

Estimating Poverty in India Without Expenditure Data

A Survey-to-Survey Imputation Approach

David Newhouse^a and Pallavi Vyas^{a,b}

ASSA Annual Meeting 2021

^a World Bank Group; ^b Ahmedabad University

January 4, 2021



WORLD BANK GROUP



Ahmedabad
University

The Problem

No recent Indian poverty data based on consumption expenditure.

- Poverty rates in India are calculated using consumption expenditure from household socioeconomic surveys conducted by the National Sample Survey Organization (NSSO) every five years.
- The latest data on consumption expenditure is available from the 2011-2012 (68th round) socio-economic survey.
- The next NSSO survey to gather consumer expenditure data was the 75th round from July 2017 to June 2018. The data has not been made available.
- An estimated 25% of the world's extreme poor lived in India as of 2013, meaning that an updated poverty measure for India is critical for regional and international poverty statistics.



WORLD BANK GROUP



Ahmedabad
University

Current Method to Estimate Poverty

Using Consumption Expenditure Data

- The typical method used by the World Bank to adjust poverty to a later year is the “line-up” method.
- As example, to “line-up” a survey from 2012 to 2015, each household’s per capita consumption, measured in 2012, is assumed to grow at the same rate.
- The growth rate is calculated between the last consumption expenditure survey and the Household Final Consumption Expenditure (HFCE) calculated from the National Accounts data.
- The “line-up” method tends to overestimate the pace of poverty decline.



Main Findings

- Poverty Estimates for 2014-15: 12.7% at \$1.90/day per capita.
 - Predict that the urban poverty rate decreases from 13.4% to 10.4%.
 - Predict that rural poverty rate decreases from 24.8% to 13.8%.
- Compare with Predictions of the Other Models.
- Compare the Accuracy of the Model with Other Models.
 1. Forward Prediction into 2011-12: Using 2004-05 and 2009-10 surveys.
 2. Backward Prediction into 2004-05: Using 2009-10 and 2011-12 surveys.
- Robustness Checks.



Survey to Survey Imputation (SSI) Methodology

- Use Elbers, Lanjouw, and Lanjouw (2003) [ELL] model to generate results.
- Predict probability of being poor in the 72nd round (2014-15) of the NSSO data.
 1. Common variables in the three previous rounds: 2004-05, 2009-10, and 2011-12.
 2. Have similar wording and recall periods in the survey.
 3. Sampling frame derived from 2001 and 2011 censuses.
 - a. Geographic areas
 - b. Survey methodology (e.g., stratification, enumeration blocks)
 - c. Sample size
 - d. Similarity in households surveyed



Data

- Dependent Variable: Welfare aggregate for national poverty.
 - Monthly Per Capita Expenditure in 2011 US\$.
- Explanatory Variables: At the Household and District level.
 1. Demographic: household size, ages of household members, religion, caste, gender.
 2. Labor Market : industry, occupation, type of labor (selfemployed, casual, regular wage, other).
 3. Expenses on Miscellaneous Services:
 - Household services: domestic help, barber, beauty , laundry, priest, grinding, tailor.
 - Recreation: movies, theatre, picnics, clubs, photography, hiring of equipment.
 - Transport: fares, petrol, diesel and other conveyance related expenses.
 4. Rainfall: District level. Standard deviation from the mean (1981-2017).

SSI: First Step

1. First Step [Household (HH) Consumption Expenditure Data]:

$$\ln(y_{cht}) = X_{cht}^h \beta_1 + Z_{cht}^h t \beta_2 + X_{ct}^d \beta_3 + Z_{ct}^d t \beta_4 + \beta_5 t + u_{cht} \quad (1)$$

where,

$$u_{cht} = \eta_{ct} + \epsilon_{cht}$$

2. Variables chosen for equation (1) using the Least Absolute Shrinkage and Selection Operator (LASSO) method. The LASSO specification selects a γ vector to minimize:

$$\underset{\gamma}{\text{Min}} \quad \frac{1}{N} \left[\ln(y_{cht}) - x_{cht}^h \gamma_1 - x_{cht}^h t \gamma_2 - x_{ct}^d \gamma_3 - x_{ct}^d t \gamma_4 - \gamma_5 t \right]^2 + \lambda \sum_{j=1}^5 |\gamma_j| \quad (2)$$



Generalized Least Squares

- Because of heteroskedasticity in the error term and spatial correlation from the introduction of η_{ct} , we re-estimate equation (1) using Generalized Least Squares (GLS).
- The weights for the GLS specification are the predicted variances of the error terms from the OLS regression.
- The predicted variances are used to re-estimate equation (1) using GLS (3).

$$\ln(y_{cht}) = X' \beta_{GLS} + u_{cht} \quad (3)$$



Imputation into the Target Dataset [2014-15]

- A simulated consumption level, $\tilde{y}_{cht}$, is calculated for each household:

$$\tilde{y}_{cht} = x' \tilde{\beta} + \tilde{\eta}_{ct} + \tilde{\epsilon}_{cht} \quad (4)$$

- Assumptions for the simulations:

$$\tilde{\beta} \sim N(\hat{\beta}_{GLS}, \text{Var}(\hat{\beta}_{GLS}))$$

- The residuals η_{ct} and ϵ_{cht} are then simulated for each household.

$$\tilde{\eta}_{ct} \sim N(0, \hat{\sigma}_{\eta}^2)$$

$$\hat{\sigma}_{\eta}^2 \sim \text{Gamma}(\bar{\sigma}_{\eta}^2, \text{Var}(\hat{\sigma}_{\eta}^2))$$

- ϵ_{cht} drawn from the empirical distribution of the household residuals.

Welfare Imputation in the Target Dataset [2014-15 Data]

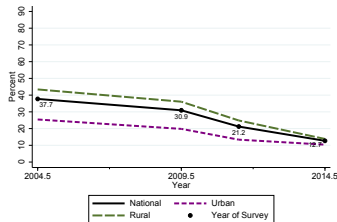
- The mean of 100 simulations of $\tilde{y}_{cht}$ gives the point estimate of household expenditure.
- Standard error is calculated using Rubin's rules using variation within and across households.
- Headcount poverty rates predicted according to the \$1.90 per day per person and \$3.10 and \$5.50 international poverty lines.



Poverty Rates: \$1.90 Per Day

	2004-05	2009-10	2011-12	2014-15
National				
Poverty rate	38.9	31.7	21.6	12.7
Standard Error	0.4	0.4	0.4	0.7
Urban				
Poverty rate	25.4	19.8	13.4	10.4
Standard Error	0.6	0.5	0.4	0.7
Rural				
Poverty rate	43.4	36.1	24.8	13.8
Standard Error	0.4	0.5	0.5	0.8

Sources: India National Sample Survey Office (NSSO) Surveys and staff estimates.



Sources: India National Sample Survey (NSSO) Surveys.



WORLD BANK GROUP



Ahmedabad University

Accuracy of the Models

Forward Prediction into 2011-12

	Actual	Model 1	Model 2	Model 3	Model 4	Line-Up
National	21.1	21.1	24.3	19.7	25.1	22.9
Urban	13.4	16.5	18.7	15.9	23.9	14.6
Rural	24.8	23.2	27.0	21.6	25.7	27.0

Backward Prediction into 2004-05

	Actual	Model 1	Model 2	Model 3	Model 4	Line-Up
National	37.5	40.7	48.9	64.1	28.5	47.3
Urban	25.4	30.4	37.0	48.4	19.5	32.0
Rural	43.4	45.8	54.7	71.7	32.9	54.8

Sources: India National Sample Survey Office (NSSO) Surveys.

Model 1: Final Model

Model 2: District dummies*Time Trend

Model 3: Expenditures at the Extensive Margin

Model 4: Constant Coefficient Model



WORLD BANK GROUP



Ahmedabad University

Predicted Poverty Rates (\$1.90 per Day) from Different Models

	Model 1	Model 2	Model 3	Model 4
National	12.7	18.8	17.4	15.4
Urban	10.4	13.1	15.4	10.0
Rural	13.8	21.6	18.4	18.0

Sources: India National Sample Survey Office (NSSO) Surveys.

Preferred Model (Model 1) vs. Line-Up Method Predictions for 2014-15

	Model	95% C.I		Line-up
National	12.7	11.2	14.2	12.1
Urban	10.4	9.1	11.7	7.7
Rural	13.8	12.3	15.4	14.2

Source: India National Sample Surveys
(NSSO) Surveys.

Robustness Check 1

Elasticity of Poverty to Growth by Model

	Elasticity	Semi-Elasticity
2004-05 to 2009-10	-0.6	-21.4
2009-10 to 2011-12	-3.0	-85.4
Model 1	-2.8	-50.6
District dummies*Time Trend	-0.7	-14.3
Expd at the Extensive Margin	-1.1	-22.7
Constant Coefficient Model	-1.8	-34.6
Typical line-up method	-3.4	-59.4

Sources: India National Sample Survey Office
(NSSO) Surveys.

Robustness Check 2

Changes in Poverty Across States

2004 to 2011 vs. 2011 to 2014

- Relationship between growth rate of per capita state GDP and state level poverty rates is the same for the 2004-2011 and 2011-14 time periods.
- A growth rate of 10% is associated with a reduction of about two percentage points in poverty rates per year.

Robustness Check 3

Actual and Predicted Poverty Rates at \$3.20 Per Day

	2004-05	2009-10	2011-12	2014-15*
National	74.6	69.7	60.4	44.2
Urban	58.4	51.5	43.3	33.4
Rural	82.1	78.1	68.3	49.4

Sources: India National Sample Survey Office (NSSO) Surveys.

*Preferred Model predictions.

Actual and Predicted Poverty Rates at \$5.50 Per Day

	2004-05	2009-10	2011-12	2014-15*
National	95.7	94.3	89.7	77.3
Urban	95.4	92.1	84.5	65.1
Rural	95.8	95.4	92.1	83.2

Sources: India National Sample Survey Office (NSSO) Surveys.

*Preferred model predictions.



WORLD BANK GROUP



Ahmedabad University

