CESifo Working Papers

www.cesifo.org/wp

A Behaviour-based Approach to the Estimation of Poverty in India

Ingvild Almås
Anders Kjelsrud
Rohini Somanathan

CESIFO WORKING PAPER NO. 4122
CATEGORY 3: SOCIAL PROTECTION
FEBRUARY 2013

CESifo *Working Paper No. 4122*

A Behaviour-based Approach to the Estimation of Poverty in India

Abstract

Since the late 1970s, the price indices underlying the poverty lines in India have been updated using aggregate indices. Widespread criticism of these indices led to the adoption of a new official methodology in 2011 based on unit values from consumption survey data. We propose an alternative approach that identifies poverty from consumer behaviour, based on the notion that equally poor households spend the same proportion of their incomes on food. Compared with official estimates, we find higher levels of poverty in eastern India, and generally, smaller reductions in poverty from 2005 to 2010. Our poverty numbers are validated by the calorie composition of households around the poverty lines and self-reported hunger.

JEL-Code: D100, E310, F010, I320.

Ingvild Almås
Department of Economics
Norwegian School of Economics and
Business Administration
Helleveien 30
Norway - 5045 Bergen
ingvild.almas@nhh.no

Anders Kjelsrud
ESOP / University of Oslo
Oslo / Norway
a.g.kjelsrud@econ.uio.no

Rohini Somanathan
Delhi School of Economics
University of Delhi
Delhi / India
rohini@econdse.org

1 Introduction

Most debates about poverty estimates are either about the consumption basket that defines adequacy or variations in its cost. In India, poverty lines were derived based on calorie norms in the late 1970s and have since been adjusted only to account for price changes. The controversy surrounding the poverty counts is largely about appropriate methods of cost comparisons across states and over time. Alternative approaches have led to very different results. In 2005, a little over a quarter of the Indian population was estimated to live in poverty in both urban and rural areas. These figures were recently updated using new methods for estimating price differences, which resulted in rural head counts being almost 50 per cent higher. These new poverty lines have been challenged and the search for a clear and defensible methodology for poverty measurement continues. The poverty debate in India is particularly charged because of the mismatch between the rhetoric of poverty eradication and performance and also because government programmes are increasingly being targeted at families that are officially classified as poor.

In this paper, we use an alternative approach to estimating spatial and temporal variations in poverty that does not rely on the direct measurement of prices. We identify the poor through their share of total expenditure on food. Since Ernst Engel's work (Engel, 1857, 1895), the empirical regularity of a downward sloping Engel curve for food has been well established. Our empirical strategy exploits the Engel relationship for Indian data to determine cost-of-living differences across states over a five-year period, 2005–10. We assume that households with the same demographics and occupational characteristics who face the same relative prices, spend the same proportion of their income on food. This enables us to use differences in nominal expenditures of households with the same food share and the same characteristics, as estimates of differences in aggregate price levels.

Our data come from the National Sample Surveys (NSS), which have been the standard source for poverty studies in India. We normalize our price estimates to yield the same aggregate price level for 2005 as official methods use to allow for a meaningful comparison of the poverty rates.

Our results point to three main conclusions. First, the dispersion in price levels across

Indian states exceeds that suggested by the official poverty lines. The consequent poverty counts indicate a greater dispersion also in poverty rates among states. Second, poverty rates in the eastern states of Assam, Bihar, Orissa and West Bengal are consistently higher than those implied by official figures and exceed 50 per cent in both survey years. Third, the decrease in overall poverty over our five-year period is much more modest than the one suggested by official statistics.

We compare our results with existing estimates by examining the behaviour of the poor around alternative poverty lines. One might reasonably expect that the poor get their calories from relatively cheap sources while the less poor substitute towards more expensive calories that have other favourable attributes such as taste or status (Jensen and Miller, 2010). Our survey data indicate that cereals are the cheapest sources of calories and that their share of calories falls as household expenditure increases. If our estimated poverty lines do represent the same level of real household income, we would expect the households around these lines to have similar consumption patterns. To test for this, we limit our sample to households in a symmetric five per cent interval around the poverty lines for each state and for each of the two time periods. We find that households clustered around our estimated lines get roughly the same share of their total calories from cereals. This is not true for the official lines, which suggests that we have been able to more accurately identify equally poor households across states and time periods. The same result is found for edible oils, which is an example of an expensive source of calories. As a final validation check, we examine rates of self-reported hunger across states and find the highest rates in many of the states that we classify as the poorest.

Several studies have used estimates of Engel curves to correct for biases in inflation series. Hamilton (2001) and Costa (2001) pioneered this strand of literature through their studies of consumer price indices in the United States. Beatty and Larsen (2005); Larsen (2007); Gibson *et al.* (2008); Olivia and Gibson (2012); Barrett and Brzozowski (2010); Chung *et al.* (2010); Carvalho Filho and Chamon (2006) have applied this method to other countries. More recently, it has also been used to estimate regional price differences (Almås, 2012; Gong and Meng, 2008). Our main contribution is to apply the methodology to Indian data and to use it to derive poverty measures. Our data are richer than those used in many of these studies and they allow us to control for demographic, occupational and relative price variables that are likely to influence the budget share for food. This makes our identifying assumption of a stable Engel curve much more plausible. In addition, we are able to validate our estimates by making an ex post comparison of the behaviour of households that we estimate as being similar.

The recent changes in the methodology for estimating poverty in India seem to constitute a significant improvement on previous techniques. For example, we show that the new estimates of price levels within states exhibit much higher correlations than before, and the spatial distribution of poverty seems more consistent with those from other studies (see e.g. Deaton and Tarozzi, 2000; Deaton, 2003). However, the current methodology remains ad hoc in many respects and we believe that the Engel approach provides a viable and theoretically grounded alternative.

The rest of this paper is organized as follows. In Section 2, we outline a brief chronology of poverty measurement in India. In Section 3, we describe our empirical methodology in detail and discuss the data and the variables used in the main analysis. The key findings are presented in Section 4. In Section 5, we present results from a range of specification checks. In Section 6, we evaluate the various poverty outcomes. Concluding remarks are provided in Section 7.

2 Poverty measurement in India: a brief chronology

Several poverty lines were constructed in India during the colonial period (Srinivasan, 2007). For independent India, the first poverty lines were set by the Planning Commission Working Group in 1962. These were 20 and 25 rupees per capita per month for rural and urban areas, respectively, in 1960–61 prices. These rather arbitrary lines persisted for almost 20 years.

An attempt to link poverty estimates to nutritional needs was made in the late 1970s. Calorie norms of 2400 and 2100 calories per capita per day for the rural and urban sector were used, and the expenditure equivalents of these norms were identified through the empirical expenditure distribution observed in the NSS survey of 1973–74 (Government of India, 1979). The resulting poverty lines were 49 rupees in rural areas and 57 rupees in urban areas. No attempt was made to capture differences in prices or spending across states and the above poverty lines were applied to the entire country.[1]

Regionally disaggregated prices were first used in the mid-1990s to derive separate poverty lines for rural and urban sectors within each state. This was done by another official expert group, led by D.T. Lakdawala (Government of India, 1993). Spatial price indices had been computed for the 1960s in two previous studies based on NSS data. These series were extended for later years by using the consumer price index for agricultural labours (CPIAL) for rural areas and the consumer price index for industrial workers (CPIIW) for urban areas. Both indices were reweighted to reflect the consumption patterns of the poor in 1973–74.[2]

Over the years, this methodology came under substantial criticism (see e.g. Deaton and Tarozzi, 2000; Deaton, 2003, 2008). The implied rural-to-urban price differentials were considered too large to be credible; the consumption weights used in the price indices were only infrequently updated; and the poverty lines failed to preserve the original calorie norms. A new expert committee was formed in late 2005, chaired by Suresh Tendulkar, and most of its suggestions were adopted by the Planning Commission in 2011 (Government of India, 2009, 2011). The new approach no longer anchors the poverty lines to any form of calorie intake norm. However, for the sake of continuity, the all-India urban head count for 2004–05 of 25.7 per cent was taken as given and, as a result, the new estimates are best thought of as providing new relative poverty lines across states and urban and rural sectors.

The new method derives a set of updated prices for poverty measurement through unit values computed from the same NSS data that are used to estimate household expenditure. In this sense, it is similar in spirit to the methods proposed by Angus Deaton and his co-authors. Although unit values may differ from prices because they do not adjust for differences in quality, it has been argued that these biases are quite small (Deaton, 1988). More importantly, it is only possible to construct unit values for a proportion of

[1]All lines are stated in terms of per capita expenditures per month.

[2]For the rural sector, the spatial price index in Bhattacharya and Chatterjee (1974) was used, and for the urban sector, the spatial index in Minhas and L.R (1989) was used.

the goods and services consumed by a household; namely, those for which survey data can provide meaningful quantities. This includes food items and fuels, but excludes education, health care and other services. For categories of consumption for which unit values cannot be readily computed, the Tendulkar methodology uses price information from a variety of sources. The cost of school attendance is derived from the NSS employment–unemployment survey; health care costs are calculated from the NSS Morbidity and Health Care survey; and prices for the remainder of households' consumption bundles (including entertainment, services and durables) are derived from the price data underlying the CPIAL and CPIIW. This makes the new procedures somewhat ad hoc as future lines, and poverty counts are likely to depend on another set of myriad sources that are not necessarily comparable with the current ones.[3]

We now describe our alternative approach based on estimating food shares using the Engel relationship.

3 Empirical methodology: the Engel method

The identification strategy of the Engel method is to assume that similar households with the same real income, facing the same relative prices, also have the same budget share for food. Hence, if two comparable households in different Indian states have the same expenditure levels, we attribute any difference in their budget shares for food that cannot be explained by differences in relative prices, to price level differences between the states. Because of potential unobservable differences between rural and urban areas of India, we estimate on the rural and urban sample separately.

There are several advantages of using food as the indicator good. First, because the income elasticity differs substantially from unity, the budget share is sensitive to the level of household real income and, thereby, to the price deflator for nominal income. Second, because food is a perishable good, expenditures in one period cannot provide a flow of consumption goods in another period. Third, studies of different countries, and over different time periods, suggest that the Engel curve for food is log-linear and stable

[3]Subramanian (2011) provides a critical review of the new methodology.

(Banks *et al.*, 1997; Beatty and Larsen, 2005; Blundell *et al.*, 1998; Leser, 1963; Working, 1943; Yatchew, 2003).

In our main estimation, we use the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980a) to estimate the Engel curves. This system extends the specifications by Working (1943) and Leser (1963). Deaton and Muellbauer showed that the AIDS is consistent with utility maximization under a budget constraint. AIDS preferences are flexible; they allow for non-homothetic tastes as well as substitution in consumption (see Deaton and Muellbauer, 1980b). The AIDS is of rank two. The extended Quadratic Almost Ideal Demand System (QAIDS) of Banks *et al.* (1997) is of rank three and is hence even more flexible. Whereas it is clear that the QAIDS is preferable to AIDS for many goods, this is not clearly the case for food; e.g., Banks *et al.* (1997) themselves cannot reject the null hypothesis of a log-linear relationship for food. Because it is standard to use the AIDS in the Hamilton tradition, we use this system in our main estimation. However, our robustness analysis shows that the QAIDS generates similar results. The robustness analysis also contains a more general discussion of the functional form assumptions.

3.1 The AIDS model

The Engel curve in the AIDS is given by:

$$m_{hdst} = a + b(\ln y_{hdst} - \ln P_{st}) + \gamma(\ln P^{f}_{dst} - \ln P^{n}_{dst}) + \theta X_{hdst} + \epsilon_{hdst}, \quad (1)$$

where m_{hdst} is the budget share for food, y_{hdst} is the nominal household expenditure level, and X_{hdst} is a vector of household-specific control variables, such as demographics, religion and occupation, for household h in district d in state s at time t. P_{st} is the composite price of consumption in state s at time t. P^{f}_{dst} is the price of food and P^{n}_{dst} is the price of non-food items in district d in state s at time t.[4]

The only unknown variable in this regression is the overall state price level P_{st}. This is also the only variable measured at the *state* level. Hence, it can be identified through state- and time-specific dummy variables:

[4]This is the standard specification used in the literature following Hamilton (2001) and Costa (2001).

$$m_{hdst} = a + b \ln y_{hdst} + \gamma(\ln P^f_{dst} - \ln P^n_{dst}) + \theta X_{hdst} + \sum_{s=2}^{N} d_{s1} D_{s1} + \sum_{s=1}^{N} d_{s2} D_{s2} + \epsilon_{hdst}, \quad (2)$$

where D_{st} is the state level dummy variable for state s in period t, and N is the total number of states. State 1 in period 1 is taken as the base and, hence, D_{11} is not included in the estimation. The state dummy coefficient, d_{st}, is a function of the overall state price level, P_{st}, and the coefficient for the logarithm of household expenditures, b:

$$d_{st} = -b \ln P_{st}. \quad (3)$$

From Equation (3), it follows that the overall price level is given by:

$$P_{st} = e^{-\frac{d_{st}}{b}}. \quad (4)$$

This price level is measured relative to the base state in the base time period.[5]

3.2 Data and control variables

The NSS publishes large consumer expenditure surveys for all states and major union territories in India, typically every five years. We use data from the two latest large survey rounds, obtained in 2004–05 (the 61st round) and 2009–10 (the 66th round). Our analysis is based on data from all 28 Indian states, in addition to Delhi and Puducherry.[6] Summary statistics, covering 222,558 households, are shown in Table 1. All mean values in the table are calculated using the sample weights given by the NSS.

The expenditure surveys are conducted with a 30-day recall period for most consumption goods, while expenses on some infrequently purchased items are evaluated using a 365-day recall period.[7] The measure of total monthly consumption contains all expenditures

[5] This is a normalization. All results are invariant to the choice of base state and period.

[6] We exclude the union territories of A&N Islands, Chandigarh, Daman & Diu, D&N Haveli and Lakshadweep, which together constitute barely one per cent of the NSS sample. Our final sample of 30 states and union territories is identical to the one used in the construction of the official poverty lines.

[7] This applies to clothing, bedding, footwear, education, institutional medical expenses and durable goods. The 66th NSS round is published as two separate surveys, each with different recall periods. To

incurred by households on consumption goods and services, including in-kind and home production. The NSS values items received in-kind at their average local retail price, while home production is evaluated at the ex-farm or ex-factory rate (which implies that distributive service charges are excluded.)

Most of the control variables used in the estimation are taken directly from the NSS. This includes the household's female proportion, occupation, religion, land ownership, number of free meals and the age of the household head. In addition to these variables, we construct a unit value index to control for potential differences in relative prices for food and non-food items. The unit values are calculated from NSS data on the quantity and value of consumption, simply by dividing value by quantity. This is possible for 127 food items and 41 non-food items. We then compute median unit values for all these 168 consumption items at the district level.[8] Aggregate measures for food and non-food prices are subsequently derived by applying the weighted country-product-dummy method of Rao (1990). Finally, the relative price indices used in our analysis are calculated as food prices divided by non-food prices (see Appendix A for details of the construction of the relative price variable).

To avoid potential misspecifications arising from variations in family composition, we use only households comprising two children and two adults in the main estimation. Although this is the most frequently observed family composition in the NSS dataset, the restriction reduces our sample size by almost 90 per cent. As a robustness check, we therefore run the estimation on the full sample including controls for the number of children and the number of adults. All main findings are robust to this change of sample.

obtain a comparable sample for the two time periods, and for comparability with the official poverty counts, we use the "type 1" survey version.

[8]We use the median rather than the mean because it is less sensitive to outliers.

Table 1: Household descriptive statistics

		Rural				Urban			
		2004–05		2009–10		2004–05		2009–10	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Monthly per capita expenditure (MPCE)	Rupees	578.90	409.64	952.51	723.59	1101.63	923.59	1851.60	1788.62
Children in HH	#	2.48	1.87	2.17	1.71	1.88	1.68	1.68	1.55
Adults in HH	#	3.62	1.88	3.61	1.82	3.72	1.95	3.66	1.89
HH head's age	#	46.06	13.23	46.62	12.96	46.14	13.14	46.38	13.30
Proportion of females in HH	Share	0.49	0.16	0.48	0.16	0.48	0.18	0.48	0.18
Self-employed (non-agriculture)	Share	0.17	0.37	0.16	0.37				
Self-employed (agriculture)	Share	0.39	0.49	0.35	0.48				
Agricultural labour	Share	0.25	0.43	0.25	0.43				
Self-employed	Share					0.43	0.50	0.42	0.49
Salaried labour	Share					0.39	0.49	0.37	0.48
Casual labour	Share					0.12	0.32	0.14	0.35
Hindu	Share	0.84	0.37	0.84	0.37	0.78	0.42	0.78	0.41
Muslim	Share	0.11	0.32	0.12	0.32	0.16	0.37	0.16	0.37
Christian	Share	0.02	0.14	0.02	0.15	0.02	0.15	0.03	0.16
Other religion	Share	0.03	0.17	0.03	0.16	0.04	0.19	0.04	0.19
Cultivated land	Hectares								
None	Share	0.36	0.48	0.40	0.49	0.90	0.29	0.91	0.28
Less than 2	Share	0.51	0.50	0.48	0.50	0.08	0.26	0.07	0.25
2 or more	Share	0.13	0.34	0.12	0.32	0.02	0.14	0.02	0.14
Free meals outside HH	#	0.08	0.18	0.09	0.17	0.06	0.19	0.06	0.18
No. of households	1000		78.64		58.61		44.40		40.91

Note: The table displays summary statistics for 30 Indian states and union territories. All variables are weighted by the multipliers provided by the NSS.

4 Findings

4.1 Regression estimates

Table 2 reports the key results from the estimation of the AIDS model given in Equation 2, separately for the rural and urban sector. The table shows that the logarithm of total monthly expenditure has a significantly negative effect on the budget share for food. This is as expected from Engel's Law. The size of the coefficients implies expenditure elasticities of +0.77 and +0.70 in the rural and urban sectors, respectively.[9] These values are consistent with estimates from previous studies of other countries (see e.g. Almås, 2012; Beatty and Larsen, 2005; Carvalho Filho and Chamon, 2006; Costa, 2001). For the relative price effect different studies show substantially different elasticities. We find a significantly positive effect in the urban sector. The point estimate is consistent with a price elasticity of -0.82, which is consistent with, e.g., the estimate obtained by Costa (2001) for 1960–90 for the United States.[10] For the rural sector, however, there is no significant effect of relative prices, which is consistent with the finding of, e.g., Hamilton (2001).

Table 2: Pooled AIDS regression model

Dep. var.: Budget share for food (percentage)	**Rural**	**Urban**
Log of household expenditure	-12.628	-13.761
	(0.215)	(0.181)
Log of relative food/non-food prices	-0.085	1.902
	(0.623)	(0.786)
Observations	14257	9112
R^2	0.379	0.523

Note: Robust standard errors are in parentheses.The controls not shown in the table are: the age of the household head; the proportion of females in the household; three occupation dummies for each sector (urban and rural); three religion dummies; the number of free meals taken outside the home; dummies for cultivated land; and dummies for every state in each time period.

[9] The expenditure elasticities are calculated as $1 + \frac{\beta}{m}$. The estimates are derived by using the mean food share (m) from the sample used for estimation.

[10] This elasticity is calculated as $-1 + \frac{\gamma - mb}{m}$.

4.2 Constructing price and poverty measures

The estimation results from the previous subsection, and specifically the coefficients for the log of expenditure and the dummy coefficients, are used to calculate updated price measures (see Equation 4). This is possible only up to a normalization; all prices are relative to the base state and time period. However, for the purpose of presentation, we reweight the indices so that they are relative to the population-weighted all-India index value (which is normalized to 100). We include *all* the 30 states and union territories used for estimation. In presenting the results, however, we focus on the spatial price indices and the poverty estimates for the 17 largest states, those labelled "major states" by the NSS. These states cover roughly 80 per cent of the Indian population.[11] Table 3 includes a set of spatial indices for 2004–05 and 2009–10. The "Engel" columns show the prices derived from the estimates from Section 4.1, whereas the columns labelled "IPC" report the prices implied by the official poverty lines. The latter are calculated by dividing the state-specific poverty lines by the all-India poverty line for each sector and time period, and normalizing such that the weighted average equals 100.

The last row of the table summarizes the dispersion in the indices, measured by the coefficient of variation (CV). The CV suggests that there is more price variation across rural than across urban areas. Furthermore, the Engel prices imply somewhat more dispersion than is suggested by the official measures. This is true for both sectors.

One simple validity check of the various price estimates is to investigate the correlation between the rural and urban indices. Given the belief that markets are relatively well integrated within states, we would expect to see a substantial positive correlation. That is, we expect states with a high (low) price level relative to the all-India average in one sector to have a relatively high (low) price level also in the other sector (Deaton and Tarozzi, 2000). It transpires that the Engel indices are strongly positively correlated between rural and urban areas (given correlation coefficients of 0.92 and 0.83 in 2004–05 and 2009–10, respectively). The corresponding correlation coefficients for the implicit IPC prices are also positive, but somewhat lower (0.81 and 0.72, respectively). A striking

[11]We focus on these 17 states both for presentational clarity and because we consider the price and poverty estimates for the other 13 states to be less reliable. This is because relatively few households are covered by the NSS in each of these states; e.g., in rural Delhi, only 59 households are covered.

contrast is that the old IPC measures exhibit a *negative* correlation between spatial prices in rural and urban areas (−0.34 in 2004–05).[12] This seems implausible and suggests that the price measures in use until recently were out of date.[13]

We also construct a set of intertemporal indices. These numbers are derived by comparing the dummy coefficients for each state and sector for 2004–05 and 2009–10. Summary statistics are presented in Table 7. The Engel analysis suggests an overall cost-of-living increase of about 60 per cent for the five-year period in both sectors. This implies an average annual increase of approximately 10 per cent. By comparison, the implicit IPC price measures indicate an overall increase of 50 per cent, which is equivalent to an average annual increase of approximately 9 per cent.[14] In particular, the Engel analysis identifies relatively higher cost-of-living increases for some western and south-western states, such as Karnataka, Maharashtra and Rajasthan.

Given the spatial and intertemporal indices, it is relatively straightforward to compute updated poverty rates. However, since our price measures are identified only up to a normalization, we are forced to choose a base year. This choice is necessarily somewhat arbitrary. We proceed by normalizing our set of indices so that they match the all-India poverty lines for 2004–05, as they appear in the Tendulkar Expert Group report (Government of India, 2009).[15]

Specifically, we compute updated state poverty lines for both time periods using the Engel price indices and the poverty line base of 2004–05.[16] This procedure implies that our estimated all-India head-count ratios for 2004–05 differ from the official ones only because of different spatial prices. However, the head counts for 2009–10 are expected to

[12]This correlation coefficient is based on the spatial prices implied by the old official poverty lines. Note that these poverty lines were estimated for only 22 of the 30 states and union territories. The spatial price index is therefore based on these 22 states only.

[13]See Deaton and Tarozzi (2000) and Deaton (2003) for similar findings for earlier years.

[14]The official consumer price indices are closer to the Engel estimates for the rural sector but not for the urban sector.

[15]This is the most attractive normalization because we want our measures to be comparable with the official ones. An alternative would be to use 2009–10 as the base year. This alternative normalization yields a slightly larger decrease in urban poverty and a slightly smaller decrease in rural poverty compared with the estimated decreases based on the normalization actually used. Note that if the income distributions were uniform, the choice of normalization would have no impact on the poverty measures.

[16]The 2004–5 poverty line for state s is given by $l_s^{4-5} = o^{4-5} * \frac{SPI_s}{100}$, where o^{4-5} is the official all-India poverty line in 2004–05 and SPI_s is the SPI for state s. The 2009–10 poverty line for state s is given by $l_s^{9-10} = l_s^{4-5} * \frac{i_s}{100}$, where i_s is the price increase for state s.

Table 3: Spatial price indices

	2004–05				2009–10			
	Rural		**Urban**		**Rural**		**Urban**	
	Engel	*IPC*	*Engel*	*IPC*	*Engel*	*IPC*	*Engel*	*IPC*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Andhra Pradesh	96.9	96.0	91.7	96.6	102.6	96.0	98.0	105.8
Assam	171.4	105.8	158.5	103.0	141.9	105.8	129.1	99.5
Bihar	126.4	96.0	121.8	90.3	117.1	96.0	117.2	88.6
Chhattisgarh	71.0	88.3	96.2	88.2	62.6	88.3	68.1	92.1
Gujarat	109.7	111.1	112.7	113.1	122.6	111.1	108.8	108.7
Haryana	97.9	117.2	94.5	107.5	123.5	117.2	114.6	111.4
Jharkhand	117.0	89.6	117.8	91.2	85.1	89.6	114.2	94.9
Karnataka	82.7	92.5	89.5	100.9	95.3	92.5	94.2	103.7
Kerala	93.5	119.0	110.4	100.3	82.0	119.0	101.4	94.9
Madhya Pradesh	63.8	90.4	80.3	91.3	64.6	90.4	73.0	88.1
Maharashtra	78.7	107.4	87.6	108.4	92.6	107.4	97.6	109.8
Orissa	93.3	90.3	100.3	85.3	91.5	90.3	99.3	84.1
Punjab	88.7	120.4	95.1	110.3	91.9	120.4	104.7	109.7
Rajasthan	100.6	105.8	97.7	97.5	114.5	105.8	99.8	96.6
Tamil Nadu	107.7	97.8	90.4	96.1	88.8	97.8	95.1	91.5
Uttar Pradesh	83.8	96.4	98.5	91.3	87.8	96.4	97.0	91.4
West Bengal	133.1	98.6	125.6	98.3	129.0	98.6	113.5	94.9
All India	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
CV	0.26	0.11	0.18	0.08	0.22	0.11	0.15	0.09

Note: CV denotes the coefficient of variation. The all-India values are population-weighted averages of state-level prices, normalized to 100.

deviate because of both spatial and intertemporal dimensions.

Table 4 compares the poverty rates derived from Engel analysis with those based on the old and new official methodology. This comparison reveals three main findings. First, the Engel method suggests more geographical variation in poverty than either of the official measures. This is true for both sectors and both time periods. Second, there are consistently higher concentrations of poverty in eastern India, in states such as Assam, Bihar, Orissa and West Bengal. In each of these states, more than 50 per cent are classified as poor. Third, the Engel estimates indicate that most areas experienced some poverty alleviation over the five-year period. However, the reduction is substantially more modest than the one suggested by the official measures.

TABLE 4: Head-count ratios

	2004–05						**2009–10**			
	Rural			**Urban**			**Rural**		**Urban**	
	Engel	*IPC*	IPC_{old}	*Engel*	*IPC*	IPC_{old}	*Engel*	*IPC*	*Engel*	*IPC*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Andhra Pradesh	32.1	32.3	10.5	20.3	23.4	27.4	27.8	22.7	15.7	17.7
Assam	85.8	36.2	22.1	44.8	21.8	3.6	73.7	39.9	42.4	25.9
Bihar	82.2	55.7	42.7	60.4	43.7	36.1	76.9	55.3	64.8	39.4
Chhattisgarh	29.3	55.1	40.8	34.9	28.4	42.2	17.3	56.1	12.5	23.6
Gujarat	37.4	39.1	18.9	19.5	20.1	13.3	46.6	26.6	19.4	17.7
Haryana	11.6	24.8	13.2	15.6	22.4	14.5	26.4	18.6	26.5	23.0
Jharkhand	78.1	51.8	46.3	36.9	23.8	20.3	40.1	41.3	46.9	31.0
Karnataka	23.4	37.5	20.7	18.8	25.9	32.6	34.5	26.1	16.6	19.6
Kerala	8.3	20.2	13.2	22.7	18.4	20.0	3.5	12.0	18.2	12.1
Madhya Pradesh	17.4	53.6	36.8	25.5	35.1	42.7	18.3	41.9	16.7	23.4
Maharashtra	20.3	47.9	29.6	15.4	25.6	32.1	20.2	29.5	14.2	18.3
Orissa	63.0	60.8	46.9	46.4	37.6	44.7	53.2	39.2	40.7	25.9
Punjab	4.6	22.1	9.0	10.1	18.7	6.3	3.1	14.6	17.5	17.9
Rajasthan	30.0	35.8	18.3	29.6	29.7	32.3	33.9	26.4	24.2	19.9
Tamil Nadu	46.3	37.5	23.0	16.4	19.7	22.5	20.5	21.2	16.8	12.8
Uttar Pradesh	27.8	42.7	33.3	38.2	34.1	30.1	32.4	39.3	38.8	31.7
West Bengal	66.8	38.3	28.4	40.4	24.4	13.5	70.3	28.8	36.2	21.9
All India	39.7	41.8	28.3	25.6	25.7	25.7	37.7	33.3	23.9	20.9

Note: The head-count ratio measures the percentage of the population with an expenditure level below the poverty line. The all-India rates are weighted averages of the state-level poverty head counts, using the NSS multipliers.

5 Robustness analysis

In this section, we undertake several robustness checks. All confirm our main findings. First, we estimate the Engel relationship using data on *all* available households rather than data on only those comprising two children and two adults. Second, we evaluate our log-linear functional form assumption by conducting a semi-parametric analysis. Third, we obtain estimates from a more flexible quadratic specification. The key regression coefficients from the first and third robustness checks are reported in Table 5.

5.1 The AIDS model for all households

Our main specification is estimated using data on households comprising two children and two adults only. Relaxing this restriction increases the sample size considerably. Compared with those from the main specification, the all-households regression estimates generate similar expenditure elasticities for the urban sector, but slightly larger elasticities for the rural sector.[17] As we found from the main specification, the price of food relative to non-food items has an insignificant effect in the rural sector and a significantly positive effect in urban areas.

The implied spatial prices are presented in the second and fifth columns of Table 6. Although the point estimates differ somewhat from the indices obtained from the main analysis, the major geographical patterns remain the same. The second row of Table 7 presents the implied all-India intertemporal price indices. These values are somewhat lower than those obtained from our main specification, but they indicate, as before, higher cost-of-living increases than suggested by the official measures. Thus, our sampling restrictions do not drive our main findings.

[17]The extended sample's expenditure elasticity of +0.79 versus +0.77 from the main specification.

5.2 Functional form analysis

Our main estimation relies on a log-linear relationship between the budget share for food and total expenditures. To investigate whether this functional form appropriately explains the observed behaviour, we present estimates from a semi-parametric kernel analysis. The analysis is based on removing any effects on the budget share for food from the covariates used in the main analysis, except that from the logarithm of nominal expenditures. This is done by taking differences, using the tenth-order optimal differencing weights proposed by Yatchew (2003).

The residual parts of the budget shares—the parts not explained by the covariates—are plotted against the logarithm of nominal expenditures in Figure 1, separately for each of the major states and time periods. So, while the analysis forces the partial effects of the covariates to be linear and similar over time and between states, the effect of the log of expenditure is allowed to have a more flexible functional form and to vary across states. However, the plotted lines are close to being log-linear and there is little variation, both over time and between states. Hence, the kernel analysis suggests that our main results are not driven by our functional form assumptions.

To investigate our functional form assumptions further, we estimate the QAIDS and present subsequent alternative price measures. The QAIDS counterpart to the AIDS empirical specification, given by (Equation 1), can be written as:

$$m_{hdst} = a + b_1(\ln y_{hdst} - \ln P_{st}) + b_2(\ln y_{hdst} - \ln P_{st})^2 + \gamma(\ln P^f_{dst} - \ln P^n_{dst}) + \theta X_{hdst} + \epsilon_{hdst}. \quad (5)$$

The overall price component, P_{st}, is identified directly using non-linear iteration and state- and time-specific dummy variables. As in the main analysis, we restrict the sample to households comprising two children and two adults. For both urban and rural sectors, the coefficients for the squared expenditure terms are statistically significant but small. The other coefficients are comparable with those from the linear specification.

The third and sixth columns of Table 6 report the corresponding spatial price measures, which confirm, or even strengthen, our first two findings. There is more price dispersion between states than is implied by the official measures, and the price indices indicate a relatively high cost of living in the eastern states. The third row of Table 7 reports the implied intertemporal price measures. These are very similar to those from our main specification. This similarity validates our third main finding: the official poverty lines understate the cost-of-living increase over our study period. Overall, the similarities with the two sets of price indices from the linear specification are consistent with the small size of the coefficients for squared expenditures.

Table 5: Regressions for robustness checks

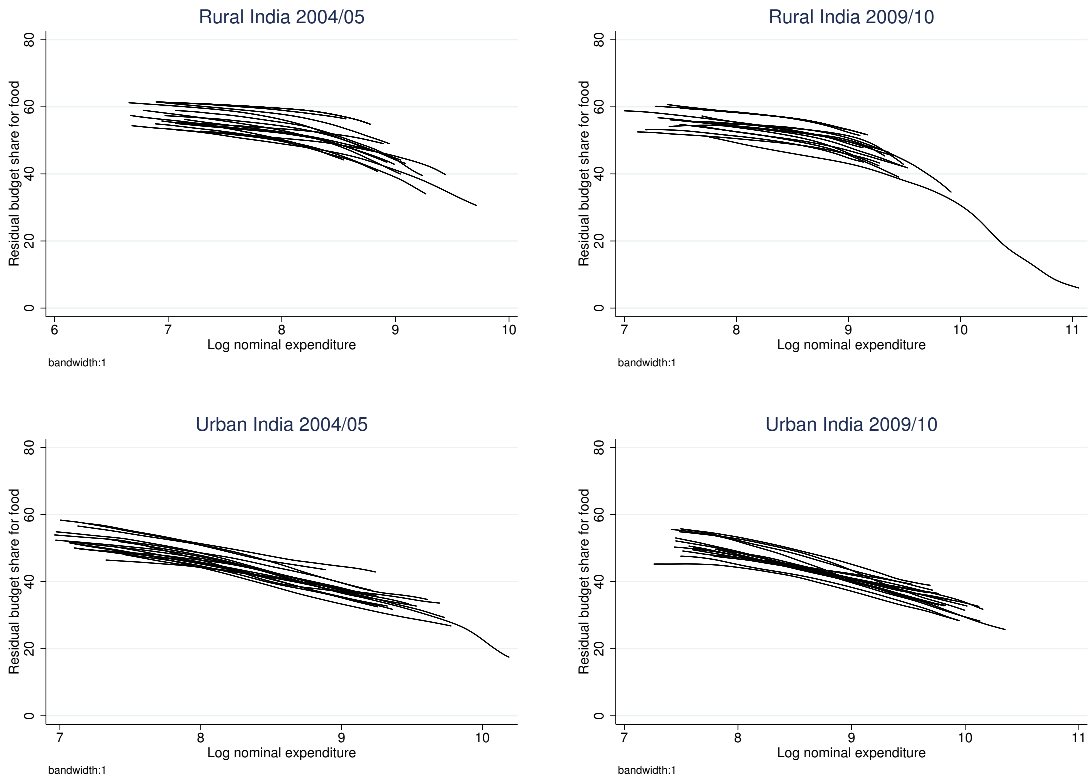
	Rural		**Urban**	
Dep. var.: Budget share for food (percentage)	AIDS_{full}	QAIDS	AIDS_{full}	QAIDS
Log of household expenditure	-11.163	38.083	-13.103	6.165
	(0.0852)	(2.523)	(0.079)	(3.001)
Log of household expenditure squared		-3.216		-1.220
		(0.161)		(0.183)
Log of relative food/non-food prices	0.267	0.123	2.670	1.868
	(0.215)	(0.595)	(0.284)	(0.769)
Observations	137152	14260	85300	9110
R^2	0.348	0.395	0.472	0.525

Note: Standard errors are given in parentheses. Controls that are included but not shown in the table are: the age of the household head; the proportion of females in the household; three occupation dummies for each sector (urban and rural); three religion dummies; the number of free meals taken outside the home; and dummies for cultivated land; In addition, AIDS_{full} includes controls for the number of children and the number of adults, and their squares.

6 Evaluating the poverty outcomes

The Engel method yields spatial prices that differ—substantially for some states—from the official measures. In particular, the measured price levels, and thus the measured poverty rates, are consistently higher for some states in eastern India: Assam, Bihar, Orissa and West Bengal, especially in the rural sector. In this section, we show that an alternative behaviour-based measure, and a measure of self-reported hunger, support these updated estimates.

Figure 1: Semi-parametric analysis



Note: The figures display estimates from the Epanechnikov kernel smoother, obtained using a bandwidth of unity and based on data on households comprising two children and two adults from the 17 major states. For the purpose of presentation, the figures are constructed based on excluding the top and bottom one per cent of the expenditure distribution in each state and sector. The line with the long tail in the upper-right figure represents Kerala. The tail is essentially driven by nine households for which the log of expenditure is 10 or more.

Table 6: Spatial prices indices—robustness

	Spatial prices, 2004–05			Spatial prices, 2009–10		
	AIDS	$AIDS_{full}$	*QAIDS*	*AIDS*	$AIDS_{full}$	*QAIDS*
	(1)	(2)	(3)	(4)	(5)	(6)
Rural						
Andhra Pradesh	96.9	99.4	92.1	102.6	105.1	101.3
Assam	171.4	183.0	186.5	141.9	165.3	152.4
Bihar	126.4	129.8	133.8	117.1	122.3	122.5
Chhattisgarh	71.0	76.3	68.0	62.6	60.6	65.4
Gujarat	109.7	111.4	104.1	122.6	116.7	115.9
Haryana	97.9	98.3	99.3	123.5	119.0	124.1
Jharkhand	117.0	116.9	132.8	85.1	103.1	84.4
Karnataka	82.7	84.5	78.9	95.3	86.4	93.4
Kerala	93.5	89.1	94.1	82.0	75.8	91.8
Madhya Pradesh	63.8	59.8	62.6	64.6	65.2	66.3
Maharashtra	78.7	71.4	76.5	92.6	81.6	89.8
Orissa	93.3	101.7	98.1	91.5	95.2	94.2
Punjab	88.7	87.1	89.0	91.9	95.9	90.3
Rajasthan	100.6	98.4	97.9	114.5	103.4	111.1
Tamil Nadu	107.7	100.2	105.7	88.8	82.6	85.5
Uttar Pradesh	83.8	81.3	80.8	87.8	90.3	85.0
West Bengal	133.1	139.2	135.6	129.0	131.5	131.1
All India	100.0	100.0	100.0	100.0	100.0	100.0
CV	0.26	0.29	0.30	0.22	0.26	0.23
Urban						
Andhra Pradesh	91.7	90.1	90.5	98.0	101.8	98.1
Assam	158.5	147.4	160.5	129.1	125.5	130.2
Bihar	121.8	130.7	123.9	117.2	113.2	119.4
Chhattisgarh	96.2	83.7	95.8	68.1	70.4	70.7
Gujarat	112.7	115.0	111.6	108.8	108.8	107.4
Haryana	94.5	93.6	95.5	114.6	108.2	115.1
Jharkhand	117.8	132.5	121.3	114.2	118.3	116.7
Karnataka	89.5	88.1	88.4	94.2	98.2	92.9
Kerala	110.4	100.5	108.6	101.4	93.2	101.3
Madhya Pradesh	80.3	76.9	79.8	73.0	69.9	73.3
Maharashtra	87.6	89.0	87.9	97.6	99.1	97.9
Orissa	100.3	113.1	101.0	99.3	101.0	99.2
Punjab	95.1	94.2	95.7	104.7	105.8	104.1
Rajasthan	97.7	101.5	98.0	99.8	104.2	99.6
Tamil Nadu	90.4	95.3	89.8	95.1	92.1	94.7
Uttar Pradesh	98.5	96.2	99.3	97.0	94.7	96.7
West Bengal	125.6	122.4	124.9	113.5	110.8	114.3
All India	100.0	100.0	100.0	100.0	100.0	100.0
CV	0.18	0.19	0.19	0.15	0.15	0.14

Table 7: All-India intertemporal prices

	Rural	Urban
Engel		
AIDS	161.4	158.2
AIDS$_{full}$	155.1	155.6
QAIDS	158.8	158.3
UV (IPC)	150.6	150.3
CPI*	155.0	145.1

Note: * The CPIAL and CPIIW are used for the rural and urban sectors, respectively.

6.1 Calorie composition

An adequate intake of calories and nutrition is a central feature of individual well-being. This is why perceived calorie needs were given a key role in defining the Indian poverty lines in the 1970s. However, under the new methodology, the poverty counts are derived with no reference to calorie intake, partly because calorie thresholds are difficult to define. There exists no consensus on what is needed for subsistence. Moreover, any calorie norm is likely to vary considerably both across individuals and over time; calorie needs depend on characteristics such as age, sex and occupation, and also on factors not typically recorded in household surveys, such as body mass, stress levels and altitude.

Thus, a potentially more robust way of identifying hunger and starvation is to study the *composition* of individuals' calorie intakes. If we believe that poor families, living in hunger, are likely to maximize their calorie intakes, one would also expect them to rely on cheap sources of calories. Better-off households are likely to substitute away from the cheapest sources towards calories that come from more expensive foods that have other attributes such as taste (Behrman and Deolalikar, 1988). Thus, it seems plausible that there exists a negative relationship between a household's share of its total calories that come from cheap sources and household real income. Jensen and Miller (2010) formalize this simple and intuitive idea within a theoretical consumer choice framework.

In Appendix B, we plot the share of total calories from cereals versus the logarithm of total expenditure.[18] Of the main food groups reported in the NSS data, cereals have by far the

[18]For each food item, the calorie data are obtained by multiplying the quantity consumed by the corresponding calorie conversion factor from the NSS. These widely used factors are based on work by

cheapest price per calorie. Because cereals are also widely consumed all over India, using this food group seems appropriate to proxy "cheap calories". Not surprisingly, we find a monotonic negative relationship between cereal shares and the log of total expenditure.

In what follows, we use this monotonic negative relationship to evaluate the validity of the Engel-based and official poverty counts. We do this by looking more closely at households with expenditure levels that are 2.5 per cent above and below the poverty lines. *If* the state-wise poverty lines represent the same real expenditure level across states, one would also expect these households to have similar cereal calorie shares, despite the fact that their nominal expenditure levels vary. This hypothesis is investigated in Figure 2. Because the figure is based on households within a limited range of the expenditure distribution, we restrict the analysis to the 12 states with the most rural households in the NSS data. These are the states with more than 2000 rural households in both survey rounds.[19]

The first two graphs in the figure display simple fitted lines for the cereal calorie shares and nominal expenditure levels, separately for the rural and urban sectors. The fitted lines representing families close to the Engel lines are almost horizontal. Thus, there are hardly any systematic differences between these households; i.e., they seem to behave as if they were equally poor. Interestingly, households from states such as Assam, Bihar, West Bengal and Orissa, which should have relatively high nominal poverty lines according to the Engel analysis, do by no means diverge from the other households.

The figure also graphs corresponding fitted lines for families around the official poverty lines, based on the new and old methodologies. These households do not behave as if they were equally poor. In particular, based on their higher cereal shares, households from Assam, Bihar, West Bengal and Orissa seem to act as if they were poorer than households close to the poverty lines in other states.[20]

The second set of graphs in Figure 2 presents an alternative set-up for evaluating the calorie composition of households close to the poverty lines. Even though the comparison

the National Institute of Nutrition (Gopalan *et al.*, 1971).

[19]For consistency, we use the same 12 states for the urban sector. With a few exceptions, these are the states with the most urban households.

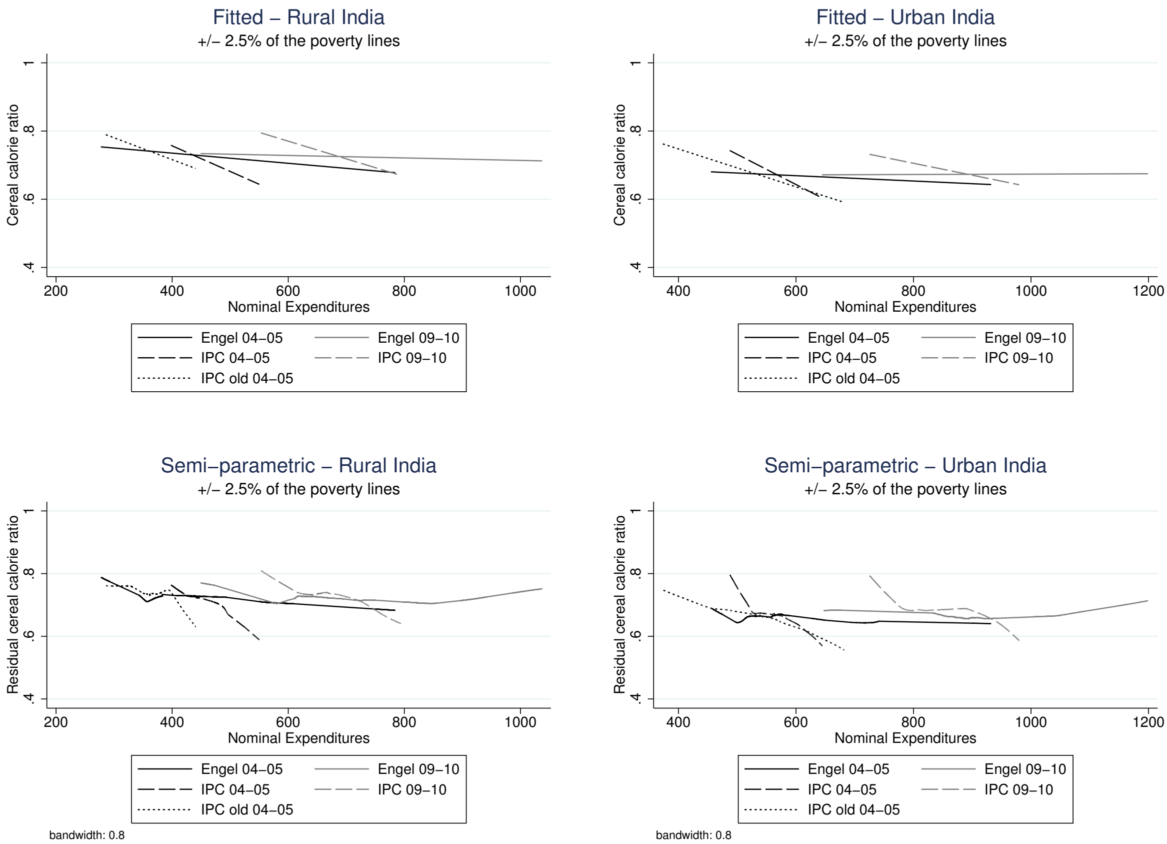
[20]Alternatively, we could have used data from all the 17 major states. This changes all the slope coefficients somewhat. However, households close to the Engel poverty lines are still substantially more homogenous in terms of cereal shares than those close to either the new or the previous official lines.

of cereal calorie ratios is likely to be relatively robust to variations in household characteristics, we cannot exclude the possibility that the observed pattern stems from systematic differences across states in household characteristics and/or in relative food prices. To investigate this more carefully, we undertake a semi-parametric analysis.

The controls comprise variables capturing demographics, occupation and the numbers of meals taken outside the home (for which we do not observe the calorie content). In addition, we construct a relative cereal calorie price index similarly to how we constructed the food/non-food price index, as explained in Section 3.2. Finally, we include a dummy variable for households consuming more than 50 per cent of their calories through the Public Distribution System (PDS). The PDS is a potential source of interstate variation in cereal consumption. This is because the PDS is implemented unevenly across the country and is centred around the distribution of cheap rice and wheat. However, our main interest is in the nominal expenditure variable, which is shown non-parametrically in the bottom two graphs. Not surprisingly, the conclusion drawn from the fitted lines remains unchanged: there are no systematic differences across households close to the Engel-based poverty lines. Moreover, this similarity in calorie composition is not apparent for households close to either of the official poverty lines. This indicates that the official methodologies fail to capture real cost-of-living differences across Indian states.

Similarly, given that we expect households to decrease their cereal calorie shares as they grow richer, we also expect them to *increase* their shares of more expensive calories. Thus, consumption of expensive calories could potentially be used to perform an additional validity check of the poverty estimates. However, this sort of test requires a food item whose relative consumption increases monotonically with real income. It must also be consumed by most households all over India. Edible oils is a potential candidate. In Appendix B we show that there is a clear positive relationship between edible oil calorie shares and the logarithm of total expenditure. However, this relationship is much weaker than the one for cereals, which leads to a weaker test of the poverty counts. Recall that Figure 2 exhibits a downward sloping relationship between cereal shares and nominal expenditure levels for households around the official poverty lines. Based on this finding, and the hypothesis that the estimates fail to identify equally poor households across states, we expect to observe an upward sloping relationship between edible oil shares and total

Figure 2: Cereal calorie shares and nominal expenditure levels



Note: The graphs in the upper part of the figure display simple fitted lines using only observations on households with expenditure levels that are 2.5 per cent above and below the relevant poverty line. The graphs in the lower part of the figure display estimates from the Epanechnikov kernel smoother, using the same set of households and a bandwidth of 0.8.

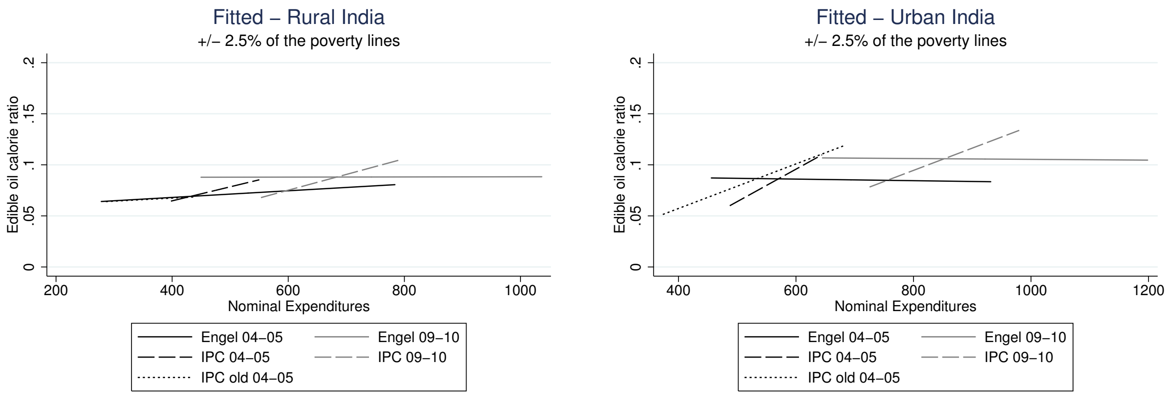
expenditure for the same set of households. This is exactly what is shown in Figure 3. By contrast, the Engel estimates provide no indication of any systematic differences across states.[21]

6.2 Self-reported hunger

As a second validity check of our estimated poverty counts, we investigate the households' own perceptions of hunger. In the NSS survey, respondents are asked whether every

[21]Sugar is another suitable candidate for an expensive source of calories. However, the relationship between sugar and total expenditure is not as marked as the one between edible oils and expenditure. Hence, using sugar data would provide a more 'noisy' validity check. Nevertheless, a comparison of the sugar calorie shares for households around the various poverty lines yields roughly similar results to those based on edible oils.

FIGURE 3: Calorie shares from edible oils & fats and nominal expenditure levels



member of the household gets "enough food every day". Given that this is a *self-reported* measure of hunger, it should be interpreted as such.[22] However, a priori, we have little reason to expect any systematic errors across states.

For each state and sector, Figure 4 shows the proportion of all households reporting lack of food.[23] These numbers are plotted against two sets of head-count ratios: those from the Engel analysis and the new official poverty rates. For presentational purposes, we combine the rural and urban head counts, using population weights. The graphs show that self-reported hunger rates are consistently below the poverty rates. However, given that poverty and hunger are different concepts, there is no reason to expect them to coincide, although one would expect them to be correlated. Interestingly, the graphs reveal that four of the five states with the highest levels of self-reported hunger are Assam, Bihar, Orissa and West Bengal. This finding is consistent with the Engel-based poverty counts.

Table 8 shows that the Engel head counts are quite strongly correlated with the self-reported hunger ratios. This supports our spatial poverty pattern. Although the poverty counts based on the new official methodology are positively correlated with the self-reported hunger ratios, the correlations are smaller than those for the Engel counts. Strikingly, for urban areas, there is a negative correlation between the hunger rates and the head counts based on the old official methodology. This constitutes further evidence

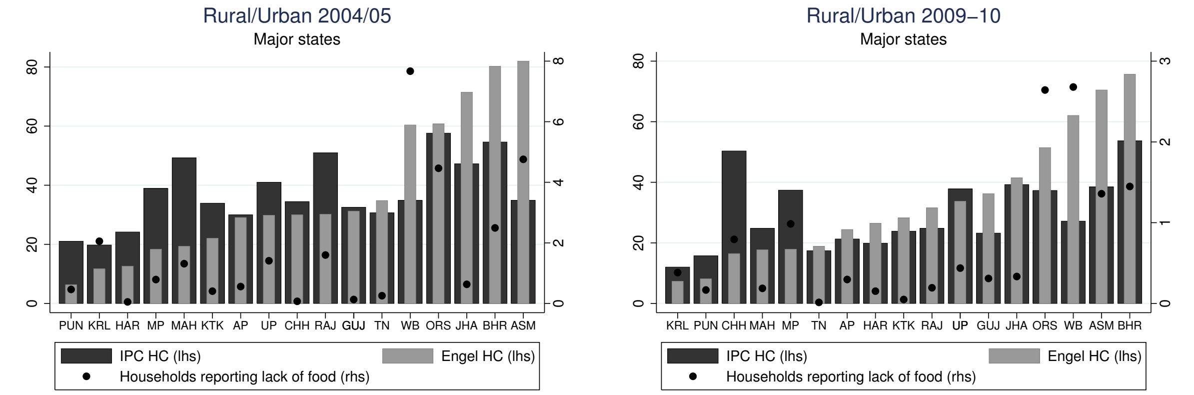
[22]See, e.g., Deaton and Tarozzi (2000) for a critical evaluation of this subjective measure.
[23]These proportions are taken from the "type 2" NSS survey, because, for some unknown reason, the question is excluded from the "type 1" survey.

Table 8: Correlations between self-reported hunger and head counts

Correlation coefficients						
	Engel		**IPC**		**IPC$_{old}$**	
	Rural	Urban	Rural	Urban	Rural	Urban
2004–05	0.55	0.73	0.20	0.28	0.30	-0.15
2009–10	0.64	0.53	0.41	0.42		

that the old IPC measures are misleading and out of date.

Figure 4: Self-reported hunger and head counts of poverty



Note: Hunger ratios are plotted against head-count ratios. The ratios are defined as the share of households reporting an inadequate level of food consumption. The hunger questions from the 61st and 66th NSS surveys are not entirely consistent with each other. In NSS61, respondents are asked "Do all members of your household get enough food every day?", and are asked to choose between: "yes: every month of the year"; "some months of the year"; and "no: no month of the year". In NSS66, respondents are asked "Do all members of your household get two square meals every day?, and are asked to choose between: "yes: every month of the year"; "some months of the year"; and "no: no month of the year". This discrepancy could explain the relatively large drop in the number of households reporting hunger over time. However, the discrepancy is not a concern because we do not compare households between survey rounds.

7 Concluding remarks

In this paper, we advocate an alternative approach for price and poverty comparisons, in which price levels are indirectly estimated based on the behavioural assumption that equally poor households spend the same proportion of their incomes on food.

Our analysis reveals three main findings. First, our estimated price differences between Indian states exceed those suggested by the official poverty lines. Second, although the recently adopted official methodology seems more credible than the previous one, it still underestimates the degree of poverty in eastern India. Third, our estimate of the decrease in poverty from 2004–05 to 2009–10 is more modest than the official estimate. These three findings are robust to various empirical specifications.

We find that the behaviour of households close to the poverty lines supports our derived spatial poverty profile. These families consume similar shares of calories from cereals. This is consistent with the idea that poor households are likely to rely on the cheapest sources of calories. The same similarity is not found for families close to the official lines. We also find that rates of self-reported hunger are much more highly correlated with the Engel poverty counts than with the official figures.

References

Almås, I. (2012). International income inequality: Measuring ppp bias by estimating engel curves for food. *American Economic Review*, **102**(1), 1093–1117.

Banks, J., Blundell, R., and Lewbel, A. (1997). Quadratic engel curves and consumer demand. *Review of Economics and Statistics*, **79**(4), 527–539.

Barrett, G. and Brzozowski, M. (2010). Using engel curves to estimate the bias in the australian cpi. *The Economic Record, The Economic Society of Australia*, **86**(272), 1–14.

Beatty, T. and Larsen, E. (2005). Using engel curves to estimate bias in the canadian cpi as a cost of living index. *Canadian Journal of Economics/Revue canadienne d'économique*, **38**(2), 482–499.

Behrman, J. and Deolalikar, A. (1988). Health and nutrition. *Handbook of development economics*, **1**, 631–711.

Bhattacharya, N. and Chatterjee, G. (1974). Between state variations in consumer prices and per capita household consumption in rural india. *Poverty and Income Distribution in India. Calcutta: Statistical Publishing Society*.

Blundell, R., Duncan, A., and Pendakur, K. (1998). Semiparametric estimation and consumer demand. *Journal of Applied Econometrics*, **13**(5), 435–461.

Carvalho Filho, I. and Chamon, M. (2006). *The myth of post-reform income stagnation in Brazil*. Number 2006-2275. International Monetary Fund.

Chung, C., Gibson, J., and Kim, B. (2010). Cpi mismeasurements and their impacts on economic management in korea. *Asian Economic Papers*, **9**(1), 1–15.

Costa, D. (2001). Estimating real income in the united states from 1888 to 1994: Correcting cpi bias using engel curves. *Journal of political economy*, **109**(6), 1288–1310.

Deaton, A. (1988). Quality, quantity, and spatial variation of price. *The American Economic Review*, pages 418–430.

Deaton, A. (2003). Prices and poverty in india, 1987-2000. *Economic and Political Weekly*, pages 362–368.

Deaton, A. (2008). Price trends in india and their implications for measuring poverty. *Economic & Political Weekly*, pages 43–49.

Deaton, A. and Muellbauer, J. (1980a). An almost ideal demand system. *The American Economic Review*, **70**(3), 312–326.

Deaton, A. and Muellbauer, J. (1980b). *Economics and Consumer Behavior*. Cambridge University Press.

Deaton, A. and Tarozzi, A. (2000). Prices and poverty in india. *Princeton, July*.

Engel, E. (1857). Die productions-und consumtions-verhältnisse des königreichs sachsen. *Bulletin de lInstitut International de la Statistique*, (9).

Engel, E. (1895). *Die Lebenskosten belgischer Arbeiter-Familien früher und jetzt*. C. Heinrich.

Gibson, J., Stillman, S., and Le, T. (2008). Cpi bias and real living standards in russia during the transition. *Journal of Development Economics*, **87**(1), 140–160.

Gong, C. and Meng, X. (2008). *Regional Price Differences in Urban China 1986-2001: Estimation and Implication*. IZA.

Gopalan, C., Sastri, B., and Balasubramanian, S. (1971). Nutritive values of indian foods. *Hyderabad, National Institute of Nutrition*.

Government of India (1979). Report of the task force on projections of minimum need and effective consumption demand.

Government of India (1993). Report of the expert group on estimation of the proportion and number of poor.

Government of India (2009). Report of the expert group to review the methodology for estimation of poverty.

Government of India (2011). Press note on poverty estimates.

Hamilton, B. (2001). Using engel's law to estimate cpi bias. *The American Economic Review*, **91**(3), 619–630.

Jensen, R. and Miller, N. (2010). A revealed preference approach to measuring hunger and undernutrition. Technical report, National Bureau of Economic Research.

Larsen, E. (2007). Does the cpi mirror the cost of living? engel's law suggests not in norway. *The Scandinavian Journal of Economics*, **109**(1), 177–195.

Leser, C. (1963). Forms of engel functions. *Econometrica: Journal of the Econometric Society*, pages 694–703.

Minhas, B.S., K. S. and L.R, J. (1989). Incidence of urban poverty in different states. *ISI Delhi*, **Technical Report No.8902**.

Olivia, S. and Gibson, J. (2012). Using engel curves to measure cpi bias for indonesia. *Monash Econometrics and Business Statistics Working Papers*, **13**.

Rao, P. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American statistical association*, pages 163–171.

Srinivasan, T. (2007). Poverty lines in india: Reflections after the patna conference. *Economic and Political Weekly*, pages 4155–4165.

Subramanian, S. (2011). The poverty line: Getting it wrong again. *Economic & Political Weekly*, **46**(48), 37.

Working, H. (1943). Statistical laws of family expenditure. *Journal of the American Statistical Association*, **38**(221), 43–56.

Yatchew, A. (2003). *Semiparametric regression for the applied econometrician*. Cambridge Univ Pr.

Appendix A Relative food and non-food prices

The 61st NSS survey round provides information on quantities and values for 187 consumption goods. Eleven of these items are found to be of insignificant value and are excluded in the Tendulkar Expert Group methodology.[24] We exclude the same 11 items. To derive a comparable set of unit value items for the two survey rounds, we make additional adjustments. Items that appear in the questionnaires from the 66th round but not in those from the 61st are either excluded or aggregated with a relevant item. Items without readily available quantities in either of the two rounds, or items with non-comparable unit measures, are excluded.[25] Furthermore, following the Tendulkar methodology, we aggregate PDS items with the relevant non-PDS items before calculating unit values.[26]

Overall, we end up calculating unit values for 127 food items and 41 non-food items. The relative food and non-food price index is calculated at the district level, separately for the rural and urban sector. There are some discrepancies between the district divisions in the two NSS survey rounds. For the 30 largest states included in our main regression, there are 24 more districts in the 66th round compared with the 61st.[27] Moreover, some districts have no surveyed households in one of the two sectors. Thus, we end up computing the relative price index for 578 rural districts in NSS 61 versus 595 in NSS 66, and 573 urban districts in NSS 61 versus 587 in NSS 66.

Because items included in the surveys are different in nature, there is no single suitable unit of measurement for all of them. Hence, we must work with different units for different items.[28] What is important is that we use the same units of measurements for all districts

[24] These are khoi, barley, singara, berries, misri, ice, katha, snuff, cheroot, ganja and cotton.

[25] Ice cream, other milk products and other intoxicants are excluded because of missing quantities in NSS66. Dhoti and sari are excluded because of non-comparable units (meters in NSS61 and numbers in NSS66). Soya beans are excluded from NSS61 because of their exclusion from NSS66. Petrol and diesel are excluded from NSS66 because of their exclusion from NSS61. Supari and lime are aggregated into other ingredients for pan in NSS61 because of their exclusion from NSS66. Second-hand footwear is aggregated into other footwear, and cooked meals received as assistance or payment are aggregated into cooked meals purchased in NSS66.

[26] This applies for rice, wheat, sugar and kerosene.

[27] The numbers of extra districts in each state are as follows: Arunachal Pradesh 3; Assam 4; Bihar 1; Haryana 1; Jharkhand 4; Madhya Pradesh 3; Nagaland 3; Punjab 1; Tamil Nadu 1; Uttarakhand 2; and West Bengal 1.

[28] Food items are typically given in kilograms, except, e.g., eggs and some fruits, which are given in numbers. Textiles are given in meters, whereas shoes are in pairs, and some clothes are reported in numbers.

and for both survey rounds.

Summary statistics are shown in Table 9. The indices used in the analysis are at the district level, but because of space considerations, the table reports average state index values. The table suggests that there are relatively large differences in relative food/non-food prices across Indian states. It can also be seen that food unit values generally increased by more than did non-food unit values during the five-year period from 2004–05 to 2009–10. The relative price index increased by roughly 15 per cent in the rural sector, and by 12 per cent in the urban sector. For comparison, the corresponding ratios increased by 21 per cent and 17 per cent in the CPIAL (rural) and CPIIW (urban) indices, respectively.[29]

TABLE 9: Relative food and non-food prices

	Rural				**Urban**			
	2004–05		2009–10		2004–05		2009–10	
Andhra Pradesh	1.04	(0.12)	1.11	(0.12)	0.96	(0.11)	1.04	(0.11)
Assam	1.09	(0.11)	1.29	(0.10)	1.11	(0.13)	1.29	(0.14)
Bihar	0.89	(0.12)	1.07	(0.09)	0.90	(0.10)	1.01	(0.11)
Chhattisgarh	1.05	(0.06)	1.11	(0.20)	1.05	(0.08)	1.02	(0.14)
Gujarat	1.12	(0.12)	1.29	(0.15)	1.03	(0.10)	1.22	(0.12)
Haryana	0.99	(0.11)	1.17	(0.12)	0.94	(0.11)	1.11	(0.13)
Jharkhand	1.02	(0.10)	1.15	(0.09)	1.02	(0.15)	1.01	(0.10)
Karnataka	0.93	(0.07)	1.06	(0.13)	0.96	(0.07)	1.01	(0.12)
Kerala	1.07	(0.12)	1.15	(0.10)	1.01	(0.09)	1.06	(0.10)
Madhya Pradesh	1.01	(0.12)	1.12	(0.17)	1.04	(0.13)	1.04	(0.13)
Maharashtra	1.11	(0.15)	1.23	(0.18)	1.07	(0.12)	1.19	(0.12)
Orissa	0.98	(0.10)	1.23	(0.20)	0.97	(0.09)	1.12	(0.10)
Punjab	0.83	(0.10)	1.03	(0.16)	0.84	(0.09)	1.01	(0.16)
Rajasthan	0.95	(0.15)	1.24	(0.14)	0.93	(0.12)	1.12	(0.12)
Tamil Nadu	1.17	(0.11)	1.21	(0.10)	1.13	(0.09)	1.24	(0.12)
Uttar Pradesh	0.96	(0.10)	1.13	(0.11)	0.96	(0.13)	1.16	(0.15)
West Bengal	0.96	(0.07)	1.11	(0.10)	0.95	(0.05)	1.15	(0.08)
All India	1.00	(0.14)	1.15	(0.15)	1.00	(0.13)	1.12	(0.15)

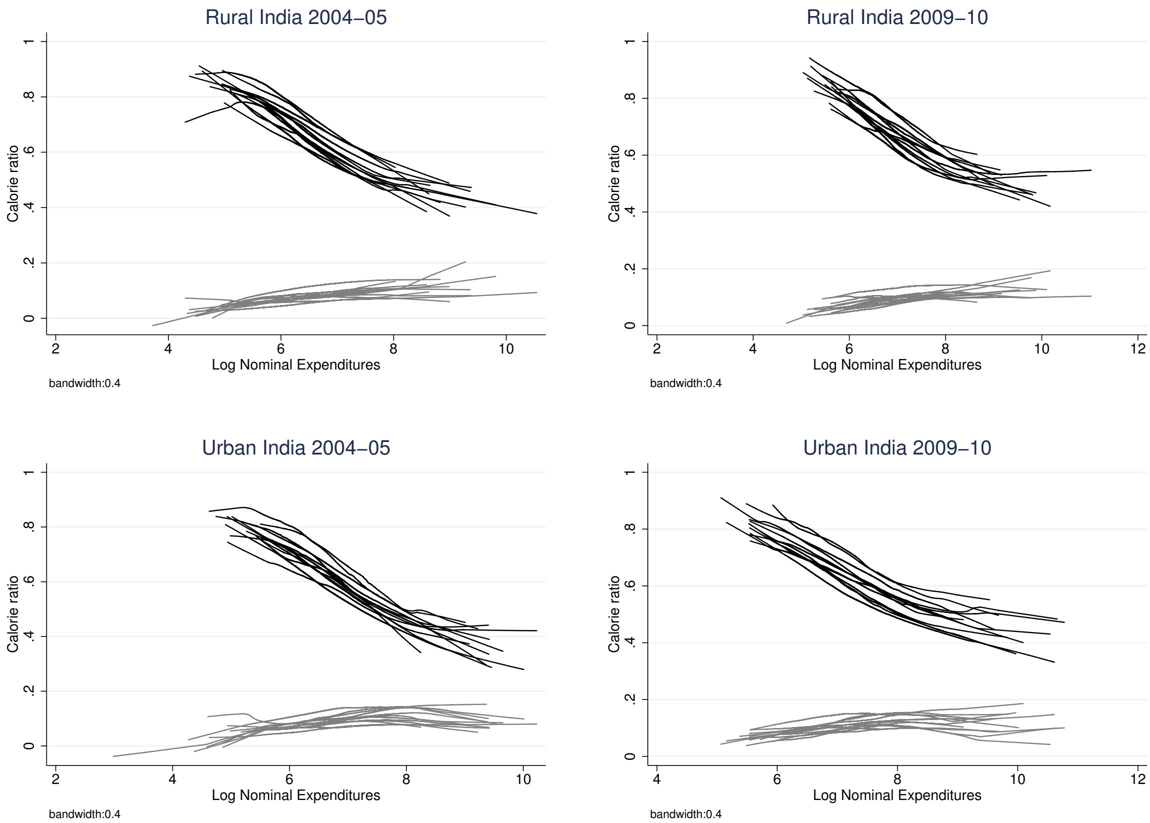
Note: The table shows food unit values relative to non-food unit values. All values given are population-weighted state averages, obtained using the multipliers from the NSS data. The weighted all-India average for 2004–05 is normalized to unity for the rural and urban sectors separately. The standard deviations are clustered at the district level.

[29]These figures are obtained by comparing the food component with the non-food component in the two price indices.

Appendix B Calorie consumption

Figure 5 shows the proportion of all calories consumed obtained from cereals (black lines) and edible oils (grey lines), separately for each sector and time period, for all states used in the calorie analysis is Section 6.1.

Figure 5: Cereal calorie shares and the log of nominal expenditure



Note: The figures display estimates from the Epanechnikov kernel smoother, using a bandwidth of 0.4.