% CI, Predicted
5% CI, Actual
95% CI, Actual

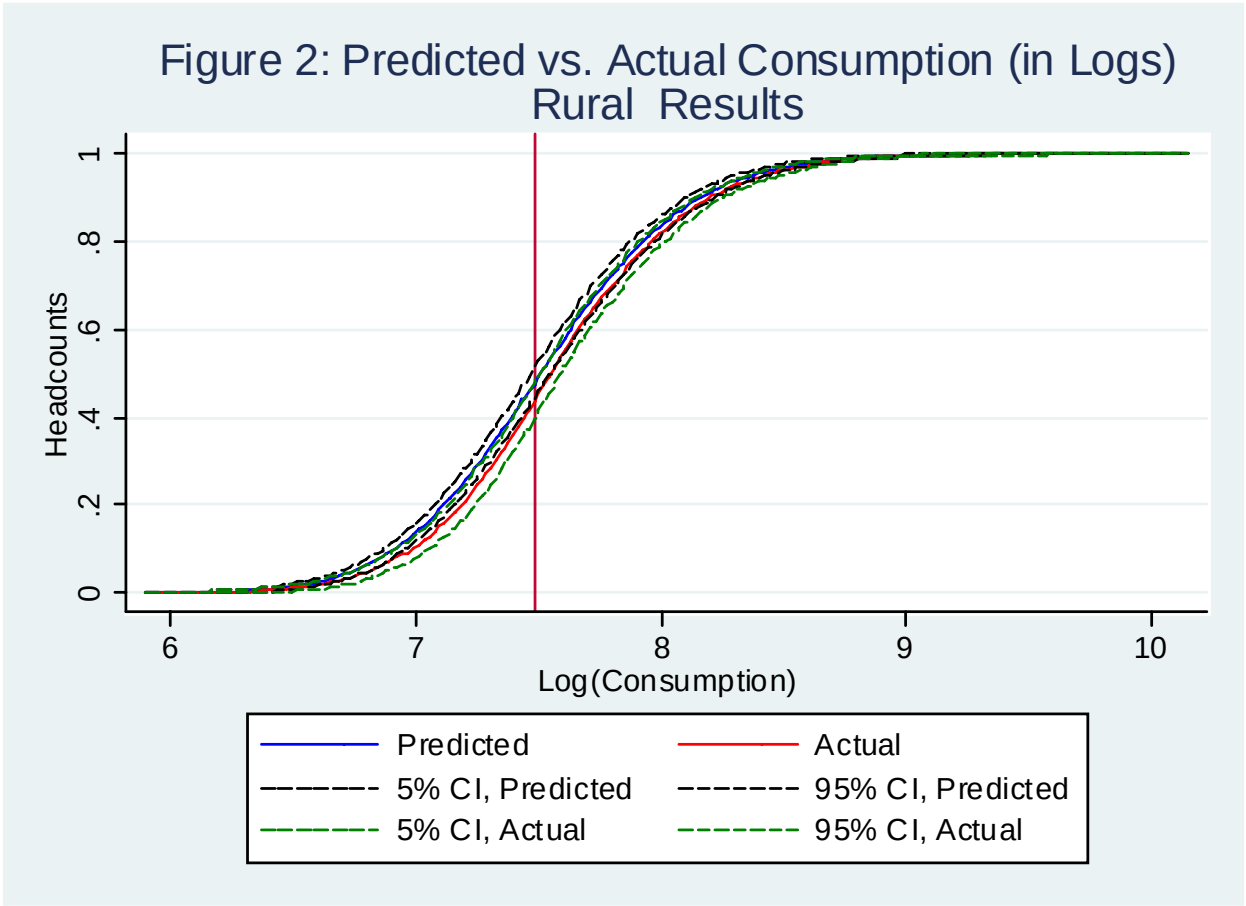
Figure 2: Predicted vs. Actual Consumption (in Logs)
Rural Results
Headcounts
0
.2
.4
.6
.8
1
6
7
8
9
10
Log(Consumption)
Predicted
Actual
5% CI, Predicted
95% CI, Predicted
5% CI, Actual
95% CI, Actual

Table A1
Vietnam: National Results from VLSS 1992-1993

National	Point Estimate	Robust S.E.	p-value
Geographic indicators:			
Dummy: HH located in Central Coast region	0.135	0.019	0.000
Dummy: HH located in South East region	0.269	0.020	0.000
Dummy: HH located in Mekong River region	0.355	0.016	0.000
Household demographics:			
Household size	-0.151	0.007	0.000
Square of household size	0.006	0.001	0.000
Dummy: HH has no children aged 0-5	0.109	0.012	0.000
Education and Profession Variables:			
Average adult male education (in levels)	0.058	0.006	0.000
Dummy: HH does not operate a farm	0.167	0.016	0.000
Housing quality indicators:			
Dummy: House has permanent floor material	0.103	0.014	0.000
Dummy: HH has electricity in home	0.046	0.015	0.003
Dummy: HH lives in one room house	-0.078	0.013	0.000
Additional Wealth Indicators:			
Dummy: HH does not own a motorbike	-0.294	0.021	0.000
Dummy: HH owns a sewing machine	0.141	0.017	0.000
Dummy: HH does not own a radio or stereo	-0.176	0.014	0.000
Dummy: HH owns a bicycle	0.070	0.013	0.000
Dummy: HH owns a color television	0.284	0.024	0.000
Dummy:HH owns an electric fan	0.114	0.017	0.000
Number of handtools owned by HH	0.011	0.001	0.000
Intercept	7.823	0.038	0.000
R-squared	0.543		
Number of observations	4799		

Table A2
Rural Results from VLSS 1992-1993

Rural: All Variables	Point Estimate	Robust S.E.	p-value
Geographic indicators:			
Dummy: Household located in Central Coast region	0.188	0.021	0.000
Dummy: Household located in Central Highlands region	0.190	0.033	0.000
Dummy: Household located in South East region	0.282	0.023	0.000
Dummy: Household located in Mekong River region	0.360	0.018	0.000
Household demographics:			
Household size	-0.158	0.008	0.000
Square of household size	0.006	0.001	0.000
Dummy: Household has no children aged 0-5	0.098	0.012	0.000
Education and Profession Variables:			
Average adult male education (in levels)	0.065	0.006	0.000
Dummy: HH operates a farm	-0.116	0.017	0.000
Dummy: HH does not operate a self-employed business	-0.091	0.013	0.000
Housing quality indicators:			
Dummy: Number of rooms in house is one	-0.085	0.014	0.000
Dummy: House does not have permanent floor material	-0.069	0.014	0.000
Dummy: House shares toilet with other households	-0.074	0.013	0.000
Dummy: House fuel source is coal/bottled gas/electricity/kerosene	0.219	0.032	0.000
Additional Wealth Indicators:			
Dummy: HH does not own a motorbike	-0.264	0.027	0.000
Dummy: HH does not own a sewing machine	-0.115	0.019	0.000
Dummy: HH does not own a radio or stereo	-0.137	0.015	0.000
Number of handtools owned by HH	0.014	0.001	0.000
Dummy: HH owns a bed	0.132	0.022	0.000
Dummy: HH does not own a bicycle	-0.075	0.013	0.000
Dummy: HH does not own a black and white television	-0.106	0.020	0.000
Dummy: HH does not own a color television	-0.234	0.036	0.000
Dummy: HH does not own a wall or table clock	-0.054	0.015	0.000
Intercept	8.494	0.060	0.000
R-squared	0.468		
Number of observations	3839		

Table A3
Urban Results from VLSS 1992-1993

Urban: All Variables	**Point Estimate**	**Robust S.E.**	**p-value**
Geographic indicators:			
Dummy: Household located in Northern Uplands Region	-0.163	0.039	0.000
Dummy: Household located in Mekong River Region	0.207	0.033	0.000
Household demographics:			
Household size	-0.207	0.017	0.000
Square of household size	0.009	0.001	0.000
Dummy: Household has three female adults (15+)	0.147	0.036	0.000
Education and Profession Variables:			
Average adult female education (in levels)	0.027	0.012	0.019
Dummy: Household operates self-employed business	0.084	0.026	0.001
Housing quality indicators:			
Dummy: House has permanent floor material	0.112	0.037	0.003
Dummy: House has private drinking tap	0.106	0.030	0.000
Dummy: House fuel source is wood/grass/leaves/straw	-0.138	0.027	0.000
Dummy: House does not have electric lighting	-0.145	0.044	0.001
Additional Wealth Indicators:			
Dummy: HH owns a radio or stereo	0.186	0.027	0.000
Dummy: HH does not own a camera	-0.284	0.068	0.000
Dummy: HH does not own a wall or table clock	-0.116	0.028	0.000
Dummy: HH owns a color television	0.204	0.033	0.000
Dummy: HH owns a refrigerator	0.185	0.038	0.000
Dummy: HH does not own a motorbike	-0.299	0.033	0.000
Intercept	8.821	0.106	0.000
R-squared	0.630		
Number of observations	960		

Table A4
Northern Uplands Results from VLSS 1992-1993

Northern Uplands: All Variables	Point Estimate	Robust S.E.	p-value
Household demographics:			
Household size	-0.197	0.023	0.000
Square of household size	0.010	0.002	0.000
Dummy: HH has no children	0.098	0.032	0.002
Dummy: HH has 2 children aged 0-5	-0.099	0.029	0.001
Dummy: HH has 2 children aged 5-10	-0.083	0.027	0.002
Education and Profession Variables:			
Average adult female education (in levels)	0.048	0.013	0.000
Average adult female education (in levels)	0.058	0.011	0.000
Dummy: HH operates a farm	-0.117	0.038	0.002
Housing quality indicators:			
Dummy: House does not have a toilet	-0.079	0.025	0.002
Dummy: House has permanent floor material	0.116	0.026	0.000
Dummy: House does not have permanent roof material	-0.111	0.024	0.000
Dummy: HH owns second house	0.249	0.075	0.001
Additional Wealth Indicators:			
Dummy: HH does not own a bicycle	-0.103	0.024	0.000
Dummy: HH owns a black and white television	0.119	0.036	0.001
Dummy:HH owns an electric fan	0.109	0.035	0.002
Dummy: HH owns a refrigerator	0.572	0.104	0.000
Dummy: HH owns a motorbike	0.156	0.045	0.000
Dummy: HH owns a radio or stereo	0.121	0.027	0.000
Number of handtools owned by HH	0.019	0.002	0.000
Intercept	7.595	0.083	0.000
R-squared	0.616		
Number of observations	736		

Table A5
Red River Delta Results from VLSS 1992-1993

Red River Delta: All Variables	Point Estimate	Robust S.E.	p-value
Geographic indicators:			
Dummy: HH located in urban area	0.191	0.039	0.000
Household demographics:			
Household size	-0.074	0.006	0.000
Dummy: HH has no children aged 0-5	0.129	0.020	0.000
Education and Profession Variables:			
Dummy: HH does not operate a farm	0.110	0.026	0.000
Average adult female education (in levels)	0.026	0.011	0.018
Housing quality indicators:			
Dummy: House has 4 or more rooms	0.096	0.047	0.040
Dummy: House fuel source is wood/grass/leaves/straw	-0.246	0.035	0.000
Dummy: House has private toilet	0.097	0.021	0.000
Dummy: House has permanent roofing material	0.084	0.025	0.001
Additional Wealth Indicators:			
Dummy: HH owns a motorbike	0.343	0.042	0.000
Dummy: HH does not own a radio or stereo	-0.121	0.025	0.000
Dummy: HH owns a bicycle	0.114	0.022	0.000
Dummy: HH does not own a black and white television	-0.117	0.029	0.000
Dummy: HH owns a wall or table clock	0.098	0.022	0.000
Dummy: HH does not own a color television	-0.311	0.041	0.000
Intercept	7.957	0.080	0.000
R-squared	0.620		
Number of observations	1216		

Table A6
North Central Results from VLSS 1992-1993

North Central: All Variables	Point Estimate	Robust S.E.	p-value
Geographic indicators:			
Dummy: Household located in rural area	0.116	0.046	0.013
Household demographics:			
Household size	-0.052	0.008	0.000
Dummy: HH has no children aged 0-5	0.140	0.026	0.000
Dummy: HH has no children aged 5-10	0.147	0.038	0.000
Dummy: HH has one child aged 5-10	0.083	0.039	0.032
Education and Profession Variables:			
Average adult male education (in levels)	0.049	0.013	0.000
Dummy: HH operates self-employed business	0.087	0.026	0.001
Housing quality indicators:			
Dummy: House does not have permanent floor material	-0.112	0.029	0.000
Dummy: House is not private structure	-0.135	0.054	0.013
Additional Wealth Indicators:			
Dummy: HH does not own a radio or stereo	-0.197	0.034	0.000
Dummy: HH does not own a bicycle	-0.101	0.030	0.001
Dummy: HH owns a wall or table clock	0.109	0.035	0.002
Dummy: HH does not own a color television	-0.384	0.131	0.004
Dummy: HH owns an electric fan	0.155	0.038	0.000
Dummy: HH owns a motorbike	0.334	0.066	0.000
Intercept	7.672	0.149	0.000
R-squared	0.451		
Number of observations	639		

Table A7
Central Coast Results from VLSS 1992-1993

Central Coast: All Variables	Point Estimate	Robust S.E.	p-value
Household demographics:			
Household size	-0.124	0.017	0.000
Square of household size	0.004	0.001	0.000
Dummy: HH has no children	0.127	0.054	0.019
Dummy: HH has no children aged 0-5	0.123	0.039	0.002
Education and Profession Variables:			
Average adult female education (in levels)	0.081	0.021	0.000
Dummy: HH head does not work in agriculture	0.138	0.042	0.001
Housing quality indicators:			
Dummy: House fuel source is coal/bottled gas/electricity/kerosene	0.331	0.058	0.000
Dummy: House has electric lighting	0.163	0.045	0.000
Additional Wealth Indicators:			
Dummy: HH does not own a motorbike	-0.363	0.054	0.000
Dummy: HH owns a radio or stereo	0.150	0.041	0.000
Number of handtools owned by HH	0.019	0.004	0.000
Dummy: HH owns a bed	0.402	0.062	0.000
Dummy: HH owns a bicycle	0.145	0.044	0.001
Dummy: HH owns an electric fan	0.207	0.052	0.000
Intercept	7.210	0.107	0.000
R-squared	0.605		
Number of observations	544		

Table A8
Central Highlands Results from VLSS 1992-1993

Central Highlands: All Variables	Point Estimate	Robust S.E.	p-value
Household demographics:			
Number of children aged 0-15 in household	-0.117	0.017	0.000
Number of female adults aged 15+ in household	-0.075	0.030	0.013
Dummy: HH head has a spouse	-0.194	0.071	0.007
Education and Profession Variables:			
Dummy: HH head is retired or unemployed	0.340	0.110	0.003
Dummy: HH head has no schooling	-0.332	0.076	0.000
Dummy: Household does not operate self-employed business	-0.177	0.069	0.012
Housing quality indicators:			
Dummy: House has electric lighting	0.158	0.074	0.036
Dummy: House has no permanent toilet	-0.332	0.125	0.009
Dummy: House uses public toilet	-0.587	0.117	0.000
Number of rooms in house	0.148	0.036	0.000
Additional Wealth Indicators:			
Dummy: HH does not own a motorbike	-0.352	0.106	0.001
Dummy: HH does not own a bed	-0.257	0.093	0.007
Dummy: HH does not own a bicycle	-0.194	0.068	0.006
Intercept	8.585	0.225	0.000
R-squared	0.765		
Number of observations	128		

Table A9
South East Results from VLSS 1992-1993

South East: All Variables	Point Estimate	Robust S.E.	p-value
Geographic indicators:			
Dummy: Household located in urban area	-0.106	0.055	0.055
Household demographics:			
Household size	-0.189	0.022	0.000
Square of household size	0.007	0.002	0.000
Education and Profession Variables:			
Dummy: Household operates self-employed business	0.081	0.035	0.021
Dummy: HH head ever attended school	0.162	0.051	0.002
Housing quality indicators:			
Number of rooms in house	0.115	0.020	0.000
Dummy: House has permanent wall material	0.103	0.044	0.021
Dummy: House fuel source is coal/bottled gas/electricity/kerosene	0.308	0.049	0.000
Dummy: House has electric lighting	0.189	0.049	0.000
Additional Wealth Indicators:			
Dummy: HH owns a washing machine	0.396	0.145	0.006
Dummy: HH owns a color television	0.214	0.052	0.000
Dummy: HH owns an electric fan	0.133	0.056	0.018
Dummy: HH owns a motorbike	0.359	0.044	0.000
Intercept	7.695	0.088	0.000
R-squared	0.613		
Number of observations	544		

Table A10
Mekong River Results from VLSS 1992-1993

Mekong River: All Variables	Point Estimate	Robust S.E.	p-value
Household demographics:			
Household size	-0.097	0.007	0.000
Dummy: Household has two children aged 5-10	-0.148	0.037	0.000
Education and Profession Variables:			
Dummy: HH has member that works in agriculture	-0.151	0.041	0.000
Dummy: HH head never attended school	-0.209	0.036	0.000
Dummy: HH operates a self-employed business	0.133	0.027	0.000
Housing quality indicators:			
Number of rooms in house	0.073	0.018	0.000
Dummy: House has permanent wall material	0.123	0.037	0.001
Dummy: Household does not own second home	-0.332	0.102	0.001
Additional Wealth Indicators:			
Dummy: HH does not own a motorbike	-0.189	0.058	0.001
Dummy: HH does not own a sewing machine	-0.182	0.031	0.000
Dummy: HH does not own a radio or stereo	-0.179	0.031	0.000
Number of handtools owned by HH	0.020	0.003	0.000
Dummy: HH does not own an electric fan	-0.290	0.050	0.000
Dummy: HH owns a refrigerator	0.294	0.102	0.004
Intercept	8.887	0.127	0.000
R-squared	0.476		
Number of observations	992		

	National			Rural			Other Urban			Nairobi		
	1997	2005	*Diff*	1997	2005	*Diff*	1997	2005	*Diff*	1997	2005	*Diff*
Household Demographic Variables												
Household Size	6.2	6.5	*0.3*	6.4	6.7	*0.3*	5.0	5.9	*0.9*	4.8	4.9	*0.0*
HH head - age	44.9	46.1	*1.2*	46.1	47.5	*1.4*	37.9	41.4	*3.5*	38.3	38.8	*0.4*
No. of HH members - female	3.2	3.3	*0.2*	3.3	3.4	*0.2*	2.6	3.0	*0.4*	2.4	2.4	*0.1*
HH head - female	0.24	0.27	*0.02*	0.25	0.28	*0.03*	0.19	0.25	*0.06*	0.19	0.17	*-0.01*
No. of HH members - adults	3.2	3.6	*0.4*	3.2	3.7	*0.4*	2.8	3.5	*0.7*	3.0	3.0	*0.0*
No. of HH members - kids	3.0	2.9	*-0.1*	3.2	3.1	*-0.1*	2.2	2.4	*0.1*	1.8	1.9	*0.0*
No. HH members - ages under 5	1.0	1.2	*0.2*	1.0	1.2	*0.3*	0.9	1.0	*0.0*	0.8	0.9	*0.1*
No. HH members - ages 5-10	1.2	1.0	*-0.2*	1.3	1.1	*-0.2*	0.8	0.8	*0.0*	0.6	0.6	*0.0*
No. HH members - ages 10-15	0.8	0.7	*-0.1*	0.9	0.8	*-0.1*	0.5	0.6	*0.1*	0.5	0.4	*0.0*
No. HH members - ages over 60	0.1	0.2	*0.0*	0.2	0.2	*0.1*	0.0	0.1	*0.1*	0.0	0.0	*0.0*
No. HH members - female kids	1.5	1.5	*0.0*	1.6	1.5	*0.0*	1.1	1.2	*0.1*	0.9	0.9	*0.0*
No. HH members - female adults	1.7	1.9	*0.2*	1.7	1.9	*0.2*	1.5	1.8	*0.4*	1.5	1.5	*0.1*
No. HH members - male adults	1.5	1.7	*0.2*	1.5	1.8	*0.2*	1.3	1.7	*0.4*	1.5	1.5	*-0.1*
Household Education Variables												
HH head - no education	0.26	0.23	*-0.04*	0.30	0.26	*-0.04*	0.09	0.13	*0.04*	0.05	0.03	*-0.03*
HH head - primary	0.45	0.47	*0.02*	0.47	0.50	*0.03*	0.33	0.37	*0.04*	0.40	0.31	*-0.09*
HH head - secondary	0.24	0.28	*0.04*	0.20	0.23	*0.03*	0.50	0.45	*-0.05*	0.39	0.54	*0.15*
HH head - post secondary	0.05	0.02	*-0.03*	0.04	0.01	*-0.03*	0.08	0.05	*-0.04*	0.15	0.09	*-0.06*
HH females - no education	0.35	0.31	*-0.04*	0.40	0.35	*-0.04*	0.15	0.20	*0.05*	0.08	0.04	*-0.04*
HH females - primary	0.62	0.65	*0.04*	0.64	0.69	*0.05*	0.46	0.56	*0.10*	0.55	0.44	*-0.11*
HH females - secondary	0.26	0.32	*0.06*	0.23	0.28	*0.05*	0.46	0.44	*-0.01*	0.36	0.55	*0.20*
HH females - post secondary	0.05	0.01	*-0.03*	0.04	0.01	*-0.03*	0.06	0.02	*-0.05*	0.10	0.06	*-0.03*
HH males - no education	0.17	0.17	*-0.01*	0.20	0.19	*0.00*	0.06	0.10	*0.04*	0.03	0.01	*-0.02*
HH males - primary	0.56	0.61	*0.05*	0.59	0.65	*0.07*	0.36	0.49	*0.13*	0.43	0.33	*-0.10*

HH males - secondary	0.34	0.37	*0.03*	0.30	0.33	*0.03*	0.53	0.50	*-0.02*	0.51	0.58	*0.07*
HH males - post secondary	0.06	0.03	*-0.04*	0.04	0.01	*-0.03*	0.10	0.06	*-0.04*	0.23	0.11	*-0.12*
Housing Quality Indicators												
House floor of low quality	0.70	0.64	*-0.06*	0.79	0.75	*-0.04*	0.18	0.23	*0.05*	0.29	0.15	*-0.15*
House roof of low quality	0.32	0.22	*-0.09*	0.36	0.26	*-0.10*	0.12	0.08	*-0.03*	0.02	0.01	*-0.01*
Household access to piped water	0.32	0.28	*-0.03*	0.22	0.17	*-0.05*	0.80	0.59	*-0.20*	0.90	0.95	*0.06*
Household water from river, lake,	0.34	0.36	*0.02*	0.40	0.44	*0.04*	0.03	0.08	*0.04*	0.00	0.00	*0.00*
Household has flush toilet	0.07	0.09	*0.02*	0.01	0.01	*0.00*	0.35	0.24	*-0.11*	0.42	0.64	*0.22*
Household has no toilet	0.15	0.00	*-0.15*	0.17	0.00	*-0.17*	0.03	0.01	*-0.02*	0.08	0.00	*-0.08*
Household access to electricity	0.08	0.13	*0.05*	0.02	0.04	*0.02*	0.37	0.41	*0.04*	0.44	0.72	*0.27*
Household Durable/Asset Ownership												
HH owns - Sewing Machine	0.08	0.06	*-0.02*	0.06	0.05	*-0.01*	0.19	0.09	*-0.10*	0.16	0.06	*-0.09*
HH owns - Sofa	0.26	0.47	*0.22*	0.20	0.42	*0.22*	0.60	0.63	*0.03*	0.49	0.77	*0.28*
HH owns - Bicycle	0.27	0.32	*0.05*	0.29	0.35	*0.07*	0.26	0.22	*-0.03*	0.10	0.12	*0.02*
HH owns - Animal Cart	0.03	0.02	*-0.01*	0.03	0.03	*-0.01*	0.02	0.00	*-0.02*	0.01	0.01	*0.00*
HH owns - Motorcycle	0.00	0.01	*0.00*	0.01	0.01	*0.00*	0.00	0.01	*0.01*	0.00	0.00	*0.00*
HH owns - Radio	0.62	0.57	*-0.05*	0.59	0.59	*-0.01*	0.77	0.44	*-0.33*	0.73	0.62	*-0.11*
HH owns - Refrigerator	0.03	0.03	*0.00*	0.01	0.01	*0.00*	0.14	0.10	*-0.04*	0.18	0.22	*0.05*
HH owns - Gas/electric stove	0.06	0.06	*0.00*	0.03	0.02	*-0.01*	0.19	0.12	*-0.07*	0.25	0.32	*0.07*
HH owns - TV	0.09	0.19	*0.10*	0.04	0.13	*0.08*	0.30	0.37	*0.07*	0.33	0.58	*0.24*
HH owns - Stereo	0.05	0.02	*-0.03*	0.03	0.01	*-0.02*	0.14	0.05	*-0.08*	0.12	0.08	*-0.03*
HH owns - Fan	0.01	0.01	*0.00*	0.00	0.00	*0.00*	0.09	0.08	*-0.02*	0.05	0.03	*-0.02*
HH owns - Car	0.03	0.02	*0.00*	0.01	0.01	*0.00*	0.06	0.04	*-0.02*	0.16	0.12	*-0.04*
HH owns - Phone	0.02	0.00	*-0.02*	0.01	0.00	*0.00*	0.07	0.00	*-0.06*	0.16	0.01	*-0.15*