NORGES HANDELSHØYSKOLE

Bergen, 11.12.2009

Cost of Living Differences and Poverty Reduction

New Evidence for India Based on Micro Data

Anders Grøn Kjelsrud

Supervisor: Ingvild Almås

Master Thesis, Department of Economics

NORWEGIAN SCHOOL OF ECONOMICS

AND BUSINESS ADMINISTRATION

Abstract

The Indian Planning Commission's (IPC) state-wise poverty lines are widely used. Yet, the underlying price adjustments are based on outdated studies as well as consumer price indices (re-weighted versions of the CPIAL and the CPIIW), which may not serve as good proxies for the cost of living. Hence, we have reasons to expect the poverty measures to be biased.

This thesis adopts the Engel methodology proposed by Hamilton (2001), and fully incorporates the quadratic extension suggested by Costa (2001), a thanks goes to everyoneto estimate new sets of consistent cost of living indices. Subsequently, new poverty estimates are provided. My findings suggest that the official price indices have overstated the increase in cost of living in the mid eighties and early nineties, and hence, understated the poverty decline. The Engel methodology also suggests larger state-wise price differences than implied by the official state-wise poverty lines. (*JEL*: D1, E31, F01, I32)

Acknowledgement

I would like to thank my supervisor, Ingvild Almås, for her useful guidance during the process of writing this master thesis. I will also like to thank her for giving me the opportunity to work with this project during the summer of 2009. Second, I would like to thank Professor Rohini Somanathan for helpful comments, in addition to providing me with the data used in this thesis. Finally, a thanks goes to everyone that has been giving me comments and advices under the way.

Bergen, December 2009

Anders Grøn Kjelsrud

Contents

List of Tables

1 Introduction

Poverty reduction is central to achieving many of the outcomes of the United Nations (UN) millennium goals. According to the UN's measures, India consists of 17 percent of the total world population[1] and has a large share of the world's poor. Hence, the evolution of poverty in India has a great impact on the evolution of global poverty. Since the liberalization of the economy in 1991, India has seen a strong and robust economic growth measured in GDP. However, economic growth provides no guarantee for a corresponding reduction in poverty. Hence, poverty estimates for India are important measures in a global development perspective. Accurate poverty estimates are also central for many domestic Indian issues. They determine the intensity of the poverty reduction effort and resources transferred from the Indian central government to the states. They also help evaluate the performance of state governments. Despite the importance of these measures, there are reasons to expect that the official Indian poverty measures inherit substantial biases, mainly caused by biases in the price indices used for comparison. The price indices are based on outdated spatial price indices as well as consumer price indices, which may not serve as good proxies for the true cost of living.

In this master thesis I study the official Indian poverty measures compiled by the Indian Planning Commission (IPC), and calculate a range of new price indices and poverty estimates. I believe these new price indices, and not just the poverty rates have a value in themselves, because price comparisons between states and over time are central in the Indian poverty debate. My main analysis focuses on the eighties and early nineties. The official poverty estimates declined modestly in the eighties, but declined steeply during the nineties. Many attribute this to the liberalization. An important policy issue is to assess whether the poverty reduction started already before the liberalization, or whether it mainly was a consequence of it.

My estimated price and poverty measures are based on estimation of Engel curves

[1] According to the UN's "The World Population Prospects 2004".

for food. Bruce Hamilton (2001) first proposed this method for consumer price index bias measurement in the United States. I provide a reformulation of this method, which enables me to identify price indices that are consistent both across Indian states and over time. I also fully incorporates the quadratic extension suggested by Costa (2001). The identification is based on observed consumer behavior at the household level.

The IPC operates with state-specific poverty lines for both the rural and the urban sector in India. The state- and sector-wise poverty lines are derived through two academic studies, which measured price differences across Indian states in the early 1960s. These spatial poverty lines are deflated over time using state- and sector specific price indices. There are two potential sources of bias in this methodology; one is the potential miss-measurement stemming from the spatial indices constructed by using data from the early 1960's, and the other is the potential miss-measurement in the price indices which is used to transform the spatial differences in the 1960's, into spatial differences for other years. The latter are fixed basket indices of the Laspeyres type. The biases in these types of price indices, such as the substitution bias, the outlet bias, the quality bias, and the new goods bias, are well known in the literature (see e.g., Hausman 2003). All the mentioned biases occur because the price indices fail to be consistent with consumer preferences. As I use Engel's Law, and identifies the price indices through two demand systems with utility maximization under a budget constraint, my method is consistent with consumer preferences.

This thesis reports five main findings. First, there is a larger spread in price levels across Indian states than what is suggested by the IPC, according to the Engel analysis. Second, there is strong correlation between the spatial prices in the rural sector and the urban sector. This is as expected due to arbitrage. Yet the IPC's state-wise poverty lines imply a negative correlation between the spatial prices in the two sectors, clearly suggesting that the official poverty lines are outdated. According to the Engel analysis, the poverty lines are most outdated in the urban sector. Third, my estimates indicate that the price increase in the mid eighties and early nineties is overestimated. Because of this, the officially published figures for the mid eighties and early nineties underestimate the poverty decline. This is most prominent in urban areas. Fourth, the Engel analysis suggests that the official poverty measures for the northeast part of India are too optimistic. This finding is consistent with anecdotes and casual observation from these areas. Finally, official estimates indicate that the poverty decline in the mid eighties is larger than in the early nineties for rural India, whereas the opposite is true for urban India. The Engel methodology however, reveals a much more steady development of decreased poverty in both rural and urban India for the whole period under study.

The indian poverty measures, and especially the prices indices used for comparisons, have been in focus in recent studies (Deaton and Tarozzi 2000; Deaton and Kozel 2005; Himanshu and Murgai 2009; Palmer-Jones and Dubey 2008). A prominent researcher is Professor Angus Deaton. In a series of papers he calculates new sets of state-wise price indices, which is used to update the official poverty measures. All these price indices are based on so-called unit values (UV). Deaton's work is a natural reference point for my estimates, and is described in more detail in Section 2.5.

The rest of the thesis is organized as follows. Section 2 presents the Indian poverty monitoring system, and discusses potential sources for biases as well as alternative approaches for poverty measurement. Section 3 describes two demand systems, which make the theoretical foundations for the cost of living identification discussed in Section 4. In Section 5 the data and descriptive statistics are discussed, while Section 6 presents the estimation results. Section 7 gives calculations of the implied spatial prices, as well as price indices over time and the corresponding poverty measures. Concluding remarks are given in Section 8.

2 Poverty Monitoring in India

The Indian Planning Commission (IPC) is in charge of the official poverty measures in India. Ever since 1962, the IPC have based their measures on consumer expenditure surveys conducted by the National Sample Survey Organization (NSS), i.e., the poverty estimates are based on expenditures and not income. The NSS conduct large household surveys once every fifth year, with smaller annual surveys in the years in between (since the 42nd round in 1986-87). The IPC uses only the large rounds to estimate poverty rates, which means that they are not published on a yearly basis.

2.1 Poverty Measures

The IPC's poverty rates are measured as headcount ratios (HCR) – the ratio of poor people to the total population. A poor household is defined as a household with an expenditure level below a specific poverty line. The main advantage of the HCR is that it is easy to interpret. However, it has some limitations as a poverty index. First, it ignores how far a household falls short of the poverty line. This is clearly problematic, and could in some situations result in undesirable poverty outcomes. For instance, a situation where an already poor household becomes even poorer would not result in any change in the poverty measure. Another example would arise if a very poor household transfers some of its income to a household just below the poverty line. This would result in a decline in the measured poverty, even if most people would agree that the poverty situation has worsened, not improved. Second, the HCR is extremely sensitive to the density of the population around the poverty line. Since changes in poverty over time are driven entirely by the number of households crossing the poverty line, a minor change in income per capita could lead to a spuriously large poverty decline if many household are "bunched" near the poverty line. It is therefore difficult to compare poverty changes in different areas by just looking at the changes in the HCR without further information of the initial density of the population around the poverty lines (Deaton and Dréze 2002).

Even if the use of HCR is problematic in many respects, the IPC still prefers it at the expense of more sophisticated poverty indices such as the Foster-Greer-Thorbecke indices or the Sen index. The choice of poverty measure is not the focus in this thesis, and hence, I base my main poverty estimates on the HCR. However, for comparisons reasons, I also estimate the poverty gap index (PGI). In contrast to the HCR, the PGI adjusts for the depth of poverty in addition to the incident[2]. The index can be interpreted as the aggregated shortfall of the poor's consumption from the poverty line, normalized by the population size. More specifically, it measures the per capita shortfall as a percentage of the poverty line (Deaton and Dréze 2002).

2.2 Evolution of The Indian Poverty Lines

Since the official poverty measures are based on HCRs, the choice of poverty line is an issue of importance. This section gives a brief overview of the evolution of the Indian poverty lines. The IPC has changed its methodology for the construction of poverty lines several times since the 1960s. From the late seventies to the mid-nineties, the IPC used the methodology proposed by The Task Force 1979 (Government of India 1979). The first step in this committee's proposal was to define a minimum requirement for calories. Based on calculated nutrition needs, the "calorie norm" was set to 2.100 and 2.400 calories per person per day for the urban and the rural sector, respectively. The second step was to identify a monthly total expenditure level in each sector, consistent with these calorie norms. For this they used the 28th round of the NSS consumer expenditure survey (1973-74). From the distribution of the calorie intake in each sector, The Task Force identified an average expenditure level, including non-food expenditures, where the calorie requirements were met. These specific expenditure levels happened to be 49 Rupees (Rs) per capita per month and 57 Rs per capita per month in the rural and the urban sector, respectively, and were used as poverty lines. In terms of 1973-74 US Dollars (USD), the poverty lines corresponded to USD 5.9 and USD 6.8[3]. The methodology implied that the chosen poverty lines were partly structural (based on the calorie norm), and partly behavioral, since they used the observed expenditure patterns to identify the equivalents to the calorie norm in terms of Rs. The two poverty lines were held constant in real terms

[2] $PGI \equiv (1/z)[\sum(z-y_i)/n]$, where y_i is the expenditure level of person i, z is the poverty line and n is the population size.

[3] According to United Nations database, the average exchange rate in 1973-74 were 0.12 USD per Rs.

and deflated with the private consumption deflator from the National Accounts Statistics (NAS). Implicitly (since they were calculated separately), the poverty lines implied a price differential between the rural and the urban sector of 14 percent. This price differential turned out to be consistent with independent estimates provided by Bhattacharya and Chatterjee (1971).

In 1989 the IPC constituted an independent expert group to look into potential improvements of the The Task Force's methodology, resulting in the Expert Group Report (Government of India 1993). This panel of different experts recommended to keep the existing poverty lines from 1973-74, but to disaggregate them to the state level to account for different cost of living in the Indian states. In order to achieve state-specific poverty lines that reflected real price differences, The Expert Group recommended to use two earlier academic studies that calculated inter-state price differences relative to all-India. For the rural sector they proposed a study by Bhattacharya and Chatterjee (1974), and for the urban sector a study by Minhas, Jain, Kansal, and Saluja (1988). Both studies were based on household surveys conducted by the NSS in the early 1960s. The second main proposal was to use state- and sector specific price deflators to adjust the poverty lines in the base year, into poverty lines for other years. Based on this proposal, the IPC ended up using re-weighted versions of the consumer price index for agricultural labourers (CPIAL) and the consumer price index for industrial workers (CPIIW) for the rural and the urban sector, respectively[4]. These re-weighted price indices were constructed to better reflect the cost of living of the poor, using national average consumption patterns of people around the poverty line in 1973-74. The modified indices are based on the same price date as the original price indices, but reflect that the poorest have a higher budget share for food than the average population. The proposals from The Expert Group make the basis for today's official poverty lines, and have been used since 1983 (back-casting the methodology).

[4]The Expert Group originally suggested to use a weighted average of the CPIIW and the consumer price index for non-manual employees (CPINM) for the urban sector.

2.3 Poverty Debate

There has been a huge debate regarding the poverty development in India during the 1990s (see e.g., Datt, Kozel, and Ravallion 2003; Deaton and Kozel 2005; Sen and Himanshu 2005). After a modest decline in the 80s, the official poverty rates steeply declined during the 90s. This is especially true for the second part of the decade, where the poverty incident decreased from 36 percent of the total population in 1993-94 to 26 percent of the total population in 1999-00. Several methodology issues have been fiercely debated, both amongst politicians and academics. Besides a general discussion regarding the methodology used to construct the poverty measures, two special issues have been the main topics in the debate.

The first issue is related to disparity in consumption levels between estimates from the NAS and the NSS household surveys, with the NAS data consistently higher than the other. Generally this is as expected, since national accounts usually include a range of items that are not reported in household surveys. However, the disparity between the two estimates seemed to increase during the 1990s, and as a consequence, data from the NAS implied a much steeper poverty decline. This generated a debate, which was not only theoretically, but also politically motivated. The opponents of the liberalization generally argued that the NAS data was no more reliable than the NSS data, while pro-reformers generally argued the other way around (see e.g., Deaton and Kozel (2005) for a discussion). Before the 1990s the IPC used the mean from the NAS to scale up the expenditure distribution from the NSS data. This method is practiced by some organizations and researchers (e.g., Sala-i-Martin 2006), but generally there is no reason to assume that national accounts are superior to households surveys (Anand and Segal 2008; Milanovic 2005).

The second issue is related to the specific design of the 55th NSS household survey round (1999-00). Many believe this survey is not consistent with the other household surveys conducted by the NSS, because of a different reporting period for consumption. Although the survey were accepted by the Indian Government, most independent researchers argue that the special design of the survey resulted in an overestimation of the poverty reduction (see Deaton and Kozel (2005) for a review). A more comprehensive discussion of the 55th NSS survey is given in Appendix B.

2.4 Potential Bias in The Poverty Lines

In addition to the special issues concerning the poverty rates discussed above, there are probably more general biases in the IPC poverty measures. The IPC's methodology to calculate poverty lines have been criticized both on normative and technical grounds. Perhaps the most powerful normative critic is that the poverty lines today fail to preserve the original calorie norms (see e.g., Holla, Guha, and Krishna Kumar 2008; Deaton and Dréze 2008). The calorie norm, calculated by The Task Force 1979, is only used to peg the state-wise poverty lines to the all-India poverty lines in 1973-74. After that, changes in consumption patterns, which potentially could lead to a different calorie intake, are ignored. The choice of structural poverty lines are certainly an important issue, but since the main focus in this thesis is to study differences in cost of living across time and space and the consequences for poverty measurement, I will not focus on this normative perspective. Instead, I focus on the technical critics, which is related to how the IPC makes adjustments in the poverty lines for different states and time periods. There are at least two potential sources of bias in the IPC's methodology, one that concerns the spatial dimension and one that concerns the time dimension. In the next subsections I discuss these in detail.

2.4.1 The Spatial Dimension – Unit values

As the state-wise poverty lines are based on price indices constructed with household survey data from the early 1960s, they are carried forward to the base year of the poverty lines (1973-74) using state specific inflation rates. Both the index for the rural sector and the index for the urban sector are based on so-called unit values (UV). UVs are calculated by simply taking the ratio of total value of consumption to the total quantity of consumption for different goods. For instance, if a household reports a purchase of four apples, and the total value of these four apples are reported to be eight Rs, then the value of one apple is easily calculated as two Rs. Under certain assumptions, among them that utility is weakly separable in the commodity groups, UVs can be used to construct cost of living indices for different states and sectors, and potentially also for different socio-economic groups (Deaton 1988).

It is difficult to evaluate the quality of the NSS household surveys from the early 1960s without the necessary background material, which is not available. However, what is known is that they were not collected specifically for the purposes of construction price indices, and hence, it is uncertain how appropriate they are. The household surveys provide necessary information of quantity and value on most high-frequency consumption items, such as food and fuels, but not for other kinds of items, such as for example transportation, housing or education. Thus, the UV indices have to assume that the relative prices between the covered goods and the non-covered goods are the same in all states. Another concern with the use of UVs as a measure of cost of living is quality differences in goods. Since very few consumption goods are perfectly homogenous, the UVs will reflect differences in quality as well as differences in prices. For instance, richer households usually report higher UVs than less rich households, simply because they buy goods with a higher quality. This quality problem is likely to persist, even with fine tuned consumption categories. The required assumption in the inter-state indices used for the poverty lines, is however that the average quality of each consumption good is the same in all the Indian states.

Even if one is willing to ”accept” the UV indices from the 1960s as consistent measures of real price differences between states at that time, it is much harder to be convinced by the deflated price differences for other years. These will of course, in addition to the other requirements, rest on the state-wise price deflators’ ability to consistently capture the relative increase in the cost of living. However, the direction of the spatial bias in the poverty lines is hard to asses a priori.

2.4.2 The Time Dimension – Laspeyres Bias

Both the CPIAL and the CPIIW, compiled by the Labour Bureau, are based on a fixed basket Laspeyres index formula, with the known biases inherited in these type of price indices. The price deflators used for the poverty lines are re-weighted versions of these indices, and thus, inherit the same type of problems as the original indices. The Laspeyres formula compares prices in different time periods by using the price structural in the first period. Hausman (2003) discusses four sources of bias in a fixed basket approach consumer price index (CPI). First, as fixed basket calculations rely on a reference price vector for comparison, that is, the actual consumption levels are evaluated using a different set of prices than those faced by consumers, they fail to pick up substitution in consumption. If the consumers face the reference prices, they would have chosen differently; they would have substituted away from goods that are relatively more expensive and towards goods that are relatively less expensive at these reference prices. The failure to incorporate the substitution in consumption is referred to as a substitution bias. Second, the fixed basket approaches fail to take into account that consumers substitute away from stores that are relative expensive towards stores that are relative less expensive. This is referred to as an outlet bias. Third, the fixed basket calculations fail to pick up differences in quality. This is referred to as a quality bias. Interestingly, all these biases are likely to systematically lead to an overestimation of the increase in cost of living with the Laspeyres formula[5]. Fourth, the fixed basket approaches meet a challenge when new goods are introduced. This is referred to as the new goods bias.

The bias in the fixed basket CPI is likely to be more serious when the consumption weights are infrequently updated, which indeed is the case for the two current price indices. The base year and weights in the CPIIW were updated in 1988 (base year 1982) and again in December 2005 (base year 2001). Prior to these updates, the base year and the weights used were as old as from 1960. The CPIAL was revised in 1995, with 1986 as

[5] Since the Paasche formula evaluates prices using the price structural in the final time period, the opposite result will occur.

the new base year. Before this revision the weights steamed from a consumption survey from 1960-61. The weights in the modified indices used for the poverty lines, are based on consumption patterns in 1973-74, and have not been updated for later years. This means that these weights are even more outdated than the weights used in the original price indices. Hence, the advantage of being tailored specifically to the poor comes at the expense of more outdated consumption weights.

2.5 Alternative Approaches

The importance of the poverty estimates, combined with the problems in the methodology used to compile them, have stimulated researchers to produce independent poverty estimates. A prominent researcher in that respect is Professor Angus Deaton, who in a series of papers calculates new sets of state-wise price indices using UVs. Although Deaton is the most cited, other researchers have constructed UV price indices for India (e.g., Palmer-Jones and Dubey 2008; Himanshu and Murgai 2009). However, I mainly focus on Deaton's estimations, both in this section and when I evaluate my own estimated price indices in Section 7.

Deaton and Tarozzi (2000) calculate UV indices with the Laspeyres, Paasche, Fisher and the Törnqvist formulas for the largest Indian states, with data from the 43rd (1987-88) and 50th round (1993-94) of the NSS household surveys. Deaton (2003) updates these price indices for the 55th NSS survey round (1999-00). The new price indices are used to update the official poverty lines for each state and sector, resulting in a new set of poverty rates. One of the main findings in these two papers is that the price differential between the rural and the urban sector implied by the official poverty lines is too large. The official estimates are not comparing prices in the two sectors directly, but instead, the implied price differential has its origin in nutrition needs and consumption patterns in 1973-74. In the IPC's most recent measures (2004-05), the poverty lines for urban areas are roughly 50 percent higher than the corresponding poverty lines in rural areas[6]. According to Deaton and Tarozzi's UV indices, the true price differential is instead much closer to the 15 percent implied by the poverty lines used before the Expert Group Report 1993.

To update the official poverty lines with the new sets of price indices, Deaton and Tarozzi anchor the different poverty lines to the all-India rural poverty line for 1987-88. All the other poverty lines are then calculated with reference to the price differences to this base group. As a consequence of this and the finding of less price differences between the two sectors, the urban poverty rates as well as the all-India poverty rates decline, compared to the IPC's estimates. This might very well be a reasonable conclusion in itself. However, the conclusion is based on an arbitrary choice of base group, without any plausible justifications. If Deaton and Tarozzi instead had used the urban poverty line as a base[7], the opposite result would have occurred – poverty rates for the rural sector, and hence, also for all-India would have increased. This dramatically changes the conclusion. As pointed out by Himanshu, the urban poverty line may be less controversial to use as references, because the rural poverty rates are widely considered to be too low (Himanshu and Murgai 2009). Perhaps more seriously; the change in poverty for a specific area over time depends on the level of the poverty line in itself, it is not a unique function of the change in the nominal poverty line (Deaton and Tarozzi 2000). This means that the base group will influence the outcome over time, even if the focus are entirely on one specific area. The share of UV-goods of total consumption falls overt time. Since most UV-items consist of food, a falling covered share is consistent with real growth in expenditures and Engel's Law. A direct consequence for Deaton estimation however, is that the price indices are calculated on an increasingly thinner data material.

Deaton (2008) calculates price indices and poverty rates based on the 55th (1999-00) and 61st (2004-05) round of the NSS households surveys. The methodology used to con-

[6]For states like Andhra Pradesh, Karnataka and Maharashtra, the urban poverty line are about 85 percent higher than the corresponding rural poverty line i 2004-05.

[7]This is done in Himanshu and Murgai (2009).

struct these estimates differs in some respects from that of Deaton (2003) and Deaton and Tarozzi (2000). Among other things, the UV price indices are now combined with non-UV items from the CPIIW and the CPIAL. One advantage of this is that the price indices cover a broader range of goods. Another is that it is no longer necessary to assume that relative prices between UV-items and non-UV items are the same over different survey rounds. A second methodology change compared to earlier studies, is the use of both the rural and the urban poverty lines as base. As a consequence, the price differential between the two sectors implied by the IPC's measures are accepted also in Deaton's estimates. This contrasts with the findings in Deaton and Tarozzi (2000) and Deaton (2003).

3 Engel's Law

Engel's Law (Engel 1857), named after the German statistician Ernst Engel, is one of the most established connections in economics. It states that the budget share for food is negatively related to real income, everything else equal. Thus, when a household becomes richer, it uses a smaller share of its total budget on food items. This is certainly not the same as an absolute decrease in food expenditure, but instead the law states that households increase their food spendings by less (in percentage) than their increase in total expenditure. Working (1943) formalized this hypothesis, and argued that there was a linear relationship between budget share for food and real income. Leser (1963) later proposed a log-linear relationship. In 1980, Deaton and Muellbauer presented a theoretical demand system of great importance; An Almost Ideal Demand System (AI) (Deaton and Muellbauer 1980). Until then, Engel's Law were really just an empirical finding, without theoretical foundation. However, Deaton and Muellbauer showed that Engel's Law is consistent with consumer preferences, that is, it is consistent with utility maximization under a budget constraint. Banks, Blundell, and Lewbel (1997) further proposed a quadratic extension of this system; The Quadratic Almost Ideal Demand System (QAI). In the following I describe both the AI and the QAI demand system more detailed.

3.1 An Almost Ideal Demand System

Deaton and Muellbauer's AI-model is based on a cost function that represents the so-called "Price-Independent Generalized Logarithmic" (PIGLOG) preferences;

$$\ln C(u,\mathbf{p}) = (1-u)\ \ln a(\mathbf{p}) + u\ \ln b(\mathbf{p}), \tag{1}$$

where u and $\mathbf{p}$ are utility and the vector of consumption prices, respectively, and $\ln a(\mathbf{p})$ and $\ln b(\mathbf{p})$ are differentiable functions of the consumption prices. The corresponding indirect utility function is given by:

$$\ln V(\mathbf{p},y) = \left[\frac{\ln y\ -\ \ln a(\mathbf{p})}{b(\mathbf{p})}\right], \tag{2}$$

where y is nominal income. Deaton and Muellbauer further define a specific form for both $\ln a(\mathbf{p})$ and $\ln b(\mathbf{p})$;

$$\ln a(\mathbf{p}) = \alpha_0 + \sum_i \alpha_i\ \ln p_i + \frac{1}{2}\sum_i\sum_j \gamma_{ij}^*\ \ln p_i\ \ln p_j\ , \tag{3}$$

$$\ln b(\mathbf{p}) = \ln a(p) + \beta_0 \prod_i\ p_i^{\beta_i}\ . \tag{4}$$

Substituting these two expressions into the cost function given in Equation (1) gives:

$$\ln C(u,\mathbf{p}) = \alpha_0\ + \sum_i \alpha_i\ \ln p_i\ + \frac{1}{2}\sum_i\sum_j \gamma_{ij}^*\ \ln p_i\ \ln p_j\ + u\ \beta_0 \prod_i\ p_i^{\beta_i}\ . \tag{5}$$

Using Shepard's Lemma, the demand function for good i can be found directly from the cost function:

$$\frac{\partial\ C\ (u,\mathbf{p})}{\partial\ p_i} = q_i\ .$$

By multiplying both sides with $p_i\ /\ C(u,\mathbf{p})$ we get:

$$S_i = \frac{p_i\ q_i}{C\ (u,\mathbf{p})} = \frac{\partial\ ln\ C(u,\mathbf{p})}{\partial\ \ln p_i}\ .$$

Hence, to attain the budget share for each good i, we can just differentiate Equation (5) with respect to log of the price of good i. Doing that gives:

$$S_i = \alpha_i \; + \sum_j \frac{1}{2} \, (\gamma_{ij}^* + \gamma_{j\,i}^*) \; \ln p_j \; + \; \beta_i \, u \, \beta_0 \prod p_i^{\beta_i} \;, \tag{6}$$

which is the budget share equation for the AI-system. For a utility maximizing consumer the value of the cost function, $C(u, \mathbf{p})$, must equal total income, y. From Equation (5) this gives:

$$\ln y = \ln C \; (u, \mathbf{p}) \quad \Longleftrightarrow \quad \ln y - \; \ln a(\mathbf{p}) = u \; \beta_0 \prod \; p_i^{\beta_i} \; .$$

Using this finding, we could therefore rewrite the budget share equation given in Equation (6) as:

$$\begin{aligned} S_i &= \alpha_i \; + \sum_j \gamma_{ij} \; ln \; p_j \; + \; \beta_i \; \ln \frac{y}{a(\mathbf{p})} \;\; , \text{ where:} \\ \gamma_{\,ij} &= \frac{1}{2} \; (\gamma_{ij}^* + \gamma_{j\,i}^*) \; . \end{aligned} \tag{7}$$

From Equation (7), we can see that the demand system is consistent with a log-linear relationship between budget share for food and real income, as suggested by Leser (1963). A convenient property of the system, is that the coefficient of real income, β_i, in the budget share equation is constant.

3.2 The Quadratic Almost Ideal Demand System

Banks et al. (1997) present a generalization of the AI-model of Deaton and Muellbauer, starting from an indirect utility function and a cost function of the form:

$$\ln V(\mathbf{p}, y) = \left\{ \left[\frac{\ln y \; - \; \ln a(\mathbf{p})}{b(\mathbf{p})} \right]^{-1} + \; \lambda(\mathbf{p}) \right\} \;\; , \;\; \text{and} \tag{8}$$

$$\ln C(u, \mathbf{p}) = \ln a(\mathbf{p}) \; + \; \frac{u b(\mathbf{p})}{1 - u\lambda(\mathbf{p})} \; , \tag{9}$$

where $\lambda(\mathbf{p})$ is a differentiable function of the consumption prices. When $\lambda(\mathbf{p})$ is independent of prices, the two functions are reduced to that of PIGLOG preferences, and hence, also the AI-model (see Equation (1) and (2)). Thus, the QAI-model have the AI-model nested as a special case. By Roy's Identity, the budget shares for good i can be calculated as:

$$S_i = \frac{\partial \ln a(\mathbf{p})}{\partial \ln\ p_i} + \frac{\partial \ln b(\mathbf{p})}{\partial \ln p_i}\ \ln \frac{y}{a(\mathbf{p})} + \frac{\partial \lambda}{\partial \ln p_i}\ \frac{1}{b(\mathbf{p})} \left[\ln \frac{y}{a(\mathbf{p})} \right]^2 . \tag{10}$$

Contrary to the AI-model, where the budget shares are being linear in $\ln[\, y/a(\mathbf{p})]$, the budget shares in the QAI-model are quadratic in $\ln[\, y/a(\mathbf{p})]$. Following Deaton and Muellbauer (1980) the QAI-model sets $\ln a(\mathbf{p})$ to have the translog form in Equation (3), and $b(\mathbf{p})$ is defined as a Cobb-Douglas price aggregator:

$$b(\mathbf{p}) = \prod_i p_i^{\beta_i} .$$

To complete the QAI-model, Banks et al., finally define $\lambda(\mathbf{p})$ as:

$$\lambda(\mathbf{p}) = \sum_i \lambda_i\ \ln p_i ,$$

$$\text{where}: \ \sum_i \lambda_i = 0 .$$

From the expressions for $\ln a(\mathbf{p})$, $b(\mathbf{p})$ and $\lambda(\mathbf{p})$ the corresponding budget share equation for good i can be obtain from Equation (10) as:

$$S_i = \alpha_i + \sum_j \gamma_{ij}\ \ln p_j + \beta_i\ \ln \left[\frac{y}{a(\mathbf{p})} \right] + \frac{\lambda_i}{b(\mathbf{p})} \left\{ \ln \left[\frac{y}{a(\mathbf{p})} \right] \right\}^2 , \tag{11}$$

which is a non-linear function of the consumption prices. Blundell, Pashardes, and Weber (1993) propose a similar model where the coefficients of real income are constant, as in the AI-model. However, Banks et al. (1997) prove that this is not possible with utility maximization; the coefficient of the quadratic term must be price dependent, as it is in Equation (11).

4 Identification of Cost of Living Indices

Given the theoretical platform described above, together with the empirical evidences for Engel's Law, it should be possible to identify differences in cost of living between areas or between time periods. If one is able to control for all relevant factors that have an effect on food expenditures, one should in principle also be able to infer movements in real income from the movements in the budget share for food. Hamilton (2001) was the first to use this basic idea to measure the degree of bias in the CPI in the United States. The main assumption in his method is that there exists a stable Engel relation over all situations being compared – across geographical areas, as well as across different time periods. Hence, if we observe that two households with identical characteristics living in two different areas have equal budget shares for food, they should also have identical real incomes. If their nominal incomes are different, this reveals a price differential between the two areas. Costa (2001) extends Hamilton's method with a quadratic term of real income, and thus makes the system more flexible to potential functional form miss-specifications. However, Costa's approach is not fully exploring the interaction that arises within the system when the quadratic term is included, which might help in the identification[8].

The Engel approach identifies the cost of living from the budget share for food and the observed consumer behavior, and hence, it should not be affected by aggregation problems. Given that the assumption of a unique Engel relation is valid, the method would avoid at least some of the biases associated with the fixed basket price indices (see Section 2.4.2). As discussed in Hausman (2003), the Hamilton-Costa approach is likely to capture the outlet bias and the substitution bias, but less likely the other sources of bias. The Engel method could therefore be interpreted as a measure of a lower bound of the bias in CPI (quality improvement is likely a major source of bias over time). However, as pointed out by Beatty and Røed Larsen (2005), the Engel curve approach will capture the quality changes that improve the durability of goods.

[8]See Section 4.1 for a more detailed description.

In the following I propose a reformulation of the Hamilton-method, which makes it possible to calculate spatial price indices (SPI). Instead of relating the consistent real income levels to a specific price index, I identify consistent price indices in levels directly. Since one of the main motivations here is to study spatial price differences across Indian states, none of the official Indian price indices are appropriate to use in the estimation in itself. As in Hamilton and Costa's earlier work, the procedure is based on a two-goods version of the demand systems, where the two goods are food and non-food items. Because of the adding up property of the demand systems, it will then only be necessary to use one of the goods to complete the systems.

4.1 Empirical Specifications

The empirical investigation of the AI-model is based on the following specification:

$$m_{h,s,r,j} = \alpha + \beta(\ln\, y_{h,s,r,j} - \ln\, P_{s,j}) + \gamma(\ln\, P_{f,s,r,j} - \ln\, P_{n,s,r,j}) + \theta X_{h,s,r,j} + \varepsilon_{h,s,r,j}, \quad (12)$$

where $m_{h,s,r,j}$ is the budget share for food, $y_{h,s,r,j}$ is the nominal household expenditure, and $X_{h,s,r,j}$ is a vector of demographic control variables including the age of the household head and the number of children and adults in the household, for household h in state s and region r at time j. $P_{f,s,r,j}$ is the price of food and $P_{n,s,r,j}$ is the price of non-food in region r in state s at time j. $P_{s,j}$ is the composite price of consumption, a weighted average of food and non-food prices, in state s at time j, and $\varepsilon_{h,s,r,j}$ are the residuals. If region-wise cross-state-and-time comparable food and non-food price data are unavailable for the periods under study, the coefficient for relative prices, γ, cannot be estimated. Consequently, an alternative estimation equation excludes relative prices between food and non-food items and, therefore, implicitly assumes that the budget share for food is unaffected by relative prices. However, as shown in Almås (2007), the results are very robust to relative price effects. When excluding the relative price effect, Equation (12) can be simplified to:

$$m_{h,s,r,j} = \alpha + \beta(\ln\ y_{h,s,r,j} - \ln\ P_{s,j}) + \theta X_{h,s,r,j} + \varepsilon_{h,s,r,j}, \tag{13}$$

The Hamilton method can be used to derive price indices in levels directly, both cross-section and over time. To identify the unbiased consumption price for state s in period j, estimate the following expression:

$$m_{h,s,r,j} = \alpha + \beta(\ln\ y_{h,s,r,j}) + \theta X_{h,s,r,j} + \sum_s \sum_j\ \delta_{s,j} D_{s,j} + \varepsilon_{h,s,r,j}. \tag{14}$$

The dummy coefficient, $\delta_{s,j}$, is a function of the unbiased macro consumption price, $P_{s,j}$, and the coefficient for the logarithm of household expenditure, β:

$$\delta_{s,j} = -\beta\ \ln\ P_{s,j}, \tag{15}$$

The macro consumption price is thus given by:

$$P_{s,j} = e^{-\frac{\delta_{s,j}}{\beta}}. \tag{16}$$

The corresponding empirical specification of the QAI-model, without relative food and non-food prices, is given by:

$$m_{h,s,r,j} = \alpha + \beta_1(\ln\ y_{h,s,r,j} - \ln\ P_{s,j}) + \beta_2(\ln\ y_{h,s,r,j} - \ln\ P_{s,j})^2 + \theta X_{h,s,r,j} + \varepsilon_{h,s,r,j}, \tag{17}$$

Solving Equation (17) gives:

$$\begin{aligned} m_{h,s,r,j} = \alpha + \beta_1(\ln y_{h,s,r,j}) + \beta_2(\ln y_{h,s,r,j})^2 + \theta X_{h,s,r,j} - \beta_1 \ln P_{s,j} + \beta_2(\ln P_{s,j})^2 \\ -2\beta_2 \ln y_{h,s,r,j}\ \ln P_{s,j} + \varepsilon_{h,s,r,j}, \end{aligned} \tag{18}$$

which simplifies to:

$$m_{h,s,r,j} = \alpha + \beta_1(\ln y_{h,s,r,j}) + \beta_2(\ln y_{h,s,r,j})^2 + \theta X_{h,s,r,j} + \sum_s \sum_j \delta_{1,s,j}\, D_{s,j} + \sum_s \sum_j \delta_{2,s,j} \ln y_{h,s,r,j} D_{s,j} + \varepsilon_{h,s,r,j}\ . \tag{19}$$

When introducing the quadratic component, we can see from Equation (19) that an interaction term between the dummy variables and expenditure arises. This contrasts with the quadratic specification in Costa (2001). In that case, an interaction term between the time dummies and total expenditure should be included (Costa 2001, 1294). The coefficients in Equation (19) correspond to the following parameters in the theoretical model in Equation (11) as follows:

$$\begin{aligned}
&\alpha = \alpha && \beta_1 = \beta \\
&\beta_2 = \frac{\lambda}{b(\mathbf{p}_{s,j})} && \delta_1 = \frac{\lambda}{b(\mathbf{p}_{s,j})}\left\{\ln P_{s,j}\right\}^2 - \beta \ln P_{s,j} \\
&\delta_2 = 2\,\frac{\lambda}{b(\mathbf{p}_{s,j})} \ln P_{s,j} && P_{s,j} = a(\mathbf{p}_{s,j})
\end{aligned}$$

The model in Equation (19) is overidentified and the price component cannot be efficiently identified through ordinary least square estimation (OLS). However, the model can be estimated using non-linear estimation techniques. To identify the price component, I proceed by rewriting Equation (17) as follows:

$$m_{h,s,r,j} = \alpha + \beta_1(\ln y_{h,s,r,j} - \sum_s \sum_j \delta_{s,j} D_{s,j}) + \beta_2(\ln y_{h,s,r,j} - \sum_s \sum_j \delta_{s,j} D_{s,j}) + \theta X_{h,s,r,j} + \varepsilon_{h,s,r,j}, \tag{20}$$

The price component can be identified directly from the dummy coefficients. To estimate the expression in Equation (20), I am using a modified Gauss-Newton iteration technique. As starting values for the dummy variable coefficients in the iteration procedure, I use the estimated prices from the linear specification (the AI-model). For the expenditure and the control variable coefficients, I use estimated coefficients from the specification given in Equation (19).

The identification strategies described in this section use the same cost of living concept as proposed in Hamilton (2001). The procedure gives consistent measures of the cost of living when relative prices are equal in all situations. However, as shown in detail in Appendix A.1, Hamilton's identification do not capture the complete cost of living in situations where relative prices differ. There are also other theoretical challenges with Hamilton's method. The method identifies a single price index for each area and time period, and uses this price index to calculate the real income level in each situation. However, this will only be valid if preferences are homothetic, that is, they are equal for all real income levels. Yet the two theoretical demand systems and Engel's Law imply nonhomthetic tastes, since the budget share for food falls in real income. This is a conceptual problem, not only for this thesis, but also for other papers using the Engel method proposed by Hamilton (2001).

5 Data and Descriptive Statistics

5.1 The NSS Household Surveys

Compared to other developing countries, India has a long and proud tradition in collecting quality databases for addressing socio-economic issues. The NSS collects data on a wide range of subjects, many of them related to poverty. However, my Engel Curve analysis is based solely on the household expenditure surveys. The NSS publishes large consumer expenditure surveys for all the major states and union territories in India, once every fifth year. These surveys are constructed separately for the urban and the rural sector. In this thesis I make use of the household surveys from 1983, 1987-88 and 1993-94[9]. The number of households in each survey are shown in Table 1. A household is defined as a group of persons normally living together and taking food from a common kitchen. Total

[9]In Appendix B I also make us of the surveys from 1999-00 and 2004-05.

expenditure includes all domestic consumer expenditure during the last 30 days. The way the surveys are compiled makes it possible to distinguish food and non-food expenditures, and hence, study these two consumption groups. Food expenditure consists of all food groups, and includes non-alcoholic beverages. It also includes both home-cooked food and restaurant meals. Since food at home and food at restaurants not necessary are perfect substitutes, I should ideally have separated them in the estimation. However, this is not possible given my data sample. Although the effect of restaurant meals on the budget share for food should be investigated further, the effect is not likely to be very large (see Hamilton 2001). Consumption from cash purchase is evaluated at the purchase price, while consumption out of home production is evaluated at ex farm or ex factory rate. The value of in-kind consumption items is evaluated at the average local retail prices during the reference period.

TABLE 1: Number of Households in Each Survey Round

Year	Rural	Urban	Total
1983	75821	38426	114247
1987-88	82597	44649	127246
1993-94	69206	46148	115354

Source: My own calculations based on the NSS unit record data.

5.2 Control Variables

In order to identify a ceteris paribus Engel relation between budget share for food and real income, it is important to control for households' demographic (see e.g., Blow 2006). In principle one can think of many variables that could have an effect on the budget share spent on food items. However, in this analysis I am limited to the information given in the household surveys. I proceed by using three demographic control variables, namely number of children, number of adults belonging to the household, and age of household head. These control variables are consistent with the relevant literature. A child is defined as a person less than 16 years of age, while an adult is a person with an age of 16 or above.

The classification of the household head is not mechanical, but instead self-reported in the household surveys.

In order to control for relative food and non-food prices, and to identify the γ-coefficient, I should ideally have comparable relative food and non-food prices in levels between different states. However, the CPIAL and the CPIIW measure the percentage change in prices from the base period – they do not report cross-section data on price levels. I therefore proceed by estimating the models without including relative prices. This procedure is valid if relative prices are equal in all states and time periods, or alternatively, they do not affect the preferences for food and non-food items. As a robustness check I control for relative prices for food and non-food items by using relative inflation rates as proxies for relative prices in levels. These calculations are shown in Appendix A.1. Although the specific price estimates differ somewhat, my main findings are not affected.

5.3 Sample Weights

The household surveys from the NSS are constructed in sample strata. That is, in each survey the sample households are divided into different groups. The surveys are designed to be representative for the whole Indian population, with a specific sample weight given to each stratum. The sample weights reflect the probability of selection in the surveys. If the mean of a variable differs between strata, an unweighted average will give a biased measure of the mean for the whole population. When doing descriptive analysis this could easily be fixed by weighting each stratum with its corresponding sample weight. This will give an unbiased measure of the population mean. Analogous problems could arise when doing regression analysis, but matters are somewhat more complicated. For instance, consider a linear regression where the structural parameters differ between strata. One alternative is to proceed by using the sample weights in the estimation, and get the average population parameters. This is a reasonable strategy if one sees the regression model mainly as descriptive and not structural. However, when the model is seen as structural, this strategy seems less reasonable. The weights will make the data sample look like the true population, and hence, make the estimation independent of the specific sample design. Even so, when the structural parameters differ between subgroups, a weighted regression will still be inconsistent. The heterogeneity in the parameters is characteristics of the population, not of the specific sample design. Since both unweighted (OLS) and weighted estimators will be inconsistent in the case of heterogeneity, and since OLS is the most efficient estimator in the case of homogeneous parameters, there is support for OLS on econometric grounds (Deaton 1997). In this thesis I am mainly interested in the structural Engel curve. Therefore, I proceed by using an unweighted estimator and ignore the sample weights in the estimation.

5.4 Descriptive Statistics

Summary statistics for the data sample used in the main analysis are presented in Table 2, separately for the rural and the urban sector. When calculating the mean of each variable, I am using the sample weights given in the NSS surveys. It is much more likely that the mean differs between strata than the parameters differ. It can be seen from the table below that the budget share for food falls over time for both sectors. This is consistent with growth in real income and Engel's Law. The development in the demographic variables over time could also be interpreted as a sign of growth in real income and general improvement in the living conditions. The age of the household head increases over time, which probably reflects a higher life expectancy. In addition, it can be seen that the average number of children decreases over time, which could be linked to a more modern way of living. Finally, note that the average nominal expenditure level is higher in the urban sector than in the rural sector. This could reflect higher real income levels, or alternatively, just higher price levels in the urban sector.

TABLE 2: Summary Statistics

	1983	1987-88	1993-94
RURAL			
Total Expenditure	111.97 (0.381)	157.83 (0.577)	281.43 (1.143)
Food Share	0.70 (0.001)	0.68 (0.001)	0.68 (0.001)
No of Children	2.98 (0.011)	2.83 (0.010)	2.57 (0.010)
No of Adults	3.60 (0.012)	3.57 (0.010)	3.51 (0.010)
Age of Household Head	45.24 (0.067)	44.98 (0.063)	45.08 (0.067)
Ln Relative Prices	0.09 (0.000)	0.47 (0.001)	0.47 (0.001)
URBAN			
Total Expenditure	163.91 (0.878)	245.26 (1.506)	462.97 (3.321)
Food Share	0.65 (0.001)	0.63 (0.003)	0.61 (0.001)
No of Children	2.65 (0.017)	2.49 (0.016)	2.12 (0.013)
No of Adults	3.69 (0.017)	3.65 (0.017)	3.50 (0.015)
Age of Household Head	44.07 (0.098)	44.04 (0.090)	44.25 (0.093)
Ln Relative Prices	-0.00 (0.000)	0.00 (0.000)	0.17 (0.000)

Note: Standard errors in parentheses. These are corrected for the sample design using a Taylor-linearized variance estimation. Ln Relative Prices are calculated in the same way as in Equation (14) and (17), i.e., food inflation divided by non-food inflation in the CPIAL and the CPIIW for the rural and the urban sector, respectively.

Source: My own calculations based on the NSS unit record data.

5.5 Data Exclusion

I limit my Engel analysis to cover the 17th largest Indian states, plus Delhi (urban only). I choose to do this, partly because of small samples sizes for the smallest states, and partly to make my resulting price indices comparable with the official as well as Deaton's alternative estimates. This delimitation reduces the sample size with approximately nine percent. Further steps are taken to limit the final data sample used in the estimation. First, I exclude all households which report zero or negative total expenditure or food expenditure[10]. Second, households with a food share below zero or above unity are excluded[11]. Third, and finally, observations with a reported age of household head above 120 years is excluded[12]. Combined, all these steps result in a final data sample consisting of 322.318 households in total. This means that the original sample is reduced by approximately 10 percent.

[10]This step results in 702 observations being dropped.

[11]Resulting in a sample reduction of 1579 observations.

[12]Reduces the sample by another 335 observation. Over 70 percent of these households report an age above 1000 years. These observations, which obvious are measured with error, will make trouble for the iteration procedure used in the identification of the QAI-model.

5.6 Sample Pooling

Before I turn to the empirical results, I address one last consideration; namely sample pooling. In the following empirical analysis, I proceed by pooling data from 1983, 1987-88 and 1993-94 in the estimation. A critical assumption for pooling survey data from different areas or different time periods is that they are completely comparable and harmonized. Another important assumption is that each sub-population can be represented with the same structural model. In this setting, that is equivalent to assuming that the preferences are identical in each sub-population, i.e., the existence of a unique Engel relation over all situations being compared. If these assumptions are violated, the identified prices would not only include real price differences, but also noise from the data or differences in preferences, and hence, they would be biased price measures. We have to be aware of these potential problems when pooling data from different sub-populations.

My main concern is the rural to urban comparison. The distinction between the two sectors is a major structural feature of developing countries (Ray 1998). A large share of the population in rural areas is often employed in agricultural activities. For many of them, agricultural will define their way of living. People in urban areas, on the other hand, are more often employed in advanced industries and services. Another feature of urban areas is that they generally have access to a wider range of consumption goods than rural areas. Although it is reasonable to assume that individuals located in rural and urban areas have the same innate preferences, different supply of consumption goods can influence how they allocate their spendings between food and non-food items. Because of this concern, as well as concerns of how comparable the survey data are, I proceed by analyzing the rural and the urban sector separately[13]. This strategy is the safest with respect to non-comparable data sets, but it is not optimal because it rules out the possibility to investigate the price differential between the two sectors. This is a major disadvantage in this case, because the implied price differential from the official poverty lines is perceived as one of

[13] In Appendix C I discuss the pooling of rural and urban household in detail, and present estimation results of a pooled regression.

the failures of the IPC's measures (see Section 2.5).

6 Empirical Results

The estimation results from the regression of the AI- and the QAI-model (Equation (14) and (20)) is presented in Table 3. First, consider the linear Engel specification (AI), given in the first two columns. As expected, the logarithm of total monthly expenditure has a significant negative effect on the budget share for food in both sectors. Everything else equal, richer households seem to use a smaller share of their total budget on food items. This finding is consistent with Engel's Law. However, we also see that the Engel relation seems to be steeper for the urban sector compared to the rural sector. That is, budget share for food decreases at a faster rate when households become richer in urban areas. This gives some support for the decision to estimate the model separately for the two sectors. Interestingly, the coefficients for the quadratic extension, given in the third and fourth columns, are highly significant for both sectors. In the relevant literature where Engel curves for food is estimated, it is mostly assumed that the Engel relation is log-linear, and hence the AI-model is used (e.g., Almås 2007; de Carvalho Filho and Chamon 2007; Gibson, Stillman, and Le 2007; Hamilton 2001). However, in this case the significant squared terms suggest that the quadratic specification is most appropriate for India in the period under study[14]. Finally, we can see that all the three demographical control variables have a significant effect. Especially number of children and number of adults influence the budget share spent on food. The age of the household head is statistically significant, but the small magnitude of the coefficients suggests that the variable barely has effect on the budget share for food. The state-dummies are reported separately in Table 4[15]. Figure 1 presents the estimated relation between budget share for food and log expenditure at the sample mean of the other control variables.

[14] As a robustness check for the iteration procedure I have estimated the quadratic model with a range of different starting values. All parameters seem extremely robust.

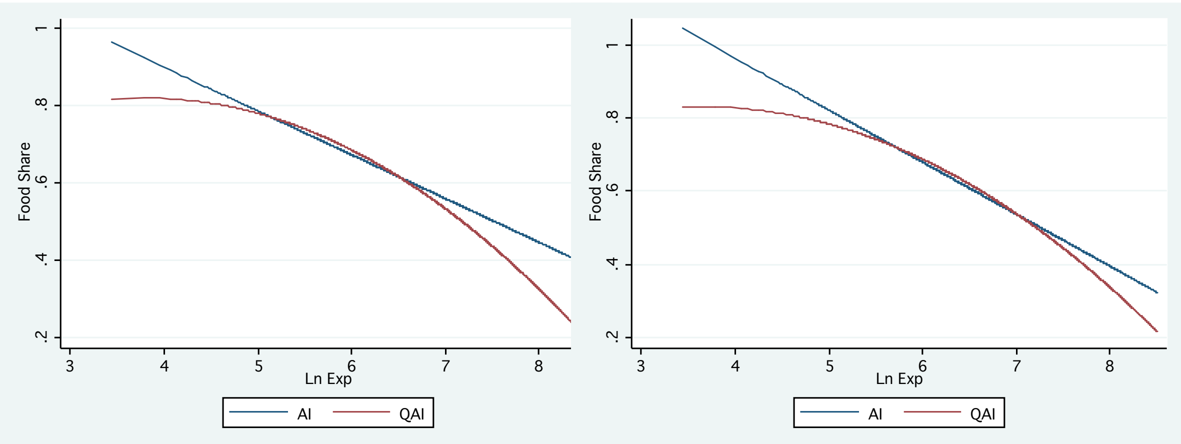
[15] Due to space considerations, I order these coefficients by survey years.

TABLE 3: Regression – Main Analysis

	AI		QAI	
	Rural	Urban	Rural	Urban
Ln Exp	-0.1133	-0.1420	0.2171	0.1914
	(0.00073)	(0.00075)	(0.00335)	(0.00479)
$(\text{Ln Exp})^2$	–	–	-0.0284	-0.0262
			(0.00028)	(0.00037)
No of Children	0.0165	0.0203	0.0162	0.0184
	(0.00018)	(0.00021)	(0.00016)	(0.00021)
No of Adults	0.0200	0.0256	0.0215	0.0249
	(0.00023)	(0.00026)	(0.00020)	(0.00024)
Age of Household Head	0.0002	0.0004	0.0003	0.0005
	(0.00002)	(0.00003)	(0.00002)	(0.00003)
Constant	1.2459	1.4002	0.2877	0.3517
	(0.00439)	(0.00517)	(0.01002)	(0.01535)
Observations	204137	118219	204137	118219
Adj. R^2	0.259	0.387	0.294	0.412

Note: Robust standard errors in parentheses. Dummy coefficients are reported separately in Table 4.

FIGURE 1: Regression Plots



(a) Rural Sector (b) Urban Sector

TABLE 4: Dummy Coefficients

State	AI			QAI		
	1983	1987-88	1993-94	1983	1987-88	1993-94
Rural						
Andhra Pradesh	–	0.030	0.104	–	1.386	2.540
Assam	0.109	0.116	0.186	2.539	2.597	4.942
Bihar	0.086	0.102	0.157	2.270	2.429	4.141
Gujarat	0.067	0.107	0.165	1.703	2.534	4.125
Haryana	0.052	0.068	0.134	1.627	1.973	3.589
Himachal Pradesh	0.057	0.063	0.126	1.584	1.757	3.038
Jammu & Kashmir	0.090	0.089	0.146	2.043	2.150	3.584
Karnataka	0.022	0.051	0.103	1.289	1.674	2.612
Kerala	0.054	0.077	0.140	1.582	2.055	3.487
Madhya Pradesh	0.045	0.055	0.098	1.413	1.666	2.470
Maharashtra	-0.003	0.043	0.086	0.996	1.503	2.286
Orissa	0.095	0.082	0.133	2.452	2.059	3.100
Punjab	0.023	0.067	0.129	1.323	1.959	3.363
Rajasthan	0.023	0.066	0.133	1.269	1.938	3.350
Tamil Nadu	0.036	0.035	0.135	1.429	1.400	3.301
Uttar Pradesh	0.012	0.011	0.104	1.130	1.202	2.678
West Bengal	0.101	0.120	0.162	2.584	2.937	3.978
Urban						
Andhra Pradesh	–	0.031	0.108	–	1.278	2.153
Assam	0.074	0.103	0.165	1.727	2.034	3.212
Bihar	0.072	0.092	0.169	1.709	1.927	3.357
Gujarat	0.055	0.100	0.168	1.428	1.986	3.215
Haryana	0.020	0.062	0.121	1.171	1.563	2.409
Himachal Pradesh	0.034	0.065	0.141	1.235	1.563	2.688
Jammu & Kashmir	0.061	0.113	0.170	1.491	2.163	3.231
Karnataka	0.022	0.048	0.122	1.191	1.455	2.388
Kerala	0.050	0.083	0.138	1.452	1.822	2.678
Madhya Pradesh	0.022	0.050	0.109	1.168	1.460	2.207
Maharashtra	0.026	0.077	0.137	1.237	1.769	2.689
Orissa	0.062	0.082	0.122	1.577	1.805	2.382
Punjab	0.019	0.053	0.120	1.139	1.470	2.416
Rajasthan	0.008	0.065	0.139	1.099	1.606	2.658
Tamil Nadu	0.028	0.057	0.127	1.229	1.504	2.516
Uttar Pradesh	-0.001	0.043	0.118	1.024	1.384	2.366
West Bengal	0.040	0.072	0.148	1.346	1.673	2.780
Delhi	0.017	0.075	0.177	1.183	1.822	3.549

Note: The dummy coefficients from the QAI-model are not comparable with those from the AI-model. The way the QAI-model is specified here (see Equation 20), the dummy coefficients present the the prices compared to the base group directly. All coefficients are significantly different from the base group at a 1% level, except Maharashtra rural (AI and QAI) and Rajasthan urban (AI).

7 Findings

From the estimation discussed above, I calculate a range of different measures. First, spatial price indices (SPI) for all the survey years are calculated, separately for the rural and the urban sector. These indices give a measure of real price differences in levels between the Indian states. Second, state- and sector-wise price indices over time are calculated. These indices measure the real increase in the cost of living between the various survey rounds. If my only concern in this thesis was to calculate new poverty measures, I could have calculated these poverty measures directly without going through the estimation of different price indices. However, I believe the price measures have a value in themselves, because price comparisons between states and time periods are central in the way the official poverty measures are compiled, and in the general poverty discussion. Third, I use the estimated cost of living indices, both across time and space, to update the IPC's official poverty lines for each state and sector, and calculate the corresponding headcount ratios (HCR) and poverty gap index (PGI). Since I found a significant squared term of total expenditure for both sectors in the empirical investigation, my preferred indices are those of the quadratic specification. The prices from the linear Engel curves are presented mostly for comparison. As can be seen from the different tables, the two sets of price indices differ, but generally they are quite similar[16].

7.1 Spatial Price Indices

This section investigates the spatial dimension of the IPC's poverty lines, and compares it to the state-wise SPIs resulting from the Engel analysis. Tables 11 to 13 present SPIs relative to all-India for all the survey years, for the rural and the urban sector, separately. The first column in each table presents the implied SPIs from the log-linear Engel specification (AI), while the second column presents SPIs from the quadratic Engel specification (QAI). To construct these indices, I proceed by weighting each state by its population

[16]Due to space considerations, I only present the summary tables in the text. The different state-wise measures are presented in the end of this thesis.

compared to the all-India population in the same sector. This weighting procedure closely mimics the weights used in the official Indian price indices and Deaton's alternative UV price indices. Alternatively, I could have used total expenditure as weights[17]. The third column in the tables presents the state-wise price differences implied by the IPC's state-wise poverty lines. These prices are calculated by dividing the state-specific poverty line by the all-India poverty line for the same year and sector. I hereafter refer to these indices as IPC's SPI. For the surveys in 1987-88 and 1993-94, it is also possible to compare the estimated price indices to those of Deaton and Tarozzi (2000), derived through UV calculations. Their SPIs are presented in the fourth column of Table 12 and Table 13.

When evaluating my calculated SPIs, I report several findings. These findings are structured in three parts. The first part concerns the price variation in the different SPIs, the second concerns the correlation between the SPI in the rural and the urban sector, and third concerns the correlation between my estimated SPIs and the IPC's and Deaton and Tarozzi's SPIs.

7.1.1 Price Variation

Table 5 summarizes the dispersion in the different SPIs, measured by the coefficient of variation (CV)[18]. The first finding is that there seems to be more price variation in the rural sector than in the urban sector. This is true for all the survey years. Generally this is as expected, since rural areas probably are characterized by less standardized supply of consumption goods compared to the urban sector. Also note that the CV is lower in the SPIs from the QAI-model than in the AI-model, except for the urban sector in 1983. The second finding is that my SPIs include substantially larger price variation than the IPC's SPIs. Compared to Deaton and Tarozzi's UV SPIs the differences in variation are even bigger. According to the UV SPIs the differences in aggregated price levels across states are small. However, this finding can be influenced by the consumption goods that is not

[17] Such a change would not affect the main findings, although the price estimates would differ slightly.

[18] CV is a normalized measure of dispersion in a distribution. It is defined as the standard deviation over the mean.

included in the UV indices, for instance housing, which is likely to include more price variation than the high-frequency items included in Deaton and Tarozzi's SPIs.

The large price variation in my SPIs compared to the other is by large driven by a few outlier states, which in some of the years are relative extreme (especially rural areas of Assam in 1983 and 1993-94 with a high cost of living compared to all-India). Excluding them from the analysis reduces the CV value towards the values seen for the IPC's price indices. It is therefore important to assess whether these states actually have a different cost of living than previously thought, or whether there is something else that influences the results.

Looking more closely at the state-wise price indices from the QAI-model (Tables 11 to 13), the outlier states can be identified. In the rural sector, especially three of the states have a high cost of living compared to the all-India average in all survey rounds. These states are Assam, Bihar and West Bengal. Assam and Bihar also have a relatively high cost of living in urban areas. Common for the three states is their location northeast in India. Assam is part of the so-called Northeast states[19] and is connected to the rest of India only through a narrow strip in West Bengal. Because of this geographical location, in addition to large cultural differences with the rest of India, Assam has for a long time been isolated from the other Indian states. This is reflected by the fact that Assam is an important part of the Indian Government's "Look East Policy" program[20]. Assam, Bihar and West Bengal also have in common a relatively weak economic growth during the 80s and 90s compared to the rest of India. Although it is hard to track data on the magnitude of trade between the Indian states, the northeast part of the country is probably less integrated with the domestic Indian market, and hence, the trade with these states is probably smaller. This could be a potential explanation for the high cost of living found in the Engel analysis. There are other states that have a substantially different cost of living than the all-India average in one or more of the survey rounds. For instance, Andhra

[19]The description is used for the northeastern-most states in India. More specifically; Arunachal Pradesh, Assam, Meghalaya, Manipur, Mizoram, Nagaland, Sikkim and Tripura.

[20]A program to integrate with the neighbours in the Northeast Asia.

Pradesh have a low cost of living in both sectors, according to my estimates[21]. It will go beyond the scope of this thesis to analyze all the outlier states, but they should be carefully investigated in future research.

When evaluating the price differences across states, it must be remembered that I have not controlled for relative food and non-food prices. Hence, different relative prices could potentially spuriously influence the price variation found in the Engel analysis. As a robustness check I control for relative prices in Appendix A.1. As shown in that exercise, the overall disparity in the prices seems to be very robust against relative prices.

TABLE 5: Coefficient of Variation for Different SPI

(a) Rural Sector

Year	AI	QAI	IPC	UV
1983	0.318	0.316	0.101	–
1987-88	0.245	0.214	0.082	0.054
1993-94	0.242	0.212	0.104	0.067

(b) Urban Sector

Year	AI	QAI	IPC	UV
1983	0.168	0.168	0.088	–
1987-88	0.159	0.147	0.098	0.044
1993-94	0.163	0.155	0.113	0.051

Source: My own calculations based on the NSS unit record data, the IPC's poverty lines and Deaton and Tarozzi (2000).

7.1.2 Correlation Between Rural and Urban SPI

From the brief analysis of the outlier states in the previous section, it seems like the high and low cost states generally are the same in both sectors. In this section I investigate this more formally. More specifically, I compute Pearson's correlation coefficient between the SPI in the rural and the urban sector[22]. The correlation between the two sectors is expected to be high, since there probably is more interaction between the rural and the urban sector in one particular state than there is between the different Indian states. We therefore expect, given arbitrage arguments, that prices are more similar between rural and urban areas of each state than they are across different states (Deaton and Tarozzi

[21] Other rural areas with a low cost of living are Uttar Pradesh (1983, 1987-88) and Madhya Pradesh (1993-94), while rural areas of Orissa (1983), Jammu & Kashmir (1983) and Gujarat (1987-88 and 1993-94) have high cost of living compare to the all-India average.

[22] The formula for the Pearson correlation coefficient between variable Y and X is: $\frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{(\sum X^2 - \frac{(\sum X)^2}{n})(\sum Y^2 - \frac{(\sum Y)^2}{n})}}$

2000). The correlation analysis will thus give a test of how reasonable the different SPIs really are. The results from the correlation analysis of the different SPI are shown in Table 6[23].

The correlation between the two sectors in the SPIs from the Engel analysis is relatively strong and significantly different from zero. The magnitude of the coefficients is also comparable to those from Deaton and Tarozzi's UV indices presented in the fourth column. Thus, these two sets of SPIs look quite similar in respect to this correlation measure. The rural to urban correlation for the IPC's SPIs on the other hand, is low. In 1983 the correlation is positive, but very weak, for the other survey years the correlation is strongly negative. This means that states with a high price level in the rural sector generally have a low price level in the urban sector, and vice versa. The state-wise poverty lines in the base year (1973-74) were calculated separately for the rural and the urban sector using two different UV SPIs constructed with data from the early 1960s. Since then, the poverty lines in the two sectors have been deflated over time with two independent price indices (re-weighted versions of the CPIAL and the CPIIW, see Section 2.2). The original SPIs from the 1960s inherit a positive correlation between rural and urban prices, although it is not very strong. The correlation analysis here shows that the SPIs have diverged over time, suggesting that the official poverty lines are outdated.

TABLE 6: Correlation Between Rural and Urban SPI

Year	AI	QAI	IPC	UV
1983	0.859***	0.834***	0.019	–
1987-88	0.716***	0.634***	-0.387	0.723***
1993-94	0.773***	0.761***	-0.601**	0.729***

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: My own calculations based on the NSS unit record data, the IPC's poverty lines and Deaton and Tarozzi (2000).

[23] When calculating the significant level I have approximated the price estimates as constants, although the prices have been estimated themselves.

7.1.3 Correlation Between Different SPI

Table 7 presents correlation coefficients between the different SPI for various survey rounds. The correlation between the SPIs resulting from the Engel analysis and IPC's SPIs for the urban sector is actually negative, although it is not significant different from zero. This is a strange result, and implies that the two sets of indices are independent, or even that they move in opposite directions. However, my SPIs for the urban sector are positive correlated with Deaton and Tarozzi's indices, but not significant at any reasonable level. The third column presents the correlation between the IPC's and Deaton and Tarozzi's SPIs. This correlation is weakly positive, but not significant. In contrast to for the urban sector, there is a relatively strong correlation between my and the IPC's SPIs for the rural sector. Although the correlation seems to fade over time, it is still significantly different from zero at a five percent level in all survey years. This finding implies that the states generally are measured in the same direction compared to the all-India average cost of living level in the two indices. The correlation between the SPIs from the Engel analysis and Deaton and Tarozzi's SPIs is somewhat weaker, but also positive. Again, there is no significant correlation between the IPC's and Deaton and Tarozzi's SPIs.

According to the Engel SPIs and the correlation analysis discussed above, the bias in the IPC's state-wise poverty lines seems to be most serious in the urban sector.

TABLE 7: Correlation Between Different SPI

(a) Rural Sector

Year	QAI/IPC	QAI/UV	IPC/UV
1983	0.802***	–	–
1987-88	0.633***	0.462*	0.325
1993-94	0.545**	0.419*	0.343

(b) Urban Sector

Year	QAI/IPC	QAI/UV	IPC/UV
1983	-0.057	–	–
1987-88	-0.085	0.237	0.348
1993-94	-0.212	0.441*	0.112

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: My own calculations based on the NSS unit record data, the IPC's poverty lines and Deaton and Tarozzi (2000).

7.2 Price Indices Over Time

This section investigates the time dimension of the IPC's poverty lines. Tables 14 and Table 15 show state-wise price indices compared to the previous survey round, i.e., they measure the increase in cost of living for each state and sector over time. Similar price indices from the IPC and Deaton and Tarozzi (2000) are presented for comparison. Note that Deaton and Tarozzi's measures for all-India include some price information from small states and union territories not listed in the tables[24]. Table 8 summarizes the increase in cost of living in the all-India indices. First, we can be see that the Engel specifications measure a smaller increase in cost of living for both time periods compared to the IPC's poverty line deflator. This finding is as expected, given that the IPC's poverty lines are deflated over time with fixed basket consumer price indices (re-weighted versions of the CPIAL and the CPIIW). The likely biases in these kind of indices are discussed in Section 2.4.2. My finding that the fixed basket indices seems to be biased upwards, is consistent with similar studies for other countries (e.g., Hamilton 2001; Costa 2001; Beatty and Røed Larsen 2005; de Carvalho Filho and Chamon 2007).

Deaton and Tarozzi's price indices are based on the so-called Törnqvist formula, which use information from consumption patterns in all time periods to construct consumption weights. Because of this, the index formula is thought to better capture substitution in consumption than the Laspeyres index formula used in the IPC's poverty deflator (Deaton and Tarozzi 2000). Yet, the Engel analysis suggests that the UV indices overestimate the increase in cost of living. This could be due to the consumption items not including in Deaton and Tarozzi's price indices, for instance housing and services, or because of other types of index biases.

Table 14 presents the state-wise price indices for 1987-88 with 1983 as a base. From this table we see that the overestimation of the increase in cost of living found in the

[24]Strictly speaking are these numbers not completely comparable to neither my or the IPC's estimates. However, the weighting procedure I use almost reproduces Deaton and Tarozzi's all-India measures when applied to their state-wise prices. I therefore proceed by regarding the all-India numbers from the different price indices as comparable.

official figures for all India does not apply uniformly for all states. In two of the states in urban areas, and seven of the states in rural areas, the Engel specifications measure a higher increase in cost of living[25]. According to the Engel analysis, rural areas of Orissa actually experienced a decrease in cost of living in this time period. Orissa is an outlier in the SPI for 1983 with a cost of living of 54 percent higher than all-India rural. Much of this difference to all-India is therefore compensated for with a decrease in prices over time. Without studying Orissa more specifically, it is hard to conclude whether data issues or excluded variables influences my estimates. This should be investigated more closely in future work. Table 14 shows the state-wise price indices for 1993-94 with 1987-88 as a base. The upward bias in the official poverty line deflator seems to apply more uniformly for this time period. For all states and sectors except five, the Engel method estimates a smaller increase in the cost of living, compared to the IPC[26].

TABLE 8: Price Level Compared to the Previous Survey Round

(a) Rural Sector

Year	AI	QAI	IPC	UV
1987-88	121.2	121.5	128.7	–
1993-94	166.7	165.5	178.7	169.8

(b) Urban Sector

Year	AI	QAI	IPC	UV
1987-88	129.5	129.8	140.2	–
1993-94	164.8	164.0	173.5	173.8

Note: The values for 1987-88 compare prices to those of 1983, i.e., 1983=100, while the values for 1993-94 compare prices to those of 1987-88, i.e., 1987-88=100.
Source: My own calculations based on the NSS unit record data, the IPC's poverty lines and Deaton and Tarozzi (2000).

7.3 Poverty Estimation

Given the new sets of price indices, both cross-section and over time, it is possible to calculate new poverty lines with the corresponding HCRs and PGIs. According to the cost of living estimates derived through the Engel analysis, it seems that both the spatial dimension and the time dimension in the IPC's poverty lines are measured with substantial biases. Since HCRs are extremely sensitive to the price indices used to adjust the

[25]These are rural areas Andhra Pradesh, Gujarat, Karnataka, Maharashtra, Punjab, Rajasthan and Tamil Nadu, and urban areas of Rajasthan and Delhi.
[26]These are rural areas of Andhra Pradesh, Assam and Tamil Nadu, and urban areas of Bihar and Delhi.

poverty lines, it is not surprising that the poverty measures from the Engel analysis tell a different store than the official estimates. For a meaningful comparison of the different poverty estimates, it is necessary to ”anchor” my updated poverty lines to a specific IPC poverty line. Ideally I would have chosen the base poverty lines in the IPC's methodology as an anchor (1973-74). However, this is not possible with my data set. Instead I proceed by using the all-India poverty lines in 1983 as an anchor. These poverty lines are 89.50 Rs and 115.65 Rs per capita per month for the rural sector and the urban sector, respectively. From these two poverty lines I calculate new state-specific poverty lines in 1983 and state-specific poverty lines for other years, using the price indices resulting from the QAI-model. The choice of the official all-India poverty lines in 1983 as anchor is not optimal, since they themselves may be measured with bias. My poverty estimates for 1983 will therefore only capture the spatial differences between the Engel and the IPC's price measures, not the likely bias stemming from the price indices used to deflate the poverty lines in the base year (1973-74) forward to 1983.

The first column in Tables 16 to 18 presents poverty rates that are updated with the Engel price indices. The second column in these tables shows my attempt to reproduce the IPC's official poverty rates. There are minor differences in these estimates and those published by the IPC, mainly because they use interpolation rather than computations from the unit record data (see Deaton and Dréze 2002). In the calculation of the state-wise poverty rates, I use the sample weights provided in the NSS surveys. The weights correct for the specific sample design, and make the sample represent the Indian population. The IPC calculates the all-India poverty rate implicitly from the state-wise poverty rates, by adding up weighted poverty rates for each state. In addition, the IPC assigns poverty lines or poverty ratios to the smallest states and union territories[27]. The procedure to assign

[27]The IPC proceeds as follows: Poverty lines for Maharashtra are used for Goa and Dadra & Nagar Haveli. Poverty ratios for Assam are used for Sikkim, Arunachal Pradesh, Meghalaya, Mizoram, Manipur, Nagaland and Tripura. Poverty ratios for Tamil Nadu are used for Pondicherry and Andaman & Nicober Island. Urban poverty ratios for Punjab are used for both urban and rural poverty of Chandigarh. Poverty ratios for Goa is used for Daman & Diu. Poverty ratios for Kerala are used for Lakshadweep, and finally, in 1993-94 the poverty ratios in Himachal Pradesh used for Jammu & Kashmir.

neighboring price levels look more reasonable than to assign neighboring poverty ratios. Still I choose to replicate the IPC's methodology, both in my reproduction of the IPC's HCRs and in my updated HCRs derived through the Engel analysis[28].

I structure my poverty findings in two parts. First, I discuss the implications for the aggregate poverty rates, and second, I discuss the poverty rates for the different states.

7.3.1 All-India Poverty Estimates

Table 9 summarizes the all-India poverty measures for both sectors. Interestingly, the updated HCRs show a lower incident of poverty for both sectors in 1983 compared to the IPC's measures. This is perhaps an unexpected result, since the only thing that distinguishes the two sets of estimates is differences in the SPIs for 1983. The difference is largest in the rural sector, where the Engel HCR measures a four percent points lower poverty ratio. However, HCRs could give misleading conclusions in this kind of comparisons, since they only consider the incident of poverty and not the depth. Interestingly, the PGI shows the opposite result for the rural sector. According to this index, the degree of poverty is actually measured as higher with the price estimates from the Engel analysis than in the IPC's official figures.

The differences between the Engel poverty measures and the IPC's poverty measures in 1987-88 and 1993-94 will, in contrast to in 1983, be affected by different measures of changes in cost of living over time in addition to differences in the SPIs. As can be seen from Table 9, this results in larger differences between the two estimates. From 1983 to 1987-88 the Engel HCRs imply a reduction in the poverty rates of approximately eight percent points in both sectors, while the reduction in the IPC's measures are roughly seven and two-three percent points in the rural and the urban sector, respectively. Between 1987-88 and 1993-94 the Engel analysis implies a overall poverty reduction of seven and 11 percent points for the two sectors, while the corresponding reduction in the

[28]However, in contrary to in the IPC's measures, I calculate the HCR for Jammu & Kashmir in 1993-94 directly, without assuming the same poverty ratios as in Himachal Pradesh.

IPC's estimates are roughly two and six percent points for the rural and the urban sector, respectively. Thus, the Engel analysis show a much more steady decrease in poverty during the 1980s and early 1990s, with an acceleration in the urban poverty reduction between 1987-88 and 1993-94. The estimates from the PGI confirm this finding.

It is not equally straight forward to compare the Engel poverty estimates with the estimates from Deaton and Tarozzi (2000), because they use official figures from 1987-88 as an anchor. Even if we only are interested in the poverty trend, it is hard to compare the estimates, since I use a different anchor and because changes in measured poverty will depend critically on the poverty line in itself. One potential procedure is to use the implied all-India poverty lines in 1987-88 from the Engel analysis as an anchor for the UV price indices, and then calculate the corresponding HCRs and PGIs. Another is to use the official poverty lines in 1987-88 as base for the Engel price estimates[29]. Table 10 presents HCRs and PGIs from Deaton and Tarozzi's UV indices based on these two strategies[30]. When using the Engel all-India poverty lines in 1987-88 as a base, we can see that the HCRs for the rural sector are quite similar to my estimates. The poverty reduction in the urban sector is however smaller, compared to the Engel method.

TABLE 9: Poverty Measures – QAI & the IPC

(a) HCR

	Rural		**Urban**	
Year	QEC	IPC	QEC	IPC
1983	41.78	45.74	40.42	41.41
1987-88	33.49	38.94	32.42	38.90
1993-94	26.19	36.51	21.88	32.51

(b) PGI

	Rural		**Urban**	
Year	QAI	IPC	QAI	IPC
1983	14.08	13.24	11.68	11.63
1987-88	8.49	9.38	8.11	10.38
1993-94	6.00	8.42	5.01	8.29

Source: My own calculations based on the NSS unit record data and IPC's poverty lines.

[29] Another caveat is that Deaton and Tarozzi compare prices in rural and urban areas, and use only the all-India rural poverty line as a base. Hence, their estimates also adjust for price differences between the rural and the urban sector. Since this thesis analyzes the two sectors separately, it implicitly assume that the price differential in the official estimates are correct. To compare my estimates with Deaton and Tarozzi, I proceed by calculating new poverty lines using their price indices, but use both the rural and the urban poverty lines as base.

[30] Because these poverty measures have been calculated in a different way than in Deaton and Tarozzi (2000) they are not equal to the measures reported there. The exception is the column for the rural sector labeled "IPC". These numbers should be identical to the ones in the original papers. More specifically, Tables 9 and 10 in Deaton and Tarozzi (2000) (HCR), and Table 2b in Deaton and Dréze (2002). Yet, I am not able to reproduce the exact values. This could be due to rounding errors.

TABLE 10: Poverty Measures – Deaton and Tarozzi

(a) HCR

	Rural		Urban	
Year	IPC	QAI	IPC	QAI
1987-88	38.15	32.99	38.92	33.06
1993-94	31.53	26.43	31.77	26.41

(b) PGI

	Rural		Urban	
Year	IPC	QAI	IPC	QAI
1987-88	9.04	7.43	10.30	8.21
1993-94	6.74	5.39	7.83	6.10

Note: The columns labeled ”IPC” use the official all-Indian poverty rates as a base, while the columns labeled ”QAI” use the all-Indian poverty rates implied by the Engel analysis.
Source: My own calculations based on the NSS unit record data and Deaton and Tarozzi (2000).

7.3.2 State-wise Poverty Estimates

From Tables 16 to 18 we see that the incident of poverty varies largely between the Indian states. We also see that the poverty rates generally are lower in the urban sector than in the rural sector. The Engel method reveals that the rural areas with the highest proportion of poor are Assam, Bihar, Orissa and West Bengal. This is true for all survey years. Not surprisingly, the same four states are among the states with highest proportion of poor also in the urban sector, as well as in the IPC's poverty measures for both sectors. However, the IPC's figures suggest that the poverty ratios in these areas are lower, in addition to the differences to the all-India average are being smaller. Anecdotes and casual observations from Bihar and West Bengal indicate that the official poverty estimates are too optimistic for the northeast part of India[31]. This gives support to the poverty measures from the Engel analysis.

The third and fourth columns in Table 17 and Table 18 present changes in poverty rates from the previous survey round implied by the Engel HCRs and the IPC's HCRs, respectively. First consider the rural sector. For some of the states the poverty reduction is extraordinary. This is especially true for Assam, Jammu & Kashmir and Orissa between 1983 and 1987-88, and for West Bengal between 1987-88 and 1993-94, who all experienced a poverty reduction of roughly 30 percent points[32]. Notice that the poverty rate in

[31] Professor Rohini Somanathan (Delhi School of Economics, University of Delhi) pointed this out to me.

[32] As discussed earlier, Orissa is a special case with a negative growth in cost of living between 1983 and 1987-88. This influences the poverty estimates and should be investigated further in future research.

rural areas of Assam increases with 12 percent points between 1987-88 and 1993-94, and hence, some of the poverty reduction from 1983 seems to be lost. In the urban sector both Assam and Orissa experienced a poverty reduction of over 20 percent points between each survey round. Steep poverty reductions were also seen in Gujarat, Jammu & Kashmir and Kerala between 1987-88 and 1993-94, according to the Engel analysis. It is out of the scope of this thesis to study each state separately, but since regional differences in poverty reduction is an important policy issue, it should be carefully investigate in future research.

8 Concluding Remarks

In this thesis I have derived state-wise price indices, both across Indian states and over time, through Hamilton's Engel curve approach, and calculated the corresponding poverty measures. This method of revealing changes in the cost of living is consistent with consumer preferences.

The Engel analysis reveals five main findings. First, there seems to be more variation in the cost of living between the Indian states than implied by the official state-wise poverty lines. Second, the spatial dimension in the IPC's poverty lines is more seriously outdated for the urban sector than for the rural sector, according to the Engel analysis. Third, the Engel methodology suggests that the official published figures overestimate the increase in cost of living in the mid eighties and early nineties, and consequently, underestimate the poverty reduction in the same time period. Fourth, the official poverty estimates are too optimistic for the northeast part of India. Fifth, and finally, the Engel methodology reveals a much more steady poverty reduction than the IPC both in the rural and in the urban sector for the period under study.

This thesis uses the all-India poverty lines in 1983 as base for its poverty measures. Because of this, it will not capture the likely biases in the official price deflator used to transform the poverty lines in the base year (1973-74) to poverty lines in 1983. Future research should make use of all the household surveys back to 1973-74, to fully capture the bias introduced by the IPC's fixed basket price deflator. Household surveys for later years should also be incorporated.

References

Almås, Ingvild, 2007. "International Income Inequality: Measuring PPP bias by estimating Engel curves for food." Discussion paper SAM 17/2007, Norwegian School of Economics and Business Administration.

Anand, Sudhir and Paul Segal, 2008. "What Do We Know about Global Income Inequality?" *Journal of Economic Literature*, 46(1): 57–94,.

Banks, James, Richard Blundell, and Arthur Lewbel, 1997. "Quadratic Engel Curves and Consumer Demand." *Review of Economics and Statistics*, 79(4): 527–539.

Beatty, Timothy and Erling Røed Larsen, 2005. "Using Engel Curves to Estimate Bias in the Canadian CPI as a Cost of Living Index." *Canadian Journal of Economics*, 38(2): 482–499.

Bhattacharya, N. and G.S. Chatterjee, 1971. "On Rural-Urban Consumer Prices and per Capita Household Consumption in India by Levels of Living." *Sankhya*, 33: 3–4.

Bhattacharya, N. and G.S. Chatterjee, 1974. "Between State Variations in Consumer Prices and Per Capita Household Consumption in Rural India." *Srinivasan, Poverty and Income Distribution in India*, Statistical Publishing Society.

Blow, Laura, 2006. "Explaining Trends in Household Spending." IFS Working Papers 03/06.

Blundell, Richard Blundell, Panos Pashardes, and Guglielmo Weber, 1993. "What do we Learn About Consumer Demand Patterns from Micro Data?" *American Economic Review*, 83(3): 570–597.

de Carvalho Filho, I. and M. Chamon, 2007. "The Myth of Post-Reform Income Stagnation: Evidence from Brazil and Mexico." International Monetary Fund Working Paper.

Costa, Dora L., 2001. "Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves." *Journal of Political Economy*, 109: 1288–1310.

Datt, Gaurav, Valerie Kozel, and Martin Ravallion, 2003. "A Model-Based Assessment of India's Progress in Reducing Poverty in the 1990s." *Economic and Political Weekly*, 38(4): 355–361.

Deaton, Angus, 1988. "Quality, Quantity, and Spatial Variation of Price." *American Economic Review*, 78(3): 418–430.

Deaton, Angus, 1997. *The Analysis of Household Surveys : A Microeconometric Approach to Development Policy*. Johns Hopkins University Press.

Deaton, Angus, 2003. "Prices and Poverty in India, 1987-2000." *Economic and Political Weekly*, January 25: 362–368.

Deaton, Angus, 2008. "Price Trends in India and Their Implications for Measuring Poverty." *Economic and Political Weekly*, February 9: 43–49.

Deaton, Angus and Jean Dréze, 2002. "Poverty and Inequality in India: A Reexamination." *Economic and Political Weekly*, September 7: 3730–3748.

Deaton, Angus and Jean Dréze, 2008. "Nutrition In India: Facts And Interpretations." Centre for Development Economics, Delhi School of Economics, Mimeo.

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

Deaton, Angus and John Muellbauer, 1980. "An Almost Ideal Demand System." *American Economic Review*, 70(3): 312–326.

Deaton, Angus and Alessandro Tarozzi, 2000. "Prices and Poverty in India." Princeton University, Mimeo.

Engel, Ernst, 1857. "Die Productions-und Consumtionsverhltnisse des Knigreichs Sachsen." *Zeitschrift des Statistics Breaus des Kniglich Schsischen Ministeriums des Innern*, 8(1): 1–54.

Gibson, John, Steven Stillman, and Trinh Le, 2007. "CPI Bias and Real Living Standards in Russia During the Transition." *Journal of Development Economics*, 87(1): 140–160.

Government of India, 1979. "Report of the Task Force on Projections of Minimum Needs and Effective Consumption Demand." New Delhi.

Government of India, 1993. "Report of the Expert Group on Estimation of the Proportion and Number of Poor." New Delhi.

Hamilton, Bruce, 2001. "Using Engel's Law to Estimate CPI Bias." *American Economic Review*, 91(3): 619–30.

Hausman, Jerry, 2003. "Sources of Bias and Solutions to Bias in the Consumer Price Index Sources of Bias and Solutions to Bias in the Consumer Price Index." *The Journal of Economic Perspectives*, 17: 23–44.

Himanshu, Abhijit Sen and Rinku Murgai, 2009. "Towards New Poverty Lines for India." *Economic and Political Weekly*, Forthcoming.

Holla, Jayarama, Puja Guha, and T. Krishna Kumar, 2008. "Engel Curve Method for Measuring Poverty." *Economic and Political Weekly*, 43(30): 115–123.

Leser, C. E. V., 1963. "Forms of Engel Functions." *Econometrica*, 31(4): 694–703.

Meenakshi, J.V. and Ranjan Ray, 1999. "Regional Differences in India's Food Expenditure Pattern: A Complete Demand Systems Approach." *Journal of International Development*, 11: 47–74.

Milanovic, Branko, 2005. *Worlds Apart: Measuring International and Global Inequality*. Princeton, NJ: Princeton University Press.

Minhas, B., L.R. Jain, S.M. Kansal, and M.R. Saluja, 1988. "Measurement of General Cost of Living for Urban India - All India and Different States." *Sarvekshana*, 12: 1–23.

Palmer-Jones, Richard.W. and Amaresh Dubey, 2008. "Poverty Incidence in India since 1993: Another View." The School of Development Studies, University of East Anglia, Working Paper 08.

Radhakrishna, R. and C. Ravi, 2004. "Malnutrition in India. Trends and Determinants." *Economic and Political Weekly,*, 34(7): 671–676.

Ray, Debraj, 1998. *Development Economics*. Princeton University Press.

Sala-i-Martin, Xavier, 2006. "The World Distribution of Income: Falling Poverty and … Convergence, Period." *Quarterly Journal of Economics*, 121(2): 351–397.

Sen, Abhijit and Himanshu, 2005. "Poverty and Inequality in India. Getting Closer to the Truth." Jawaharlal Nehru University, Mimeo.

Sundaram, K. and Suresh Tendulkar, 2003. "Poverty in India in the 1990s: An Analysis of Changes in 15 Major States." *Economic and Political Weekly*, 38(14): 1385–1393.

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

TABLE 11: Spatial Price Indices 1983

(a) Rural Sector

State	AI	QAI	IPC
Andhra Pradesh	64.7	63.1	81.2
Assam	169.2	159.7	109.9
Bihar	137.0	142.7	108.9
Gujarat	116.4	107.3	93.1
Haryana	102.3	102.7	99.0
Himachal Pradesh	107.0	99.8	99.0
Jammu & Kashmir	143.2	128.5	102.5
Karnataka	78.8	81.4	93.1
Kerala	104.2	100.0	111.0
Madhya Pradesh	95.7	88.9	93.4
Maharashtra	63.1	62.9	98.6
Orissa	148.8	154.0	118.7
Punjab	79.0	83.7	99.0
Rajasthan	79.0	80.0	89.7
Tamil Nadu	88.6	90.1	107.4
Uttar Pradesh	71.5	71.2	93.7
West Bengal	157.3	162.4	117.9
All-India Rural	100.0	100.0	100.0
Coefficient of Variation	0.318	0.316	0.101

(b) Urban Sector

State	AI	QAI	IPC
Andhra Pradesh	82.2	80.9	92.0
Assam	137.9	139.7	84.3
Bihar	136.1	138.2	96.7
Gujarat	120.9	115.5	106.5
Haryana	94.6	94.8	89.5
Himachal Pradesh	104.4	99.9	88.4
Jammu & Kashmir	126.0	120.6	86.1
Karnataka	95.8	96.3	103.9
Kerala	116.5	117.5	106.0
Madhya Pradesh	95.8	94.5	106.2
Maharashtra	98.8	100.1	109.4
Orissa	127.2	127.6	107.9
Punjab	94.0	92.1	87.4
Rajasthan	87.1	88.9	98.2
Tamil Nadu	99.9	99.4	104.0
Uttar Pradesh	81.9	82.8	95.3
West Bengal	108.7	108.9	91.6
Delhi	92.9	95.7	106.6
All-India Urban	100.0	100.0	100.0
Coefficient of Variation	0.168	0.168	0.088

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines.

TABLE 12: Spatial Price Indices 1987-88

(a) Rural Sector

State	AI	QAI	IPC	UV
Andhra Pradesh	69.8	72.2	79.8	94.0
Assam	148.2	134.7	110.6	106.7
Bihar	131.5	126.0	104.5	104.6
Gujarat	136.9	130.8	99.8	110.5
Haryana	97.5	102.6	106.7	98.9
Himachal Pradesh	92.8	91.4	106.7	101.6
Jammu & Kashmir	115.0	110.0	107.9	95.1
Karnataka	83.7	87.2	90.7	99.3
Kerala	101.7	103.6	113.4	104.9
Madhya Pradesh	86.7	86.6	92.9	94.2
Maharashtra	75.9	76.4	100.4	103.8
Orissa	110.2	107.0	105.4	96.6
Punjab	96.2	102.2	106.7	94.2
Rajasthan	95.5	100.8	102.0	103.9
Tamil Nadu	98.1	98.7	102.6	105.5
Uttar Pradesh	73.5	76.7	99.5	91.4
West Bengal	153.5	152.5	112.2	99.2
All-India Rural	100.0	100.0	100.0	100.0
Coefficient of Variation	0.245	0.214	0.082	0.054

(b) Urban Sector

State	AI	QAI	IPC	UV
Andhra Pradesh	79.1	79.7	93.7	94.0
Assam	131.0	126.7	78.1	103.0
Bihar	121.1	120.1	92.7	100.5
Gujarat	128.4	123.7	106.8	109.5
Haryana	97.9	97.4	88.3	101.5
Himachal Pradesh	100.0	97.4	88.9	98.2
Jammu & Kashmir	140.3	134.8	91.5	92.2
Karnataka	88.7	90.6	105.6	98.2
Kerala	113.5	113.5	100.7	97.6
Madhya Pradesh	90.1	91.0	110.0	98.2
Maharashtra	109.4	110.3	116.7	107.8
Orissa	113.0	112.5	102.0	94.0
Punjab	92.0	91.6	89.4	96.6
Rajasthan	100.1	100.1	102.0	101.5
Tamil Nadu	94.5	93.7	102.3	100.8
Uttar Pradesh	85.7	86.2	95.1	98.8
West Bengal	105.2	104.3	92.5	100.1
Delhi	107.9	113.5	109.1	102.8
All-India Urban	100.0	100.0	100.0	100.0
Coefficient of Variation	0.159	0.147	0.098	0.044

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines. The fourth column is taken from Table 6 in Deaton and Tarozzi (2000); the Törnqvist Index.

TABLE 13: Spatial Price Indices 1993-94

(a) Rural Sector

State	AI	QAI	IPC	UV
Andhra Pradesh	80.5	79.9	79.2	97.9
Assam	165.3	155.1	112.7	109.3
Bihar	127.7	130.0	103.1	98.1
Gujarat	138.0	129.6	98.2	116.5
Haryana	104.2	113.0	113.6	103.3
Himachal Pradesh	97.5	95.6	113.6	104.5
Jammu & Kashmir	116.4	112.7	–	104.1
Karnataka	79.6	82.3	90.7	103.5
Kerala	110.9	109.9	118.5	112.7
Madhya Pradesh	75.7	77.7	93.8	94.2
Maharashtra	68.3	72.0	94.7	105.7
Orissa	103.4	97.3	94.3	92.8
Punjab	100.4	106.1	113.6	105.0
Rajasthan	103.3	105.3	104.9	105.5
Tamil Nadu	105.4	103.8	95.5	107.0
Uttar Pradesh	80.4	84.2	103.5	91.8
West Bengal	134.0	124.9	107.2	96.6
All-India Rural	100.0	100.0	100.0	100.0
Coefficient of Variation	0.242	0.212	0.104	0.067

(b) Urban Sector

State	AI	QAI	IPC	UV
Andhra Pradesh	82.5	81.8	98.9	94.0
Assam	123.3	122.1	75.5	105.9
Bihar	127.0	127.6	84.8	95.7
Gujarat	126.1	122.2	105.6	105.2
Haryana	90.4	91.5	91.8	100.9
Himachal Pradesh	103.6	102.1	90.1	99.3
Jammu & Kashmir	127.3	122.8	–	95.7
Karnataka	90.9	90.7	107.7	99.4
Kerala	101.5	101.8	99.7	100.5
Madhya Pradesh	83.0	83.9	112.7	94.8
Maharashtra	100.7	102.2	116.8	110.6
Orissa	90.7	90.5	106.0	90.6
Punjab	89.7	91.8	90.1	101.7
Rajasthan	102.4	101.0	99.8	99.7
Tamil Nadu	94.4	95.6	105.4	100.4
Uttar Pradesh	88.2	89.9	91.9	94.1
West Bengal	108.8	105.7	88.0	100.0
Delhi	133.5	134.9	110.0	106.3
All-India Urban	100.0	100.0	100.0	100.0
Coefficient of Variation	0.163	0.155	0.113	0.051

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines. The fourth column is taken from Table 6 in Deaton and Tarozzi (2000); the Törnqvist Index.

TABLE 14: Price Indices for 1987-88 Relative to 1983

(a) Rural Sector

State	AI	QAI	IPC
Andhra Pradesh	130.6	138.9	126.5
Assam	106.2	102.5	129.6
Bihar	116.3	107.2	123.5
Gujarat	142.5	148.1	138.1
Haryana	115.5	121.4	138.8
Himachal Pradesh	105.1	111.3	138.8
Jammu & Kashmir	97.3	104.0	135.5
Karnataka	128.8	130.0	125.4
Kerala	118.2	125.9	131.5
Madhya Pradesh	109.9	118.3	128.0
Maharashtra	145.7	147.4	131.0
Orissa	89.8	84.4	114.2
Punjab	147.5	148.3	138.8
Rajasthan	146.4	152.9	146.5
Tamil Nadu	134.1	133.0	123.0
Uttar Pradesh	124.5	130.7	136.6
West Bengal	118.2	114.1	122.4
All-India Rural	121.2	121.5	128.7

(b) Urban Sector

State	AI	QAI	IPC
Andhra Pradesh	124.7	127.8	142.7
Assam	123.0	117.8	129.8
Bihar	115.2	112.8	134.4
Gujarat	137.4	139.0	140.5
Haryana	134.1	133.4	138.4
Himachal Pradesh	124.1	126.6	140.9
Jammu & Kashmir	144.1	145.1	148.9
Karnataka	119.9	122.1	142.4
Kerala	126.1	125.5	133.1
Madhya Pradesh	121.8	125.0	145.2
Maharashtra	143.4	143.0	149.6
Orissa	115.1	114.5	132.5
Punjab	126.8	129.0	143.5
Rajasthan	148.8	146.1	145.6
Tamil Nadu	122.5	122.4	137.8
Uttar Pradesh	135.5	135.2	139.8
West Bengal	125.4	124.3	141.6
Delhi	150.3	154.1	143.5
All-India Urban	129.5	129.8	140.2

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines.

TABLE 15: Price Indices for 1993-94 Relative to 1987-88

(a) Rural Sector

State	AI	QAI	IPC	UV
Andhra Pradesh	192.3	183.2	177.3	175.9
Assam	185.8	190.5	182.1	173.7
Bihar	161.8	170.7	176.3	159.7
Gujarat	168.1	163.9	175.7	170.6
Haryana	178.1	182.2	190.2	174.2
Himachal Pradesh	175.1	172.9	190.2	167.1
Jammu & Kashmir	168.8	169.4	–	181.5
Karnataka	158.4	156.1	178.7	175.1
Kerala	181.7	175.4	186.7	172.3
Madhya Pradesh	145.6	148.4	180.5	171.9
Maharashtra	149.9	156.0	168.6	172.6
Orissa	156.3	150.5	159.8	164.6
Punjab	174.0	171.7	190.2	190.7
Rajasthan	180.2	172.9	183.7	166.9
Tamil Nadu	179.1	174.1	166.2	167.7
Uttar Pradesh	182.4	181.7	185.9	167.9
West Bengal	145.5	135.5	170.8	166.5
All-India Rural	166.7	165.5	178.7	169.8

(b) Urban Sector

State	AI	QAI	IPC	UV
Andhra Pradesh	171.8	168.4	183.1	177.2
Assam	155.1	157.9	167.8	177.7
Bihar	172.9	174.2	158.7	165.2
Gujarat	161.9	161.9	171.6	165.4
Haryana	152.2	154.1	180.3	177.6
Himachal Pradesh	170.7	172.0	176.0	175.2
Jammu & Kashmir	149.5	149.4	–	178.5
Karnataka	168.8	164.1	176.9	177.1
Kerala	147.4	147.0	171.8	173.5
Madhya Pradesh	151.9	151.2	177.8	170.9
Maharashtra	151.8	152.0	173.7	181.1
Orissa	132.3	131.9	180.3	167.8
Punjab	160.7	164.4	174.9	187.1
Rajasthan	168.6	165.5	169.8	171.8
Tamil Nadu	164.6	167.3	178.9	170.5
Uttar Pradesh	169.6	171.0	167.8	165.4
West Bengal	170.5	166.2	165.1	170.6
Delhi	204.0	194.8	174.9	175.7
All-India Urban	164.8	164.0	173.5	173.8

Note: These price indices compare prices to those of 1987-88, i.e., 1987-88=100.
Source: My own calculations based on the NSS unit record data and the IPC's poverty lines. The fourth column is taken from Table 3 in Deaton and Tarozzi (2000); the Törnqvist Index.

TABLE 16: Headcount Ratios 1983

(a) Rural Sector

State	QAI	IPC
Andhra Pradesh	11.12	26.46
Assam	83.54	43.53
Bihar	83.84	64.85
Gujarat	42.57	29.33
Haryana	25.43	23.28
Himachal Pradesh	19.83	18.93
Jammu & Kashmir	54.81	27.84
Karnataka	27.06	35.81
Kerala	31.22	39.16
Madhya Pradesh	45.08	49.58
Maharashtra	12.92	46.29
Orissa	85.66	68.18
Punjab	7.87	14.25
Rajasthan	26.84	34.06
Tamil Nadu	40.79	54.48
Uttar Pradesh	24.54	46.87
West Bengal	82.43	63.77
All-India Rural	41.78	45.74

(b) Urban Sector

State	QAI	IPC
Andhra Pradesh	27.10	37.52
Assam	67.14	22.90
Bihar	73.62	48.65
Gujarat	46.85	39.73
Haryana	34.06	29.32
Himachal Pradesh	17.27	10.27
Jammu & Kashmir	56.21	17.09
Karnataka	37.95	42.34
Kerala	53.33	45.70
Madhya Pradesh	42.02	52.45
Maharashtra	34.62	40.48
Orissa	64.45	49.60
Punjab	26.54	23.35
Rajasthan	30.98	38.48
Tamil Nadu	45.77	49.21
Uttar Pradesh	38.15	50.92
West Bengal	45.09	32.72
Delhi	19.73	27.68
All-India Urban	40.42	41.41

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines.

TABLE 17: Headcount Ratios 1987-88

(a) Rural Sector

State	QAI	IPC	Δ QAI (38-43)	Δ IPC (38-43)
Andhra Pradesh	12.21	21.00	1.08	-5.47
Assam	55.45	39.45	-28.08	-4.08
Bihar	66.60	53.91	-17.24	-10.94
Gujarat	50.99	28.59	8.42	-0.74
Haryana	12.35	15.34	-13.08	-7.93
Himachal Pradesh	5.37	16.68	-14.46	-2.25
Jammu & Kashmir	22.02	25.90	-32.79	-1.94
Karnataka	25.40	32.61	-1.66	-3.21
Kerala	18.46	29.40	-12.77	-9.76
Madhya Pradesh	31.16	42.01	-13.92	-7.56
Maharashtra	14.89	40.88	1.97	-5.41
Orissa	54.20	58.66	-31.47	-9.52
Punjab	7.79	12.75	-0.08	-1.50
Rajasthan	28.26	33.30	1.42	-0.76
Tamil Nadu	38.47	46.30	-2.32	-8.18
Uttar Pradesh	15.92	41.83	-8.62	-5.05
West Bengal	71.97	48.80	-10.45	-14.97
All-India Rural	33.49	38.94	-8.29	-6.80

(b) Urban Sector

State	QAI	IPC	Δ QAI (38-43)	Δ IPC (38-43)
Andhra Pradesh	22.89	41.10	-4.21	3.59
Assam	43.67	11.32	-23.48	-11.58
Bihar	65.48	51.89	-8.15	3.24
Gujarat	45.00	38.49	-1.85	-1.24
Haryana	20.12	18.38	-13.94	-10.94
Himachal Pradesh	7.23	7.23	-10.04	-3.05
Jammu & Kashmir	42.37	14.96	-13.84	-2.12
Karnataka	30.37	49.19	-7.58	6.86
Kerala	43.72	40.04	-9.61	-5.66
Madhya Pradesh	28.59	47.15	-13.43	-5.30
Maharashtra	31.61	40.26	-3.00	-0.22
Orissa	44.06	42.58	-20.39	-7.02
Punjab	10.69	13.70	-15.85	-9.65
Rajasthan	31.08	37.89	0.10	-0.59
Tamil Nadu	29.31	40.20	-16.47	-9.01
Uttar Pradesh	31.96	44.94	-6.19	-5.98
West Bengal	37.69	33.74	-7.40	1.02
Delhi	13.54	15.54	-6.18	-12.14
All-India Urban	32.42	38.90	-6.24	-2.51

Note: The third and fourth columns present the difference in the HCRs between the 38th and the 43rd NSS household survey.
Source: My own calculations based on the NSS unit record data and the IPC's poverty lines.

TABLE 18: Headcount Ratios 1993-94

(a) Rural Sector

State	QAI	IPC	Δ QAI (43-50)	Δ IPC (43-50)
Andhra Pradesh	9.35	15.89	-2.85	-5.10
Assam	67.40	44.90	11.95	5.46
Bihar	67.30	57.95	0.70	4.04
Gujarat	34.22	22.16	-16.76	-6.43
Haryana	17.01	28.26	4.67	12.91
Himachal Pradesh	8.38	30.36	3.00	13.68
Jammu & Kashmir	9.54	30.36	-12.49	+4.46
Karnataka	13.67	30.11	-11.74	-2.50
Kerala	12.84	25.38	-5.61	-4.02
Madhya Pradesh	14.98	40.72	-16.18	-1.30
Maharashtra	10.39	37.91	-4.51	-2.97
Orissa	39.48	49.81	-14.72	-8.85
Punjab	4.72	11.69	-3.07	-1.06
Rajasthan	16.62	26.40	-11.64	-6.91
Tamil Nadu	28.30	32.92	-10.18	-13.38
Uttar Pradesh	15.97	42.31	0.05	0.48
West Bengal	42.72	41.18	-29.25	-7.61
All-India Rural	26.19	36.51	-7.3	-2.43

(b) Urban Sector

State	QAI	IPC	Δ QAI (43-50)	Δ IPC (43-50)
Andhra Pradesh	14.64	38.82	-8.25	-2.28
Assam	32.80	7.93	-10.86	-3.40
Bihar	58.19	34.84	-7.29	-17.05
Gujarat	29.16	28.28	-15.85	-10.21
Haryana	9.01	16.47	-11.10	-1.91
Himachal Pradesh	9.01	9.26	1.78	2.03
Jammu & Kashmir	15.64	9.26	-26.73	-5.70
Karnataka	21.23	39.90	-9.14	-9.29
Kerala	17.61	24.31	-26.11	-15.72
Madhya Pradesh	16.34	48.08	-12.25	0.93
Maharashtra	19.50	34.99	-12.12	-5.27
Orissa	21.49	40.64	-22.56	-1.94
Punjab	6.20	10.90	-4.50	-2.80
Rajasthan	21.76	31.02	-9.32	-6.87
Tamil Nadu	22.45	39.91	-6.86	-0.29
Uttar Pradesh	24.46	35.09	-7.51	-9.84
West Bengal	25.49	22.95	-12.20	-10.79
Delhi	19.86	16.09	6.32	0.55
All-India Urban	21.88	32.51	-10.54	-6.39

Note: The third and fourth columns present the difference in the HCRs between the 43rd and the 50th NSS household survey.
Source: My own calculations based on the NSS unit record data and the IPC's poverty lines.

TABLE 19: Poverty Gap Index

	QAI			IPC		
	1983	1987-88	1993-94	1983	1987-88	1993-94
Rural						
Andhra Pradesh	2.31	2.24	1.58	5.91	4.31	2.88
Assam	26.96	12.67	16.29	8.77	7.45	8.21
Bihar	33.24	18.68	19.11	20.17	12.93	14.66
Gujarat	10.20	12.05	7.29	6.25	5.49	4.07
Haryana	7.38	2.54	3.07	6.74	3.62	5.60
Himachal Pradesh	5.56	0.81	1.25	5.44	2.63	5.55
Jammu & Kashmir	13.33	3.75	1.53	6.07	4.52	5.55
Karnataka	6.68	5.71	2.44	9.77	7.86	6.27
Kerala	7.31	3.56	2.44	10.07	6.34	5.61
Madhya Pradesh	12.29	7.05	2.72	13.97	10.58	9.48
Maharashtra	3.03	2.50	2.08	12.64	9.57	9.28
Orissa	35.35	14.51	8.41	22.70	16.27	11.96
Punjab	2.55	1.04	0.60	3.79	1.92	1.90
Rajasthan	6.93	7.02	2.96	9.45	8.64	5.21
Tamil Nadu	11.61	9.54	6.08	17.47	12.60	7.30
Uttar Pradesh	5.44	2.67	2.69	12.71	9.85	10.37
West Bengal	35.82	22.48	8.87	21.16	11.58	8.29
All-India Rural	*14.08*	*8.49*	*6.00*	*13.24*	*9.38*	*8.42*
Urban						
Andhra Pradesh	6.49	4.82	2.69	9.60	10.57	9.27
Assam	20.94	10.02	6.14	5.65	1.49	0.91
Bihar	28.08	20.71	17.41	13.47	13.00	7.86
Gujarat	11.94	10.51	6.49	9.30	8.23	6.22
Haryana	9.53	3.89	1.58	8.23	3.56	3.05
Himachal Pradesh	5.63	0.83	1.18	4.47	0.73	1.25
Jammu & Kashmir	12.38	9.43	2.50	3.09	2.41	1.25
Karnataka	10.74	7.30	4.35	12.87	14.09	11.36
Kerala	17.62	11.80	3.72	14.16	10.49	5.53
Madhya Pradesh	10.67	6.05	2.98	14.72	13.56	13.43
Maharashtra	9.42	8.87	4.92	11.84	12.29	10.15
Orissa	20.50	11.79	4.42	13.85	11.12	11.41
Punjab	7.58	1.74	0.88	6.63	2.28	1.69
Rajasthan	8.63	7.14	4.47	11.10	9.64	7.00
Tamil Nadu	13.05	7.37	4.81	14.59	11.47	10.23
Uttar Pradesh	9.98	7.20	5.53	14.54	12.22	9.03
West Bengal	13.05	8.64	5.49	8.12	7.44	4.55
Delhi	4.01	2.43	4.87	6.08	2.87	3.90
All-India Urban	*11.68*	*8.11*	*5.01*	*11.63*	*10.38*	*8.29*

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines.

Appendices

A Sensitivity Analysis

A.1 Relative Food and Non-Food Prices

As a sensitivity test I control for relative food and non-food prices. The standard empirical specification of the AI-model with relative prices is given by Equation (12). The corresponding budget share equation in the QAI-model is expressed as:

$$\begin{aligned} m_{h,s,r,j} = \alpha + \beta(\ln\ y_{h,s,r,j} - \ln\ P_{s,j}) + \frac{\lambda}{b(\mathbf{p}_{s,r,j})}(\ln\ y_{h,s,r,j} - \ln\ P_{s,j})^2 \\ + \gamma(\ln\ P_{f,s,r,j} - \ln\ P_{n,s,r,j}) + \theta X_{h,s,r,j} + \varepsilon_{h,s,r,j}, \end{aligned} \tag{21}$$

The Indian price indices for the rural and the urban sector do not report relative food and non-food prices in levels. However, the price indices are separated in different consumption categories, which make it possible to distinguish between relative inflation rates. Thus, I proceed by using the relative inflation rates as a proxy for relative price levels. However, this is only possible for the urban sector. The price index for the rural sector (the CPIAL) is limited to the state level, and since variation in relative inflation rates between different states will be perfectly correlated with the state dummy variables, there is no unique way to identify the γ-coefficient. For the urban sector on the other hand, the price data is compiled from 78 centres all over India, which may provide enough variation to identify γ for this sector. When using relative inflation rates as proxies for relative price levels, I implicitly assume that the relative food and non-food prices were equal across states in the 1960 (the base year for the CPIIW). If relative prices since then have diverged, there should be possible to identify the price effect. On the contrary, if relative prices have converged over time, it is harder to assess what the $\gamma-$coefficient would pick up. However, it seems more likely that relative prices in India have diverge since 1960, due to uneven economic progress and exposure to international trade across states and regions. To estimate the relative price effect for urban sector, the CPIIW price data is connected to the NSS consumer expenditure surveys. I proceed by matching each NSS region to the nearest geographical price centre in the same state in the urban price index[33].

Table 20 presents estimation of the AI-model and the QAI-model for the urban sector when relative food and non-food prices are included. We can see that relative prices have a significant effect. A positive coefficient means that the share of food expenditure in the total budget increases when relative prices for food increase. Since food is a necessary good, this seems economically plausible. Compared to similar studies for other countries, the estimated price effect is relatively large. For instance, Hamilton finds a relative price coefficient of 0.0368 based on household data from the United States and the linear Engel specification (AI) (Hamilton 2001). When comparing the magnitude of the relative price effect with other studies, it is useful to compute elasticities. The price effect for the linear specification found here implies an approximate uncompensated price elasticity of -0.72[34]. The price elasticity in the quadratic specification will depend on the size of the household's total real expenditure[35]. As an approximation, I calculate the uncompensated price elasticity at the mean real-income level. This gives an elasticity of -0.83. These elasticities seem large compare to calculated elasticities for other countries. Hamilton (2001) calculates a price elasticity of -0.65 in his data sample, although this elasticity covers only food at home, while my price elasticity covers total food. Meenakshi and Ray (1999) uses quartile-data for India (based on the NSS surveys) from 1972-73 to 1987-88 and the AI-model, and calculates a price elasticity of -0.78 for cereals and cereals substitutes. The price elasticities for other food groups are generally lower. Although

[33]Each state is divided into region in the NSS households surveys. The merging of prices and household data could be done in many different ways. However, without detailed knowledge of the price structure in the different geographical regions in India, it seems most reasonable to assign relative prices according to this mechanical rule. I do not have price data from Himachal Pradesh. Instead I assign relative prices from the nearest price centre in Punjab.

[34]The price elasticity in the budget share equation can be calculated as $\frac{\partial m}{\partial p}\frac{1}{m}-1$. For the AI-model this gives $-1+[\frac{\gamma-\alpha b}{m}]$, where α is the share of food in the total price index. As an approximation, I substitute α and m with the population weighted average of food share.

[35]For the QAI the price elasticity can be expressed as $\frac{\gamma}{m}-[b_1+2b_2\ln(\frac{y}{p})]\frac{\alpha}{m}-1$.

my estimated price elasticities seem large compared to the other studies, they are not implausible large.

TABLE 20: Regression – Relative Prices

	AI	QAI
Ln Exp	-0.1422	0.1884
	(0.00075)	(0.00477)
$(\text{Ln Exp})^2$	–	-0.0259
		(0.00037)
Relative Prices	0.0907	0.0532
	(0.00926)	(0.00814)
No of Children	0.0203	0.0184
	(0.00021)	(0.00021)
No of Adults	0.0256	0.0250
	(0.00026)	(0.00024)
Age of Household Head	0.0004	0.0005
	(0.00003)	(0.00003)
Constant	1.4014	0.3582
	(0.00517)	(0.01532)
Observations	118219	118219
Adj. R^2	0.388	0.412

Note: Robust standard errors in parentheses. Dummy coefficients are not reported.

A.1.1 Cost of Living When Relative Prices Differ

The cost function in the AI-model in the two-goods case is given by:

$$\ln\ C(p_f, p_n, u) = a_0 + a_f \ln p_f + a_n \ln p_n + \gamma_{fn} \ln p_f \ln p_n + u\beta_0 b(\mathbf{p})\ , \tag{22}$$

where b($\mathbf{p}$) is a price index:

$$b(\mathbf{p}) = p_f^{\beta} p_n^{-\beta}. \tag{23}$$

The price index included in the estimated budget share equation in (12), equals $\ln a(\mathbf{p})$, and is given by:

$$\ln\ P = \ln a(\mathbf{p}) = a_0 + a_f \ln p_f + a_n \ln p_n + \gamma_{fn} \ln p_f \ln p_n\ . \tag{24}$$

The true cost of living index (COLI) comparing price situation $\mathbf{p}^1$ with price situation $\mathbf{p}^0$ can generally be defined as:

$$COLI(\mathbf{p}^1, \mathbf{p}^0, \bar{u}) = \frac{C(\bar{u}, \mathbf{p}^1)}{C(\bar{u}, \mathbf{p}^0)}, \tag{25}$$

where $\bar{u}$ is a reference utility level. Hence, the logarithmic COLI in the two-goods case can be expressed through the cost function as follows:

$$\ln\ COLI(\mathbf{p}^1, \mathbf{p}^0, \bar{u}) = [\ln P^1 - \ln P^0] + \bar{u}\ \beta_0 [b(\mathbf{p}^1) - b(\mathbf{p}^0)]\ . \tag{26}$$

As $b(\mathbf{p})$ is homogenous of degree zero in prices, it will only differ across states if relative prices differ. Hence, if relative prices and preferences (i.e., the coefficients) are the same in all situations being compared, the second term in the expression above would only have a level effect (equal for all states at a reference utility level), and $\ln P$ would pick up the relative differences in cost of living. This is what has been utilized in the main analysis of this thesis. However, when relative food and non-food prices differ between states and time periods, the COLI will depend on the relative prices as well as the utility level used for comparisons. The first component of the expression in (26) can be interpreted as the uniform or average change in cost of living, while the second component capture the marginal effect of changes in relative prices. With comparable food and non-food prices it should be possible to calculate the complete COLI for a given references utility level, $\bar{u}$, using the expression in (26) and the corresponding indirect utility function given by:

$$u = \frac{\ln y - \ln P}{\beta_0 b(\mathbf{p})} \tag{27}$$

The cost function in the two-goods version of the QAI-model is given by:

$$\ln\ C(u,p_f,p_n) = a_0 + a_f \ln p_f + a_n \ln p_n + \gamma_{fn} \ln p_f \ln p_n + \frac{ub(\mathbf{p})}{1-u\lambda(\mathbf{p})}\ , \tag{28}$$

where:

$$\lambda(\mathbf{p}) = \lambda \ln p_f - \lambda \ln p_n$$

The COLI for a given reference utility level, $\bar{u}$, can therefore be expressed through the cost function as:

$$\ln\ COLI(\mathbf{p}^1,\mathbf{p}^0,\bar{u}) = [\ln P^1 - \ln P^0] + \bar{u}[\frac{b(\mathbf{p}^1)}{1-\bar{u}\lambda(\mathbf{p}^1)} - \frac{b(\mathbf{p}^0)}{1-\bar{u}\lambda(\mathbf{p}^0)}]\ . \tag{29}$$

Since both $b(\mathbf{p})$ and $\lambda(\mathbf{p})$ are homogenous of degree zero in prices, these will only differ across states if relative prices differ. The corresponding indirect utility function follows as:

$$u = \frac{\ln y - \ln P}{\lambda(\mathbf{p})(\ln y - \ln P) + b(\mathbf{p})} \tag{30}$$

It is not straightforward to estimate the full COLI in the QAI-model, since I do not have estimates of the parameter λ. Here I use an approximation of λ, based on the coefficient for the quadratic term in Equation (21) ($\lambda/b(\mathbf{p})$), and the population weighted average of $b(\mathbf{p})$ over the full sample.

In the literature inspired by Hamilton (2001), the marginal relative price component present in the COLI is generally ignored, even if relative food and non-food prices are used as a control variable in the estimation. Yet the identified prices are referred to as "cost of living" measures. In the following I present estimates of this conventional identification strategy, which I hereafter referred to as "Hamilton-prices", as well as the full COLI identification described above. I proceed by using the mean utility level as a references in the estimation of the COLI[36].

[36]When calculation the COLI for the AI-model I am ignoring the β_0-parameter present in the indirect utility expression. This will only have an level effect.

The calculations are presented in Tables 21 to 23. For the two first survey rounds the differences between the Hamilton-prices and the COLI are negligible. This is not surprising, given the small differences in food and non-food inflation rates shown in the summary statistics in Table 2. However, for the third survey round (1993-94), the two estimates differ. For instance, the COLI measures a smaller increase in the cost of living from 1987-88 to 1993-94 compared to the Hamilton-prices. This finding could be due to the Hamilton-prices not adequately capturing substitution in consumption. As the COLI measures a lower increase in cost of living over time, the corresponding poverty lines also increase by less, and hence, the measured poverty decrease is steeper. This finding strengthens my poverty results in the main analysis. However, we should keep in mind that I have used relative inflation rates instead of relative price levels. This might cause an inaccurate identification of the COLI. The Hamilton-prices based on the estimation in Table 20, would be more robust to the inclusion of relative prices, since they do not use the relative prices directly in the identification. The poverty measures with the Hamiton-prices are roughly comparable to those found without relative prices, although the poverty reduction in the second time period is somewhat steeper. Again, this strengthens the poverty findings in the main analysis of this thesis.

TABLE 21: SPI Urban Sector – Relative Prices

	Hamilton			COLI		
	1983	1987-88	1993-94	1983	1987-88	1993-94
AI						
Andhra Pradesh	82.2	79.2	81.3	82.2	79.2	72.9
Assam	137.9	131.1	119.8	137.9	131.1	108.4
Bihar	136.1	121.2	139.0	136.2	121.3	147.3
Gujarat	121.0	128.4	123.2	121.0	128.5	106.8
Haryana	94.6	97.9	94.6	94.5	97.8	93.9
Himachal Pradesh	104.4	100.0	107.5	104.3	99.9	93.2
Jammu & Kashmir	121.7	128.6	121.8	115.8	113.6	106.6
Karnataka	95.8	88.8	88.2	95.8	88.9	80.3
Kerala	116.5	113.6	96.1	116.4	113.6	83.6
Madhya Pradesh	95.8	90.2	83.6	95.7	90.2	77.2
Maharashtra	98.9	109.6	102.0	98.9	109.7	109.7
Orissa	127.1	113.0	91.5	127.1	113.0	86.1
Punjab	94.1	92.1	86.1	94.1	92.2	77.1
Rajasthan	87.1	100.2	103.9	87.1	100.3	108.1
Tamil Nadu	100.0	94.6	92.0	100.1	94.7	84.9
Uttar Pradesh	81.9	85.8	89.8	81.8	85.8	94.4
West Bengal	108.7	105.3	111.5	108.7	105.4	105.5
Delhi	93.0	108.0	126.4	92.9	108.0	108.8
All-India Urban	100.0	100.0	100.0	100.0	100.0	100.0
Coefficient of Variation	0.166	0.148	0.164	0.164	0.140	0.183
QAI						
Andhra Pradesh	80.8	79.6	81.3	80.8	79.7	72.0
Assam	140.0	127.1	119.6	140.0	127.1	107.0
Bihar	138.7	120.4	134.1	138.8	120.5	143.1
Gujarat	115.6	124.0	120.6	115.7	124.1	102.9
Haryana	94.6	97.3	93.8	94.5	97.2	92.9
Himachal Pradesh	99.8	97.3	104.2	99.7	97.2	89.0
Jammu & Kashmir	118.3	128.5	120.2	111.7	111.6	103.6
Karnataka	96.3	90.6	89.5	96.2	90.7	80.5
Kerala	117.4	113.6	99.0	117.3	113.6	84.8
Madhya Pradesh	94.4	90.9	84.3	94.4	90.9	77.1
Maharastra	100.0	110.3	102.7	100.0	110.4	111.5
Orissa	127.9	112.7	90.9	127.9	112.7	84.9
Punjab	92.1	91.5	90.2	92.1	91.5	79.7
Rajasthan	88.8	100.1	101.8	88.8	100.2	106.4
Tamil Nadu	99.5	93.8	94.2	99.6	93.9	86.1
Uttar Pradesh	82.8	86.3	90.8	82.7	86.3	96.0
West Bengal	109.0	104.4	106.9	108.9	104.4	100.4
Delhi	95.5	113.3	131.4	95.4	113.3	111.2
All-India Urban	100.0	100.0	100.0	100.0	100.0	100.0
Coefficient of Variation	0.168	0.141	0.155	0.167	0.132	0.175

Source: My own calculations based on the NSS unit record data.

TABLE 22: Prices Over Time Urban Sector – Relative Prices

	1983 to 1987-88		1987-88 to 1993-94	
	Hamilton	COLI	Hamilton	COLI
AI				
Andhra Pradesh	124.7	124.7	152.4	126.2
Assam	123.0	123.0	135.6	113.3
Bihar	115.2	115.2	170.3	166.6
Gujarat	137.4	137.4	142.4	113.9
Haryana	134.0	133.8	143.5	131.7
Himachal Pradesh	124.0	123.8	159.4	128.0
Jammu & Kashmir	136.8	126.9	140.5	128.6
Karnataka	119.9	119.9	147.5	124.0
Kerala	126.2	126.1	125.6	98.6
Madhya Pradesh	121.9	122.1	137.6	115.1
Maharashtra	143.4	143.5	138.1	130.1
Orissa	115.0	114.9	120.2	104.6
Punjab	126.7	126.7	138.6	114.7
Rajasthan	148.9	148.9	153.9	147.8
Tamil Nadu	122.5	122.4	144.2	123.0
Uttar Pradesh	135.6	135.7	155.4	150.8
West Bengal	125.4	125.4	157.2	137.2
Delhi	150.4	150.4	173.7	138.1
All-India Urban	129.4	129.3	148.4	137.1
QAI				
Andhra Pradesh	127.7	127.6	158.3	128.1
Assam	117.6	117.7	145.9	119.3
Bihar	112.6	112.5	172.6	168.3
Gujarat	139.0	138.9	150.8	117.6
Haryana	133.4	133.2	149.4	135.6
Himachal Pradesh	126.4	126.2	165.9	129.8
Jammu & Kashmir	140.8	129.4	145.0	131.5
Karnataka	122.0	122.1	153.0	125.9
Kerala	125.4	125.4	135.1	103.1
Madhya Pradesh	124.8	125.1	143.7	117.5
Maharashtra	143.0	143.1	144.4	134.9
Orissa	114.2	114.1	125.0	106.9
Punjab	128.8	128.7	152.8	123.5
Rajasthan	146.1	146.1	157.6	150.6
Tamil Nadu	122.2	122.1	155.6	130.0
Uttar Pradesh	135.1	135.2	163.2	157.6
West Bengal	124.2	124.2	158.8	136.3
Delhi	153.9	153.8	179.7	139.2
All-India Urban	129.6	129.5	155.0	141.8

Source: My own calculations based on the NSS unit record data.

TABLE 23: Poverty Measures Urban Sector – Relative Prices

	HCR		**PGI**	
Year	Hamilton	COLI	Hamilton	COLI
1983	40.40	40.36	11.68	11.67
1987-88	32.31	32.22	8.08	8.06
1993-94	18.63	11.51	4.14	2.50

Source: My own calculations based on the NSS unit record data.

A.2 Uniform Households

In my main model-specification I control for the household size and composition in the estimation by including variables for number of children and number of adults belonging to the household. In this section I perform another sensitivity test by using only households consisting of two adults and two children. If the cost of living differ between population groups, it could be preferable not to aggregate them, but instead analyze them separately (Hamilton 2001). In addition, if expenditure patterns and preferences differ between sub-groups, the Engel methodology with aggregation will give inaccurate price estimates. The sub-population consisting of two children and two adults includes 18.286 rural households and 10.819 urban households, which equals roughly nine percent of the sample used in the main model.

Table 24 presents the regression results from this specific population group. Although the estimated parameters differ from the ones in Table 3, they are quite similar. Table 25 and Table 26 show the implied SPIs for the rural and the urban sector, respectively. The last row in these tables presents the correlation with the corresponding SPIs from the main model (Tables 11 to 13). We see that the correlation is strong for all SPIs, although the state-wise prices differ somewhat. Table 27 and Table 28 show the increase in cost of living over time. For the first time period the change in cost of living is basically the same as in the main model. For the second period on the other hand, there seems to be differences. For the rural sector the estimation from the sub-sample suggests a 10 percent higher increase in the cost of living compared to the main model, while for the urban sector the estimate are roughly 14 percent lower. As a consequence, the poverty decline in the urban sector is steeper, while the poverty decline in the rural sector is slightly weaker. Still, the main findings in this thesis are not seriously affected.

TABLE 24: Regression – Uniform Households

	AI		QAI	
	Rural	Urban	Rural	Urban
Ln Exp	-0.1253	-0.1550	0.2456	0.2655
	(0.00235)	(0.00236)	(0.01466)	(0.02093)
$(\text{Ln Exp})^2$	–	–	-0.0321	-0.0322
			(0.00127)	(0.00159)
Age of Household Head	0.0011	0.0007	0.0010	0.0008
	(0.00009)	(0.00014)	(0.00009)	(0.00013)
Constant	1.3565	1.5652	0.2884	0.1967
	(0.01463)	(0.01595)	(0.04240)	(0.06877)
Observations	18286	10819	18286	10819
Adj. R^2	0.294	0.448	0.316	0.468

Note: Robust standard errors in parentheses. Dummy coefficients are not reported.

TABLE 25: SPI Rural Sector – Uniform Households

	AI			QAI		
	1983	1987-88	1993-94	1983	1987-88	1993-94
Andhra Pradesh	66.9	70.3	86.2	59.3	69.5	82.8
Assam	166.7	147.8	158.9	166.9	142.9	162.9
Bihar	130.8	134.8	116.8	144.2	134.1	123.6
Gujrat	108.8	138.1	138.4	96.3	138.0	133.1
Haryana	114.7	105.7	120.7	110.5	106.4	136.8
Himachal Pradesh	117.8	96.4	91.8	108.5	91.1	89.8
Jammu and Kashmir	152.1	103.7	100.5	137.1	95.3	96.6
Karnataka	79.0	85.9	84.0	77.0	84.7	82.7
Kerala	116.4	113.5	112.9	107.5	113.3	112.8
Madhya Pradesh	96.3	83.3	76.4	85.2	81.6	76.6
Maharastra	68.3	77.1	70.8	62.7	74.5	71.0
Orissa	139.3	107.4	106.0	151.4	105.3	104.8
Punjab	87.2	104.7	103.6	82.3	100.5	103.6
Rajasthan	73.1	82.1	120.4	72.6	86.3	121.0
Tamil Nadu	96.2	103.6	109.9	92.7	102.6	104.9
Uttar Pradesh	69.4	71.9	75.2	64.5	72.2	74.9
West Bengal	158.2	149.8	130.2	185.4	160.1	127.4
All-India Rural	100.0	100.0	100.0	100.0	100.0	100.0
Coefficient of Variation	0.302	0.244	0.226	0.361	0.257	0.241
Correlation	0.978	0.967	0.943	0.984	0.971	0.936

Note: The row labeled ”correlation” presents the correlation coefficients with the corresponding price indices in the main analysis.
Source: My own calculations based on the NSS unit record data.

TABLE 26: SPI Urban Sector – Uniform Households

	AI			QAI		
	1983	1987-88	1993-94	1983	1987-88	1993-94
Andhra Pradesh	81.8	77.2	80.7	79.5	76.3	79.5
Assam	177.8	133.8	121.5	208.4	125.6	122.8
Bihar	136.0	124.5	133.0	145.8	131.7	134.6
Gujrat	127.9	119.7	119.4	121.3	115.0	116.5
Haryana	81.7	124.5	97.8	78.7	116.5	95.8
Himachal Pradesh	106.1	112.4	113.6	100.7	105.6	106.2
Jammu and Kashmir	118.9	135.9	139.7	115.6	146.2	140.2
Karnataka	89.3	86.5	87.7	90.1	85.7	86.9
Kerala	122.4	119.6	100.2	129.8	122.9	101.2
Madhya Pradesh	97.8	94.4	88.7	99.1	93.5	88.8
Maharastra	101.2	109.0	102.3	101.5	109.9	106.8
Orissa	130.1	114.8	92.6	131.4	113.8	90.2
Punjab	93.3	90.1	93.9	90.1	88.1	94.6
Rajasthan	89.8	105.7	107.4	90.8	110.3	105.1
Tamil Nadu	101.8	88.2	90.1	99.7	85.6	87.9
Uttar Pradesh	77.9	86.2	96.4	78.1	85.8	97.4
West Bengal	105.0	104.1	101.6	103.1	102.7	99.4
Delhi	90.6	112.4	126.9	88.4	122.0	130.4
All-India	100.0	100.0	100.0	100.0	100.0	100.0
Coefficient of Variation	0.232	0.161	0.161	0.291	0.175	0.168
Correlation	0.928	0.895	0.936	0.896	0.917	0.937

Note: The row labeled ”correlation” presents the correlation coefficients with the corresponding price indices in the main analysis.
Source: My own calculations based on the NSS unit record data.

TABLE 27: Prices Over Time Rural Sector – Uniform Households

	1983 to 1987-88		1987-88 to 1993-94	
	AI	QAI	AI	QAI
Andhra Pradesh	130.9	139.7	214.8	204.3
Assam	110.4	102.2	188.5	195.3
Bihar	128.3	111.0	151.8	158.0
Gujrat	158.0	171.0	175.7	165.3
Haryana	114.7	114.9	200.0	220.4
Himachal Pradesh	101.9	100.2	166.8	168.9
Jammu and Kashmir	84.8	83.0	169.9	173.7
Karnataka	135.4	131.3	171.4	167.4
Kerala	121.4	125.7	174.3	170.7
Madhya Pradesh	107.7	114.1	160.7	160.9
Maharastra	140.6	141.8	160.9	163.3
Orissa	96.0	82.9	172.9	170.7
Punjab	149.5	145.6	173.4	176.7
Rajasthan	139.8	141.8	257.0	240.4
Tamil Nadu	134.1	132.1	185.8	175.1
Uttar Pradesh	128.9	133.6	183.3	177.6
West Bengal	117.8	103.0	152.4	136.4
All-India Rural	124.5	119.3	175.3	171.4

Source: My own calculations based on the NSS unit record data.

TABLE 28: Prices Over Time Urban Sector – Uniform Households

	1983 to 1987-88		1987-88 to 1993-94	
	AI	QAI	AI	QAI
Andhra Pradesh	123.2	125.1	156.9	158.6
Assam	98.1	78.5	136.4	148.9
Bihar	119.4	117.8	160.5	155.5
Gujrat	122.0	123.6	149.9	154.2
Haryana	198.7	192.7	118.0	125.2
Himachal Pradesh	138.2	136.6	151.9	153.1
Jammu and Kashmir	149.0	164.8	154.4	146.0
Karnataka	126.4	124.0	152.3	154.3
Kerala	127.4	123.4	125.9	125.3
Madhya Pradesh	126.0	122.9	141.0	144.6
Maharastra	140.4	141.1	141.0	147.9
Orissa	115.1	112.8	121.1	120.7
Punjab	126.0	127.4	156.7	163.6
Rajasthan	153.5	158.4	152.7	145.0
Tamil Nadu	113.0	111.9	153.5	156.3
Uttar Pradesh	144.2	143.1	168.1	172.8
West Bengal	129.3	129.8	146.6	147.3
Delhi	161.7	179.7	169.6	162.8
All-India Urban	130.4	130.3	150.2	152.3

Source: My own calculations based on the NSS unit record data.

TABLE 29: Poverty Measures – Uniform Households

	Rural		**Urban**	
Year	HCR	PGI	HCR	PGI
1983	40.35	14.16	39.92	11.87
1987-88	31.81	8.41	32.35	8.25
1993-94	27.52	6.44	17.68	3.96

Source: My own calculations based on the NSS unit record data.

B Extended Data Sample

In this section I extent the data sample to include all the five latest NSS household surveys. Thus, the sample covers the time period 1983 to 2004-05. Table 30 presents summary statistics for 1999-00 and 2004-05. As can be seen from that and the summary statistics for the first three survey rounds presented in Table 2, the average food share falls quite drastically between the survey rounds in 1993-94 and in 1999-00; roughly seven percent points in both the rural and the urban sector. The increase in total monthly expenditure is also substantial for both sectors, but still, we might suspect that there is something else that causes the reduction in food share.

Tabel 31 present the estimation with the full data sample, and Table 32 presents the changes in the cost of living over time for all-India, separately for the rural and the urban sector. From Table 32 we can see that the QAI-model suggests that the cost of living increased by marginally 5.2 percent for the rural sector and 8.1 percent for urban sector in the time period 1993-94 to 1999-00. In comparison, the IPC's poverty deflator grows by roughly 60 percent for both sectors in the same period. The fourth column in Table 32 presents the UV price indices from Deaton and Tarozzi (2000), Deaton (2003) and Deaton (2008). These indices measure a smaller price increase than the IPC in the period 1993-94 to 1999-00, but still, the prices increase with over 50 percent for both sectors. The disparity between the Engel estimates and the IPC's price measures are substantially smaller for the time period 1999-00 to 2004-05, but the Engel analysis still suggests very little increase in the cost of living.

Based on these findings I discuss potential explanations for the large disparity in the different estimates. First of all, the possibility that the cost of living in India did not grow by much in the late 1990s should not be totally excluded. It is well known that the Indian economy experienced liberalization in this time period, in terms of increased openness and trade with the rest of the world. This could have influenced the cost of living in a negative direction. That said, the implied poverty rates from the Engel analysis look less convincing than the price estimates themselves. These poverty measures are presented in Table 33. For the rural sector the poverty ratio falls from 24 percent of the total population in 1993-94 to only three percent of the total population in 1999-00. The poverty ratio in the urban sector falls from 17 percent of the total population in 1993-94 to two percent of the total population in 1999-00. These poverty rates seem unrealistically low, and it is therefore necessary to look for alternative explanations for why my estimates differ by this much.

First, I have not controlled for relative food and non-food prices. Thus, changes in relative prices could be one source behind the marginal increase in prices between 1993-94 and 1999-00. However, as a robustness check I control for relative prices for the urban sector[37]. The estimates presented in this section seems very robust to the inclusion of relative prices. Based on this finding, relative food and non-food prices are therefore not likely to be the main source.

A second potential explanation is the special design of the 55th NSS household survey. The problem associated with the household survey can briefly be described as follows[38]. After the 50th round in 1993-94, the NSS experimented with the reporting periods in the so-called in-between rounds 51 through 54. The experiment consisted of using a different reporting period for some randomly drawn household; 7 days for high-frequency items (food, pan, tobacco) and 365 days for low-frequency items (durable goods). For the rest of the households they used a uniform reporting period of 30 days as before. The conclusion from these experiments was that the experimental households systematically reported a higher expenditure level than the other households. In the large 55th survey round, the NSS chose to use both the 7 and 30 days reporting period for the high-frequency items, while for low-frequency items the reporting period were 365 days only. As it turned out, the stable ratio between expenditure levels from the 7 and the 30 days reporting period from the experimental rounds disappeared in the large 55th survey. Some authors

[37] I use the same strategy as described in more details in Appendix A.1. The results are not reported.

[38] A more detailed description of the 55th NSS survey round can be found in Deaton and Kozel (2005).

argue that this reflects households trying to answer consistently between the two different reporting periods, since they were asked for both. This probably boosted the 30 days expenditure level, and consequently reduced the official headcount ratios (Deaton and Kozel 2005). Various attempts to correct the inconsistency of the 55th survey round have been put through, all based on more or less controversial assumptions (see e.g., Deaton and Dréze 2002; Sen and Himanshu 2005; Sundaram and Tendulkar 2003).

A uniform increase in both food and total expenditures could not by itself explain the low increase in cost of living found in the Engel analysis. However, if the special sample design spuriously lead to a smaller budget share for food for a given real income level, households in 1999-00 would look richer than they really were. Hence, the identified price level would have to be low to justify the artificial high real income level (to fit my model specification). Whether this effect is present is difficult to assess, because it is not really known which consumption groups that were most affected by the change in reporting period. However, if the special design of the household survey is the main reason for the low increase in cost of living between 1993-94 and 1999-00, we would except to see a large increase in the cost of living in the final time period (1999-00 to 2004-05), to "make up" for the spuriously low increase in the preceding time period. Since the NSS survey in 2004-05 is consider being consistent with earlier surveys, the effect of the 55th survey should applied with the opposite sign. As can be seen from Table 32, this effect is not present in the data.

This lead me to a third potential explanation; that there could have taken place a structural break in the Engel relation. This would violate my main assumption in this thesis; namely that there exist a unique Engel relation over all situations being compared. Whether a structural break took place in the 1990s is discussed in the literature. Radhakrishna and Ravi (2004) argue that they found evidence for a shift in the Engel curve, which was driven by changes in preferences. Sen and Himanshu further pointed out that they believe it exists; *"strong corroborative external evidence that Engel shifts did occur from food to non-food"* (Sen and Himanshu 2005, 15). According to these writers, the budget share for food apparently plunged even within fixed income classes. If this is true, it will lead to spuriously high real income levels with my method, and consequently spuriously low price levels. It is out of the scope of this thesis to fully investigate the potential break in the Engel relation in the late nineties, but it should be an issue for future research.

TABLE 30: Summary Statistics – Extended Sample

	1999-00	2004-05
RURAL		
Total Expenditure	485.87 (1.438)	558.88 (2.565)
Food Share	0.61 (0.001)	0.60 (0.001)
No of Children	2.65 (0.014)	2.48 (0.012)
No of Adults	3.69 (0.014)	3.62 (0.012)
Age of Household Head	45.82 (0.074)	46.06 (0.073)
Ln Relative Prices	0.40 (0.001)	0.23 (0.001)
No of Households	71385	79284
URBAN		
Total Expenditure	854.69 (7.418)	1052.41 (9.069)
Food Share	0.53 (0.001)	0.50 (0.001)
No of Children	2.11 (0.015)	1.88 (0.017)
No of Adults	3.70 (0.017)	3.72 (0.021)
Age of Household Head	45.20 (0.100)	46.14 (0.117)
Ln Relative Prices	0.10 (0.001)	-0.06 (0.001)
No of Households	48924	45340

Note: Standard errors in parentheses. These are corrected for the sample design using a Taylor-linearized variance estimation. Ln Relative Prices are calculated in the same way as in Equation (14) and (17), i.e., food inflation divided by non-food inflation in the CPIAL and the CPIIW for the rural and the urban sector, respectively.

Source: My own calculations based on the NSS unit record data.

TABLE 31: Regression – Extended Sample

	AI		QAI	
	Rural	Urban	Rural	Urban
Ln Exp	-0.1129	-0.1405	0.1603	0.1009
	(0.00055)	(0.00057)	(0.00233)	(0.00330)
$(\text{Ln Exp})^2$	–	–	-0.0228	-0.0183
			(0.00019)	(0.00024)
No of Children	0.0173	0.0210	0.0172	0.0197
	(0.00014)	(0.00016)	(0.00015)	(0.00016)
No of Adults	0.0197	0.0258	0.0215	0.0256
	(0.00019)	(0.00020)	(0.00020)	(0.00018)
Age of Household Head	0.0001	0.0002	0.0002	0.0003
	(0.00002)	(0.00002)	(0.00001)	(0.00002)
Constant	1.2482	1.3952	0.4279	0.6073
	(0.00355)	(0.00426)	(0.00718)	(0.01090)
Observations	332954	198509	332954	198509
Adj. R^2	0.355	0.477	0.382	0.491

Note: Robust standard errors in parentheses. Dummy coefficients are not reported.

TABLE 32: Prices Over Time – Extended Sample

(a) Rural Sector

Year	AI	QAI	IPC	UV
1987-88	121.1	117.8	128.7	–
1993-94	166.8	165.3	178.7	169.8
1999-00	98.3	105.2	159.1	154.5
2004-05	100.6	102.0	108.8	114.2

(b) Urban Sector

Year	AI	QAI	IPC	UV
1987-88	129.4	127.9	140.2	–
1993-94	153.6	152.7	173.5	173.8
1999-00	103.5	108.1	161.4	157.7
2004-05	104.3	104.7	118.6	–

Source: My own calculations based on the NSS unit record data, the IPC's poverty lines and Deaton and Tarozzi (2000), Deaton (2003) and Deaton (2008).

TABLE 33: Headcount Ratios – Extended Sample

	Rural		Urban	
Year	QAI	IPC	QAI	IPC
1983	40.79	45.74	40.32	41.41
1987-88	31.22	38.94	31.38	38.90
1993-94	24.35	36.51	17.23	32.51
1999-00	3.01	26.32	2.29	23.93
2004-05	1.83	26.21	1.44	25.73

Source: My own calculations based on the NSS unit record data and IPC's poverty lines.

C Rural and Urban Households

In this section I pool data for the rural and the urban sector in the estimation. The advantage of this, is that it enables me to compute comparable prices between the two sectors. However, the main concern is that the data sets from the two sub-populations not are comparable and harmonized (as discussed in Section 5.6). Table 34 presents the estimated parameters from the pooled regression, and Tables 35 to 37 present the implied state-wise prices in the urban sector compared to the prices in the rural sector. The first two columns show the estimates from the AI-model and the QAI-model, respectively, while the third column shows the implied prices from the IPC's state-wise poverty lines. The fourth column in Table 36 and Table 37 presents the price differences implied by Deaton and Tarozzi's UV price indices.

The estimates from the Engel analysis suggest that the average all-India cost of living is lower in the urban sector than in the rural sector. This contrasts with both the IPC's and Deaton and Tarozzi's measures, as well as what is generally thought of the price differences between the two sectors. Based on this, we might suspect that there is problems associated with the data pooling. One hypothesis is that non-food items are relatively more important in urban areas, because of the greater variety of available consumption goods. If this is true, households in urban areas will devote a smaller share of their total budget on food items for a given real income level than households in rural areas. Thus, the main assumption for the identification of consumption prices in this thesis will be violated. More specifically, since it is expected that the budget share for food falls in real income, urban households will look richer than they actually are, and hence, the identified consumption prices will be biased downwards, compared to rural households, to ”justify” the spuriously high real income level in urban areas.

It is hard to evaluate the price estimates from the Engel analysis without further investigate how comparable the household data are for the two sectors. This should be carefully investigated in future work, especially because the price differential between the two sectors implied by the official poverty lines is consider to be too large (Deaton and Tarozzi 2000; Himanshu and Murgai 2009).

TABLE 34: Regression – Rural and Urban Pooled

	AI	QAI
Ln Exp	-0.1258	0.2119
	(0.00037)	(0.00261)
$(\text{Ln Exp})^2$	–	-0.0280
		(0.00021)
No of Children	0.0186	0.0172
	(0.00013)	(0.00013)
No of Adults	0.0224	0.0226
	(0.00015)	(0.00015)
Age of Household Head	0.0003	0.0004
	(0.00002)	(0.00002)
Constant	1.3088	0.2990
	(0.00263)	(0.00802)
Observations	323816	323816
Adj. R^2	0.343	0.377

Note: Robust standard errors in parentheses. Dummy coefficients are not reported.

TABLE 35: Urban relative to Rural 1983

States	AI	QAI	IPC
Andhra Pradesh	105.4	108.3	146.5
Assam	77.9	75.8	99.2
Bihar	97.2	84.5	114.7
Gujarat	93.9	92.8	147.9
Haryana	77.2	79.6	116.8
Himachal Pradesh	81.6	84.8	115.5
Jammu & Kashmir	81.2	81.7	108.6
Karnataka	104.1	102.3	144.3
Kerala	99.2	102.1	123.4
Madhya Pradesh	89.5	91.6	146.9
Maharashtra	130.4	136.5	143.3
Orissa	82.7	71.4	117.4
Punjab	96.1	94.4	114.1
Rajasthan	92.3	95.8	141.5
Tamil Nadu	100.3	94.0	125.1
Uttar Pradesh	98.2	100.4	131.5
West Bengal	65.2	58.1	100.3
All-India Urban	89.3	86.3	129.2

TABLE 36: Urban relative to Rural 1987-88

States	AI	QAI	IPC	UV
Andhra Pradesh	97.6	100.2	165.2	110.7
Assam	85.2	86.0	99.3	108.0
Bihar	91.6	88.7	124.8	108.1
Gujarat	89.7	87.2	150.6	105.4
Haryana	87.4	87.3	116.5	112.1
Himachal Pradesh	91.2	96.2	117.2	104.8
Jammu & Kashmir	113.7	113.7	119.3	103.8
Karnataka	94.7	96.1	163.9	110.0
Kerala	100.8	101.3	125.0	103.5
Madhya Pradesh	93.3	97.0	166.7	113.0
Maharashtra	126.4	132.6	163.6	114.1
Orissa	98.6	96.5	136.2	110.2
Punjab	80.7	82.4	118.0	113.2
Rajasthan	93.0	91.4	140.7	106.7
Tamil Nadu	88.0	86.4	140.3	109.0
Uttar Pradesh	103.1	103.7	134.5	118.1
West Bengal	65.8	63.1	116.1	112.7
All-India Urban	91.4	92.1	140.8	111.4

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines. The fourth column in Table 36 is taken from Table 4 in Deaton and Tarozzi (2000).

TABLE 37: Urban relative to Rural 1993-94

States	AI	QAI	IPC	UV
Andhra Pradesh	86.8	91.6	170.6	110.5
Assam	70.9	71.5	91.5	111.6
Bihar	97.2	90.2	112.4	112.5
Gujarat	83.4	85.5	147.1	105.2
Haryana	72.7	73.5	110.5	115.6
Himachal Pradesh	89.1	95.9	108.5	108.1
Jammu & Kashmir	95.5	99.2	–	107.0
Karnataka	97.6	100.3	162.3	110.6
Kerala	78.9	84.0	115.1	104.2
Madhya Pradesh	93.3	98.2	164.2	115.8
Maharashtra	123.3	128.2	168.5	118.2
Orissa	79.5	84.1	153.7	110.5
Punjab	72.3	78.5	108.5	114.2
Rajasthan	86.3	86.9	130.1	111.3
Tamil Nadu	80.0	82.6	150.9	109.7
Uttar Pradesh	94.6	97.8	121.4	116.5
West Bengal	74.6	76.4	112.1	117.5
All-India Urban	88.2	90.7	136.7	115.6

Source: My own calculations based on the NSS unit record data and the IPC's poverty lines. The fourth column in Table 36 is taken from Table 4 in Deaton and Tarozzi (2000).