POLICY RESEARCH WORKING PAPER 7270

Poverty Dynamics in India between 2004 and 2012

Insights from Longitudinal Analysis Using Synthetic Panel Data

Hai-Anh H. Dang
Peter F. Lanjouw

Abstract

Recent National Sample Surveys point to significant poverty reduction in India since 2004/05, with a marked acceleration between 2009/10 and 2011/12. This paper enquires into important aspects of income mobility between 2004/05 and 2011/12, based on new statistical methods to convert the three pertinent National Sample Survey rounds into synthetic panels. The analysis draws on the synthetic panels to derive a vulnerability line for India that can be used to separate out a population subgroup comprising non-poor households facing a heightened risk of falling into poverty. The paper documents a strong pattern of upward mobility out of poverty and vulnerability into the middle class, with a noticeable acceleration between 2009/10 and 2011/12. The paper further undertakes a careful investigation into the comparability of the survey rounds, prompted by the observation that fairly significant modifications had been made to survey questionnaires. The findings suggest that changes in questionnaire design have not compromised the comparability of the data.

**Poverty Dynamics in India between 2004 and 2012:
Insights from Longitudinal Analysis Using Synthetic Panel Data**

Hai-Anh H. Dang and Peter F. Lanjouw[*]

World Bank

JEL: C15, I32, O15

Key words: poverty dynamics, mobility, imputation, synthetic panel, India

I. Introduction

Poverty has steadily decreased in India over the past decade. Since India makes up a quarter of the world's poor (i.e., those living under $1.25 a day), which is roughly half again its share of the world's population (17 percent), reducing poverty in this country would not only impact its welfare alone but would also register a significant impact on global poverty estimates.[1] What is particularly striking is the acceleration of poverty reduction that appears to be taking place. Between 2004/5 and 2009/10 poverty declined from 37.7 percent to 29.9 percent. Over the subsequent two years, poverty declined by a further 10 percentage points, to 20.0 percent. These achievements in poverty reduction have been widely remarked on and celebrated.[2]

The aim of this paper is to consider the recent experience of poverty decline from two perspectives that have not historically received a great deal of attention in the Indian context – due most likely to the scarcity of nationally representative panel survey data.[3] First, we ask whether there is any suggestion that those who are currently non-poor remain at a heightened risk of falling into poverty. Second, we investigate to what extent one can discern a core subset of the population

---

[1] We use the poverty rates and population data respectively from the World Bank's PovCalNet database (http://iresearch.worldbank.org/PovcalNet/index.htm) and Development Indicators database. All figures are estimated averages for the two years 2011 and 2012.

[2] There have been nevertheless some concerns raised around the credibility of the most recent episode of poverty decline, we come back to more discussion in the next section. Unless otherwise noted, all the poverty rates are based on the national poverty lines.

[3] Smaller panel surveys have been fielded for India, but none of these provide nationally representative data; see Dercon and Shapiro (2007) for a recent review. For recent studies that use these panel surveys, see, e.g., Munshi and Rosenzweig (2009), Krishna and Shariff (2011), and Dercon, Krishnan, and Krutikova (2013) respectively for analysis of the REDS panel between 1982 and 1999, the NCAER panel between 1993-94 and 2004-05, and the ICRISAT panel between 1975 and 2006. While panel surveys allow more in-depth analysis of mobility, Rosenzweig (2003) discusses potential issues that can bias these surveys (that are not nationally representative) such as split-offs or attrition. A new nationally representative panel survey (IHDS) fielded by the University of Maryland and NCAER promises much improvement over the previous panels (http://ihds.umd.edu/index.html). But note that compared to the NSSs, the IHDS has less than half the sample size and collects a much reduced version of household consumption data (i.e., 47 consumption items in the latter vs. more than 400 items in the former).

that remains stuck in poverty and is somehow unable to participate in the processes of upward mobility.

Our analysis studies the period between 2004/5 and 2011/12, and draws on three "thick" rounds of National Sample Survey data referring to 2004/5, 2009/10, and 2011/12. As noted above, these data sources indicate a substantial decline in poverty over the entire 7-year period, with a sharp acceleration occurring after 2009/10. Our analysis suggests that during the first episode of poverty reduction between 2004/5 and 2009/10, poverty decline was accompanied by an increase in the share of the population that can be considered vulnerable, or facing a heightened risk of falling into poverty. Between 2009/10 and 2011/12, poverty decline accelerated further, and this was now accompanied also by a decline in the fraction of the population that was particularly at risk of falling into poverty.

We show further that aggregate trends in poverty reduction mask a considerable degree of entry into, and exit out of, poverty, but that a substantial core of the poor have remained poor over the duration of the study period. We document some of the key household characteristics of those who have managed to escape poverty and vulnerability, and contrast these with those who have fallen into this undesirable welfare status during this period.

Before turning to a detailed discussion of our empirical findings, we confront in this paper several methodological challenges that have typically held back investigations of the kind we are attempting here. The key difficulty is that analysis of poverty transitions and of the likelihood of escaping, or falling into, poverty depends on the availability of panel data that permit the analyst to follow households over time. Yet in India, as in many other countries, nationally representative panel data are not available. The existing data sources underpinning poverty analysis—the NSS

surveys—are high quality cross-sectional data sources that offer at best a snapshot of living conditions at specific moments of time.

In order to overcome this limitation, we implement in this paper a methodology for converting the NSS cross-sectional surveys into synthetic panels. The approach we follow has been recently introduced into the literature (Dang, Lanjouw, Luoto and McKenzie, 2014a; Dang and Lanjouw, 2013) and a number of studies which validate the method have generally yielded encouraging findings (Dang et al., 2014a; Dang and Lanjouw, 2013; Cruces et al., 2014; Martinez, 2015). We highlight the assumptions imposed by the methodology and discuss their applicability to the Indian context.[4]

The methodology for constructing synthetic panels is predicated on strict comparability of the underlying cross-section surveys. It has already been noted that India's NSS surveys are generally regarded as high-quality data sources. We also focus our attention here on the "thick" rounds that involve larger sample sizes and that are designed to be representative at the rural/urban and state level. Nonetheless we investigate whether the 2009/10 and 2011/12 rounds are strictly comparable, since the possibility of a breakdown in comparability is prompted by the remarkable rate of poverty decline as well as evidence that there are some noticeable changes in the design of the consumption questionnaire between these two years. We note that there had been intensive and contentious debate around the comparability of the 1999/00 round of the NSS survey with earlier NSS rounds, after a certain number of changes and adjustments had been made to the

---

[4] Synthetic panels constructed using the Dang et al. (2014a) and Dang and Lanjouw (2013) methods have been applied to study poverty dynamics in various settings including multi-country analysis for Latin America (Ferreira et al., 2013; Vakis, Rigolini, and Lucchetti, 2015), South Asia (Rama et al., 2015), and Europe and Central Asia (Cancho et al., 2015). Specific country case studies using synthetic panels investigate countries including the Kyrgyz Republic (Bierbaum and Gassmann, 2012), Bhutan (World Bank, 2014), and Senegal (Dang, Lanjouw, and Swinkels, forthcoming). Another promising use of synthetic panels is to evaluate program impacts (Garbero, 2014).

questionnaire (Deaton and Kozel, 2005). We ask therefore, whether there is any call for similar disquiet with the surveys examined in this paper.

We tackle this question with an imputation-based method recently explored in Dang, Lanjouw and Serajuddin (2014b) that builds on a number of earlier studies (Elbers, Lanjouw, and Lanjouw, 2003; Tarozzi, 2007).[5] Our findings suggest that the 2009/10 and 2011/12 survey rounds do not appear to suffer from serious comparability issues. The observation of a sharply accelerated poverty decline after the 2009/10 round, from 29.9 to 20 percent in 2011/12, seems robust. We also appear to be on solid footing with respect to the data underpinnings for converting these three NSS rounds into synthetic panels.

We start, in Section II, with a brief discussion of poverty trends during the late 2000s and explore further the question of whether the 2009/10 and 2011/12 NSS rounds are comparable. Section III describes our efforts to assess the comparability of the 2009/10 and 2011/12 surveys. Besides offering validation evidence for the recent poverty decline, these two sections also describe the preparatory data work required to construct synthetic panels with which to study poverty dynamics. Section IV then implements our approach to convert the three most recent NSS rounds between 2004/05 and 2011/12 into synthetic panels. We also describe and implement in this section an approach to construct "vulnerability lines" that permits a richer mobility analysis by allowing us to identify a particular portion of the population that is non-poor but that faces a heightened risk of falling into poverty. We implement a procedure proposed in Dang and Lanjouw (2014) for specifying vulnerability lines anchored explicitly to the observed incidence of non-poor

---

[5] Elbers et al. (2003) provide a method that imputes household consumption from a survey into a population census. Adapting this approach for survey-to-survey imputation, Christiaensen et al. (2012) impute poverty estimates using data from several countries including China, Kenya, Russia, and Vietnam; other studies analyze data from Morocco (Douidich et al., 2014) and Uganda (Mathiassen, 2013). See also Tarozzi and Deaton (2009) and Rao (2003) for other studies on survey-to-census imputation.

households falling into poverty. This procedure makes light demands on data and can be straightforwardly applied to synthetic panel data.

We then turn in Section V to a discussion of mobility between the three population segments that derive from this analysis: the poor, the vulnerable, and the middle class, and we produce some basic profiles of the population in different transition categories. We end in Section VI with concluding remarks.

II. Poverty Trends and Data

Steady GDP per capita growth has helped drive down poverty rates in India in the late 2000s.[6] In particular, GDP per capita increased by almost half (47 percent) during the period 2004-2009 (World Bank, 2015), and poverty decreased by 21 percent over the same period. The country's continued economic growth resulted in a further increase of GDP per capita over the subsequent two years, by almost one-fifth (19 percent) in 2011/12. While this robust growth rate should be expected to bring more poverty reduction, the contemporaneous fall in poverty rates turned out to be much larger than expected. To quite a few observers, the fall in poverty has been startling.[7]

Figure 1 plots the annual growth rate of GDP per capita (left axis) and the headcount poverty rate (right axis) between 2004 and 2012. Since a large share of the labor force is employed in agriculture, the figure also displays the annual growth rate of the value added per worker of the agricultural sector. The disconnect between GDP per capita growth rates and poverty reduction is

---

[6] See, e.g., Datt and Ravallion (2011) and Ravallion (2011) for comprehensive discussions on economic growth and poverty in India for earlier periods.
[7] For example, Dutta and Panda (2014) observe that there is much controversy around the (arbitrariness) of the specification of the poverty line. Saxena (2013) points out a couple of inconsistencies such as the share of the population that need food subsidies or the slum population in major cities are much larger than the reported poverty rate, and that the specified poverty lines may be too low and may potentially be distorted due to political motives. In addition to these last two issues, Himanshu (cited in Rao, 2013) voices the concern that imputed spending values for certain social transfer programs may not be calculated correctly. See also the BBC (Limaye, 2013), New York Times (Gupta, 2013), and Washington Post (Lakshmi, 2012) for related discussion on the debates on poverty in this period.

brought out sharply where, despite a remarkably weaker growth of the former, the slope of the line representing the latter is much steeper in the second period than that in the first period. An even weaker growth of the agricultural sector further highlights this difference.

Despite the various arguments for or against this swift fall in poverty, one simple but perhaps not unreasonable hypothesis is that the questionnaire design of the consumption module in the 2011/12 (68th) round of the NSS is not comparable to that in the 2009/10 (66th) round (and 2004/05 or 61th round), which in turn leads to inconsistently constructed and incomparable consumption data. Indeed, there are several major changes to the questionnaire in the 68th round that include: i) changing the consumption code, ii) aggregating consumption items in broader groups, iii) disaggregating consumption items in smaller groups, iv) using/ providing somewhat different item names, v) dropping some consumption items in previous rounds, and vi) adding new consumption items. These changes may not be harmless in affecting the comparability of the consumption data over time.[8]

To further investigate whether these changes may lead to different consumption aggregates over time, we explore the raw item-by-item consumption data at the household level and examine a variety of alternative consumption aggregations over time. The results shown in Appendix 1, Table 1.1 confirm that these questionnaire revisions could be a source of concern. While most consumption groupings make up rather similar shares in total household consumption, the share of the items with some change in code (grouping number 2) are two percentage points lower, and the share of the new items added in the 68th round are three percentage points higher than those

---

[8] We use data from Type 1 Schedule for all survey rounds. An Excel file that provides a comparison and detailed tracking of the change to each consumption item for the 61th, 66th, and 68th rounds of the NSS are available upon request. Survey design issues that compromise the comparability of poverty estimates are found in various countries such as China (Gibson, Huang, and Rozelle, 2003), Tanzania (Beegle et al., 2012), and Vietnam (World Bank, 2012). See also Deaton and Grosh (2000) and Crossley and Winter (forthcoming) for general reviews on the influence of survey design on the quality of consumption data in developing and richer countries respectively.

items in the 66th round that are dropped.[9] While these differences may balance out on average, and may not result in any significant change to the total consumption aggregate, they may also point to potentially deeper comparability issues with the consumption data. Moreover, even if mean values are not much affected, these changes could affect different parts of the consumption distribution differently, and could thus still have a bearing on poverty estimates.

The discussion above evokes a similar, but much larger, poverty debate that took place in India in the early 2000s. In the late 1990s, the National Sample Survey Office (NSSO) revised the questionnaire of the NSS in 1999/2000 (55th round) in an attempt to bring estimates of household consumption from the survey in line with those from national accounts. In particular, these revisions include changing the recall period for household durables and education expenses from a 30-day interval to a 365-day interval, and using both the traditional 30-day recall period as well as a new 7-day recall period for food items. The Government of India published estimates showing that the headcount poverty rate fell by 10 percentage points between 1993/1994 and 1999/2000. Independent researchers, however, noted the possibility of non-comparability of the published consumption data, and applied a variety of methods to adjust for this. A variety of estimates were produced with some suggesting a rate of decline ranging from only somewhat lower than the official estimates (Deaton and Dreze, 2002; Tarozzi, 2007) to one estimate suggesting a mere three percentage point decline in poverty during the decade of the 1990s (Sen and Himanshu, 2005; see also Kijima and Lanjouw, 2003). As is powerfully argued in the book "The Great Indian Poverty Debate" (Deaton and Kozel, 2005), concerns about comparability can greatly complicate assessments of poverty trends.

---

[9] Compared with the 61th round, the share of the new items added in the 66th round is approximately 0.1 percent and equals the share of the items that are dropped from the former. This implies greater comparability between these two survey rounds.

We describe in the next section a method for gauging comparability between the 2009/10 and the 2011/12 rounds of the NSS.

III. Predicted Poverty Trends Using Imputation

We provide here a brief overview of the survey-to-survey imputation method described in Dang et al. (2014b) before discussing results. Further discussion on technical details and estimation procedures are available in the cited paper.

III.1. Overview of the Imputation Method

Let x_j be a vector of characteristics that are commonly observed between the two surveys, where j indicates survey round, j= 1, 2.[10] These characteristics can include household variables such as the household head's age, sex, education, ethnicity, religion, language, occupation, household assets or incomes, and other community or regional variables. Household consumption (or income) data exist in one survey round but are missing in the other survey round, thus without loss of generality, let (survey) round 1 and round 2 respectively represent the survey round with and without household consumption data, and y_1 represent household consumption in round 1. Alternatively, we can also refer to round 1 as the base survey, and round 2 as the target survey.

To further operationalize our estimation, we assume that the linear projection of household consumption on household and other characteristics (x) in both survey rounds—if such consumption data were also available in period 2—are given by a cluster random-effects model[11]

$$y_j = \beta_j' x_j + \mu_j + \varepsilon_j \qquad (1)$$

---

[10] To make notation less cluttered, we suppress the subscript for each household in the following equations.

[11] This assumption assumes that the returns to the characteristics x_j in both periods are captured by equation (1) and precludes the (perhaps exceptionally) rare situations where there could be no correlation between these characteristics and household consumption due to unexpected upheavals in the economy or calamitous disasters. Contexts where there are sudden changes to the economic structures (e.g., overnight regime change) may also introduce noise into the comparability of the estimated parameters.

where β_j are the vector of coefficients, and the cluster random effects μ_j and the error term ε_j are assumed uncorrelated with each other and to follow a normal distribution, conditional on household characteristics. Equation (1) thus provides a standard linear random effects model that can be estimated using most available statistical packages. Let z_2 be the poverty line in period 2, if y_2 existed the (headcount) poverty rate P_2 in this period could be estimated with the following quantity

$$P(y_2 \leq z_2) \tag{2}$$

where P(.) is the probability (or poverty) function that gives the percentage of the population that are under the poverty line z_2 in round 2.

Assume that the sampled data in round 1 and round 2 are representative of the population in each respective time period, such that estimates based on the same characteristics x in these two survey rounds are consistent and comparable over time (Assumption 1). And assume further that given the estimated consumption parameters from round 1, the changes in the distributions of the explanatory variables x between the two periods can capture the change in poverty rate in the next period (Assumption 2).[12] Given these two assumptions, Dang et al. (2014b) propose an approach to impute the poverty rate for round 2, where the parameter estimate $\hat{\beta}_1$ and the distributions of the cluster random effects and the error term estimated from data in round 1 can be imposed on the data in round 2. Note that the standard errors of the imputation-based estimates can in fact be even smaller than that of the true (or design-based) rate if there is a good model fit (or the sample size in the target survey is larger than that in the base survey; see, e.g., Matloff, 1981).

---

[12] While this assumption may seem counterintuitive, it may be especially relevant to economies where the returns to characteristics do not change or simply change little over time (i.e. involving survey rounds that are implemented close in time, assuming the returns to characteristics in most economies do not normally change much within a short time interval).

If consumption data are available from both the base and target surveys, we can use an Oaxaca-Blinder type decomposition to formally test for Assumption 2 to shed further light on model selection. In particular, the change in poverty between the survey rounds can be broken down into two components, one due to the changes in the estimated coefficients (the first term in square brackets in equation (3) below) and the other the changes in the *x* characteristics (the second term in square brackets in equation (3) below). Assumption 2 would be satisfied if the poverty change is mostly explained by the latter component. This can be expressed as

$$P(y_2) - P(y_1) = \left[P(y_2) - P(y_2^1)\right] + \left[P(y_2^1) - P(y_1)\right] \qquad (3)$$

Furthermore, if we make a stricter assumption about the error term in equation (1) following a standard normal distribution, that is $\varepsilon_j \mid x_j \sim N(0,1)$, we can estimate equation (1) by a random effects probit model instead of the linear random effects model.

$$P(y_j) = \Phi(\beta_j' x_j + \mu_j + \varepsilon_j) \qquad (4)$$

But the standard modelling tradeoff holds: if our stricter assumption is correct, estimation results are more accurate and vice versa. For comparison purposes, we will later present estimates using both the linear random effects and random effects probit models.[13]

Following the estimation procedures in Dang et al. (2014b), our empirical implementation involves a two-stage process. First, we apply the estimated parameters from the 2004/05 round on the 2009/10 data to impute poverty for the latter. Since the questionnaires remain the same over these two survey rounds, their consumption data are comparable, and we can thus validate these estimated poverty rates against those based on the actual consumption data for the 2009/10 round.

---

[13] We provide a Stata ado program named "*povimp*" that automates the proposed estimation process (Dang and Nguyen, 2014). Type "*ssc install povimp*" from within Stata (StataCorp., 2013) to download this program from the SSC Archive, which is maintained by Christopher F. Baum at Boston College. Our Stata program automatically allows for complex survey designs by offering an option to specify the variables indicating the clusters and the strata.

Second, we produce imputation-based poverty estimates for 2011/12 using the same (model) specifications as with the first step, but with the estimated parameters from the 2009/10 round on the data from the 2011/12 round. Put differently, the first step would offer further validation that this imputation method works in the context of India, as well as provide the appropriate specification to use for the imputation; these two steps would satisfy the two assumptions discussed earlier.

III.2. Estimation Results

Since changes in household (heads') characteristics may indicate the corresponding changes in household consumption, it can be useful to examine as a preliminary check the distributions of household characteristics across the two survey rounds in 2009/10 and 2011/12. The summary statistics provided in Appendix 1, Table 1.2 show that these changes appear rather negligible with most of the differences being not statistically significant. Some characteristics that are associated with higher levels of household welfare (e.g. heads with completed post-graduate education, household members with regular salary incomes, or urban residents receiving regular wages) show a statistically significant improvement over time; but others that have opposite effects (e.g., backward classes and radio ownership) also have statistically significant changes.[14] The picture provided from considering the pairwise changes in the distributions of these variables over time thus seems mixed at best.

We then proceed to impute poverty for the target survey in 2009/10, using the estimated parameters from the base survey in 2004/05. Assumption 1 on survey comparability is satisfied since the questionnaires (and sample design) for these two survey rounds remain the same. To satisfy Assumption 2, we can then consider five different household consumption model

---

[14] We can infer the direction of the correlation between these characteristics and household consumption from the regression results in Appendix 1, Tables 1.3 and 1.4.

specifications where the changes in the distributions of the explanatory variables x between the two periods can capture to varying degrees the change in poverty over time. These specifications are built on a cumulative basis for comparison purposes (and robustness checks), with later specifications sequentially adding more variables to earlier specifications.

Specification 1 is the most parsimonious specification and consists of household size, household heads' age, gender, and dummy variables indicating whether the head is Hindu or Islam, whether the head belongs to a scheduled tribe, a scheduled caste or backward classes, whether the head is literate, and the head's education levels. Specification 2 adds to Specification 1 household demographics such as the shares of household members in the age ranges 0-14, 15-24, and 25-59 (with the reference group being those 60 years old and older). Specification 3 adds to Specification 2 employment variables, which include dummy variables indicating whether the household has any member working for a regular salary, whether the head is self-employed in the agricultural sector or the non-agricultural sector (for rural residents), and whether the head works for regular wage, is self-employed or engaged in casual work or other type of work (for urban residents). Specification 4 adds to Specification 3 a variable indicating home ownership. Finally, Specification 5 adds a more detailed list of asset variables, which include the energy sources for lighting and cooking, whether the household has a radio, television set, electric fan, sewing machine, freezer, air conditioner, bicycle, motorbike, and a car. However, slightly more than 5,000 and 1,000 households are missing these assets variables in the 2004/05 and 2009/10 rounds, respectively. Full model specifications and regression results are provided in Appendix 1, Table 1.3.

Estimation results using the linear random effects model shown in Table 1 (row 1) indicate that all the imputation-based poverty estimates in Specifications 1 to 4 fall within the 95 percent

13

confidence interval of the true poverty rate estimated from the actual consumption data for 2009/10. Put differently, these estimates are not statistically significantly different from the true poverty rate of 29.9 percent. The exception is Specification 5 where the imputation-based estimate is half a percentage point outside this confidence interval, which can be due to either model overfitting or smaller sample sizes with both the base and target surveys. Estimation results using the random effects probit model (row 2) are broadly similar, with estimates from Specifications 2 to 4 falling within the 95 percent confidence interval of the true poverty rate.

Thus for our purpose of finding a good model specification to impute poverty in the 2009/10 round, assuming consumption data in this round were not available, Specifications 1 to 4 with the normal linear regression models and Specifications 2 to 4 with the random effects probit models can all be employed. But among these specifications, our preferred specifications for interpretation are Specification 2 with the normal linear regressions and Specifications 3 and 4 with the random effects probit model since these three specifications provide better estimates that are within one standard error of the true rate.

It is useful to note that the standard errors for the imputation-based estimates are progressively smaller in the normal linear regression models and random effects probit models than that of the design-based poverty estimate. This is consistent with our earlier discussion since assuming the specification is correct, a good model fit can help bring down the standard errors. Similarly, the random effects probit models make a stricter (modelling) assumption on the error term than the linear random effects models, thus their standard errors are consequently smaller.

As a further check on the model specification, we provide a decomposition test of the changes in poverty due to the changes in the household characteristics and the estimated coefficients in Table 2 based on equation (3). (Note that we are now working with consumption data in both

surveys, rather than working with consumption data in only the base survey, as with the estimates for Table 1.) Estimation results confirm that, for the model specifications that provide estimates within the 95 percent confidence interval of the true poverty rates, the changes captured by the characteristics are closer to 100 percent. For example, under Specification 4 with the random effects probit regression, the change due to the coefficients is the most negligible; this specification also provides a point estimate of poverty (29.7 percent, Table 1) that is closest to the true poverty rate.

We turn next to impute poverty for 2011/12 with the estimated parameters from the 2009/10 survey round.[15] We have preferred specifications for analysis but we also show estimates for all the other specifications for comparison in Table 3. Our preferred specifications show that the imputation-based poverty estimates can be range from 22.9 percent (Specifications 3 and 4, the probit model) to 25 percent (Specification 2, the linear regression model). Interestingly enough, except for Specification 5 that could be excluded due to overfitting concerns, all other estimates— including even Specification 5 with the probit model—fall within this range.

These imputation-based estimates are larger than the design-based estimates of 22 percent, and the differences are statistically significant (outside the 95 percent confidence of the latter). However, considering all specifications together, the difference between the probit estimates and the design-based estimate is between one and two percentage points, while that between the normal linear regression estimates and the design-based estimates is between two and three percentage points. Thus according to our imputation-based estimates, while the design-based estimate may

---

[15] We use estimated parameters from the 2009/10 round, rather than the 2004/05 round, to impute poverty in the 2011/12 round since these parameters may change over time. Indeed, the null hypothesis of the equality of the estimated parameters in these two survey rounds is rejected with significantly large value from a Wald test (results available upon request). More generally, survey rounds that are closer in time are more appropriate for imputation.

underestimate poverty in 2011/12, it appears that this underestimation may in practice be not that large.

IV. Constructing Synthetic Panels

Our findings in the previous section suggest that the sharp decrease in poverty rate between 2009/10 and 2011/12 is reasonably captured by the 66th and 68th rounds of the NSSs. Put differently, these two survey rounds provide comparable consumption data for most practical poverty measurement purposes, which is a prerequisite for constructing synthetic panel data. We next provide a brief overview of the methods that will be used.

IV.1. Overview of the Synthetic Panel and Vulnerability Analysis Methods[16]

Let x_{ij} be a vector of household characteristics observed in survey round j (j= 1 or 2) that are also observed in the other survey round for household i, i= 1,…, N. These household characteristics include variables that may be collected in only one survey round, but whose values can be inferred for the other round. These variables may be roughly categorized in three types i) time-invariant variables such as ethnicity, religion, place of birth, or parental education; ii) deterministic variables such as age (which given the value in one survey round can then be determined given the time interval between the two survey rounds),[17] and ii) time-varying household characteristics if retrospective questions about the values of such characteristics in the first survey round are asked in the second round.

---

[16] We provide an overview of the methods that construct synthetic panels and vulnerability lines developed by Dang et al. (2014a) and Dang and Lanjouw (2013, 2014) in this section. For more details, interested readers are encouraged to read the original papers.

[17] To reduce spurious changes due to changes in household composition over time, we usually restrict the estimation samples to household heads age, say 25 to 55 in the first cross section and adjust this age range accordingly in the second cross section. This restriction also helps ensure certain variables such as heads' education attainment remains relatively stable over time (assuming most heads are finished with their schooling). This age range is usually used in traditional pseudo-panel analysis but can vary depending on the cultural and economic factors in each specific setting.

Then let y_{ij} represent household consumption or income in survey round j, j= 1 or 2. The linear projection of household consumption (or income) on household characteristics for each survey round is given by

$$y_{ij} = \beta_j' x_{ij} + \varepsilon_{ij} \tag{5}$$

Let z_j be the poverty line in period j. We are interested in knowing such quantities as

$$P(y_{i1} < z_1 \ and \ y_{i2} > z_2) \tag{6a}$$

which represents the percentage of households that are poor in the first period but nonpoor in the second period, or

$$P(y_{i2} > z_2 \mid y_{i1} < z_1) \tag{6b}$$

which represents the percentage of poor households in the first period that escape poverty in the second period. In other words, for the average household, quantity (6a) provides the joint (unconditional) probabilities of household poverty status in both periods, and quantity (6b) the conditional probabilities of household poverty status in the second period given their poverty status in the first period.

If true panel data are available, we can straightforwardly estimate the quantities in (6a) and (6b); but in the absence of such data, we can use synthetic panels to study mobility. To operationalize the framework, we make two standard assumptions. First, we assume that the underlying population being sampled in survey rounds 1 and 2 are identical such that their time-invariant characteristics remain the same over time. More specifically, coupled with equation (5), this implies the conditional distribution of expenditure in a given period is identical whether it is conditional on the given household characteristics in period 1 or period 2 (i.e., $x_{i1} = x_{i2}$ implies $y_{i1}|x_{i1}$ and $y_{i1}|x_{i2}$ have identical distributions). Second, we assume that ε_{i1} and ε_{i2} have a bivariate

normal distribution with positive correlation coefficient ρ and standard deviations σ_{ϵ_1} and σ_{ϵ_2} respectively. Quantity (6a) can be estimated by

$$P(y_{i1} < z_1 \text{ and } y_{i2} > z_2) = \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, -\frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) \qquad (7)$$

where $\Phi_2(.)$ stands for the bivariate normal cumulative distribution function (cdf)) (and $\phi_2(.)$ stands for the bivariate normal probability density function (pdf)). In equality (7), the parameters β_j and σ_{ε_j} are estimated from equation (5), and ρ can be estimated using an approximation of the birth-cohort-aggregated household consumption between the two surveys. Note that in equality (7), the estimated parameters obtained from data in both survey rounds are applied to data from the second survey round (x_2) (or the base year) for prediction, but we can use data from the first survey round as the base year as well. It is then straightforward to estimate quantity (6b) by

dividing quantity (6a) by $\Phi\left(\dfrac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}\right)$, where $\Phi(.)$ stands for the univariate normal cumulative

distribution function (cdf).[18]

Using the given poverty lines z_j, quantities (6a) and (6b) classify the population into two groups, one is poor and the other nonpoor. But we can obtain richer analysis by further disaggregating the nonpoor group into two additional groups: the vulnerable (those that are nonpoor but still face a significant risk of falling into poverty) and the middle class (the remaining group with higher consumption levels). The dividing vulnerability line v_j that separates these groups can be derived from a specified vulnerability index $\mathcal{P}$, which is defined as the percentage of the non-poor population in the first period that fall into poverty in the second period. This

---

[18] Further asymptotic results and formulae for the standard errors are provided in Dang and Lanjouw (2013).

vulnerability index can be anchored to, say, social protection targets, within the bounds given by the data (Dang and Lanjouw, 2014). We will further discuss the vulnerability line in the next section.

Given v_j, we can extend expression (6a) to analyze the dynamics for these three categories: poor, vulnerable, and middle class. For example, the percentage of poor households in the first period that escape poverty but still remain vulnerable in the second period (joint probability) is

$$P(y_{i1} < z_1 \ and \ z_2 < y_{i2} < v_2) = \Phi_2\left(\frac{z_1 - \beta_1'x_{i2}}{\sigma_{\varepsilon_1}}, \frac{v_2 - \beta_2'x_{i2}}{\sigma_{\varepsilon_2}}, \rho \right) - \Phi_2\left(\frac{z_1 - \beta_1'x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2'x_{i2}}{\sigma_{\varepsilon_2}}, \rho \right)$$

(8)

IV.2. Setting Vulnerability Lines
Table 4 shows a range of values of the vulnerability line that correspond to different vulnerability indexes for the two periods 2004-09 and 2009-11. The vulnerability index falls within the range [17, 34] for the first period, but this range both narrows and decreases to [15, 29] for the second period, suggesting that the population as a whole are better off in the latter. Put differently, at the same vulnerability index of, say, 20 percent in both periods, the vulnerability line is higher at 1,035 rupees per month in the first period, but shrinks by 20 percent to 830 rupees per month in the second period.[19]

Do we also have lower vulnerability for the longer period 2004-11? In general, this is an empirical question since a longer period is likely to be associated with larger vulnerability indexes unless household consumption grows so fast that it can offset this trend (Dang and Lanjouw, 2014). Figure 2 then plots the vulnerability line against the vulnerability index for the two periods 2004-

---

[19] All numbers are converted to 2004 price for all rural India. Note that we provide more detailed estimation results for India for the period 2004-09 in this paper than those in our other paper (Dang and Lanjouw, 2014). Our estimates are also different from those in the latter, which deflate all numbers to a population-weighted monthly national poverty line instead.

09 and 2009-11 on the same graph, and adds that for the period 2004-11 as well.[20] The curve for the period 2009-11 lies everywhere below that for the period 2004-09 and is closer to the origin, which provides a graphical illustration of lower vulnerability in the latter period as discussed above. The curve for the period 2004-11 lies below both those for the other two periods, even though far more so compared with that for the period 2009-11, thus indicating that vulnerability is lowest for this period.

What is then an appropriate vulnerability line to use? A common, but rather *ad hoc*, approach is to arbitrarily scale up the poverty line by a certain factor to obtain the vulnerability line. In particular, vulnerability has been defined as simply occurring within a fixed income range between 1.25 times and twice the national poverty line in India (NCEUS, 2007). Other countries similarly define the vulnerability line as twice (Pakistan; Lopez-Calix et al., 2014) or 30 percent above the national poverty line (Vietnam; World Bank, 2012). This approach has the advantage of being simple and easy to understand, but appears to be based on no underlying welfare theoretical framework.

The recent approach proposed in Dang and Lanjouw (2014) instead derives the vulnerability line from a specified vulnerability index $\mathcal{P}$ in the spirit of vulnerability to poverty.[21] This vulnerability index can in turn be obtained in several different ways. One way is to identify the percentage of the population that can be supported with the available social transfer budget which can provide a guideline to the associated level of vulnerability.[22] Another way is to consider the

---

[20] As discussed earlier, we consider the cohorts that range from 25 years old in the first year in each period, which is 2004 for the periods 2004-09 and 2004-11, and 2009 for the period 2009-11. While we may also adjust the age range so that the cohorts studied for the period 2009-11 are the same as those for the period 2004-09 (i.e., considering the cohorts that range from 30 years old in 2009), this setup is more natural if we are to compare mobility in these two periods (since it keeps ages and other associated characteristics fixed for each periods).

[21] See Hoddinott and Quisumbing (2010) for a recent review of approaches to measuring vulnerability.

[22] For a (very) simplistic example, assume that the available social transfer budget can help, say, 30 percent of the vulnerable population from falling into poverty, we can then identify the vulnerability index and vulnerability line

highest risk of falling into poverty, say, 15 percent, that is deemed socially acceptable; another is

to put forward a social protection target that aims to reduce vulnerability to below a certain level

(which can be similar to the recent goal of reducing the global headcount poverty to 3 percent or

less by 2030 proposed by the World Bank). We will employ a vulnerability index of 20 percent

and the associated vulnerability line for our welfare analysis in the next section, but we will also

provide, for comparison purposes, some estimates that are based on twice the national poverty line

(i.e., 893.4 rupees per month in 2004 price for all rural India) as with the existing practice in the

country. Table 4 indicates that doubling the national poverty line would roughly translate into a

vulnerability index of between 21 and 22 percent for 2004-2009, and between 19 and 20 percent

for 2009-2011.

V. Welfare Dynamics Analysis

We have discussed the changes in poverty over time in the previous section, thus will focus on

discussing the other dynamics with vulnerability in this section. We start first with showing the

welfare transitions for all the population before delving further into population groups.

V.1. All Population

The welfare transition matrixes for the three consumption groups for the two periods 2004-09

and 2009-11 are respectively shown in Panel A and Panel B in Table 5, where the vulnerability

index is fixed at 20 percent for both periods.[23] Together with the decrease in poverty, there is an

expansion of the vulnerable and the middle class (categories) in the period 2004-09. This trend

---

associated with this figure. For India during the period 2004-09, the associated vulnerability index and vulnerability
line would be respectively 26 percent and 675 rupees per month (Table 4).

[23] As noted earlier, we restrict the data to households whose head's age is between 25 and 55 in the first survey round
and adjust accordingly for the second survey round (e.g., age ranges 25-55 for 2004/05 and 30- 60 for 2009/10 in the
period 2004-09) to keep household units stable. This results in some slight differences with poverty rates based on
this data compared to the full data. We are using the first and second survey rounds respectively as the base and target
surveys for constructing the synthetic panels. For these reasons, the estimated poverty rate for 2011/12 slightly change
from 23.7 percent (Tables 5 and 6) to 25 percent (Table 7) below. This slight difference appears not very large in
practice and is consistent with the imputation-based estimates of poverty in section III.

seems to continue in the second period 2009-11, but with a faster shrinkage of the poor and growth of the middle class and almost no change to the vulnerable. Specifically, the fall in poverty rises from 14 percent (i.e., =1- (31.9/36.9)) during the first period to 22 percent during the second period, while the middle class growth increases from 24 percent to 28 percent over the same time interval. In terms of absolute numbers, the vulnerable category decrease by roughly ten percentage points and shrink from making up more than half of the population in 2004-09 to less than half of the population in 2009-11; the middle class is two and a half times as large in the latter compared to the former (e.g., =30.6/12.2).

Another useful way to gauge welfare mobility in the two periods is to look at the percentage of the population that change their welfare status over time. In 2004-09, 23 percent of the population move up one or two welfare categories (i.e., the sum of the upper off-diagonal cells) while 17 percent move down one or two welfare categories (i.e., the sum of the lower off-diagonal cells). The corresponding figures in 2009-11 are larger and respectively 30 percent and 19 percent, suggesting that the population as a whole are both better off and more mobile in this period.

For further comparison, we also provide analysis with the vulnerability line being twice the poverty line and show estimation results in Table 1.5, Appendix 1. Estimates are unsurprisingly rather similar for the period 2009-11, since a vulnerability line equal to twice the poverty line results in a vulnerability index between 19 and 20 percent in this period. The results for the period 2004-09 somewhat vary more, but appear qualitatively similar as well. For example, 26 percent of the population move up one or two welfare categories and 18 percent move down one or two welfare categories in this period.

Since vulnerability may change over time, the different lengths of time in the period 2004-09 and 2009-11 may not provide perfectly comparable comparison. Offering another angle at studying

mobility in these periods, we still fix the vulnerability index at 20 percent in the first period, but use its associated vulnerability line of 1,035 rupees per month as the vulnerability line in the second period. In other words, once the vulnerability index is given in the first period, we hold constant the vulnerability line in both periods.[24]

Estimation results are shown in Table 6, where we keep Panel A the same as with Table 5 for convenience and provide the new estimates using the same vulnerability line in Panel B. There is also a similar shrinkage of the poor and growth of the middle class in 2009-11 as seen before, even though the middle class growth now speeds up from 24 percent in 2004-09 to 36 percent in 2009-11. Another difference is that the vulnerable now expands in the second period, but at a slower rate (3 percent) than that (5 percent) in the first period. In terms of absolute numbers, the vulnerable are slightly larger but the middle class is around 35 percent to half a time larger in the second period. With respect to mobility, the population as a whole are still better off and more mobile in this period, with 27 percent and 17 percent of the population moving up and down one or two welfare categories respectively. These results are qualitatively similar with those obtained from keeping fixed the vulnerability index at 20 percent for both periods as shown in Table 5.

We now turn to looking at the welfare transition over the longest interval 2004-11, and provide estimation results in Table 7. For comparison purposes, we show results for two approaches with constructing the vulnerability line where Panel A provides estimates based on the vulnerability line associated with a vulnerability index of 20 percent, and Panel B shows estimates for the vulnerability line being equal to twice the national poverty line. The vulnerable category is smaller while the middle class is larger with the first approach, but both the vulnerable and the middle

---

[24] For a graphical illustration, keeping fixed the vulnerability index for one period corresponds to drawing a vertical line in Figure 2 at the specified index, while keeping fixed the vulnerability line corresponds to drawing a horizontal line at the specified line for this period. The intersections of such lines with the curve for the other period are likely to differ, which creates different results depending on whether the vulnerability index or the vulnerability line is fixed.

class grow faster with the second approach. In particular, while the vulnerable remain almost the same in Panel A, this category grows by 6 percent in Panel B, and the middle class expansion climbs up from 50 percent in Panel A to 60 percent in Panel B. Mobility is rather similar, however, with both approaches: roughly 35 percent of population move up one or two welfare categories and 15 percent move down one or two welfare categories.

V.2. Profiling of Population Groups

Keeping the vulnerability fixed at 20 percent, Figure 3 plots the percentage of the poor or vulnerable in the first year that move up one or two welfare categories in the second year in the two periods 2004-09 and 2009-11.[25] The transitions are disaggregated by education levels (i.e., less than primary education, primary education, middle education, secondary education, and college), occupation (which is further broken down into two categories of residence: i) rural areas: self-employment in non-agriculture, agricultural labor, other labor categories, self-employment in agriculture, remaining categories, and ii) urban areas: self-employment, wage workers, and remaining categories), and socio-ethnic groups (i.e., scheduled tribe, scheduled caste, other backward groups, and remaining groups).[26]

Two remarks are in order for Figure 3. First, more education achievement, urban residence, wage work, and belonging to social groups other than the scheduled or backward groups are positively associated with higher-than-average chances of upward mobility. For example, these results are shown for the period 2009-11 with the orange dots representing these probabilities lying above the orange dashed line that represents the national average. Second, the period 2004-09

---

[25] We show the conditional, rather than the joint, probabilities for Figures 3 to 6 since this provides larger numbers that help bring out more clearly the transition patterns for the different population groups. For example, a small percentage of the population with secondary or higher education are usually found in poverty or vulnerability in the first period to start with, consequently their transitions to higher income categories are smaller.

[26] An additional assumption required for producing these graphs is that mobility for each population group/ profile should generally follow that for the whole population.

shows qualitatively similar albeit weaker mobility than in the period 2009-11. For example, ceteris paribus, jumping from a middle education to a secondary education is associated with having one percentage point higher for upward mobility in the first period, but as large as having six percentage points higher for upward mobility in the second period. This generally concurs with our earlier findings that the period 2009-11 exhibits more mobility than the period 2004-09.

Figure 4 presents a similar graph where upward mobility is disaggregated at the state level, where, for better presentation purposes, states' mobility in the period 2009-2011 is ranked in an ascending order. While this figure indicates that certain states maintain a similar level of performance in both periods (e.g., Chandigarh and Delhi are strong performers but Lakshadweep and Dadra & Nagar Haveli are weak performers), this may change over time. For example, such states as Meghalaya and Arunachal Pradesh are strong performers in the first period but become weak performers in the second period.

Factors that are positively correlated with upward mobility are in general related to those associated with escaping downward mobility, but this may not always hold (see, e.g., Dang, Lanjouw, and Swinkels (forthcoming) for an analysis of mobility in Senegal). We thus produce two figures for downward mobility for the same population groups (Figures 5 and 6). Interestingly, for India it is generally true that the same factor can be associated with both increasing upward mobility and decreasing downward mobility. For example, out of all occupational categories, wage workers living in urban areas have the largest and smallest chance of upward mobility and downward mobility, respectively.[27]

VI. Conclusion

---

[27] But also note that the ranking of states in terms of mobility change slightly between the two periods. For example, Delhi and Chandigarh switch their place from Figure 4 (upward mobility) to Figure 6 (downward mobility).

We investigate in this paper the poverty and vulnerability dynamics in India between 2004/05 and 20011/12 using three rounds of the NSSs. In the absence of actual panel data, we construct synthetic panels using statistical methods that were recently developed by Dang et al. (2014a) and Dang and Lanjouw (2013). We present analysis using vulnerability lines that correspond to a vulnerability index of 20 percent and that are also close to twice the national poverty line.

Estimation results point to faster poverty reduction and more upward mobility in the period 2009-11 than in the period 2004-09. In particular, while the vulnerable category makes up more than half of the population in 2004-09, it accounts for less than half of the population in 2009-11, and the middle class grows two and a half times as large in the second period. While 23 percent and 17 percent of the population, respectively, experience upward and downward mobility in the first period, the corresponding figures in the second period are larger at respectively 30 percent and 19 percent. This pattern of stronger upward mobility is qualitatively similar when considered over the longer period 2004-11.

Factors including more educational achievement, urban residence, wage work, and belonging to socio-ethnic groups other than the scheduled or backward groups are positively associated with higher-than-average chances of upward mobility and lower-than-average changes of downward mobility.

Our paper also presents a two-step analysis procedure where careful checks should be done in the first step to ensure data comparability across survey rounds before synthetic panels can be constructed in the second step. This procedure may be relevant to quite a few other contexts since situations where data are not comparable across survey rounds—leading to, for example, the recent debate on poverty decline in 2011/12 in India—appear to occur more frequently than one might think. We discuss a statistical method (Dang et al., 2014b) that can be employed for this checking

purpose. Estimation results show that the poverty decline between 2009/10 and 2011/12 is not severely over-estimated (or equivalently, the design-based poverty estimate using the 2011/12 survey round is practically comparable to those from previous rounds).

Our methods are promising of richer analysis for welfare dynamics that can further exploit the richness of the NSS data. For example, future research can provide more disaggregated analysis within each state, and analyze either more survey rounds to study transition trajectories between more than two periods or survey rounds that are farther apart to investigate longer-term transitions. Another direction is to make better use of the "thin" rounds, in addition to the "thick" rounds, to build a more comprehensive picture of these dynamics over time.

References

Beegle, Kathleen, Joachim De Weerdt, Jed Friedman, and John Gibson. (2012). "Methods of Household Consumption Measurement through Surveys: Experimental Results from Tanzania". *Journal of Development Economics*, 98(1): 3-18.

Bierbaum, Mira and Franziska Gassmann. (2012). "Chronic and transitory poverty in the Kyrgyz Republic: What can synthetic panels tell us?" *UNU-MERIT Working Paper #2012-064.*

Cancho, Author César, María E. Dávalos, Giorgia Demarchi, Moritz Meyer, and Carolina Sánchez Páramo. (2015)."Economic Mobility in Europe and Central Asia: Exploring Patterns and Uncovering Puzzles". *World Bank Policy Research Paper No. 7173.*

Christiaensen, Luc, Peter Lanjouw, Jill Luoto, and David Stifel. (2012). "Small Area Estimation-based Prediction Models to Track Poverty: Validation and Applications." *Journal of Economic Inequality*, 10 (2): 267-297.

Crossley, Thomas F. and Joachim K. Winter. (forthcoming). "Asking Households About Expenditures: What Have We Learned?" in Carroll, C., T. F. Crossley and J. Sabelhaus. (Eds.). *Improving the Measurement of Consumer Expenditures. Studies in Income and Wealth*, Volume 74. Chicago: University of Chicago Press.

Cruces, Guillermo, Peter Lanjouw, Leonardo Lucchetti, Elizaveta Perova, Renos Vakis, and Mariana Viollaz. (2014). "Estimating Poverty Transitions Repeated Cross-Sections: A Three-country Validation Exercise". *Journal of Economic Inequality*, DOI 10.1007/s10888-014-9284-9.

Dang, Hai-Anh and Peter Lanjouw. (2013). "Measuring Poverty Dynamics with Synthetic Panels Based on Cross-Sections". *World Bank Policy Research Working Paper No. 6504.*

---. (2014). "Welfare Dynamics Measurement: Two Definitions of a Vulnerability Line and Their Empirical Application". *World Bank Policy Research Paper No. 6944.*

Dang, Hai-Anh and Minh Cong Nguyen. (2014). "POVIMP: Stata Module to Provide Poverty Estimates in the Absence of Actual Consumption Data." *Statistical Software Components S457934*. Boston College, Department of Economics.

Dang, Hai-Anh, Peter Lanjouw, Jill Luoto, and David McKenzie. (2014a). "Using Repeated Cross-Sections to Explore Movements in and out of Poverty". *Journal of Development Economics*, 107: 112-128.

Dang, Hai-Anh, Peter Lanjouw, Umar Serajuddin. (2014b). "Updating Poverty Estimates at Frequent Intervals in the Absence of Consumption Data: Methods and Illustration with Reference to a Middle-Income Country." Policy Research Working Paper No.7043. Washington DC: The World Bank.

Dang, Hai-Anh, Peter Lanjouw, and Rob Swinkels. (forthcoming). "Who Remained in Poverty, Who Moved Up, and Who Fell Down? An Investigation of Poverty Dynamics in Senegal in the Late 2000s." In Machiko Nissanke and Muna Ndulo. (Eds). *Festschrift for Erick Thorbecke*.

Datt, Gaurav and Martin Ravallion. (2011). "Has India's Economic Growth Become More Pro-Poor in the Wake of Economic Reforms?" *World Bank Economic Review*, 25(2): 157-189.

Deaton, Angus and Margaret Grosh. (2000). "Consumption". In Margaret Grosh and Paul Glewwe. (Eds.), *Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of LSMS Experience*. Washington, DC: The World Bank.

Deaton, Angus and Jean Dreze. (2002). "Poverty and Inequality in India: A Re-Examination". *Economic and Political Weekly,* 37(36): 3729-3748.

Deaton, Angus and Valerie Kozel. (2005). *The Great Indian Poverty Debate*. New Delhi: Macmillan.

Dercon, Stefan, and Joseph S. Shapiro. (2007). "Moving On, Staying Behind, Getting Lost: Lessons on Poverty Mobility from Longitudinal Data." In D. Narayan and P. Petesch. (Eds.) *Moving Out of Poverty: Cross-Disciplinary Perspectives*. Washington, DC: World Bank.

Dercon, Stefan, Pramila Krishnan, and Sofya Krutikova. (2013). "Changing Living Standards in Southern Indian Villages 1975–2006: Revisiting the ICRISAT Village Level Studies". *Journal of Development Studies*, 49(12): 1676-1693.

Douidich, Mohamed, Abdeljaouad Ezzrari, Roy van der Weide, and Paolo Verme. (2013). "Estimating Quarterly Poverty Rates Using Labor Force Surveys: A Primer." Policy Research Working Paper No. 6466. Washington DC: The World Bank.

Dutta, Bhaskar and Manoj Panda. (2014). "Social Welfare and Household Consumption Expenditure in India – 2004-05 to 2011-12". *Economics and Political Weekly*, XLIX(31): 113-122.

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

Ferreira, Francisco H. G., Julian Messina, Jamele Rigolini, Luis-Felipe López-Calva, Luis Felipe López-Calva, and Renos Vakis. (2012). *Economic Mobility and the Rise of the Latin American Middle Class*. Washington DC: World Bank.

Garbero, Alessandra. (2014). "Estimating poverty dynamics using synthetic panels for IFAD-supported projects: a case study from Vietnam". *Journal of Development Effectiveness*, 6(4): 490-510.

Gibson, John, Jikun Huang, and Scott Rozelle. (2003). "Improving Estimates of Inequality and Poverty from Urban China's Household Income and Expenditure Survey". *Review of Income and Wealth*, 49(1): 53–68.

Gupta, Monobina. (2013). "How Not to Talk About Poverty in India". Accessed on the Internet in January 2015 at the New York Times http://india.blogs.nytimes.com/2013/08/01/how-not-to-talk-about-poverty-in-india/?_r=0

Hoddinott, John and Agnes R. Quisumbing. (2010). "Methods for Microeconometric Risk and Vulnerability Assessment". In Ricardo Fuentes-Nieva and Papa A. Seck. (Eds). *Risk, Vulnerability and Human Development: On the Brink.* London: Palgrave Macmillan- United Nations Development Programme.

Kijima, Yoko and Peter Lanjouw. (2003). "Poverty in India during the 1990s: A Regional Perspective." Policy Research Working Paper No. 3141. Washington DC: The World Bank.

Krishna Arinudh and Abusaleh Shariff. (2011). "The Irrelevance of National Strategies? Rural Poverty Dynamics in States and Regions of India, 1993–2005". *World Development,* 39(4): 533–549.

Lakshmi, Rama. (2012). "Is India misrepresenting its poverty numbers". Accessed on the Internet in January 2015 at the Washington Post http://www.washingtonpost.com/blogs/worldviews/post/is-india-misrepresenting-its-poverty-numbers/2012/03/20/gIQA1dtXPS_blog.html

Limaye, Yogita. (2013). "Debate as figures suggest decrease in Indian poverty". Accessed on the Internet in January 2015 at the BBC http://www.bbc.com/news/business-23547479

Lopez-Calix, Jose, Carolina Mejia, David Newhouse and Carlos Sobrado. (2014). Pakistan Poverty Trends, Scenarios and Drivers. *Policy Paper Series on Pakistan, PK 23/12*. Washington DC: World Bank.

Martinez, Arturo Jr., Mark Western, Michele Haynes, Wojtek Tomaszewski. (2013). "Measuring Income Mobility Using Pseudo-Panel Data". *Philippine Statistician*, 62(2): 71-99.

Mathiassen, Astrid. (2013). "Testing Prediction Performance of Poverty Models: Empirical Evidence from Uganda". *Review of Income and Wealth*, 59(1): 91–112.

Matloff, Norman S. (1981). "Use of Regression Functions for Improved Estimation of Means". *Biometrika*, 68(3): 685-689.

Munshi, Kaivan and Mark Rosenzweig. (2009). "Why Is Mobility in India So Low? Social Insurance, Inequality, and Growth". NBER Working Paper 14850.

National Commision for Enterprises in the Unorganised Sector (NCEUS). (2007). *Report on Conditions of Work and Promotion of Livelihoods in the Unorganised Sector*. New Delhi, India.

Rama, Martin, Tara Béteille, Yue Li, Pradeep K. Mitra, and John Lincoln Newman. (2015). *Addressing Inequality in South Asia*. Washington, DC: World Bank.

Rao, J. N. K. (2003). *Small Area Estimation*. New Jersey: Wiley.

Rao, Kirthi. (2013). "India's poverty level falls to record 22%: Planning Commission". Accessed on the Internet in January 2015 at http://www.livemint.com/Politics/1QvbdGnGySHo7WRq1NBFNL/Poverty-rate-down-to-22-Plan-panel.html

Ravallion, Martin. (2011). "A Comparative Perspective on Poverty Reduction in Brazil, China, and India." *World Bank Research Observer*, 26 (1): 71-104.

Rosenzweig, Mark R. (2003). "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility". *American Economic Review*, 93(2): 112-117.

Saxena, Naresh Chandra. (2013). "A poor poverty debate in India". Accessed on the Internet in January 2015 at the Business Standard http://www.business-standard.com/article/opinion/a-poor-debate-on-poverty-in-india-113072501165_1.html

Sen, Abhijit and Himanshu. (2005). "Poverty and Inequality in India". In Angus Deaton and Valerie Kozel. (Eds). *The Great Indian Poverty Debate*. New Delhi: Macmillan.

StataCorp. (2013). *Stata: Release 13. Statistical Software*. College Station, TX: StataCorp LP.

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

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

Vakis, Renos, James Rigolini, and Leonardo Lucchetti. (2015). *Left Behind: Chronic Poverty in Latin America and the Caribbean*. Washington, DC: World Bank.

World Bank. (2012). "Well Begun, Not Yet Done: Vietnam's Remarkable Progress on Poverty Reduction and the Emerging Challenges". *Vietnam Poverty Assessment Report 2012*. Hanoi: World Bank.

---. (2014). *Bhutan Poverty Assessment Report*. Poverty Reduction and Economic Management, South Asia Region. Washington DC: World Bank Group.

---. (2015). *World Development Indicators Database*. Washington DC: World Bank.

Table 1: Predicted Poverty Rates Based on Imputation, India 2009/10 (percentage)

Method	Estimated rate					True rate
	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5	
1) Normal linear regression model	29.3	29.6	30.4	30.6	31.2	29.9
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.4)
2) Direct estimation of poverty rate using probit model	28.9	29.2	29.5	29.7	28.2	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
Control variables						
Parsimonious	Y	Y	Y	Y	Y	
Demographics	N	Y	Y	Y	Y	
Employment	N	N	Y	Y	Y	
Owning home	N	N	N	Y	Y	
Household assets	N	N	N	N	Y	
N (base survey, 2004/05)	124,543	124,543	124,374	124,340	119,292	
N (target survey, 2009/10)	100,832	100,832	100,798	100,595	99,469	100,853

Note: Standard errors in parentheses are adjusted for complex survey design. All estimates are obtained with population weights. Method 1 uses the normal linear regression model with the theoretical distribution of the error terms and Method 2 uses a probit regression. Both specifications use state random effects. Imputed poverty rates for 2009/10 use the estimated parameters based on the 2004/05 data, with 1,000 simulations. The underlying regression results are provided in Appendix 1, Table 1.3. True rate is the direct estimate based on survey data.

Table 2: Decomposition of Changes in Poverty over Time, India 2004/05- 2009/10 (percentage)

Method	Estimated rate				
	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5
A. Using theoretical distribution of the error terms					
Due to characteristics	108.0	103.9	93.2	90.5	83.8
Due to coefficients	-8.0	-3.9	6.8	9.5	16.2
Total	100	100	100	100	100
B. Direct estimation of poverty rate using probit model					
Due to characteristics	113.3	109.3	104.6	103.2	121.9
Due to coefficients	-13.3	-9.3	-4.6	-3.2	-21.9
Total	100	100	100	100	100
Control variables					
Parsimonious	Y	Y	Y	Y	Y
Demographics	N	Y	Y	Y	Y
Employment	N	N	Y	Y	Y
Owning home	N	N	N	Y	Y
Household assets	N	N	N	N	Y
N	100,832	100,832	100,798	100,595	99,469

Note: The decomposition of the changes in poverty for Panels A and B are implemented using the estimation results in Table 1. The underlying regression results are provided in Appendix 1, Table 1.3.

Table 3: Predicted Poverty Rates Based on Imputation, India 2011/12 (percentage)

Method	Estimated rate					True rate
	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5	
1) Normal linear regression model	24.4	25.0	24.3	24.5	27.1	22.0
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.3)
2) Direct estimation of poverty rate using probit model	23.7	24.3	22.8	22.9	24.1	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
Control variables						
Parsimonious	Y	Y	Y	Y	Y	
Demographics	N	Y	Y	Y	Y	
Employment	N	N	Y	Y	Y	
Owning home	N	N	N	Y	Y	
Household assets	N	N	N	N	Y	
N (base survey, 2009/10)	100,832	100,832	100,798	100,595	99,469	
N (target survey, 2011/12)	101,639	101,639	101,603	101,596	101,525	101,662
Note: Standard errors in parentheses are adjusted for complex survey design. All estimates are obtained with population weights. Method 1 uses the normal linear regression model with the theoretical distribution of the error terms and Method 2 uses a probit regression. Both models use state random effects. Imputed poverty rates for 2011/12 use the estimated model parameters based on the 2009/10 data, with 1,000 simulations. The underlying regression results are provided in Appendix 1, Table 1.4. True rate is the direct estimate based on survey data.						

Table 4: Setting Vulnerability Lines at Different Values of Vulnerability Indexes, India 2004/05- 2011/12

No	Vulnerability index (%)	2004/05- 2009/10			2009/10- 2011/12		
		Vulnerability line (rupee)	Increase (%)	Pop. share with consumption above poverty line but less than V-line in first period (%)	Vulnerability line (rupee)	Increase (%)	Pop. share with consumption above poverty line but less than V-line in first period (%)
1	34	470	5	3.5	N/A	N/A	N/A
2	33	495	11	7.3	N/A	N/A	N/A
3	32	510	14	9.4	N/A	N/A	N/A
4	31	525	18	11.6	N/A	N/A	N/A
5	30	550	23	15.0	N/A	N/A	N/A
6	29	575	29	18.3	460	3	1.7
7	28	605	35	22.1	495	11	6.5
8	27	635	42	25.6	510	14	8.5
9	26	675	51	30.0	530	19	11.1
10	25	715	60	33.8	565	26	15.6
11	24	750	68	37.0	610	37	21.2
12	23	805	80	41.3	660	48	26.9
13	22	875	96	45.9	705	58	31.6
14	21	950	113	50.0	770	72	37.6
15	20	1035	132	53.5	830	86	42.5
16	19	1175	163	57.8	915	105	48.4
17	18	1410	216	62.2	1025	129	54.4
18	17	1945	335	66.1	1165	161	59.9
19	16	N/A	N/A	N/A	1390	211	65.6
20	15	N/A	N/A	N/A	1905	326	71.3

Note: Vulnerability lines are in 2004 price for all rural India, which can be inflated to those in 2009 and 2011 respectively with a scale factor of 1.51 and 1.83. Relative increases of the vulnerability line from the poverty line for each country is shown under the columns "Increase" (columns 4 and 7). All numbers are estimated with synthetic panel data and weighted with population weights. Household head's age range is restricted to between 25 and 55 for the first survey and adjusted accordingly for the second survey in each period. Estimation sample sizes are 91,751 and 76,479 households for the first period, and 73,681 and 75,159 households for the second period. The incremental value for iteration is five rupees. The exchange rate is US$1 for 45.3 rupees in 2004 (World Bank, 2015).

Table 5: Welfare Transition Dynamics Based on Synthetic Panel Data at Similar Vulnerability Index, India 2004/05- 2011/12 (percentage)

Panel A: Vulnerability line corresponding to V-index= 0.2		2009			
		Poor	**Vulnerable**	**Middle class**	**Total**
2004	**Poor**	20.0	16.2	0.6	36.9
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	11.6	35.1	6.6	53.3
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	0.3	4.6	5.0	9.8
		(0.0)	(0.0)	(0.0)	(0.1)
	Total	31.9	55.9	12.2	100
		(0.1)	(0.0)	(0.1)	
Panel B: Vulnerability line corresponding to V-index= 0.2		2011			
		Poor	**Vulnerable**	**Middle class**	**Total**
2009	**Poor**	12.8	14.0	3.4	30.2
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	9.5	23.5	12.9	45.9
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	1.5	8.2	14.3	23.9
		(0.0)	(0.0)	(0.1)	(0.1)
	Total	23.7	45.7	30.6	100
		(0.1)	(0.0)	(0.1)	

Note: The vulnerability index is defined as $P(Y_1 < Z_1 | Z_0 < Y_0 < V_0) = 0.2$ for both periods yielding a monthly vulnerability line of 1,035 rupees per capita in the first period and 830 rupees per capita in the second period; both lines are in 2004 price for all rural India. The all rural India poverty lines are 446.68 rupees, 672.8 rupees, and 816 rupees respectively for 2004/05, 2009/10, and 2011/12 in each year's price. All numbers are estimated with synthetic panel data and weighted with population weights, where the first survey round in each period is used as the base year. Bootstrap standard errors in parentheses are estimated with 1,000 bootstraps adjusting for the complex survey design. Household head's age range is restricted to between 25 and 55 for the first survey and adjusted accordingly for the second survey in each period. Estimation sample sizes are 91,751 and 76,479 households for the first period, and 73,681 and 75,159 households for the second period.

Table 6: Welfare Transition Dynamics Based on Synthetic Panel Data at Similar Vulnerability Line, India 2004/05- 2011/12 (percentage)

Panel A: Vulnerability line corresponding to V-index= 0.2		2009			
		Poor	Vulnerable	Middle class	Total
2004	Poor	20.0	16.2	0.6	36.9
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	11.6	35.1	6.6	53.3
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	0.3	4.6	5.0	9.8
		(0.0)	(0.0)	(0.0)	(0.1)
	Total	31.9	55.9	12.2	100
		(0.1)	(0.0)	(0.1)	
Panel B: Same vulnerability line in both periods		2011			
		Poor	Vulnerable	Middle class	Total
2009	Poor	12.8	16.1	1.3	30.2
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	10.5	36.4	9.7	56.6
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	0.5	5.8	6.9	13.2
		(0.0)	(0.0)	(0.1)	(0.1)
	Total	23.7	58.3	17.9	100
		(0.1)	(0.1)	(0.1)	

Note: The vulnerability index is defined as $P(Y1<Z1|Z0<Y0<V0)= 0.2$ for the first period yielding a monthly vulnerability line of 1,035 rupees per capita in 2004 price in the first period for all rural India. We use this same vulnerability line for the second period. The all rural India poverty lines are 446.68 rupees, 672.8 rupees, and 816 rupees respectively for 2004/05, 2009/10, and 2011/12 in each year's price. All numbers are estimated with synthetic panel data and weighted with population weights, where the first survey round in each period is used as the base year. Bootstrap standard errors in parentheses are estimated with 1,000 bootstraps adjusting for the complex survey design. Household head's age range is restricted to between 25 and 55 for the first survey and adjusted accordingly for the second survey in each period. Estimation sample sizes are 91,751 and 76,479 households for the first period, and 73,681 and 75,159 households for the second period.

Table 7: Welfare Transition Dynamics Based on Synthetic Panel Data, India 2004/05-2011/12 (percentage)

Panel A: Vulnerability line corresponding to V-index= 0.2		2011			
		Poor	Vulnerable	Middle class	Total
2004	**Poor**	15.3	15.9	5.7	37.0
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	8.2	18.2	13.8	40.3
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	1.5	6.7	14.6	22.8
		(0.0)	(0.0)	(0.1)	(0.1)
	Total	25.0	40.8	34.1	100
		(0.1)	(0.0)	(0.1)	
Panel B: Vulnerability line equals twice poverty line		2011			
		Poor	Vulnerable	Middle class	Total
2004	**Poor**	15.3	18.4	3.2	37.0
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	9.0	26.6	12.1	47.7
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	0.7	5.5	9.1	15.3
		(0.0)	(0.0)	(0.1)	(0.1)
	Total	25.0	50.5	24.5	100
		(0.1)	(0.0)	(0.1)	

Note: The vulnerability index is defined as that corresponding to a vulnerability index of 0.2 (i.e., 770 rupees) in Panel A, and twice the poverty line (i.e., 893.4 rupees) in Panel B. All numbers are in 2004 price for all rural India. The all rural India poverty line is 446.68 rupees for 2004/05. All numbers are estimated with synthetic panel data and weighted with population weights, where the first survey round in each period is used as the base year. Bootstrap standard errors in parentheses are estimated with 1,000 bootstraps adjusting for the complex survey design. Household head's age range is restricted to between 25 and 55 for the first survey and adjusted accordingly for the second survey in each period. Estimation sample sizes are 91,751 and 75,159 for the first and second period respectively.

Figure 1: Annual Growth of GDP per capita vs. Headcount Poverty Rate, India 2004- 2012
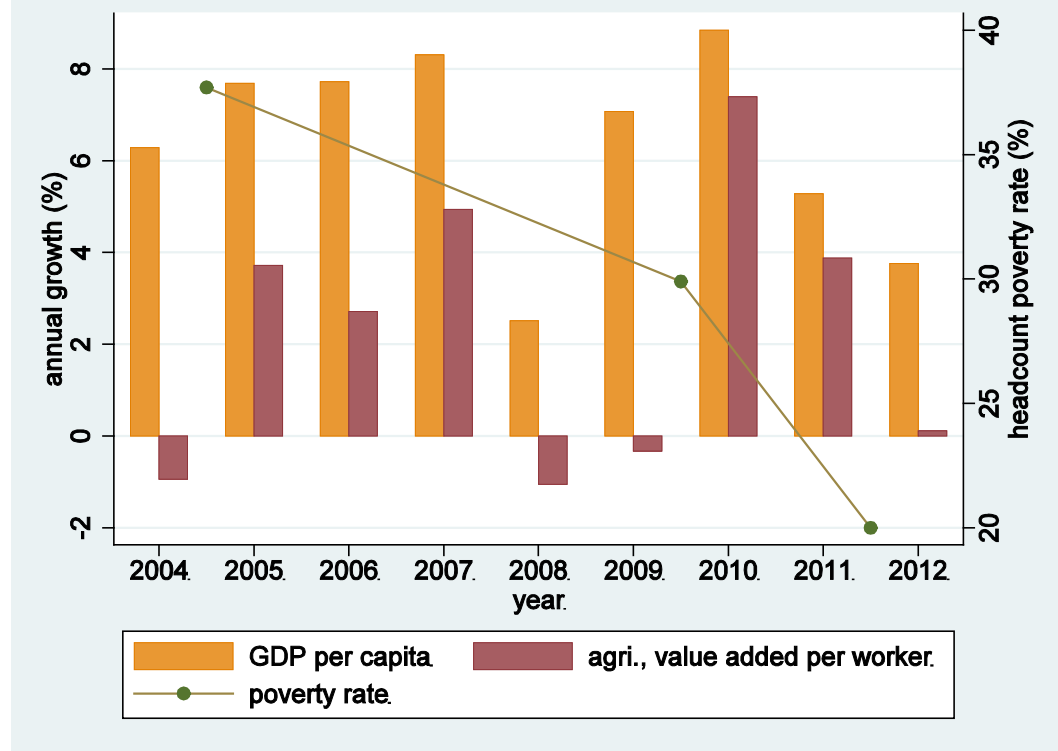
Figure 2: Vulnerability Index vs. Vulnerability Lines, India 2004/05- 2009/10

Note: dash line represents twice the 2004 all rural India poverty line (i.e., 893.4 rupees).
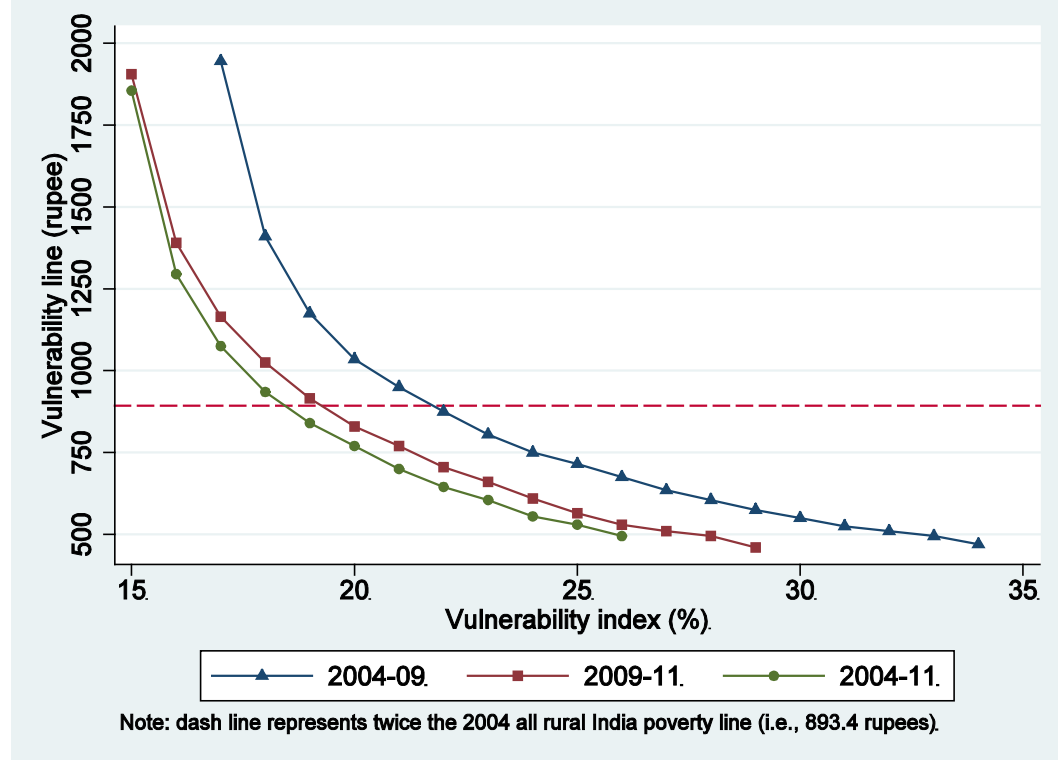
Figure 3: Profiling of the Population that Moved up One or Two Income Groups, India 2004/05- 2011/12

Note: dashed lines represent the national average for each period (i.e., 25.9% for 2004-09 and 39.8% for 2009-11).
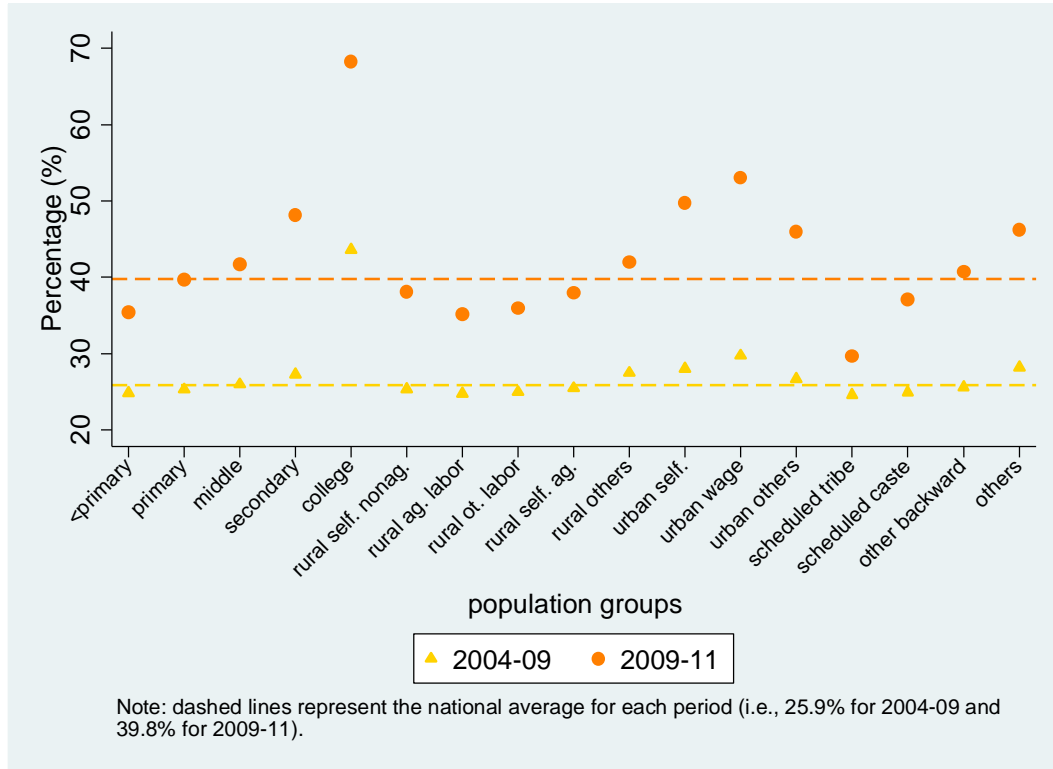
Figure 4: Profiling of the Population that Moved up One or Two Income Groups by State, India 2004/05- 2011/12

Note: dashed lines represent the national average for each period (i.e., 25.9% for 2004-09 and 39.8% for 2009-11).
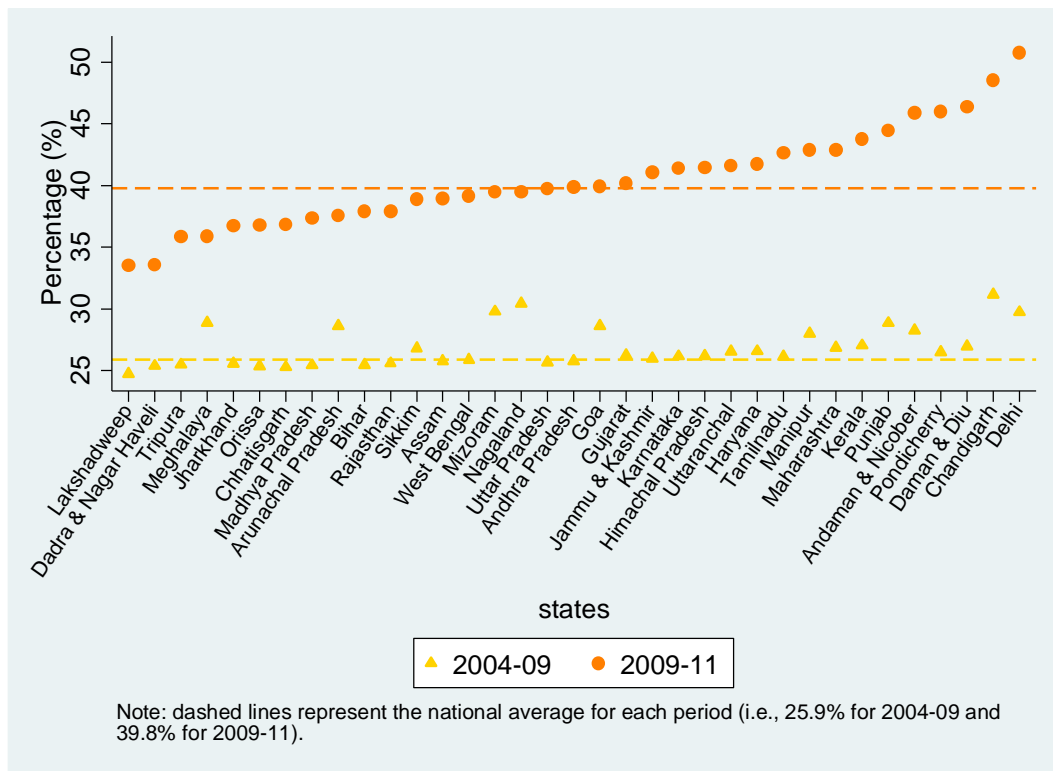
Figure 5: Profiling of the Population that Moved down One or Two Income Groups, India 2004/05- 2011/12

Note: dashed lines represent the national average for each period (i.e., 26.1% for 2004-09 and 27.5% for 2009-11).
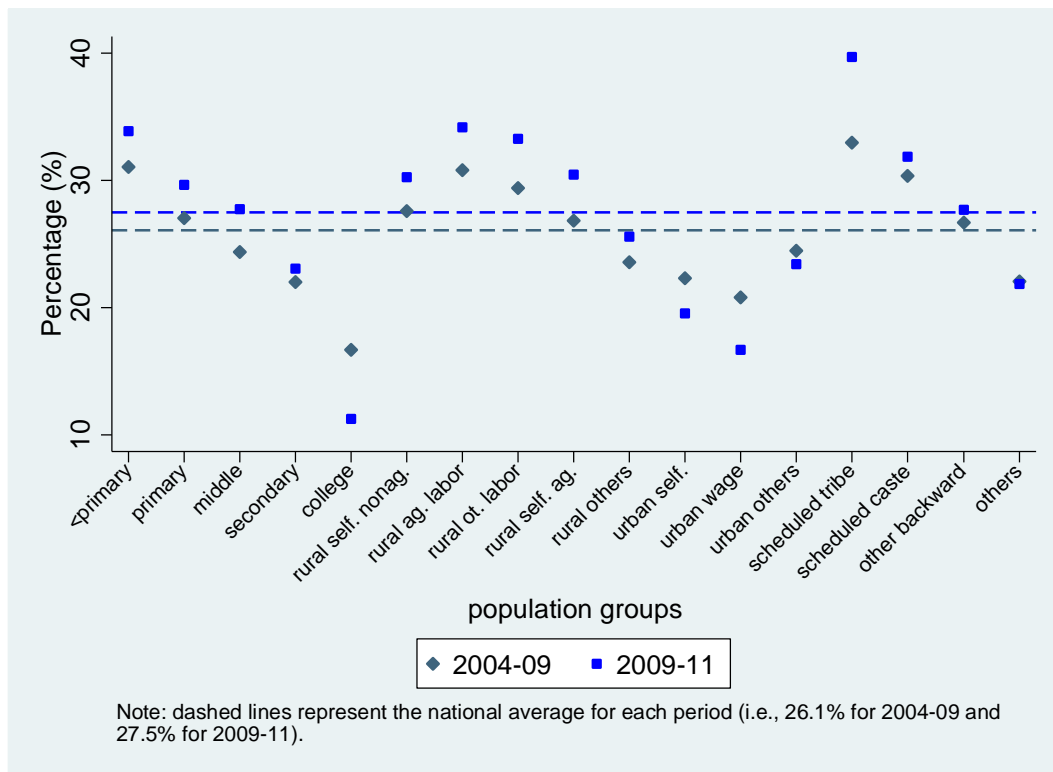
Figure 6: Profiling of the Population that Moved down One or Two Income Groups by State, India 2004/05- 2011/12

Note: dashed lines represent the national average for each period (i.e., 26.1% for 2004-09 and 27.5% for 2009-11).
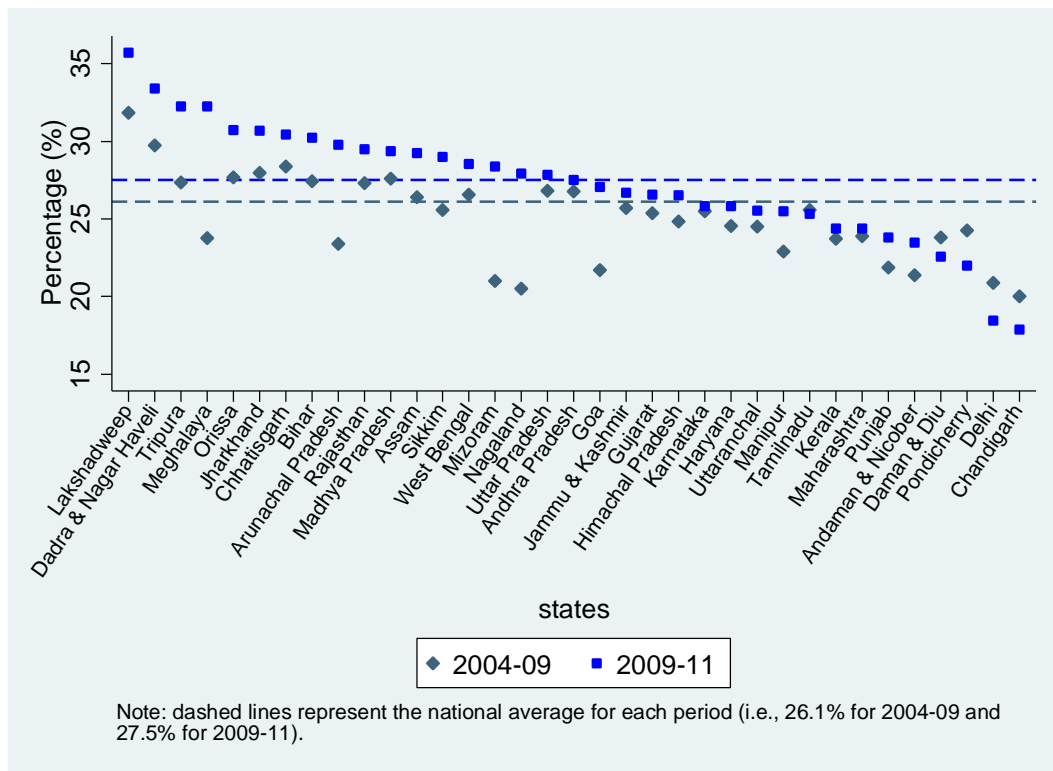
Appendix 1: Additional Tables

Table 1.1: Disaggregation of Total Household Consumption by Item, India 2009/10- 2011/12

No	Total expenditure category	2009/10		2011/12	
		Mean expenditure	Share of expenditure (%)	Mean expenditure	Share of expenditure (%)
1	same item & same code in 2009 and 2011	3275.8	63.2	4335.1	63.0
2	same item but different code in 2009 and 2011	1605.1	31.0	1981.0	28.8
3	item more disaggregated in 2009	19.8	0.4	26.7	0.4
4	item more disaggregated in 2011	75.4	1.5	76.5	1.1
5	item partly different between 2009 & 2011	80.2	1.5	114.2	1.7
6	item found in 2009 only	**128.1**	**2.5**	**N/A**	**N/A**
7	item found in 2011 only	**N/A**	**N/A**	**346.5**	**5.0**
8	Other household expenditure	0.0	0.0	0.0	0.0
	Total household expenditure	**5184.4**	**100**	**6880.0**	**100**

Note: All expenditure data are adjusted for state and sector deflators and obtained with household weight.

Table 1.2: Summary Statistics, India 2009/10 and 2011/12

	2009/10	2011/12	Differrence
Household size	5.67	5.53	-0.13***
	(2.65)	(2.54)	(0.03)
Age	46.59	46.69	0.10
	(13.05)	(13.04)	(0.13)
Female head	0.08	0.09	0.01***
	(0.27)	(0.28)	(0.00)
Hindu	0.82	0.81	-0.01
	(0.38)	(0.39)	(0.00)
Islam	0.13	0.14	0.01**
	(0.33)	(0.34)	(0.00)
Scheduled tribe	0.09	0.09	0.00
	(0.28)	(0.29)	(0.00)
Scheduled caste	0.20	0.19	-0.01***
	(0.40)	(0.39)	(0.00)
Other backward classes	0.42	0.44	0.02***
	(0.49)	(0.50)	(0.01)
Literate	0.11	0.12	0.01***
	(0.32)	(0.33)	(0.00)
Complete primary education	0.14	0.13	-0.01***
	(0.34)	(0.33)	(0.00)
Complete middle education	0.15	0.15	0.00
	(0.35)	(0.36)	(0.00)
Complete secondary education	0.11	0.12	0.00
	(0.32)	(0.32)	(0.00)
Complete senior second education	0.06	0.06	0.00
	(0.24)	(0.24)	(0.00)
Have a diploma/ certificate	0.01	0.01	-0.00
	(0.09)	(0.09)	(0.00)
Complete graduate education	0.05	0.06	0.00**
	(0.22)	(0.23)	(0.00)
Complete post-graduate education	0.02	0.02	0.00***
	(0.13)	(0.15)	(0.00)
Share of household members age 0-14	0.31	0.30	-0.01***
	(0.22)	(0.22)	(0.00)
Share of household members age 15-24	0.19	0.19	-0.00
	(0.21)	(0.21)	(0.00)
Share of household members age 25-59	0.42	0.43	0.01***
	(0.19)	(0.19)	(0.00)
Any household member has regular salary income	0.18	0.21	0.03***
	(0.38)	(0.41)	(0.00)
Rural self-employed in non-agriculture	0.12	0.12	0.00
	(0.32)	(0.33)	(0.00)
Rural self-employed in agriculture	0.26	0.27	0.01***
	(0.44)	(0.44)	(0.01)
Urban self-employed	0.11	0.12	0.00
	(0.32)	(0.32)	(0.00)
Urban regular wage/ salary earning	0.10	0.11	0.01***
	(0.30)	(0.32)	(0.00)
Urban casual labor	0.04	0.04	0.00
	(0.19)	(0.19)	(0.00)
Urban other work	0.02	0.02	0.00
	(0.13)	(0.13)	(0.00)
Own home	0.89	0.89	-0.00
	(0.32)	(0.32)	(0.00)
Main lighting source is electricity	0.73	0.78	0.06***
	(0.45)	(0.41)	(0.01)
Mainly use firewood for cooking	0.62	0.53	-0.09***
	(0.49)	(0.50)	(0.01)
Mainly use LPG for cooking	0.27	0.30	0.03***
	(0.44)	(0.46)	(0.00)
Own a radio	0.27	0.20	-0.07***
	(0.44)	(0.40)	(0.00)
Own a television	0.53	0.61	0.08***
	(0.50)	(0.49)	(0.01)
Own an electric fan	0.65	0.72	0.07***
	(0.48)	(0.45)	(0.01)
Own a sewing machine	0.16	0.19	0.03***
	(0.37)	(0.40)	(0.00)
Own a freezer	0.17	0.21	0.04***
	(0.38)	(0.41)	(0.00)
Own an air-conditioner	0.11	0.13	0.02***
	(0.31)	(0.33)	(0.00)
Own a bike	0.57	0.58	0.01**
	(0.49)	(0.49)	(0.01)
Own a motor bike	0.22	0.27	0.05***
	(0.42)	(0.45)	(0.00)
Own a car	0.03	0.04	0.01***
	(0.18)	(0.21)	(0.00)
N	99,469	101,525	

Note: * p<0.05, ** p<0.01 *** p<0.001. Standard deviations/ errors are in parentheses.
Differences are estimated with t-tests that takes into account complex survey design with cluster
sampling and stratification. All estimates are obtained with population weights.

Table 1.3: Model Specifications of Household Consumption, India 2004/05

	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5
Household size	-0.064***	-0.041***	-0.045***	-0.044***	-0.055***
Age	0.015***	0.001	0.000	0.002***	-0.002***
Age squared	-0.000***	0.000***	0.000***	0.000***	0.000***
Female head	0.014***	0.046***	0.046***	0.047***	0.030***
Hindu	-0.104***	-0.108***	-0.098***	-0.100***	-0.054***
Islam	-0.171***	-0.158***	-0.135***	-0.136***	-0.021***
Scheduled tribe	-0.229***	-0.217***	-0.216***	-0.212***	-0.111***
Scheduled caste	-0.224***	-0.208***	-0.178***	-0.175***	-0.086***
Other backward classes	-0.108***	-0.099***	-0.092***	-0.091***	-0.046***
Literate	0.105***	0.097***	0.084***	0.083***	0.042***
Complete primary education	0.164***	0.152***	0.132***	0.132***	0.069***
Complete middle education	0.274***	0.259***	0.226***	0.227***	0.113***
Complete secondary education	0.423***	0.400***	0.356***	0.356***	0.157***
Complete senior second education	0.548***	0.519***	0.469***	0.468***	0.237***
Have a diploma/ certificate	0.724***	0.696***	0.634***	0.630***	0.312***
Complete graduate education	0.785***	0.752***	0.697***	0.697***	0.348***
Complete post-graduate education	0.928***	0.895***	0.835***	0.836***	0.421***
Share of household members age 0-14		-0.341***	-0.334***	-0.340***	-0.349***
Share of household members age 15-24		0.100***	0.088***	0.079***	0.020**
Share of household members age 25-59		0.281***	0.266***	0.257***	0.158***
Any household member has regular salary income			0.139***	0.134***	0.079***
Rural self-employed in non-agriculture			0.120***	0.122***	0.068***
Rural self-employed in agriculture			0.164***	0.166***	0.143***
Urban self-employed			0.066***	0.052***	-0.147***
Urban regular wage/ salary earning			0.009*	-0.016***	-0.160***
Urban casual labor			-0.168***	-0.187***	-0.221***
Urban other work			0.166***	0.147***	0.005
Own home				-0.079***	-0.129***
Main lighting source is electricity					0.043***
Mainly use firewood for cooking					-0.076***
Mainly use LPG for cooking					0.088***
Own a radio					0.080***
Own a television					0.090***
Own an electric fan					0.081***
Own a sewing machine					0.023***
Own a freezer					0.202***
Own an air-conditioner					0.041***
Own a bike					0.003
Own a motor bike					0.208***
Own a car					0.340***
Constant	6.272***	6.489***	6.448***	6.476***	6.615***
σ_u	0.10	0.10	0.10	0.11	0.00
σ_e	0.43	0.42	0.41	0.41	0.37
R2 (overall)	0.05	0.05	0.06	0.07	0
ρ	0.38	0.41	0.43	0.43	0.56
N	124543	124543	124374	124340	119292

Note: * p<0.05, ** p<0.01 *** p<0.001. Standard errors are not shown for lack of space. All model specifications use normal linear regression with state random effects.

Table 1.4: Model Specifications of Household Consumption, India 2009/10

	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5
Household size	-0.074***	-0.054***	-0.057***	-0.055***	-0.070***
Age	0.006***	-0.007***	-0.008***	-0.005***	-0.010***
Age squared	-0.000	0.000***	0.000***	0.000***	0.000***
Female head	0.019***	0.045***	0.041***	0.044***	0.039***
Hindu	-0.114***	-0.117***	-0.107***	-0.110***	-0.043***
Islam	-0.164***	-0.155***	-0.130***	-0.131***	-0.019***
Scheduled tribe	-0.197***	-0.190***	-0.195***	-0.191***	-0.113***
Scheduled caste	-0.208***	-0.197***	-0.168***	-0.164***	-0.086***
Other backward classes	-0.104***	-0.098***	-0.092***	-0.090***	-0.054***
Literate	0.104***	0.094***	0.080***	0.081***	0.053***
Complete primary education	0.152***	0.141***	0.119***	0.121***	0.065***
Complete middle education	0.259***	0.245***	0.212***	0.214***	0.109***
Complete secondary education	0.401***	0.384***	0.334***	0.337***	0.147***
Complete senior second education	0.538***	0.516***	0.455***	0.456***	0.209***
Have a diploma/ certificate	0.723***	0.705***	0.624***	0.621***	0.307***
Complete graduate education	0.729***	0.708***	0.640***	0.642***	0.294***
Complete post-graduate education	0.876***	0.852***	0.776***	0.777***	0.377***
Share of household members age 0-14		-0.264***	-0.253***	-0.263***	-0.295***
Share of household members age 15-24		0.148***	0.139***	0.122***	0.061***
Share of household members age 25-59		0.277***	0.266***	0.253***	0.161***
Any household member has regular salary income			0.150***	0.143***	0.077***
Rural self-employed in non-agriculture			0.106***	0.105***	0.036***
Rural self-employed in agriculture			0.188***	0.190***	0.124***
Urban self-employed			0.049***	0.031***	-0.175***
Urban regular wage/ salary earning			0.003	-0.029***	-0.160***
Urban casual labor			-0.214***	-0.236***	-0.268***
Urban other work			0.149***	0.122***	-0.006
Own home				-0.112***	-0.177***
Main lighting source is electricity					0.038***
Mainly use firewood for cooking					-0.086***
Mainly use LPG for cooking					0.028***
Own a radio					0.035***
Own a television					0.098***
Own an electric fan					0.069***
Own a sewing machine					0.040***
Own a freezer					0.206***
Own an air-conditioner					0.064***
Own a bike					-0.029***
Own a motor bike					0.185***
Own a car					0.335***
Constant	7.028***	7.194***	7.161***	7.190***	7.407***
σ_u	0.12	0.10	0.08	0.05	0.00
σ_e	0.44	0.43	0.42	0.42	0.37
R2 (overall)	0.07	0.05	0.04	0.01	0.00
ρ	0.37	0.39	0.42	0.42	0.56
N	100832	100832	100798	100595	99469

Note: * $p<0.05$, ** $p<0.01$ *** $p<0.001$. Standard errors are not shown for lack of space. All model specifications use normal linear regression with state random effects.

Table 1.5: Welfare Transition Dynamics Based on Synthetic Panel Data at a Vulnerability Line Equal to Twice the Poverty Line, India 2004/05- 2011/12 (percentage)

Panel A: Vulnerability line equals twice poverty line		2009			
		Poor	Vulnerable	Middle class	Total
2004	**Poor**	20.0	15.4	1.4	36.9
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	11.2	27.9	8.7	47.9
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	0.7	6.1	8.4	15.3
		(0.0)	(0.0)	(0.1)	(0.1)
	Total	31.9	49.5	18.6	100
		(0.1)	(0.0)	(0.1)	
Panel B: Vulnerability line equals twice poverty line		2011			
		Poor	Vulnerable	Middle class	Total
2009	**Poor**	12.8	14.9	2.5	30.2
		(0.0)	(0.0)	(0.0)	(0.1)
	Vulnerable	9.9	28.0	12.1	50.0
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	1.0	7.5	11.3	19.9
		(0.0)	(0.0)	(0.1)	(0.1)
	Total	23.7	50.4	25.9	100
		(0.1)	(0.1)	(0.1)	

Note: The vulnerability line is defined as twice the poverty line yielding a monthly vulnerability line of 893.4 rupees per capita in 2004 price in the first period for all rural India. We use this same vulnerability line for the second period. The all rural India poverty line is 446.68 rupees, 672.8 rupees, and 816 rupees respectively for 2004/05, 2009/10, and 2011/12 in each year's price. All numbers are estimated with synthetic panel data and weighted with population weights, where the first survey round in each period is used as the base year. Bootstrap standard errors in parentheses are estimated with 1,000 bootstraps adjusting for the complex survey design. Household head's age range is restricted to between 25 and 55 for the first survey and adjusted accordingly for the second survey in each period. Estimation sample sizes are 91,751 and 76,479 households for the first period, and 73,681 and 75,159 households for the second period.