Food demand, uncertainty and investments in human capital. Three essays on rural Andhra Pradesh, India.

Edoardo Masset

D.Phil. Development Studies

Institute of Development Studies
University of Sussex

Author’s declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature………………………………………………………………………………

Food demand, uncertainty and investments in human capital. Three essays on rural Andhra Pradesh, India.

Edoardo Masset
Institute of Development Studies, University of Sussex

Abstract

This dissertation provides some explanations of the causes of poverty in rural India, by investigating poverty determinants that are too often neglected in the literature and in policy debates. It proceeds in three main chapters, each addressing a specific research question.

The first chapter focuses on the process of agricultural transformation in the state of Andhra Pradesh. In the early stages of economic development, all countries undergo a process of transformation of their production and employment structure. As a result, agricultural output as a share of total GDP decreases, as does rural employment as a share of total employment. Over the last 50 years, the share of agriculture in total output has considerably declined in Andhra Pradesh. However, the agricultural sector continues to employ the great majority of the labour force. The theoretical section of this chapter shows how structural change is affected by the characteristics of food demand and by income inequality. The empirical analysis, using novel semiparametric methods, estimates food Engel curves and food elasticities, which are used to simulate the effects on changes in income distribution on the composition of demand.

The second chapter analyses the stabilising effect of irrigation on household expenditure. The expansion of irrigation infrastructure, together with the introduction of hybrid seeds and chemical fertilisers, was the most important technological advancement in Indian agriculture of the last 50 years. The positive impact of irrigation on income of rural households has been extensively documented, but its stabilising effect has been largely neglected. The first part of the chapter builds a theoretical model that establishes the causal links between access to irrigation, income stability, and consumption smoothing over the seasonal cycle. The empirical analysis assesses the stabilising impact of irrigation on expenditure using modern impact evaluation techniques. The findings indicate that consumption patterns of households with access to irrigation are more stable over the seasonal cycle and over the years.

The third chapter studies the effect of income uncertainty on educational choices made by the rural poor. It investigates the demand side of education in order to understand why a large number of rural children do not enrol or complete primary education. The theoretical part of the chapter presents an inter-temporal consumption model that shows how the expectation of income variability negatively affects household expenditure on education. The empirical analysis uses a duration model with time covariates in order to estimate the determinants of child progress in school, and provides evidence that income variability negatively affects investments in education.

Table of contents

Acknowledgments

I am indebted with several people who helped me to complete my doctoral work. First of all I would like to thank my supervisor, Dr. Howard White, for stimulating my analytical thinking and for providing constant support and encouragement.

This dissertation would not have been possible without the availability of two important datasets. I would like to thank Nalini Kumar, Peter Lanjouw, and Rinku Murgai of the World Bank for providing me the Andhra Pradesh expenditure data of the National Sample Survey Organisation, and the staff of the regional Office of the Director of Census Operations of Andhra Pradesh for providing me the data of the Census of Villages of Andhra Pradesh. I would also like to extend my thanks to the staff of the British Library of Development Studies for their help with searching other relevant data sources.

This dissertation has greatly benefited from discussions with a number of experts based at the University of Sussex and elsewhere. In particular, I would like to thank Sherman Robinson, Xavier Cirera, and Rachel Sabates-Wheeler of the Institute of Development Studies for helpful comments on earlier drafts and for many useful discussions. I would also like to thank Dr. D.V. Ramana of the Centre for Developmental Research in Hyderabad, Dr. M.J. Mohan Rao of the National Institute of Rural Development in Hyderabad, and K. Vijay (private consultant) of Mahaboobnagar for their invaluable comments and suggestions. Finally, I would like to thank all those ordinary people in Andhra Pradesh who unintentionally with their observations opened my eyes on a number of fundamental issues.

My warmest and most grateful thanks go to my wife Barbara for her illuminating suggestions and extraordinary support during these three years of intense work.

List of Tables

List of Figures

Acronyms

APEP	Andhra Pradesh Primary Education Project
BPL	Below the Poverty Line
CARA	Constant Relative Risk Aversion
CPIAL	Consumer Price Index for Agricultural Labourers
CPIIW	Consumer Price Index for Industrial Workers
CRRA	Constant Relative Risk Aversion
DES	Directorate of Economics and Statistics
DfID	Department for International Development
DHS	Demographic and Health Surveys
DPEP	District Primary Education Programme
EGS	Employment Guarantee Scheme
GDP	Gross Domestic Product
ILO	International Labour Organisation
LES	Linear Expenditure System
LOWESS	Locally Weighted Scatterplot Smoothing
MEASURE	Monitoring and Evaluation to Assess and Use Results
MOL	Ministry of Labour
MSP	Minimum Support Price
NCERT	National Council of Educational Research and Training
NCO	National Classification of Occupations
NFHS	National Family and Health Survey
NSSO	National Sample Survey Organisation
OBB	Operation Black Board
ODA	Overseas Development Administration
OLS	Ordinary Least Square
PDS	Public Distribution System
UNESCO	United Nations Educational Scientific and Cultural Organisation

1 Introduction

1.1 Purpose of the study

Poverty in India has been the focus of a heated debate in recent times (Deaton and Kozel, 2005). Official poverty statistics are published by the Planning Commission of the government of India based on expenditure data collected by the National Sample Survey Organisation (NSSO). A number of technical controversies regarding the data, the indicators, and the methodologies used to measure poverty have resulted in estimates that diverge from those presented in official statistics. Bhalla (2002) using consumption data from the National Accounts, rather than from the NSSO, finds that poverty in India is much less than what is reported by the Indian government. Ray and Lancaster (2005), using a poverty line based on caloric requirements, rather than on the purchase of a basic basket of goods like the NSSO, find that poverty in India is much worse than what is revealed in official statistics. Deaton (2008), after adjusting the expenditure data and the official poverty line with price indices different from the official Consumer Price Index for Agricultural Labourers (CPIAL) used by the NSSO, finds that poverty has not decreased by the amount declared by the Indian government.

In spite of these controversies a general consensus has emerged that poverty in India has decreased over the last two decades, but that the rate of decrease is substantially lower than the one depicted by official statistics (Ravallion, 2008). The decrease in poverty is particularly low once we consider the sustained rates of growth in the Gross Domestic Product (GDP) witnessed by India in recent times. Not all Indians have shared in the benefits of this process of overall economic growth. Ravallion (2004) shows that though absolute poverty in India fell since the early 1990s, the rate of pro-poor growth was appreciably lower than the rate of economic growth, resulting in a distributional shift against the poor. Banerjee and Picketty (2005) considering a much longer period of time (from 1922 to 2000) show that since the mid 1980s much of the economic growth reported at the country level was captured by a top income group representing less than 1% of total population, a group whose income is often not recorded in official statistics.

There is also considerable evidence that poverty in India has not decreased in some geographic areas and among some disadvantaged social groups. First, a large number of studies have found that the patterns of economic growth among Indian states have diverged over the last two decades, and that the divide between forward and backward states has increased (see for example Kurian, 2000, Rao et al., 1999, Sachs et al., 2002). Second, a significant increase in inequality between rural and urban areas has been documented (see for example Deaton and Dreze, 2002, Jha, 2000). Finally, the most vulnerable social groups, the scheduled castes and the scheduled tribes in particular, have not substantially improved their living conditions over the last decades (see for example Gang et al., 2008, Sundaram and Tendulkar, 2003).

The uneven process of economic growth and the persistence of poverty in some Indian states, in rural areas, and among disadvantaged groups, suggest that not all Indian citizens have the same opportunities for escaping from poverty. This is also supported by qualitative evidence. For example, Krishna et al. (2004) using qualitative indicators of poverty in 36 villages of rural Andhra Pradesh found that nearly 80% of poor households reported as poor in 1978-79 were still poor in 2003-04. This body of evidence on the persistence of poverty in India suggests that there are some factors, largely out of the control of single individuals that prevent the poor from escaping poverty. There is a large literature in developing economics that provides an explanation to these phenomena by positing the existence of poverty traps (see for example Azariadis and Stachurski, 2005, Bowles et al., 2006b, Carter and Barrett, 2006). Poverty traps are self-reinforcing mechanisms that cause poverty to persist. The identification of these mechanisms casts doubts on the ability of markets to secure economic progress for all, and provides a possible explanation for the observed large income inequalities between countries and individuals.

This dissertation draws on economic models of poverty traps in order to explain the persistence of poverty among rural households in India. The three chapters comprising this dissertation investigate the economic mechanisms determining the persistence of poverty at both the macro and microeconomic level. The three sections that follow this introduction set the context of the study and briefly illustrate the data and the methodology used. The dissertation then proceeds in three main chapters. Each chapter addresses a specific research question and can be read separately from the others as a free-standing essay. The concluding chapter discusses the findings of the three main chapters and the explanations they offer of the persistence of poverty in rural India in light of the economic theory of poverty traps.

1.2 Setting the context: the state of Andhra Pradesh

The present study focuses on the Indian state of Andhra Pradesh rather than on India for three main reasons. First, a large amount of micro and macro data was available for this state. Empirical research occupies a central part of each chapter, and the availability of data was a decisive factor. Second, in order to answer the questions asked in each chapter, the data required a series of adjustment that could hardly be made at the country level. These adjustments include, for example, the correction of expenditure data for regional variation in prices, or the combination of household consumption data with wage and rainfall data obtained from other sources at the regional level. Finally, over the last four years I have had the opportunity to conduct research in many districts of Andhra Pradesh, and this has deepened my understanding of the most urgent development issues and of the policy options available.

The state of Andhra Pradesh was officially founded in 1956 by uniting nine districts of the Telangana region, located in the north of the state and surrounding the capital Hyderabad, to other 11 districts that were formerly part of the Madras presidency, located south of Hyderabad and on the coast of modern Andhra Pradesh. These 20 districts shared a common language (*Telugu*), but had different histories. The Telangana region had been independently administered by a *Nizam* (governor) since the occupation by the Moghul Empire in the seventeenth century, while the 11 districts of the Coastal and Rayalseema regions had been under British rule since the eighteenth century. In the late 1950s, soon after the foundation of the state, three additional districts were created from the existing 20 and, as a result, modern Andhra Pradesh is composed of 23 administrative districts.

Researchers and public officials often group contiguous and homogeneous areas of the state into regions. The most common grouping is the subdivision in administrative districts described above, which will be widely used in the present study. Very common, and frequently used in this study, is also the grouping into three historical macro regions: Coastal Andhra, Rayalseema, and Telangana, whose main characteristics can be summarised in the following way. Coastal Andhra is the richest region of the state in terms of per capita income and other development indicators. It is served by good infrastructure and benefits from good soils and rainfall. Telangana, with the exception of the state capital Hyderabad, is poor and infrastructure is inadequate, though part of the region has good soils and rainfall. Rayalseema is very dry and with poor prospects of agricultural growth. It is the poorest area of the state in terms of income and other development indicators.

Andhra Pradesh is the fifth largest state in India with a population of 76 million according to the census of 2001. The state GDP has grown at an average rate of 5% per year between 1960 and 2005 in absolute terms, and of 3% in per capita terms. These growth rates are similar to those experienced by India as a whole, and in terms of per capita income, Andhra Pradesh normally ranks in the richer half of Indian states. However, a large fraction of the population lives in poverty, particularly in rural areas.

According to official data published by the Planning Commission (Government of India, 2007), poverty in Andhra Pradesh is lower than in the rest of India and has rapidly declined over the last 30 years. Table 1.1 shows official poverty rates of Andhra Pradesh and India over the last 30 years based on data collected by the NSSO. Poverty rates in Andhra Pradesh are not only lower than those of India, but are much lower in rural areas as compared to urban areas, which is at odds with common perceptions and empirical evidence elsewhere in the country. Official poverty headcounts however are incorrect for two reasons. First, they are based on caloric poverty lines set in 1973-74 and it is believed that the cost of these lines was set too low in Andhra Pradesh, particularly in rural areas (Subrahmanyam, 2003). Since the poverty lines are updated using the consumer price index, the initial error is carried over at each new survey round. Second, the consumer price indices used to adjust the poverty lines are based on household consumption shares of 1973-4. As a larger share of expenditure was spent on food in 1973-74 and food prices have decreased since then, the result is an underestimation of poverty (Deaton, 2008).

When poverty lines are adjusted and alternative price indices are used, poverty rates turn out to be very different. For example Deaton (2003) reports poverty headcounts for Andhra Pradesh that are nearly twice those of official statistics, and in which the relation between rural and urban poverty is reversed. Rural poverty appears to be at least twice the size of urban poverty. In addition, if a caloric poverty line is calculated at each round, either in terms of number of calories or in terms of the cost of the same calories, the number of poor increases dramatically, and poverty in Andhra Pradesh appears to have increased over the last 30 years (see Palmer-Jones and Sen, 2001, Patnaik, 2007, Ray and Lancaster, 2005). It should also be mentioned that there are doubts regarding the reliability of the survey data used in the calculation of official poverty lines. Bhalla (2002) finds that household consumption reported by the survey method was only 55% of consumption reported by National Accounts in 1999-00 and that the poverty count based on the latter data was only 13%.

Table 1.1 Official poverty headcount in Andhra Pradesh and India

	1973-74 (%)	1977-78 (%)	1983-84 (%)	1987-88 (%)	1993-94 (%)	1999-00 (%)	2004-05 (%)
Andhra Pradesh							
rural	48.4	38.1	26.8	21.0	15.9	11.0	10.8
urban	52.6	46.5	41.2	41.1	38.8	26.6	27.1
India							
rural	56.4	53.1	45.6	39.1	37.3	27.1	28.3
urban	49.0	45.2	40.8	38.2	32.4	23.6	25.7

Source: Government of India(2007), Reddy et al. (2003b), and Himanshu (2007).

In a series of papers, Dubey and Palmer-Jones (2005a, 2005b, 2005c) question the use of official poverty lines for the assessment of living standards, including their adjusted versions, and recommend the use of revised or alternative welfare indicators. They point to major flaws of Deaton’s price adjustments based on the calculation of unit values and on the anchoring of poverty lines to a specific poverty line (2005c). They provide their own estimates of poverty headcounts and find that in Andhra Pradesh, contrary to what is reported by official statistics, poverty is much higher in rural areas than in urban areas, though poverty in urban areas turns out to be slightly higher than the estimates based on Deaton adjustments (2005a). Finally, they observe that the neglect of environmental variables, such as the quality of education and health services, police, water, and justice, produce poverty lines that are not truly representative of standards of living (2005b). They also suggest that this might explain the observed lack of correlation between poverty measures and other indicators of well-being. The results of their analysis call for substantial revisions of the poverty indicators used, and for the adoption of multiple indicators of well-being in conducting welfare analysis.

Official poverty figures are not matched by an analogous improvement in other human development indicators. Malnutrition rates are very high, despite improvements in the last 15 years, though they are lower than in the rest of India. Table 1.2 reports the percentage of stunted and underweight children under five calculated by the National Family and Health Survey (NFHS). Stunting is an indicator of long-term food deprivation, while wasting is an indicator of short-term deprivation. A child is stunted when height-for-age is more than two standard deviations away from the reference norm, and is underweight when weight-for-age is below two standard deviations from the norm. More than a third of children in Andhra Pradesh were reported stunted or underweight, a level of malnutrition which is very high not only in absolute terms, but also in comparison to other countries. For example the percentage of underweight children in sub-Saharan Africa in 2005 was 29.6%. It has also been observed that the reduction in malnutrition rates is much less than might be expected given the sustained growth in per capita incomes over the period. This can only be attributed to a decline in the consumption of calories per capita, a phenomenon that is not well understood and for which a convincing explanation has yet to be put forward (Deaton and Dreze, 2008).

Table 1.2 Stunting and wasting in Andhra Pradesh and India

	1992-93 (%)	1998-99 (%)	2005-06 (%)
Andhra Pradesh			
stunting	n/a	38.6	33.8
wasting	49.1	37.6	36.4
India			
stunting	47.1	45.5	38.4
wasting	51.9	47.0	45.9

Source: MEASURE DHS online STATcompiler.

Child mortality has declined considerably over the last 15 years, but little progress has been made with respect to infant mortality, which is now higher than in the rest of India. Table 1.3 shows mortality rates of infants and children in Andhra Pradesh and India reported by the NFHS. Note that these rates are very high both in absolute and relative terms. Countries outside Asia with similar infant mortality rates are Zimbabwe (60% in 2005), Tanzania (68% in 2004) and Liberia (71% in 2007). Interestingly, high infant mortality in Andhra Pradesh is largely driven by high neonatal mortality, while postnatal mortality follows the same decreasing pattern shown by child mortality. Child and postnatal mortality reflect exogenous causes of death and are normally correlated with factors like income per capita, education and vaccination. Neonatal mortality reflects endogenous causes of death and is correlated with factors like health care and mother's status at the time of delivery. This suggests that in Andhra Pradesh ante-natal care and mothers' health status at delivery are particularly poor.

Table 1.3 Infant and child mortality in Andhra Pradesh and India

	1992-93 (‰)	1998-99 (‰)	2005-06 (‰)
Andhra Pradesh			
infant mortality	73.1	70.8	68.4
child mortality	24.6	22.0	11.1
India			
infant mortality	86.3	73.0	65.0
child mortality	35.5	30.5	21.8

Source: MEASURE DHS online STATcompiler.

Illiteracy rates are much higher in Andhra Pradesh than in the rest of India. Indeed Andhra Pradesh has the fifth highest illiteracy rate in the country. Table 1.4 show the data on education levels in Andhra Pradesh and India over the last 15 years calculated by the NFHS. Despite the significant progress made, large differences in educational achievements still exist between urban and rural areas, boys and girls, social groups, and developed and underdeveloped regions (Reddy and Rao, 2003).

Table 1.4 Educational attainments in Andhra Pradesh and India

	1992-93 (%)		1998-99 (%)		2005-06 (%)	
	male	female	male	female	male	female
Andhra Pradesh						
Illiterate	37.0	58.8	29.4	50.2	27.0	45.1
Primary	29.3	22.7	31.2	27.2	28.0	23.7
Secondary	25.9	14.7	27.9	18.2	36.9	27.3
Higher	6.4	2.4	11.5	4.4	7.9	3.7
India						
Illiterate	29.2	54.7	21.6	44.4	21.9	41.5
Primary	36.3	26.9	33.0	28.0	28.7	25.1
Secondary	25.5	13.4	32.1	20.8	40.9	28.5
Higher	6.1	2.6	13.3	6.8	8.3	4.7

Source: MEASURE DHS online STATcompiler.

1.3 Data used

This study makes intensive use of data both at the micro and the aggregate level. Each chapter contains an empirical section, where the theoretical hypotheses formulated are tested using econometric methods. This section briefly discusses the data and more detailed information on the data and the adjustments made for the empirical analysis is provided in each chapter.

The expenditure survey data collected by the NSSO are the most important data for this study. NSSO surveys are a legacy of the survey experiments initiated by the statistician P.C. Mahalanobis in the late 1930s. Mahalanobis pioneered the now popular method of collecting consumption data from random samples of households. The Planning Commission of India rapidly adopted the survey data collection method in order to monitor living standards and deal with distributional issues. Every year since 1944, the NSSO has been interviewing a very large number of household throughout India, collecting detailed information on consumption of food and non-food items (the 'thin' rounds). In addition, every five years, larger surveys are conducted in an exercise that sees the participation of more than 100,000 Indian households (the 'thick' rounds). This dissertation uses the last four thick rounds conducted in 1983-83 (38^{th} round), 1987-88 (43^{rd} round), 1993-94 (50^{th} round), 1999-00 (55^{th} round) and 2004-05 (61^{st} round). The data collected during the 55^{th} survey round have generated a heated debate, in which their reliability has been questioned due to changes made to the questionnaire. The availability of the latest large survey round however, allows the analysis of variables that are comparable across surveys and that span a period of nearly 25 years.

Second in importance are the census data. The Ministry of Home Affairs has conducted a population census 14 times every ten years since 1881. Scope and coverage of the census have changed over time and the information collected now includes, alongside demographic characteristics of every Indian household, data on village facilities, education, fertility and migration. This dissertation uses data from the Census of Villages of Andhra Pradesh, which aggregates population census data at the village level, for the years 1961, 1971, 1981, 1991 and 2001.

Another important source of data for this study is the Directorate of Economics and Statistics (DES). The DES is part of the Department of Agriculture and Cooperation and operates under the Ministry of Agriculture. The DES collects a vast amount of data at the state level on a wide range of issues, from agricultural added value, to crop yields, land use and agricultural prices. This study makes extensive use of two DES publications. The first is *Golden Jubilee of Andhra Pradesh: 1956-2005- 50 years* (DES, 2005), a collection of state and district level data on a large number of socioeconomic variables over the last 50 years. The second publication is *Agricultural Wages in India* (DES, 1975-2005), which contains data on wages of skilled and unskilled labour by month and by district for every Indian state since 1955.

Finally, the website *indiastat*, inaugurated by the Government of India in 2000, has been an important and easily accessible source for some state and district level data, such as district rainfall, price indices, crop yields and land use.

1.4 Structure of the study

This dissertation comprises three main chapters, that follow a similar structure, and a final concluding chapter. In each chapter, after a short introduction to the topic, a review of the literature is presented. A theoretical model establishing causal relationships is formulated algebraically, based on a set of reasonable hypotheses and assumptions. The data used to empirically test the hypotheses made are presented and descriptive statistics are shown. The econometric methods used to estimate the hypothesised causal relationships are illustrated and results are presented of the estimations and tests performed. A final section concludes.

Chapter 2, *Food demand and the agricultural transformation in Andhra Pradesh*, provides a theoretical and empirical contribution to models of structural change. Historical accounts of the agricultural transformation in the economic literature point to two main drivers of structural change: technological progress in agriculture and inelastic food demand. Economic models of structural change however, neglect the importance of demand factors in shaping the process of agricultural transformation. Food demand depends on both the current level of living standards and on the distribution of income. The extent to which structural change is affected by the assumptions made regarding consumers' preferences and existing income inequality is analysed. The effects of changes in income distribution on food consumption are assessed by estimating the elasticity of food demand using novel semiparametric methods.

The results of the empirical analysis suggest that the large share of agricultural employment observed in Andhra Pradesh can be partially explained by the poor operation of the Engel's law of demand. Living standards are very low in Andhra Pradesh, and the vast majority of households are not able to consume the minimum recommended number of calories per day. A large proportion of any increase in per capita income is therefore spent on food. In addition, as the distribution of income becomes more equal, food demand further increases, because the poor, who spend proportionally more on food, have more resources. The high level of food demand helps sustaining food prices and labour demand in agriculture, thus impairing the prospects of industrialisation and wider economic growth.

Chapter 2 proceeds in seven sections. After a brief introduction in Section 2.1, Section 2.2 reviews the literature on the agricultural transformation. Section 2.3 builds a simple general equilibrium model showing that the size of employment in agriculture is determined in the long term by technological progress in agriculture, while the pace of structural change in the employment of the labour force depends on the characteristics of food demand and on the existing income distribution. Section 2.4 presents some stylised facts on agricultural output and employment, technological progress, food demand and income inequality in Andhra Pradesh. Section 2.5 illustrates the advantages of semiparametric methods for estimating food demand. Section 2.6 estimates food expenditure elasticities in order to assess the effect of changes in income and income distribution on the pattern of structural change of the economy. Section 2.7 concludes.

Chapter 3, *The stabilising effect of irrigation on households' expenditure*, analyses the stabilising effect of irrigation on household expenditure. Research on the benefits of irrigation has primarily focused on its ability to increase farmers' and agricultural labourers' incomes, while disregarding its effect on income and consumption variability. In this chapter the hypothesis is made that irrigation reduces income uncertainty and the need for precautionary savings as an insurance mechanism. The present study is probably the first attempt to assess the stabilising impact of irrigation on household expenditure and saving decisions. A theoretical model establishing the causal links between access to water, income stability, and consumption smoothing over the seasonal cycle is constructed, and modern impact evaluation techniques are employed to measure the stabilising effect of irrigation.

The empirical analysis finds that households without irrigation have more unstable consumption patterns, that their consumption tracks income more closely, and that they engage in precautionary savings in order to smooth income fluctuations. The combination of high production risk and liquidity constraints is the source of three negative welfare effects. First, high consumption instability is a welfare loss in itself. Second, higher income instability implies higher precautionary savings, which are an inefficient form of savings. Finally, vulnerability to production shocks poses the risk of falling into chronic poverty for extremely poor households.

Chapter 3 proceeds in eight sections. Section 3.1 is the introduction. Section 3.2 reviews the literature on poverty and seasonality. Section 3.3 reviews the literature on consumption smoothing and on the impact of irrigation on rural income. Section 3.4 builds an intertemporal consumption model over the seasonal cycle that incorporates uncertainty of future income. Section 3.5 describes in detail the data used in the analysis and the adjustments made to expenditure figures. Section 3.6 presents some descriptive statistics on irrigation in the state, rainfall patterns and characteristics of farmer and agricultural labour households with and without irrigation. Section 3.7 describes the Roy model used to detect the difference in consumption patterns of irrigated and non-irrigated households. Section 3.8 presents the findings on the impact of irrigation and output uncertainty on household expenditure. Section 3.9 concludes.

Chapter 4, *Income uncertainty and investments in human capital*, studies the effect of production risk and uncertainty on educational choices made by the rural poor. In rural India, investment opportunities are limited and for the majority of household the most important and profitable investment is their children's education. The state of education in rural India is however very poor and the main causes for this on the supply side are well known. Chapter 4 investigates the demand side of education in order to understand why a large fraction of rural children do not enrol or do not complete primary education. The analysis conducted is innovative in two ways. First, this is the first study of the effect of income uncertainty on educational choices in developing countries, as the economics of education literature has restricted the study of uncertainty to the effect of the variability of future earnings on schooling decisions. Second, the empirical analysis employs new econometric techniques by estimating child progress in school using a duration model with time covariates.

The empirical analysis finds that income uncertainty negatively affects schooling decisions of rural households and expenditure on education of farmers and agricultural labourers. The empirical analysis also finds that education expenditure by households with non-irrigated farms and agricultural labourers living in non-irrigated villages is not affected by income uncertainty. These results point to production risk and uncertainty as factors that contribute to determine the low levels of investments in human capital by Indian rural households.

Chapter 4 proceeds in eight sections. Section 4.1 introduces the topic. Section 4.2 reviews the literature on education in Andhra Pradesh and the effect of uncertainty on household schooling decisions in developing countries. Section 4.3 constructs an intertemporal utility maximisation model of investments in education, which incorporates a precautionary motive for saving. Section 4.4 describes the data and the variables used in the empirical analysis. Section 4.5 discusses the econometric methods employed for the estimation of a school attainment model and an education expenditure model. Section 4.6 presents descriptive statistics on education in Andhra Pradesh across regions, gender and social classes using birth cohort analysis. Section 4.7 presents the findings on the effect of uncertainty on investments in human capital. Section 4.8 concludes.

Chapter 5, *The conclusions*, revisits the empirical findings of the three preceding chapters in light of the theory of poverty traps and provides some explanations for the persistence of poverty in rural Andhra Pradesh. Firstly, the theoretical elements of poverty trap models are discussed. Secondly, it is discussed how the findings of each chapter constitute elements of poverty traps. Finally, a number of directions for future research are described.

Three main poverty traps are identified: an under-consumption trap; an irrigation and liquidity trap; and an educational trap. At the macroeonomic level it is suggested that the inability of Andhra Pradesh to industrialise is a result of the extreme poverty of its rural population, which restricts the market for manufacturing products and prevents the exploitation of economies of scale. At the microeconomic level it is maintained that production risk, combined with a failure of credit and insurance markets, prevents households from escaping poverty. High production risk forces poor household to save for precautionary reasons and to underinvest in production activities and educational choices that offer higher rates of return.

2 Food demand and the agricultural transformation in Andhra Pradesh

2.1 Introduction

Andhra Pradesh is sometimes called ‘the rice bowl of India’, and is commonly referred to as an agricultural state. Over the last 50 years however, agricultural output as a share of GDP has decreased from more than 60% per cent to less than 30%, and in the late 1980s the service sector has become the leading sector of the economy. The relative decline of agriculture in total GDP is the combined result of poor growth of the agricultural sector and of a more sustained growth in manufacturing and services. The bulk of labour employment however, continues to be concentrated in agriculture. Today, nearly 70% of the total workforce is employed by the agricultural sector and the share has not significantly decreased over time. Moreover, while the absolute number of farmers has remained fairly stable over the last 50 years, the number of agricultural labourers has increased, and they now represent more than 40% of the total workforce.

Official poverty statistics report low and declining poverty rates in rural Andhra Pradesh over the last 25 years. Official poverty headcounts however are plagued by technical problems that raise doubt about the reliability of poverty estimates. Calculations made by independent researchers show much higher poverty rates, particularly in rural areas, and less progress over time. Furthermore, living standards of rural households in Andhra Pradesh, as measured by various socioeconomic indicators like literacy rates and infant mortality, are low not only by international standards, but also in comparison to other Indian states. The prospects of improving living conditions of rural households largely rely on changing the structural composition of the labour force. However, the development of the non-farm sector in rural areas is modest, while migration to urban areas is negligible and labour employed in agriculture is either stable or growing.

This chapter provides an account of the process of agricultural transformation in Andhra Pradesh. Several factors may explain the evolution of the sectoral composition of the labour force, but two principal forces shaping the agricultural transformation process are technological progress in agriculture and inelastic demand for food products. The present analysis focuses on the latter, and assesses the extent to which inelastic food demand can explain the pattern of structural change in the economy. It argues that in the early stages of development, when living standard are very low and household income is spent mainly on food, the process of structural change is slow unless there is a large increase in income inequality.

This chapter proceeds in seven sections. Section 2.2 reviews the literature on the characteristics and the determinants of the process of agricultural transformation; the literature on the relationship between inequality and economic growth; and the literature on the impact of under-consumption on economic stagnation. Section 2.3 develops a simple general equilibrium model of the economy and of the effects that different assumptions regarding food demand have on agricultural employment. Section 2.4 presents descriptive statistics on the process of structural change in Andhra Pradesh. Section 2.5 introduces the econometric methods employed in the empirical estimation of food demand. Section 2.6 presents the results of the estimation of parametric and semiparametric food Engel curves. Section 2.7 concludes.

2.2 Literature review

2.2.1 The agricultural transformation

In the early stages of economic growth, countries undergo a process of sectoral transformation of their production structure both in terms of output and employment. All countries begin the development process with a production and employment base that is predominantly rural, and become rich by expanding the modern sector. Traditionally, at least in the experience of European countries, the modern sector was the manufacturing sector. However, in more recent times, some developing countries have modernised through the expansion of the service sector without having to undergo an intensive process of industrialisation.

The process of transition of the economy from an agricultural to a modern state is known as agricultural transformation, and is one of the most documented processes of growth economics. At the time of Kuznets' writings, the cross-country negative relationship between the size of the agricultural sector and per capita income, and the positive relationship between the size of the manufacturing sector and per capita income were already well known. Kuznets (1957) collected a large amount of evidence in support of this observation, and also documented the simultaneous decline over time of the labour force employed in agriculture, the large increase in the share of the labour force employed by the service sector, and the modest increase of employment in manufacturing. Other surveys of the sectoral development process conducted in more recent times have confirmed the validity of the patterns described by Kuznets. The surveys by Chenery and Syrquin (1975) and Mundlak et al. (1997) are particularly extensive both in terms of countries and of time periods covered.

Given that the agricultural transformation has been almost universally documented across different countries and time periods, it seems obvious that explanations of economic growth should also explain the process of agricultural transformation. However, modern economic growth theory, which begun with the work of Solow (1956), has focused entirely on aggregate growth, neglecting the sectoral composition of the economy and the transformation the economy undergoes in the course of development. Perhaps, as suggested by Temple (2005), this neglect results from the need for an eclectic approach in studying the sectoral transformation. Or perhaps sectoral models are not popular because they deal with issues that are considered intractable, like increasing returns to scale.

Economists have attempted to explain the process of agricultural transformation by using dual economic models. In a dual model, the economy is divided in two sectors. One is poor, rural, and traditional, whilst the other is rich, modern, and industrialised. The two sectors are characterised by different production processes and growth dynamics. Dual models analyse how the two sectors interact in the course of the growth process. The formulation of dual economic models dates back at least to Lewis (1954), and Ricardo’s interpretation of the industrial revolution in England (see Hayami and Godo 2005) can be considered a very early example of this type of approach. Recent surveys of duality in economics can be found in Ranis (1988), Bardhan and Udry (1999), and Temple (2005).

The main characteristics of dual economic models are described by Dixit (1973). First, the agricultural and the modern sector are characterised by different production processes. While agricultural production occurs under constant or decreasing returns to scale, manufacturing production occurs under increasing returns to scale. Second, the agricultural sector produces mostly consumption goods, while investment goods are produced by the manufacturing sector. Third, consumer demand for agricultural goods is inelastic with respect to income, while there are no restrictions on the income elasticity of demand of manufactured goods.

Two other elements could be added to this list. First, the characteristics of the organisation of production are very different in the two sectors. While production in manufacturing tends to concentrate in large production units, in agriculture, even in modern times and countries, the family farm is the unit of production. This has implications for the market structure of the two sectors, particularly with respect to the determination of prices. Second, contrary to common belief, most of the world poor live in rural areas, and the processes of migration and urbanisation are strongly associated with poverty reduction (Ravallion et al., 2007). There is therefore a strong welfare argument in favour of studying the interaction of agriculture and the rest of the economy in the course of the development process.

Historical and economic interpretations of the process of agricultural transformation focus on two main elements: technological progress in agriculture, and the Engel's law of demand. See for example the classical treatment of this topic by Hayami and Ruttan (1971), Timmer (1988), and Mundlak (2005). Technological progress in agriculture has often been considered the trigger of the industrial revolution in England and Europe. A classical example is the introduction in England of the Norfolk system which started the agricultural revolution (Hayami and Godo, 2005). This was a new system of crop rotation, whereby fallow land was planted with forage crops that would increase livestock production, and ultimately crop production via the increase of manure production. The application of science to agriculture in modern times has led to much greater achievements in agricultural production. The discovery of hybrid varieties in the 1930s led to extraordinary increases in yields of wheat and corn in particular. Some of this technology was subsequently passed on to some developing countries in the late 1960s leading to what is now known as the green revolution.

Whatever the source and the characteristics of technological advancement in agriculture, the result has always been an increase in productivity of either labour or land or both. Productivity gains in agriculture in the post second world-war period were particularly impressive. Mundlak et al. (1997) find that in 80% of the 130 countries studied over the period 1960-1990, labour productivity grew faster in agriculture than in the rest of the economy. Overall, the median difference of productivity growth rates between agriculture and non-agriculture was above 1%. Similar data disaggregated by geographic areas and with reference to older periods can be found in Timmer (1988).

Technological progress in agriculture generates an output surplus, and this is where the Engel's law of demand comes into play. There is an obvious physical limit to human consumption of food. Even if populations are growing in body size and consuming increasing quantities of food, there is still a clear limit to the amount of food that can be consumed. As a result, as income increases food consumption increases less than proportionally. This phenomenon is commonly known as Engel's law. In principle, food consumption could further increase if new products, for example new food varieties or better qualities of existing varieties, were introduced. Indeed new food products are constantly introduced, but this does not shift household budgets towards the consumption of more food because even more new non-food products are constantly introduced.

The production of an agricultural surplus has in turn obvious effects on labour demand and prices. In the course of economic development, both employment in agriculture and agricultural prices decrease. Mundlak et al. (1997) in a sample of 130 countries over the period 1950-1990 observed an average annual decline in the agricultural labour force of about 2%. As for agricultural prices, Binswanger et al. (1987) showed that real international agricultural prices declined between 0.5 and 0.7% per year over the period from 1900 to 1984. In the short term, the labour surplus thus generated may or may not find employment in the modern sector, but in the long term the migration of labour to more productive non-agricultural enterprises appears inevitable. The description of this process can be enriched by adding more details, like the role played by governments in sustaining agriculture, the role of education in helping the allocation of labour in non-agricultural activities, the role of endogenous versus exogenous technological change, and so on. However, technological progress and Engel's law alone explain much of the process that Mundlak (2000) calls the curse of agriculture; through its own development, agriculture becomes unimportant.

While the effects of technological progress on economic growth have been extensively studied in models of economic growth, the effects of Engel's law have been much less investigated. The central role of food demand in dual models of economic development has been highlighted by Dixit (1973). The Engel's law operates differently depending on the level of living standards in the economy and on the existing distribution of income. Economists building dual economic models often make crucial assumptions about preferences and consumers' demand neglecting the role that these have in driving the results of their models.

This chapter describes the effects of different characteristics of food demand on the development of the agricultural sector. This is accomplished by first building a dual model of the process of agricultural transformation in Andhra Pradesh. The model solution shows how the characteristics of consumers' preferences and of food demand shape the pace of the transformation process. [1] The change in employment in rural areas following technological change in agriculture critically depends on the income elasticity of food demand in the economy. The poorer the population and the larger the income elasticity o food demand, the lower is the impact of technological progress on agricultural employment and overall economic growth. The chapter then proceeds to estimate the specific form assumed by food demand in Andhra Pradesh in order shed further light on the characteristics of the development process in the state. The estimation of food demand elasticities offers the opportunity to simulate the impact of income inequality on economic growth. It is found that, within the neoclassical model of agricultural transformation, changes in the income distribution affect the pace of economic growth. The characteristics of this process and the links to the theory of under-consumption are discussed in the following sections.

2.2.2 Inequality, food demand and economic growth

The study of the relationship between inequality and economic growth was pioneered by the path-breaking work of Kutznets (1955). Kuznets analysed the correlation between inequality and income across a number of countries and formulated what became known as the inverted-U hypothesis. According to this hypothesis – in fact the observation of an empirical regularity - in the early stages of economic development income inequality increases, but at later stages income disparities tend to disappear.

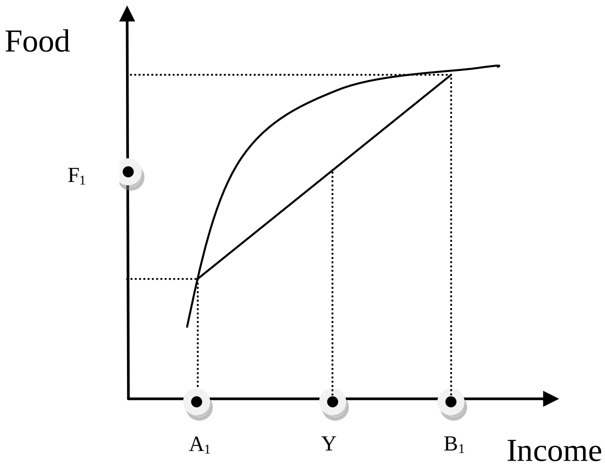
[1] See equation (2.21) in Section 2.3.5.

This hypothesis was formalised by Robinson (1975) in the context of a dual economic model. Robinson's model is based on two assumptions. The first assumption is that there are large differences in productivities and wages between urban and rural occupations. The second assumption is that there is a larger variety of occupations, and therefore wages, in urban areas compared to rural areas. Under these assumptions, as the development process forces the labour force to migrate from rural to urban areas, overall inequality increases. Further, the number of workers employed in highly paid jobs and the variability of wages in urban areas also increase. In spite of its great popularity, the inverted-U hypothesis has found little empirical support. Early empirical work by Ahluwalia (1975) confirmed the presence of an inverted-U hypothesis, but more recent studies have not found a clear relationship between income and inequality across countries (Anand and Kanbur, 1993, Deininger and Squire, 1996, World Bank, 2006).

There is also a body of literature that investigates the opposite direction of causality running from inequality to income growth. Some of this work studies how inequality may generate political demands for redistributions that affect long term growth (Bertola, 1993, Perotti, 1992, Persson and Tabellini, 1994). There is also empirical evidence from cross country studies that inequality, in particular land inequality, negatively affects income growth (Alesina and Rodrik, 1994). This latter argument was made strongly in the 2006 World Development Report on equity and development (World Bank, 2006). The World Bank report argues that initial levels of inequality affects economic growth in at least two ways. First, unequal access to markets results in individual economic choices that are inefficient for the economy as a whole. Second, unequal societies give rise to institutions that tend to perpetuate inequality and that miss opportunities for innovation and investments (World Bank, 2006).

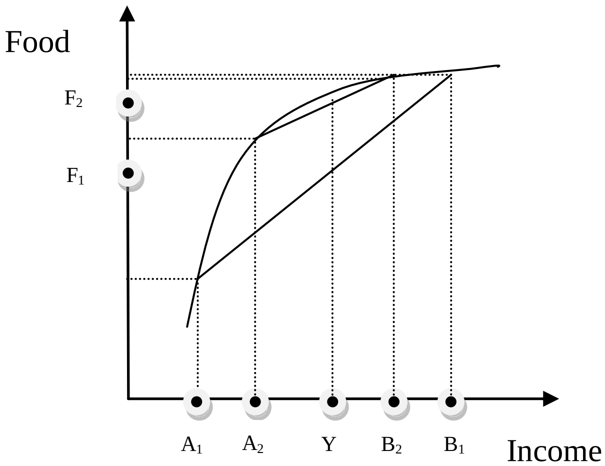
There is yet another way in which inequality affects income distribution. Income not only determines the level of consumption but also its composition. This follows from the non-linearity of Engel curves. As income increases, households' demand for food increases less than proportionally. Since individuals with different incomes consume food in different proportions, changes in the distribution of income affect food demand. This is known in demand theory as the aggregation problem: the transition from micro to macroeconomics of consumer behaviour (Deaton and Muellbauer, 1980c). The literature on the aggregation problem, that dates back to Antonelli (1886), has shown that exact aggregation of individual consumer demand is possible only if Engel curves are linear (Blundell and Stoker, 2005). Given that food Engel curves have been proved to be non-linear by hundreds of empirical studies, food consumption cannot be expressed as a function of income levels ignoring the distribution of income.

Figure 2.1 Food Engel curve



In order to illustrate the relevance of the distribution of income in determining food consumption patterns consider the chart in Figure 2.1. Suppose the economy is composed of only two individuals: the first individual (*A*) is poor while the second (*B*) is rich. Their average income is *Y*, while their average food consumption is F_1.

Figure 2.2 Food Engel curve with equalising transfer



Consider now a reduction in income disparity between the two individuals (Figure 2.2), by operating an income transfer from the rich individual (*B*) to the poor (*A*). While average income *Y* remains unchanged after the transfer, average food consumption increases to F_2. After the transfer, food consumption of the rich has only marginally decreased, while food consumption of the poor has increased dramatically.

This example shows that a change in the distribution of income affects food consumption. A decrease in inequality increases aggregate food demand, while an increase in inequality reduces aggregate food demand. Since one of the conclusions of this chapter is that the rate of urbanisation is a function of the income elasticity of demand, it follows that an increase in inequality accelerates the process of agricultural transformation.

There have been attempts in the literature to model the relationship between economic growth and inequality through demand patterns. De Janvry and Sadoulet (1983) formulate a model showing that high initial inequality in Brazil and Mexico, via consumption patterns, determines inequitable economic growth over the 1960s and the 1970s. Similarly, Baland and Ray (1991) developed a model in which the levels of output and employment are affected by the distribution of income via the structure of demand. In their model, income inequality generates high demand for luxury goods, which diverts resources from the production of necessities, and the market clears by reducing overall employment. The role of demand factors in the development of a dual economy is also the focus of a study by Bourguignon (1990), who concludes that an unambiguous reduction in inequality is unlikely at an early stage of development.

This chapter investigates the role of demand factors empirically, rather than modelling the interplay of inequality and economic growth via consumption patterns as in the studies mentioned above. Given the estimates of income elasticities of food demand, the effects of changes in income distribution on the food expenditure share of the economy are simulated. The simulations are then used to estimate the share of the changes in food demand that are explained by an increase in income and by a change in the distribution of income.

2.2.3 Economic growth and under-consumption theories

The main hypothesis formulated in this chapter is that extreme poverty and high food demand retard the process of economic growth. Weak demand for the goods produced in the non-agricultural sector prevents the economy from taking off, unless this demand is generated outside the domestic economy. The effect can be so large to constrain the economy into a low equilibrium poverty trap.

The idea that low levels of demand produce economic stagnation is not new in economics and dates back at least to Malthus (Bleany, 1976). Bleany (1976) reviews several ‘under-consumption’ theories that explain economic stagnation with poor consumers’ demand, and distinguishes two main strands of this literature. The first is the Malthusian type which is concerned with savings, while the second is the Sismondian type which is concerned with income distribution. The argument put forward by Malthus runs in the following way (Screpanti and Zamagni, 2005). Incomes of workers and landowners are entirely spent in the purchase of consumer goods, and completely resolve in effective demand. Profits, however, are mostly saved, and if the profits share in the economy increases more rapidly than the wage share, then the incomes paid to workers will not be able to provide a level of effective demand sufficient to absorb the value of the supply they produce. This arguments was further developed by Hobson in order to explain the business cycle (Bleany, 1976). Similarly to Malthus, Hobson argues that in expansion phases real wages decrease because do not fully adjust to the rise in prices (Bleany, 1976). At the same time, the profits share in the economy, and therefore savings and investments, increase. The result is that production increases more rapidly than demand, unsold inventories accumulate, prices and profits drop and the economy enters into depression.

While the Malthusian theory of underconsuption proposes a suggestive explanation of the business cycle, the Sismondian theory advances an explanation of economic stagnation that is closer in spirit to the one proposed in this chapter. Sismondi believed that the division of society between rich and poor was at the origin of economic crises and stagnation (Screpanti and Zamagni, 2005). Sismondi begins with the observation that the poor buy necessities while the rich buy luxuries. If the income distribution is skewed, the market for luxuries shrinks or never takes off. The poverty of workers is therefore the ultimate cause of economic stagnation, which should be addressed by redistributive policies. In a similar vein, Baran and Sweazy formulated a theory in which the unequal accumulation of income generated by the oligopolistic structure of the economy creates a problem of lack of demand (Bleany, 1976). This in turn creates a tendency of capitalist economies to stagnate.

More recently, Murphy et al. (1989a) built a model in which the equality of income distribution is a precondition for growth. This model predicts that in order for the non-agricultural sector to exploit increasing returns to scale, a minimum level of aggregate level of demand for manufactured goods is required. Weak demand originating from widespread poverty prevents the economy from taking off. These conclusions are obtained while assuming decreasing or constant returns to scale in agriculture and increasing returns to scale in manufacturing. These assumptions have been shown to generate economy-wide poverty traps (Bloom et al., 2003, Graham and Temple, 2006). The interpretation of the model presented in this chapter as a poverty trap model will be further discussed in Chapter 5, while the validity of the model in the open economy case is discussed in Section 2.7 of the present chapter.

2.3 A model of agricultural transformation in Andhra Pradesh

This section builds a simple general equilibrium model to explain the process of agricultural transformation in Andhra Pradesh. Several similar models can be found in the literature and the one presented here is particularly akin to the one formulated by Matsuyama (1992). Matsuyama formulated a general equilibrium model to show that a country's path to industrialisation varies depending on the economy being closed or open, and depending on its comparative advantage in the production of agricultural and manufacturing goods. The purpose of the model presented here is different, but the design of the structure of the economy and the assumptions regarding the functional relationship that govern production and consumption are very similar to those employed by Matsuyama (1992).

An early model of the process of agricultural transformation that stood the test of time is the one developed by Jorgenson (1961, 1967). Jorgenson modelled the joint dynamics of population, agricultural productivity, and demand for food. His model shows that agricultural transformation can be obtained either through technological advancement in agriculture or through the introduction of measures of fertility control. In more recent times, models in the same vein have been produced by Eswaran and Kotwal (1993) who showed the welfare effect on the poor of technological progress in agriculture; Echevarria (1997), who showed how the sectoral composition of the economy affects per capita income growth and vice versa; Leitner (2000), who modelled the effects of agricultural transformation on the saving ratio; and Gollin and Parente (2002), who aimed to explain differences in the rates of industrialisation between countries. Two other similar models deserve a separate mention for their originality. The first is a dual economic model built by Bourguignon (1990) that shows the possible effects of sectoral economic growth on inequality. The second is a model by Caselli and Coleman (2001) which shows how the agricultural transformation of the economy of the United States was accelerated by heavy investments in education in the southern agricultural states.

The model presented in this section shares with all these models a simplified characterisation of the structure of the economy, some basic assumptions regarding production and consumption processes, and the belief that technological progress in agriculture is the trigger of wider economic progress. Unlike the models mentioned above however, the one presented here focuses on the implications that different assumptions regarding the consumption function have on the results of this class of economic models.

The structure of the model is extremely simplified in order to focus attention on two determinants of economic development: technological progress in agriculture and inelastic food demand. The economy consists of only two sectors: agriculture, which produces food, and the modern sector, which produces a non-food item. Disaggregating the modern sector into manufacturing and services would bring the model closer to reality, but would not change the specification of the main economic relationship considered. The characteristics of the production process in the two sectors are very different. While it is assumed that production in the agricultural sector operates under constant returns to scale, it is assumed that the modern sector is characterised by increasing returns to scale. Agriculture employs land and labour, while the modern sector employs labour and capital. There is exogenous technological progress in agriculture that raises agricultural productivity and the production of food. The inelastic income demand for food brings agricultural prices down, and consequently labour demand in agriculture decreases. Workers move from the agricultural sector to the modern sector, where productivity is higher because of the presence of increasing returns to scale. The benefits of growth are distributed according to the contribution of each factor to the production process. Society is composed of four income classes:

farmers, agricultural labourers, capitalists, and urban workers. The dynamics of population, capital, and savings are ignored, and the attention is focused on the joint effect of technological progress and food demand on the sectoral composition of employment.

The model is built on a series of simplifying assumptions. The first is that the economy is closed and that the state engages in limited trade with the rest of the world. A second assumption is that prices are determined by the market, and that the state does not intervene in agricultural markets. Third, technological progress in agriculture is exogenous, and lastly, workers are allowed to move freely across sectors. The model is described in more detail in the following sections.

2.3.1 Agricultural production

All farmers cultivate plots of land of the same size and share the same production function. Production uses only one variable factor: labour (N), in the proportion a (with $0<a<1$), and one fixed factor: land (L). Farmers in Andhra Pradesh typically use very little capital, and the exclusion of capital from the agricultural production function is hence not too unrealistic. Total agricultural factor productivity (B_a) measures the current state of technology in agriculture. This variable can be thought of as the size of irrigation infrastructure, which together with the diffusion of modern seeds is the main determinant of increased agricultural productivity in Andhra Pradesh. Agriculture produces only one good (A) which is consumed as food. This good could be rice, which makes up to 30-40% of total food expenditure of an average household in Andhra Pradesh. The production function is the standard Cobb-Douglas with constant returns to scale:

$$A = B_a \left(aN\right)^{\alpha} L^{1-\alpha} \tag{2.1}$$

Notice that the price of the agricultural good (P_a) is used as the *numéraire*, which means that all equations are divided throughout by the price P_a. As a result, the agricultural price cancels out from all equations, while the price of the manufactured good (P_m) becomes the terms of trade between manufacturing and agriculture, that is the price of the manufactured good relative to the price of the agricultural good.

Agricultural labour is paid a wage w_a. The farm's cost function is simply $C = w_a aN$, while the profit function is $\pi = A - w_a aN$. Farmers employ labour to maximise profits, and the equilibrium wage per worker in agriculture is:

$$w_a = \alpha \frac{A}{aN} \tag{2.2}$$

Farmers' profits consist of an implicit land rent (r_l). Rent is the value of agricultural output left after farmers pay all hired labour including their own: $r_l = A - w_a aN$, substituting the equilibrium wage for w_a, the land rent is $r_l = A(1-\alpha)$ and the land rent per unit of land is:

$$r_l = (1-\alpha)\frac{A}{L} \tag{2.3}$$

Note that farmers also employ their own labour in the farm, that there is no unemployment, and that all agricultural labour force works for the same number of days. This implies that farmers' income is larger than agricultural labourers' income.

2.3.2 Manufacturing production

In the modern sector each firm (F_i) uses one variable factor (labour N_i), and one fixed factor (capital K_i), to produce the manufactured good M_i. The manufacturing sector in Andhra Pradesh is rather underdeveloped. The service sector produces the largest share of non-agricultural GDP and is the largest employer in the modern sector. Therefore, there is no dominant manufacturing product playing the role that rice plays in the agricultural sector. The good M can be thought of as cotton textile, or more generally as cloth, because the production of this good employs the largest share of the manufacturing workforce in Andhra Pradesh. Given that there are only two sectors in the economy, manufacturing production employs $(1 - a)N$ units of labour, where a is the share of labour employed in agriculture defined in the section above.

In manufacturing there are increasing return to scale determined by production externalities. The assumption that increasing returns to scale are external to the firm is adopted. This assumption is common in the growth economics literature (see for example Romer, 2006). In this formulation, each firm operates under constant returns to scale, but production externalities are generated at the level of the entire economy so that aggregate output grows under increasing returns to scale. Given the poor level of capitalisation of manufacturing in Andhra Pradesh, the externality is assumed to emerge from the workers' learning-by-doing during the production process and from specialisation in the division of labour. Technological progress can thus be modelled as an increasing function of the number of manufacturing workers.

Firms decide the levels of factors employment based on the current state of technology (T_i), and on the current marginal productivity of labour and capital (β and 1-β respectively). As a by-product of the production process, firms end up with more output than expected. Output is then distributed to factors in proportion to their contributions to production. The production function of the single firm is a Cobb-Douglas with constant returns to scale:

$$M_i = T_i N_i^{\beta} K_i^{1-\beta} P_m \quad . \tag{2.4}$$

Total factor productivity (T) is a function of manufacturing employment in the economy, where the parameter λ is the increasing returns to scale parameter:

$$T_i = B_i \left[(1-\alpha) N \right]^{\lambda} \tag{2.5}$$

Substituting this expression for total factor productivity in the manufacturing production function we have:

$$M_i = B_i N_i^{\beta} K_i^{1-\beta} \left[(1-\alpha) N \right]^{\lambda} P_m \tag{2.6}$$

By aggregating over all firms total output is:

$$M = B_m \left[(1-\alpha) N \right]^{\beta+\lambda} K^{1-\beta} P_m \tag{2.7}$$

in which the returns to scale are increasing, as the sum of labour and capital productivities is larger than one, and equal to 1+ λ. Note that the inclusion of increasing returns to scale in manufacturing implies that a movement of labour from agriculture has the effect of increasing workers' productivity in manufacturing and per capita output of the entire economy.

All firms maximise profits, employ the same technology, and use the same level of capital and labour (that is: $M_i=M/F$, $N_i = [(1-\alpha)N]/F$, and $K_i=K/F$). Factors payments to capital (r_k) and labour (w_m) are:

$$w_m = \beta \frac{M}{(1-\alpha)N} P_m \tag{2.8}$$

$$r_k = (1-\beta)\frac{M}{K} P_m \tag{2.9}$$

As in the case of the agricultural sector, capitalists earn a capital rent (r_k) in addition to a wage (w_m) for their own work in the firm, from which it immediately follows that they attain higher incomes than manufacturing workers.

2.3.3 Income

There are four income classes in the economy, two in manufacturing and two in agriculture. In manufacturing workers earn a wage w_m, while capitalists earn a rent r_k for each unit of capital in addition to their labour income. Similarly, in agriculture workers earn a wage w_a, while farmers earn an implicit rent for their land r_l in addition to the remuneration for the use of their own labour. Table 2.1 summarises the expressions for incomes accruing to each class.

Table 2.1 Income distribution by class

Income class	Income
Capitalists	$r_k K + (1-a)Nw_m$
Manufacturing workers	$(1-a)Nw_m$
Farmers	$r_l L + aNw_a$
Agricultural labourers	aNw_a

Aggregating over all income classes and assuming that in the long term wages in agriculture and in manufacturing are equal ($w_a = w_m = w$), total income (Y) in the economy is simply:

$$Y = wN + r_l L + r_k K \tag{2.10}$$

2.3.4 Consumption

Consumption of agricultural and manufactured goods are functions of income and prices. The demand equations in general form are: $C_a = f(Y, P_i)$, and $C_m = f(Y, P_i)$. The particular forms assumed by the demand functions depend on the assumptions made regarding consumers' preferences. Three cases will be considered: homothetic preferences, hierarchical preferences and quasi-homothetic preferences. The demand equations for each system of preferences will be derived, and their ability to describe observed consumers' behaviour will be discussed.

Notice that demand equations derived from the theory of consumer behaviour must satisfy three basic requirements (Deaton and Muellbauer, 1980c). These are the adding-up condition, homogeneity, and a reasonable representation of observed consumer behaviour.[2] The adding-up condition stipulates that consumers cannot spend more than their income ($C_a+C_m=Y$) and that the budget constraint must be satisfied. It also implies some testable restrictions on the derivatives of the demand functions with respect to income that are known as Engel's equations.[3] Homogeneity means that changes in income and prices by the same amount do not change the quantity demanded of any of the goods. Formally this is equivalent to demand equations that are homogenous of degree zero: $f_i(\vartheta y, \vartheta p) = f_i(y, p)$. Finally, demand equations should be able to represent luxuries (goods whose demands increases more than proportionally with

---

[2] There are two other properties that demand equations must satisfy: symmetry and negativity. These however are not relevant for the present discussion.

[3] Engel's equations state that changes in income and prices determine changes in the composition of the budget constraint, leaving its value unchanged. Formally, Engel's equations state that the sum of the derivatives with respect to income of all goods are equal to one:

$\sum_i p_i \frac{\partial q_i}{\partial y} = 1$. In budget shares form the Engel's equations are: $\sum_i w_i \eta_i = 1$, where w is the share consumed of good i, and η is the income elasticity for the same good.

income), necessities (whose demand increases less than proportionally with income), and inferior goods (whose demand decreases as income increases).

Homothetic preferences

Homothetic preferences mean that preferences are independent of total expenditure (Deaton and Muellbauer, 1980c). For a factor $\theta>0$ the utility function is homogenous of degree one:

$$U(\theta q_i) = \theta U(q_i) \tag{2.11}$$

In other words, doubling the quantities consumed of each good doubles utility, and utility is produced under constant returns to scale. The composition of the budget is therefore independent of the scale of total expenditure. The Cobb-Douglas and the Constant Elasticity of Substitution are typical examples of utility functions with homothetic preferences. Appendix A shows how to obtain a set of demand equations from the maximisation of a Cobb-Douglas utility function. The derived quantities demanded of the two goods are:

$$C_a = cY \qquad \text{and} \qquad C_m = (1-c)\frac{Y}{P_m} \tag{2.12}$$

It can be easily verified that the adding-up and homogeneity conditions are satisfied. These demand equations however perform rather poorly in representing consumer behaviour. All income elasticities are equal to one and no luxuries or necessities are allowed. The Engel functions are straight lines through the origin, and the consumption shares of all commodities are constant and independent of total expenditure. Most importantly, homotheticity is contradicted by empirical budget studies. Expenditure elasticities are rarely equal to one. The budget composition varies across households of increasing expenditure, and over time as total expenditure increases for all households.

Hierarchic preferences

Hierarchic preferences mean that consumers have a hierarchy of wants over different goods. A hierarchic utility function implies that the set of goods purchased changes with total expenditure. As income increases the number of goods purchased also increases, as some new goods are purchased while others are abandoned, and in general the rich purchase a larger number of goods than the poor. The concept of hierarchical preferences has received little attention by demand theorists, with the notable exception of Fraser Jackson (1984), but it has been widely used in agricultural transformation models for its simplicity and tractability.

With only two goods, hierarchic preferences mean that consumers spend all their income on the inferior item (food) until they reach a level of satisfaction (g_a), after which any additional income is spent on the superior good (the manufactured good). The demand equations for the two goods in the two states are:

$$C_a = Y \quad \text{and} \quad C_m = 0 \quad \text{if} \quad Y \le g_a \tag{2.13}$$

$$C_a = g_a \quad \text{and} \quad C_m = Y - g_a \quad \text{if} \quad Y > g_a$$

Adding-up and homogeneity are obviously satisfied. Hierarchic demand equations represent a considerable improvement in terms of representing consumer behaviour compared to the homothetic case. While in the homothetic case all income elasticities are equal to one, in the hierarchic case they are allowed to switch from the value of zero to the value of one as expenditure increases. This however is a credible representation of consumers demand for only a small number of goods, and an alternative formulation that lies between the homothetic and the hierarchical case is required to better approximate consumer behaviour.

Quasi-homothetic preferences

Quasi-homothetic preferences imply a minimum level of expenditure on at least one of the goods. A typical utility function for this type of preferences is:

$$U = c\ln(C_a - g_a) + (1-c)\ln(C_m - g_m) \tag{2.14}$$

where g_a and g_m represent the minimum expenditure on food and on the manufactured good respectively. The derived demand functions, assuming for simplicity that there is no minimum expenditure on the manufacturing good, are: [4]

$$C_a = g_a + c(Y - g_a) \quad \text{and} \quad C_m P_m = (1-c)(Y - g_a) \tag{2.15}$$

Adding-up is satisfied provided that the sum of the propensities to consume of the two goods is equal to one, and in the equations above this was imposed by construction. Income elasticities can now take any value, and are given by:

$$\eta_a = \frac{c}{w_a} \quad \text{and} \quad \eta_m = \frac{c-1}{w_m} \tag{2.16}$$

Luxuries will have elasticities greater than one, while necessities will have elasticity less than one. As expenditure increases, the share of expenditure on luxuries over total expenditure increases, while the share of expenditure on necessities decreases.

While homothetic Engel curves are lines through the origin, quasi-homothetic Engel curves are straight lines that do not go through the origin. The elasticities vary with the share of each good consumed over total expenditure, and therefore vary with total expenditure. In the long term however, as expenditure grows to infinity all income elasticities become equal to one, budget shares become equal to c and c-1, and the system becomes homothetic.[5]

Quasi-homothetic demand functions represent a considerable improvement in terms of characterising consumer behaviour compared to the homothetic and the hierarchic cases. There are however two characteristics of these demand equations that contradict observed consumer behaviour. First, the propensities to consume can only be positive, which rules out inferior goods. Secondly, in cross-sectional studies of household budgets Engel curves have been found to be non-linear.

---

[4] See Appendix A for the derivation of these demand equations.

[5] This can be seen by applying l'Hôpital rule. Both the denominator and the numerator of the elasticities in (2.16) grow to infinity as expenditure increases. Differentiating both terms with respect to 1 we have c/c=1.

2.3.5 Model solution

The model can be solved by setting three equilibrium conditions. The first condition is the equality of the agricultural and urban wages, which follows from the assumptions of free movement of labour and absence of unemployment. The second and third conditions are the equality of supply and demand in the markets for the agricultural and manufactured goods:

$$w_a = w_m, \qquad A = C_a, \qquad \text{and} \qquad M = C_m \tag{2.17}$$

By equating the wage equations (2.2) and (2.8) and the output equations (2.1) and (2.7) to the respective demand equations, the model can be reduced to a system of three equations in three unknowns: the share of agricultural employment in total employment (a), total income (Y), and the sectoral terms of trade (P_m):

$$\begin{aligned} P_m &= f(a) \\ Y &= f(a) \\ a &= f(P_m, Y) \end{aligned} \tag{2.18}$$

A sequential solution can be obtained by solving the equality in the labour market for the terms of trade, solving the equality in the market for the agricultural product for income, and then substituting the expressions for income and the terms of trade in the equilibrium of supply and demand of the manufactured good. Appendix B illustrates the solutions obtained under the three different hypotheses regarding the consumption function.

The model solution proposed produces an expression for the share of employment in agriculture in terms of the exogenous variables and parameters: $a = f(c, \alpha, \beta, g_a, L, K, N, B_a)$. The effect of technological progress in agriculture on employment in the agricultural sector can be seen in comparative statics by differentiating the agricultural employment share with respect to the total factor productivity: $\frac{\partial a}{\partial B_a}$. In what follows it will be shown that the sign and size of this derivative depend on the assumptions made regarding the consumption function.

The homothetic case

In the case of homothetic preferences there is a unique solution for the share of employment in the agricultural sector that depends only on the production elasticity of labour in the two sectors, and on the income elasticity of food. The share of employment in agriculture increases with the average propensity to consume food and with labour productivity in agriculture, and it decreases with the increase in labour productivity in the manufacturing sector:

$$a = \frac{c\alpha}{\beta - \beta c + \alpha c} \tag{2.19}$$

The derivative $\frac{\partial a}{\partial B_a}$ with respect to total factor productivity in agriculture is zero, and the share of employment in agriculture is independent of technological change. With homothetic preferences technological change has no effect on the sectoral composition of output, as consumers are willing to absorb any increase in food supply. Homothetic preferences are never used in models of agricultural transformation, and they are shown here only to illustrate an extreme case. Demand equations derived from homothetic preferences however, are sometimes used in general equilibrium models, probably ignoring the strong implications that they impose on the working of the economy.

The hierarchic case

In the case of hierarchical preferences the derivative $\frac{\partial a}{\partial B_a}$ with respect to total factor productivity in agriculture is:[6]

$$\frac{\partial a}{\partial B_a} = -\frac{1}{\alpha} g_a^{\frac{1}{\alpha}} N^{\frac{1-\alpha}{\alpha}} B_a^{-\frac{1+\alpha}{\alpha}} \tag{2.20}$$

The value of this derivative is with no doubt negative, because all the terms on the right-hand-side are positive. Once all consumers have achieved the desired minimum level of

[6] See Appendix B for the derivation of (2.20).

food consumption g_a, agricultural production remains fixed, and increases in total factor productivity in agriculture make agricultural employment immediately redundant, and the share of agricultural labour decreases. This function has been widely used in modelling the process of agricultural transformation (see for example Gollin et al., 2002, Jorgenson, 1967, Laitner, 2000).

The quasi-homothetic case

In the case of quasi-homothetic preferences the solution to $\frac{\partial a}{\partial B_a}$ is rather complex and can be shown to be equal to:

$$\frac{\partial a}{\partial B_a} = -\frac{-\frac{\beta a g_a N(1-c)}{B_a}}{(\beta c - c\alpha - \beta)A + g_a N\beta(1-c)(1-\alpha)} \tag{2.21}$$

This derivative is negative, and therefore the share of employment in agriculture decreases with technological change in agriculture, provided that the agricultural sector produces enough to satisfy the minimum requirement of food per capita.[7] When per capita production in agriculture is small, the share of employment in agriculture may increase, as consumers demand larger quantities of the agricultural good until reaching the minimum requirement g_a. Notice that the larger the propensity to consume food out of current income (*c*), the smaller is the value of the derivative, and hence the smaller is the effect of technological change on agricultural employment. Similarly the larger the minimum requirement of food g_a, the smaller is the value of the derivative. Notice also that this derivative is in absolute value smaller that the derivative produced by hierarchic preferences. This can be shown by considering that by reducing the value of *c*, the size of the derivative decreases, and when *c* is reduced to its minimum value (zero), we are back to the hierarchic preferences case.

The results of this theoretical review of the effects of consumer preferences on agricultural employment can be summarised in the following way. Homothetic preferences and unitary elasticities of food consumption rule out any effect of

[7] See Appendix B for a derivation of this derivative, and for a proof that it is negative.

technological change on agricultural employment, and under these demand conditions the process of agricultural transformation cannot take place. Hierarchical preferences imply a faster process of transformation compared to quasi homothetic preferences, but they oversimplify economic behaviour. Quasi-homothetic preferences introduce more flexibility in the model, by allowing the income elasticity of food and the minimum food requirement to determine the size of the effect of technological progress on agricultural employment.

2.3.6 Income distribution and food consumption

One shortcoming of quasi-homothetic preferences is that, as already noted, the derived system of demand equations consists of Engel curves that are linear in expenditure. However, empirical studies of household budgets have found non-linear Engel curves for many commodities as well as for broad aggregates of commodities.

In principle, the non-linearity of consumer demand is not problematic. First, linear demand equations can be a good approximation if what is investigated is the short-term behaviour of the model. Second, non-linear demand equations in cross sections do not necessarily contradict linear demands because budget studies observe consumption of different households, while what is needed is observing consumption of the same households over time. As Deaton and Muellbauer (1980c) point out, it is not obvious that Engel curves should be non-linear in panel data and for broad categories of goods like food. Finally, consumption aggregates can become larger over time, as new items are introduced in the groups. Therefore the change in demand brought about by a change in income over time may differ from the change in demand brought about by the same change in income over a cross section of households.

The non-linearity of Engel curves however is problematic in one respect, which is known in demand theory as the aggregation problem (Deaton and Muellbauer, 1980c). Demand theory models consumption of individuals or households, and in order to obtain the market demand for a commodity, individual demands are aggregated into a single demand function. Certainly there are differences between the individual consumption functions, but it is assumed that these differences average out and that final demand only depends on prices and income. The procedure of summing up all individual demands in order to calculate market demand is correct if individual demand equations are linear, but it may produce unexpected results if demand equations are non-linear.

Consider for example the linear demand equation of a single household implied by quasi homothetic preferences:

$$C_{ai} = (1-c)g_{ai} + cY_i \tag{2.22}$$

Summing over all households in the economy the market food demand is:

$$C_a = (1-c)g_a N + cY \tag{2.23}$$

where $g_a = \frac{\sum_i g_{ai}}{N}$ and $Y = \sum_i Y_i$. The marginal propensity to consume food (*c*) is the same for all households, and the subscripts *i* for each household have disappeared. Suppose now of operating a transfer of income from a rich household to a poor one. With linear Engel curves nothing changes, as all households consume food in the same proportions. With non-linear Engel curves however food consumption increases because the poor spend a larger share of expenditure on food.

Linear demand equations like (2.23) fail to detect distributional effects on food consumption. Changes in food consumption are not only determined by changes in aggregate income, but also by changes in income distribution. One simple way to incorporate distributional effect in the food demand equation is adding a term that summarises the inequality of the distribution. For example Chambers and Pope (1992) use the following specification:

$$C_a = g_a N + b_1\sigma^2 + b_2 Y + b_3 Y^2 \tag{2.24}$$

where σ^2 is the variance of income distribution. The quantity of food consumed varies with expenditure in a non-linear way and is linearly related to income distribution. The more equal is the distribution of income, the larger is the consumption of food, and the more unequal is the distribution the smaller is the market demand for food.

A demand function like the one in (2.24) could be incorporated in the model outlined above. With its inclusion, the demand for food and the effect of technological progress on agricultural employment would also depend on changes in income distribution. An increase in inequality would reduce food demand and accelerate the agricultural transformation, while a reduction in inequality would have the opposite effect. One additional complication is that the distribution of income is endogenous to the model and therefore the direction of the change cannot be determined a priori. Before doing further modelling, the effect of a change in income distribution on food consumption is assessed empirically in order to establish its relevance (see Section 2.6).

2.4 Data and descriptive statistics

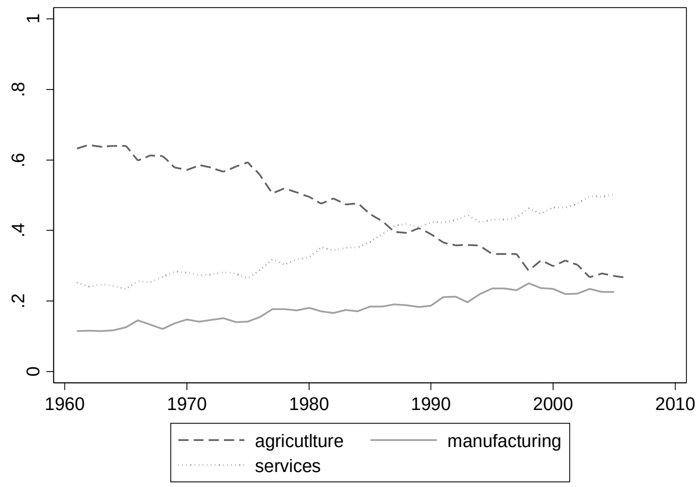
This section presents some stylised facts about the process of agricultural transformation in Andhra Pradesh. It draws on national level data to show changes over time in agricultural production as a percentage of total GDP, agricultural employment as a share of total employment, technological progress and productivity in agriculture, food demand and income inequality. The extent to which these facts fit with the model of agricultural transformation outlined above is discussed at the end of this section.

2.4.1 Agricultural production and employment

The size of the agricultural sector in Andhra Pradesh has considerably decreased over the last 45 years. Figure 2.3 shows the trend in the share of agricultural output over total GDP for the period between 1961 and 2006, using data collected by the DES (2005). Figures are calculated at 1993-94 constant prices and reflect the output of the primary sector, which according to the DES definition includes agriculture, forestry and logging, fishing, and mining and quarrying. Disaggregated data for the period 1981-1993 show that more than 90% of the primary sector output is generated in agriculture, and the data can therefore be considered representative of the agricultural sector.

The chart in Figure 2.3 clearly shows that the share of agricultural output over total GDP has decreased over time, from above 60% in 1960 to less than 30% in more recent times. Though the common perception is that Andhra Pradesh is a predominantly agricultural state, it should be recognised that in terms of output produced, agriculture now represents less than one third of the total value of the economy.

Figure 2.3 Sectoral composition of GDP in Andhra Pradesh (1960-2005)



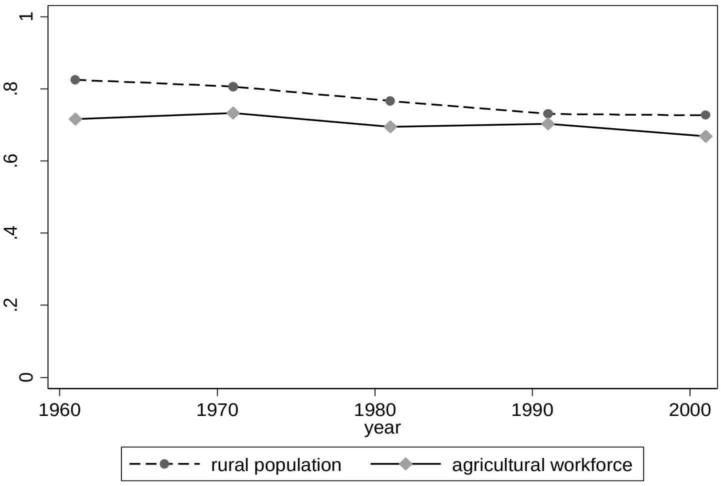
Source: calculated from DES (2005).

If Andhra Pradesh is not an agricultural state in terms of output produced, it certainly is in terms of employment of its workforce. Figure 2.4 shows the trends in the shares of rural population over total population and in the share of agricultural employment over total employment for the period from 1961 to 2001. The chart uses data collected by the last five rounds of the population census of India. The ratios were calculated for the years 1961, 1971, 1981, 1991, and 2001, and the lines between the data points were drawn by simple linear interpolation. Rural population in the state has declined only marginally over the period considered, from above 80% to 75% of total population. Similarly, the share of agricultural employment over total employment has decreased only slightly and has remained rather stable over the last 50 years.

Some limitations of the population census data that have a bearing on these numbers need to be acknowledged. First, there is some degree of arbitrariness in the definition of rural and urban agglomerations. The Census of India classifies an agglomeration as urban if it satisfies three criteria: population is above 5,000 inhabitants, at least 75% of the population is employed in non-agricultural activities, and population density is at least 400 inhabitants per squared kilometre. Second, rural population is not a good indicator of employment in agriculture, because by the census own definition, up to 25% of population in urban areas may be employed in agriculture, while population of villages below 5,000 inhabitants may well be employed in non-agricultural activities.

Third, the agricultural workforce was calculated using data on the numbers of cultivators, agricultural labourers, and livestock workers. The census definitions of economic activities are however problematic, because they have changed over time, and data on female employment have not always been collected or included in the data reported.

Figure 2.4 Percentage of rural population and agricultural workforce (1961-2001)



Source: calculated from Census of India (2001).

Figure 2.5 further disaggregates the share of agricultural workforce reported in Figure 2.4 by farmers and agricultural labourers. The absolute number of farmers (ranging from seven to eight million) has not substantially changed over the period considered, though their relative number has decreased. On the other hand, the absolute number and the share of agricultural labourers have increased, and they now represent more than 40% of total employment. At the same time, Andhra Pradesh, as the rest of India, has witnessed a process of feminisation of the agricultural labour force. The Indian census reports that the percentage of women classified as agricultural labourers has doubled from 25.6% to 49.6% over the period from 1961 to 1981 (Da Corta and Venkateshwarlu, 1999). In 2001 women constituted 46.9% and 53.5% of the agricultural workforce in India and Andhra Pradesh respectively. This process of feminisation of the labour force has been differently explained by the neoliberal and the Marxists schools of thought (Garikipati, 2006). The neoclassical economic literature has stressed the operation of demand factors. According to this school, the increase in female employment is the result of an increase in labour demand for those tasks, like rice weeding and harvesting, that are primarily performed by women (Walker and Ryan, 1990). This interpretation is supported by the observed reduction in the male-female wage differential for agricultural labourers during the green revolution (Hazell and Ramaswamy, 1991). This interpretation is challenged by authors writing in the Marxist tradition. Da Corta and Venkateshwarlu (1999) conducted a series of interviews in 25 villages of the Chittoor district in Andhra Pradesh and concluded that the increasing participation of women in the agricultural labour force was a result of farmers' attempts to reduce labour costs at a time when male workers were increasingly engaged in off-farm activities. Garikipati (2006) conducted extensive field studies in the Mahaboobnagar district of Andhra Pradesh in 2001 and 2002 and concluded that the feminisation of the agricultural labour force resulted in an increase in the women's workload both within and outside the home, which was accompanied by little improvement of their living standards.

Figure 2.5 Percentage of farmers and agricultural labourers (1961-2001)

Source: calculated from Census of India (2001).

The large share of rural population is the result of limited migration from rural to urban areas. The agricultural labour force grew by an average of 1.6% per year compared to a rate of growth of 1.9% of total population. Assuming that fertility rates in rural and urban areas are similar, the difference between these two rates can be considered an indicator of the rate of migration from the agricultural to the non-agricultural sector, expressed as a percentage of the agricultural labour force. A rate of migration of 0.3%

per year compares rather poorly with the median value of 2% found over the period 1959-1990 by Mundlak et al. (1997) for a sample of 130 countries.

A low rate of migration is confirmed by census data of 2001 (Office of the Director of Census Operations, 2005), which show that 70 % of total population was born in the place of enumeration, and only 6% was enumerated in a district different from the district of birth. The percentage of people enumerated in a location different from the place of birth but within the same district was 22%. It is useful at this point to disaggregate the data on migrants by origin and destination, and by sex (see Table 2.2). A large component of female migration occurs within rural areas, and other data in the census indicate that the main reason for female migration is marriage. Male migration is mainly motivated by economic reasons, but only about 25% of this migration occurs from rural to urban areas.

These data should be treated with caution because the 2001 census classified persons as migrants if they were enumerated at a place different from their place of birth or from their place of last residence. This definition of migration is problematic in several ways. First, this method measures migration as a stock rather than a flow. The flow of migrant over a given period of time cannot be calculated because respondents are defined as migrants if they ever migrated in the course of their life, irrespectively of when this happened. Secondly, this definition ignores temporal migration. People who migrated, even for long periods of time, and then returned to their place of origin are not classified as migrants. Finally, part of the migrant population will have died by the time of the interview, and if mortality rates of migrants and non-migrants differ, migration rates might be biased.

Table 2.2 Percentage of migrants by area of origin and destination

Migration flow	Male migrants	Female migrants
Rural to rural	45.3	73.0
Rural to urban	26.7	12.4
Urban to urban	20.2	8.7
Urban to rural	7.7	5.9
Total	100.0	100.0

Source: calculated from census migration data (Office of the Director of Census Operations, 2005).

2.4.2 Agricultural productivity

The model outlined in Section 2.3 identified technological progress in agriculture as one of the two main drivers of the process of agricultural transformation. Descriptive evidence on the operation of technological progress in Andhra Pradesh will now be briefly discussed.

Table 2.3 shows rates of growth of output in agriculture and in the modern sector between 1961 and 2006. It also reports per capita growth rates, measured over rural and urban populations, and productivity growth rates, measured over employment in agriculture and the modern sector. All indicators show that the performance of the agricultural sector was poorer than that of the modern sector over the period considered. These data however do not adequately capture productivity growth in agriculture for two reasons. First, the growth rates are aggregated over a long time period, and much of the productivity growth in agriculture occurred during the last three decades. Secondly, growth rates per capita and per worker are good indicators of productivity growth from a welfare point of view, as they show the output produced per person. However, from a technical point of view growth in output per unit of land is a better indicator of technological progress.

Table 2.3 Growth rates in agriculture and in the modern sector (1961-2006)

	Overall growth	Per capita growth	Growth per worker
Agriculture	2.4	0.5	0.9
Modern sector	6.3	4.4	3.9
Total GDP	4.6	2.7	2.7

Source: calculated from DES (2005).

Figure 2.6 shows the trends of agricultural output per hectare and per worker in Andhra Pradesh over the last 50 years measured at 1993-94 constant prices. Starting from the mid 1970s, output per hectare has more than doubled, while output per worker has only slightly increased over the same period. The different growth patterns of land and labour productivity result from the specific characteristics of the technological progress in agriculture. In the mid 1970s, improved hybrid seeds were introduced together with the expansion of irrigation and the use of fertiliser (the green revolution package). This technology is eminently land saving and produced a considerable increase in production per unit of land. Productivity per worker however, has lagged behind, and Andhra

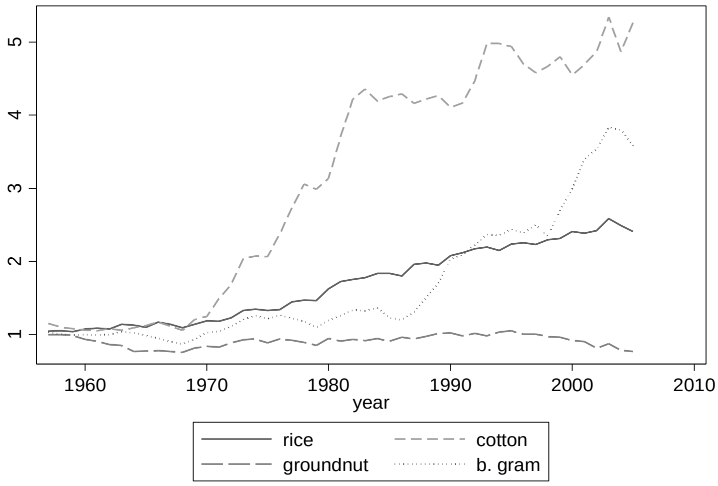
Pradesh seems to have followed what Hayami and Ruttan (1971) call the Asian path to agricultural growth. Land productivity increased substantially after the start of the green revolution, but high rates of population growth and poor absorption of excess labour by the non-agricultural sector resulted in a large increase of the agricultural labour force, which in turn is reflected in low productivity per worker.

Figure 2.6 Labour and land productivities in agriculture (1960-2005)

Source: calculated from DES (2005).

It should also be emphasised that the technological progress was not even across all cultivated crops. Figure 2.7 shows the productivity indices of four main crops in Andhra Pradesh: rice, cotton, groundnut, and Bengal gram (a pulse). Rice is by far the most important crop in Andhra Pradesh, with 30% of cultivated land under this crop in 2005. Cotton has become an important commercial crop and is cultivated on nearly 10% of all cultivated land. Groundnut is produced in the dry land areas of the south of the state, and grows on some 15% of the cultivated area. The chart shows that while yields of rice and cotton have increased considerably, yields of groundnut have remained stable and even decreased in recent times. The chart also clearly shows that the early 1970s marked the beginning of agricultural productivity growth in Andhra Pradesh.

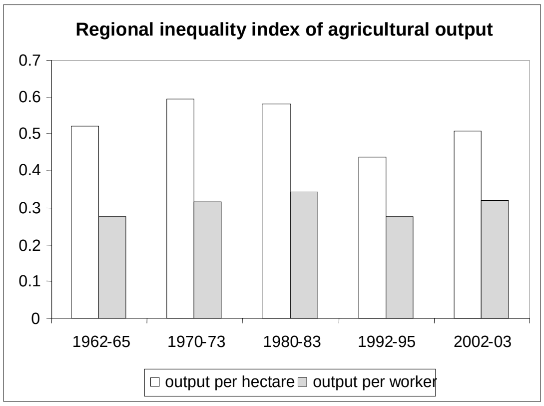
Figure 2.7 Productivity indices of main agricultural crops (1956-2005)



Source: calculated from DES (2005).

There was also considerable unevenness across region in productivity growth, which was the result of differences in agro-ecological conditions and in irrigation infrastructure. As discussed in Section 1.2, Andhra Pradesh can be subdivided into three main macro-regions that are differently endowed in terms of soil characteristics, rainfall patterns and irrigation infrastructure. The Coastal region is served by good irrigation infrastructure and benefits from good soils and rainfall. Irrigation in the Telangana is poor, but the region enjoys relatively good soils and rainfall. The southern region of Rayalseema lacks both good soils and infrastructure and it is mainly rainfed.

Figure 2.8 District level disparities in Andhra Pradesh: 1962-2003



Source: Bhalla and Singh (2001) for the series up to 1992-95 and DES (2005) for the 2002-03 series.

The green revolution had a disequalising effect on the regions of Andhra Pradesh. Figure 2.8 shows the coefficients of variations of output per hectare and output per worker across districts and over time. The diffusion of the HYV (high yield variety) technology was confined to the production of wheat in Northern India over the period from 1962-73, but over the 1970s and the 1980s it spread to the cultivation of rice in Coastal Andhra Pradesh (Bhalla and Singh, 2001). This is reflected in the chart of Figure 2.8, which shows an increase in district level inequality in the value of agricultural output per hectare over the same period.

The 1980s witnessed and expansion in the use of fertiliser in all areas of the State, an increase in the use of water-pump irrigation in areas not served by canal infrastructure, and a diversification of production patterns away from cereals to oilseed and cotton (Bhalla and Singh, 2001). In addition, the area irrigated by canals declined over the 1990s (IEG, 2007). The simultaneous operation these factors resulted in a substantial decrease in the productivity differentials across districts.

Substantial differences however persist between the three regions in terms of agricultural productivity and access to irrigation. Table 2.4 illustrates the historical patterns of growth of agricultural productivity and of irrigated areas by region. In Table 2.4 the Coastal region is further subdivided into a North Coastal and a South Coastal region as the two are characterised by significant differences in soil and infrastructure characteristics - the South Coastal being the better endowed of the two. The figures in Table 2.4 show the large advantage of the Coastal areas over the rest of the state in terms of both agricultural productivity and access to irrigation. The Telangana region has made considerable progress over the period considered, while productivity growth in the Rayalseema region has been modest. The last two columns of Table 2.4 show the change occurred in the percentage of the area irrigated, and provide an indication of the link between water provision and agricultural productivity. Growth in agricultural productivity was highest in Telangana where the area irrigated doubled over the period. Only modest increases in productivity were observed in the Coastal region where the increase in the share of irrigated area was modest. The Rayalseema region experienced little productivity growth in spite of almost doubling the share of irrigated area, which could be explained by the increasing degradation of soils and the poverty of the production technology employed (Subrahmanyam, 2003).

Table 2.4 Regional breakdown of agricultural productivity and irrigated area in Andhra Pradesh: 1955-1999

	Crop output value per hectare					Percentage of irrigated area	
Region	1955-58	1969-72	1979-82	1989-92	1996-99	1955-58	1996-99
North Coastal Andhra	148.7	121.0	94.9	99.2	87.5	50.1	44.8
South Coastal Andhra	194.0	149.7	154.1	144.6	149.1	46.2	62.6
Telangana	49.0	58.6	61.8	71.5	77.9	15.8	36.0
Rayalseema	58.7	81.0	73.4	81.3	70.9	12.6	23.3

Source: Subrahmanyan (2003)

Palmer-Jones and Sen (2003), have stressed the importance of initial conditions in determining the rate of agricultural growth in rural India. If agro-ecological conditions influenced the location of irrigation facilities in Andhra Pradesh, this would have resulted in an uneven pattern of agricultural growth at the regional level and much of the impact on agricultural growth attributed to irrigation would be in fact the consequence of an initial agro-ecological advantage. If this is true, then the success of irrigation policies in advantaged areas cannot be replicated in areas that are not favourably endowed. This in turn could explain the poor growth of Rayalseema in spite of the increase in the share of irrigated area.

2.4.3 Food demand

The second driver of the process of agricultural transformation according to the model outlined in Section 2.3 is the diminishing demand for food. Table 2.5 shows the shares of food expenditure over total expenditure calculated using household survey expenditure data collected by the NSSO over the period from 1987 to 2005. Average food shares were calculated in two ways. The household food shares are simple averages of households' shares. The market food shares are weighted averages of households' food shares, where the weights are households' expenditure shares of total expenditure in the sample:

$$\overline{w} = \sum_i \frac{y_i}{Y} w_i \qquad (2.25)$$

where y_i is household expenditure and Y is total expenditure in the sample. Market food shares are less representative of household behaviours, but more representative of market behaviour.

Table 2.5 Food expenditure shares (1987-2005)

	1987-88	1993-94	1999-00	2004-05
Household food shares				
all sample	61.9	61.2	54.5	53.9
Rural	62.8	63.0	57.4	57.4
Urban	58.5	56.0	47.2	43.1
Market food share				
all sample	58.2	58.3	49.7	51.1
rural	59.3	60.2	53.9	55.8
urban	55.0	53.9	42.6	41.1

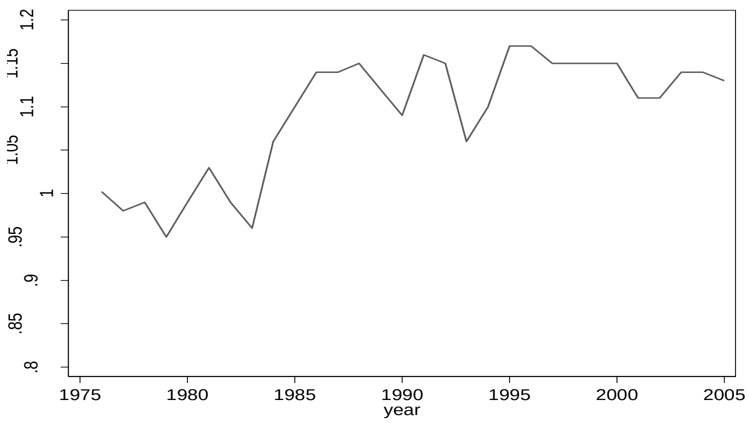
Source: calculated from NSSO data.

The food shares shown in Table 2.5 are high even for a low income country. Seale et al. (2003), replicating a study conducted earlier by Theil et al. (1989), calculated food budget shares of 114 countries using data collected by the International Comparison Programme between 1993 and 1996. Food shares ranged from a minimum of 9%, in the United States, to a maximum of 73% in Tanzania. Low income countries, defined as countries with per capita expenditure less than 15% the expenditure in the United States, were found to have an average food share of 52%. In Andhra Pradesh around the same time (1993-94 survey) reported food shares were about 60%.

Food shares have only slightly declined over the period from 1987 to 1993. They are smaller in urban areas, where incomes are higher, and decreased faster in urban areas than in rural areas. Table 2.5 shows a sizeable drop between the survey of 1993-94 and the survey of 1999-00. This might be the consequence of a bias in the measurement of household expenditure, which resulted from changes made to the survey questionnaire used in the 1999-00 round. The standard NSSO questionnaire uses a 30-day recall period for high-frequency expenditures, namely food, stimulants and intoxicants, and miscellaneous goods and services. In addition, the questionnaire uses both a 30-day and a 365-day recall period for low-frequency purchases of clothing and footwear, education, health, and durable goods. However, in 1999-00 the NSSO tested a new questionnaire in which food data were collected using both a 7-day and a 30-day recall period, whilst some low-frequency purchases used only the 365-day recall. These modifications are believed to have biased the comparability of the data in two ways (Deaton and Kozel, 2005). The first source of bias derives from the fact that the 7-day recall has been shown in past experiments to produce larger expenditures than the 30-day recall. In addition, the 7-day recall was conducted on the same page and simultaneously to the 30-day recall. Therefore it is likely that both the respondent and the interviewer would reconcile the amounts during the interview. The result is that the 30-day food expenditure in the 1999-00 round is biased upward. The second source of bias is that from experiments with the 'thin' rounds have shown that the poor report larger expenditure on non-food items when using a 365-day recall, while the rich report a lower expenditure on a 365-day recall. Since the expenditure of the rich on non-food items is much higher than that of the poor, the general result is an underestimation of expenditure. Since these two biases increase both food and non-food expenditure, it is difficult to predict their effect on expenditure shares, as these can vary in both directions. However, given that in general the food share decreases as income per capita increases, and given that the food share of 2004-05 is only slightly smaller than the one of 1999-00, it is likely that the latter was underestimated, and that the decrease in the food share over the period was smoother than it appears in Table 2.5.

The stability of food demand in Andhra Pradesh seems to be confirmed by trends in agricultural prices. If agricultural supply exceeds demand, then in the absence of state interventions in agricultural markets, prices of agricultural products should fall. Figure 2.9 shows a price index constructed using DES data on prices of all agricultural commodities produced in the state for the period from 1975 to 2005. Nominal prices were deflated by the series of the Consumer Price Index for Agricultural Labourers (CPIAL). The chart shows that prices of agricultural goods increased in real terms in the second half of the 1980s and remained stable thereafter.

Figure 2.9 Real price index of agricultural goods (1975-2005)



Source: calculated from *indiastat*.

2.4.4 Income inequality

Over the period from 1987 to 2005, Andhra Pradesh experienced an increase in per capita income and in the equality of income distribution both in urban and rural areas. Table 2.6 shows the value of median per capita expenditure in the four survey rounds. Medians were calculated taking into account probability sampling weights following the methodology outlined by Deaton (1997).[8] In addition, medians of urban and rural areas were deflated by the Consumer Price Index for Industrial Workers (CPIIW) and by the CPIAL, in order to allow comparability across sectors and over time. Real per capita income increased by less than 20% over the period considered, at the rate of 1% per year. In urban areas, per capita income growth was nearly double that of rural areas.

Two limitations of growth rates calculated from NSSO survey data need to be pointed out. First, data from the National Accounts report a much larger increase in per capita income over the same period. According to National Accounts data reported by the DES (2005), per capita income more than doubled in Andhra Pradesh between 1987-88 and 2004-05. Given the size of the difference between the survey data and the National Accounts data, it is likely that the survey data underestimate per capita expenditure. Possible reasons for this underestimation are the introduction of new consumption goods not reported in the questionnaires, and the increasing rate of refusal in

[8] Household observations are sorted in ascending order of the value of per capita expenditure. The running sum of the individual sampling weights (household sampling weights times household size) is calculated. The per capita expenditure of the household occupying the central position of this running sum is then selected as the weighted median of the sample.

participating in the interview or in providing accurate answers as the population becomes richer or more urbanised (Bhalla, 2002). Second, figures were adjusted using the CPIAL and the CPIIW. These indices have a number of shortcomings, which are discussed at length in Deaton and Tarozzi (2000) and Deaton and Dreze (2002). The three main shortcomings of official price indices are: they are based on outdated household expenditure shares; they are not fully representative of all areas of the country; and they rely on the Laspeyres formula that tends to understate inflation over time.

Table 2.6 Median per capita expenditure and inequality indices (1987-2005)

	1987-88	1993-94	1999-00	2004-05
Median per capita expenditure[a]				
All sample	44.1	48.0	54.1	51.3
Rural	41.6	44.4	48.4	47.1
Urban	55.9	62.5	77.8	68.8
Gini coefficient				
All sample	0.284	0.262	0.294	0.256
Rural	0.275	0.248	0.259	0.238
Urban	0.295	0.276	0.313	0.259
Standard deviation of the logarithm of per capita expenditure				
All sample	0.495	0.458	0.495	0.451
Rural	0.480	0.437	0.443	0.417
Urban	0.517	0.483	0.532	0.459

Notes: [a]median per capita expenditure is in thousands of 2004-05 Rupees.
Source: calculated from NSSO data.

Table 2.6 also shows the values of two inequality indices: the Gini coefficient and the standard deviation of the logarithm of per capita expenditure. Both indices were calculated taking into account the survey sampling weights.[9] Income inequality is not high in Andhra Pradesh, where the Gini coefficient is between 0.25 and 0.30. Data of

---

[9] The weighted Gini was calculated following the methodology outlined in Deaton (1997), while the standard deviation of the logarithm of per capita expenditure was calculated using the following formula:

$$\sigma_w^2 = \frac{N}{N-1}\sum_i^N w_i\left(\ln x_i - \ln \bar{x}_w\right)^2$$

where N is the number of households, and $\ln \bar{x}_w = \sum_i w_i \ln x_i$ is the weighted mean where the weights are defined by $w_i = \frac{weight_i}{\sum_i weight_i}$, with $\sum_i w_i = 1$.

The weights ($weight_i$) are the probability sampling weights provided by NSSO.

the *World Development Indicators* (World Bank, 1990-2005), report Gini coefficients in a range from a minimum of 0.24 in Hungary (1997) to a maximum of 0.71 in Namibia (1993). Countries with inequality indices similar to those of Andhra Pradesh are Bangladesh (0.32 in 2000), and Pakistan (0.33 in 1999), while the Gini coefficient for the whole of India in 1999-00 was 0.33.

Expenditure inequality has decreased in Andhra Pradesh in both rural and urban areas, but more so in rural areas. Notice that the survey data of 1999-00 report a surge in inequality. Some authors, Deaton and Dreze (2002) and Sen and Himanshu (2004), have interpreted this surge as part of a pattern of increasing inequality in India between rural and urban areas, and within urban areas. The latest round of 2004-05 however reports a Gini coefficient that is lower than the one calculated for the survey of 1993-94. This suggests that the surge in inequality in 1999-00 might be a consequence of the changes made to the survey questionnaire. Leaving aside the 1999-00 data, a pattern of decreasing inequality emerges, which is more pronounced in rural areas.

2.4.5 Summary of the descriptive evidence

The descriptive evidence presented in this section on the process of agricultural transformation in Andhra Pradesh can be summarised in the following way. During the last 50 years, agricultural output as a share of total GDP has decreased considerably. Employment in the agricultural sector however, has only slightly decreased, and there has been an increase in the share of agricultural labourers. Technological progress in agriculture has resulted in a significant increase in output per hectare beginning in the mid 1970s. The growth in productivity resulted in substantial increases in yields of rice and cotton. Despite a substantial increase in per capita income, the share of food demand in total consumption has only slightly decreased and it remains at a very high level even today. Inequality in the distribution of expenditure has decreased over the period, and this has potentially contributed to the high levels of food demand in the state.

2.5 Econometric methods

This section discusses the econometric methods used for the estimation of expenditure elasticities of food consumption, and for the simulation of the effect of a change in income distribution on food consumption. This section proceeds in three parts. The first discusses the theoretical evolution over time in the econometric estimation of Engel curves. The second illustrates the semiparametric methods employed to estimate food Engel curves and to calculate expenditure elasticities. The third presents the strategy adopted to simulate changes in income distribution.

2.5.1 Estimation of food Engel curves

The Engel curve describes a functional relationship between consumption of a given commodity and income or household expenditure. Using *q* for the quantity consumed of a given commodity and *y* for income of the *i* household, the function can be written as:

$$q_i = f(y_i) \tag{2.26}$$

The function was named after Ernst Engel, a German economist who in 19th century studied the food consumption of the Belgian working class using a nonparametric method known today as a regressogram (Hardle, 1990). Engel plotted food consumption on household income, dividing the horizontal axis in intervals of equal size, and calculating the average food consumption of households with income falling in the same interval. This method generated a discontinuous step function which showed that consumption of food increases less than proportionally with income.

After the original work of Engel, many other empirical observations of food expenditure patterns from household budget studies have found that food consumption increases less than proportionally with income, and that food consumption as a share of total consumption decreases as income increases. This observation has been found so frequently in empirical studies that it has acquired the status of law in the economic literature, and it is known as the Engel's law. For a review of recent estimation of food Engel curves in more than 100 countries see the work of Seale et al. (2003).

Since Engel's original work, economists have estimated innumerable Engel functions for the most different purposes, normally employing parametric methods. Prais and Houthakker (1971) wrote a comprehensive review of these experiments, and performed their own estimations of the functional forms most commonly used up to their times. These forms were the linear, the hyperbolic, the semilogarithmic, the double logarithmic, and the logarithmic reciprocal. All these forms were showed to have some advantages over the other forms for some of the goods or for part of the range of the relationship. Prais and Houthakker (1971) concluded that the widely used double logarithmic and the semilogarithmic forms performed better than the others in terms of goodness of fit.

Prais and Houthakker (1971) however were also concerned with the functional relationship being consistent with observed consumer behaviour. Engel curves should be able to represent luxuries (commodities whose consumption increases more than proportionally with income), necessities (whose consumption increases less than proportionally with income), and inferior goods (commodities whose consumption decrease as income increases). In addition, Engel curves should allow the same commodity to be a luxury for the poor but a necessity for the rich. Based on these criteria, they concluded that the double logarithmic form was too restrictive, since it implied a constant income elasticity, and opted for the semilogarithmic form as the best functional form available. This form is able to represent necessities, luxuries, and inferior goods, and allows the income elasticity to vary with income levels:

$$q_i = a + b \ln y_i \tag{2.27}$$

The choice of the functional form of the Engel curve has often been based on practical criteria of goodness of fit, sometimes disregarding basic principles of demand theory. One of these principles in particular, the adding-up condition, is violated by all the form listed above including the semilogarithmic. Adding-up requires that consumers do not spend more than their income. This principle places some restrictions on the demand elasticities of each goods, known as Engel's and Cournot's equations. Simply put, these equations state that changes in income and prices determine changes in the composition of the budget constraint but leave its value unchanged. One functional form that satisfies adding-up, and that is able to represent closely consumer behaviour was originally proposed by Working (1943), elaborated by Leser (1963), and widely used after Deaton and Muellbauer (1980b). This form is known as the Working-Leser, and relates the commodity budget shares to the logarithm of per capita expenditure:

$$w_i = a_i + b_i \ln y \tag{2.28}$$

This form satisfies the adding-up condition provided that the sum of the parameters *a* estimated over all commodities in the household budget is equal to one, and that the sum of the parameters *b* is equal to zero. It allows for luxuries, necessities and inferior goods, and for elasticities to vary with income. Finally, the form is linear in the logarithm of expenditure, and is easily estimated by ordinary least square (OLS) equation by equation, with the adding-up restrictions being automatically satisfied.

One disadvantage of the Working-Leser form is that necessities and luxuries are represented by different curves, which means that the same good, for example food, cannot be a luxury for some household and a necessity for another. Each commodity can only have an elasticity which is either above or below one. A second shortcoming of this form is that though it allows for varying elasticities, these are bound to vary always in the same direction. Elasticities can only decrease as income increase. If the good is a necessity, the elasticity decreases down to minus infinity as income increases, while if the good is a luxury the elasticity decreases down to $1+b$. Finally, the Working-Leser form postulates that the change in elasticity is linear with income, which is not necessarily true.

More recent research on Engel curves has focused on the estimation of the Working-Leser equation using polynomials in the expenditure term (by adding quadratic or cubic terms), and nonparametric or semiparametric methods. Unlike the standard parametric Working-Leser form, these methods allow expenditure elasticities to vary in any direction for the same good over the entire range of household expenditure. Most studies have found evidence of non-linear budget shares and elasticities for a large number of goods.

In the case of the estimation of food Engel curves, a difference appears to emerge in the empirical literature between studies conducted with data from developed countries, which find a linear relationship between the food share and the logarithm of household expenditure (Atkinson et al., 1990, Banks et al., 1997), and studies of developing countries, which find non-linear elasticities (Bhalotra and Attfield, 1998, Gong et al., 2005, Kedir and Girma, 2007). Food expenditure elasticities increased with income at very low income levels in Pakistan (Bhalotra and Attfield, 1998), were higher than one for very poor households in Ethiopia (Kedir and Girma, 2007), and decreased at different rates at different levels of household expenditure in China (Gong et al., 2005).

2.5.2 Semiparametric estimation of Engel curves

This section illustrates the econometric methods employed by the present study for estimating food elasticities and addresses the following issues: semiparametric estimation of Engel curves; measurement error and endogeneity in the estimation of Engel curves; specification test of functional forms; and calculation of expenditure elasticities from nonparametric curves.

Semiparametric estimation of Engel curves

The present study estimates food Engel curves both parametrically and semiparametrically. The parametric model is a polynomial in per capita expenditure up to the third grade:

$$w = a + b_{i1} \ln y_i + b_{i1} \ln^2 y + b_{i1} \ln^3 y + \sum_j c_j X_{ij} + e_i \tag{2.29}$$

This model is estimated in its linear, quadratic, and cubic form of per capita expenditure. The estimation is carried out using OLS corrected for the use of sampling weights. In addition to the logarithm of per capita expenditure, the model includes a large set of control variables (*X*). These include a set of demographic variables that describe the demographic composition of the household, which are normally found to explain a large part of the variation in food expenditure across households. Following Deaton (1997), the demographic variables include the age of the head of household, the logarithm of household size, and the ratios of six demographic groups over household size disaggregated by gender. The demographic groups consists of three child-age groups (zero to four, five to nine, and ten to 17), adults (individuals aged 18 to 59) and elderly (people aged 60 and above).

Other variables in *X* control for the social class (scheduled castes and scheduled tribes) to which the household belongs, and its main occupation (whether farmers, agricultural labourers, or other). These variables capture different dietary habits and requirements. For example, members of unscheduled caste are largely vegetarian, while farmers require more energy food for more demanding physical work, and a different type of clothing compared to urban workers.

The model also includes a number of geographic variables. These consist of a dummy variable for urban areas, and four dummy variables for the regional subdivision of the state used by the NSSO. The NSSO subdivides each Indian state in macro regions by adjoining administrative districts that are homogeneous in terms of climate and socioeconomic characteristics.[10] As in the case of social and occupational variables, location variables capture different expenditure habits that vary with climate or custom.

The semiparametric model combines the advantages of parametric and nonparametric estimation. Engel curves may be non-linear, and a nonparametric smoother has the advantage of not imposing any specific functional form, like the logarithmic or the quadratic, to the data. In a nonparametric model, the data define the functional form. On the other hand, a nonparametric smoother cannot control for determinants of food consumption other than per capita expenditure, which is the advantage of a parametric model. The semiparametric model expresses the relationship between food consumption and per capita expenditure nonparametrically, controlling at the same time for other determinants of food consumption. The semiparametric model used has the form:

$$w = a + m_i(\ln y) + \sum_j c_j X_{ij} + e_i \qquad (2.30)$$

where m_i is the non parametric smoother calculated over per capita expenditure, and the X variables are those already included in the parametric model. In other words, this model allows the estimation of a different coefficient of per capita expenditure (m_i) for each household in the sample, and it estimates coefficients (c_j) of the other explanatory variables that are common to all households in the sample.

[10] The four macro regions used by NSSO in Andhra Pradesh group the following districts together: 1) Srikakulam, Vizianagaram, Vishakhapatnam, East Godavari, West Godavari, Krishna, Guntur, Prakasam, Nellore; 2) Chittor, Cuddapah; 3) Anantapur, Kurnool; 4) Mahaboobnagar, Rangareddy, Hyderabad, Medak, Nizamabad, Adilabad, Karimnagar, Warangal, Khamman, Nalgonda.

Equation (2.30) is estimated using the method of first differencing, which was chosen for its simplicity and ease of computation. First differencing is clearly described in Johnston and Dinardo (1997), and discussed in detail by Yatchew (2003). The model aims to remove from the dependent variable the portion that is determined by explanatory variables whose coefficients are estimated parametrically. Once the parametric component has been removed, the dependent variable can be regressed over per capita expenditure using a nonparametric method of choice. The nonparametric function estimated is:

$$w_i - X_i \hat{b}_{diff} = m_i(y_i) + e_i \tag{2.31}$$

where the food share of equation (2.30) is replaced by the food share minus its parametric component. First differencing finds estimates of the coefficients $\hat{b}_{diff}$ in three simple steps. First, the observations in the sample are sorted in ascending order of the logarithm of per capita expenditure. Second, the following first differences are computed for all observations in the sample:

$$dw_i = w_i - w_{i-1} \quad \text{and} \quad dx_i = x_i - x_{i-1} \tag{2.32}$$

Third, the differenced variables are used to run the OLS regression:

$$dw_i = b_{diff} dX_i + u_i \tag{2.33}$$

The estimated coefficients obtained from this regression are then multiplied by the original values of the explanatory variables, and their product is subtracted from the dependent variable as in equation (2.31). The advantage of this method over the more commonly used one developed by Robinson (1988), is its ease of computation. While Robinson's method requires the estimation of a different nonparametric regression for each explanatory variable, the method of first differencing only requires the estimation of one nonparametric regression (2.31).

Values of m_i in (2.31) are obtained using a nonparametric smoother. The present analysis uses the Cleveland's Locally Weighted Scatterplot Smoothing (LOWESS), which was preferred to the more common kernel estimators. A full description of the LOWESS algorithm can be found in Hardle (1990) and Fan and Gijbels (1996). Perhaps the simplest way to think of this method is that of performing a polynomial regression at each data point of the dependent variable giving lower weight to observations that are more distant from the point and that produce large residuals. Putting together the lines estimated at each point generates the line estimated by LOWESS.

This method is computationally intensive but has a series of advantages over kernel smoothers. The latter are potentially affected by two sources of bias, a curvature bias, and a boundary problem bias, which are discussed by Deaton (1997) and Cameron and Trivedi (2005). The curvature bias arises when the function to estimate is concave. Kernel estimators evaluate the function at each point applying decreasing weights to more distant values. Suppose that the true function is concave down and that the kernel estimator uses at each point two other equally spaced points to evaluate the function. By Jensen inequality, the evaluation will be biased downwards.[11] This bias increases the slope of the true function. The second source of bias arises from the fact that at the two end points of the distribution of the explanatory variable, the kernel smoother uses values that are either mostly decreasing or increasing depending on the slope of the function at the endpoint. If the dependent variable is negatively sloped, as is likely to be the case for the food share, the function will be underestimated to the left of the distribution and overestimated to its right. This bias results in a flattening of the curve at the extremes points.

The LOWESS estimator is not affected by either the curvature bias or the boundary problem bias. In addition, it uses a robustified method that makes it less sensitive to outliers compared to other nonparametric smoothers. This is a great advantage in the use of survey data that are affected by considerable measurement error.

---

[11] The Jensen inequality states that for a concave down function: $\frac{\sum_i^n f(x_i)}{n} \leq f\left(\frac{\sum_i^n x_i}{n}\right)$

Measurement error and endogeneity

Expenditure data are often measured with error, and measurement error generates a bias in the parameter estimates. In this particular case, the bias arises because the error in the measurement of per capita expenditure (the independent variable) is correlated with the error in the measurement of the budget food share (the dependent variable). The sign of this bias cannot be determined a priori, because measurement error can appear either in the numerator or in the denominator of the food share or in both.

This problem can be addressed by using instrumental variables. In order to do so, one or more instruments are needed that are correlated with per capita expenditure, but uncorrelated with the measurement error. Once one or more plausible instruments have been found, different instrumental variables approaches can be followed. The approach followed here is the estimation of an augmented regression of the Hausman type. In this approach the variable measured with error (per capita expenditure) is regressed using OLS on the instruments. After estimation, the estimated residual are calculated and included in the semiparametric model (see Gong et al., 2005, Yatchew, 2003). The parameter estimate of the estimated residual tests for the endogeneity of per capita expenditure, and provides a correction for the measurement error bias in the estimated regression.

Specification test

Parametric estimation can provide a good representation of food Engel curves, and a statistical test is needed to assess whether the semiparametric estimation differs from parametric functional forms. One way of comparing the shapes of parametric and semiparametric Engel curves relies on the estimation of confidence intervals. However, the computation of confidence interval for the LOWESS smoother is not straightforward. In principle, confidence intervals could be obtained using bootstrapping methods, but LOWESS is already computationally intensive, and the bootstrapping of confidence intervals would require a large amount of computer time.

The present study uses a specification test in order to compare the parametric and semiparametric estimation of the Engel curves. The semiparametric technique of first differences offers a straightforward method of testing among alternative functional forms. This test is described by Yatchew (2003) and has the following form:

$$V = \frac{\sqrt{n}\left(\sigma_{res}^2 - \sigma_{diff}^2\right)}{\sigma_{diff}^2} \tag{2.34}$$

where n is the number of observations, and the terms σ_{res}^2 and σ_{diff}^2 are the residual variances of the parametric and the semiparametric model respectively:

$$\sigma_{res}^2 = \frac{1}{n}\sum_i^n e_i^2 \quad \text{and} \quad \sigma_{diff}^2 = \frac{1}{n}\sum_i^n \left(dw_i - dX_i b\right)^2 \tag{2.35}$$

This test is normally distributed, and one-sided.

Calculation of elasticities

The ultimate purpose of the estimating Engel curves is to calculate expenditure elasticities. Computation of expenditure elasticities is a relatively simple matter in the case of parametric models, but is rather more complicated in the case of semiparametric models, and the topic requires some discussion.

The parametric elasticities for the linear, quadratic, and cubic Working-Leser forms are:

$$\eta_l = 1 + \frac{b_1}{w} \tag{2.36}$$

$$\eta_q = 1 + \frac{b_1 + 2b_2 \ln y}{w}$$

$$\eta_c = 1 + \frac{b_1 + 2b_2 \ln y + 3b_3 \ln^2 y}{w}$$

where the b_s are the coefficient estimates of higher order terms of the logarithm of per capita expenditure, and w is the share of food expenditure over total expenditure.

Elasticities are normally calculated using the sample mean of households' food shares, but they can also be calculated for food share values corresponding to different percentiles of the expenditure distribution.

A nonparametric equivalent to the parametric food elasticity is:

$$\eta_i = 1 + \frac{\hat{m}_i'(x_i)}{\hat{m}_i(x_i)} \quad (2.37)$$

where the denominator is the food share predicted by the semiparametric model, and the numerator is the derivative of the nonparametric function estimated by LOWESS. There are several methods to calculate the derivative in the numerator of (2.37). The method chosen for this study is the finite difference approach described in Cameron and Trivedi (2005):

$$\hat{m}_i' = (\hat{m}_i - \hat{m}_{i-1}) / (\ln x_i - \ln x_{i-1}) \quad (2.38)$$

This derivative can be described as a difference quotient, and is calculated at each data point. It is the ratio between the change in the fitted value of the dependent variable and the change in the explanatory variable. Before taking the first differences the data need to be sorted in ascending order of the explanatory variable (the logarithm of per capita expenditure). It is also advisable, in order to avoid very large values of the derivative, to produce oversmoothed nonparametric predictions of food shares.

Notice that both the derivatives and the elasticities are indexed by *i*, which means that a different value of the derivative and of the elasticity is calculated for every household. There are several ways in which these household elasticities can be summarised in order to compare them with those estimated by parametric models. Two approaches were followed here. The first consists of calculating the weighted mean of all the elasticities in the sample, where the weights are the probability sampling weights. This estimator provides the average household response to changes in expenditure. A second approach consists of calculating a market elasticity, which is the response of the sample data to an increase in per capita expenditure of 1%. This is simply the change in food consumption for the entire sample resulting from a change in per capita expenditure by 1%. The change is simulated using observed food and total expenditure data, and the estimated individual elasticities. The computation of this elasticity is equivalent to calculating the average elasticity of the sample while applying the shares of household expenditures over total expenditure in the sample as weights. This second elasticity provides an estimate of the response of the market to a change in expenditure.

2.5.3 Simulations of changes in income distribution

As discussed in Section 2.3.6, the aggregation problem poses serious difficulties to the modelling of food demand as a function of total expenditure alone. If the food Engel curve is non linear, an income transfer from a rich household to a poor one will increase the aggregate consumption of food. One way of tackling this problem is formulating an analytical model that relates food consumption to both total expenditure and to the inequality of the expenditure distribution. The present study follows a simpler approach. Changes in the expenditure distribution will be simulated, and market elasticity and food consumption will be calculated at different levels of expenditure inequality using the household elasticities estimated semiparametrically.

The simulations performed consist of mean-preserving changes in the expenditure distribution of one of the survey rounds. In order to achieve higher and lower inequality, while preserving average and total expenditure in the sample, money is transferred from one household to another. For example, in order to increase inequality, a sum is taken from a poor household and given to a rich household. Households are selected randomly for this transfer, after assigning a probability of selection that is proportional to household per capita expenditure.

In practice the simulation of transfers operates in the following way. Household per capita expenditure is expressed as a share of total expenditure in the sample:

$$\pi = \frac{x_i}{\sum_i x_i} \tag{2.39}$$

These shares are then sorted in ascending order and the cumulated sum is calculated, so that the value of the running sum for the richest household in the sample is one. Two random numbers between zero and one are then drawn from a uniform distribution. The first number identifies the beneficiary household, while the second number identifies the household that is affected by the transfer. A beneficiary is selected when the random number is between the household probability of selection and the probability of selection of the household that is ranked one position below in terms of per capita expenditure. At each draw, households are reordered in terms of per capita expenditure, the selection probabilities are recalculated and new random numbers are drawn.

The simulation performed more than 200,000 transfers, each of a value of 1,000 Rupees. The sum of all transfers corresponds to the redistribution of 30% of total household expenditure in the sample. Inequality indices, market elasticities, food shares, and food consumption were calculated at each 1% redistribution step in one direction or the other.

2.6 Empirical results

This section assesses the operation of Engel's law in Andhra Pradesh using four large nationally representative household surveys conducted by the NSSO between 1987 and 2005. The aim is to quantify the responsiveness of food consumption to changes in living standards and income distribution. This will shed some light on the slow pace of agricultural transformation in the state.

This section proceeds in three parts. The first presents results of the estimation of semiparametric food shares and elasticities. The second repeats the same analysis for seven food categories, in order to explain some puzzling characteristics of the estimated Engel curves. The third simulates changes in income distribution in order to assess the extent to which food consumption changes as income inequality increases or decreases.

2.6.1 Estimated food shares and elasticities

The results of the estimation of models (2.29) and (2.32) for the four survey rounds are shown in Tables 2.8 and 2.9. Model (2.29) estimates the following parametric Engel curves: linear, quadratic, and cubic. Table 2.8 only shows the results of the cubic form because it fitted the data better than the other forms. The dependent variable is the household food share, while the explanatory variables consist of the control variables described in Section 2.5.2. Model (2.32) runs a regression of the food share on the explanatory variables both expressed in first differences. Per capita expenditure is not included in this model because it will be estimated nonparametrically after the removal of its parametric component following the methodology outlined in Section 2.5.2. Observations with food shares equal to zero or one were dropped, because these values were considered non-admissible. Observations with per capita expenditure values four standard deviations above the mean were also dropped in order to remove outliers. Both regressions were run correcting the standard errors for the probability sampling weights.

The significance and the sign of each estimated coefficient will not be discussed here, as the main objective of this exercise is to assess the relationship between the food share and per capita expenditure. Overall, the estimated coefficients have the expected signs, and the statistical significance is high, particularly in the case of the cubic model. Demographic and geographic variables explain the largest part of the variance in household food expenditure, while occupation and social class play a minor role. The share of the total variance explained by the cubic model is rather high for a cross-sectional study, with R-squares ranging between 0.35 and 0.51. Statistical significance of the estimated parameters and the R-squares are much lower in the case of the model estimated in first differences, as can be expected.

As explained in Section 2.5.2, errors in the measurement of households' expenditure may generate a spurious correlation between the dependent variable and the explanatory variables. In order to test, and at the same time correct for this potential bias, a regression of per capita expenditure on a series of instruments was run, and the estimated residuals were included among the explanatory variables of models (2.29) and (2.32).The results of this regression are presented in Table 2.7. The variables included in the instrumenting regression consist of the demographic composition of the households (number of adults and children by sex, and the age of the reported head of the household), the education of the male and female head of household, the size of land owned, household belonging to a scheduled caste or tribe, and location variables for urban areas and macro-regions. The residuals from these regressions were calculated and then included in the estimation of the food expenditure models. The parameter estimates of the residuals provide a test for endogeneity of per capita expenditure, and at the same time provide a correction for the measurement error bias in the estimation of the food share models (Yatchew, 2003).

Table 2.7 Hausman instrumenting regression for the food share model

Variables	1987-88	1993-94	1999-00	2004-05
Education of male head of household	0.116***	0.098***	0.112***	0.093***
	(0.008)	(0.006)	(0.006)	(0.006)
Education of female head of household	0.081***	0.071***	0.065***	0.062***
	(0.010)	(0.009)	(0.008)	(0.007)
Land owned (acres)	0.027***	0.021***	0.033***	0.018***
	(0.003)	(0.003)	(0.005)	(0.005)
Ratio of male children (0 to 4)	-0.025	-0.097	-0.104	-0.122*
	(0.066)	(0.070)	(0.073)	(0.076)
Ratio of male children (5 to 9)	0.183***	0.057	0.146**	-0.004
	(0.054)	(0.059)	(0.059)	(0.067)
Ratio of male children (10 to 14)	0.507***	0.313***	0.425***	0.235**
	(0.064)	(0.065)	(0.068)	(0.076)
Ratio of female children (0 to 4)	-0.022	-0.206**	-0.067	-0.270***
	(0.065)	(0.068)	(0.077)	(0.082)
Ratio of female children (5 to 9)	0.098*	0.106*	0.119**	0.014
	(0.061)	(0.061)	(0.061)	(0.066)
Ratio of female children (10 to 14)	0.405***	0.214**	0.334***	0.129*
	(0.066)	(0.072)	(0.071)	(0.071)
Ratio of male adults	0.615***	0.463***	0.533***	0.358***
	(0.051)	(0.051)	(0.051)	(0.054)
Ratio of female adults	0.278***	0.239***	0.364***	0.227***
	(0.050)	(0.047)	(0.049)	(0.049)
Ratio of male elderly	0.227**	0.260***	0.174**	0.107
	(0.074)	(0.074)	(0.069)	(0.076)
Household size (log)	-0.341***	-0.281***	-0.323***	-0.301***
	(0.013)	(0.014)	(0.015)	(0.016)
Age of head of household	0.002***	0.002**	0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Scheduled caste	-0.058***	-0.076***	-0.094***	-0.107***
	(0.014)	(0.014)	(0.014)	(0.015)
Scheduled tribe	-0.169***	-0.116***	-0.065**	-0.185***
	(0.022)	(0.023)	(0.023)	(0.024)
Macro-region two	-0.124***	-0.120***	-0.093***	-0.148***
	(0.015)	(0.014)	(0.015)	(0.015)
Macro-region three	0.069***	0.062***	0.049**	-0.009
	(0.017)	(0.017)	(0.019)	(0.018)
Urban area	0.037**	0.051***	0.136***	0.209***
	(0.015)	(0.013)	(0.014)	(0.014)
Constant	9.577***	10.196***	10.559***	10.775***
	(0.052)	(0.054)	(0.056)	(0.057)
Observations	9111	8164	8765	8106
R-square	0.351	0.355	0.444	0.434

Notes: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%; estimated standard errors in parentheses.

The estimated residuals were found significant in all regressions, except the 1999-00 survey in the case of the cubic model, and the 1999-00 and 2004-05 surveys in the case of the model in first differences. The size of the estimated coefficient on the residuals is always large and negative, pointing to the presence of measurement error and of a spurious correlation between the food share and total per capita expenditure.

Table 2.8 Food share regression: parametric cubic form

Variables	1987-88	1993-94	1999-00	2004-05
Logarithm of per capita expenditure	12.153***	20.482***	9.160***	18.687***
	(1.597)	(2.871)	(1.804)	(2.643)
Logarithm^2 of per capita expenditure	-1.230***	-1.929***	-0.809***	-1.688***
	(0.164)	(0.275)	(0.162)	(0.236)
Logarithm^3 of per capita expenditure	0.041***	0.060***	0.023***	0.050***
	(0.006)	(0.009)	(0.005)	(0.007)
Male children (0 to 4)	-0.023	-0.033*	0.038**	-0.051**
	(0.019)	(0.020)	(0.019)	(0.023)
Male children (5 to 9)	0.039**	0.003	0.066***	-0.016
	(0.017)	(0.016)	(0.016)	(0.021)
Male children (10 to 14)	0.070***	0.046***	0.086***	0.023
	(0.020)	(0.018)	(0.017)	(0.022)
Female children (0 to 4)	-0.017	-0.052**	0.029	-0.036
	(0.019)	(0.020)	(0.019)	(0.023)
Female children (5 to 9)	0.030*	-0.026	0.027	0.001
	(0.016)	(0.017)	(0.017)	(0.022)
Female children (10 to 14)	0.036*	0.003	0.062***	0.005
	(0.018)	(0.019)	(0.018)	(0.019)
Male adults	0.001	0.009	0.024*	0.023
	(0.015)	(0.015)	(0.014)	(0.014)
Female adults	0.014	0.030**	0.025*	-0.006
	(0.014)	(0.013)	(0.013)	(0.012)
Male elderly	-0.057**	-0.004	0.008	-0.037**
	(0.023)	(0.021)	(0.020)	(0.018)
Logarithm of household size	-0.019***	-0.011**	-0.051***	-0.022***
	(0.005)	(0.005)	(0.005)	(0.007)
Age of head of household	0.001***	0.000**	0.001***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Scheduled caste	-0.005	0.003	0.006*	-0.000
	(0.004)	(0.004)	(0.004)	(0.004)
Scheduled tribe	-0.002	0.011*	0.012**	0.001
	(0.007)	(0.006)	(0.006)	(0.007)
Agricultural labour household	0.003	0.013***	0.009**	0.005
	(0.004)	(0.004)	(0.004)	(0.004)
Farmer household	0.016***	0.031***	0.026***	0.017***
	(0.004)	(0.004)	(0.004)	(0.004)
NSSO region no. 2	-0.025***	-0.021***	-0.048***	-0.017***
	(0.003)	(0.003)	(0.003)	(0.003)
NSSO region no. 3	-0.008	0.007	-0.037***	-0.026***
	(0.005)	(0.005)	(0.005)	(0.005)
NSSO region no. 4	-0.021***	-0.004	-0.025***	-0.011**
	(0.006)	(0.005)	(0.005)	(0.006)
Urban area	-0.019***	-0.018***	-0.021***	-0.082***
	(0.004)	(0.004)	(0.004)	(0.004)
Estimated residual	-0.081***	-0.070***	0.000	-0.026***
	(0.009)	(0.010)	(0.009)	(0.010)
Constant	-38.949***	-71.238***	-33.237***	-67.618***
	(5.168)	(9.985)	(6.670)	(9.838)
Observations	9111	8164	8765	8106
R-square	0.356	0.397	0.500	0.518

Notes: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%; estimated standard errors in parentheses.

Table 2.9 Food share regression: first differences

Variables	1987-88	1993-94	1999-00	2004-05
Male children (0 to 4)	-0.019	-0.010	-0.011	-0.044*
	(0.019)	(0.020)	(0.020)	(0.024)
Male children (5 to 9)	0.056***	0.011	0.046***	-0.020
	(0.016)	(0.016)	(0.016)	(0.019)
Male children (10 to 14)	0.059***	0.053***	0.061***	0.003
	(0.020)	(0.018)	(0.016)	(0.023)
Female children (0 to 4)	-0.005	-0.036*	0.009	-0.029
	(0.020)	(0.020)	(0.020)	(0.024)
Female children (5 to 9)	0.044**	-0.018	0.016	-0.001
	(0.017)	(0.017)	(0.020)	(0.021)
Female children (10 to 14)	0.044**	0.027	0.040**	0.022
	(0.019)	(0.021)	(0.018)	(0.020)
Male adults	0.007	0.017	0.006	0.037**
	(0.015)	(0.015)	(0.015)	(0.015)
Female adults	0.019	0.052***	0.024*	-0.001
	(0.014)	(0.013)	(0.013)	(0.013)
Male elderly	-0.030	0.020	-0.009	-0.020
	(0.021)	(0.022)	(0.019)	(0.018)
Logarithm of household size	-0.030***	-0.012**	-0.049***	-0.018***
	(0.005)	(0.005)	(0.005)	(0.006)
Age of head of household	0.001***	0.000*	0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Scheduled caste	-0.007*	0.009**	0.005	0.002
	(0.004)	(0.004)	(0.004)	(0.005)
Scheduled tribe	-0.013*	0.011*	0.011*	-0.006
	(0.007)	(0.006)	(0.005)	(0.007)
Agricultural labour household	0.000	0.005	0.010**	0.007
	(0.004)	(0.004)	(0.004)	(0.005)
Farmer household	0.021***	0.031***	0.024***	0.016***
	(0.004)	(0.004)	(0.004)	(0.004)
NSSO region no. 2	-0.026***	-0.018***	-0.046***	-0.016***
	(0.003)	(0.003)	(0.003)	(0.003)
NSSO region no. 3	-0.000	0.013**	-0.040***	-0.022***
	(0.005)	(0.005)	(0.005)	(0.005)
NSSO region no. 4	-0.017***	-0.001	-0.025***	-0.004
	(0.005)	(0.005)	(0.005)	(0.006)
Urban area	-0.019***	-0.023***	-0.027***	-0.088***
	(0.004)	(0.004)	(0.004)	(0.004)
Estimated residual	-0.056***	-0.069***	0.001	-0.009
	(0.010)	(0.010)	(0.009)	(0.011)
Observations	9110	8163	8764	8105
R-square	0.085	0.092	0.152	0.192

Notes: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%; estimated standard errors in parentheses.

The three parametric forms estimated were tested against the semiparametric form using the specification test described in Section 2.5.2. Table 2.10 reports the value of the test statistic together with the probability value for a one-sided test. Stars represent different significance levels for the rejection of the hypothesis of the equality between the models. With the exception of the 1999-00 survey, the parametric specifications are strongly rejected by the tests. There is no clear superiority of the quadratic form over the linear and vice versa. The cubic form performs better than the other forms, with lower statistics in all surveys, and could not be rejected in the 1993-94 and 1999-00 surveys. Even the cubic form however was strongly rejected in the 1987-88 and the 2004-05 surveys, suggesting that the use of a nonparametric specification is preferable to its parametric approximations.

Table 2.10 Specification test of parametric Engel curves

	1987-88	1993-94	1999-00	2004-05
Linear	5.48***	3.96***	1.58*	5.46***
	(0.000)	(0.000)	(0.056)	(0.000)
Quadratic	5.02***	4.12***	1.34*	5.57***
	(0.000)	(0.000)	(0.091)	(0.000)
Cubic	4.49***	1.12	0.75	3.68***
	(0.000)	(0.131)	(0.225)	(0.000)

Notes: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%; probability values in parentheses.

The parametric and semiparametric Engel curves were used to calculate food expenditure elasticities following the formulas described in Section 2.5.2. Table 2.11 presents the set of estimated elasticities for the four survey rounds. Two versions of the semiparametric elasticities are shown: average elasticities, which represent the average sample response to changes in income, and market elasticities, which assess the response of total demand to a change in income.

Semiparametric elasticities are larger than parametric ones and are very high, ranging from just below 0.8 to nearly 0.9. These elasticities are very high even for a low income country. For example, the study of Seale et al. (2003), using data of the early 1990s, finds an average food elasticity of 0.73 for low income countries, while the largest elasticity found is the one of Tanzania, equal to 0.8. Food elasticities are decreasing over time, which can be related to the increase in living standards. This pattern is interrupted by the elasticities of the 1999-00 survey, which stand out for their very low values compared to estimates for neighbouring years. This latter result however is probably due the changes introduced in the survey questionnaire for this round.

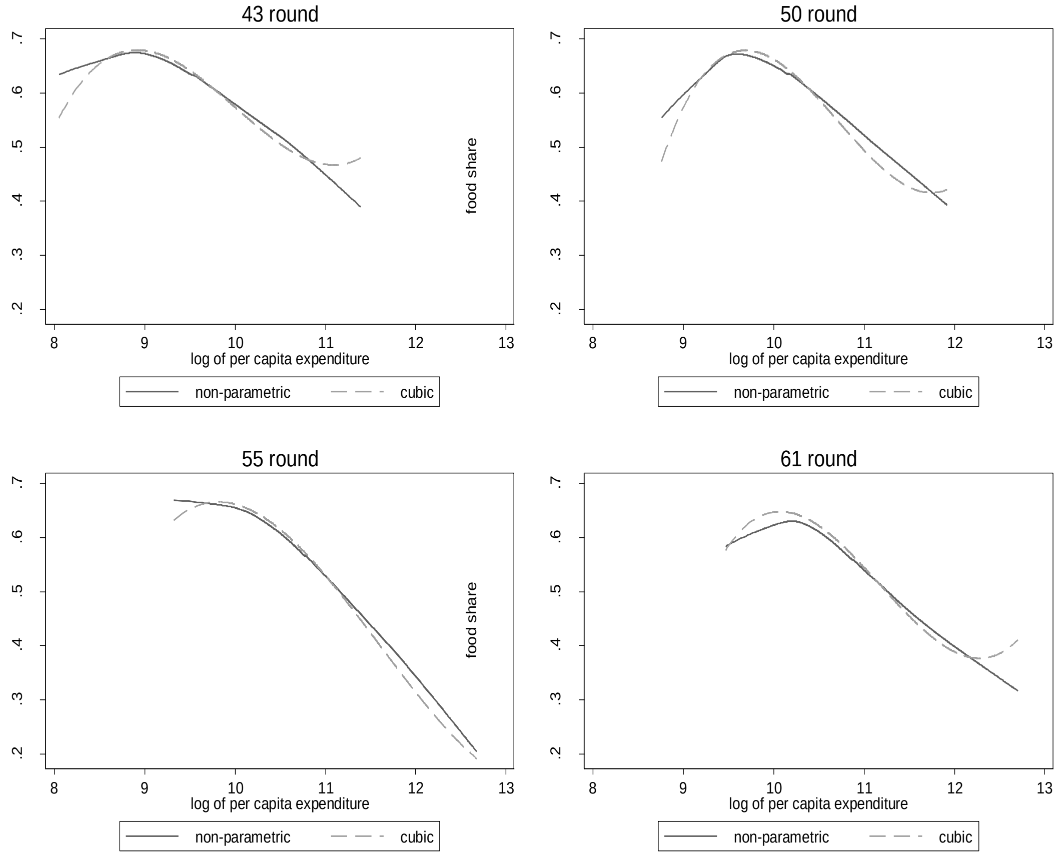
Table 2.11 Parametric and semiparametric expenditure elasticities of food

Model	1987-88	1993-94	1999-00	2004-05
Linear	0.819	0.766	0.654	0.717
Quadratic	0.847	0.825	0.696	0.729
Cubic	0.851	0.831	0.705	0.741
Semiparametric (average)	0.875	0.864	0.721	0.789
Semiparametric (market average)	0.846	0.832	0.678	0.769

Source: calculated from NSSO data.

The food shares estimated semiparametrically were smoothed again using LOWESS, and plotted against per capita expenditure together with the estimated cubic shares (see Figure 2.10). There is a clear shift of the curves to the right over time, as per capita expenditure increases. Note that the variable reported on the *x* axis is nominal expenditure, and therefore part of this shift is simply the result of an increase in prices. There is also a downward shift of the curves, particularly in 2004-05, which is a result of the improvement in living standards, as households consume proportionally less food as income increases. With the exception of the curve of the 55th round, all curves have similar shapes. The linear portions of the curves and their slope have not changed over time. Notice also that there was little change over time in the shape of the first upward sloping tract of the curves.

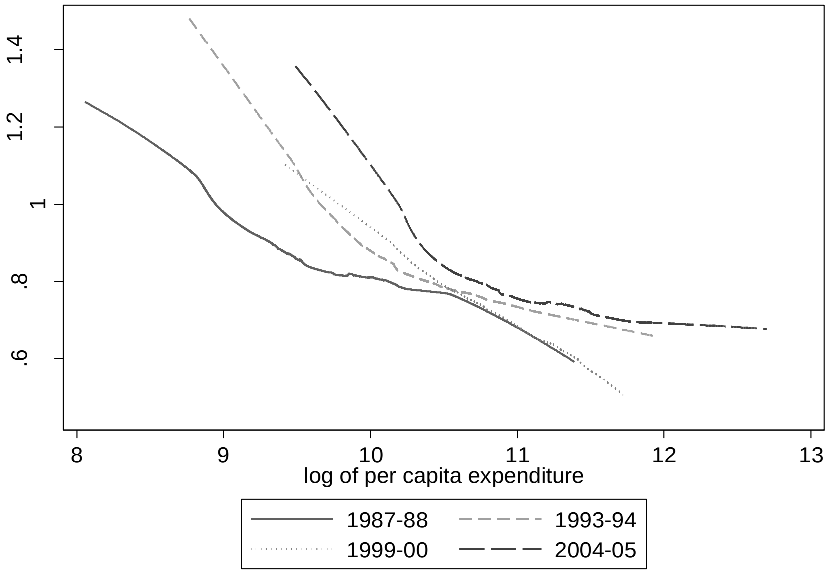
Figure 2.10 Semiparametric food shares and per capita expenditure



Source: calculated from NSSO data.

The semiparametric elasticities were smoothed again using LOWESS and plotted against per capita expenditure in Figure 2.11. There is a visible shift of the curves to the right over time, which is simply the result of the increase in nominal income. The elasticity curves have similar shapes with the exception of the curve based on the 1999-00 data. Food elasticities are above one over a large portion of the curve, they decrease rapidly up to nearly half of the expenditure distribution, and become nearly flat thereafter. Notice that the declining section of the curves are becoming steeper over time, suggesting that the relative reduction in consumption expenditure as income increases occurs at a faster rate when living standards are higher.

Figure 2.11 Semiparametric food elasticities and per capita expenditure

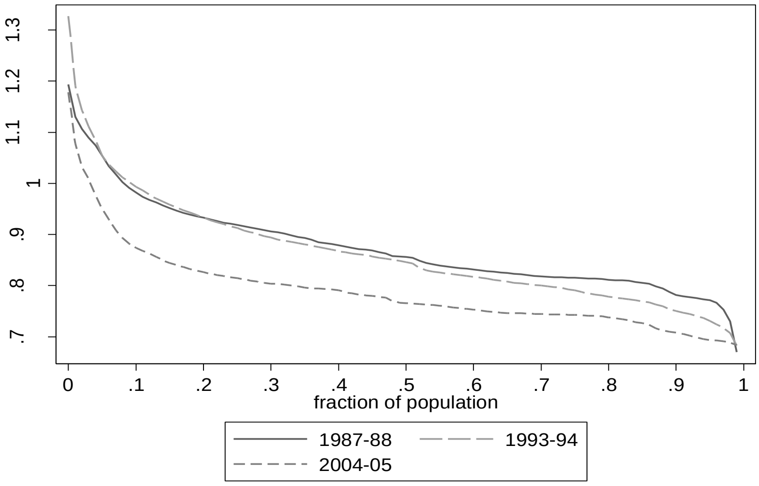


Source: calculated from NSSO data.

In order to show the size of the sample with a given value of food elasticity, Figure 2.12 plots the cumulative distribution function of the individual eslasticities for three of the four survey rounds.[12] The fraction of the population with the elasticity reported on the vertical axis can be read on the horizontal axis. The range of the estimated elasticities is relatively narrow, being between 0.7 and one for at least 90% of the sample. The majority of households show very high elasticities in all surveys. More than 30% of households have an elasticity above 0.9 until 1993-94, while nearly 40% have an elasticity above 0.8 in 2004-05. Notice that up to 1993-94, nearly 10% of the population has a food elasticity above one. This fraction declined to less than 5% in the survey of 2004-05.

[12] The distribution function of the elasticities estimated for the 55th round was omitted because its pattern is too different from those observed in the other survey rounds.

Figure 2.12 Food elasticities by population percentiles



Source: calculated from NSSO data.

2.6.2 Elasticities of different food categories

The Engel curves in Figure 2.10 exhibit two rather puzzling features. The first is the upward sloping section of the food share on the left of the charts. Even in the latest survey of 2004-05, there is a considerable fraction of households for whom food is a luxury good. The second is the similarity of the shape of the curves, despite the considerable increase in per capita income over the period covered by the four surveys. In order to find an explanation to these unexpected characteristics of the Engel curves, food expenditure was further disaggregated in seven broad categories, and the exercise was repeated for each of them.

The seven broad categories were identified through the standard procedure of grouping goods that are close substitutes in consumption. The expenditure shares and the semiparametric elasticities of each category for all survey rounds are reported in Table 2.12 and 2.13. Cereals, the first category, includes rice, a staple food in Andhra Pradesh, wheat, which is scarcely consumed, plus a number of coarse cereals like *jowar* (sorghum), *ragi* (finger millet), and *bajra* (pearl millet). Pulses include beans and lentils like *tur*, *moong*, and *gram*. Dairy products mainly consist of milk, *curd*, *ghee*, and butter. Oils and fats includes margarine, coconut oil, groundnut oil, and other edible oils. Other food includes sugar, salt, spices, beverages and processed foods. The remaining two categories, meat egg and fish, and vegetables and fruit, are self-explanatory.

Table 2.12 Expenditure shares of seven food categories

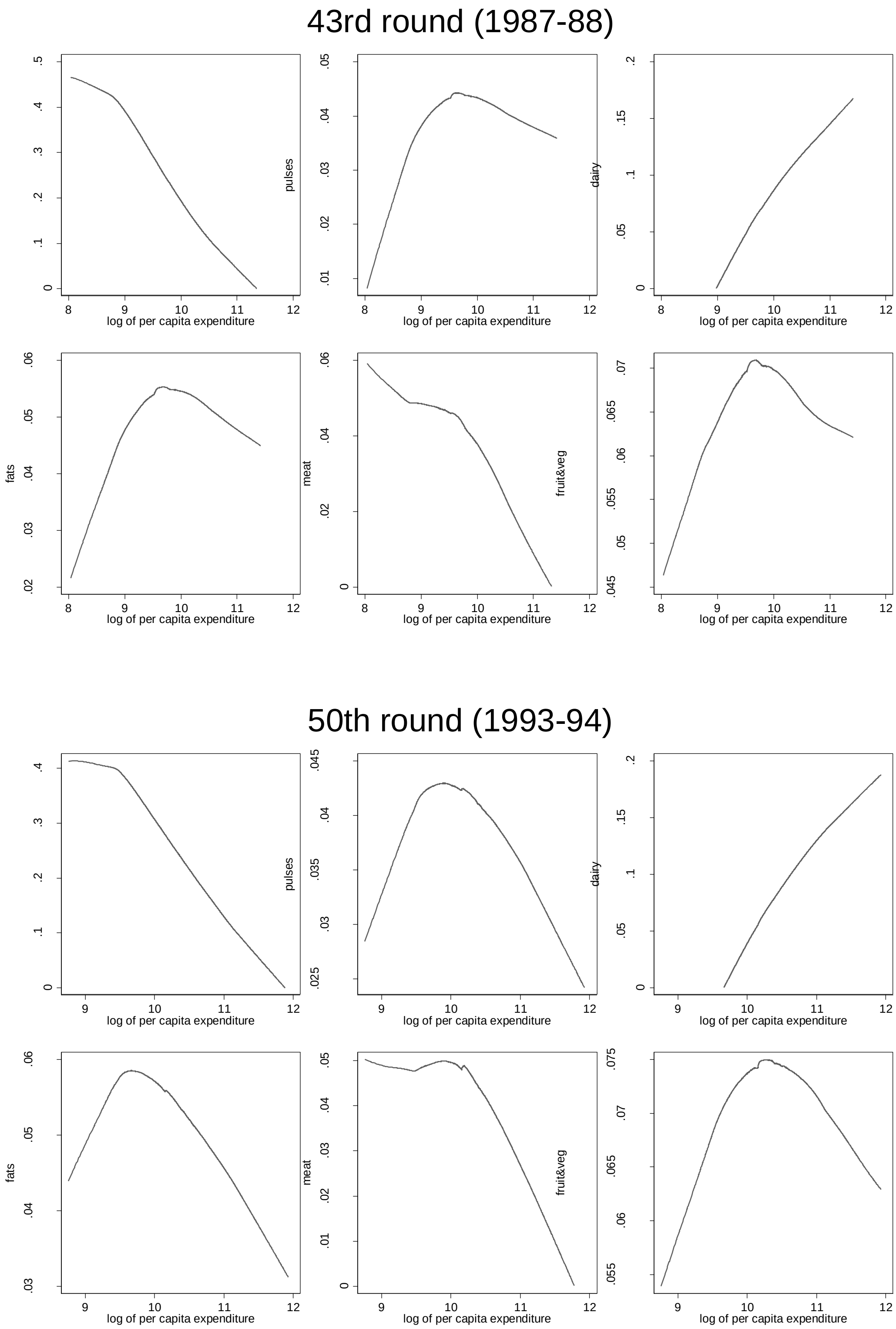
	1987-88	1993-94	1999-00	2004-05
Cereals	28.3	27.2	23.7	20.9
Pulses	4.3	4.1	3.7	3.5
Dairy products	4.9	5.5	5.6	5.4
Oils and fats	5.3	5.5	3.6	5.0
Meat, egg, and fish	4.3	4.5	4.1	4.2
Vegetable and fruit	6.8	7.3	7.1	8.1
Other	8.5	10.7	7.4	7.6

Table 2.13 Semiparametric elasticities of seven food categories

	1987-88	1993-94	1999-00	2004-05
Cereals	0.199	0.166	0.399	0.495
Pulses	1.074	0.889	0.764	0.718
Dairy products	3.006	3.156	2.061	2.208
Oils and fats	1.077	0.841	0.695	0.575
Meat, egg, and fish	0.189	0.320	0.243	0.038
Vegetable and fruit	1.018	1.032	0.702	0.769
Other	1.131	0.885	0.525	0.780

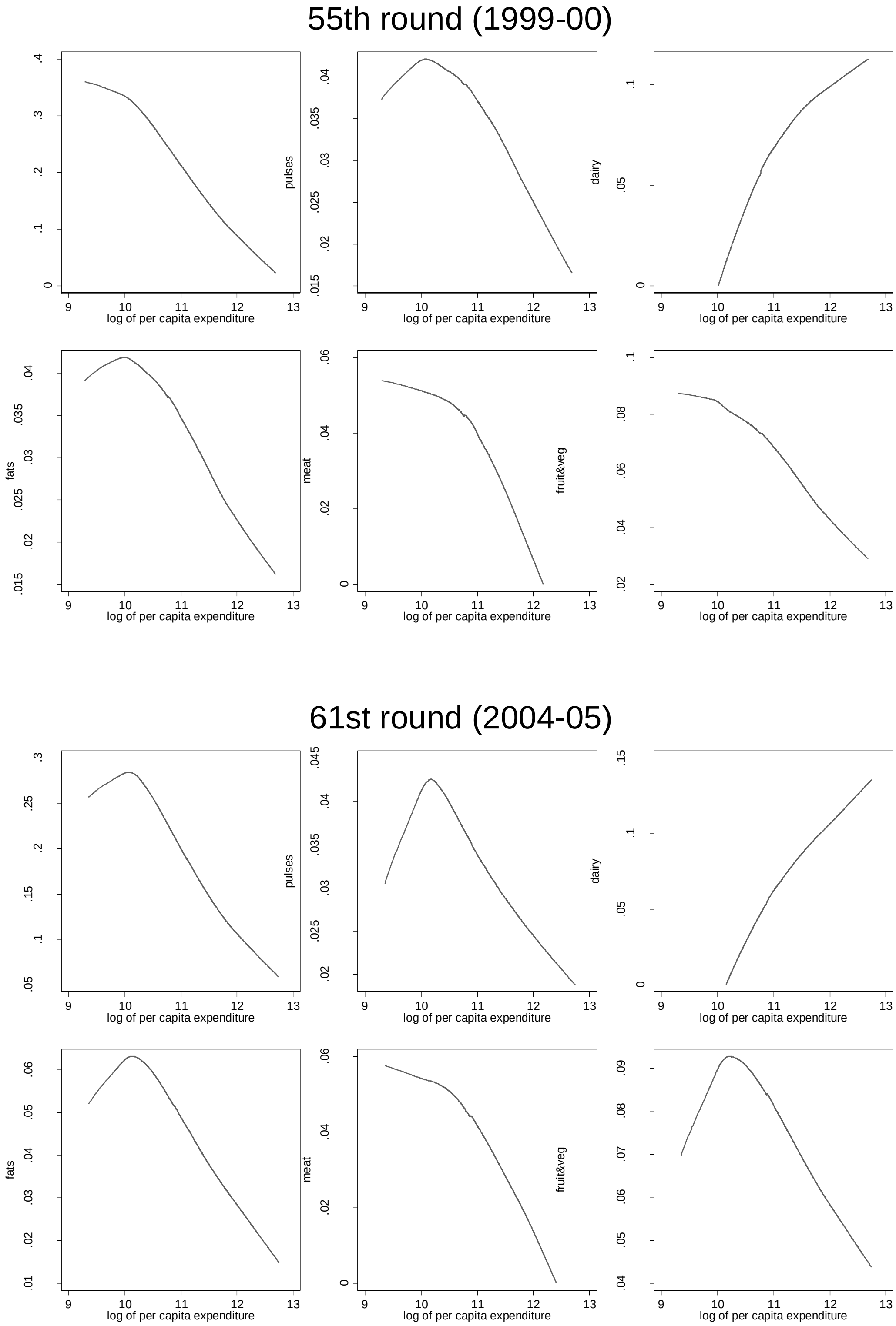
The semiparametric food shares of each category are plotted in Figures 2.13 and 2.14 for the four survey rounds, while the plots of the semiparametric elasticitities are in Appendix C. Two categories, cereals and ‘meat, egg and fish’, are necessities for all households, and display declining expenditure shares. Dairy products are clearly luxury goods. Their consumption share increases with total expenditure and is near zero for poor households. Three categories, oils and fats, pulses, and vegetable and fruit, are luxuries for the poor but necessities for the rich. The shapes of the food category shares suggest that the upward sloping section of the food share curves in Figure 2.10 are determined by the upward sloping section of the expenditure shares of oils and fats, pulses, and vegetable and fruit. Dairy products do not contribute to turning the slope of the food expenditure curve upwards because they are not affordable for the poor. This offers a simple explanation to the fact that food is a luxury good for many poor households. The diet of the poor is largely composed of cheap cereals, particularly rice. As income increases, the poor diversify their diet by increasing their consumption of pulses, vegetables, and oils. The increase in the consumption of pulses reflects the need of increasing the energy contents of the diet, while the increase in the consumption of oils and fats, and vegetable and fruit is probably explained by taste and social status respectively.

Figure 2.13 Semiparametric food shares by food category (43rd and 50th rounds)



Source: calculated from NSSO data.

Figure 2.14 Semiparametric food shares by food category (55th and 61st rounds)



Source: calculated from NSSO data.

Explaining the invariance of the food share curves over time is more difficult. In theory, changes in the shape of the Engel curve over time should mirror changes over a single cross section. For example, the change in food expenditure between a poor and a rich household in a cross section, should be equivalent to the change in food expenditure over time as the same poor household becomes as rich as the rich household. In practice this may not happen for at least two reasons. First, there are other determinants of food expenditure that change over time in addition to income. In particular, consumers' tastes and the introduction of new food and non-food products may alter food consumption patterns. Second, consumers may be influenced by habits, and therefore adjust consumption over time more slowly than the change in income suggests. Nevertheless, if these two effects are not strong, the Engel curves over time should track the cross-sectional Engel curves rather closely, and by observing Engel curves over different surveys, we should be able to see a movement of observations along the curve from the left to the right as consumers become richer. As income increases, food items that were luxuries should gradually become necessities for an increasing fraction of the population. The charts in Figures 2.13 and 2.14 seem to confirm this for some of the seven food categories. The Engel curves of pulses, fats, and vegetables and fruit seem to lose observations on the left and acquire observations on the right along the same curve as income increases over time. The estimated elasticities shown in Table 2.13 provide additional evidence that food items are becoming necessities. The size of all elasticities is decreasing over time.

One important exception to this pattern of decreasing elasticity is the behaviour of cereals, whose elasticities have increased over the last two surveys. Cereals have become luxury good for a considerable fraction of the population in the latest survey round of 2004-05. Table 2.13 shows that expenditure elasticity of cereals nearly doubled over the last two surveys compared to the previous period. This can also be seen in the charts of Figures 2.13 and 2.14. The estimated cereals expenditure share was a downward sloping line in the 43rd round, it became flat at the top in the 50th round, and became upward sloping for the poor in the 61st round. Note that since cereals still account for more than 25% of total expenditure of the poor, this drives the aggregate result that the food shares are invariant, and that food is a still a luxury good for a considerable fraction of the population in 2004-05.

The fact that rice, millet and sorghum have become luxury goods in Andhra Pradesh for a considerable number of households is puzzling and deserves at least a tentative explanation. The status of luxury for a commodity like rice suggests an increase in poverty levels. The data however exclude this, as per capita income increased and expenditure inequality decreased over this period. The explanation advanced here is that recent modifications in the public distribution system of food (PDS) have resulted in rice becoming a luxury good for some households for whom it was previously a necessity.

India has a long tradition of food subsidies. In the late 1960s the government of Subramanian introduced a procurement system, whereby the state would by food, mainly wheat and rice, in times of shortage in order to stabilise prices and to redistribute the same food to the poor (Varshney, 1995). After the success of the green revolution, the procurement system has become a *de facto* support system for producers of cereals. The supply of cereals now exceeds demand, and surplus states generate large deficits by paying farmers procurement prices higher than market prices (Jha et al., 2007).

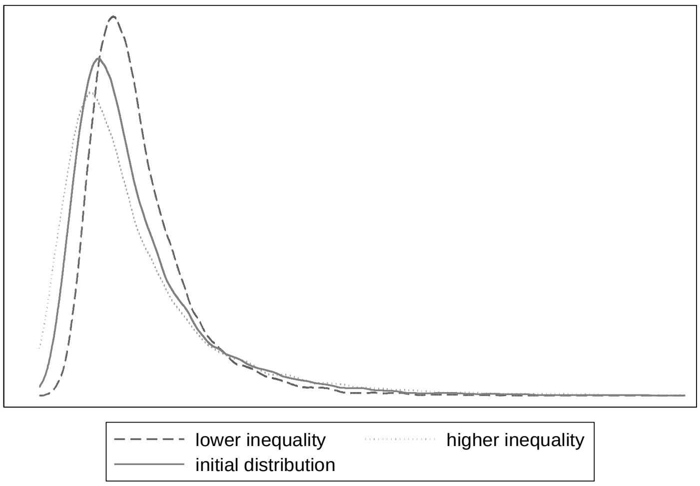
In the early 1980s the government of Andhra Pradesh began selling to the poor procured rice at two Rupees per kilogram, well below the market price. In 1992, the government increased the price of rice distributed through the PDS to 3.5 Rupees per kilogram and introduced a ceiling of 20kg per household per month. Tarozzi (2005) using anthropometric data provided by the NFHS found no effects of this price increase on malnutrition rates of children. However, the flattening of the cereals expenditure share for the poor that is visible in the 50^{th} round of 1993-94 might be a consequence of this price change.

Another important change in the PDS occurred in 1997, when the government of India introduced a new targeting system for all welfare programmes. Households below the poverty line (BPL) were issued a card that gave them access to most welfare programmes, including the PDS. The procedures for assigning the card were rather complex and changed repeatedly. Recent studies have found evidence of mistargeting in the BPL system. Dutta and Ramaswami (2001), using NSSO data of 1999-00, found a moderate targeting error in the PDS of Andhra Pradesh. The size of the type I error, the share of households who received the food and should not have, was 22%, while the type II error, the share of households that did not receive the food but should have, was 20%. Jalan and Murgai (2007) using the latest NSSO survey of 2004-05 found a type I error of 46% and a type II error of 63% among rural households of 16 major Indian states. In Andhra Pradesh, the errors were much higher: 74% and 87% respectively. According to this study, three out of four poor rural households in Andhra Pradesh are not classified as BPL and therefore do not have access to the rice distributed through the PDS. This might explain why cereals have become a luxury for the poor in 2004-05, after being a necessity for nearly all households during the previous 15 years.

2.6.3 Food demand and income distribution

This section assesses the extent to which changes in income distribution alter the aggregate consumption of food for the NSSO samples of households. This assessment is made by simulating changes in the expenditure distribution, and then calculating market elasticities and food shares at each step of the simulation. Figure 2.15 shows a kernel density plot of the expenditure distribution observed in the survey of 1987-88, together with two simulations performed by redistributing 30% of the existing expenditure using the method described in Section 2.5.3. A shift of the distribution to the right represents a decrease in income inequality, while a shift to the left is an increase in inequality. The Gini coefficient of the middle (initial) distribution is 0.281, while the Gini of the density on the right is 0.211 and the on the left is 0.336. Notice that in both cases, as income distribution changes, the sample mean and the total household expenditure in the sample remain the same.

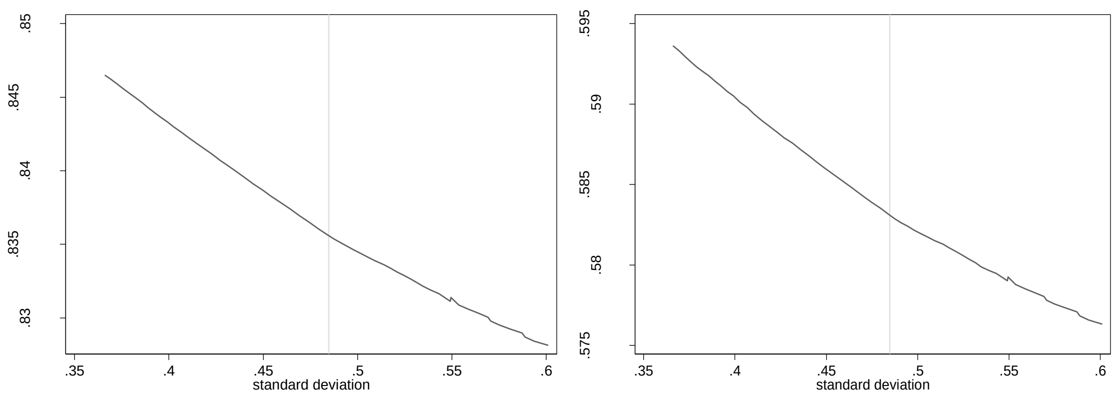
Figure 2.15 Simulated changes in the expenditure distribution



Source: calculated from NSSO data.

The simulated food elasticities and food shares for varying levels of income inequality are reported in Figure 2.16 (left and right chart respectively). The solid vertical lines in the middle of the charts indicate the value of the standard deviation of the logarithm of per capita expenditure of the initial distribution. The two curves show that a more egalitarian income distribution, while leaving income unchanged, has the effect of increasing aggregate food elasticity, aggregate food share, and total food consumption. This follows from the fact that food Engel curves in Andhra Pradesh are non-linear and that the poor spend proportionally more on food than the rich. The curves in Figure 2.16 are linear, though their slopes are slightly steeper on left-hand-side of the vertical line representing the initial expenditure distribution, which implies that food consumption changes more rapidly after a decrease in inequality than after an increase inequality.

Figure 2.16 Simulated changes in food elasticity and food share



In order to quantify the effect of a change in income distribution on food demand, the slope of the line of the chart on the right of Figure 2.16 and the elasticity of the food share with respect to the standard deviation of per capita expenditure were calculated. The first was estimated running a regression of the food share on the standard deviation of expenditure and was found to be equal to -0.075. The second was estimated by regressing the logarithm of the food share on the logarithm of the standard deviation of per capita expenditure, and was found to be equal -0.13. This implies that a 10% decrease in the standard deviation of expenditure would increase market food consumption by 1.3%. As an example, consider that the effect of the reduction in inequality observed over the 43rd and the 50th round would have resulted in an increase in the market food share from 58.2 to 58.8 The observed food share in the 50th round was 58.3 which is the result of the combined effect of the reduction of inequality and of an increase of per capita income. This means that over the period from 1987-88 to 1993-94 the income effect almost entirely compensated the income distribution effect thus leaving the food shares unchanged. This suggests that the income distribution effect can produce changes in food consumption that are similar to those produced by changes in mean income.

2.7 Conclusions

This chapter provided an account of the process of agricultural transformation in Andhra Pradesh. The theoretical part focused on the implications of different assumptions regarding the consumption function for the size of agricultural employment. It was shown in equation 2.21 that the rate of change in the share of employment in the agricultural sector is a function of the income elasticity of food demand, and that the larger this elasticity the slower is the pace of the agricultural transformation. The empirical analysis tested the hypothesis that the pace of the process of transformation of the structure of the labour force is shaped by the characteristics of the aggregate demand for agricultural goods and by the inequality of income distribution. Expenditure elasticities of food, and of seven food categories, were estimated using semiparametric Engel curves, and simulations of changes in food consumption produced by changes in income distribution were performed. The main results of the empirical analysis are as follows:

- Food expenditure elasticities are very high in Andhra Pradesh both in absolute terms and compared to other countries. Despite a considerable increase in per capita income over the last 20 years, elasticities and food expenditure shares have only slowly declined, from 0.85 in 1987-88 to 0.77 in 2004-05, and from 0.65 in 1987-88 to 0.54 in 2004-05 respectively. For a large part of the population food is a luxury good or close to be a luxury good.
- Food elasticities are very high for many households due to the poor quality of their diet. Cereals, and rice in particular, occupy the largest portion of food expenditure of poor households. As households become richer, they tend to purchase more of pulses, oils, vegetables, and eventually dairy products, for reasons of taste, social status, and energy requirements.

- An increase in the equality of the income distribution has the effect of increasing the aggregate consumption of food, as poor households, who spend proportionally more on food, have more resources. Conversely, an increase in income inequality reduces the aggregate demand for food. The income distribution elasticity of food demand (the change in the consumption of food for a change in the variance of the expenditure distribution), was found to be equal to 0.13. This effect is considerable since a simple simulation exercise over the period from 1987-88 to 1993-94 shows that the income distribution effect produced by a decrease in inequality almost entirely compensated the income effect produced by an increase in per capita incomes.
- Despite the considerable increase in per capita income and the reduction in income inequality during the period under study, cereals have become a luxury good for many poor households. Weaknesses in the new BPL targeting system may have precluded the poor from accessing food distributed through the PDS. Given that the share of expenditure on cereals among the poor is very high, this also explains why in 2004-05 food is still a luxury good for more than 5% of households.

In order to better analyse the effects of food demand on structural change, the theoretical model was built on two rather strong assumptions: the closeness of the economy, and the absence of government interventions in agricultural markets. These two assumptions will now be discussed in turn as they may affect the validity of the model.

The theoretical model of Section 2.3 was designed assuming a closed economy. This assumption is common in the literature modelling the agricultural transformation that was reviewed in Section 2.2. However, Andhra Pradesh trades with the rest of India, and with the rest of the world, and the close-economy assumption is rather unrealistic, unless it is shown that the size of openness is relatively small. As illustrated by Matsuyama (1992), when the dual economic model is considered in the context of an open economy, its results can be quite different. If the economy is open to trade, the prices of manufactured goods and of agricultural goods are no longer domestically determined. The model no longer reaches its equilibrium by setting the equality between the agricultural wage and the manufacturing wage, but by setting the equality between the ratio of the domestic wages in the two sectors and the ratio of the international wages in the same sectors. To see this, consider a country that has a comparative advantage in agriculture. This country will specialise in agriculture and will import manufactured goods from the rest of the world until the domestic ratio of the productivities in the two sectors is equal to the ratio in the productivities in the rest of the world. However, in order to achieve this equilibrium the country will have to transfer labour from the modern sector to the rural sector. In other words, in an open economy the demand for agricultural goods includes, in addition to domestic demand, international demand by countries that specialise in manufacturing. If the country has a comparative advantage in agriculture, its process of industrialisation may be delayed. The lack of data on the openness of the Andhra Pradesh economy makes a test of this hypothesis impossible.

State interventions in agricultural markets are another significant omission of the model. The state of Andhra Pradesh subsidises agricultural producers in three major ways (Gulati and Narayan, 2003). First, farmers can purchase fertiliser at a subsidised price. Second, they obtain electricity free of charge in order to operate water pumps. Finally, they can sell their produce, particularly rice, to state agencies at a minimum support price (MSP). The state runs large deficits in order to purchase rice that is then either stockpiled or sold through the PDS (Jha et al., 2007). Since rice production is labour intensive, these interventions keep labour demand artificially high. State interventions therefore result in an increase in the production of staple food and labour demand in agriculture that, according to the model of Section 2.3, contribute to the slow pace of the agricultural transformation in the state.

A number of competing explanations of the process of agricultural transformation in India can be found in the literature. Because the present analysis is conducted within a single state and because detailed data are not available at the regional level, it is not possible to assess the comparative relevance of each factor as done, for example, by Besley et al. for all India (2007). Some competing explanations of the process however have been extensively investigated in the literature and deserve to be mentioned here. These include: the impact of land reform policies and of tenancy institutions; the role played by the public education of the labour force in rural areas; and the impact of economic growth in the non-agricultural sector of the economy.

Since there is considerable evidence of an inverse relationship between farm size and productivity (see for example Berry and Cline, 1979, Binswanger et al., 1995, Carter, 1984), land reform policies are often recommended for their productivity enhancing effect. Indeed, the first Indian governments after independence largely relied on land reforms and other institutional measures in order to increase productivity in rural areas (Varshney, 1995). Besley and Burgess (2000) analyse the impact of redistributive policy in Indian states and find that land reforms were associated with poverty reduction but not with agricultural output growth. The index of land reform calculated by Besley and Burgess (2000) for Andhra Pradesh is only half the value of the all-India index. In addition, land reform in Andhra Pradesh consisted of tenancy reform and of the abolition of intermediaries. These policies are shown to have a significant impact on poverty reduction, but none on agricultural productivity. Banerjee and Iyer (2005) analyse the relationship between agricultural growth and the land tenancy structure that Indian states inherited from the British rule. They find that landlord-based areas grew more slowly during the period after the independence compared to non-landlord based areas. Most of the difference depends on the different adoption of new technologies and public investments in infrastructure, which in turn depends on the prevailing political environment. Landlord-based states like Andhra Pradesh are highly conflictive, and this prevents state governments from adopting policies promoting agricultural growth. The conclusions reached by these studies suggest that the poor pace of land reform in Andhra Pradesh, and the resulting high political conflict in rural areas, may have contributed to the poor rates of agricultural growth.

One of the main elements of the model proposed in this chapter is that the growth in agricultural productivity helps the expansion of the non-farm sector of the economy. This hypothesis was validated for India by Ravallion and Datt (1996), who used time series data of Indian states from 1950 to 1991 to test the impact of sectoral economic growth on poverty reduction. These results have been challenged by Foster and Rosenzweig (2004) who, using panel data over the period 1971-1999, show that growth in the non-farm sector was highest in areas where agricultural yields were the lowest. Conversely, areas of high agricultural productivity saw little growth of the non-farm sector. These findings are consistent with a model in which capital is free to move among areas and is invested in those areas where wages are lowest because of low agricultural productivity. The conclusion reached by Foster and Rosenzweig is that the growth of the agricultural sector did not promote growth of the non-farm sector. Burgess and Pande (2005) analysed the effect of changes in geographic coverage of Indian banks on economic growth over the period from 1977 to 1990. They found that economic growth was largely a result of an increase in non-agricultural output, while the agricultural output remained unaffected. They also found and increase in rural wages resulting from a contraction of labour supply following the increasing demand for labour in the secondary and tertiary sectors. These findings of these studies suggest that the non farm sector, rather than the agricultural sector, could be the driver of economic development in India.

Another factor that may contribute to explain the slow process of structural change in Andhra Pradesh is the poor qualification of the labour force. As shown in Section 1.2, literacy rates in rural Andhra Pradesh are very low and skilled labour is scarce. Castelli and Coleman (2001) investigate the role of the education of the labour force in the process of economic development and find that the rural labour force decreased more rapidly in the southern states of the United States than what the operation of the Engel law would predict. They attribute this result to a reduction in education costs and to the migration of rural population to the skilled sector of the economy. In Andhra Pradesh poor educational levels prevent the rural labour force from finding employment outside agriculture. Trivedi (2006) analyses the relationship between secondary school enrolment and economic growth in India using state level data for the period from 1965 to 1992. He finds that there is a positive correlation between male and female secondary enrolment rates and economic growth, and that improving female education in rural areas would have a large impact on economic growth. Foster and Rosenzweig (1996) investigate the relationship between technological change and schooling in India using panel data of NCAER covering the period 1968-1981. They find that technological change positively affects returns to schooling, either because better educated people are more able to manage new technologies, or because they become aware of new technologies more quickly. In either case, technical change is more effective and successful among an educated rural population. The relationship between economic development and education in rural areas of Andhra Pradesh will be further investigated in Chapter 4.

3 The stabilising effect of irrigation on household expenditure

3.1 Introduction

More than 30% of the GDP of Andhra Pradesh is generated in the agricultural sector, and more than 70% of the workers are employed in agriculture, either as farmers or agricultural labourers. The majority of the population living in poverty and extreme poverty resides in rural areas and, as shown in Chapter 2, the prospects of increasing living standards and reducing poverty are strongly related to developing the agricultural sector. Improvements in farms' productivity are critical to this end. Over the last 30 years, substantial technical innovation has taken place in agriculture in India. Technological progress has included the introduction of new seeds and modern inputs, and increased access to irrigation. The latter in particular has proved crucial for the adoption of new crop varieties and the use of fertiliser. It is now well documented that the adoption of a package consisting of new crop varieties, fertiliser, and irrigation – the green revolution – has determined a considerable improvement in living standards of farmers and agricultural labourers.

Not all areas and farmers benefit equally from access to irrigation. First, there are geographic areas that for their location and soil characteristics cannot be reached by canal infrastructure. Second, even in irrigated areas, not all farms have equal access to irrigation. Finally, the equitable operation of irrigation infrastructure, particularly with respect to the timing and the frequency of waterings, has proved difficult. All these arguments have been investigated at length in the literature and they will not be repeated here. Instead, this chapter focuses on a rather neglected aspect of irrigation: its stabilising effect on the income of rural households.

The benefits of income stabilisation are twofold. First, a reduction in income uncertainty and a more stable consumption pattern constitute welfare gains in themselves. Secondly, a reduction in income uncertainty may release household resources that can be invested in more productive enterprises. Potential benefits of income stabilisation over the years and over the seasonal cycle will be separately investigated for farmers and agricultural labourers.

This chapter is structured as follows. Section 3.2 reviews the literature on poverty and seasonality in developing countries; on the strategies rural households adopt to stabilise consumption; and on the impact of irrigation on agricultural income. Section 3.4 builds a theoretical model of optimal intertemporal allocation of consumption for households that are poor, face substantial income uncertainty, and have little access to credit. Section 3.5 describes the data used in the empirical analysis and the necessary adjustments made to expenditure figures. Section 3.6 presents some descriptive statistic regarding access to irrigation in the state, rainfall patterns, and household classification. Section 3.7 describes the econometric methods used in order to identify the causal effect of irrigation on household expenditure. Section 3.8 presents the results of the empirical analysis for farmers and agricultural labourers separately. Section 3.9 concludes.

3.2 Literature review

3.2.1 Poverty and seasonality

India has a highly seasonal pattern of rainfall, with 50% of total precipitation falling in just 15 days, and 90% of river flows running in just four months (World Bank, 2005). Seasonality of climate and agricultural production has a strong impact on social and economic life of rural households. A typical life scenario for a southern Indian state can be depicted in the following way (Chambers et al., 1981, Devereux et al., 2008). Rural households are employed as small farmers or agricultural labourers in an environment where a dry season follows a wet season. Food and income produced in the wet season are stored for the lean season. Toward the end of the dry season food becomes scarcer and more expensive as stocks are running down. In the pre-harvest period, when rains begin, land must be prepared and crop sown. Households are burdened by exceptional work at this time of the year which coincides with highest food prices. Labourer households migrate to areas where labour is demanded. This is the time when food is most needed for hard physical work and families are malnourished. This is also the time in which diseases are more common, particularly diarrhoea and malaria, which further debilitates poor households. With the new harvest, food becomes available again and a new cycle begins.

Historically, people have responded to this seasonal cycle of income and hunger by adapting their modes of life or by storing food. Adaptation to seasonality however can be very costly. There are a number of strategies that households can adopt in order to smooth income and consumption over the seasons, each of which bears a cost (Fafchamps, 2003). Households may store grains, thus incurring in post-harvest losses. They can borrow from local moneylenders, thus paying high interest rates. They can withhold spending for the next season, thus having their savings reduced by inflation. They can grow crops that are more resistant to water scarcity, like millets, but that yield lower income on average.

When the seasonal cycle hits particularly hard and interacts with health shocks, like a sudden death or illness, households' responses may generate vicious cycles of poverty (Gill, 1991). The families more vulnerable to economic and health seasonal shocks are also those least able to respond to shocks, and if negative seasonal shocks recurr year after year the result may be dramatic. Households affected may be forced to sell their land or livestock, thus impairing their future ability to generate income. Alternatively, they may negotiate new loans, thus falling in vicious circles of indebtedness. In extreme cases, a combination of repeated seasonal shocks, distress sales and indebtedness may lead to family dissolution, migration and destitution.

Gill (1991) draws on anthropological and medical literature from a large number of countries to show how disadvantaged households and individuals may be dragged into destitution after series of seasonal shocks. Behrman and Deolalikar (1989) show for a sample of Indian households, that if seasonal calories deficits are particularly severe the labour productivity of labour may fall to such an extent that households fall in a poverty trap. Devereux et al. (2008) describe how a series of seasonal shocks destroyed labour markets in rural Malawi, thus setting the stage for the famine of 2002.

Even if households are able to escape poverty traps, seasonal fluctuations are likely to persist. These fluctuations bear a cost, as risk-averse households prefer a stable consumption stream to an unstable one. The size of the cost will depend on initial household wealth, on the size of fluctuations, and on the degree of risk aversion.

In spite of the considerable welfare cost of seasonality, seasonal poverty analysis has been neglected in the development literature. There are two main reasons for this neglect. The first, underlined by Chambers (1981) is a general bias against seasonal analysis. The second, put forward by Sahn (1989) is the irrelevance of seasonality in the wider process of agricultural development. These two explanations are now discussed in turn.

Chambers (1981) attributes the neglect of seasonal analysis to a number of researchers' biases. First, poverty analysis is mostly conducted in urban areas with infrequent travels to rural areas and is based on a very limited knowledge of rural life. Second, all-year-round data are difficult to collect and process. Finally, the specialisation of researchers prevents seasonal analysis, because an understanding of seasonality requires knowledge in the fields of economics, anthropology, health, nutrition and demography. Gill (1991) maintains that the neglect of seasonality in the social sciences has been particularly strong among economists, because of a macroeconomics focus of development economics and because of economic curricula based on the study of developed economies. While the researchers' specialisation bias may still be present today, data availability is no longer a constraint to seasonal analysis. Many institutes like the NSSO and the Living Standard Measurement Surveys of the World Bank collect socioeconomic data over the agricultural year and pay attention to seasonal cycles of consumption and production.

Sahn (1989) maintains that a seasonal analysis of poverty is to some extent irrelevant. He observes that there is no seasonality of expenditure in developed countries, in spite of strong seasonal agricultural production patterns. According to Sahn, seasonal poverty in developing countries is likely to disappear through the processes of agricultural growth and market development. There is some truism in this argument because clearly seasonal poverty will disappear if rural poverty disappears. The effects of economic development on seasonality will now be analysed in more detail.

First, economic development and the rise in incomes reduce the welfare cost of consumption fluctuations. This happens because, assuming the concavity of utility functions, richer households are less affected by income fluctuations. India has indeed witnessed a considerable increase of rural incomes over the last two decades. However, there is a large share of the population living in extreme poverty, for which fluctuations still have a welfare cost.

Second, as income increases, the share of income produced in agriculture decreases. Therefore the income component that is sensitive to seasonal fluctuations decreases over time during the process of economic development. However, it has been shown in Chapter 2 that the process of agricultural transformation has been rather slow in India, and that a large share of income and employment is still generated in the agricultural sector. There are also signs that climate change is increasing the size of fluctuations, by making rains more erratic and unpredictable (Devereux et al., 2008).

Third, as markets develop in rural areas, thanks to the development of infrastructure, the cost of coping with seasonal fluctuations decreases. Much of the cost of seasonality derives from the inability to transport agricultural products from surplus areas to deficit areas; from the inability to store production efficiently; and from household inability to insure production or borrow. While there has been considerable progress in road infrastructure in India, financial markets and agricultural technologies are still rather underdeveloped.

For the reasons outlined above, Indian households still bear a considerable cost for seasonality. While market development forces are in operation, there are two main policies that can be adopted to help households smoothing the seasonal cycle: social protection interventions and technological development of agricultural production. This chapter is concerned with the second type of policies. In particular it investigates to what extent improvements in water storage systems reduce seasonal welfare fluctuations.

3.2.2 Seasonality and irrigation

Irrigation has an obvious positive impact on agricultural production. First, irrigation determines an expansion of the area cropped. This expansion occurs both spatially, by bringing under cultivation previously uncultivated soils, and temporally, by allowing the same area to be cultivated more than once during the same agricultural year. Secondly, irrigation increases farm yields, because the availability of water encourages the adoption of more productive crop varieties and production technologies. Lipton et al.

(2003) offer an exhaustive review of the evidence gathered to date on the links between irrigation, income, and poverty reduction.

There is also a considerable body of research on the impact of irrigation on the income of all sectors of rural society, not just farmers, as well as on the wider economy. By increasing agricultural production, irrigation determines changes in prices, wages, and in the structure of the labour force. These changes affect the income of all sectors of the economy in a way that is not easily predictable. Studies analysing irrigation within the agricultural sector, or the entire economy, include Quizon and Binswanger (1986), Datt and Ravallion (1998), Fan and Hazell (2001), and Palmer-Jones and Sen (2003). All these studies found important effects of investments in irrigation on poverty reduction.

Much less attention has been devoted to the stabilising effect of irrigation on rural income. Farmers can gain considerably from a greater stability of income. First, a more stable income is in itself a welfare gain. Second, income stabilisation helps farmers to make more investments, by releasing resources previously held as precautionary savings, or by allowing the undertaking of riskier and more profitable enterprises. This point was originally made by Lipton (1989) when considering the combined effect of irrigation, new varieties and technologies (the green revolution) on poor farmers. The subject however was no further investigated, and the present study is the first attempt to empirically assess the stabilising effect of irrigation on farmers' income.

Strictly related to the issue of income volatility is the volatility of consumption, which is the subject of a vast literature on poverty dynamics, thoroughly reviewed by Baulch and Hoddinott (2000), and by the World Development Report 2000/2001 *Attacking poverty* (2001). This literature analyses the variability of income, consumption and poverty in developing countries, often using panel data. The methods most commonly used are transition matrices of income and of poverty distributions, and regression analysis of changes in income, expenditure, and poverty status. The remainder of this section summarises the main findings of this body of research.

Transition matrices show that income and expenditure volatility in developing countries are extremely high, especially in the lower and middle part of the income distribution (see for example Carter and May, 2001, McCulloch and Baulch, 2000). Early studies of income and consumption variability (Townsend, 1994, Wolpin, 1982) found evidence in support of the permanent income hypothesis and of perfect consumption smoothing. However, the large fluctuations in consumption found by more recent research indicate that households are much less able to insulate expenditure from income fluctuations than previously thought.

Many determinants of income and consumption variability have been identified with the use of regression analysis. Among these, a distinction is usually made between covariant factors, which affect a large number of households at the same time, and idiosyncratic factors, which affect a single household. Factors of the covariant type include droughts (see for example the studies by Fafchamps et al., 1998, Hoddinott and Kinsey, 2001) and macroeconomic shocks (see for example Glewwe and Hall, 1998, McKenzie, 2003). Factors of the idiosyncratic type include individual shocks, like health shocks (Dercon and Krishnan, 2000a), and composite indices of multiple shocks (Carter and May, 2001). In regression analysis, covariant and idiosyncratic factors often explain a large part of total variation in household expenditure.

It has also been found that different households have different ability to insure against income shocks. Models normally assume households' inability to borrow, and look at the use of liquid assets as buffer stocks. Assets are broadly defined and include livestock (Rosenzweig and Wolpin, 1993; Fafchamps, Udry, Czukas, 1998), risk sharing through social networks (Udry 1994), and social capital (Carter and Maluccio, 2003). The results of these studies vary with the context and the specification of the models, but some general conclusions can be drawn. First, households are more efficiently protected from idiosyncratic shocks than from covariant ones. Second, households that are differently endowed in terms of assets used as buffer stocks show different consumption patterns in response to income shocks. Third, if the assets used as buffer stocks are also used for productive purposes, poverty traps can emerge. This last theme has been the subject of a series of studies reviewed in a special issue of the Journal of Development Studies by Carter and Barrett (2006). Finally, it was found that households can reduce consumption volatility by smoothing income fluctuations directly. The theoretical background to this type of research has been provided by Morduch (1994). Examples of income smoothing include: the use of household labour supply against idiosyncratic shocks (Kochar, 1999); the choice of a less risky crop mix

(Kurosaki and Fafchamps, 2002), and the choice of less risky livelihood strategies (Barrett et al., 2001). As in the case of consumption smoothing, the ability to smooth income varies with households' assets, and poverty traps can emerge if income smoothing choices determine lower, though more stable, future income streams.

The present chapter is concerned with the analysis of seasonal fluctuations in income and consumption, rather than year-to-year fluctuations. This choice is partly dictated by practical reasons and partly by theoretical ones. The NSSO data used in this study are collected every year, but never on a scale sufficiently large to build time series. On the other hand, the surveys are composed of seasonal sub-rounds, which lend themselves well to a seasonal analysis. The last section of this chapter attempts to assess year-to-year fluctuations in consumption of irrigated and non-irrigated farmers and labourers. This however should be considered as a descriptive exercise, and its limitations will be carefully outlined.

There are also some theoretical advantages in analysing expenditure over the seasonal cycle. First, seasonal variations in economic quantities, even dramatic ones, are a common characteristic of rural areas in developing countries, and households responses to seasonal changes are themselves of interest. Secondly, it is possible that the determinants of income and consumption fluctuations are more easily observed and identified over the seasonal cycle than over the years. To see this, consider the large number of factors that have to be taken into account in order to analyse year-to-year fluctuations: weather, government policies (introduction and withdrawal of support to producers and consumers, insurance schemes, and income support policies), and trends in national and international prices. When focusing on seasonal series, many of these factors lose importance, assuming their impact is distributed over all seasons and is not concentrated on any particular one. As a consequence, the seasonal series can offer a clearer picture of the links between weather, income and consumption in isolation from the influence of other factors. Finally, the impact of irrigation on rural incomes has an obvious seasonal component, as irrigation allows the cultivation of the same area in more than one season.

Consumption seasonality in developing countries has been the subject of a small number of studies, which include Paxson (1993), Jacoby and Skoufias (1998), Dercon and Krishnan (2000b), and Pitt and Khandker (2002). The main aim of the first two studies is testing the validity of the permanent income model of consumption over the seasonal cycle. Paxson (1993) compares seasonal consumption patterns of household groups with different income patterns in Thailand, and finds that consumption is independent of seasonal income fluctuations. Jacoby and Skoufias (1998) use panel data of Indian households and cannot reject the hypothesis that households are able to smooth consumption. They attribute households' perfect smoothing ability to the working of credit markets, and the exchange of gifts. The other two studies formulate and test behavioural assumptions regarding seasonal consumption. Dercon and Krishan (2000) test the dependence of seasonal consumption on income shocks of Ethiopian households and find that households are unable to insulate consumption from income shocks. Pitt and Khander (2002) find that the participation in micro-credit programmes in Bangladesh is strongly motivated by their effectiveness in reducing seasonal consumption fluctuations. The present study is closer in spirit to these latter two studies, as it pursues an alternative analysis to testing the permanent income hypothesis.

There is virtually no research on income and consumption seasonality in developed countries. The only exception is Miron (1996) who points out that after an initial strong interest in seasonal fluctuations, economists working after the great depression of the 1930s have ignored the study of seasonality in favour of year-to-year fluctuations. Seasonality is seen as a nuisance to be removed from the data, rather than a subject of research in itself. This is so despite the fact that the seasonal variations in production and consumption in developed countries are very large. Diewert (2004) reports that seasonal expenditure accounts for one-fifth to one-third of total expenditure of a typical developed country. Climate explains seasonal expenditure on many food and non-food items, including clothing, beverages, electricity, recreational transport and others, while custom explains the concentration of expenditures in certain times of the year, particularly Christmas. The reason seasonal variations have been neglected in developed countries is probably that they are considered a matter of choice (preferences) rather than of necessity. Output seasonality in developed countries is not believed to have an effect on seasonal consumption, and certainly not to the extent that it does in countries where incomes are strongly dependent on the vagaries of the weather.

3.3 A model of irrigation and consumption variability

This section discusses the theoretical relationship between irrigation and income variability and presents a model of the impact of irrigation on seasonal consumption variability.

3.3.1 Irrigation and income variability

The positive effects of irrigation on income derive from its ability to insulate production from weather variability. A brief description of the sources of weather variability is therefore required. Andhra Pradesh, as much of the rest of India, has two rainy seasons (see Chatterji, 1992). The first and most important occurs from June to October, with winds blowing inland from the sea (southwest monsoon), while the second, of less importance, occurs from November to April, with winds blowing from inland towards the sea (northeast monsoon). In the agricultural calendar these are known as the *Kharif* and the *Rabi* seasons. For the most part, the timing and duration of monsoon is unpredictable, and no clear correlation is known between *Kharif* and *Rabi* rains. Drought conditions may result from the late onset, early withdrawal, or prolonged breaks during the monsoon. Recently cyclical pattern of drought have been discovered. Gadgil et al. (2002) report that during the period 1918-1998 there were 23 *El Niño* years, which considerably affected seasonal rainfall and yields in southern Andhra Pradesh.

By providing water to areas and at times when it is scarce, irrigation has a stabilising effect on agricultural production. Dhawan (1988), using a considerable amount of Indian data, defines two stabilising roles of irrigation: the protective role and the stabilisation role. The protective role is the smoothing of year-to-year fluctuations. Irrigation reduces the impact of droughts by recharging the water-table and by giving access to reliable sources of water, either through canal or ground water. In any poor year and season the output of irrigated areas is higher than that of dry areas. The stabilisation role is the smoothing of seasonal rainfall variations: irrigation allows one or two additional cropping seasons. Typically dry areas have a good *Kharif* season and a poor and uncertain *Rabi* season, while irrigated areas have two fairly reliable seasons plus a potential third season (the summer season). It should be noted that the stabilisation power of irrigation also depends on the source of irrigation. At one extreme there is drought-proof irrigation, in areas that have permanent access to canal water from a safe reservoir, while at the other extreme there is the traditional village reservoir (the tank), which is highly dependent on local rains.

The positive effects of irrigation on income are not limited to farmers. Irrigated farms demand more labour because they cultivate larger areas and crops that require more labour. Rice, in particular, is very labour intensive at the time of transplanting and harvesting. In addition, because agricultural production is more stable, labour demand and agricultural labour wages over the seasons and over the years are also more stable. Irrigation therefore has also a positive impact on employment and wages. A discussion of these effects, including an exhaustive list of references, can be found in Lipton et a. (2003).

3.3.2 Income variability and consumption variability

Hypotheses on the impact of irrigation on consumption variability depend on the postulated relationship between consumption and income. This is one of the oldest, and not yet resolved, issues in economics. In this section the standard intertemporal consumption model that has remained popular until the mid 1990s will be described. More recent developments in consumption theory will then be discussed, and the precautionary saving model of consumption will be introduced. Finally, the implications of this latter model for the analysis of the effect of irrigation on consumption decisions will be outlined.

In the classical intertemporal model, consumption in the current period is not related to current income but to a permanent or lifetime income. In the simplest version of the model given by Modigliani (1966), individuals save for their retirement during their working life and choose to consume every year a nearly constant fraction of their lifetime income. Friedman’s (1957) version of the model is simply based on the behavioural hypothesis that people prefer a stable consumption stream to a variable one, but the conclusions reached in terms of the relationship between consumption and income are the same as those of Modigliani’s model. Formally, it is assumed that individuals maximise a utility function made of the sum of utilities from consumption in all t years from one to T:

$$\max U = \sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)u(c_t) \tag{3.1}$$

where ρ is the rate of time preference. Maximisation of household utility (U) is subject to the constraint that consumption (c_t) in each period cannot exceed income in the same period (y_t) plus the current period assets (A_t) minus the assets carried over the second period and discounted by the interest rate (i):

$$c_t = y_t + A_t - \frac{A_{t+1}}{(1+i)} \tag{3.2}$$

Assets carried over in the second period are discounted by the interest rate, because the consumer earns an interest over the accumulated assets which can be spent in the current period. Assets in the future period A_{t+1} can also be negative if the consumer is allowed to borrow. The simplest version of this model considers only two periods and assumes that assets will be zero at the end of the second period. Given that everything is consumed in the second period, consumption in the second period is $c_2=A_2+y_1$ and the intertemporal budget constraint over the two periods is:

$$c_1 + \frac{c_2}{1+i} = A_1 + y_1 + \frac{y_2}{1+i} \tag{3.3}$$

This easily generalises to many periods in the form:

$$\sum_{t=0}^{T}\frac{c_t}{(1+i)^t} = A_t + \sum_{t=0}^{T}\frac{y_t}{(1+i)^t} \tag{3.4}$$

The solution to the maximisation problem in (3.1) subject to the budget constraint (3.4) can be found by setting the Lagrangean:

$$L_t = \sum_{t=0}^{T}\left(\frac{1}{1+\rho}\right)u(c_{t+1}) - \lambda\left[\sum_{t=0}^{T}\frac{c_{t+1}}{(1+i)^t} - A_t - \sum_{t=0}^{T}\frac{y_{t+1}}{(1+i)^t}\right] \tag{3.5}$$

From the first order condition for c_t and c_{t+1}, the so-called Euler equation can be obtained:

$$u'(c_t) = \frac{1+i}{1+\rho} u'(c_{t+1}) \qquad (3.6)$$

This equation states that the consumer is indifferent between consumption in the current period and consumption in any other period, because both alternatives, given the values of the interest rate and of the time preference, provide the same utility.

In this formulation income is known with certainty. A permanent increase in income determined by the access to irrigation has the effect of increasing consumption over all periods. The consumption pattern over time however, is entirely determined by the ratio between the interest rate ($1+i$) and the individual time preference ($1+\rho$). Irrigation has no effect on this pattern, simply because it is assumed that income variability has no effect on consumption or that consumers know income variability with certainty.

The extension of this model to seasons rather than years is straightforward. It is sufficient to substitute seasons for years in the utility function and in the budget constraint, and to add seasonal consumption preferences in the utility function.[13] The expressions (3.1), (3.4), and (3.6) can be indexed over seasons rather than years without producing any change in the functional relationship between income and consumption. The ability of consumers to allocate consumption over time is unaffected by the change in the spacing of time. Irrigation has no more impact on seasonal expenditure than it has on yearly expenditure, and seasonal consumption fluctuations, if any, are the result of consumers' choices independently of incomes.

The assumptions of the classical intertemporal consumption model have been rejected by many empirical studies and its theoretical plausibility has often been put into question (Frederick et al., 2002). The poor applicability of this model to developing countries is discussed in Deaton (1997). The failure of the model to account for fluctuations in consumption has often been attributed to the neglect of borrowing

[13] The inclusion of preferences in the utility function is required if consumption is higher in some seasons compared to others, because for example individuals spend more during festivities.

constraints and uncertainty. If consumers are unable to borrow against income fluctuations, as it is often the case in rural economies, the conclusion of the model are unwarranted. If households cannot borrow, there will be times when more consumption in needed than the one provided by current income and assets, and consumers will not be able to equalise the marginal utilities across time as in (3.6).

More recently, the basic behavioural assumptions of the intertemporal consumption model have also been questioned and this has led to the formulation of what is known as the precautionary savings model of consumption. The precautionary model of consumption departs from the standard intertemporal consumption model in two ways. First, it assumes that individuals cannot predict future incomes with precision and that they are risk-averse. Second, it postulates that poor people are impatient, in the sense that they prefer consumption now to higher consumption in the future. This is because their current level of consumption is so low that they cannot reduce it even if this implies an even lower consumption in the future. These two behavioural assumptions will now be discussed in turn.

Income uncertainty is a reasonable assumption in rural areas of developing countries, where production is strongly linked to weather conditions. One way of introducing income uncertainty in the model is to use a multiplicative risk (Newbery and Stiglitz, 1981). For example, consider an income that each year, or each season, is multiplied by a risk factor ϑ with expectation $E\vartheta = 1$, and variance $Var\vartheta = \sigma^2$:

$$y = \vartheta q \tag{3.7}$$

The expected agricultural income is equal to y because $E\vartheta = 1$, but the certainty equivalent income (which is the expected utility for that level of income) varies with the risk variance σ^2. Assuming risk-averse consumers with utility functions that are concave down, the certainty equivalent income, denoted by a circumflex accent, is lower than expected income ($\hat{y} < Ey$) for any variance that is larger than zero, and more so the larger the variance of expected income. In the present formulation the income risk consists of rainfall variability, which affects both farmers' output and agricultural labourers' wages via demand effects.

As it was discussed in Section 3.2, irrigation has a double effect on income. The first effect is increasing agricultural output and labour demand so that income increases by a scalar δ. The second effect is reducing income variability, so that the risk variance and the certainty equivalent income become functions of irrigation (I): $Var\vartheta = \sigma^2(I)$ and $\hat{y}(I)$. With the expansion of access to irrigation, expected income increases by the factor δ, and the certainty equivalent income also increases via a reduction in the risk variance. Note that it is assumed that the income process is entirely exogenous to the decisions of farmers and agricultural labourers. This is only partly true, because farmers can reduce the variability of income by growing drought resistant crops and by purchasing water pumps that extract water from the water table. Similarly, agricultural labourers can reduce the variability of wages, over seasons or years, by migrating to irrigated areas. The assumption of the exogeneity of the income process is made in order to simplify the exposition of the model, and it will be dealt with in Section 3.6, where the econometric specification of the model is discussed.

The combined result of income uncertainty and aversion to risk is that people save in good times for fear of the occurrence of hard times. Formally, this assumption implies a concave utility function and a convex derivative function of the utility function (the marginal utility of consumption $u'(c)$). Income uncertainty is introduced in the model by adding the expectation operator (E) to future utilities, and maximisation of utility with respect to the budget constraint gives the Euler equation with the expectation operator:

$$u'(c_t) = \frac{1+i}{1+\rho} E_t u'(c_{t+1}) \tag{3.8}$$

According to this expression, an increase in uncertainty of future consumption reduces consumption in the present time. To see how this happens, consider an increase in the variance of future consumption that leaves the future mean unchanged. Because marginal utility is convex this has the effect of increasing the utility of future consumption compared to present utility. To the consumer, future consumption is now more valuable, and he will equate the marginal utilities by consuming less at the present time. In other words, consumers respond to an increase in the expected income variability by saving more.

Models of precautionary saving have often rested on the hypothesis of liquidity constraints (see for example Deaton, 1991). If consumers were able to borrow, they would not need this sort of saving. More recently Carroll (2001) has shown that liquidity constraints are not strictly necessary in order to have precautionary behaviour and that people may adopt this form of saving even if able to borrow. Intuitively this happens because consumers self-impose a reluctance to borrow for fear of not being able to repay in the future. Formally, this is again implied by the convexity of the derivative of the utility function and it is illustrated by Carrol (2001) using dynamic programming methods.

The estimation of time preferences and discount factors in developing countries is a lively area of research and the evidence gathered so far, and summarised in Cardenas and Carpenter (2008), is mixed. The hypothesis that poor people are impatient was first formulated by Fisher (1930) and it is seems a rather plausible one. People who cannot afford to buy a basic basket of goods or cannot afford what is considered a minimum socially acceptable amount of items have very little incentive to save. Formally this means that the value of ρ, the time preference discounting factor, is large. Unlike permanent income models, which assume the equality between the interest rate and the rate of time preference, it is here assumed that the ratio $\frac{1+i}{1+\rho}$ is less than one. Everything else equal, individuals with a time preference of this sort show a decreasing pattern of consumption over time.

The combination of impatience and uncertainty drives the results of the precautionary saving model (Deaton, 1997). On the one hand, poor people want to spend all their money in order to satisfy their immediate needs. On the other hand, they fear the consequences of very bad times and therefore save some of the money earned in good times. There are two interesting results of this model for the present study. First, though individuals are partially able to smooth consumption fluctuations, consumption tracks income rather closely. The consumption function derived by the model is of the Keynesian type. At low levels of income and assets consumers spend all they have.

When income and assets reach their mean level, consumers begin saving and do so increasingly as income increases. Second, an increase in uncertainty shifts the consumption function downwards, because consumers need to save more for precautionary reasons.

Within the precautionary savings model, irrigation not only increases consumption in response to the increase in expected future income, but also stabilises consumption over the years. The first effect is rather obvious, while the second needs some additional explanation. The reduction of income fluctuations allows farmers to keep the same smoothed consumption pattern at a lower level of precautionary savings. In principle, farmers might decide to invest rather than consume all extra precautionary savings, but some of these savings will indeed be used to stabilise consumption. This can be seen again from the Euler equation. A decrease in the variance of expected consumption, brought about by a decrease of income variability, has the effect of decreasing the value of future marginal utility of consumption. The consumer equates the marginal utilities by increasing current consumption, which happens at the expenses of precautionary savings.

3.3.3 A model of seasonal consumption smoothing

This section presents an intertemporal consumption model with a precautionary saving motive over the seasonal cycle. The model is solved analytically over a period of three seasons using dynamic programming methods and the implications of the model for a rural economy composed of irrigated and non-irrigated farms are discussed.

The precautionary savings model can be extended to incorporate different seasons by exploiting the additivity and separability of the intertemporal utility function. Additivity and separability are routinely assumed in the solution of intertemporal consumption models and are maintained here. Separability means that overall utility from consumption can be considered as a function of sub-utilities u_i, where the aggregate consumption c_i represents bundles of similar goods (like food, housing and entertainment) or goods consumed in different years:

$$U = F\left(u_1(c_1), u_2(c_2), \ldots, u_n(c_n)\right) \tag{3.9}$$

Additivity means that the consumption groups enter the utility function additively, so that the additive and separable utility function takes the form:

$$U = F\left(u_1(c_1) + u_2(c_2) + \cdots + u_n(c_n)\right) \tag{3.10}$$

which is the form normally used in the classical intertemporal consumption model. The separation of utility in a number of sub-utilities implies that consumers make spending decisions in stages. In the first stage they decide how to allocate expenditure between groups of goods, and in the second stage they allocate expenditure within each group. Decisions taken at each stage can be considered as solutions to different maximisation problems. The implication is that a change in prices in one group will affect the distribution of expenditure within that group and between groups, but will have no effect on the distribution of expenditure within the other groups.

In the intertemporal consumption model, the consumer first maximises utility between years and then maximises utility within each year. The introduction of seasons is equivalent to the introduction of another layer in the decision process. Utility in each year can be thought of being composed of seasonal sub-utility functions u_{si} (with a number of s seasons running from one to N) so that the separable utility function becomes:

$$U = F\left[u_1(u_{s1}(c_{s1})), u_2(c_{s2})), \ldots, u_n(c_{sn})\right] \tag{3.11}$$

while the additively separable utility function is:

$$U = F\left[u_1(u_{s1}(c_{s1}) + u_{s2}(c_{s2}) \ldots) + \ldots + u_n(c_{sn})\right] \tag{3.12}$$

The consumer first maximises utility over years, then maximises utility over seasons and finally maximises utility within each season. The implication is that the allocation of expenditure over seasons is independent of the allocation of expenditure over the years. A change in the distribution of consumption over the years does not change the way consumption is distributed over the seasons in each year. This simplifying assumption allows the consideration of the utility maximisation problem over the seasons independently of the maximisation problem over the years. If there are three seasons the seasonal utility function will take the form:

$$\max u_{seas} = \sum_{s=1}^{3} \left(\frac{1}{1+\rho_{seas}} \right) u(c_s) \tag{3.13}$$

where the arguments are now indexed over seasons rather than over years, and the discount and the interest rates are now seasonal rates. Similarly, the seasonal Euler equation will take the form:

$$u'(c_s) = \left(\frac{1+i_{seas}}{1+\rho_{seas}} \right) E_s u'(c_{s+1}) \tag{3.14}$$

This equation states that consumption in each season is related to the level and the variance of consumption in the next season. Higher future consumption reduces the marginal utility of consumption in the future, and the consumer equates present and future marginal utilities by increasing consumption in the present. A larger variance of future consumption, with expected consumption invariant, increases the marginal utility of future consumption, and the consumer equates marginal utilities by decreasing consumption in the present. Irrigation has the effect of increasing the expected value of consumption in future seasons and reducing the variance of future consumption. Therefore, irrigation increases and stabilises seasonal consumption. This will be further illustrated with a dynamic programming example.

Dynamic programming is a technique based on the principle of backward induction, which is used in the solution of multistage intertemporal decision problems. The method first finds the optimum choice in the last period considered, and then works backwards to find the optimal choices in previous periods that have the final optimum choice as their outcome. Dynamic programming is meant here to model consumer behaviour rather than mimic the way individuals make rational decisions. However, given the very small time frame considered (three seasons) and with some reasonable assumptions, the solution might well approximate the farmers' decision process.

Consider a farmer that maximises utility over three seasons: *Kharif*, *Rabi* and the lean season. In order to model risk-averse farmers, the constant relative risk aversion utility function (CRRA) is used:[14]

$$u(c_s) = \sum_{s=1}^{3} \frac{1}{1+\rho_{seas}} \left[\frac{c_s^{1-\gamma} - 1}{1-\gamma} \right] \tag{3.15}$$

Utility is maximised with respect to the following budget constraint:

$$A_{s+1} = (1+i_{seas})(A_s + y_s + c_s) \tag{3.16}$$

where A_s are assets at the beginning of the season, y_s is income earned during season *s*, and c_s is consumption during the same season. The first season of the decision process is *Kharif*, when the larger fraction of annual income is made. It is assumed that in *Kharif* the consumer has no assets inherited from the previous seasons, and income is known with certainty, because already realised. The second season is *Rabi*, when an uncertain income is made and assets are those saved in *Kharif*. The third season is the lean season, when we assume that no income is earned and consumers spend any asset they were able to save in the previous two seasons. Formally we have the following restrictions: $A_1 = 0$, $A_4 = 0$, and $y_3 = 0$. The model is solved for the values of consumption in the three seasons, and of assets in the second and the third season. The application of dynamic programming to the problem yields the following consumption values for the three seasons: [15]

[14] The characteristics of this utility function are illustrated in Appendix D.

[15] These solutions were obtained following a similar worked dynamic programming example in Leonard and Long (1992). See Appendix E for a step-by-step derivation of the consumption equations in (3.17). In all the expressions, $d = \left(\frac{(1+i_{seas})^{1-\gamma}}{1+\rho_{seas}} \right)^{\frac{1}{\gamma}}$.

$$c_3 = A_3$$

$$c_2 = \frac{A_2 + y_2}{d+1}$$

$$c_1 = \frac{y_1 + \dfrac{\hat{y}_2}{(1+i_s)}}{1+d+d^2} \tag{3.17}$$

In the lean season (c_3) the solution is obtained by assumption. Nothing is saved for the following season and nothing is earned, and the consumer spends all assets previously saved plus the interest rate. In the *Rabi* season (c_2) income is known and is distributed between *Rabi* and the following lean season depending on the values of the parameters i_{seas}, ρ_{seas}, and γ that are contained in d. In the *Kharif* season (c_1), the farmer decides the expenditure level based on his expectation of income in the *Rabi* season. This result could be presented in the form of an Euler equation showing that given the convexity of the marginal utility function, consumption in the current period is determined by the expectation and by the variance of expected consumption in the following season. This result can be shown equivalently by substituting expected income in the next season by its certainty equivalent ($\hat{y}_2$). In this way, the dynamic problem under uncertainty is transformed into a problem under equivalent certainty. An increase in the variance of *Rabi* expected income reduces its certainty equivalent value in *Kharif*, thus reducing consumption in *Kharif* in order to build precautionary savings. Conversely, a decrease in the variance of *Rabi* expected income, brought about by easier access to irrigation, increases its certainty equivalent and hence consumption in *Kharif*. In this example all precautionary saving occurs in *Kharif*. Irrigated farmers can spend more in *Kharif* as they need less precautionary savings.

The model was solved without imposing a borrowing constraint, but liquidity constraints are not strictly necessary if consumers self-impose restrictions to borrowing out of fear of disastrous events. This argument is based on Carroll (2001) and proceeds in the following way. In the third season consumers spend everything. As income in the third season is zero, consumers need to make positive savings in the second season in order to be able to spend in the third season. In order to consume less than income and assets in the second season, farmers need to make sure that they arrive to the second season with positive assets. If income is zero in the second season, farmers will not be able to consume anything. The same reasoning applies to disastrous events that produce an income larger than zero but still very low in the *Rabi* season. If income is always positive in the *Rabi* season, the farmer can borrow in *Kharif*, but he will not borrow more than what he would be able to repay in the case of the worst income outcome in *Rabi*. In these circumstances farmers will borrow, but never beyond a certain amount that it is limited by their ability to repay in the future. This seems to offer another argument in favour of the stabilising effect of irrigation on consumption. To the extent that irrigation reduces the variance of future income, it makes farmers less reluctant to borrow and therefore more able to smooth consumption.

3.4 Data

This section describes the data used in the empirical analysis and discusses the limitations of the expenditure data. In three separate subsections it details the type of adjustments made in order to correct for seasonal, demographic and price variability.

3.4.1 NSSO data

India has been a world pioneer in conducting household budget studies. Since the 1940s the NSSO has been collecting information on consumption of Indian households with the purpose of monitoring living standards and guiding the policies of the Planning Commission of the Government of India. The design of these surveys has changed very little over time. Every five years the NSSO runs very large survey rounds (‘thick’ rounds), while surveys of smaller size are implemented every year (‘thin’ rounds). During the large rounds, between 100,000 and 200,000 households are interviewed throughout the country.

The sample of the large rounds is drawn from several strata that allow the analysis of the data at a highly disaggregate level. Separate surveys are conducted in each state and in each state the sample is stratified by season, urban-rural location and district. In rural areas ten households are surveyed in each of the selected villages. Within these ten households, a sub-stratum is composed of affluent households in order to ensure that these are sufficiently represented in the sample. The present study uses rural data of the latest four large survey rounds for the state of Andhra Pradesh. Table 3.1 shows the break-down of the number of observations by survey round and urban-rural location. Given that the state is divided in 23 administrative districts, there are on average between 200 and 250 observations per district in rural areas.

Table 3.1 Number of observations by NSSO survey round

Observations	43^{rd} (1987-88)	50^{th} (1993-94)	55^{th} (1999-00)	61^{st} (2004-05)
Rural households	6,016	4,908	5,181	5,555
Villages	602	492	432	556
Urban households	3,423	3,644	3,806	2,876
All households	9,439	8,552	8,987	8,431

Source: calculated from NSSO data.

The survey rounds follow a common design that has its core in the expenditures modules. An introductory module collects information on household characteristics, which include the social group to which the household belongs, religion, cultivated and irrigated land, main occupation, source of energy for cooking and lighting, and a number of other characteristics that vary by round like: income sources, crop produced, housing conditions, participation in public works and others. A second module (the household roster) collects demographic and education data on all members of the household. A varying number of modules collect extremely detailed expenditure data grouped by broad categories including food, education, health, and durable goods among others.

As discussed in Section 2.4.3, changes made to the survey questionnaire of the 55^{th} round (1999-00) have biased the comparability of the data in two ways. The 30-day food expenditure is biased upward, whilst annual expenditure on non-food items is underestimated. In the rounds following the 55^{th}, the disputed changes were removed and the problem of comparability is therefore limited to the data collected in 1999-00. These affect the results of the analysis conducted in this chapter only marginally, as will be discussed in the next section.

3.4.2 Household expenditure adjustments

The NSSO surveys collect very detailed expenditure data. Quantities and values of purchases of hundreds of food and non-food items are reported, and estimated values of home produced goods are included. This section describes the series of adjustments made in order to carry out the analysis of seasonal consumption patterns. First, household expenditure is divided by the number of adult equivalents rather than by household size. Second, items that are purchased at particular times of the year due to climate or custom are dropped from the computation of household expenditure. Third, household expenditure is divided by household specific price indices in order to correct for spatial and temporal price variation and for differences in consumer preferences.

Comparability of NSSO expenditure surveys

Expenditure data collected by NSSO surveys can be grouped in eight broad categories: food; stimulants and intoxicant; clothing and footwear; miscellaneous goods and services; education; health; and durable goods. Table 3.2 shows the availability of expenditure data for these eight broad categories by recall period in each of the survey rounds.[16] A tick in the table means the data were collected.

Table 3.2 Expenditure categories by recall period and survey round

	43^{rd} round		50^{th} round		55^{th} round[a]		61^{st} round	
	month	year	month	year	month	year	Month	year
Food	✓		✓		✓		✓	
Stimulants & intoxicants	✓		✓		✓		✓	
Clothing and footwear	✓	✓	✓	✓		✓	✓	✓
Misc. goods and services	✓		✓		✓		✓	
Education	✓		✓	✓		✓	✓	✓
Health	✓		✓	✓	✓		✓	✓
Durables	✓	✓	✓	✓		✓	✓	✓

All expenditure categories are reported over a 30-day recall period that is comparable across surveys, with the exception of the 55^{th} round. During the 55^{th} round, expenditure on clothing and footwear, education, and durables were collected only on an annual basis. However, expenditure on education and durable goods will be dropped from the computation of total household expenditure for reasons that will be clarified later, while expenditure on clothing and footwear represent only a small percentage of total expenditure (see Table 3.3). Expenditure on food and clothing was overestimated in the 55^{th} round (as discussed in Section 2.4.3). This will not affect the present analysis, which compares expenditure across groups rather than across surveys.

[16] Only comparable recall periods are reported. In the 55th round data on consumption of food, and stimulants and intoxicant were collected using both a 7-day and a 30-day recall period. As this is the only case in which a 7-day recall period was used, it is not reported in the table.

Table 3.3 Expenditure shares of rural households by broad category

	43rd round	50th round	55th round	61st round
Food	62.1	63.0	57.4	53.9
Stimulant and intoxicants	7.6	7.6	6.2	6.2
Clothing and footwear	6.3	4.3	6.7	3.4
Misc. goods and services	17.5	18.0	20.0	29.4
Education	0.6	0.7	1.0	1.4
Health	4.6	5.5	4.9	4.5
Durables	1.4	1.0	5.8	1.2

Source: calculated from NSSO data.

Demographic adjustments

Several factors affect the demographic composition of households including migration, the cost of raising children, life expectancy and cultural norms. Households of different income, social group or geographical location have different demographic composition, which makes comparisons based on per capita expenditure across different groups of households not appropriate. The reason is that children consume less than adults and that there are economies of scale inside the household for goods that are collectively used. (See Deaton (1997) for a detailed discussion of this topic, and White and Masset (2003) for a discussion of the methods used to adjust household expenditure.)

Table 3.4 Demographic characteristics of farmers and agricultural labourers

Demographic characteristics	Farmers		Agricultural Labourers	
	Irrigated	Non-irr.	Irrigated	Non-irr.
Size	5.0	4.8**	4.0	4.3***
Children under 5	0.4	0.4	0.4	0.5**
Children from 5 to 10	0.9	0.9	0.7	0.9***
Children from 10 to 15	0.4	0.4**	0.3	0.3
Adults	3.1	2.9***	2.4	2.5*
Elderly adults	0.2	0.2	0.2	0.2
Age of head of household	46.2	45.2**	41.7	42.0
Sex ratio	1.2	1.2	1.2	1.2

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.
Source: calculated from NSSO data.

The present analysis focuses on four household groups, namely farmers and agricultural labourers of irrigated and non-irrigated areas. Table 3.4 shows the main demographic characteristics of these four groups. Average age and household size are smaller among labourers compared to farmers. Within farmers, households with irrigated farms are larger in size, have more adult members and the head of household is on average older. Labourers living in irrigated villages tend to have smaller families and fewer children compared to labourers living in non-irrigated areas. Ignoring these different demographic characteristics would lead to making wrong inferences on the welfare differences between groups of households.

In order to compare expenditure data of households of different demographic composition two adjustments were made. First, expenditure is divided by adult equivalent members rather than by household size. The equivalence scale used to calculate the number of adult equivalents is arbitrary and assigns the value of 0.4 to the expenditure of children under age five, 0.7 to children between five and ten, and 0.8 to children between the age of ten and 15. Second, a small arbitrary adjustment for economies of scale is included by raising the number of adult equivalent members of the household to the power of 0.9.

Seasonal adjustments

The NSSO data are particularly effective for the analysis of seasonal expenditure for two reasons. First, the survey is conducted over a 12-month agricultural year starting on the 1st of July and ending on the 30th of June of the following year. The data are stratified by three-month sub-rounds, each composed of an equal amount of households and clusters. Each sub-round is statistically representative of the population at the state level and can be analysed separately from the others. Second, the NSSO questionnaire uses a 30-day recall period for all expenditure items, with the exception of the 55th round, and hence household expenditure can be easily assigned to a particular season.

Seasonal expenditure analysis is better performed after subtracting from total expenditure those items whose fluctuations are not related to income. There are two main sources of non-income related seasonal fluctuations in the quantities demanded of goods: climate and custom (ILO, 2004). Seasonal changes in climate affect, for example, the demand for clothing and health care, while custom determines the times at which certain payments are made, for example school fees, ceremonies and holidays.

In order to assess the extent to which each expenditure category is sensitive to seasonal patterns that are not related to income variations, a series of regressions were run of household expenditure by category on seasonal dummies. The dependent variable in the regressions is the logarithm of household expenditure on each broad category, while the explanatory variables consist of three seasonal dummies and a set of control variables, which include demographic characteristics, socioeconomic status and total household expenditure. The coefficient estimates of the three seasonal dummies are reported in Table 3.5, and represent the percentage change in consumption of each category with respect to the omitted season (July to September). Seasonal coefficients are particularly large and statistically significant for education, health, and durables. Seasonal variation in the consumption of stimulants and clothing is not statistically significant, while the variation in consumption of food and miscellaneous items is very small. In order to avoid using expenditure data that reflect seasonal consumption patterns related to climate and custom, rather than seasonal variation in income, expenses on education, health and durables were excluded from the computation of total household expenditure.

Table 3.5 Seasonal consumption by expenditure category

	Food	Stimulants	Clothing	Misc.	Education	Health	Durables
Oct-Dec	-0.012**	-0.021	0.282	0.038***	-0.625***	-0.451**	1.149***
Jan-Mar	-0.017**	0.056**	0.382	0.026**	-1.642***	-1.328***	1.944***
Apr-Jun	-0.004	0.031	-0.514	0.062***	-3.367***	-1.054***	1.925***
R-square	0.91	0.37	0.13	0.86	0.11	0.04	0.11
Observations	21271						

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.
Source: calculated from NSSO data.

Indian households spend considerable sums for marriages and other festivities. Table 3.6 shows the percentage of households that reported celebrating a ceremony during the month preceding the interview. There is no discernible seasonal pattern in the ceremonies performed. No statistically significant difference was found between groups of households with and without access to irrigation. Ceremonies expenditure was nevertheless excluded from the computation of total expenditure for two reasons. First, this expenditure item was not included in the 55^{th} survey round. Second, the inclusion of ceremonies is a potential source of outliers, because large outlays can be made particularly for marriages.

Table 3.6 Percentage of household celebrating a ceremony by season

Season	43rd round	50th round	61st round
Jul-Sep	3.5	0.7	0.8
Oct-Dec	1.5	1.6	1.6
Jan-Mar	1.7	0.6	1.9
Apr-Jun	1.3	0.8	2.7

Source: calculated from NSSO data.

Price adjustments

In order to analyse seasonal consumption patterns, seasonal price variations have to be removed from the expenditure data. This is performed dividing households expenditure by a price index, which can be obtained in two ways. The first is using the official consumer price index of Andhra Pradesh; the second is calculating the price index using the survey data. In this section advantages and disadvantages of the two indices are discussed and the reasons for choosing the second method instead of the first are explained.

The consumer price index adopted by Indian statistical authorities is based on the Laspeyres formula. The Laspeyres index (L) estimates the change in the cost of living between two time periods (zero and one) by calculating the ratio of the cost of acquiring the same quantities of goods (q^R) in two periods at different prices (p^0 and p^1). The reference basket of quantities is that in period zero, and the index is:

$$L = \frac{p^1 q^R}{p^0 q^R} \tag{3.18}$$

There are two problems with using this index. First, it ignores substitution effects, thus correcting prices in excess of the true change in the cost of living. When relative prices change, consumers shift demand from the good that increased in price to its closer substitute. For example, an increase in the price of coffee will induce an increase in the demand for tea and other stimulants whose price has not changed. This means that part of the negative effect on utility of an increase in price is absorbed by the consumer through a change in consumption patterns. This is the substitution effect of a price change. The Laspeyres index calculates the income effect of a price change, but ignores the substitution effect. As a result, the use of the Laspeyres index overcompensate consumers for the change in utility caused by a price increase. The index however is a very good approximation in cases where all prices change proportionally and therefore changes in relative prices and substitution effects are limited. Similarly, when substitution possibilities are limited, for example because goods are consumed in fixed proportions, the index works perfectly well. In general however, substitution effects are strong, and division of expenditure by this index overestimates the reduction in utility brought about by an increase in prices.

The second shortcoming of the Laspeyres index is that it adjusts the cost of living in different ways for households of different income levels. Consumers' preferences vary with income, and household budget composition varies with total household expenditure. Typically, poorer households consume more food, while richer ones consume a variety of non-food items. The implication is that changes in relative prices affect households in different ways depending on the composition of their budgets. The Laspeyres index calculates budget shares using the average consumption of all households, thus assuming that all households are equal. Note that since expenditures of rich households have more weight in the calculation of the overall budget shares, these shares tend to over-represent the budget composition of richer households. For example, an increase in food prices might have little consequence on the value of the index because at the country level the food share is rather small, though the negative impact on the utility of the poor is large. It follows that the Laspeyres index compensates for the reduction in utility determined by changes in prices by giving more weight to changes in prices of goods consumed by the rich. Dividing household expenditure by this index may therefore overestimate or underestimate the change in utility depending on which prices are increasing and which household is considered.

The Government of India publishes monthly series of the consumer price index for the whole of India and for each Indian state separately. Two different series are published, one for the industrial workers (the CPIIW), and a second for the agricultural labourers (the CPIAL). The CPIAL is representative of price variations in rural areas, and is therefore more relevant for the present study.

The CPIAL is obtained using the following methodology. First, prices of a well specified set of food and non-food items are collected at the village level. Second, village prices for each item are averaged within the same enquiry zone thus producing a zonal price for each item. A zone consists of a number of contiguous villages. Third, the zonal prices are aggregated at the state level to produce a state price for each item. Zonal prices are weighted by the expenditure share of each zone in the state. Finally, a monthly state index is calculated using the Laspeyres formula.

Until 1995, data for the index were collected by the Field Operation Division of the NSSO on the 1st week of every month. Prices were always collected from the same shops and markets and for the same items. The sample was composed of 15 states, 39 Agricultural Labour Enquiry Zones, and 422 fixed sample villages. The items surveyed ranged from 27 to 38, including food, *pan* and tobacco, fuel, clothing and miscellaneous. The weights used in the calculation of the index were obtained in the Second Agricultural Labour Enquiry of 1956-57 and were never revised since. The index base was fixed at the agricultural year 1960-61.

Since 1995, the Ministry of Labour has published series of the CPIAL calculated using a substantially revised methodology. Data collection is still carried out by the Field Operation Division of the NSSO every month, but prices are now collected over a period of four weeks, rather than in the 1st week of the month, and the sample size has been expanded. Data are now collected in 20 states, divided in 66 enquiry zones, and 600 villages (of which 54 are in Andhra Pradesh). Prices are now taken for a larger number of items (up to a maximum of 106). The weights for the calculation of the Laspeyres index have been updated using data of the NSSO expenditure survey of 1983-84. The index base has been reset to the agricultural year 1985-86. The two series can now be linked using a multiplier provided by the Ministry of Labour and published on the Indian Labour Journal (MOL, several issues).[17]

The new series of the CPIAL obviate some of the problems of the old series, but the use of this index is still problematic in some respects. First, series prior to 1995 are still based on outdated expenditure weights calculated in the 1950s. Second, although the new series have updated the weights using more recent survey data, these are not periodically modified to account for changes in consumers' preferences. Third, the

[17] The value of the multiplier for Andhra Pradesh is 4.84. In order to obtain figures of the CPIAL after 1995 using the old base of 1960-61, the data need to be multiplied by this factor.

ability of the CPIAL to represent regional price variation is questionable. Finally and most importantly, the CPIAL relies on the Laspeyres formula.

Table 3.7 Variations in food prices by district and season

Districts	Paddy	PDS rice	*Jowar*[a]	*Ragi*[a]	*Dal*	Milk	Tomato
Nizamabad	-0.07***	0.00	0.30***		0.05***	0.05**	0.05*
Karimnagar	-0.01	0.00	0.24***		0.05***	0.09***	0.22***
Medak	-0.05***	0.00	0.22***		0.08***	0.12***	-0.01
Rangareddy	-0.01	-0.01**	0.34***	0.68***	0.08***	0.23***	-0.16***
Mahaboobnagar	-0.06***	-0.01**	0.27***	0.18	0.06***	0.10***	-0.05**
Nalgonda	-0.06***	-0.01	0.16***		0.03***	0.18***	-0.03
Warangal	0.00	-0.01**	0.07*		0.13***	0.15***	0.25***
Khamman	0.01	-0.01	0.46***		0.09***	0.20***	0.33***
Srikakulam	-0.13***	0.00		-0.35*	0.08***	0.18***	0.25***
Vizianagaram	-0.12***	0.00		-0.40**	0.07***	0.22***	0.18***
Vishakhapatnam	-0.04**	-0.01		-0.23	0.08***	0.29***	0.08***
East Godavari	0.03*	-0.01*	0.36***	0.38**	0.14***	0.15***	0.15***
West Godavari	0.07***	-0.01**		0.50**	0.12***	0.22***	0.23***
Krishna	0.08***	0.00	-0.02	0.29	0.14***	0.42***	0.12***
Guntur	0.16***	0.01**			0.05***	0.49***	0.15***
Prakasam	0.11***	0.03***	0.24***	0.15	0.12***	0.25***	0.07***
Nellore	0.03	0.01**		0.40**	0.15***	0.24***	-0.12***
Cuddapah	0.03	0.00	0.31***	0.34*	0.06***	0.00	-0.09***
Kurnool	0.00	-0.01***	0.33***		0.05***	0.08***	-0.26***
Anantapur	-0.13***	0.01**	0.19***	0.07	0.03***	-0.11***	-0.14***
Chittor	0.02	0.00	0.27*	0.27	0.08***	-0.02	-0.26***
Coefficient of variation	0.06	0.01	0.09	0.25	0.03	0.10	0.15
Oct-Dec	0.02***	0.00	0.01	0.03	0.04***	0.01	-0.21***
Jan-Mar	0.02***	0.00	0.02	0.01	-0.04***	0.01	-0.53***
Apr-Jun	0.05***	0.00	0.09***	0.03*	-0.04***	0.05***	-0.31***

Notes: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%. Coefficient estimates are not available for those districts where the commodity is not consumed.

Poor spatial integration of markets and high transport costs determine large price differences across locations. This puts into question the representativeness of the sample of 54 villages used for the calculation of the CPIAL in Andhra Pradesh. Table 3.7 shows the results of regressions of the logarithm of price for some essential food items in Andhra Pradesh against district and seasonal dummies. The regressions use data of 556 villages, with ten price observations per village, from the NSSO survey of 2004-05.[18] The estimated coefficients measure the percent price difference in each district and season. The coefficients of variation of the estimated parameters of the district dummies

[18] The methodology for the analysis of spatial variation of prices draws on Deaton (1997).

reported at the bottom of the table range from 0.05 to 0.25, which point to substantial variation in food prices across districts. The coefficients of seasonal dummies are smaller than district coefficients, meaning that the spatial variation in prices is larger than the temporal variation. The analysis of price variation by item reveals that the variation in the price of rice is modest compared to other goods. This is particularly true for the rice sold through the PDS, whose price is controlled by the state government. Spatial price variations are quite large for *jowar* (Sorghum), *ragi* (Millet), milk and tomatoes. The first two are goods that are traditionally marketed locally, while the other two are perishable and their transportation is difficult.

In order to avoid the shortcomings outlined above and those related to using a price index based on the Laspeyres formula, the present study uses a price index calculated for each survey round, rather than the CPIAL. Household specific price indices were calculated using the methodology described in Deaton and Zaidi (2002). There are two main advantages in using household specific price indices. The first is that they adjust expenditure for temporal and spatial price variation simultaneously. This is particularly important given that the spatial variation was found to be larger than the temporal variation. The second advantage is that household specific indices are tailored to each household, thus reflecting individual household's income and preferences, and do not overstate or understate changes in the cost of living in the way that the Laspeyres index does.

The starting point for the calculation of household price indices is the calculation of a household Paasche price index (P^h):[19]

$$P^h = \frac{p^h q^h}{p^0 q^h} \tag{3.19}$$

The superscripts indicate that prices and quantities are measured at the household level (h). This index can be rewritten in the form:

[19] Paasche price indices are analogous to Laspeyres indices, but unlike the latter they refer to a basket of goods in the present rather than in the past.

$$P^h = \left[\sum_i^n w_i^h \frac{p_i^0}{p_i^h} \right]^{-1} \tag{3.20}$$

where w_i are the shares of household's budget spent on the *i* item. For the purpose of practical computation, the index can be approximated by:

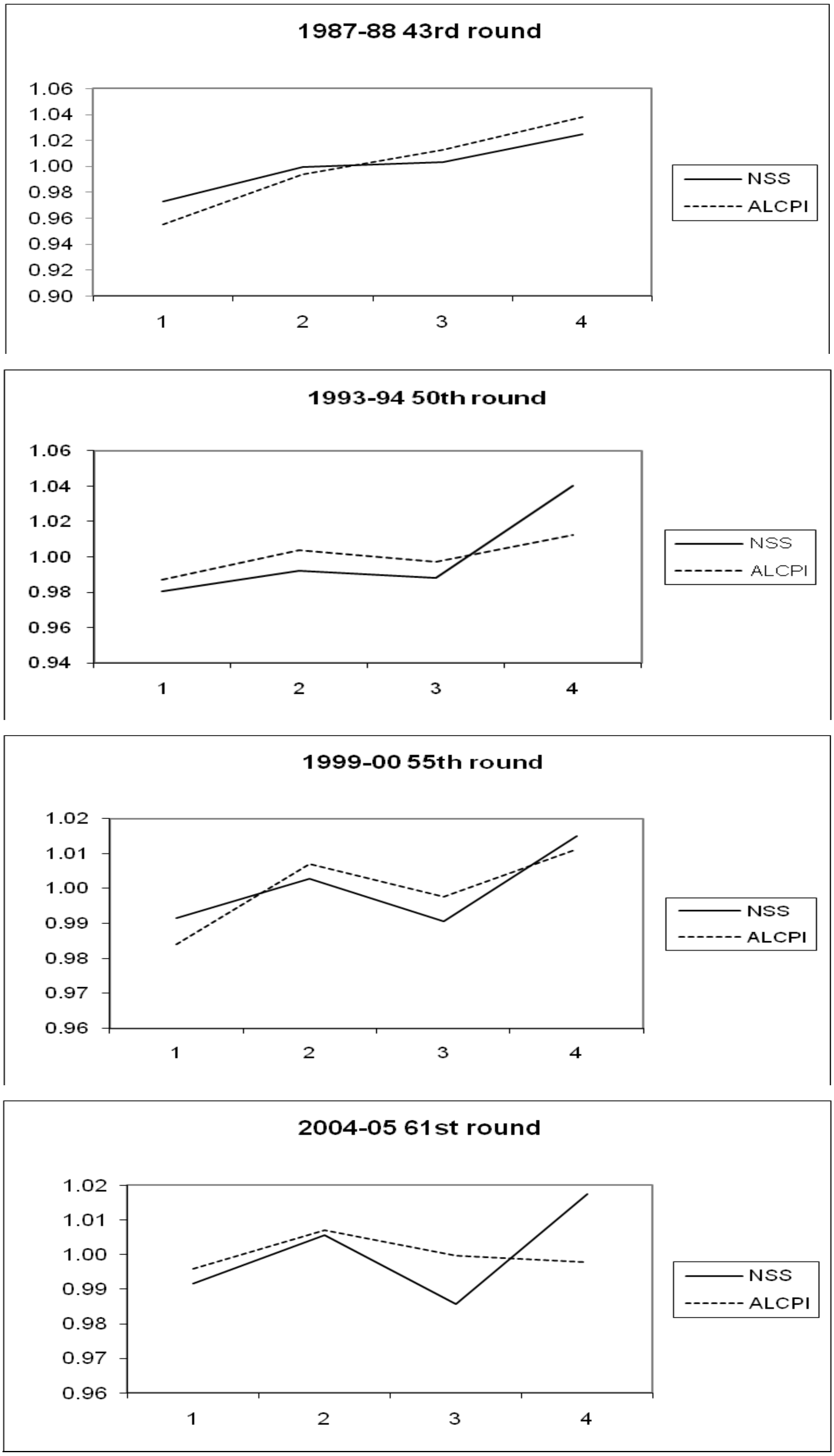
$$\ln P^h = \sum_i^n w_i^h \ln\left(\frac{P_i^h}{P_i^0} \right) \tag{3.21}$$

In practice household indices were calculated in the following way. First, all goods that could be included in the index where selected. For example, in the case of the 61st survey round, 97 items were selected and subdivided in the following categories: food (76); tobacco and *pan* (6); intoxicants (3); fuel and light (7); clothing and bedding (4); footwear (1). Items whose prices could not be calculated were dropped. This was the case where quantities consumed were not uniquely defined, or not comparable, or did not represent a homogeneous category. Several items that were rarely purchased, and those whose consumption is highly seasonal, were also dropped from the computation. Secondly, the unit values were derived by dividing the value of household consumption on an item by the quantity consumed of the same item. The shares of the expenditure of each item on total expenditure at the household level were also calculated. When unit values took unreasonable values (+/- 2.576 standard deviations from the mean), they were substituted with the mean unit value in the district where the observation was collected. Third, and finally, the median price in the cluster was taken as the price paid by the household. This rests on the assumption that prices for the same item do not vary substantially within the same village. The mean price over the sample of all households was taken as the base price. These two prices were inserted in formula (3.21), and the price indices were calculated for all household in the survey.

This method of calculating price indices is not without drawbacks. First, for practical reasons, the index can be calculated only over a limited number of goods, like food, clothing and miscellaneous, leaving out several non-food items. As household expenditure is geared towards the consumption of more non-food items over time, the index becomes increasingly less representative of the true cost of living. Second, the index is not based on true prices, but on unit values obtained by dividing the value of consumption by the quantity consumed by each household. Unit values do not take into account variations in quality. Rich household pay more for the same items because buy higher quality products, and richer villages buy better quality products than poorer ones. Consequently, price differences between households and villages are partly due to differences in the quality of the purchased goods. Using deflators based on unit values and interpreting differences in quality as differences in prices reduces the expenditure value of the rich and increases that of the poor. This happens because the expenditure of rich households is deflated by a high price index, which mistakenly takes the difference in the quality of the good purchased as a difference in the price paid by the consumer. Conversely the expenditure of the poor is deflated by a price index that is too low, because it mistakenly confuses low quality with low prices.

The official CPIAL of Andhra Pradesh and the household specific price indices calculated from survey data are compared in the charts of Figure 3.1. The CPIAL was averaged by trimester and normalised to one in order to make the two indices comparable. The household specific price indices were also normalised to one and averaged by season. There is a good deal of similarity in the seasonal patterns of the two indices: prices increase in the second season, decrease in the third, and increase again in the fourth. The main difference between the indices emerges in the last season, where the two give very different prices for two of the four rounds.

Figure 3.1 Seasonal price indices calculated from NSSO data and CPIAL



Source: solid line calculated from NSSO data, dotted line obtained from *indiastat*.

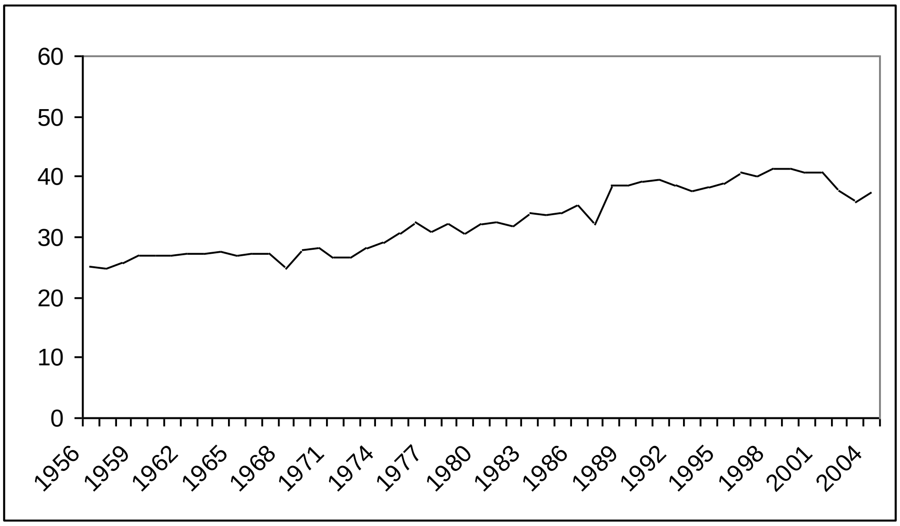
3.5 Descriptive statistics

This chapter assesses the impact of irrigation on the consumption of farmers and agricultural labourers with and without access to irrigation. Classifying households in these four categories is not a straightforward exercise and this section discusses in detail the methodology used to produce the household classifications used in the empirical analysis. This section also presents some descriptive statistics on irrigation, rainfall patterns, and agricultural seasons in Andhra Pradesh.

3.5.1 Irrigation in Andhra Pradesh

Andhra Pradesh is well endowed with water sources having four large rivers: Godavari, Krishna, Pennar and Vamsadhara. Despite large investments in the construction of canals and minor irrigation structures however, the irrigated area has reached less than 50% of the irrigation potential of the state (Reddy, 2003). According to DES data (DES, 2005), the percentage of net cropped area under irrigation has slowly increased from 25% in 1956 to nearly 40% in 1989. Since then, the amount of irrigated land has remained more or less stable at just below 40%. Percentages of irrigated area calculated using NSSO data are very similar to those reported by the DES: 36% in 1987-88; 42% in 1993-94; 43% in 1999-00; and 36% in 2004-05 (see Figure 3.2).

Figure 3.2 Percentage of cultivated area under irrigation (1956-2005)



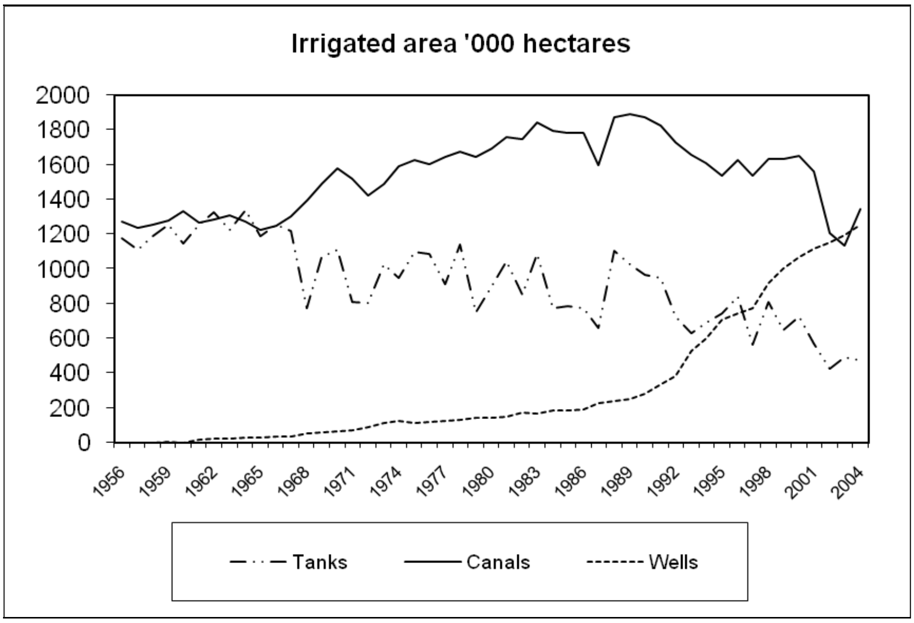
Source: calculated from DES (2005).

Since the mid 1950s, important changes have occurred in the types of irrigation used (see Figure 3.3). The use of traditional tanks has decreased steadily since the mid 1960s, while the use of canal water has diminished during the last 15 years. This reduction has been entirely compensated by the expansion in the use of ground water. Irrigation by means of water pumps is now as important as canal irrigation.

The area covered by canal irrigation over the period 1956-2005 reflects state investments in irrigation infrastructure. These increased largely in the 1970s, remained stable during the 1980s, and dropped in the 1990s (Reddy, 2003). Interestingly, the area irrigated by canals actually decreased in absolute terms over the last ten years, probably due to poor operation and maintenance. In the late 1990s, new institutional arrangements (Water Users Associations) were introduced in order to promote a better operation and maintenance of canals, but these do not seem to have prevented the deterioration of the existing infrastructure.

Three main factors explain the diffusion of ground water extraction. First, in the 1980s a new generation of water pumps that were both simple to operate and relatively inexpensive became available. Second, rising living standard in rural areas made this technology affordable even to very small farmers. Finally, the state government granted electricity free of charge to a large portion of rural areas in the state, thus subsidising the operation of water pumps.

Figure 3.3 Irrigated area by irrigation source (1956-2005)



Source: calculated from DES (2005).

There are two important consequences of the changes in the main sources of irrigation. First, the shift from tanks and canals to private wells has decreased farmers' dependence on total rainfall for irrigation. Tank irrigation consists of water collected in reservoirs that range from less than one acre to several hectares in size, and is entirely dependent on local rains. The availability of canal water depends on total rainfall in the area of the main reservoir and is not assured in times of drought. In contrast, ground water is always available, at least until the water table is exhausted, which requires prolonged or consecutive droughts. The shift towards the use of ground water against more traditional sources of irrigation has therefore in principle increased the stabilisation role of irrigation. It should be emphasised however that although ground water is a more reliable source than canals, this is strictly true only in areas were rainfall is abundant, though irregular. In dry land areas, ground water extraction reaches its maximum potential in a limited time, and is not a reliable source of irrigation. In addition, ground water extraction has benefited so far from the access to large water reservoirs. However, the extraction of ground water in the long term might be hampered by the extinction of these reservoirs or by their deterioration due to soil erosion. In conclusion, while ground water extraction has provided a good short-term solution to the shortage of water for a large number of farmers in Andhra Pradesh, it cannot be considered a reliable source of irrigation in dry land areas nor a sustainable source of irrigation in any area.

A second implication of the shift in irrigation sources is a potential change in the distribution of irrigation over geographic regions and farm classes. Private wells might have brought irrigation to areas not served by canals and to farmers that can afford the purchase of water pumps. The analysis of NSSO data however indicates that this was not the case. Table 3.8 shows the fraction of net irrigated area by district across the four surveys. It appears that in the districts of the coastal belt traditionally irrigated by canals (East and West Godavari, Krishna and Guntur), the share of irrigated area has not decreased but possibly increased over the 20 years considered. On the other hand, districts traditionally not served by canal water in the south of the state (like Anantapur, Kurnool, and Mahaboobnagar) do not show any significant increase in the share of net irrigated area.

Table 3.8 Percentage of irrigated area by district (1987-2005)

District	1987-88	1993-94	1999-00	2004-05
Srikakulam	63.2	50.5	47.3	62.2
Vizianagam	54.1	42.7	56.2	44.1
Vishakhapatnam	24.7	32.7	27.6	25.5
East Godavari	65.8	54.0	69.8	76.6
West Godavari	79.1	79.4	78.4	86.2
Krishna	83.9	76.7	98.0	94.9
Guntur	56.0	51.9	77.9	57.5
Prakasam	30.3	22.0	38.8	53.0
Nellore	69.0	74.0	70.4	57.3
Chittor	42.6	40.2	40.5	26.9
Cuddapah	25.9	28.7	57.1	48.0
Anantapur	15.1	20.0	25.2	11.1
Kurnool	11.6	20.4	13.6	11.8
Mahaboobnagar	18.2	16.4	13.3	11.3
Rangareddy	21.7	35.0	17.9	52.1
Medak	29.8	56.9	55.7	34.0
Nizamabad	51.8	53.0	75.5	59.0
Adilabad	11.2	11.2	22.1	9.3
Karimnagar	50.8	60.8	70.8	40.3
Warangal	45.8	42.5	47.7	30.7
Khamman	34.0	35.2	31.5	69.5
Nalgonda	33.7	26.8	28.2	21.7

Source: calculated from NSSO data.

Similarly, the shares of irrigated land by land holding classes indicate that the expansion of the use of ground water and the reduction in the use of canal and tank water has not brought about substantial changes (see Table 3.9).

Table 3.9 Percentage of irrigated area by farm size

	1987-88	1993-94	1999-00	2004-05
Marginal farms	54.7	54.5	52.9	45.4
Medium farms	41.6	40.8	46.8	39.2
Large farms	30.2	30.3	37.7	31.7

Note: marginal farmers cultivate less than one acre; small farmers cultivate between one and five acres; large farmers cultivate five acres or more.
Source: calculated from NSSO data.

3.5.2 Household classification

This study focuses on two types of rural households, farmers and agricultural labourers, because they are the ones more directly affected by weather related income fluctuations. Abstracting from prices, farmers' income is affected by fluctuations in agricultural output, while income of agricultural labourers is affected by fluctuations in labour demand by farmers. Fluctuations in output and labour demand are clearly related, but do not necessarily happen at the same time and season. For example a poor output for farmers in the harvesting season may be preceded by a poor labour demand in the sowing season. This requires a separate analysis of the two household classes.

Any subdivision by social classes is an oversimplification of rural society. In reality, many farmers do some paid work for other farms, while many labourers cultivate small plots of land. This section describes the methodology used by the NSSO to differentiate rural classes, highlighting its shortcomings, and outlines the methodology used in this study.

The NSSO subdivides households in five categories by main economic activity (see Table 3.10). The methodology used is rather complex and has been modified since the 55^{th} round (NSSO, 1999). Up to the 50^{th} round the classification was produced in two stages (NSSO, 1993): first at the level of household listing, and then at the stage of the household interview. At the time of listing the enumerator would take note of net household income under four categories: self-employment non-agriculture (y_1), self-employment agriculture (y_2), wage manual labour (y_3), and wage non-manual labour (y_4). The enumerator would then classify households in the following way: rural labour household (code two) if $y_3>(y_1+y_2)$, and $y_3> y_4$; self-employed non-agriculture (code one) if $y_1>y_3$, and $y_1>(y_2+y_4)$; other households (code nine) all households not otherwise classified. At the time of the survey interview the enumerator would code one the households classified self-employed non-agriculture during the listing, and would further split the remaining two categories. Rural labour would become agricultural labour (code 2) if more than 50% of income was obtained from wages earned in agriculture, and would become other labour (code 3) otherwise. Other households would become self-employed agriculture (code 4) if more than 50% of income was obtained from self-employment in agriculture, and simply ‘other’ (code 5) otherwise.

Table 3.10 NSSO classification of rural households

NSSO household category	43rd round (%)	50th round (%)	55th round (%)	61st round (%)
(1) Self-employed non-agriculture	13.8	12.8	12.7	18.3
(2) Agricultural labour	39.6	37.8	45.2	37.1
(3) Other labour	8.6	5.8	5.7	8.5
(4) Self-employed agriculture	27.3	24.2	24.6	27.4
(5) Other	10.7	6.9	11.8	8.7
Unclassified		12.5		

Source: calculated from NSSO data.

Since the 55th round, the process of classification has been reduced to one stage and households are assigned to the classes of Table 3.10 at the time of the interview. Each household is assigned to any one category if it earns 50% or more of its income from the corresponding economic activity. If the household cannot be classified based on these criteria, the enumerator revises the household income sources and classifies households in the following way: self-employed non-agriculture if $y_1 > y_2$, and $y_1 > (y_3 + y_4 + y_5)$; other labour if $y_2 > y_1$, and $y_2 > (y_3 + y_4 + y_5)$; and ‘other’ in all other cases, including households that do not report any income.

The NSSO household classification has several limitations. First, the methodology used has changed between rounds, and the household types defined are not strictly the same. Second, this classification method is based on income, which is a complex estimate – more so than consumption – and its calculation is a fairly lengthy process. It is unrealistic that households can provide reliable income estimates in the short time available for the interview. Third, the methodology used by NSSO can lead to inconsistencies, and the same household may end up in more than one category. Finally, in the 50th round the households coding for Andhra Pradesh contains many missing values, and 12% of the rural sample is unclassified.

In order to overcome these limitations this study uses the households classification based on the National Classification of Occupation codes (NCO) instead of that of the NSSO. For the NCO classification, the enumerator assigns to each household one occupation code out of more than 100 available. In order to do so, the enumerator lists all the occupations of all household members during the year preceding the interview. The respondent is then invited to indicate the occupation, among those listed, that provided the largest amount of income during the previous year. In the case where two or more occupations were equally important in terms of income, the occupation of the most senior member of the household is chosen as the household occupation. This classification is preferable to that of the NSSO for its simplicity, and because it allows the respondent to indicate the class to which the household belongs.

Another important classification employed by this study is the distinction between irrigated and non-irrigated farms, and between labour employed in irrigated and non-irrigated areas. In building separate samples of irrigated and non-irrigated farms, a complication arises from the fact that the land of irrigated farms is irrigated in different proportions. While land is entirely irrigated in some farms, the majority of farms are only partially irrigated. Restricting the category of irrigated farms only to those with 100% of the land irrigated is not appropriate, because large farms with only a fraction of their land irrigated may have, in absolute terms, larger areas of irrigated land than small farms that are 100% irrigated. Therefore all farms with irrigated land, however small, are considered as irrigated for the purpose of this study. In order to build separate samples of labourers in irrigated and non-irrigated areas, villages with 100% of cultivated land irrigated are defined as irrigated villages. This definition is arbitrary because, as in the case of the identification of irrigated farms, villages with a smaller share of irrigated land might have a larger absolute irrigated land. However, if all villages with any share of irrigated land were considered irrigated areas, almost all villages in the survey would be classified as such. Therefore a more conservative definition of irrigated areas has been adopted for the purpose of this study.

In the empirical part of the chapter a distinction will be made between irrigated and non irrigated farms and between agricultural labourers from irrigated and non irrigated villages. Farms are considered as irrigated if any share of any size of cultivated land is irrigated by any means at the time of the survey interview. Agricultural labourers are considered as living in irrigated villages if 100% of cultivated land reported in the survey is irrigated in the cluster of residence. Table 3.11 shows the percentages of farmers and of agricultural labourers in the sample, and the share of each category that has access to irrigation. Farmers account for about 25% of the rural population, and between 65% and 75% of the farms are irrigated. These figures are only apparently in contradiction with the much smaller share of total irrigated land in the state (40%), because this is a sub-sample of farmers who have chosen farming as their main economic activity, and because the share of farms with some irrigation in a given area may exceed the share of total irrigated land in the same area.

Table 3.11 Percentages of farmers and agricultural labourers

	43rd round	50th round	55th round	61st round
Farmers	25.8	26.7	24.4	26.9
of which with irrigation	77.6	70.2	66.4	66.2
Labourers	40.8	43.3	45.6	37.6
of which in irrigated villages	18.0	20.8	29.9	25.8

Source: calculated from NSSO data.

3.5.3 Rainfall and agricultural seasons

Dividing the calendar year in agricultural seasons is not a simple task. First, the different regions of Andhra Pradesh are affected by the two monsoons at different times and with varying intensity. Second, households cultivate a multiplicity of crops that have different production cycles. Possibly, the best way to define agricultural seasons is considering the most important rains in the majority of the state and the cropping pattern of paddy rice, the dominant crop, along with some other important crops. This leads to the identification of three agricultural seasons (*Kharif, Rabi* and summer). This section discusses how these three seasons fit in the four-season classification adopted by the NSSO.

The southwest monsoon is the most important source of rainfall for Andhra Pradesh, bringing more than 75% of total annual rainfall. Rains of the northeast monsoon are not only less abundant, but also more erratic. As already mentioned, the two monsoons give rise to the two cropping season of *Kharif* and *Rabi*. The amount of rainfall in each season determines the area sown by farmers and the agricultural yields. Therefore, *Kharif* is the most important agricultural season in all areas of Andhra Pradesh.

Rice is mostly planted in the period between May and June, and harvested between November and December. Depending on rains, and on the availability of irrigation, a second crop is possible in *Rabi*, with sowing in the months of December and January, and harvesting between April and May. Harvesting of main cash crops like groundnut, castor, cotton, sugarcane and chilli takes place in the months between November and

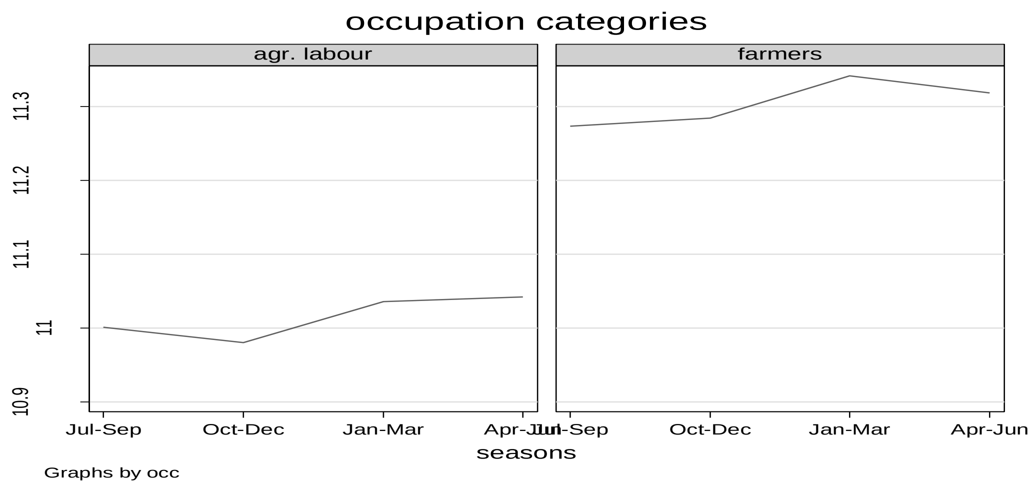
March. Occasionally a summer season is possible, in which *jowar* (sorghum), vegetables, and summer pulses are produced between March and June. The NSSO data subdivide the agricultural year in four trimesters from July to June of the following calendar year. The first trimester of the NSSO data is exclusively sowing time. No income is generated by farmers during this time, while agricultural labourers earn wages for planting operations. In the second trimester harvesting begins for rice and other cash crops, while *Rabi* sowing takes place with the arrival of the northeast monsoon. In the third trimester, the harvesting of *Kharif* rice continues and the harvesting of *Rabi* rice and of the most important cash crops take place. The fourth trimester may see the continuation of harvesting of *Rabi* rice and cash crops, and possibly of summer crops. Table 3.12 shows the distribution of seasonal agricultural activities according to the seasonal classification of the NSSO.

Table 3.12 Agricultural activities by season and main crop

	Rice	Cash crops	Other crops
Jul-Sep	Sowing	Sowing	Sowing
Oct-Dec	*Kharif* harvesting begins/ *Rabi* sowing	Harvesting begins/*Rabi* sowing	Harvesting
Jan-Mar	*Kharif* harvesting continues/ *Rabi* harvesting	*Kharif* harvesting continues/ *Rabi* harvesting	Harvesting
Apr-Jun	*Rabi* harvesting continues	*Rabi* harvesting continues	Summer harvesting for irrigated dry crops and summer pulses

Since agricultural incomes are shaped by agricultural seasons, and since household consumption tracks income to some extent, the passing of seasons is discernible in household consumption patterns. Figure 3.4 shows the average value of the logarithm of per adult equivalent expenditure of rural households by season. The charts were produced using data from all four survey rounds, normalising sampling weights to one within each survey and adjusting expenditure values across survey rounds using the official CPIAL of Andhra Pradesh. The expenditure values within each survey were adjusted for temporal and regional price differences using the methodology described in Section 3.4.2. Averages are shown separately for agricultural labourers and farmers. The following can be observed. First, seasonal expenditure fluctuations are rather small. The lines on the graph are in logarithmic scale, so that the numerical difference between two points is a percentage difference. The maximum variation over the year does not exceed 10%, and the difference between contiguous seasons never exceeds 7%. Second, there are some visible seasonal patterns. Expenditure is larger in the second half of the year when most of the harvesting is taking place. Third, farmers have much higher average expenditures than labourers (up to 30-40% higher). Fourth, labourers' expenditure fluctuations are slightly flatter than those of farmers.

Figure 3.4 Seasonal consumption of agricultural labourers and farmers



Source: calculated from NSSO data.

3.6 Econometric methods

The empirical analysis for this study estimates the difference in seasonal consumption between households (farmers and agricultural labourers) with and without access to irrigation. After adjusting for regional and temporal variation in prices, estimates of the seasonal difference in consumption patterns can be obtained in the following way:

$$c_i = a + \sum_j b_j Z_{ji} + s_2 + s_3 + s_4 + s_{2irr} + s_{3irr} + s_{4irr} + e_i \quad (3.22)$$

Household' consumption (c_i) is a function of a set of control variables (Z_j), and dummy variables for three seasons (the first season is omitted) for the non-irrigated sample (s) and the irrigated sample (s_{irr}). This specification is used to test the difference of seasonal consumption patterns of households with and without access to irrigation.

The empirical analysis also provides estimates of the effect of expected income variability on current consumption as it was modelled in Section 3.3.3 (equation 3.17). Assuming that seasonal preferences, time preferences, and interest rates are the same for irrigated and non-irrigated farms, the difference in consumption between households with and without irrigation in any season will depend on income in that season and the expectation of income in the following season. The empirical analysis uses current rainfall (r) to approximate current and expected income, and the historical variation in rainfall (σ_r) in order to approximate the variability of expected income. The model estimated is:

$$c_s = a + \sum_j b_j Z_{ji} + dr + g\sigma_{r(s+1)} + d_{irr} r_{irr} + g_{irr}\sigma_{r(s+1)irr} + e_i \qquad (3.23)$$

where c_s is household expenditure in the current season, while $\sigma_{r(s+1)}$ is expected income variability in the following season. The subscript *irr* tells whether the household has access to irrigation. The choice of the seasons in the empirical implementation of the model, and the construction of the rainfall variables is described in Section 3.7.1.

Both specifications of the model (3.22 and 3.23) suffer from two endogeneity problems. There is endogeneity in the availability or access to irrigation, and in the occupational choice of farming or working as wage labour. The first form of endogeneity overstates the impact of irrigation on consumption, while the second tends to understate the impact of irrigation. These two endogeneity issues are now discussed in turn for farmers and agricultural labourers separately.

Whether a farm is irrigated or not is largely endogenous. Irrigation might be considered exogenous if it were provided to farms by area only, as in the case of canal irrigation. Even in this case however, it could be regarded as endogenous to the extent that some farmers, with specific characteristics, might choose to buy or might inherit land in irrigated areas. Moreover, some farmers, for example the most influential and powerful, might have access to irrigation in irrigated areas while others in the same areas do not. Even allocation of irrigation by area could be endogenous, if irrigation investments are made in areas where incomes are already more stable than elsewhere. This might be the case if canal irrigation is provided to areas that are already better off in terms of geographic characteristics and infrastructure (Palmer-Jones and Sen, 2003). More serious endogeneity problems arise if farmers irrigate their farms using water pumps, because farmers investing in water pumps are likely to be different from other poorer farmers with respect to many other characteristics. In particular, farmers with water pumps might not be liquidity constrained. Although the results of the theoretical model of Section 3.3 hold even in the absence of borrowing constraints, it might be argued that these farmers have more assets that can be sold in bad times, such as seasonal downturns.[20] In all these cases, the endogeneity of irrigation implies an overestimation of the impact of irrigation on consumption stability, as the stabilising role of other factors is erroneously attributed to irrigation.

The occupational choice of farming is largely endogenous. Seasonal consumption is a function of expected income and of its variance. The variance of future income however, is not fully independent of water availability. This is true both in dry and wet areas. In dry areas households make production choices to reduce income instability. They may have multiple income sources, and grow a more diversified range of crops in order to reduce total output variability. Conversely, households in wet areas may specialise in farming and in the production of a limited number of crops. Indeed, earlier studies on the impact of the green revolution pointed to a potential increase in income instability rather than a decrease (Hazell, 1984). In this case the impact of irrigation on income and consumption stability is underestimated, because households in irrigated areas may specialise in more risky enterprises, while dry farms diversify their mix of activities and crops in order to reduce the variance of future income.

Agricultural labourers' choices of residence and occupation are also endogenous because they are associated to characteristics that are correlated with per capita expenditure patterns. Household location in irrigated or non-irrigated villages was nevertheless considered exogenous. This implies that the village of residence of an agricultural labour household is independent of that village being irrigated. This in turns

[20] On the other hand, the theoretical model of Section 3.3.3 shows that irrigation makes farmers less self-liquidity constrained, in the sense that having higher expectations about future income, allows them to more easily borrow without fearing inability to repay. An example of this behaviour is a farmer who borrows money from a trader in *Kharif* with the assurance of returning rice in *Rabi*, an option that is not so easily available to a non-irrigated farm. In this way irrigation makes farmers less credit constrained.

rules out migration from non-irrigated villages to irrigated ones, or postulates that migrants are not substantially different in characteristics from non-migrants. Though both hypotheses are rather untenable, they were nevertheless maintained for two main reasons. First, migration from irrigated to non-irrigated areas is limited. As mentioned in Section 2.4.1, data from the population census of Andhra Pradesh (Office of the Director of Census Operations, 2005) show that only 6% of the total population lives in a district different from the district of birth. It might be objected that this does not prevent labourers from migrating from dry to irrigated villages within districts. The truth is that very little is known about rural to rural migration in Andhra Pradesh. The qualitative evidence available supports the hypothesis that this type of migration is limited, though seasonal migration to irrigated areas can be substantial (Gopal, 2004). The second reason for maintaining this hypothesis is that the choice of residence is very hard to model with the limited number of variables available in the NSSO data. The consumption model of agricultural labourers therefore is only adjusted for the selection bias produced by the choice of occupation, but not for the choice of residence.

In order to overcome these endogeneity problems, a variation of the Roy model (Cameron and Trivedi, 2005) was used for the econometric analysis. In the classical Heckman selection model the dependent variable is observed only for a sub-sample of the total number of observations, for example wages among the female labour force. In our model however, the dependent variable (household expenditure) is observed for all the observations, but under two states that are mutually exclusive: the irrigated and the non-irrigated farms. A selection bias arises because the choice of the state (irrigation) is correlated with the outcome (consumption pattern). The same household cannot be observed under the two states and the selection bias needs to be corrected in order to be able to compare the two groups.

The version of the Roy model adopted here is the one Maddala (1983) defines ‘the switching regression model with endogenous switching’. The sample of irrigated and non-irrigated farmers will provide an illustrative example. A latent variable I_i^* determines whether the farm is irrigated (I_i=1) or non-irrigated (I_i=0) depending on certain Z characteristics:

$$I_i = 1 \quad \text{if} \quad \gamma Z_i \geq u_i \tag{3.24}$$

$$I_i = 0 \quad \text{if} \quad \gamma Z_i < u_i$$

The consumption model for the irrigated and non-irrigated farms takes the following form:

Irrigated farms: $c_i = \beta_1 X_{1i} + u_{1i}$ if and only if $\gamma Z_i \geq u_i$ (3.25)

Non-irrigated farms: $c_i = \beta_2 X_{2i} + u_{2i}$ if and only if $\gamma Z_i < u_i$

What is needed is the estimation of the expected values of the residuals u_{1i} and u_{2i}, which can be obtained in the following way:

$$E(u_{1i} \mid u_i \leq \gamma Z_i) = -\sigma_{1u} \frac{\varphi(\gamma Z_i)}{\Phi(\gamma Z_i)} \tag{3.26}$$

$$E(u_{2i} \mid u_i \geq \gamma Z_i) = \sigma_{2u} \frac{\varphi(\gamma Z_i)}{1 - \Phi(\gamma Z_i)}$$

The expected residuals are then substituted in the original equations. In order to simplify the substitution two new variables are defined (the inverse Mill's ratios):

$$W_{1i} = \frac{\varphi(\gamma Z_i)}{\Phi(\gamma Z_i)} \tag{3.27}$$

$$W_{2i} = \frac{\varphi(\gamma Z_i)}{1 - \Phi(\gamma Z_i)}$$

Then the original equations can be rewritten including the inverse Mill's ratios and be estimated using OLS:

$$c_{1i} = \beta_1 X_{1i} - \sigma_{1u} W_{1i} + \varepsilon_{1i} \quad \text{for} \quad I_i = 1 \tag{3.28}$$

$$c_{2i} = \beta_2 X_{2i} + \sigma_{2u} W_{2i} + \varepsilon_{2i} \quad \text{for} \quad I_i = 0$$

The estimation can be performed in two steps. First an estimate of the parameters γ is obtained using a probit model on all observations. The parameters γ are used to calculate the values of W_{1i} and W_{2i}, for each group of observations separately. Second, two equations, one for each group, are estimated separately by OLS including the inverse Mill's ratios.

An additional complication arises in the estimation of the γZ_i in the farmers' model, because the irrigated farms are a sub-sample of the total number of farms, and the sample of farms is the result of an occupational decision that is correlated with the outcome. Therefore, two selection equations were formulated; one for households that have farming as their main income generating activity, and the other for the sub-sample of irrigated farmers within the sample of farmers.

$$I_1 = 1 \quad \text{if} \quad \gamma_1 Z_1 \geq u_1 \quad \text{for the irrigated farms} \tag{3.29}$$
$$I_2 = 1 \quad \text{if} \quad \gamma_2 Z_2 \geq u_2 \quad \text{for the farmers}$$

The probability of being a farm and of being irrigated is therefore the product of two probabilities:

$$\Pr ob(I_1 > 0, I_2 > 0) = \Phi(\gamma_1 Z_1)\Phi(\gamma_2 Z_2) \tag{3.30}$$

This is a sequential choice model, where first households choose their main occupation (farming or other), and then choose whether to irrigate their farms or not. The likelihood function of the sequential decision model takes the form (see Maddala, 1983, p 280):

$$L = \prod_{I_1 I_2 = 1} [\Phi(\gamma_1 Z_1)\Phi(\gamma_2 Z_2)] \cdot \prod_{I_1 I_2 = 0} [1 - \Phi(\gamma_1 Z_1)\Phi(\gamma_2 Z_2)] \tag{3.31}$$

The simultaneous estimation of the two probit equations above (3.29) with maximum likelihood provides the parameters values required for the calculation of W_{1i} and W_{2i}.

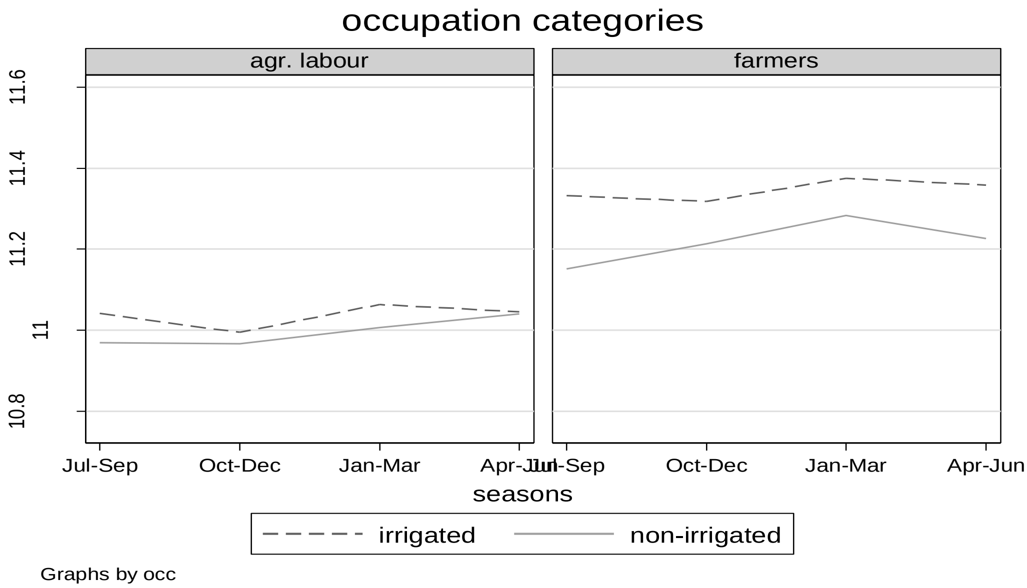
3.7 Empirical results

This section presents the results of the estimation of the causal effects of irrigation on consumption patterns of rural households. Two models are estimated. The first compares consumption patterns of households with and without access to irrigation. The second assesses the effect of income variability, measured by rainfall variability, on consumption decisions of households with and without irrigation. The econometric methods used to address issues of endogeneity and selection bias are presented in detail. The analysis is conducted both over seasons and years, and for farmers and agricultural labourers separately.

3.7.1 Seasonal consumption of farmers and agricultural labourers

There are substantial differences between expenditure of irrigated and non-irrigated farms, and between labourers of irrigated and non-irrigated villages, both in terms of the levels and seasonal patterns. Figure 3.5 shows per capita equivalent expenditure of farmers and agricultural labourers by season. Expenditure of households with access to irrigation is higher in all seasons, and the difference is particularly large between irrigated and non-irrigated farmers. For both farmers and agricultural labourers, expenditure is higher in the second part of the year. Expenditure of farmers seems to follow agricultural output, with consumption picking up in the second season and reaching its maximum in the third season. The seasonal pattern of agricultural labour households is less clear. Overall, the seasonal expenditure pattern of irrigated farms and labourers living in irrigated villages appears smoother. The consumption patterns of farmers and labourers with and without access to irrigation will now be analysed with regression analysis, and the statistical significance of these differences will be tested.

Figure 3.5 Seasonal consumption of agricultural labourers and farmers with and without irrigation



Source: calculated from NSSO data.

Farmers

Using the econometric model described in Section 3.6, separate regressions of equation 3.22, where run for samples of households operating irrigated and non-irrigated farms. The dependent variable in the regressions is the logarithm of per capita expenditure. Demographic effects are accounted for by including, among the regressors, the logarithm of household size and the ratios of five age groups by sex. The age groups are: children under five, children between the age of five and 13, children between the age of 14 and 17, adults from 18 to 59 years of age, and elderly of 60 years of age and over. Education of the head of household and spouse are included. When either of the two was not part of the household, the education of the person closest in age and of the same sex was used. The education level was divided in the following categories: illiterate, up to primary, middle, up to secondary, and higher education (graduate, post-graduate and equivalent). Other regressors include the logarithm of the size of land cultivated during the agricultural year in acres, dummy variables for each of the 23 administrative districts, and dummy variables for the four survey rounds. The different seasons of the agricultural year are included in the form of dummy variables. Each regression is adjusted by a selection term obtained in the way described in Section 3.6. The selection term is used to adjust the selection bias produced by the endogeneity of sub-sampling farmers within the rural sample of households, and further sub-sampling irrigated farms within the sub-sample of farmers. Both the choice of operating a farm and the opportunity to irrigate a farm are associated to characteristics that are correlated with per capita expenditure. The selection term is obtained by the simultaneous estimation with maximum likelihood of two probit models. Tables 3.13 and 3.14, show the probit selection equations for the farming choice, and the irrigation choice separately.

In the selection equation for the farming choice, the dependent variable is equal to one if farming is the most important production activity for the household. The selection equation must explain the household's decision to farm, using at least some variables that are not included in the final equation to be estimated. As the NSSO surveys are not particularly rich in non-expenditure data, the scope for building strong choice models is limited. The selection equation includes land owned by the household measured in acres among the regressors. The land may or may not be cultivated, as there is a sizeable lease market for land in rural Andhra Pradesh. Belonging to a scheduled caste or tribe is included because occupation in Andhra Pradesh is largely determined by caste. In particular, there is a strong correlation between agricultural labour work and scheduled caste. The historical average rainfall in the two cropping seasons (*Kharif* and *Rabi*) at the district level is included to account for environmental conditions more favourable to farming. Finally, the regression includes dummies for the four survey rounds.

Table 3.13 First selection equation: farming choice

Variable	Coefficient	St. error
Land owned	0.716***	0.012
Age of head of household	0.005***	0.001
Scheduled caste	-0.499***	0.039
Scheduled tribe	0.180***	0.052
Mean *Kharif* rainfall	0.001***	0.000
Mean *Rabi* rainfall	0.001***	0.000
50^{th} round (1993-94)	0.099**	0.037
55^{th} round (1999-00)	0.129**	0.039
61^{st} round (2004-05)	0.081*	0.043
Constant	-1.963***	0.137
Observations	21302	
F-statistic	464.2	
P-value	0.000	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

The second selection equation explains whether the farm is irrigated or not. The dependent variable is equal to one if the farm is irrigated. Farm irrigation is largely explained by the geographical location of the households (district dummies), because there is a large variation among districts in terms of access to canal water. Second, irrigation is explained by the availability of electricity at the household level, because water pumps are normally powered by electricity and the government of Andhra Pradesh has for long subsidised water extraction providing electricity free of charge to rural households. Other explanatory variables include the size of land owned by the household, and four dummies for the survey rounds. The results of this probit regression are shown in Table 3.14.

Table 3.14 Second selection equation: farm irrigation

Variable	Coefficient	St. error
Land owned	0.032***	0.009
Electricity	0.365***	0.063
Vizianagaram	0.444**	0.173
East Godavari	0.453**	0.170
West Godavari	0.854***	0.179
Krishna	0.950***	0.223
Guntur	0.229	0.169
Prakasam	-0.196	0.171
Nellore	0.527**	0.190
Chittor	0.189	0.157
Cuddapah	-0.026	0.204
Anantapur	-0.092	0.153
Kurnool	-0.408**	0.171
Mahaboobnagar	-0.265*	0.148
Rangareddy	-0.091	0.169
Medak	0.540**	0.196
Nizamabad	0.616**	0.197
Adilabad	-0.999***	0.180
Karimnagar	0.629***	0.176
Warangal	0.510**	0.163
Khamman	0.050	0.179
Nalgonda	0.237	0.167
50th round (1993-94)	-0.323***	0.080
55th round (1999-00)	-0.508***	0.086
61st round (2004-05)	-0.751***	0.086
Constant	0.388**	0.123
Observations	6454	
F-statistic	13.0	
P-value	0.000	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

The selection term obtained through the simultaneous estimation of the two selection equations is included in the estimation of equation (3.22), which tests the difference between seasonal consumption patterns of irrigated and non-irrigated farms. Table 3.15 contains the estimated coefficients, including the seasonal coefficients that test the significance of the patterns observed in Figure 3.5. The R-squares of the regressions are large (more than 70% of the total variance is explained), and most variables are highly significant, including the selection terms obtained through the selection equations. None of the seasonal dummies is significant in the regression of the irrigated farms, while two largely significant values were found in the regression of the non-irrigated farms. A seasonal pattern of consumption seems to emerge among households operating non-irrigated farms. Consumption is higher in the second and third seasons of the agricultural year. This conforms to theoretical expectations, as these are the seasons in which most of the agricultural output is realised through the *Kharif* and *Rabi* harvesting. The size of the consumption fluctuation however is not very large; its maximum fluctuation reaches the value of 11% between the first and the third seasons.

According to the theoretical model outlined in Section 3.3.3 (equation 3.17), the difference in seasonal consumption between irrigated and non-irrigated farms depends on both current and expected income. Irrigation reduces the variability of expected income caused by rainfall variability and irrigated farmers consume more both because they have higher income and because they make less precautionary savings. This hypothesis is tested in the following empirical analysis using rainfall as a predictor of expected income.

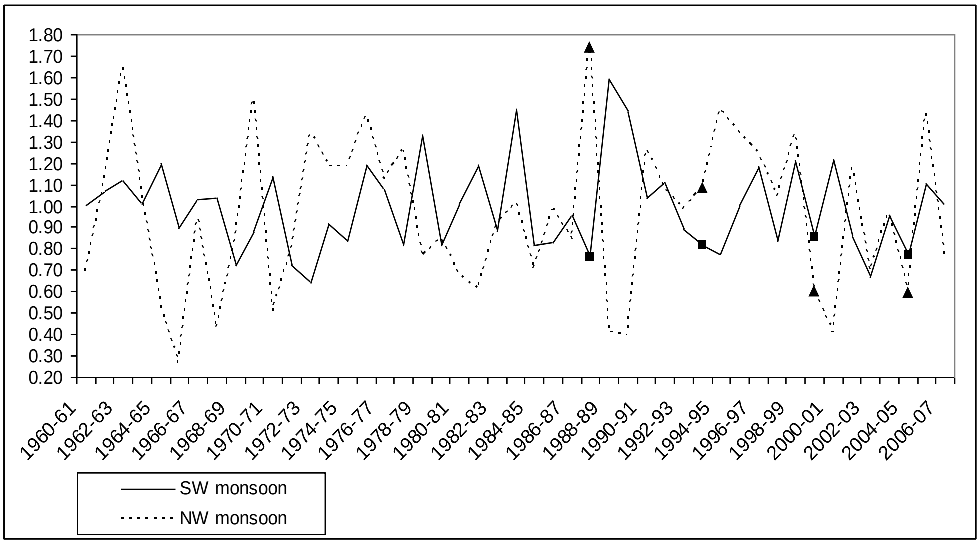
Table 3.15 Seasonal effects on consumption of farmers

Variables	Irrigated farms		Non-irrigated farms	
	Coefficient	St. error	Coefficient	St. error
Scheduled caste	-0.145***	0.026	-0.082*	0.045
Scheduled tribe	-0.138***	0.041	-0.205***	0.042
Age of head of household	0.003***	0.001	0.004***	0.001
Female children from 0 to 5	-0.078	0.115	-0.056	0.132
Male children from 5 to 10	0.120	0.086	0.155	0.133
Female children from 5 to 10	0.191**	0.086	0.106	0.132
Male children from 10 to 15	0.314***	0.094	0.434***	0.140
Female children from 10 to 15	0.298***	0.091	0.394**	0.157
Adult male	0.498***	0.091	0.394**	0.129
Adult female	0.409***	0.098	0.313**	0.148
Elderly male	0.302***	0.108	0.081	0.186
Elderly female	0.155	0.104	0.241	0.153
Logarithm of household size	-0.338***	0.021	-0.329***	0.033
Education of head of household	0.078***	0.009	0.050***	0.016
Education of spouse of head of household	0.030**	0.013	0.087***	0.028
Logarithm of cultivated land	0.095***	0.011	0.030*	0.018
2nd season (Oct-Dec)	-0.020	0.021	0.060**	0.034
3rd season (Jan-Mar)	0.024	0.022	0.114***	0.037
4th season (Apr-Jun)	0.020	0.021	0.032	0.037
Vizianagaram	-0.161***	0.045	-0.034	0.095
Vishakhapatnam	-0.042	0.066	0.147*	0.084
East Godavari	-0.035	0.055	-0.079	0.097
West Godavari	-0.126**	0.049	-0.307**	0.128
Krishna	-0.234***	0.051	-0.445***	0.139
Guntur	-0.318***	0.046	-0.273***	0.074
Prakasam	-0.097*	0.054	0.059	0.076
Nellore	-0.278***	0.060	-0.555***	0.089
Chittor	-0.286***	0.050	-0.389***	0.091
Cuddapah	-0.272***	0.052	-0.161**	0.072
Anantapur	-0.245***	0.050	-0.241***	0.076
Kurnool	-0.189**	0.061	0.048	0.074
Mahaboobnagar	-0.074	0.051	0.102	0.077
Rangareddy	-0.067	0.069	-0.016	0.069
Medak	-0.305***	0.046	-0.349***	0.079
Nizamabad	-0.181***	0.051	-0.350***	0.064
Adilabad	0.350**	0.113	0.305**	0.103
Karimnagar	-0.194***	0.047	-0.450***	0.086
Warangal	-0.155**	0.049	-0.099	0.098
Khamman	0.048	0.069	0.261**	0.091
Nalgonda	-0.203***	0.047	-0.019	0.084
Urban area	0.168***	0.043	0.253*	0.139
50th round (1993-94)	0.644***	0.024	0.837***	0.045
55th round (1999-00)	1.301***	0.036	1.639***	0.063
61st round (2004-05)	1.235***	0.022	1.380***	0.038
Selection term	-0.501***	0.092	0.510***	0.077
Constant	9.827***	0.107	8.474***	0.256
Observations	4572		1654	
R-square	0.73		0.72	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

Figure 3.6 shows the rainfall deviations from the historical average of the southwest monsoon (*Kharif* season) and northeast monsoon (*Rabi* season) for the period from 1960 to 2007. Rains of the northeast monsoon are more erratic than those of the southwest monsoon, showing higher and lower peaks in the series. Droughts, defined as deviations of more than 20% below the historical mean, are rather frequent. Drought conditions might also be concealed by these series, as the duration of dry spells and the arrival time of the monsoon also matter. Excessive rains may also have a negative impact on agricultural production, and a large positive deviation from the mean is not necessarily good. The marks on the series plotted in the chart correspond to the NSSO survey years used in this study. None of the four rounds was a particularly good year. The 55th and 61st rounds qualify as drought years, while the 43rd and the 50th rounds had bad *Kharif* rains, though these were partially compensated by good *Rabi* rains particularly in the 43rd round.

Figure 3.6 Rainfall deviations from the historical average in Andhra Pradesh



Note: marks on the series represent the four NSSO survey rounds used in this study.
Source: *indiastat*.

A version of the model in equation (3.23) was run separately for samples of irrigated and non-irrigated farmers in order to assess the effect of expected income variability on consumption patterns. Income expectations are modelled using the deviations of *Kharif* and *Rabi* rainfall from their historical average, the square of these deviations, and the historical variability of *Rabi* rains. Two models are estimated. The first includes only observations recorded in the season from July to September during *Kharif* rains. This corresponds to the final part of the lean season for farmers when planting takes place. The second model restricts the sample to observations collected between October and December, which is the period corresponding to *Kharif* harvesting and *Rabi* rains. The rainfall variables used will now be described in more detail, before proceeding to discussing the results.

The rainfall deviations are the percent deviations of total rainfall at the district level during the periods June-September (*Kharif* rains) and October-December (*Rabi* rains) over the average rainfall during the period from 1961 to 2005. Rainfall deviations affect the area sown by farmers in both *Kharif* and *Rabi*. In addition they help forming farmers' expectations regarding income of the following season, and therefore also affect current consumption. The squares of rainfall deviations are included in the regressions for two reasons. First, the positive impact of rainfall on consumption may be non-linear. Second, excessive rains may damage crops. According to data provided by the Ministry of Agriculture of Andhra Pradesh, heavy rains affected large numbers of households and districts in all years between 1987 and 2000 with the exception of drought years. The historical variability of *Rabi* rains is the coefficient of variation of *Rabi* rainfall at the district level over the period from 1961 to 2005. In order to avoid collinearity with the district dummies included in the regression, the coefficient of variation is calculated in each survey rounds over the ten years preceding the survey. This variable provides an indicator of the variability of future income, and the larger this variability, the lower is consumption in the current period because farmers save for precautionary reasons. In other words, farmers operating in districts where rainfall variability is higher, are forced to set aside some of the current income to prevent the effects of shortfalls in future consumption. This variable provides an exogenous variation in the variability of agricultural output across districts. The within-district response to this variability by irrigated and non irrigated farms is the object of the investigation. Irrigated farmers exposed to the same rainfall variability as non-irrigated ones are expected to save less for precautionary reasons, because the same variability in rainfall represents different variability of agricultural output.

Table 3.16 shows the coefficient values, and the standard errors of the rainfall variables included in the model. The dependent variable is the logarithm of per capita household expenditure, and the explanatory variables are the same included in the model of Table 3.15, with the exception of the seasonal dummies, which were not included here. Rainfall deviations do not show an impact on current consumption in the period July-October for neither farm category. When the model is run only for the observations collected in the period October-December, rainfall deviations have a positive coefficient because good rains predict higher future income thus stimulating current consumption, while the squares of the deviations have negative coefficients presumably because they reflect the negative impact of heavy rains on output. Expenditure by households in the non-irrigated sample show higher responsiveness to current rainfall, as the estimated coefficient is 50% larger than the coefficient estimated for the irrigated sample. Finally, *Rabi* rainfall variability has a slightly significant negative effect on expenditure of non-irrigated farms, but no effect on expenditure of irrigated ones.

Table 3.16 Rainfall effects on seasonal consumption of farmers

Season July to September				
	Irrigated farms		Non-irrigated farms	
	Coefficient	St. error	Coefficient	St. error
Rainfall deviation	-0.169	0.231	-0.677	0.539
Rainfall deviation squared	-0.268	0.489	-1.941*	1.020
Rabi rainfall variability	0.001	0.001	-0.001*	0.000
Selection term	-0.715***	0.174	0.843***	0.221
Observations	1163		449	
R-square	0.51		0.44	
Season October to December				
	Irrigated farms		Non-irrigated farms	
	Coefficient	St. error	Coefficient	St. error
Rainfall deviation	0.180**	0.059	0.269***	0.076
Rainfall deviation squared	-0.132***	0.034	-0.141*	0.066
Selection term	-0.871**	0.178	0.762**	0.213
Observations	1132		389	
R-square	0.46		0.57	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

These results suggest that rainfall deviations do affect current consumption in *Rabi* but not in *Kharif*. The effect is larger in *Rabi* than in *Kharif* because historically *Rabi* rains are more erratic than *Kharif* rains, and therefore cause larger fluctuations in agricultural output from its expected value. The size of the rainfall effect in *Rabi* is larger for non-irrigated farms than for irrigated farms, because expectations regarding future income are more related to current rainfall for non-irrigated farms than for irrigated ones.

Finally, the historical variability of *Rabi* rainfall has a modest negative effect on consumption of non-irrigated farms in the period July-September. *Rabi* rainfall variability has no effect on consumption of irrigated farms because irrigation reduces the variance of future expected income, and therefore reduces the precautionary motive for saving.

Agricultural labourers

The same empirical models used for farmers were used for agricultural labourers. Separate regressions of the model of equation (3.22) were run for agricultural labourers working in irrigated and non-irrigated villages. The dependent variable is the logarithm of per capita expenditure, while the explanatory variables are the same as those in Table 3.15.[21] The regression includes a selection term obtained by estimating a probit model of the choice of working as agricultural labourer against a set of explanatory variables. The standard two-step Heckman procedure was used. The inverse Mill's ratio from the probit regression for the occupational choice was calculated and included in the OLS regression. Agricultural labourers' location in irrigated or non-irrigated areas was considered exogenous.

The agricultural labour occupation is strongly correlated with belonging to one of the scheduled castes, which reflects the high stratification of the workforce by social group in rural Andhra Pradesh (see Table 3.17). Education is negatively correlated to agricultural labour work, as is the size of cultivated land. Remoteness, represented by the absence of electricity and the use of traditional fuel for cooking, and location in highly irrigated districts explains a large part of the choice of joining the agricultural labour force.

The results of the estimation of model 3.22 for agricultural labourers are presented in Table 3.18. Most variables in the regression are highly significant, including the selection terms obtained through the selection equation, and the R-squares are large. None of the seasonal dummies are significant for labourers working in irrigated villages, while two highly significant values are found for those living in non-irrigated villages. For these households, consumption is higher in the third and fourth season of

[21] Due to the large number of landless households among agricultural labourers, the size of cultivated land is expressed in acres, and not in logarithms, as was the case for farmers.

the agricultural year. The size of the consumption fluctuation however is quite small. Consumption in the second half of the year appears to be 5% higher than in the first half.

Table 3.17 Selection equation: agricultural labour choice

Explanatory variables	Coefficient	St. error
Scheduled caste	0.609***	0.030
Scheduled tribe	0.002	0.044
Age of head of household	-0.010***	0.001
Household size	0.009	0.007
Education of head of household	-0.347***	0.019
Education of spouse of head of household	-0.142***	0.029
Land owned	-0.511***	0.024
Land owned squared	0.005***	0.000
Electricity	-0.388***	0.028
Traditional fuel	0.865***	0.061
Vizianagaram	-0.106	0.078
Vishakhapatnam	-0.194**	0.080
East Godavari	0.259***	0.069
West Godavari	0.222**	0.074
Krishna	0.319***	0.080
Guntur	0.132*	0.070
Prakasam	0.009	0.076
Nellore	-0.055	0.078
Chittor	0.192**	0.072
Cuddapah	-0.061	0.084
Anantapur	0.201**	0.078
Kurnool	0.373***	0.082
Mahaboobnagar	0.349***	0.074
Rangareddy	0.105	0.088
Medak	0.197**	0.079
Nizamabad	0.040	0.082
Adilabad	0.035	0.088
Karimnagar	0.115	0.075
Warangal	0.027	0.079
Khamman	0.020	0.083
Nalgonda	0.337***	0.079
50th round (1993-94)	0.098***	0.031
55th round (1999-00)	0.263***	0.034
61st round (2004-05)	0.241***	0.038
Constant	-0.336***	0.094
Observations	21302	
F statistic	72.7	
P-value	0.000	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

The smoother consumption pattern of agricultural labour in irrigated areas is a consequence of the effect of irrigation on labour demand. Irrigation increases labour demand for a number of reasons. First, it expands the net area cropped, as more land is brought under cultivation. Second, irrigation may determine a shift from less labour demanding to more labour demanding crops. This is certainly the case in Andhra Pradesh, where irrigation promotes the cultivation of rice, which is labour demanding. Third, as yields increase, labour requirements per acre of land increase in particular for transplanting and harvesting. Irrigation has a stabilising effect on labour demand for two reasons. First, the introduction of double and triple cropping during the same agricultural year generates a more stable demand in the course of the year. Second, multiple cropping may produce a shift in the demand of labour from casual to permanent labour.

Table 3.18 Seasonal effects on consumption of agricultural labourers

Explanatory variables	Irrigated villages		Non-irrigated villages	
	Coefficient	St. error	Coefficient	St. error
Age of head of household	-0.002***	0.001	-0.003***	0.001
Female children from 0 to 5	-0.020	0.110	-0.015	0.055
Male children from 5 to 10	0.293	0.103	0.192***	0.050
Female children from 5 to 10	0.297*	0.096	0.191***	0.051
Male children from 10 to 15	0.727	0.113	0.524***	0.073
Female children from 10 to 15	0.401	0.120	0.399***	0.064
Adult male	0.680*	0.101	0.629***	0.054
Adult female	0.446*	0.105	0.400***	0.060
Elderly male	0.591	0.120	0.457***	0.073
Elderly female	0.329	0.115	0.177**	0.070
Logarithm of household size	-0.264**	0.025	-0.234***	0.017
Education of head of household	-0.008	0.016	0.000	0.010
Education of spouse of head of household	-0.010	0.020	0.009	0.014
Land cultivated	0.183***	0.030	0.011	0.011
2nd season (Oct-Dec)	-0.022	0.023	-0.006	0.012
3rd season (Jan-Mar)	0.008	0.024	0.051***	0.013
4th season (Apr-Jun)	-0.023	0.024	0.046***	0.012
Vizianagaram	-0.046	0.079	-0.063**	0.028
Vishakhapatnam	-0.038	0.057	-0.093**	0.031
East Godavari	0.028	0.049	0.046	0.033
West Godavari	-0.011	0.049	0.076**	0.031
Krishna	-0.065	0.051	0.075**	0.033
Guntur	-0.242***	0.057	-0.083**	0.030
Prakasam	-0.070	0.067	-0.073**	0.032
Nellore	-0.197***	0.053	-0.111***	0.034
Chittor	-0.144*	0.082	-0.213***	0.031
Cuddapah	-0.189**	0.079	-0.155***	0.037
Anantapur	-0.333***	0.054	-0.117***	0.031
Kurnool	-0.166*	0.090	-0.107***	0.030
Mahaboobnagar	0.080	0.100	-0.052*	0.027
Rangareddy	-0.325***	0.080	-0.051	0.032
Medak	-0.056	0.089	-0.052*	0.029
Nizamabad	-0.029	0.076	-0.074**	0.034
Adilabad	-0.105*	0.059	-0.168***	0.032
Karimnagar	-0.100*	0.058	-0.009	0.030
Warangal	-0.051	0.071	-0.027	0.032
Khamman	-0.002	0.074	-0.025	0.034
Nalgonda	-0.082	0.068	0.035	0.029
50th round (1993-94)	-0.060**	0.028	0.056***	0.013
55th round (1999-00)	0.097***	0.027	0.124***	0.012
61st round (2004-05)	0.079***	0.026	0.103***	0.014
Selection term	0.602***	0.127	0.614***	0.070
Constant	10.369***	0.119	10.283***	0.062
Observations	1592		5750	
R-square	0.35		0.28	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

In keeping with the analysis performed for farmers, rainfall variables were introduced in the model and two samples were constructed with observations of the first and the second trimester of the agricultural year. Rainfall deviations however, take on different meaning in the model for agricultural labourers compared to the farmers' model. Rainfall deviations not only represent income expectations, but also current income, as labour demand varies with the amount of rainfall in the planting season. The coefficient of variation of *Rabi* rainfall variability represents, as it did for farmers, an approximation of the variability of future income, as districts where rainfall variability is higher show a higher variability in demand for labour in the harvesting season.

Table 3.19 Rainfall effects on seasonal consumption of agricultural labourers

Season July to September				
	Irrigated villages		Non-irrigated villages	
	Coefficient	St. error	Coefficient	St. error
Rainfall deviation	0.552	0.568	0.648***	0.207
Rainfall deviation squared	0.841	1.078	1.189**	0.439
Rabi rainfall variability	-0.001	0.001	0.001	0.001
Selection term	0.515*	0.267	0.631***	0.117
Observations	317		1497	
R-square	0.40		0.35	
Season October to December				
	Irrigated villages		Non-irrigated villages	
	Coefficient	St. error	Coefficient	St. error
Rainfall deviation	0.035	0.086	0.083**	0.036
Rainfall deviation squared	-0.041	0.077	-0.045*	0.023
Selection term	0.176	0.282	0.749***	0.152
Observations	439		1401	
R-square	0.48		0.29	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

The results from the estimation of these models are shown in Table 3.19. Rainfall deviations have a highly significant effect on consumption of labourers of non-irrigated villages, but no effects on labourers of irrigated villages in both periods considered. The size of the effect is larger in the months from July to September compared to the period from October to December. The square of the rainfall deviation is also positive and very large in non-irrigated villages. The square of the rainfall deviation has a large positive sign in the first trimester but a negative sign in the second. The latter is probably a reflection of the negative effect of heavy rains. No effect is found of the *Rabi* rainfall variability on consumption in the first trimester by labourers in either irrigated or non-irrigated villages.

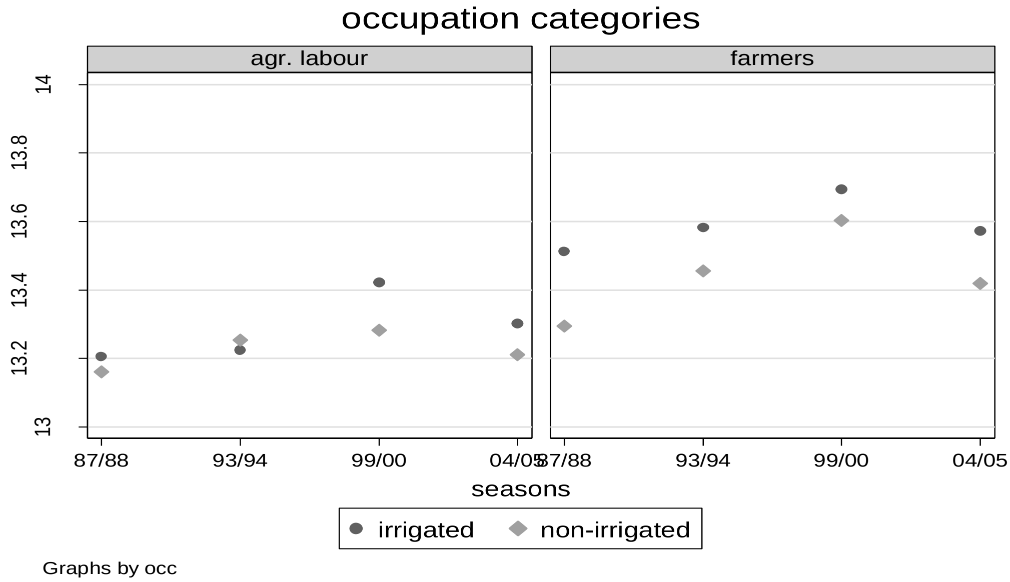
These results show that expenditure of labourers in non-irrigated villages is more responsive to current rainfall than consumption of labourers in irrigated villages. This response is either the result of higher current income or of higher expected income. Conversely, seasonal consumption of labourers of irrigated villages is not immediately affected by current and expected income. The absence of an effect of *Rabi* income variability on expenditure in the first trimester suggests that the precautionary motive for saving among both samples is not strong.

3.7.2 Consumption of farmers and agricultural labourers over the years

The impact of irrigation on consumption patterns of irrigated and non-irrigated farmers and agricultural labourers was also assessed over the four survey rounds. Expenditure is here measured in terms of annual expenditure using the data collected using the 365-day recall.[22] Figure 3.7 presents average annual expenditure by occupational category and availability of irrigation over the four survey rounds. Expenditure of irrigated farms is higher in all survey rounds, and is also smoother than that of non-irrigated farms. Expenditure of labourers living in irrigated villages is higher than that of labourers living in non irrigated villages, with the exception of the second survey round, but it does not appear to be smoother.

[22] Food and high frequency non-food items are only reported using a 30-day recall. Monthly expenditure on these items was multiplied by 12, and therefore reflects in part the expenditure pattern of the season in which data were collected.

Figure 3.7 Average annual expenditure of agricultural labourers and farmers



Source: calculated from NSSO data.

Tables 3.20 and 3.21 report the results of separate regressions of the logarithm of per capita expenditure for farmers and labourers with and without access to irrigation. The explanatory variables used in these regressions and the procedure adopted to correct the selection bias are the same used in the seasonal models. Instead of using seasonal dummies, these regressions estimate coefficients of dummies for the four survey rounds. Tables 3.20 and 3.21 report only the coefficients for the survey rounds and for the selection term. The sample of irrigated farmers seems to show a different consumption pattern, while the consumption pattern of labourers in irrigated and non-irrigated villages are very similar.

Table 3.20 Variation in farmers' expenditure by survey round

	Irrigated farms		Non-irrigated farms	
	Coefficient	St. error	Coefficient	St. error
Agricultural year 1994-94	0.060***	0.024	0.186***	0.046
Agricultural year 1999-00	0.112***	0.045	0.265***	0.046
Agricultural year 2004-05	0.097***	0.019	0.236***	0.053
Selection term	-0.141**	0.069	0.064***	0.018
Observations	4416		1553	
R-square	0.41		0.43	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

Table 3.21 Variation in agricultural labourers' expenditure by survey round

	Irrigated villages		Non-irrigated villages	
	Coefficient	St. error	Coefficient	St. error
Agricultural year 1994-94	-0.068**	0.033	0.048**	0.018
Agricultural year 1999-00	0.075**	0.032	0.100***	0.017
Agricultural year 2004-05	0.060*	0.032	0.084***	0.018
Selection term	0.611***	0.134	-0.387***	0.051
Observations	1592		5750	
R-square	0.35		0.28	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

Tables 3.22 and 3.23 report the results of regression models including rainfall variables. These models assess the effect of rainfall deviations on expenditure of households with and without irrigation. In order to do so, the samples are restricted to those households that were interviewed in the second half of the year (between January and June), and per capita expenditure is explained, among other variables, by rainfall deviations in the first half of the year (between July and December). Since most of the output is generated by the *Kharif* and *Rabi* crops, these rainfall deviations are strong predictors of annual income.

Coefficients of rainfall deviations are larger for the sample of non-irrigated farms both in *Kharif* and *Rabi*. This result suggests that expenditure by non-irrigated farmers is more dependent on current year rains, and therefore on current income. The coefficients of rainfall deviations of the agricultural labourers' regressions are more difficult to interpret. Expenditure of labourers in irrigated villages appears to be more responsive to income variations in *Kharif*, while the expenditure of labourers in non-irrigated villages is more responsive to income variations in *Rabi*.

Table 3.22 Rainfall effects on annual consumption of farmers

	Irrigated farms		Non-irrigated farms	
	Coefficient	St. error	Coefficient	St. error
Rainfall deviation (*Kharif*)	0.497**	0.246	0.696*	0.416
Deviation squared (*Kharif*)	0.752	0.503	0.208	0.846
Rainfall deviation (*Rabi*)	0.074*	0.044	0.182**	0.058
Deviation squared (*Rabi*)	-0.062**	0.029	-0.073**	0.036
Selection term	-0.025	0.080	0.061**	0.029
Observations	2216		779	
R-square	0.41		0.46	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

Table 3.23 Rainfall effects on annual consumption of agricultural labourers

	Irrigated villages		Non-irrigated villages	
	Coefficient	St. error	Coefficient	St. error
Rainfall deviation (*Kharif*)	1.094**	0.420	0.508**	0.199
Deviation squared (*Kharif*)	1.541*	0.938	0.521	0.424
Rainfall deviation (*Rabi*)	0.069	0.079	0.099**	0.039
Deviation squared (*Rabi*)	0.061	0.084	-0.074***	0.024
Selection term	0.802***	0.182	-0.359***	0.075
Observations	836		2852	
R-square	0.37		0.28	

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

The analysis of the impact of irrigation on consumption over the four survey rounds has important data limitations. First, unlike the analysis of seasonal expenditure, in which the allocation of expenditure over seasons was considered within a utility maximisation model, the analysis of expenditure over the years cannot be explained within a model of consumption and saving behaviour. The reason is that the four survey rounds available are not contiguous, hence no sequential decision modelling can be designed. A second limitation is that data on many determinants of consumption over the years are missing. Government policies and agricultural prices, for example, may affect decisions of irrigated and non-irrigated farms differently over the years and should be included as explanatory variables. Due to these two limitations, the conclusions presented in this section should be taken with caution, and be considered more as a descriptive exercise rather than a rigorous analysis.

3.8 Conclusions

This chapter assessed whether irrigation has a stabilising effect on income and expenditure of rural households. By providing water to dry areas and during times when rainfall is scarce, irrigation has a stabilising role on agricultural production, both over seasons and years. Over the years, irrigation reduces the negative impact of droughts, while within a given year it allows two or more cropping seasons. An intertemporal consumption model over the seasonal production cycle was built, which incorporates a precautionary motive for saving. In this model, seasonal consumption tracks seasonal income, though some consumption smoothing takes place as households save in the harvest season in order to spend in the lean season. Irrigation has the effect of increasing expected income and reducing expected income variability. In the empirical section of this chapter, seasonal consumption patterns and the effect of rainfall and rainfall variability on household expenditure were estimated using a Roy model. Separate regressions were run for farmers and agricultural labourers with and without access to irrigation. The main findings of the empirical analysis are as follows:

- Consumption by households with access to irrigation is more stable over the seasonal cycle. Expenditure of farmers' households without irrigation fluctuates up to 10% in the harvest season compared to the lean season. Expenditure of agricultural labourers living in non-irrigated areas fluctuates up to 5% in the harvest season compared to the lean season.

- Consumption by farmer households with access to irrigation is less affected by the size of expected income. In the lean season, current rainfall does not affect expenditure of farmers either with or without irrigation. In the harvest season, current rainfall affects expenditure of farmers without irrigation considerably.

- Consumption by agricultural labour households in areas without irrigation tracks income more closely. Current rainfall, representing labour demand for sowing and harvesting, increases consumption of agricultural labourers in non-irrigated areas, but has no effect on consumption of households in irrigated areas in both the lean and harvest seasons.

- Consumption by households with access to irrigation is less affected by expected income variability, represented by historical rainfall variability. The effect is statistically significant for farmer households but not for agricultural labour households. This is tentative evidence of precautionary saving behaviour.

- Consumption by households without irrigation is more unstable over the years and more affected by current rains. The results of the year-to-year analysis however must be taken with caution.

These results suggest that irrigation, by reducing the costs of seasonality, generates public and private savings. The size of the reduction in seasonal costs however is not easy to quantify because of the difficulty of estimating the three costs of seasonality described in Section 3.2: the welfare cost, the coping cost, and the cost of seasonal crises.

Standard household surveys do not capture the occurrence of seasonal poverty traps. Irrigation prevents seasonal crises because it reduces the occurrence of severe seasonal shocks and distress sales. However, these effects are not noticed by household surveys if seasonal shocks result in family dissolution, death or migration. The NSSO surveys interview households found in rural areas, thus missing out households that migrated or disappeared in consequence of seasonal shocks. The fluctuations in income and consumption observed by household surveys are therefore limited to fluctuations that do not produce seasonal crises and that are relatively small.

No reliable estimates of the welfare cost of seasonal fluctuations can be produced. The estimation of the welfare cost of consumption fluctuations requires data on consumers' preferences and their attitudes toward risk. Some estimates of the rate of time preference ('impatience') and of the degree of risk aversion are available for rural India (Binswanger, 1981, Pender, 1996). However, the experiments performed are too few, and the methodologies employed too questionable to provide reliable estimates (Cardenas and Carpenter, 2008).

Problems of data availability and of empirical identification prevent the estimation of the cost of coping strategies. The expenditure data of the NSSO, as those of most household survey, do not contain information on household savings, and panel data covering the seasonal cycle are extremely rare. Households' expectations of future income can be approximated by rainfall data, but rainfall data are a very imprecise measure of production risk. Finally, the empirical estimation of precautionary savings is a lively area of research and no consensus has been reached on how it should be performed (Carroll and Kimball, 2007).

While the present study cannot directly quantify the costs of seasonality and the benefits of irrigation, it does provide qualitative evidence of the fact that a) rural households without access to irrigation are not fully able to insure against seasonal fluctuations, and that b) households save for precautionary reasons. Both effects were found to be not very large. There are two possible explanations for the relatively small effect of the seasonal cycle on household expenditure: market development, and the operation of other smoothing mechanisms. These two explanations will now be discussed in turn.

It is unlikely that a rise in rural income may be responsible for a reduction in households' vulnerability to seasonal fluctuations. Over the last 25 years Andhra Pradesh has experienced a considerable increase in incomes. However, as shown in Chapter 2, this did not result in a reduction in the share of income produced in agriculture, which is the income component more sensitive to seasonal fluctuations. Nor there is evidence of a significant development of financial markets that justifies an increase in households ability to save and borrow (Pradhan et al., 2003).

It is more plausible that state interventions have helped households to smooth consumption fluctuations. Andhra Pradesh runs a policy of price support that stabilises incomes over the seasonal cycle by smoothing price fluctuations (Jairath, 2000). The state also runs an expensive programme of public distribution of food. In spite of serious inefficiencies, this programme provides most households with staple food at a fixed price over the year (Dutta and Ramaswami, 2001, Ravi and Indrakant, 2003). More recently the state also started a programme, the Employment Guarantee Scheme, which guarantees 100 days of employment to all requesting rural households (Devereux et al., 2008) and also some crop and weather insurance programmes on an experimental basis (Sinha, 2004, Vyas and Singh, 2006). Devereux et al. (2008) found large seasonal income fluctuations but no seasonal hunger among children in Andhra Pradesh, and also suggest that the impact of the seasonal cycle is smoothed by the large safety net operated by the state.

The findings of this study can also be used to perform a comparative assessment of different irrigation systems. In Andhra Pradesh there are three types of irrigation interventions, which stabilise income and consumption streams in different ways. These interventions are canal irrigation, groundwater extraction and watershed development, and will now be discussed in turn.

Investments in canal infrastructure by the Indian government have contributed to substantial poverty reduction. It is estimated that poverty in irrigated areas is one third of poverty in non-irrigated areas (World Bank, 2005). Andhra Pradesh has a long history of investments in large scale canal infrastructure with some spectacular results. The Godavari Barrage, built in the mid-19th century under British rule, transformed the famine-wracked districts of the Godavari Delta into a granary, and the builder of the barrage, Sir Arthur Cotton, into a saint revered throughout coastal Andhra Pradesh (World Bank, 2005).

Despite these successes, investments in canal infrastructure decreased substantially in the 1990s after reaching a peak in the early 1980s. The main reasons for the demise of investments in canal irrigation were inefficiencies in operation and maintenance (O&M), high investment costs, and negative social and environmental impacts (IEG, 2007). Funds for O&M were inadequate and inefficiently spent. The poor quality of irrigation set off a vicious circle as farmers were unwilling to contribute financially to service delivery (Oblitas and Peter, 1999). The government responded in the late 1990s by transferring irrigation management from the state to civil society, and Andhra Pradesh became one of the leading areas in the world in implementing participatory irrigation management. In some areas users' management of irrigation was built on existing institutions that had already performed this role very efficiently as documented by Wade (1988). However, in 1997 O&M was privatised, water charges were tripled, and water users associations (WUA) were created overnight all over the state. Impact studies of participatory irrigation management in Andhra Pradesh have found that this did not result in greater control over water use nor more equitable water distribution (IEG, 2007, Jairath, 2000).

In areas irrigated by canals the timing and quantities of water releases are operated by the irrigation department, because WUAs are unable to manage the distribution of water. The irrigation department however is not able to release water to farmers in a timely and equitable way. In addition, the release of water in the dry season depends on the recharge of the reservoirs in the wet season, which in turn depends on rainfall. In areas irrigated by canals therefore the provision of water in the dry season and the stabilising effects of irrigation are not fully guaranteed.

During the last 20 years Andhra Pradesh has witnessed a quiet revolution consisting of an increasing use of groundwater for irrigation. Today, groundwater provides the largest share of irrigated area in the state. There are several reasons for the expansion in of groundwater use (World Bank, 2005): electricity was slowly brought to rural areas; groundwater reduces waterlogging and salinity; cheap and easy to operate pumps have become available; finally, groundwater can be applied just in time bypassing canal administrators.

Groundwater extraction provides full insurance against rainfall shortfalls, because farmers can access water when they need it by operating water pumps. Protracted exploitation of groundwater however is unsustainable in the long term, because the costs of extraction are very high though mostly hidden (World Bank, 2005). First, it is estimated that between 10% and 30% of all electricity produced in India is absorbed by agriculture, and that every year it cost to the government 2.5 times the cost of maintenance of canals. Second, groundwater extraction depletes natural resources. Data show that in many areas of Andhra Pradesh water extraction is well above 100% of water recharge (Reddy et al., 2003a).

The development strategy of the Indian government for semi-arid areas where canal infrastructure and groundwater are not available consists of watershed development (MRD, 1994). Watershed interventions consist of micro infrastructures for water harvesting and farming practices that help the recharge of the water table. In Andhra Pradesh watershed projects are in operation since the early 1980s in degraded areas in terms of soils and water availability. These projects often have a strong pro-poor orientation (Springate-Baginsky et al., 2002). Watershed projects employ the poor for the realisation of infrastructures, and often include microfinance and management of common property resources components.

Watershed projects however are unlikely to help farmers smoothing seasonal fluctuations. Evaluations of watershed projects in India have found modest results in terms of both water conservation and poverty reduction (Kerr et al., 2002). The main reasons behind these disappointing results in Andhra Pradesh are the highly skewed distribution of land and the difficulty to involve marginal farmers, or the landless, in activities whose benefits mostly accrue to the land-rich (Kerr, 2002).

4 Income uncertainty and investments in human capital

4.1 Introduction

Andhra Pradesh fares poorly in terms of educational achievements. Illiteracy rates in the state are among the highest in India, and large disparities in educational attainments exist between boys and girls, rural and urban children, more developed and less developed regions, and across social groups. The causes of the poor state of education in Andhra Pradesh are well known (Prasad, 1987), though their relative importance is debated and a clear package of policies for the education sector is not available. Household poverty, discrimination in the labour market and within the household, and the paucity of state investments in education are among the main causes of the dismal state of education in Andhra Pradesh.

One consequence of the poor state of education is the low qualification of the labour force. This in turn has potential negative effects on the prospects of long-term economic growth. As shown in Chapter 2, technical progress in agriculture generates a surplus of labour in rural areas that needs to be employed elsewhere. The lack of general education and of specific skills reduces the possibility for the excess labour generated in agriculture to find employment in the modern sector. The option of migrating to urban areas is precluded to the uneducated masses that live in near subsistence conditions in the rural areas.

The weaknesses of the education system have long been recognised in India, and several programmes have been put in place to increase school attendance, particularly of girls. Possibly as a result of these programmes, considerable progress has been made over the last 15 years, and a much larger proportion of Indian children attend school today than 15 years ago. It is not at all certain however that by 2015 each child in Andhra Pradesh will be able to complete primary school, and that the disparity between boys and girls at all education levels will be eliminated as stated by the Millennium Development Goals.

This chapter assesses the impact of income uncertainty on household schooling decisions, a topic that has been largely neglected in the economics of education literature. It will be argued that high income uncertainty negatively affects investments in human capital made by poor rural households, and reduces their ability to effectively plan for the future of their children. This chapter proceeds in eight sections. Section 4.2 reviews the existing literature on the state of education in India and Andhra Pradesh, and on the effects of uncertainty and economic shocks on household schooling decisions. Section 4.3 builds an intertemporal utility maximisation model that relates child schooling to a number of determinants including income uncertainty. Section 4.4 describes the data sources that are used in the empirical analysis. Section 4.5 discusses the econometric methods for estimating a model of education attainment and a model of education expenditure. Section 4.6 presents descriptive data that illustrate the state of education in rural areas of Andhra Pradesh. Section 4.7 presents and discusses the estimation results of the models of education attainment and education expenditure that were introduced in Sections 4.3 and 4.5. Section 4.8 concludes.

4.2 Literature review

4.2.1 Basic education in India

Dreze and Sen (2002) stress five ways in which investments in human capital, and in education in particular, contribute to socio-economic development in India. First, education has an intrinsic importance, as it is valuable in itself. Second, it opens up opportunities for the individual, and allows the achievement of economic goals. Third, it opens up opportunity for society, as it promotes the discussion of social needs and their representation in political demand. Fourth, it has an empowering role, by readdressing power relations between groups, sexes, and within the family. Finally, it generates positive externalities that benefit other household members, the community and the society at large.

Despite these potential benefits, education has been largely neglected by Indian governments since the Independence. What is more striking is that educaiton has been neglected not only by governments and the elites, but also by political organisations, the unions and civil society (Dreze and Sen, 1995). Dreze and Sen (1995) point to three main reasons for this neglect. First, the conservative notion persists that lower castes do not require education and that child labour among poor people is acceptable. Second, it is believed that literacy imparted in primary school is inferior to education in a

Gandhian sense. Third, it has long been a belief among the Indian left that the existing education system is exploitative of the poor.[23]

The result of governments' neglect is that India's educational accomplishments are poor in comparison to international standards. According to UNESCO (2007), the literacy rate in India over the period 2000-2004 was just 61%, a rate similar to those reported by sub-Saharan African countries. Riboud et al. (2007) report that the proportion of the population that was illiterate in India in 2004 was similar to the one observed in China around 1970 or in Malaysia around 1960, while the fraction of the population that completed secondary school in India in 2004 (14%) is half the figure prevailing in China in 1974.

There is considerable horizontal inequality in educational achievements in India. Vaidianathan and Gopintham (2001) point to a number of regularities of the uneven distribution of educational achievements across social groups and geographic areas. First, literacy rates are much lower in rural areas compared to urban areas. The UNESCO (2007) reports that over 80% of out-of-school children in India live in rural areas. Second, literacy rates are much lower among girls. According to the UNESCO (2007) the ratio between literacy rates of female and male children is 0.65. Finally, educational achievements of children from scheduled castes and tribes are lower than those observed in the rest of the population.

Few data on learning achievements are available in India, but the data available suggest that the quality of education being imparted in primary schools is very poor. Yadav et al. (2001) review the results of the achievement tests conducted in primary schools by the National Council of Educational Research and Training (NCERT) over the decade 1990-1999. Test scores at state level show percentages around and above 40%. This is a reasonable mark if one considers that a 35% mark is a 'pass' in standard examinations across the country. However, it is a poor results if one considers that the minimum level of essential learning standard set by the NCERT requires a score of 80%. Kingdon

[23] A discussion of how education can reinforce existing power and unequal relations can be found in Levinson and Holland (1996). Jeffrey et al. (2005) are one example of an application of this type of analysis to rural India. They show how the expansion secondary education in a rural village of Uttar Pradesh resulted in a reinforcement of existing class and caste inequalities.

(2007) reports the results of two national large scale language and mathematics tests administered to primary school children in 2002 and 2005. These tests confirm that learning achievements of Indian pupils are very low. For example, nearly 55% of the students in grade five are not able to perform a simple arithmetic division, and 47% are not able to read a text of medium difficulty.

The poor state of primary education in India has been explained in several ways. First, public education expenditure is very low and there are inconsistencies between official declarations and actual government policies. Tilak (2001) documents that public expenditure as a share of GNP increased in India from an average 1.8% in the 1950s, to 2.8% in the 1960s, 3.1% in the 1970s, 3.2% in the 1980s and 4.1% in the 1990s. Despite this increasing trend, the 6% target set by the government of India in 1969 was never achieved. Lamentably, more than 90% of the education budget consists of teacher salaries rather than school buildings and teaching material. As a result, the quality of schools and teacher performance are very poor (PROBE Team, 1999), and teachers' absenteeism is widespread (Kremer et al., 2005). Dreze and Kingdon (1999) use data of five northern Indian states to illustrate the negative impact on enrolment of general indicators of school quality, like school infrastructure, teachers' training, and the provision of a midday meal.

Second, India has historically invested disproportionately in higher education and has neglected primary education (Kochar, 2003). Tilak (2001) distinguishes four phases in the allocation of educational expenditure. The first phase between 1951 and 1956 (the first 5-year plan) saw a substantial fraction (three fifths) of total expenditure going to elementary education. The second phase, from 1956 to 1969 (second and third 5-year plans), witnessed a decline in resources allocated to elementary education and a doubling or trebling of resources to higher education, in an attempt to help the industrialisation of the country. The third phase, from 1969 to 1986, showed a slight reversal of the previous trend, with a an increase in the expenditure in primary education and a decline in the expenditure in higher education. Only the last phase, after 1986, saw a return of emphasis on elementary education investments.

Third, low returns to schooling and the extreme poverty of rural households discourage parents from investing in the education of their children (Dreze and Kingdon, 1999).

Returns to schooling are particularly low for elementary education. Using NSSO data Riboud et al. (2007) show that returns to primary and middle education are well below the standard 10% observed in most countries. In addition, returns have not been affected by recent economic growth. Openness to trade, rapid growth and technological innovation have increased returns to higher education, by increasing demand of skilled labour relative to unskilled labour, but have left returns to primary education unchanged over the period from 1993 to 2004.

Fourth, parents' attitudes toward schooling and cultural factors are obstacles to child education, particularly of girls. Kaul (2001) discusses how parents' attitudes negatively affect schooling since a very early stage. A large percentage of Indian children entering primary schools are first generation learners who do not exhibit the desired level of readiness required for learning. Most children come from poor settings that do not provide sufficient stimulation in terms of adult-child interaction, sensory exposure and provision for play and learning.

Parental conservatism and lack of motivation are particularly detrimental to girls' education (Kambhapati and Pal, 2001). Dreze and Sen (2002) suggest a number of factors that shape parents attitudes against female schooling: parents are reluctant to send girls to other villages for studying; poor job perspectives and wages for women reduce the perceived benefits of female education (see also Kingdon, 1998, 2005, Kingdon and Unni, 2001); patrilocal residence and exogamy (the practice of settling in the husband's village and severing the links with the family) does not encourage investing in girls' education; the habit of marrying more educated boys (hypergamous marriage) increases the cost of the dowry if the girl is more educated.

Fifth, there are intended and unintended discriminatory practices against scheduled castes and tribes that prevent children from progressing through school (Nambissan, 1996). Nambissan and Sedwal (2001) discuss how caste dynamics at the village level, and even within the classroom, hamper the educational efforts and motivation of Dalit children. Jeffrey et al. (2004) illustrate how poverty and the lack of social connections in urban areas prevent children of Dalit families of a rural village in Uttar Pradesh from accessing salaried jobs. Hoff and Pande (2006) illustrate, by conducting an experiment in a rural Indian school, how discriminatory norms can be interiorised in such a way to reproduce existing caste inequality. Sujatha (2001) discusses two main reasons for the poor achievements of children from scheduled tribes: the absence or poor quality of educational facilities in tribal communities, and a pedagogical system disrespectful of tribal languages and cultural norms.

Finally, there are historical and social factors explaining poor educational performance and differences across Indian states. Banerjee et al. (2005) show the negative impact of British rule and of the landlord tenure system on education investments at the state level. Vaidyanathan and Gopinthanam (2001) discuss the role of the state in promoting universal education in Kerala; the pioneering role of Christian missionaries in Tamil Nadu; the success of community mobilisation and NGOs in demanding more resources for education in Rajasthan; the advantages deriving from vicinity to centres employing skilled labour like the army in Uttar Pradesh or the modern administration system in Kerala. The role played by non-agricultural development in fostering child education is also discussed by Kochar (2004), who tests the impact of wages in the nearest urban area on school completion of a sample of rural Indian students, and by Chambarbagwala (2008), who assesses the impact of expected returns to schooling, measured by the average wage of adult males, on enrolment of a sample of Indian children.

Despite all these constraints, considerable progress has been made in primary education in India over the last 10-15 years. According to the UNESCO (2007), in 2004 the net enrolment rate in India was 90% in primary school and 54% in secondary school, and many of the social and geographical disparities outlined above was substantially reduced if not eliminated. Govinda (2001) summarises the strategies – often supported or promoted by external assistance - that made the 1990s the period of most intensive primary education development in the following way. New school facilities were established in small habitations reaching population living in remote areas. Participatory school mapping, micro-planning and community management and monitoring were introduced. A district specific approach to educational planning was adopted. Alternate and part-time schools were established to reach out-of-school children. Mobilisation campaigns were initiated to change negative cultural norms and parents attitudes. Several school feeding programmes were introduced along with many special programmes targeting female children and children from disadvantaged groups.

4.2.2 Uncertainty and investments in human capital

There is considerable evidence that poor Indian households value education. Dreze and Sen (2002) discuss the myth of the lack of interest in education of Indian parents, and report a large body of evidence against it (for Andhra Pradesh see Prasad, 1987). Caldwell et al. (1985) find that in rural Karnataka most parents see education as the only hope for a different way of life, and that they strive to invest in the education of their children in order to secure a higher and more stable income in the non agricultural sector. Jeffrey et al. (2005) conducted a qualitative study in a village of Uttar Pradesh in 2000-02 and assessed parents' perceptions of schooling. They find that parents from both the high caste (*Jat*) and the low caste (*Chamar*) consider schooling as crucial in improving personal skills, social standing and employment opportunities, and that households invest large sums in the hope that they children will obtain salaried work.

The importance attached by Indian parents to education is evidenced by the substantial expenditure in education and the rapid diffusion of private schools in rural areas (Tilak, 2002). Sipahimalani (2000) reports the results of three distinct education surveys conducted by NCERT, NCAER and NSSO between 1993 and 1995. According to these surveys between 15% and 33% of Indian children are attending private primary schools. Even among very poor households (with a per capita income below Rs. 2,000 a year), 4% of 6 to fourteen year old children are attending private schools. These figures indicate that private expenditure on education is large even among the poor.

Parents spend large sum on education even for children that are enrolled in government schools, particularly in books, travel and uniforms (Sipahimalani, 2000). According to NSSO data used in this study, household in Andhra Pradesh spend on average 3% of total outlay on the education of their children. This figure is quite large if one considers the typical Indian household portfolio asset allocation. Data reported by Pradhan et al. (2003) for all Indian households show that 70% of households allocate less than 6% of their income per year to physical investments and less than 5% to financial investments. These data however include both urban and rural households, and the same report indicates that the urban saving ratio, as a multiple of rural saving ratio, is 1.7 for physical savings and 3.9 for financial savings. It would seem that for a large fraction of poor and landless rural households education is the most important form of investment.

As first observed by Stiglitz (1969), investments not only depend on expected returns but also on risk aversion, and high uncertainty and risk aversion may lead to underinvestment. Households in developing countries face a level of uncertainty that has been compared to the one prevailing during a state of war (Fafchamps, 2003). In the absence of insurance mechanisms and of access to credit, households adopt strategies that impair their ability to accumulate wealth. For example, Morduch (1990) shows that poor Indian farmers invest in technology that is less profitable but that provides a more stable income over time. There is also evidence that households store wealth in economically inefficient ways for precautionary reasons (Dercon, 2004). Banerjee (2004) builds a formal economic model showing how uncertainty and risk aversion may lead poor households to underinvest. Dela Cruz-Dona and Martina (2000) conducted a qualitative study in two rural villages of the Philippines and conclude that uncertainty of income flows and inability to insure against income uncertainty has a strong detrimental effect on parents' schooling decisions (see also Jeffrey et al., 2005).

The present study focuses on the impact of income uncertainty on schooling decisions of poor Indian households. Income uncertainty has received virtually no attention in the economics of education literature. In the original formulation of the human capital investment model by Becker (1975), little attention is paid to uncertainty, though Becker points out the potential riskiness of investments in human capital. First, investments in education are intangible and they cannot be sold in case they go wrong. Second, individuals face uncertainty about their learning ability and the quality of schooling at the time they make the schooling decisions. Third, there is the uncertainty with regard to the length of life and therefore the possibility of reaping the benefits of education. Finally, there is uncertainty regarding the future market valuation of the knowledge acquired, because the condition of market demand for the specialisation obtained cannot be known in advance.

The literature on the economics of education has focused on the effects of uncertainty of future earnings on schooling decisions. The best formal treatment of this topic has been provided by Levhari and Weiss (1974), who define the theoretical conditions under which uncertainty of future returns discourages investments in human capital. The hypotheses formulated by this literature have found few empirical applications, probably due to the difficulty of estimating uncertainty of future returns. Kodde (1986) tests the effect of uncertainty in future earnings among a sample of Dutch high school graduates using subjective perceptions of future incomes. Unexpectedly, he finds that an increase in the uncertainty of future income has a positive effect on the probability of pursuing additional education. Carneiro et al. (2003) use a sophisticated econometric technique to estimate the effect of uncertainty of future earnings on high school and college decisions among a sample of North American students. They find that if uncertainty is substantial, this has an impact on schooling decisions, but that this impact is very small compared to the effect of other determinants of schooling.

All these studies on the effects of uncertainty on schooling decisions assume that returns from physical capital are certain or at least more certain than returns from human capital. This is a reasonable assumption in developed countries, but is rather unrealistic in a poor and rural society. In rural India, households face substantial production risk as farmers and agricultural labourers, two occupations that do not require schooling. Conversely, schooling allows the undertaking of non-agricultural activities, like public jobs, which have a more stable income than agricultural activities.

The argument made in this chapter is that in the context of rural India, the uncertainty of agricultural incomes tends to reduce household investments in human capital. To see why this happens consider a rural household that is very poor and faces substantial production risk, with little opportunity to borrow. In these circumstances, uncertainty of future agricultural income has the effect of increasing the size of savings used to insure against income risk. This in turn has a negative impact on household investments. In other words, the uncertainty of future income forces households to invest their savings in an inefficient way, prioritising the protection of future consumption over investments in human capital. This theoretical construct can be seen as an application of the precautionary savings theory of consumption (Carroll, 2001, Deaton, 1991) to the standard Becker’s (1975) human capital investment model.

The analysis pursued in this chapter bears some similarity with analyses of the impact of economic shocks on school attendance. Several studies have empirically investigated the effect of macro-economic and idiosyncratic shocks on schooling in developing countries. The evidence thus far is mixed. A number of studies have found no effect of economic shocks on schooling. For example, Duryea and Arends-Kuenning (2003) investigate the effect of economic slowdowns in urban Brazil over the 1980s and the 1990s and conclude that Brazilian economic crises did not contribute to an increase in child employment or a reduction in school attendance. McKenzie (2003) analyses the strategies adopted by Mexican households to cope with the economic crisis of 1994-95 and finds that school attendance of primary school children increased during the crisis both in urban and rural areas. Similarly, Schady (2004) studies the effect of the Peruvian crisis of the early 1990s on schooling of urban children, and the evidence he collects suggests that children were more likely to be in school and complete their grades during the crisis.

Other studies have found a significant impact of economic shocks on schooling. For example, Jacoby and Skoufias (1997) find that aggregate shocks, measured by fluctuations of village income, negatively affect schooling of Indian children. Jensen (2000) compares attendance rates of rural children exposed to rainfall shocks in Cote d'Ivoire and finds that attendance rates were lower in regions more negatively affected by weather shocks during the period 1986-1987. Thomas et al. (2004) investigate the effects of the economic crisis that affected Indonesia in 1998 and find that the crisis produced a decline in enrolment rates. The focus of all these studies is the impact of economic shocks on school attendance. This however is only loosely related to grade progression. Children may be withdrawn from school temporarily during a crisis or after an idiosyncratic shock, but can return to school once the crisis is over or the shock has been overcome. This point is clearly made in Thomas et al. (2004), while Skoufias and Parker (2006) find evidence that though the Mexican economic crisis of 1995 reduced girls' school attendance, their overall grade progression remained unchanged.

The present study adopts a different perspective from the studies described above. Rather than assessing the direct effect of income variability on school attendance, it investigates the effect of expected income variability on child progression through primary school. The hypothesis made is that households make schooling decisions taking into account the future probability of economic shocks. If this hypothesis is valid, household facing higher income risk will invest less in human capital.

4.3 A model of uncertainty and investments in human capital

This section presents a version of the human capital investment model originally formulated by Becker (1975), in which households determine the level of consumption and investment in education by maximising the use of resources over the life cycle. Three crucial assumptions of the model are discussed: exogenous income risk in agriculture; diminishing returns to schooling; and precautionary saving behaviour. A formulation of the model under uncertainty is developed which will be tested empirically in Section 4.7.

4.3.1 An intertemporal model of consumption and investments in education

The model developed in this section is designed for rural households that are poor and whose income comes from agricultural activities, either as farmers or agricultural labourers. Irrigation infrastructure is poor and the agricultural income is heavily dependent on rainfall. Schooling does not increase the productivity of farmers and agricultural labourers, but it allows the undertaking of non-agricultural activities that earn higher returns. Household schooling decisions are framed as investments in human capital. Households invest in the education of their children if future benefits from schooling are larger than schooling costs. The benefits from schooling consist of the higher income, either in the form of wage income or self-employment income, that the more educated earn in the non-agricultural sector. Other benefits of schooling, like the potential increase in productivity of the farm or in the home and the pleasure of learning are ignored. Costs of schooling are of two types, direct costs and opportunity costs. Direct costs include tuition fees, and the purchase of books, stationery, transport, and uniforms. The opportunity cost consists of the unearned wage, or the farm product lost, for not engaging in agricultural activities while in school.

In order to simplify matters, the starting point is a simple two-period model. In the first period households produce an agricultural income and make decisions regarding children schooling. In the second period household income, which includes the non-agricultural income of schooled adults, is spent. A similar model was formulated by Levhari and Weiss (1974), and extended by Kodde (1986), and Snow and Warren (1990). These authors however considered the effect of uncertainty of future labour earnings on schooling investments. The model developed here takes future labour returns as certain, and focuses on the effect of uncertainty of agricultural incomes on investments in education.

Households maximise a utility function over a life cycle composed of only two periods (zero and one) of the form:

$$\max U = E\left[\frac{u}{(1+\delta)}(c_0, c_1)\right] \tag{4.1}$$

Households derive utility (U) from consuming the amount of goods c_0 in period zero and c_1 in period one. Utility in period one is discounted by the time preference factor (δ), which measures the consumer's impatience. This is the extent to which the consumer prefers consumption in period one to consumption in period two. The expectation sign means that future income is uncertain and that the consumer has an idea of its probability distribution. Education is an investment good that generates benefits in the second period. No allowance is made for the fact that people acquire education for the simple pleasure of learning. Therefore, education does not have an independent effect on household utility and it is not a consumption good.

Households invest in education by employing a proportion (l) between zero and one of their total available time (T) for studying in period one. Time spent in school in the first period earns a non-agricultural income ($f(l)$) in the second period, which is a function of the time spent on learning, and whose form will be specified later. Schooling entails two types of costs. The first consists of the direct costs of education (S), which is proportional to schooling time. The second is the opportunity cost of schooling, which is the income forgone in the agricultural activity (y_a). Income in the two periods for the household investing in education is therefore:

$$y_0 = (1-l)y_{a0} \tag{4.2}$$

$$y_1 = (1-l)y_{a1} + f(l) \tag{4.3}$$

At this point, several simplifying assumptions have already been made. First, it is assumed that household time is entirely spent either working in the farm or studying and leisure time or other work in the home is not considered. Second, all income generated in the two periods is spent and no bequest is left at the end of period one. Third, households can save in the first period and savings can be negative. However, there is an implicit restriction on borrowing originating from the fact that consumption in the second period must be positive, which in turn implies that borrowing in the first period must be lower than income in the second period discounted by the interest rate (*r*). Fourth, households have no initial assets that they can sell for consumption purposes or for education investments. The effects of different initial wealth positions on household investments are discussed later.

Investments in education can be funded by reducing current consumption or by borrowing, thereby reducing future consumption. The budget constraint takes the form:

$$c_1 = (1+r)[(1-l)y_{a0} - c_0 - lS] + [(1-l)y_{a1} + f(l)] \qquad (4.4)$$

The first term on the right-hand-side is household saving in the first period, which can be negative if household spend in consumption and education more than they earn in the agricultural activity. The second term on the right-hand-side is household income in the second period, which is equal to the sum of the agricultural and the non-agricultural incomes. Households maximise utility by choosing the optimal levels of schooling (*l*) and of consumption in the first period (c_0). Maximisation of utility (4.1) with respect to the budget constraint (4.4) yields the following first order conditions:

$$\frac{\partial U}{\partial c_0} = E\left[\frac{\partial U}{\partial c_0} - \frac{(1+r)}{(1+\partial)}\frac{\partial U}{\partial c_1}\right] = 0 \qquad (4.5)$$

$$\frac{\partial U}{\partial l} = E\left\{\frac{\partial U}{\partial c_1}\left[f_l(l) - \frac{(1+r)}{(1+\delta)}y_{a0} - (1+r)S - y_{a1}\right]\right\} = 0 \qquad (4.6)$$

The first order condition (4.5) is the solution to the intertemporal consumption problem. It states that transferring consumption from one period to the other will not increase utility, but will rather reduce it. The second order condition (4.6) equates costs and benefits of education. It states that the marginal utility value of additional benefits from education in the second period, resulting from investing in schooling, equals the marginal utility value of the cost of schooling in the first period. Therefore no utility gain can be made by reducing or increasing the amount of schooling in the first period.

Equation (4.6) can be rewritten by applying the expectations operator (*E*) to each term. The expectation operator of the agricultural income term in period zero (y_{a0}) cancels out, because when schooling decisions are made agricultural income is known. The expectation operator of the marginal returns to schooling term $f_l(l)$ also cancels out, because it is assumed that non-agricultural income is known with certainty. This assumption rules out the search for employment in the non-agricultural sector and unemployment spells, and all other uncertainty factors that were discussed in the previous section. It is a radical way to express the fact that non-agricultural incomes are more stable than agricultural ones. The expectation sign only applies to the agricultural income term of period one (y_{a1}), and equation (4.6) takes the form:

$$\frac{\partial U}{\partial c_1} f_l(l) = \frac{\partial U}{\partial c_1}\left[\frac{(1+r)}{(1+\delta)} y_{a0} + (1+r)S\right] + \frac{\partial U}{\partial c_1} E[y_{a1}] \qquad (4.7)$$

Under certain conditions, this equation shows that an increase of agricultural income risk reduces investments in schooling. These conditions consist of assumptions regarding the characteristics of the agricultural income process, the shape of the non-agricultural income function with respect to schooling, and the shape of the utility function. Before showing the effect of uncertainty of agricultural income on education investments, these three conditions will be discussed in turn.

4.3.2 Characteristics of the agricultural income process

Following Newbery and Stiglitz (1981), the present analysis considers an uncertain agricultural income characterised by multiplicative risk. Agricultural income is a function of agricultural inputs (*z*), such as labour, land and livestock, and technology (τ), like irrigation, all multiplied by a risk coefficient (*θ*), with expectation $E\tilde{\vartheta} = 1$ and variance $Var\tilde{\vartheta} = \sigma^2$. Agricultural income is therefore:

$$y_a = \tilde{\vartheta} q(z, \tau) \qquad (4.8)$$

In this formulation, the expected agricultural income is equal to y_a because $E\tilde{\vartheta} = 1$, but the certainty equivalent income (which is the expected utility for that level of income) varies with the risk variance σ^2 and with the particular shape assumed by the utility function. Assuming risk-averse consumers with utility functions that are concave down, the certainty equivalent income is lower than expected income, and more so the larger its variance.

Rainfall variability is the only source of agricultural income risk considered. For a given level of agricultural technology and irrigation infrastructure, weather variability determines the variability of agricultural output. Idiosyncratic risk in agricultural production is also important, and correlated to the inputs of the agricultural production function and to the demographic structure of the household. For example, farmers of poorer households are more likely to fall ill. However, for simplicity it is assumed that the sample of household is relatively homogenous in characteristics and that idiosyncratic agricultural production risk is randomly distributed across households.

4.3.3 Returns to schooling in rural India

This study makes the important assumption that returns to schooling are decreasing with levels of schooling. For example, returns to schooling *f(l)* have the form:

$$f(l) = a + bl - \frac{1}{2}cl^2 + lR \qquad (4.9)$$

In this formulation non-agricultural income is a quadratic function of schooling (*l*), which is the available household time spent in school. Unlike empirical Mincerian functions, which either assume linear returns to schooling or let the returns be determined by the data, equation (4.9) explicitly assumes that returns to schooling are decreasing, as $f_l = b - cl + R > 0$, but $f_{ll} = -c < 0$. Another important difference between equation (4.9) and Mincerian equations is the term *R*, which represents factors like individual ability and parental background that may increase returns to schooling if operating in conjunction with learning. Hence, in the specification of the returns to non-agricultural income used here, returns to schooling $f_l(l)$ vary with factors like parental background included in *R* and are decreasing with schooling.

The assumption that returns to schooling are decreasing is supported by a large body of evidence. Psacharopoulos (1985, 1994, 2002) produced the most extensive reviews of estimations of returns to schooling from all countries in the world. The results of his review with regard to the shape of the returns to schooling function can be summarised in the following way. First, rates of return decline with educational level and are the highest in primary school. The main reasons for this pattern are the low cost of investment in primary education and the larger productivity differential between primary graduates and illiterates. Second, rates of return decline with income across countries and over time within countries, because marginal returns to education are diminishing and as the stock of educated workers increases returns decrease. The evidence collected by the estimation of Mincerian equations in India, however, is of the opposite sign to what has been generally found in the rest of the world. Estimations of returns to schooling in India have found very low or insignificant returns to primary education and a U-shaped, or increasing, relationship between returns and educational levels. The evidence accumulated so far is so large to deserve some attention.

Table 4.1 shows the results of empirical estimations of returns to schooling in India. The table is based on Kingdon's (1998) and Dutta's (2006) reviews of earlier studies and includes Kingdon's and Dutta's own estimates. Two main points can be made about the results reported in Table 4.1. First, returns to primary education are well below the level of 10% which is normally found in other developing countries. Second, returns are often increasing with educational level and are very high for secondary and higher education.

Table 4.1 Estimated returns to schooling in India

	Year	Primary	Middle	Secondary	Higher
Urban Uttar Pradesh[a]	1995	2.6	4.9	17.6	18.2
Urban Tamil Nadu[b]	1995	2.9	9.0	17.0	15.6
Urban Madhya Pradesh[b]	1995	3.1	9.7	12.5	13.5
Dehli[c]	1975-76	2.4		6.9	11.4
Urban Tamil Nadu[d]	1989	-1.8	8.5	-0.01	16.9
Urban and rural India[e]	1983-84	3.3	2.4	5.3	9.0
Urban and rural India[e]	1993-94	2.1	1.7	4.3	9.2
Urban and rural India[e]	1999-00	2.4	2.0	4.6	10.3
Urban and rural India[f]	1983-84	6.1	7.1	13.2	12.2
Urban and rural India[f]	1993-94	6.2	6.4	12.6	12.2
Rural Andhra Pradesh[g]	1990	9.9		3.2	7.0
Urban Madhya Pradesh[h]	1987-88	1.4	6.9	14.2	9.6
Urban Tamil Nadu[h]	1987-88	1.1	6.4	12.4	17.1

Source: [a]Kingdon (1998), [b]Unni I (1995), [c]Banerji and Knight (1985), [d]Santaparaj (1996), [e]Dutta (2006), [f]Duraisamy (2002), [g]Tilak (1990), [h]Kingdon and Unni (2001).

In light of the theoretical model developed for the present study the following observations can be made regarding these estimations. First, there are several technical problems with the estimation of this type of equations which are extensively discussed by Heckman et al. (2003). These include: endogeneity of schooling; omitted variable bias; neglect of school costs; and neglect of uncertainty. Few studies address all these sources of bias. In the case of the studies in Table 4.1, the neglect of school costs seems particularly important as these are likely to increase with schooling. Second, it is by no means clear that individuals should make schooling choices based on returns estimated by Mincerian regressions. This point is convincingly made by Heckman et al. (2003), who also show that if school decisions are based on observed average returns at different schooling level, and if the earning variability is larger for the less educated, as the evidence suggests, then Mincerian equations overestimate returns to schooling. Third, most studies reported in Table 4.1 are performed on small urban samples of wage earners. The only studies of rural households are those of Tilak (1990), who finds decreasing returns to schooling in Andhra Pradesh, and Dutta (2006), who finds U-shaped returns for India. Lastly, it is not always clear whether the higher returns in secondary school are equally distributed between lower and upper secondary or are driven by higher returns in upper secondary. The latter prepares students for college and is not part of the primary education cycle that is the focus of this study.

4.3.4 Precautionary saving behaviour

In the *General Theory*, Keynes (1936) identified eight motives for saving: precaution, foresight (the life-cycle motive), calculation (intertemporal substitution), improvement, independence, enterprise, pride, and avarice. Consumption theory has traditionally focused on the life-cycle and the intertemporal substitution motives. The precautionary motive has been recently incorporated in consumption theory by the precautionary savings theory of consumption. As Keynes puts it (1936), people save 'to build up a reserve against unforeseen consequences'.

Formally, precautionary savings are introduced in consumption theory through specific assumptions about the shape of the utility function. While risk-aversion requires the utility function to be concave, prudence requires the derivative of the utility function to be convex. Prudence, as risk-aversion, can be measured and the methods and theorems of risk-aversion almost equally apply to the analysis of prudence (Kimball, 1990).

Risk-aversion means that income draws below expected income affect utility more than draws above expected income. In other words, risk-averse individuals fear losses more than they appreciate gains, and uncertainty of future income has the effect of impoverishing consumers. Formally this is represented by a concave utility function as those in the top charts of Figures 4.1 and 4.2. These two figures chart some of the most commonly used utility functions in theoretical work. All these functions are concave down, which implies consumers' risk-aversion.

Risk-aversion can be measured by the level of concavity of the utility function. A measure of risk-aversion is given by the absolute degree of risk-aversion (Newbery and Stiglitz, 1981):

$$A = -\frac{U''(y)}{U'(y)} \tag{4.10}$$

The more concave the utility function, the more weight is given by the consumer to negative income draws compared to positive income draws, and the more the consumer is impoverished by income uncertainty.

Figure 4.1 Logarithmic and quadratic utility functions

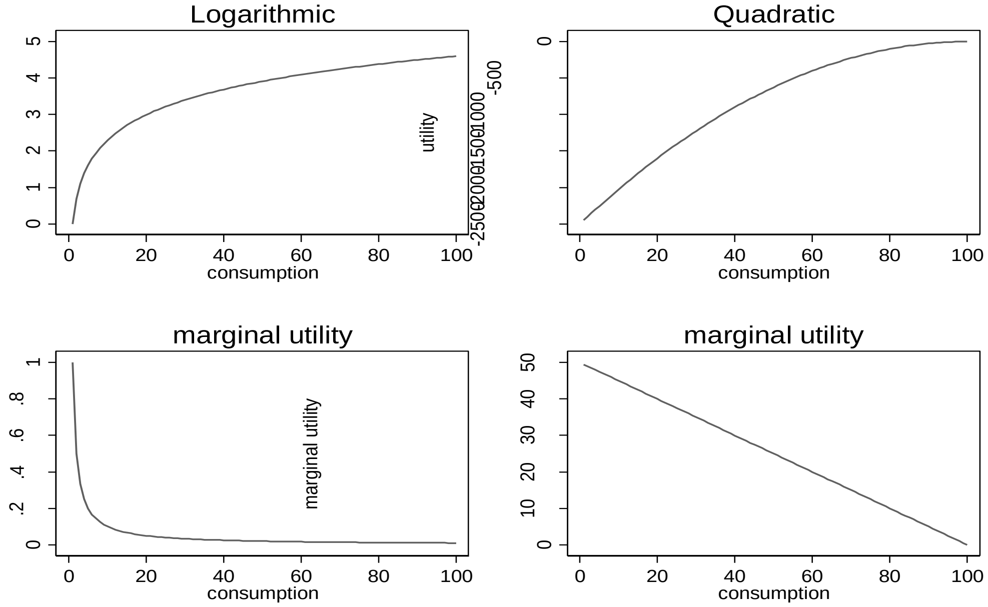
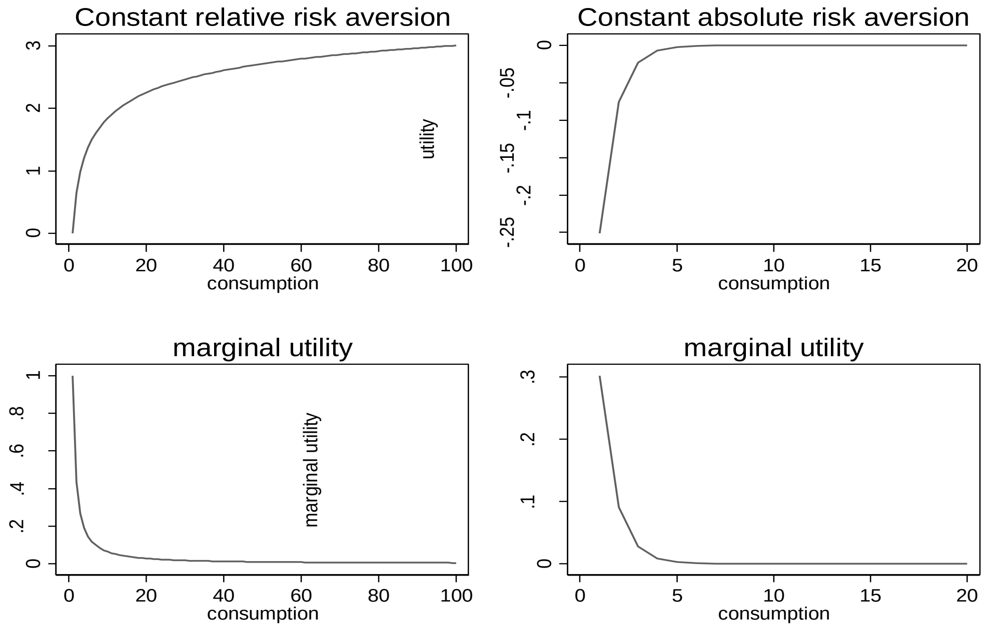


Figure 4.2 CRRA and constant absolute risk aversion (CARA) utility functions



Prudence means that the consumer is willing to forgo current consumption to compensate for future uncertainty. This is so because to the eyes of the consumer income is more valuable in uncertain times. Formally this is represented by the convexity of the derivative of the utility function, as depicted in the bottom charts of Figure 4.1 and Figure 4.2. An increase in the variance of income in a period increases the value of marginal utility of consumption in that period. This is not true for all utility functions however. For example, the quadratic utility function on the top left of Figure 4.1 has a linear derivative function, which rules out prudent behaviour.

Prudence can be measured by the level of convexity of the derivative of the utility function. A convenient measure is the absolute degree of prudence (Kimball, 1990):

$$P = -\frac{U'''(y)}{U''(y)} \tag{4.11}$$

The more convex is the derivative of the utility function, the more valuable to the consumer is income in the period in which income is uncertain. Table 4.2 summarises the values of the derivatives of the four utility functions illustrated in Figure 4.1 and Figure 4.2, together with the values of the coefficients of risk aversion and prudence.

Table 4.2 Properties of common utility functions

	Utility function (U)	First derivative (U')	Second derivative (U'')	Third derivative (U''')	Coefficient of risk aversion $(-U''/U')$	Coefficient of prudence $(-U'''/U'')$
Logarithmic	$\ln c$	$\frac{1}{c}$	$-\frac{1}{c^2}$	$\frac{2}{c^3}$	$\frac{1}{c}$	$\frac{2}{c}$
Quadratic	$-(a-bc)^2$	$2b(a-bc)$	$-2bc$	0	$\frac{1}{(a/c)-b}$	0
CRRA	$\frac{c^{1-\gamma}-1}{1-\gamma}$	$c^{-\gamma}$	$-\gamma c^{-(\gamma+1)}$	$\gamma(1+\gamma)c^{-(\gamma+2)}$	$\frac{\gamma}{c}$	$\frac{(1+\gamma)}{c}$
CARA	$-\frac{e^{-\gamma c}}{\gamma}$	$e^{-\gamma c}$	$-\gamma e^{-\gamma c}$	$\gamma^2 e^{-\gamma c}$	γ	γ

To see how prudence affects consumption behaviour, consider the result of the simplest intertemporal consumption model. Consumers allocate consumption over two periods, there are no initial assets and all income earned over the two periods is spent. In order to further simplify, assume that the rate of time preference equals the rate of interest so that their ratio is one. Maximisation of the intertemporal utility function gives the Euler equation:

$$U'(c_1) = E_1 U'(c_2) \tag{4.12}$$

The optimal allocation of consumption between the two periods is obtained when the marginal utility of consumption in the first and the second periods are equal. At this point, no gain can be made by transferring a unit of consumption from one period to the other.

Suppose now that uncertainty in the second period increases. In the case of prudent consumers this will lead to a reduction in consumption in the first period. To see how this works, suppose that the variance of future consumption c_2 increases, though its expected value $E(c_2)$ remains unchanged. This leads to an increase in the marginal utility of consumption of the second period $U'(c_2)$, because the derivative function of the utility function is convex. The term on the right-hand-side of the Euler equation increases and in order to preserve the equality the term on the left-hand-side must increase. In order for this to happen the marginal utility of consumption in period one must increase, which can only happen by reducing consumption in period one, because the utility function is concave. As consumers react to uncertainty of future income with an increase in current savings, these can only have a precautionary motive.

Note that this result does not require the imposition of borrowing constraints, though an implicit constraint follows from the fact that consumption in the second period must be positive. This alone is sufficient to create in the consumer a reluctance to borrow, or to borrow too much, because the consumer fears that in a later period he might face a very low income and will not be able to repay the loan (Carroll, 2001).

The extent to which the consumer saves against future uncertainty depends on the degree of convexity of the marginal utility function, which is measured by the coefficient of prudence (*P*). Bagliano and Bertola (2004) show how the size of this effect can be measured by applying a Taylor series expansion to the right-hand-side of the Euler equation $E_1 U'(c_2)$ around consumption in period one c_1. A second order expansion of this term, which ignores higher order terms, is:

$$E_1 U'(c_2) = U'(c_1) + \frac{U''(c_1)}{1!} E_1(c_2 - c_1) + \frac{U'''(c_1)}{2!} E_1(c_2 - c_1)^2 \tag{4.13}$$

Note that the term on the left-hand-side and the first term on the right-hand-side cancel out because they are the Euler equation. The expression can hence be rewritten as:

$$E_1(c_2 - c_1) = -\frac{1}{2}\frac{U'''(c_1)}{U''(c_1)} E_1(c_2 - c_1)^2 \tag{4.14}$$

An increase in uncertainty increases the square of the expected difference between present and future consumption, which is the last term on the right-hand-side. Note that the expected value of c_2 has not changed, but because it is the square of the difference which is calculated, the term increases in value. This increase has the effect of augmenting the difference between current and future consumption (the term on the left-hand-side) by half the value of the coefficient of absolute prudence. The latter is always positive, because for prudent consumers it is the ratio of a positive third derivative of the utility function over a negative second derivative.

4.3.5 Uncertainty and investments in human capital

In order to model the behaviour of prudent consumers, it is assumed that the utility function $U(c)$ is concave ($U'(c) > 0$ and $U''(c) < 0$), and that the marginal utility function $U'(c)$ is convex ($U'''(c) > 0$). As discussed in the previous section, prudence implies that the marginal utility function of consumption varies with the variance of expected income. Consider an increase in the variance of future income that leaves its mean unaffected. Given the convexity of the marginal utility function, this results in an increase in the marginal utility of future consumption. The value of future consumption increases when income is more uncertain and prudent households save in order to insure against future uncertainty.

The impact of uncertainty on human capital investments can be seen by analysing again equation (4.7). Consider a mean-preserving increase of the income variance in the last term on the right-hand-side of the equation. Though future expected income has not changed, the marginal utility of consumption in the second period has increased. Households are more inclined to save in the first period as consumption in the second period has become more valuable. In order for the equality to be preserved, the left-hand-side of the equation must increase. Given that marginal returns to schooling are decreasing, this can only happen by a reduction in the size of *l*, which is the proportion of household time spent in school. An increase in uncertainty of future consumption has therefore the effect of reducing education investments.

By inverting the left-hand-side of (4.7), the optimum level of household schooling ($l*$) can be expressed as a function of the marginal returns to schooling, the interest rate, the time discount rate, the expected agricultural income, and the cost of schooling:

$$l* = f\left(f_l, r, \delta, y_a, S\right) \tag{4.15}$$

For the purpose of empirical estimation, equation (4.15) can be expanded by specifying the determinants of its components. Marginal returns to schooling (f_l), by inversion of equation (4.9), consist of the wage paid to increasing level of education w_s, experience (approximated by age), and the factors included in *R*, which affect the market valuation of additional years of schooling. The main determinants of the returns to schooling included in *R* are: child's gender and social group, because women and members of scheduled castes and tribes are discriminated in labour markets; parents' background and motivation; child's learning abilities; and the quality of the school attended.

Expected agricultural income (y_a) depends on the factors determining the income process described in equation (4.8): the amount of farm input used in production (*z*); the expected production risk (θ); and the production technology (τ). The cost of schooling (*S*) consists of the opportunity cost, which is approximated by the wage for unskilled labour paid in agricultural labour markets (w_{uns}), and of the direct costs of education, which are approximated by the demographic structure of the households (*h*); distance (*d*) to school; and other locality specific factors (ξ), like the availability of school materials, school fees, etc.

The theoretical model was formulated ignoring credit constraints. While these are not strictly needed for uncertainty to have an effect on education investments, households that are liquidity constrained have fewer opportunities to insure against income uncertainty. Precautionary savings is one insurance mechanism, but it is likely to be the sole insurance mechanism only for very poor households that have no assets to sell or to use as collateral. Different levels of liquidity constraints could be included in the model by using household specific interest rates (r_i). These however are hard to define and likely to be correlated with household assets. The discount rate (δ) can also be considered to vary with household wealth by assuming, as in Fisher (1930), that poorer households are more impatient. Household specific interest rates, the time discount rate, and household ability to borrow are better summarised by household wealth measured by assets (A). For estimation purposes equation (4.15) can therefore be rewritten in expanded form as:

$$l^* = f\left(w_s, R, A, \tilde{\vartheta} q(z,\tau), S(w_{uns}, h, d, \xi)\right) \tag{4.16}$$

Equation (4.16) will be estimated by a model of progression through school levels with data specific to each child. In principle, it would be interesting to estimate progression through school separately for household categories that are more dependent on agricultural output variability, like farmers and agricultural labourers with and without access to irrigation. The model of equation (4.16) however is not appropriate for this task because current household occupation is endogenous to schooling and is observed in the present, not at the time the schooling decisions were made. During the survey interview, households report past schooling achievements, but household income categories are defined at the time of the interview. Few households are entirely specialised in farming or wage work, and classifications of households as farmers or labourers are based on the assessment of household income shares at the time of the interview. Household classifications based on income shares from different sources are meaningful when current decisions are considered, but it is questionable that the current income generating process can be used to explain schooling decisions made in the past. First, the shares of income from different activities change over time. Second, the income generating process is endogenous to schooling, because past education achievements affect the share of non-agricultural household income in the present. Therefore the income classification of households cannot be used to explain schooling decisions. The distinction between irrigated and non-irrigated farm is even more problematic because the irrigation status of a farm may change in any direction over a time span of ten to 15 years.

Education expenditure data at the household level will therefore be used to compare schooling decisions of different household groups. In doing so, it is assumed that schooling is proportional to education expenditure. An inevitable drawback of this approach is that the analysis of household expenditure on education can only be conducted at the household level, thus loosing information that is specific to the individual child.

4.4 Data

This chapter uses data from four different sources: the NSSO expenditure data; the rainfall data available on the *indiastat* website; the DES data on agricultural wages in India; and data of the Census of Villages of Andhra Pradesh. The characteristics of these data sources and their use will now be discussed in turn.

4.4.1 NSSO data

The empirical analysis uses data from the last four large NSSO survey rounds in Andhra Pradesh, namely the 43rd round of 1987-88, the 50th round of 1993-94, the 55th round of 1999-00, and the 61st round of 2004-05. Table 4.3 reports the size of the samples of rural households and children between the age of five and 18 for each survey round.

Table 4.3 Rural observations by NSSO survey round

Observations	43rd (1987-88)	50th (1993-94)	55th (1999-00)	61st (2004-05)
Rural households	6,016	4,908	5,181	5,555
Rural children	9,045	6,547	6,864	6,553
All households	9,439	8,552	8,987	8,431

Source: calculated from NSSO data.

The core of the NSSO survey is the expenditure module, where detailed information is collected of all expenditures incurred by the household over the month or the year preceding the interview. Data on education expenditure are collected with considerable detail and completeness, and will be used in the estimation of the expenditure models of Section 4.7.4. A specific survey module, the household roster, collects demographic data for each member of the household, including information on educational achievements of all household members. These data will be used in the estimation of the school attainment model of Section 4.7.2. In addition to the expenditure module and the household roster, the survey questionnaire also includes an introductory section which collects data on several socioeconomic characteristics of the household, many of which will be used as explanatory variables in all models estimated for the present analysis.

4.4.2 Rainfall data

Rainfall data are collected by the DES and are available for download on the *indiastat* website. Rainfall data are in millimetres of rain, and are reported for every month and for each district of Andhra Pradesh from 1952 onwards. These data will be used to test the hypotheses formulated in the theoretical model in Section 4.3.5 regarding the effect of uncertainty on education investments. The empirical formulation of the model postulates that rural households form their expectations regarding rainfall variability based on their past experience of this variability, and that they use these expectations to predict future variability of agricultural income.

4.4.3 DES data

Since 1955, the DES has been collecting data on rural wages of skilled and unskilled workers for each district and state in India. These data are published for every agricultural year (from June to July) in the *Agricultural Wages of Indi*a series of the DES (1975-2005). The series publishes data collected at specific locations in each district for rural skilled labour (carpenter, blacksmith and cobbler), and unskilled labour (field labour, other agricultural labour and herdsman).[24] The series are further disaggregated by male, female and child labour.

Wages are used in this study to approximate the opportunity cost of schooling and the perceived returns to schooling. Male and female wages for field labour are used to approximate the opportunity cost of schooling. Field labour wages are averages of wages paid to ploughmen, sowers, harvesters, weeders and transplanters. The reported wage of child labour was not used on account of the large number of missing

[24] Localities surveyed in the State of Andhra Pradesh are: Korasawada (Srikakulam), Vaddadi (Visakhapatnam), Bonangi (Vizianagram), Kothapalli (East Godavari), Keyyalaguden (West Godavari), Ghantasala (Krishna), Tadikonda (Guntur), Karamchedu (Prakasam), Tummur (Nellore), Sirvel (Kurnool), Bukkapatnam (Anantapur), Chitvel (Cuddapah), Gellapalli (Chittoor), Arutla (Ranga Reddy), Chandoor (Nizamabad), Sanga Reddy (Medak), Veldanda (Mahboobnagar), Athmakur (Nalgonda), Chagallu (Warangal), Cheral (Khamman), Chippial (Karimnagar), Basar (Adilabad).

observations and doubts regarding its representativeness of the value of child work, as only a small fraction of children is employed full time. The male wage of carpenters was chosen to approximate the perceived returns to schooling, because data on wages of blacksmiths and cobblers are largely incomplete and the skilled wage series are available only for male workers. In all cases, wages are reported for a normal working day of eight hours.

4.4.4 Census data

The population census of India is conducted every ten years and data of the census of rural villages of Andhra Pradesh are available for the years 1961, 1971, 1981, 1991 and 2001. These data report the population of each rural village in the state by sex, social group and occupation. They also report the number of literate individuals for each village. The data of the two more recent rounds (1991 and 2001) also report the availability of public facilities in the villages, such as schools and clinics, and detailed data on land use.

Census data cannot be matched with data collected by the NSSO, because the latter does not provide census village codes for the clusters surveyed. They are used to present descriptive statistics on literacy rates in Section 4.6 and to build district level variables for the regression models in Section 4.7. In particular, the census data are used to calculate the per capita availability of primary, middle and secondary school in each district.

4.5 Econometric methods

This section describes the methods employed for the estimation of two econometric models. The first is a model of school attainment estimated by a Cox regression with time covariates. The second is an education expenditure model in share form estimated by four different econometric techniques. This section describes the rationale and the specification of the two models in detail.

4.5.1 Estimation of the school attainment model

The school attainment model estimates an empirical version of equation (4.16), focusing on the effect of households' expectations of future income variability on children's school attainments. The specific variables included in the econometric model are discussed in detail in Section 4.7.1.

The dependent variable of the school attainment model is the highest level of education completed by the child. The NSSO surveys do not report school attendance or highest year of schooling attained. Instead, the household roster categorises education of each household member in terms of highest level of education achieved in the following way: illiterate; below primary; primary; middle; secondary; higher secondary; graduate and above. There is also a category for non-formal education which includes titles obtained through Non-Formal Education Courses, Adult Education Centres, and Total Literacy Campaign. For the purpose of the present analysis all non-formal education is considered equivalent to illiteracy.

The levels of education considered by the present analysis correspond to pre-higher education, which consists of primary, middle and lower secondary school, and are summarised in Table 4.4. These three levels cover the first ten years of schooling and upon completing lower secondary school students are awarded the Secondary School Certificate. Students may then progress to higher secondary school (grades XI and XII), which focuses on preparation for university. Upon completing higher secondary school students are awarded the Senior School Certificate in public examinations that give them access to universities. The two years of higher secondary have been excluded from the analysis because only a tiny fraction of rural children are reported to achieve this level of education.

Table 4.4 Pre-Higher Education in India

Level of education	Age range	Years of schooling
Primary (grades I to V)	6 to 11	5
Middle (grades VI to VIII)	11 to 14	3
Lower Secondary (grades IX to X)	14 to 16	2

The sample of children included in the present analysis are between the age of five and 18, rather than between six and 16, because entrance and exit from school do not necessarily respect official age limits. Setting the lower limit at age five allows the inclusion of early entrants into primary and for age heaping around number five in the reporting of age.[25] Setting the upper limit at age 18, two years after the official completion of lower secondary, allows for late entrants and repeaters.

The definition of attainment categories as in Table 4.4 implies censoring of observations. The problem of censoring derives from the fact that some of the children in the sample did not attain higher education levels simply because they were too young to do so. For example, a child aged 12 with a primary school title is censored, because she might be attending middle school. On the other hand, a child with the same title and aged 18 is a dropout, because she was not able to achieve a grade higher than primary. However, it is not obvious whether a 15-year-old child who completed primary should be considered a censored observation or a dropout. In order to operate this distinction it is necessary to use some reasonable cut-off points based on the child's age. For the purpose of the present analysis, children above the age of six who had not achieved below primary at the time of the interview are classified as illiterate, while children aged six are censored observations. Children above 11 who had not achieved primary at the time of the interview are classified as below primary, while children below 12 are censored observations. Children aged above 14 who had not achieved middle are classified as primary, while children below 15 are censored observations. Finally, children above 16 who had not completed lower secondary are classified as middle, while children below 17 are censored observations.

The structure of the dependent variable as an ordered categorical variable lends itself to estimation by an ordered choice model like the ordered probit. Estimating school attainment with a standard ordered probit model however poses two limitations, namely censoring and the presence of time-varying explanatory variables. These two limitations are now discussed in turn.

Censoring generates a spurious correlation between attainment and variables that are correlated with age. By model construction, children of older cohorts are more likely to have reached higher school levels, which is often the opposite of what happens in reality. There are several ways to address the problem of censored observations. One straightforward and radical approach is to include only the uncensored observations in

[25] Age heaping is the rounding up or down of self-reported age. It is common in survey data to observe clustering of observations at ages five, ten, 15 etc.

the sample. This implies limiting the sample to children who have reached 18 years of age at the time of the survey. This would leave no ambiguity as to whether children are still in school or dropped out, because all children aged 18 had the opportunity of attaining the highest education level considered by this study. This approach has two main drawbacks. First, it drastically reduces the size of the sample. Second, it amplifies the problem of explaining past schooling decisions using current household characteristics.

Another often used method for dealing with censored observations is the estimation of a censored ordered probit (Glewwe, 1999, King and Lillard, 1987). This method modifies the likelihood function of the standard ordered probit model in order to account for censored observations. A drawback of this approach is that the computation of the modified likelihood function can be cumbersome. In addition, the results of the ordered probit are not easy to interpret. The sign of the regression parameters can be interpreted as determining whether school attainment increases or not with the regressors. However, the marginal effects of the regressors on the probability of attaining a given level of education have to be computed separately, and the computation is not straightforward.

The presence of time-varying covariates is a more serious limitation than censoring. It is standard practice in the estimation of school attainment models to explain schooling decisions made in the past with current values of the explanatory variables. Examples include: household expenditure in the year preceding the interview, local market wages for skilled and unskilled labour, and school availability at the time of the survey. However, household expenditure, wages and school availability vary over time. Assuming that they are fixed is a misspecification error, which would lead to some of the outcomes being explained by the wrong observations. In addition, this is a potential source of endogeneity, because some of the variables measured at the time of the survey are correlated with past educational attainments in the sense that they were determined by the level of education attained by children. For example, household expenditure can be large at the time of the interview due to higher wages earned by children who went to school, rather than the other way round.

In order to overcome both the problem of censoring and time-varying explanatory variables, the Cox proportional hazard model is used to estimate the determinants of school attainment (Cameron and Trivedi, 2005). The Cox proportional hazard model estimates the determinants of dropouts rather than of school attainment. There are several advantages with using this approach. First, it deals effectively with the problem of censoring, hence all observations in the sample can be used without the risk of obtaining biased estimates. Second, the likelihood function is easy to compute compared to likelihood function of the censored ordered probit and most statistical packages have built-in routines to perform this computation. Third, the regression parameters are in the form of hazard ratios that are easy to interpret and provide not only a measure of the qualitative impact of the explanatory variable (the sign), but also the absolute value of its effect on the dependent variable. Finally, Cox proportional hazard models offer an easy solution to the problem of time-varying covariates, as these are easily incorporated into the model.

There are however also some disadvantages with using the Cox proportional hazard model. First, the model estimates the probability of failure (school dropout) rather than attainment, and the reader interested in the determinants of attainment has to read the results in reverse. Second, the estimation requires a rather complex set-up of the data in order to be performed. The data used in the Cox regression need to report the time when the observation was recorded. In the school attainment model this corresponds to the educational level attained by the child, which is the dependent variable. A second variable that specifies whether the child was successful or failed at each school level, and one that indicates whether the observation is censored are needed.

The structure of the dataset when using time-varying covariates is slightly more complex and is organised in the following way. Each child observation is repeated for every value of the independent variable from the lower until the highest value achieved or until censored. For example, an 18-year-old child dropping out of middle school would have three observations, one for the below primary decision, one for the completion of primary and another one for the failure in middle school. A 16-year-old child who completed middle school would be assigned three observations as in the case above, but with no failures and would be coded as a censored observation because at 16, the child could still complete secondary school. A child never enrolled will be classified as a failure at below primary. In other words, the dataset contains multiple observations for the same child, each relating to success or failure at each value of the dependent variable until the child fails or is censored. Once the data have been structured in this way, the time-varying variable can easily be assigned to each child. This will reflect the point in time the child observation refers to. The time dependent covariates used in the model are external and ancillary (Kalbfleisch and Prentice, (1980). They are external because they are not directly involved with the process leading to dropouts, and they are ancillary because they are the result of a stochastic process external to the individual, whose marginal probability distribution does not involve the parameters of the model.

4.5.2 Estimation of the education expenditure model

The education expenditure model estimates the determinants of schooling of different household groups. This model rests on the assumption that progression in school is correlated to education expenditure. One advantage of this model over the school attainment model is that the dependent variable and the explanatory variables are observed at the same time and no misspecification bias arises when using the second to explain the first. A second advantage of this model is that the analysis can be performed separately for different household groups because income categories are not endogoneous to school choice, as it was the case for the school attainment model. The main disadvantage of this approach is that expenditure is observed at the household level, rather than at the child level, hence the dependent variable cannot be explained by factors specific to a particular child.

A convenient form for the estimation of education expenditure is the Working-Leser specification (Deaton and Muellbauer, 1980a). The advantages of this functional form were discussed in Section 2.5.1, and are only briefly summarised here. This form satisfies standard restrictions of demand theory, such as homogeneity and adding-up, and provides a realistic representation of consumer behaviour allowing for luxuries, necessities, and inferior goods. This specification models Engel curves that relate linearly household's budget shares (w) of each good (j) to the logarithm of total expenditure ($\ln x$):

$$w_j = a_j + b_j \ln x \tag{4.17}$$

The education expenditure model estimates household education Engel curves. Shares of expenditure on education (w_e) for this model were obtained from the NSSO data by aggregating expenditures collected using a one-year recall period. The NSSO surveys collect detailed information on household expenditure on education disaggregated in the following way: purchase of books; purchase of stationery; payment of school fees; private tutoring; the cost of bus transport to school; and the payment of library cards. Some relevant education costs, like the purchase of school uniforms and the private cost of transport to school, are not explicitly coded but are likely to be included in the residual category called other education expenditure. All education expenditures, with the exception of bus transport costs, are reported on both annual and monthly recall basis. As education expenditures are made occasionally and at specific times of the year, like for example the beginning of the school year, expenditure from the annual recall rather than the monthly recall were used in order to reduce the number of observations with value zero. The shares were obtained by dividing education expenditure by total household expenditure in the course of the year.[26] Expenditure data were also adjusted for regional and seasonal variation in prices by calculating household specific price indices, using the methodology outlined in Section 3.4.2. In order to make expenditure data from the four surveys comparable in real terms, the expenditure figures were further adjusted by the annual CPIAL of Andhra Pradesh provided by the DES.

For estimation purposes the model of equation (17) is expanded to include sets of demographic and control variables:

$$w_{ei} = a + b \ln \frac{x_i}{s_i} + \sum_{j=1}^{r} c_j D_{ji} + d \ln s_i + \sum_{h}^{n} e_h Z_{hi} + \eta_i \tag{4.18}$$

Household expenditure shares are explained by the logarithm of household per capita expenditure (total household expenditure x_i over household size s_i), the demographic composition of the household in terms of the ratios D of household members of a given age and sex category (j) over household size; the logarithm of household size; and a set of control variables Z. The latter includes all the variables that are likely to have an

[26] Expenditure categories, like food and miscellaneous non-food items that are reported on a monthly recall basis only, were multiplied by 12 before adding them to other household expenditure in order to obtain total household expenditure.

impact on educational choices. These are the same variables included in the schooling equation (16): school costs; returns to schooling; determinants of returns to schooling; factors affecting the ability to borrow; and expected income variability.

A serious limitation in estimating equation (4.8) is that the majority of households in the sample reported zero expenditure on education. If we consider the whole sample of rural households of the four survey rounds, the average percentage of observations with zero expenditure on education is 62%. This percentage however has been decreasing over the years, from 81% in the 43rd round, to 60% in the 50th round, 55% in the 55th round, and 46% in the 61st round.

A model widely used in estimating demand equations with many zero values is the tobit model of Tobin (1958), which was originally developed to estimate household expenditure on durable goods. The tobit model can be seen as a hybrid of the probit model and ordinary least square. One shortcoming of the tobit model is that it produces biased parameter estimates in the presence of heteroscedasticity (Deaton, 1997), which is the norm when dealing with expenditure data from household surveys. Several alternatives to the tobit model have been put forward in the econometric literature. These are all based on different interpretations given to the observations with zero values. Some models consider these observations as censored, while other models consider zero observations as household decisions and suggest the use of selection models. None of the proposed alternatives to the tobit model is convincingly superior to the others. The strategy adopted in this study is therefore to estimate equation (4.18) using a battery of different models and identify the parameter values that are robust to different model specifications. The models used for this task are: standard OLS, the heteroscedastic tobit, a semiparametric version of the tobit model, and the two-part model. The characteristics of these models will now be discussed in turn.

The first model used is standard OLS. Using OLS implies ignoring both the problems of censoring and selectivity. This model implicitly considers zero expenditure as a genuine household's decision, but does not attribute this decision to any explanatory factors. The model estimates an average effect of the explanatory variable over both spenders and non-spenders. Given the distribution of the data points over total expenditure, with many zero observations on the left side of the expenditure distribution, the method of least square tends to produce expenditure coefficients that are low compared to those produced by other models.

The second model used is the heteroscedastic tobit. Heteroscedasticity is a constant feature of household expenditure data and arises in this particular case because richer household enjoy more discretionality in their expenditure decisions. This results in an increasing dispersion of education expenditures as total expenditure increases. Parameters estimated by the tobit model in the presence of heteroscedasticity therefore tend to be large compared to those obtained by other methods. One way of correcting the tobit model for heteroscedasticity is to make reasonable assumptions about the form of the heteroscedasticity and to substitute estimates of the standard deviation of the residuals in the likelihood function of the standard tobit model (Maddala, 1983, page 180):

$$\ln L = \sum_{0} \ln\left[1 - \Phi\left(\frac{\beta' X_i}{\sigma_i}\right)\right] + \sum_{1} \ln\left[\frac{1}{\sigma_i}\varphi\left(\frac{y_i - \beta' X_i}{\sigma_i}\right)\right] \tag{4.19}$$

where the first summation sign refers to the observations for which expenditure is zero, and the second summation refers to the observations for which expenditure is larger than zero. The standard deviation of the residuals σ_i is simultaneously estimated from the data by maximum likelihood using a linear specification $\sigma_i = a_i + \sum_j^n b_j x_i$, where the *x* are all the explanatory variables of the original model or, alternatively, those that were found significant after running an OLS regression of the square residuals on all explanatory variables of the original model.

The third model used is a semiparametric version of the tobit model discussed in Deaton (1997), which uses the censored least absolute deviation estimator. The algorithm to perform the estimation has been suggested by Buchinsky (1994, p. 412) and can be summarised in the following way. A median regression is run on the sample and the predicted values are calculated. All observations whose predicted values are negative are discarded, and a new median regression is performed on the new reduced sample. The predicted values are again calculated for the entire sample and those with negative values are discarded. This procedure is iterated until convergence is achieved, which occurs when the number of observations in two consecutive estimations is the same. When the semiparametric version of the tobit model was run on the samples used for the present analysis, convergence was not always achieved and after a number of iterations the number of observations would oscillate around a set of close values. In these cases, parameter estimates of regression with the lower values of the sum of the absolute deviations from the median were considered.

The last model used is the two-part model of Duan et al. (1983). This model separates the behaviour of spenders and non-spenders in two stages. In the first stage consumers decide whether to spend on education or not, a decision that can be estimated by a probit model. In the second stage the consumers decide how much to spend conditional on having decided to spend, which can be estimated by OLS after transforming the dependent variable in logarithmic form. The composed likelihood function is such that its estimation is equivalent to the separate estimation of the probit model for the whole sample and the least squares estimation of the logarithm of expenditure for the sample of spenders. The two-part model is therefore equivalent to running OLS regressions of the logarithm of per capita expenditure for households reporting positive expenditures.

The education expenditure model was estimated using the four econometric specifications outlined above separately for household groups sharing specific characteristics: farmers and agricultural labourers, with and without access to irrigation. In order to avoid selection bias in the estimation of the parameters, the expenditure equations were estimated using the standard Heckman two-step procedure. First, a probit model was run to explain the probability of being in a given household category. Second, the inverse Mill’s ratio obtained from the probit model was included in the final expenditure regression. The selection equation models used are the same that were estimated in Chapter 3 to model the choice of farming and agricultural labour occupation.

4.6 Descriptive statistics

This section presents recent trends of educational indicators in Andhra Pradesh. First, trends in literacy rates across gender and geographic areas are calculated using census population data. Second, completion rates of different household groups are calculated using NSSO data. Finally, trends in household education expenditure are discussed.

4.6.1 Trends in literacy rates

According to the census of India of 2001, Andhra Pradesh ranked 22nd among 28 Indian states in terms of adult literacy, despite having a per capita income above the Indian average. Three main reasons are advanced to explain the poor educational performance of the state. These are the historical legacy of the pre-independence period, the low state investments in primary education, and the poor quality of schools (Reddy and Rao, 2003).

Vaidyanathan and Gopinathan (2001) emphasise the role of historical and social factors in shaping educational policies across Indian states. State efforts, social mobilisation, and missionary work were extremely important in promoting primary education in many parts of India, but played little or no role in Andhra Pradesh (Reddy and Rao, 2003). Banerjee et al. (2005) investigate how social divisiveness produced by the British rule, landlord-based tenure systems and caste fragmentation affected public investments across Indian states during the post-independence period. They find that all these factors, and caste divisiveness in particular, had some negative impact on education investments. Banerjee et al. (2005) also recognise that some princely independent states, like the state of Hyderabad in Andhra Pradesh, were more divisive and authoritarian than the British rule. Reddy and Rao (2003) confirm that the *Nizams* of the princely state of Hyderabad, ruling over the modern Telangana region, never promoted public education, while the local landlords – the *zamindars* - never encouraged universal education, and the work of missionaries was concentrated in very small areas of the rich Godavari and Krishna regions.[27]

The task of universalising primary education was undertaken only after the formation of the state in 1956. Education spending has been traditionally neglected by the state governments of Andhra Pradesh. Data on public expenditure between 1980 and 1996 show that the state invested less than 3% of the state GDP on education, which is below

[27] *Zamindars* were the landlords of pre-independence rural India. The institution has many similarities with the lords of medieval Europe. The *Zamindars* were preserved under British rule as tax collectors. The system was abolished soon after Independence.

the Indian average of 3.8% (Reddy and Rao, 2003). In addition, only around 40% of education expenditure was directed to primary education, which indicates a preference for investments in higher education (Reddy and Rao, 2003). Sipahimalani (2000) reports that spending per student in Andhra Pradesh was the lowest among all major states except West Bengal in both 1991-92 and 1997-98, and that the proportion of GDP spent in education by Andhra Pradesh over the 1990s fell slightly.

Rural villages in Andhra Pradesh are reasonably provided with primary school buildings. The spatial concentration of rural agglomerations has helped to determine this pattern. In Andhra Pradesh, 70% of the population lives in villages above 2,000 inhabitants compared to 50% in the rest of India. Data from the *All India Education Survey* (NCERT, 1997) show that in 1993, 70% of villages in Andhra Pradesh had a primary school and nearly 90% of villages had a school within a distance of 1km. However, the number of middle schools is much lower. The survey found that only 14% of villages had a middle school, though 80% had a middle school within a distance of 3km. Investments in education infrastructure was uneven across villages and some areas of the state have been neglected. Twenty per cent of villages predominantly inhabited by tribal communities do not have a primary school within a distance of 3km.

Though the number of schools in the state is relatively large, the quality of school infrastructure is rather poor. Data of the government of Andhra Pradesh, documented by Reddy and Rao (2003), report that only 75% of primary schools are *pukka* buildings (made of durable materials); half of them have a single room; and 40% have a single teacher. The provision of basic facilities, like water, is assured only in 30% of buildings, while only 14% of schools have toilets and a mere 8% have separate toilets for girls. These percentages are well below the minimum standards set by the Government of India.

Despite this rather grim picture of the state of education infrastructure in Andhra Pradesh, considerable progress has been made particularly during the last 20 years. Table 4.5 shows the literacy rates calculated using the last five rounds of the population census of India (Office of the Director of Census Operations, 2001),[28] together with their growth rates.[29] There is a considerable gap between the literacy rates observed in urban and rural areas. This gap however has reduced over the years as literacy rates increased substantially in rural areas particularly since 1991. A gap of comparable size exists between literacy rates of males and females living in rural areas. This gap however has also narrowed significantly, and the latest census data show a dramatic increase in female literacy rates.

Table 4.5 Literacy rates in Andhra Pradesh (1961-2001)

	1961	1971	1981	1991	2001	Growth rate
Rural literacy	16.7	19.1	22.9	29.7	47.0	2.7
Urban literacy	30.9	38.5	50.8	56.2	66.8	2.0
Overall literacy	21.2	24.6	29.9	36.8	52.4	2.4
Rural male literacy	25.0	27.2	32.0	39.3	56.3	2.2
Rural female literacy	8.3	10.8	13.9	19.9	37.6	4.2

Source: calculated from data of the Census of India (2001).

Table 4.6 shows the rural literacy rates by district in order to see whether the spatial difference in educational achievements has increased or decreased over the years. There is considerable inter district variation in literacy rates, with the districts located in the more developed Coastal area of the state showing better educational performance. Differences between districts have however reduced over time. The growth rates in the last column show that literacy grew faster in districts with lower initial literacy rates. The coefficients of variation reported in the bottom row of the table show a steady decline over time in the dispersion of literacy rates across districts. The correlation of literacy growth rates with poor initial conditions, and the decline of the coefficient of variation over time imply that the districts have became more equal to each other over time.

[28] These are not the adult literacy rates reported in official statistics. The latter are calculated as the number of literate persons above the age of six. The version of the census population data used in this study does not contain information on the number of persons under six years of age. Hence, literacy rates were calculated over population of all ages. Inevitably, this results in an underestimation of the true literacy rate, because a large number of individuals were too young to be attending school at the time of the census.

[29] Growth rates are per annum and were obtained from the coefficient of regressions of the logarithm of the literacy rate over time.

Table 4.6 Literacy rates in Andhra Pradesh by district (1961-2001)

	1961	1971	1981	1991	2001	growth rate
Adilabad	8.1	10.1	13.5	20.9	38.9	3.9
Anantapur	16.3	18.9	23.4	30.0	44.6	2.5
Chittoor	17.6	21.4	26.6	37.6	54.7	2.8
Cuddapah	18.6	20.9	26.3	36.2	52.1	2.6
East Godavari	22.6	26.8	30.4	36.1	54.2	2.0
Guntur	24.2	25.6	30.8	34.1	50.9	1.8
Karimnagar	11.1	12.6	17.0	26.1	43.4	3.4
Khammam	12.2	14.6	19.8	27.4	44.4	3.2
Krishna	25.9	28.6	33.9	37.5	56.9	1.8
Kurnool	18.3	19.3	23.2	26.8	41.3	1.9
Mahbubnagar	11.2	13.2	16.1	20.3	34.3	2.7
Medak	12.2	13.6	17.7	22.0	40.2	2.9
Nalgonda	12.0	14.8	18.8	27.6	45.9	3.3
Nellore	19.0	21.6	26.0	35.0	53.7	2.6
Nizamabad	11.3	13.3	16.8	23.2	40.8	3.1
Prakasam	18.4	20.2	25.6	30.1	47.4	2.3
Ranga Reddi	14.0	14.7	19.4	25.8	44.6	2.9
Srikakulam	14.1	16.6	19.6	26.9	45.8	2.8
Visakhapatnam	10.5	12.8	15.3	23.8	41.0	3.3
Vizianagaram	13.1	15.1	16.9	23.8	39.8	2.7
Warangal	11.7	13.7	17.7	26.4	44.3	3.3
West Godavari	27.5	31.0	33.6	41.6	63.0	1.9
Coefficient of variation	0.34	0.32	0.28	0.21	0.15	

Source: calculated from data of the Census of India (2001).

Part of the progress made in education in Andhra Pradesh may be due to a series of educational projects implemented by the Government of India and by the state of Andhra Pradesh over the last 20 years. This claim cannot be backed by evidence, because none of these projects has been rigorously evaluated (Kingdon, 2007). The two most important education projects currently being implemented in Andhra Pradesh are the *Sarva Shiksha Abhiyan* (Campaign for Universal Education, former DPEP), and the midday meal scheme.[30] The *Sarva Shiksha Abhiyan* aims at reaching universal primary education by 2010, and focuses on enrolling out-of-school children and improving school quality. Project inputs include the provision of additional teachers, civil works, teacher training, grants for teaching materials, free textbooks to low caste and female pupils, grants for disable pupils, and special facilities for girls. The midday meal scheme provides every child in government and government-aided school with a meal

[30] DPEP is the District Primary Education Programme launched in 2001 by the Government of India with the support of the World Bank.

containing the minimum requirement of calories and proteins. The scheme was scaled up to the national level in 2006 and is meant to serve school children for 200 days a year. The Government of India pays for the cost of transport, food, and food preparation, while state governments provide kitchens equipped with all necessary utensils.

Other projects, now phased out but relevant for the period considered in the present analysis, include the Operation Black Board (OBB) and the Andhra Pradesh Primary Education Project (APEP). The OBB was funded by the Government of India, and focused on school infrastructure. It aimed at providing all primary schools in India with at least two *pukka* classrooms, two teachers, and blackboards, maps, charts, and a library. The APEP, which merged with DPEP after 1995, was funded by ODA and aimed at improving school quality through teacher training and school construction.[31]

It is important to note here that both past and present education programmes have focused on improving access to school for girls and children of disadvantaged communities. This was also the goal of a number of campaigns promoted by the Government of Andhra Pradesh, like *Mabadi* (our school) and *Chanduvkundam* (back to school), which were targeted to women and children living in small and disadvantaged communities.

4.6.2 Completion rates by household groups

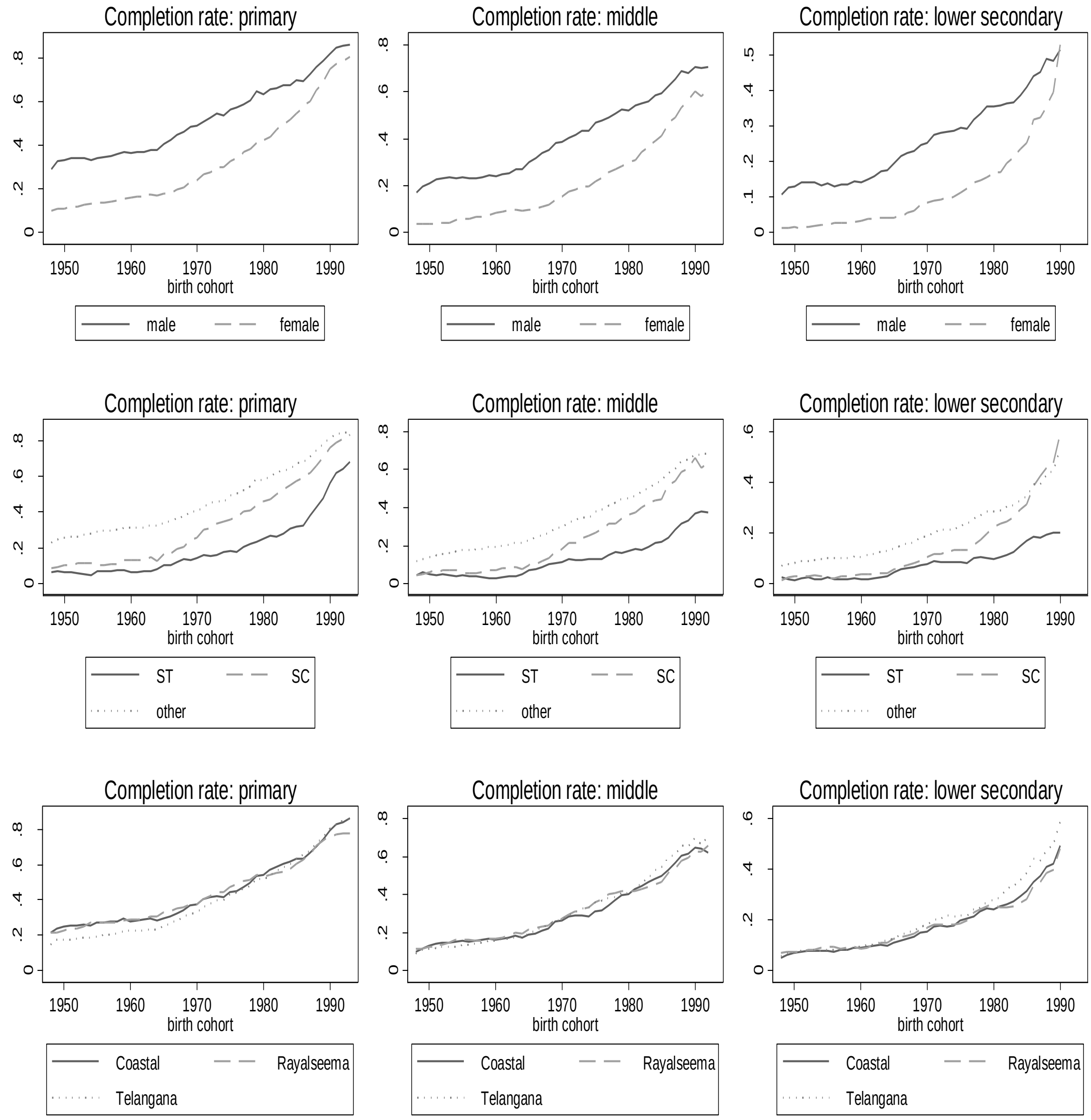
The census data do not report literacy rates by the occupational categories used in the present analysis (farmers and agricultural labourers with and without access to irrigation). In addition, the census only provides literacy rates and no other data on educational achievements. Therefore, the NSSO data, which cover a shorter period of time, will be used to analyse how different socioeconomic groups progress in school.

The charts in Figure 4.3 show completion rates in primary, middle and lower secondary school disaggregated by gender, social group, and regional location. The charts plot average completion rates over annual birth cohorts of rural individuals. By doing so, educational attainments for a period of 30 years prior to the oldest NSSO survey

[31] ODA was the United Kingdom Overseas Development Administration, which was replaced in 1997 by the Department for International Development (DfID).

available can be calculated. The birth cohorts were constructed in the following way. The data from the four survey rounds where joined and the average completion rates were computed for each birth cohort from Independence to the year 1992. In order to avoid censoring of observations, only children above 12 were considered for primary school, only children above 15 for middle school, and only children above 17 for lower secondary. The data points were also smoothed around a five-year moving average to make the reading of the chart easier.

Figure 4.3 Completion rates by birth cohort (1945-1992)



Source: calculated from NSSO data.

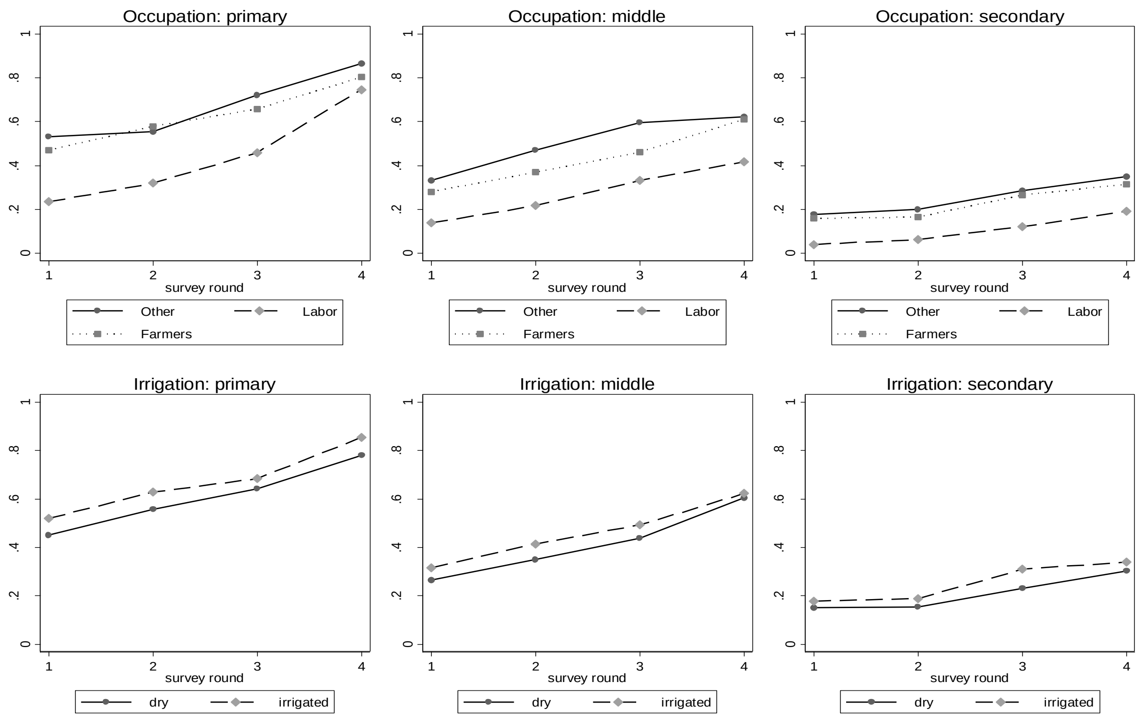
The three top charts of Figure 4.3 compare completion rates of male and female birth cohorts for the three educational levels. Girls have almost closed the large gap in primary completion rates they had with boys at the time of Independence. As the chart shows, much of the reduction in the gap occurred in the last 15 years. Differences between male and female completion rates are however still large in middle school. Interestingly, the female completion rate in secondary school has increased exponentially from the early 1980s and the gender gap has been eliminated.

The three charts in the middle of Figure 4.3 show completion rates by social groups. Theses reflect differences in income, attitudes towards education, discrimination policies (both negative and positive) and discrimination in labour markets. The charts show that children from scheduled castes have closed the gap with members of other backward castes and higher castes at all three school levels. Children of scheduled tribes however, continue to lag behind. While the gap in primary completion rates between them and other social groups has narrowed over the last 15 years, it has remained large. Worse still, the gap in completion rates in middle and secondary schools between scheduled tribes and other social groups has increased over time.

The charts at the bottom of Figure 4.3 show the completion rates in the three macro-regions of Andhra Pradesh. The data used to draw these charts should be considered with caution because they ignore migration. The region where the interview took place is not necessarily the region where the person was born or attended school, and many children who were born and studied in rural areas had migrated to towns by the time of the interview. However, assuming that migration patterns in the three macro regions are comparable, the charts show that completion rates are very similar in the three regions. The rich and more developed Coastal region does not have higher completion rates than the other two regions. It may be also observed that the Telangana region, which was lagging behind the other two regions after Independence, is now ahead at all school levels, while the Rayalseema region begun lagging behind the other two regions from the mid 1980s.

The charts in Figure 4.4 show completion rates of different income groups: farmers, agricultural labourers and other households living in rural areas. The three charts at the bottom further disaggregate data on farmers by whether the farm household has access to irrigation or not. Households were assigned to the occupation groups used in Figure 4.4 based on the criteria described in Section 3.5.2. A farm was hence considered irrigated if any portion of the cultivated land is irrigated irrespectively of size, while agricultural labour households with access to irrigation are those who live in irrigated villages.

Figure 4.4 Completion rates by occupational group (1987-2005)



Note: survey rounds are 1987-88 (1), 1993-94 (2), 1999-00 (3), and 2004-05 (4).
Source: calculated from NSSO data.

Birth cohorts of farm and agricultural labour households and of households with and without access to irrigation cannot be calculated because occupational group and irrigation status vary over time. Farms that now have access to irrigation might not have been irrigated ten years earlier, and households whose current main source of income is farming might have been agricultural labourers a few years before. In addition, the occupational status can be endogenous because household who invested in schooling in the past are less likely to be farmers and agricultural labourers today. In order to use data that is representative of each occupation category, the analysis was limited to the completion rates of school age children in the five years preceding the survey. In addition, at each data point the sample was limited to children older than the minimum age required for each school level in order to avoid censoring. The four data points corresponding to each survey round were connected by simple linear interpolation.

The top three charts in Figure 4.4 show that occupational groups can be ranked in the following order in terms of completion rates. Households working in non-agricultural activities have the highest completion rates, followed by farmer households and, at a considerable distance, by agricultural labour households. It appears that in primary school the gap between agricultural labourers and other households has decreased over the period considered. However, the gap in the completion of middle and secondary school has remained stable. The three bottom charts of Figure 4.4 show that children of households with access to irrigation have higher completion rates than children of non-irrigated farms. These differences might simply reflect income and location effects, as irrigated farms are on average richer and more likely to be located in areas that are more socioeconomically developed. The difference in completion rates for primary and middle school between the two groups has not changed over time. There was however a slight increase in the difference of completion rates at the secondary school level between the two groups.

4.6.3 Expenditure on education in Andhra Pradesh

According to the Indian Constitution, education is compulsory and free up to the age of 14, which is the official age for the completion of middle school. In practice however, rural households incur several costs in sending their children to primary school. These costs can be divided in three categories: opportunity costs, compulsory costs, and discretionary costs.

The opportunity cost of schooling consists of the income forgone because the child is not working. The size of the opportunity cost depends on the numbers of hours the child would work if not in school and on the wage paid for these hours. Data collected by the PROBE team (1999) found that a small percentage of children (20% of boys and 22% of girls) was engaged in full time work and that an even smaller percentage (5% for boys and 1% for girls) was employed in wage labour. Most child work consists of housework spent on activities like fetching wood and water, grazing and child minding. However, to the extent to which this work substitutes for work otherwise performed by adults, and therefore releases adult time for paid work, it does have a monetary value.

In principle the opportunity cost of schooling might be calculated by multiplying the number of hours spent studying by some average wage. This simple computation however ignores that children attending school also contribute to housework. The PROBE team estimated that the average number of working hours per day is four for boys and five for girls. The PROBE team also estimated that out-of-school children – both boys and girls – work approximately two hours more than children attending school. As the average number of school days in a year in Indian villages is 150, children not attending school work approximately 300 extra hours. These data suggest that the opportunity cost of schooling is relatively small. However, given that the monetary costs of schooling are small, and that many poor household do not spend on education, the opportunity cost is probably the most important component of overall school costs. In addition, opportunity costs increase over time for each household, because as children grow, they can earn higher wages in the labour market or help the household more efficiently in unpaid work.

Table 4.7 reports the average expenditures on education of rural households in Andhra Pradesh and the percentage of households with non-zero expenditures for a sample of households with children aged between six and 17. School expenditure has increased considerably over the last 20 years, as has the number of households reporting non-zero expenditure. Education expenditure is approximately 3% of average household expenditure.

Table 4.7 Average household expenditure on education

	1987-88	1993-94	1999-00	2004-05
Average expenditure[a]	173	272	392	807
Percentage with non-zero expenditure	22.3	56.5	65.6	75.6
Average non-zero expenditure[a]	775	481	597	1,067

Note: [a]expenditure is in 2005 Rupees.
Source: calculated from NSSO data.

Monetary education expenditure can be subdivided into compulsory expenditure and discretionary expenditure. Compulsory expenditure consists of the payment of school fees. However, research conducted by the PROBE team found that public primary school charge only negligible fees, and most monetary expenditure is discretionary. Table 4.8 reports the shares of education expenditure by item as reported by the NSSO expenditure survey for households with school age children. One important item, school uniform, is missing from the list and is possibly reported under the voice ‘other education expenditure’. This omission is unfortunate, because the cost of school uniforms can be substantial and is paid by the large majority of households. Total household education expenditure might therefore be underestimated, though it is difficult to say by how much.

Table 4.8 Share of education expenditures by item

	1987-88	1993-94	1999-00	2004-05
Books	25.7	32.0	28.0	16.4
Library	0.2	0.9	0.1	0.1
Stationery	21.5	15.7	16.9	15.7
Fees	27.1	32.9	44.2	60.1
Private tutor	7.8	10.2	6.5	2.0
Bus	0.1	0.7	1.0	4.6
Other	17.7	7.5	3.2	0.1

Source: calculated from NSSO data.

The purchase of books and stationery, and fee payment account for the majority of household expenditure on education. School fees today account for more than 50% of household expenditure on education and have been increasing over the period considered. Given that fees for public schools are negligible, the rise in fee expenditure over the last ten years is likely to reflect the spread of private education in rural areas. Muralindharam and Kremer (2006) conducted a large scale survey of Indian public and private schools in 2003 and found that 30% of the villages sampled in Andhra Pradesh had a private school. The NSSO data do not indicate whether children attend public or private schools, or whether there is a private school in the village surveyed. It is therefore not possible to investigate the increase in private education for the sample considered here.

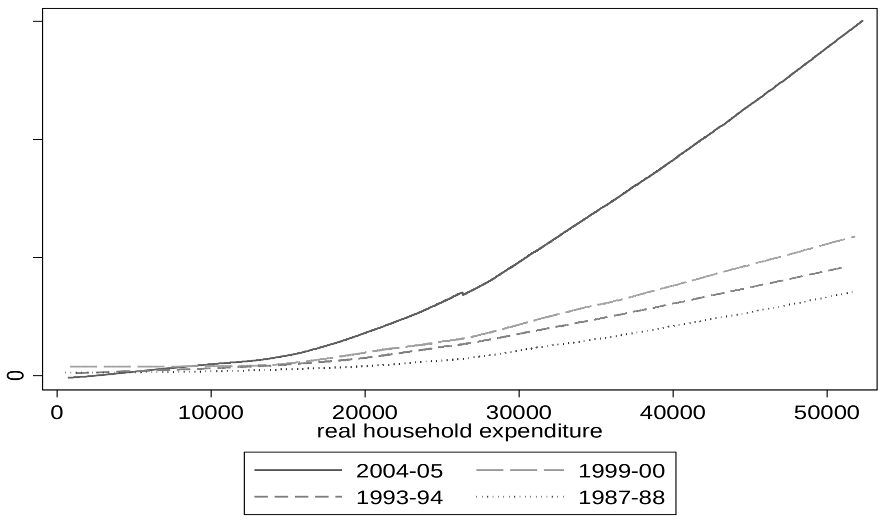
The main reason behind the expansion of private schooling is likely to be the poor quality of government schools. Muralindharam and Kremer (2006) found that the presence of private schools in rural villages was positively correlated with teacher absence in public schools and negatively correlated with income per capita in the state. This suggests that private schools develop in areas where the quality of public schools is poor, independently of income per capita. They also found that teacher absenteeism in

Andhra Pradesh was above 20% and that only 50% of the teachers present were actively engaged in teaching. These percentages are in line with those found for all India.

Household expenditure on bus transport to school has increased over the last 20 years. This may reflect the increase in attendance of schools outside the village in particular middle and secondary schools, which are not available in the majority of villages. Tilak (2002) uses data from a survey of 33,000 households conducted by the National Council of Applied Economic Research in 1,765 villages of 16 Indian states in 1994 and finds that household expenditure on education increases with the level of education. He also finds that, on average, the cost of secondary education is nearly twice that of primary education due to fee payment and transport.

The chart in Figure 4.5 shows the relationship between expenditure on education and total household expenditure in the four NSSO surveys. The curves were obtained nonparametrically using LOWESS on samples of rural households with school age children. In order to make the data comparable across surveys, the expenditure figures are all expressed in 2005 Rupees using appropriate multipliers obtained from the Andhra Pradesh CPIAL.

Figure 4.5 Education expenditure and total household expenditure



Source: calculated from NSSO data.

The upward shift of the curves shows that education expenditure increased over time, and that the largest increase was recorded between the last two survey rounds. The curves are flat on the left of the chart due to many zero values and because expenditure of rich households is several times that of poorer ones. The shape of the curves is nearly exponential, which indicates an expenditure elasticity larger than one.

An expenditure elasticity larger than one implies that the share of education expenditure over total expenditure increases with income and over time. Table 4.9 shows the shares of education expenditure for different household income groups across the NSSO surveys rounds. The share of education expenditure has increased over the four survey rounds from 0.7% to nearly 3%. Households not engaged in agricultural activities spend proportionally more on education than farmer households, and the latter spend proportionally more than agricultural labour households. Expenditure shares of all household groups are increasing over time, though the increase is less pronounced for households not engaged in agriculture. The last two rows at the bottom of Table 4.9 show the education expenditures shares of farmers with and without access to irrigation. Education expenditure of irrigated farms is larger in all survey rounds and the expenditure shares of the two groups are increasing over time at a similar rate.

Table 4.9 Education expenditure shares by occupational group

	1987-88	1993-94	1999-00	2004-05
All households	0.7	1.1	1.5	2.9
Agricultural labourers	0.4	0.6	0.8	1.7
Farmers	0.8	1.2	1.6	3.6
Others	1.2	1.6	2.7	3.7
Irrigated farms	0.9	1.5	1.9	4.9
Non-irrigated farms	0.7	1.0	1.4	3.1

Source: calculated from NSSO data.

4.7 Empirical results

This section presents the results of the estimation of determinants of schooling (equation 4.16), using the duration model with time covariates described in Section 4.5.1. Before presenting the results of the empirical analysis, the explanatory variables included in the model and their potential effect on the dependent variable are discussed.

4.7.1 School attainment model: variables used

The school attainment model estimates the determinants of children's progression through school and assesses the effect of income variability on households' education investments. The dependent variable is the highest school level completed by children between the age of five and 18 over five school level categories described in Section 4.5.1: illiterate; below primary; primary; middle; and lower secondary. The model uses a large set of explanatory variables in order to take into account all possible determinants of schooling. These are now discussed in the order in which they appear in Table 4.10.

The econometric specification chosen consists of a duration model with time covariates. The latter are: demographic variables; wages; school availability; and rainfall variability. The construction of each time-varying variable and the rationale for its inclusion in the empirical model will be discussed in detail. The first variable of the model is a dummy for female children. Girls' educational achievement in rural India are notoriously poorer than boys', though as shown in Section 4.6.2, the gap has significantly narrowed in Andhra Pradesh over the last 15 years. Parents' attitudes and conservative behaviours are often blamed for girls' poor educational achievements. For example, parents may not want their daughters to travel long distances to go to school because it is socially unacceptable. Pal (2004), using data from a sample of children collected in West Bengal in 1987-89, concludes that only one third of the of the difference in schooling between boys and girls can be attributed to the market characteristics included in the regression model. Economic reasons however are probably more important for girls' discrimination within the household. Job opportunities for women are limited in rural areas and households may not reap the benefits of girls' schooling if they leave the households once married (PROBE Team, 1999). In other words, parents may find it more profitable to invest their limited time and monetary resources on the schooling of boys.

Number of siblings captures the number of children in the household. This is a time-varying variable, whose value is measured at the time the schooling decision for each child is made. While household size changes frequently and unpredictably as new household members join the household and others leave to form new households, the number of sibling is more stable. The number of siblings can increase, as couples have more children, or decrease, as some children become adult. The past values of this variable however cannot be recovered with absolute precision from the NSSO data because we cannot account for child mortality and child fostering. While the latter is not common in rural India, mortality rates of children under five were as high as 10% in rural Andhra Pradesh in the 1990s and even higher in the 1980s, particularly for disadvantaged households. Therefore the number of sibling at the time the schooling decision was made could be underestimated, particularly for households with high rates of child mortality, like the very poor and scheduled castes and tribes. Despite this limitation, the time-varying number of siblings calculated from the household roster of the NSSO data is a better explanatory variable than the current number of siblings, though inevitably imprecise.

Table 4.10 Explanatory variables of the school attainment model

Variable	*Description*
Female child	One if the child is female
Siblings	Number of children in the household (time varying)
Birth order	Child’s rank by age among her siblings
Father’s and mother’s education	Education level of the parents: illiterate, below primary, primary, lower secondary, senior secondary, higher.
Muslim household	One if the household is Muslim
Christian household	One if the household is Christian
Scheduled caste	One if household belongs to scheduled castes
Scheduled tribe	One if household belongs to scheduled tribes
Land owned	Number of hectares of land owned, including uncultivated and leased land
Traditional fuel	Wood or dung is used as primary source for cooking
Electricity	Electricity is the primary source for lighting in the home
Skilled wage	Average wage of carpenters in the district (time varying)
Unskilled wage	Average wage of male field labour in the district (time varying)
Schools per capita	Number of primary, middle and secondary schools per head by district (time varying)
Rainfall variability	Standard deviation of rainfall during the ten years preceding the period relevant to the schooling choice (time varying)
Time of schooling decision	five-year time intervals from 1976 to 2005 in which the school choice is made
Districts	Dummies for 23 administrative district subdivision

In households with many children, each child receives a small fraction of the resources available to the household, which includes parents’ time as well as economic resources.

In environments where school quality is poor, like rural areas of Andhra Pradesh, children require considerable parents' attention, for example in the form of motivation or help in doing homework. The household time constraint may thus play unfavourably for children living in households with many children (Hanushek, 1992).

The birth order of the child captures the child's ranking in terms of birth with respect to her siblings. The first born in the household has rank one, the second has rank two, and so forth. This variable is fixed in time as the birth order cannot change. It is obtained from the household roster and not from the mother's birth history. Therefore, birth order is underestimated because it is affected by the mortality bias discussed above. Birth order is positively correlated with child schooling. In rural areas, children often contribute to the housework and girls in particular are responsible for caring for younger siblings while parents are away for work. Older children are more likely to be employed in housework and children of higher order have better opportunities of attending school (PROBE Team, 1999). Ota and Moffat (2007) use data from 100 households from six villages of Andhra Pradesh to analyse the impact of birth order on schooling decisions. In addition to the operation of a 'resource dilution' effect, which predicts that for lack of resources a family has to prioritise the education of some children over others, they suggest the operation of a 'teaching effect', whereby siblings have a positive impact on schooling achievements of younger children by learning sharing. Their empirical analysis finds that first-born children are less likely to be in school regardless of gender and that the likelihood of being in school increases with the birth order of the child. They also find that younger children are more likely to be in school if older siblings are girls, which they explain with the need of elder sister to work to provide resources for education of the younger siblings rather than with the need of caring for the infants. A different result is obtained by Krishnaji (2001) who uses household data from two districts of Andhra Pradesh (Mahaboobnagar and Adilabad) in order to assess the determinants of school enrolment at the village and household level. His results suggest that first-born children have a higher probability of being in school.

Father's and mother's education are categorical variables representing the highest school level completed. These variables are fixed in time and uncensored, as it is assumed that parents are no longer studying at the time their children are attending school. The education categories used are: illiterate, below primary, primary, middle, lower secondary, upper secondary and higher. Parents' education may affect child schooling in different ways. First, more educated parents are more favourably inclined to child schooling, while uneducated households may not perceive the advantages of education. Secondly, more educated parents may support their children by providing motivation, help with the homework and a role model. Third, and more importantly for the purpose of the present analysis, education is a strong correlate of household income. Household expenditure reported at the time of the interview cannot be used to measure household income, unless some radical assumptions are made regarding the equality between expenditure and permanent income. First, current expenditure may be very different from expenditure at the time when schooling decisions were made. Second, current expenditure is partly endogenous to past schooling decisions. In the absence of household expenditure, parents' education represents a good portion of household income generating capacity.

Religion may influence attitudes toward education. The empirical model includes a dummy variable for Muslim households and another for Christian households. Muslims and Christians are religious minorities in rural Andhra Pradesh. Ninety-two per cent of rural household in the NSSO data are Hindu, 4% are Muslim and 3% are Christian. Figures for other religious affiliations are negligible.

The model includes dummy variables for households belonging to scheduled castes and tribes. These groups are notoriously victims of discriminatory practices in rural India, and they represent 22% and 7% of our sample respectively. Children of scheduled castes and tribes are less likely to progress through primary school for several reasons. First, members of scheduled castes and tribes are discriminated in labour markets and their perceptions of returns to schooling are lower than those of members of higher castes. Second, scheduled castes and tribal communities are discriminated by state policies. An example of this is the paucity of state investments in education in areas inhabited by the scheduled tribes. Third, children of disadvantaged groups are victim of social exclusion. Schools, though accessible in the village, may be socially distant (PROBE Team, 1999) if scheduled castes and tribes live in remote and isolated hamlets, or if access to school is banned by families of higher castes. Against this backdrop of social exclusion and discrimination, it must be recognised that development projects are often targeted to improve the living conditions of disadvantaged groups and that state policies include positive discrimination in school and in the working place.

Land owned is the number of hectares owned by the households, including land that is not cultivated or is leased out. Some 80% of rural households in the sample own some land, though only 44% of households own more than one acre, which is the minimum amount for a farm to generate a minimum level of income. Land is the most important determinant of agricultural output and is a strong correlate of farm income. In addition, land is used as collateral for loans, hence it also captures households' ability to borrow.

Traditional fuel is a dummy variable for households that use woods or dung as their primary energy source for cooking instead of other methods, like charcoal, kerosene, gas and electricity. The number of households in our sample that are relying on traditional cooking fuel is very large though decreasing over time (93% in 1987-88, 94% in 1994-94, 87% in 1999-00, and 80% in 2004-05). This variable captures aspects of poor housing and living conditions and distance from urban centres and schools of good quality.

Electricity is a dummy variable for households that use electricity as their primary source of energy for lighting alternatively to other methods, like kerosene, gas and oil. The number of rural household using electricity is relatively small in our sample thought it has increased substantially over time (33% in 1987-88, 50% in 1994-94, 69% in 1999-00, and 84% in 2004-05). Electricity in rural areas is provided by the state and, though it is paid by the household, is a poor correlate of household income. The use of electricity however may be an indicator of urbanisation and of easier access to quality education. In addition, electricity helps schooling directly by allowing students to do their homework and attending classes in otherwise dark rooms.

The skilled wage is the wage for skilled labour at the district level described in Section 4.4.3. The district skilled wage is intended to represent parents' perceptions of future returns to schooling. It is assumed that parents' expectations about future wages are based on the average of current wages in the district. The skilled wage is a time-varying variable that was assigned to each child in the following way. Wage time series over the time period spanned by the four surveys (from 1975 to 2005) were obtained for each district. Three-year averages were calculated and assigned to each child observation based on when the schooling decision was made and child's age. For children who attained (or might have attained) below primary, the average wage for the period the child was aged five to seven was used. For children who completed (or might have completed) primary, the average wage for the period the child was eight to ten was used. For children who completed (or might have completed) middle school the average wage between the time the child was aged 11 and 13 was used. Finally, for children who completed (or might have completed) lower secondary school, the average wage for the period the child was 14 to 16 was used.

Unskilled wage is the wage for agricultural labour at the district level described in Section 4.4.3. This variable was obtained and assigned to each child in the same way as the skilled wage variable. Unskilled wage is intended to represent the opportunity cost of schooling. It is an indicator of the potential income obtainable if the child leaves school or if she substitute adults' work in the home thus releasing adult working time.

The number of schools per capita is the number of primary, middle and secondary school per 1,000 inhabitants calculated at the district level. These figures were obtained by dividing the numbers of school of each level by the number of individuals in the district reported by the population censuses of 1981, 1991 and 2001. The variable was then assigned to each child based on the time when the schooling decision was made and on the school level relevant to the decision. For example, if the decision to progress to middle school was made between 1995 and 2005, the number of middle school in the district from the 2001 census was used. The number of schools is intended to represent the quality of schooling at the district level. This is an imperfect indicator of school quality because it overlooks important quality characteristics like the quality of the facilities and of teaching. Moreover, most villages have a primary school and access to a middle school in the radius of one or two kilometres. However, the assumption is that other quality characteristics are correlated with the number of school buildings.

Rainfall variability is the standard deviation of district rainfall during the ten years prior to when the schooling decision was made. This is a time-varying variable which is assigned to each child based on the child's age and on the time when the schooling decision was made. The procedure followed for assigning the variable to each child observation is the same used for the assignment of the wage variables. There is no rigorous rationale for the choice of a time span of ten years for the calculation of the standard deviation. It seems reasonable to use a number of observations that provides a sufficient range of weather realisations and that at the same time does not go too far back in time. The standard deviation of district rainfall captures household expectations of rainfall variability and therefore expectations of agricultural output variability. It must be emphasised that this is a very rough indicator of household expectations of output variability. First, as already mentioned, the choice of using the standard deviation over ten years is arbitrary and not based on a psychological observation of farmers' behaviour. Second, the variability is calculated at the district level, though there might be large variability within each district. Finally, rainfall variability does not immediately translate into output variability, as other factors matter, like the intensity of rains, their spacing over the year and the monsoons onset. Despite these limitations, this variable provides a broad indication of the agricultural output variability expected by rural households. Its effect however cannot be measured with precision, because if there is an effect of income variability on education investments, this is likely to be obfuscated by the factors outlined above in the empirical estimation. The analysis however is fully legitimate from a qualitative point of view, because it allows to conclude whether there is an effect or not and to give a broad indication of the size of this effect.

The time of schooling decision is a series of variables acting as time trends dummies. They are meant to capture historical improvements in schooling not explained by other variables and determined by factors such as changes in incomes, public expenditure in education, and people's attitudes toward education. These variables were built in the following way. Children between the age of five and 18 were selected and the data from the four surveys were merged to create a single dataset spanning 30 years of schooling from 1975 to 2005. The data from the four surveys overlap in terms of schooling as, for example, 18-year-old children interviewed in the fourth survey in 2005 were in primary school at the time of the second survey in 1993-94. Given the spacing of the surveys, most observations (80%) lie in the time interval between 1985 and 2000 and fewer observations are located at the tails of the distribution, as they are covered only by the first and last surveys. The 30 years spanned by the four surveys were then divided into six five-year time intervals. Finally, each observation was assigned to a time interval based on the age of the child and her level of schooling at the time when the schooling decision was made.

The regression model also includes 23 dummies for each of the administrative districts in which the state is subdivided. The aim of these variables is to capture district specific characteristics that are not otherwise revealed by the data, like expenditure and quality, and the presence and efficacy of education projects at the district level.

4.7.2 School attainment model: coefficients estimates

Table 4.11 shows the results of the school attainment model for a sample of rural children aged from five to 18. A value of the hazard ratio larger than one means a larger probability of dropping out of school, while a value below one means a higher probability of completing school. In the case of dummy variables, like sex or social group, the hazard ratio is the ratio between the probability of dropping out of school for a child with a given characteristics and the probability of dropping out for a child without that characteristic. In the case of continuous variables, the hazard ratio is the ratio of probabilities after a one unit increase in the value of a particular characteristic. As a result, continuous variables tend to have small hazard ratios because they have large absolute average values. Small hazard ratios for continuous variables should not be confused with lack of explanatory power.

Most hazard ratios are highly significant and display dropout risk of the expected sign. Female children are at much higher risk of dropping out of school. The probability of completing school decreases with the number of siblings in the households and increases with the birth order of the child. The educational background of both parents is extremely, and equally, important for the child’s success in school. The household’s religious affiliation seems to matter only marginally. Children of Christian households are more likely to progress in school, while children of Muslim households are more likely to drop out, but the latter effect is not statistically significant. The findings on completion rates by social groups confirm those of the descriptive analysis conducted in Section 4.6.2. Children of scheduled caste households are only marginally more at risk of not completing school, and the coefficient is not statistically significant. Conversely, children of scheduled tribe households are at a very high risk of dropping out.

The size of land owned by the household has a positive effect on the probability of the child completing each school level. The value of the hazard ratio is low but notice that an increase in land holdings by three hectares, which implies a move from a small to large farm category, has a 10% impact on school progression at any school level. Children of households using traditional cooking methods and households with electricity are respectively less and more likely to progress through school. The hazard ratios associated with these variables are large and capture income effects, as well as factors like school distance and quality that are associated with urbanisation.

Wages of skilled and unskilled labour have the expected effect on school completion. The size of the skilled wage provides incentives to school progression, while the unskilled wage, which represents the opportunity cost of schooling, discourages schooling. The hazard ratios of skilled and unskilled wages are low, but it should be kept in mind that they represent changes in relative probabilities for changes in wages by one Rupee.[32] For example, an increase of the skilled wage by 50% would increase the completion probability by 9%, while an increase of the unskilled wage by 50% would decrease the completion probability by 5%. The number of schools per 1,000 individuals in the district has a significant and large effect on the probability of completing each school level. This variable shows the extent of the progress that can be made by investing more in school quality and infrastructure.

The model also includes time and location fixed effects. The time variables are important for two reasons. First, the model contains variables that are varying over time, wages in particular, and it is important to avoid spurious correlations between increases in wages and attainment levels. The diverging effects of the skilled and the unskilled wages, though the latter is not statistically significant, seem to suggest that the effects captured by these variables are not spurious. Second, the time variables represent historical changes not otherwise captured by the model. These variables show that children’s chances of staying in school have increased substantially over time. The regional dummies are highly significant and some hazard ratios are relatively high. Children living in the southern, dry and poor districts of Rayalseema and in the Telangana districts of Rangareddy, Adilabad, Mahaboobnagar and Prakasam are at

[32] The sample averages of skilled and unskilled wages are 70 and 47 Rupees respectively.

much higher risk of dropping out of school. Among all districts, the hazard ratios of Kurnool and Anantapoor are particularly high.

Table 4.11 School attainment model

	Hazard ratio	Standard error	Z-statistic	P-value
Female child	1.484***	0.028	21.2	0.000
Number of siblings	1.082***	0.007	12.6	0.000
Birth order	0.876***	0.011	-10.9	0.000
Father's education	0.775***	0.009	-21.3	0.000
Mother's education	0.776***	0.018	-10.8	0.000
Muslim household	1.042	0.041	1.1	0.290
Christian household	0.885**	0.054	-2.0	0.046
Scheduled caste	1.007	0.025	0.3	0.778
Scheduled tribe	1.351***	0.042	9.7	0.000
Land owned	0.972***	0.005	-5.8	0.000
Traditional fuel	1.753***	0.117	8.5	0.000
Electricity	0.736***	0.016	-14.5	0.000
Skilled wage	0.997**	0.001	-3.1	0.002
Unskilled wage	1.002*	0.001	1.3	0.179
Schools per capita	0.650***	0.059	-4.8	0.000
Rainfall variability	1.001**	0.000	2.9	0.004
Schooling decision time				
2001-2005	0.195***	0.014	-22.4	0.000
1996-2000	0.350***	0.017	-21.7	0.000
1991-1995	0.487***	0.019	-18.6	0.000
1986-1990	0.587***	0.021	-14.9	0.000
1981-1985	0.807***	0.032	-5.4	0.000
Vizianagaram	1.129*	0.079	1.7	0.085
Vishakhapatnam	1.023	0.074	0.3	0.754
East Godavari	1.082	0.086	1.0	0.323
West Godavari	1.050	0.077	0.7	0.503
Krishna	1.029	0.090	0.3	0.747
Guntur	1.152**	0.080	2.0	0.042
Prakasam	1.342***	0.089	4.4	0.000
Nellore	1.253**	0.091	3.1	0.002
Chittor	1.079	0.074	1.1	0.269
Cuddapah	1.491***	0.119	5.0	0.000
Anantapur	1.592***	0.112	6.6	0.000
Kurnool	1.598***	0.126	5.9	0.000
Mahaboobnagar	1.335***	0.098	4.0	0.000
Rangareddy	1.216**	0.101	2.4	0.019
Medak	1.173**	0.088	2.1	0.035
Nizamabad	0.962	0.080	-0.5	0.640
Adilabad	1.203**	0.084	2.6	0.008
Karimnagar	0.867*	0.075	-1.7	0.099
Warangal	1.027	0.078	0.4	0.724
Khamman	1.107	0.084	1.3	0.180
Nalgonda	0.901	0.081	-1.2	0.245
Observations	56000			
Log likelihood	-108048.12			
P-value (chi2)	0.000			

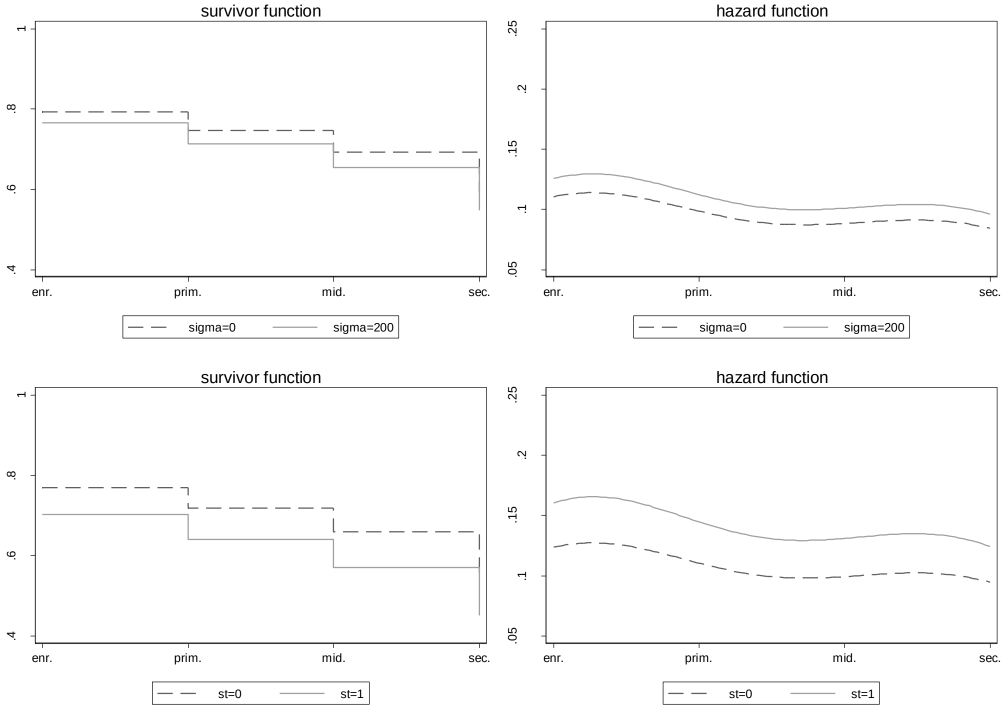
Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

The effect of rainfall variability has the expected effect on schooling decisions. More variable weather conditions reduce investments in education by rural households. The size of the hazard ratio is small, but it should be kept in mind that it represents the ratio of probabilities for a unit change in the standard deviation of district rainfall, which averages 198 for the whole sample.

One advantage of estimating the impact of uncertainty on schooling decisions by including a variable measuring uncertainty is that once the estimation has been performed, it is possible to compare the predicted values at the desired values of the explanatory variables, including the case where the uncertainty is zero. The charts of Figure 4.6 show the survivor and hazard functions obtained after the estimation of the Cox regression of Table 4.11. The survivor function represents the probability that a child completes a given school level, while the hazard function represents the risk that a child drops out at a given level conditional on having completed the previous level. There are three education levels on the chart: primary, middle and secondary. The first school level in the charts corresponds to the below primary level, which can be considered equal to the ever enrolled case.

The survivor and hazard functions of the two top charts of Figure 4.6 were calculated at the mean values of the explanatory variables, while the dotted lines were calculated after setting rainfall variability (sigma) to zero. The elimination of rainfall variability would produce an increase in completion rates by two to five percentage points at each school level. This effect is small but not negligible and does not look insignificant compared to what would be obtained by eliminating all factors that prevent children of scheduled tribes from attending school. The latter is shown in the two bottom charts which compare the functions of scheduled tribes to those of all other households.

Figure 4.6 Survivor and hazard functions with and without rainfall variability



Source: calculated from NSSO data.

4.7.3 Education expenditure model: variables used

The education expenditure model assesses the effect of rainfall variability on household education expenditure for groups of households whose incomes are differently affected by weather variability. The dependent variable is the share of education expenditure over total household expenditure. Most explanatory variables used in the school attainment model are also used in the education expenditure model. These variables affect education expenditure in the same way as they affect school progression and need not to be discussed again.

There are however some important differences in the specification of these two models. First, education expenditure is reported in the surveys at the household level, and cannot be linked to any particular child. The analysis can therefore only be household specific, rather than child specific, and the demographic variables used in this model are different from those used in the school attainment model. Second, data of the expenditure model measure household characteristics at the time of the interview and there are no time-varying covariates. Consequently, current household expenditure can now be included in the model. Third, the education expenditure model is run for subgroups of the whole sample of households, and selection terms are included in the regressions in order to reduce selection bias. These three main differences will now be discussed in more detail. Explanatory variables used in the expenditure model that were not included in the school attainment model are listed in Table 4.12.

Following Deaton (1997), the demographic structure of the household is included in the model using age-groups ratios. These ratios are obtained dividing the number of household members of each age-group by the total number of household members. It is a convenient way of including demographic characteristics in the model because it allows the identification of the effect of changing household composition while keeping household size constant. The age groups used in the model were defined based on the age ranges specific to different education levels (see Table 4.4). In the empirical specification of the model these groups are further disaggregated by sex in order to account for the unequal distribution of resources within the household. When performing the estimation, the category for female elderly is dropped in order to avoid collinearity.

Age group coefficients show how education expenditure varies as children of different age are added to the household. In principle, the coefficient values should be larger for older age groups because schooling costs are increasing by grade. However, not all children attend middle and secondary school and the coefficients for older age groups may reflect this. The disaggregation of age groups by gender is intended to assess the extent of differential treatment of boys and girls within the household. The econometric literature has often found different education expenditure coefficients for boys and girls in India, but standard tests have normally rejected the statistical significance of these effects. Kingdon (2005) found evidence for the hypothesis that this may be the result of a particular type of selection bias.

Table 4.12 Explanatory variables of the education expenditure model

Variable	*Description*
Non-school age children	Number of children aged zero to five
Primary school age children	Number of children aged six to ten
Middle school age children	Number of children aged 11 to 14
Secondary school age children	Number of children aged 15 to 18
Adults	Number of household's members aged 19 to 59
Elderly	Number of household's members aged 60 and above
Household size	Logarithm of household size
Per capita expenditure	Logarithm of total household per capita expenditure adjusted for regional and price variations within each survey and by the CPIAL across surveys
Primary schools	Number of primary schools per capita in the district
Middle schools	Number of middle schools per capita in the district
Secondary schools	Number of secondary schools per capita in the district
Survey rounds	Dummies for the four survey rounds
Selection terms	Selection term for the occupational choice calculated in Chapter 3 (see Section 3.6)

Note: demographic variables are included in the model as ratios over household size.

The demographic characteristics of the household also include the logarithm of household size. This variable shows the effect on expenditure of increasing the number of household members while keeping per capita expenditure constant. Effects of this sort can be attributed to economies or diseconomies of scale in household consumption. A share of household expenditure is for public goods, whose consumption increases less than proportionally with household size. Adding more household members at a given level of per capita income has the effect of releasing resources that can be spent on education.

The most important explanatory variable is the logarithm of per capita expenditure. The estimated coefficient of this variable describes the education Engel curve, which shows how expenditure on education varies as total expenditure increases. The functional form chosen for the specification of this model is the Working-Leser share form which allows for necessities (declining shares), and luxuries (increasing shares). The parameter estimates of the logarithm of per capita expenditure will be used to calculate the expenditure elasticity of education. The variables for skilled and unskilled wages capture the averages district level wages over the three years preceding the interview. Rainfall variability is the standard deviation of annual rainfall at the district level over the ten years preceding the survey interview. Three survey dummies were included in order to capture historical changes in education expenditure not otherwise explained by the other variables in the model.

The regressions are run for subgroups of the whole sample of households, namely farmer and agricultural labour households with and without access to irrigation. While it can be argued that the entire rural economy is to some extent dependent on agricultural output, incomes of farmer and agricultural labour households are more likely to depend on agricultural output and weather variability than those of other occupations (Walker and Ryan, 1990). Distinguishing between households with and without access to irrigation is important because irrigation reduces the output risk related to weather variability, and households with access to irrigation are more protected from drought and have access to water even when rainfall is scarce.

Once the original sample is divided in sub-samples based on households' characteristics, the sample is no longer a random sample of the rural population and parameter estimate may be biased. In order to reduce the selection bias, the model includes selection terms obtained using the econometric techniques described in Section 4.5. The characteristics of the selection bias and the empirical specifications of the selection models were discussed in Section 3.7.1 and this discussion will not be repeated here.

4.7.4 Education expenditure model: coefficients estimates

Results of the estimation of the education expenditure model are presented in Tables 4.14 and 4.15 for samples of farmer and agricultural labour households. The models were run using five different econometric specifications: standard OLS, standard tobit, tobit corrected for heteroscedasticity, semiparametric tobit, and a two-part model.[33] Most demographic coefficients corresponding to age groups of primary, middle, and secondary school age are positive and significant. Coefficient values are also slightly increasing with age group, reflecting the increasing cost of schooling by school level.

The demographic coefficients are larger for boys than for girls, pointing to the presence of gender discrimination in the allocation of education expenditure within the

[33] Regression results of the standard tobit are not shown.

household. Appropriate tests however failed to find a statistically significant difference. Demographic coefficients are larger in the farmers' models than in the labourers' one, especially the coefficients for female children. Farmers' male coefficients of secondary school age are much larger than agricultural labourers', while labourers' female coefficients are negative. The coefficient of household size is always positive and highly significant. The fact that the share of education expenditure increases with household size while keeping per capita household expenditure constant suggests that households are exploiting economies of scale in consumption.

Much of the variation in education expenditure across households is explained by the logarithm of household per capita expenditure. The estimated coefficient of this variable is better interpreted once transformed into elasticity. The elasticity is the change in education expenditure for a proportional change in total expenditure. Expenditure elasticities ε for the household group i are calculated using the following formula:

$$\varepsilon_i = \frac{b}{w_{ei}} + 1 \qquad (4.20)$$

where b is the estimated coefficient of the logarithm of per capita expenditure and w_{ei} is the average expenditure share on education for the sample of households considered. The elasticities obtained from the four specifications of the model and for the two household groups are shown in Table 4.13.

Table 4.13 Education elasticities of farmers and agricultural labourers

Econometric model	Farmers	Agricultural labourers
OLS	1.55	1.68
Tobit	1.99	1.78
Heteroscedastic tobit	1.85	1.38
Semiparametric tobit	1.44	1.14
2-part model	1.61	1.91

Source: calculated from NSSO data.

All estimated elasticities are larger than one, meaning that education is a luxury for these samples of Indian rural households and that the share of expenditure on education increases as total expenditure increases. The different values of the elasticities obtained by applying different econometric models tend to conform to econometric theory

(Deaton, 1997). Estimates obtained by the heteroscedasticity adjusted tobit and the semiparametric method yield lower values compared to those obtained by tobit estimation. Farmers' education elasticities are often, but not always, larger than those of agricultural labour households. These values are similar to values found by other studies. For example, Subramanian, quoted in Kingdon (2005), found an education elasticity of 2.14 in Andhra Pradesh, while Kingdon (2005) found values of 1.49 and 1.21 under two different specifications.

The district level skilled wage has no effect on education expenditure in any of the specifications and for any of the samples. There is some evidence of a negative effect of the unskilled wage (the opportunity cost of schooling) on education expenditure of farmers, but such effect is weak for agricultural labourers. Education expenditure increases with parents' education, reflecting both parents' attitudes towards schooling and income effects. Muslim household spend significantly less on education, while no difference in any direction is found for Christian households. Scheduled caste households spend less if they are farmers, but not if they are engaged in wage labour. There are very clear indications that tribal households are spending considerably less in education than other households, both as labourers and farmers. Land has no effect on education expenditure of labourers, while it appears to have a negative effect on expenditure of farmers. The use of traditional fuel for cooking and electricity have strong and opposite effects on education expenditure of both farmers and labourers.

Table 4.14 Education expenditure model: farmers

	OLS	Heterosced. tobit	Semi parametric	2-part
	Coeff.	Coeff	Coeff.	Coeff.
Non-school age children (M)	-0.015	-0.015	-0.022***	-1.695***
Non-school age children (F)	-0.019**	-0.023**	-0.028***	-1.942***
Primary school age children (M)	0.019*	0.060***	0.038***	0.116
Primary school age children (F)	0.011	0.036***	0.040***	0.322
Middle school age children (M)	0.018*	0.065***	0.052***	0.695
Middle school age children (F)	0.018*	0.048***	0.038***	1.038**
Secondary school age children (M)	0.020**	0.034***	0.025***	1.739***
Secondary school age children (F)	0.009	0.001	0.021***	1.279**
Adults (M)	-0.024**	-0.033***	-0.018***	-0.701
Adults (F)	0.006	0.000	0.010**	0.037
Elderly (M)	-0.004	0.003	0.005	0.311
Household size	0.001	0.012***	0.008***	0.710***
Per capita expenditure	0.008***	0.012***	0.005***	1.031***
Skilled wage	0.001	-0.001	0.001	0.003
Unskilled wage	-0.001*	-0.001	-0.001***	-0.010*
Rainfall variance	-0.001*	-0.001	-0.001***	-0.001**
Father’s education	0.001**	0.003***	0.002***	0.071***
Mother’s education	0.003***	0.002**	0.002***	0.120***
Muslim household	-0.003*	-0.007**	-0.013***	-0.422**
Christian household	0.005	0.007	-0.003	-0.120
Scheduled caste	-0.002	-0.003	-0.004**	-0.213*
Scheduled tribe	-0.003**	-0.008***	-0.009***	-0.404***
Land owned	-0.001*	0.001	-0.001***	-0.017
Traditional fuel	-0.015***	-0.010***	-0.008***	-0.516***
Electricity	0.001	0.004**	0.007**	0.142*
Primary schools	-0.003	0.002	0.003	0.362
Middle schools	-0.070	-0.120*	-0.030	-8.688**
High schools	0.092	0.257**	0.033	12.369**
50th survey round (1993-94)	0.003**	0.024***	0.022***	0.128
55th survey round (1999-00)	0.009**	0.024***	0.021***	0.820***
61st survey round (2004-05)	0.026***	0.044***	0.035***	1.588***
District dummies (22) (output omitted)				
Selection term	-0.010*	-0.011*	-0.011***	-0.346
Constant	-0.065**	-0.207***	-0.099*	-4.358***
Observations	5190	5190	2308	2701
R-square	0.21		0.08	0.44
Log-likelihood		3770.7		
P-value (F, chi^2)	0.000	0.000		0.000

Notes: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%. Convergence in the semiparametric model was achieved after 23 iterations, and the number of observations reported (2308) are those corresponding to the last iteration. (M) is for male and (F) is for female.

Table 4.15 Education expenditure model: agricultural labourers

	OLS	Heterosced. tobit	Semi parametric	2-part
	Coeff.	Coeff	Coeff.	Coeff.
Non-school age children (M)	-0.006	-0.018**	-0.015***	-1.140**
Non-school age children (F)	-0.007	-0.021**	-0.017***	-1.313**
Primary school age children (M)	0.011	0.043***	0.021***	0.761*
Primary school age children (F)	0.004	0.031***	0.013***	-0.097
Middle school age children (M)	0.017**	0.045***	0.022***	1.641***
Middle school age children (F)	0.006	0.022**	0.017***	1.132**
Secondary school age children (M)	0.016**	0.015*	-0.002	2.345***
Secondary school age children (F)	-0.005	-0.030***	-0.011**	1.243**
Adults (M)	-0.009	-0.030***	-0.015***	0.207
Adults (F)	-0.005	-0.013*	-0.011**	0.179
Elderly (M)	0.003	0.003	-0.010*	0.438
Household size	0.003**	0.016***	0.004***	0.719***
Per capita expenditure	0.005***	0.003**	0.001	0.765***
Skilled wage	-0.001	-0.001	0.001	0.003
Unskilled wage	-0.001	-0.001	-0.001	-0.009*
Rainfall variance	0.001	-0.001	-0.001*	-0.001*
Father’s education	0.003***	0.006***	0.002***	0.096**
Mother’s education	-0.001	-0.001	-0.001	0.033
Muslim household	0.001	0.001	-0.001	0.100
Christian household	-0.002	-0.003	-0.004**	-0.118
Scheduled caste	-0.001	-0.001	0.001	0.002
Scheduled tribe	-0.001*	-0.007***	-0.004***	-0.296**
Land owned	0.001	0.001	0.001	-0.047
Traditional fuel	-0.011***	-0.020***	-0.004**	-0.133
Electricity	0.003***	0.007***	0.005***	0.210**
Primary schools	-0.002	-0.006	-0.003	0.416
Middle schools	-0.034	0.001	-0.017	-6.321**
High schools	0.071	0.056	0.042	4.960
50th survey round (1993-94)	0.001	0.020***	0.014***	-0.115
55th survey round (1999-00)	0.001	0.019***	0.015***	0.581***
61st survey round (2004-05)	0.010***	0.034***	0.021***	1.297***
District dummies (22) (output omitted)				
Selection term	-0.018**	-0.028**	-0.004	-0.002
Constant	-0.036**	-0.043*	-0.019*	-1.728
Observations	5776	5776	1903	2262
R-square	0.14		0.08	0.37
Log-likelihood		3556.5		
P-value (F, chi^2)	0.000	0.000	0.000	0.000

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%. Convergence in the semiparametric model was achieved after 29 iterations, and the number observations reported (1903) are those corresponding to the last iteration. (M) is for male and (F) is for female.

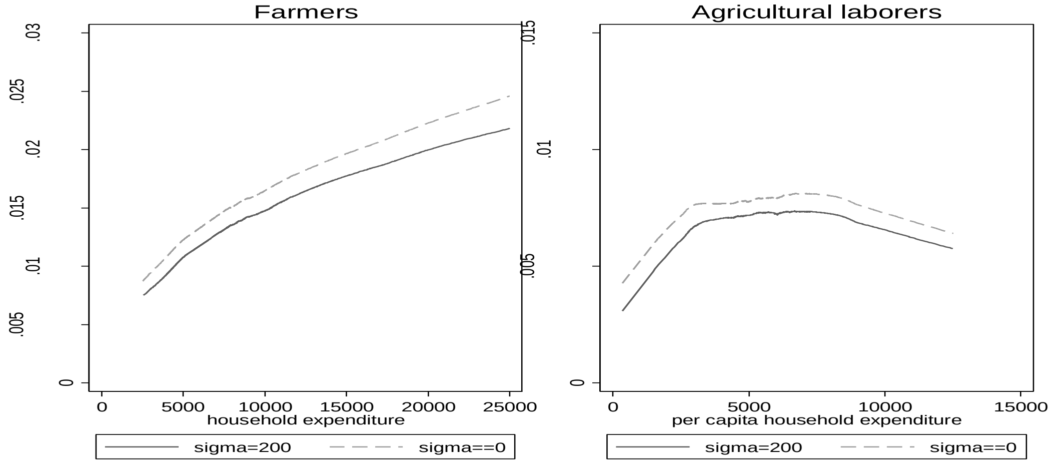
The number of primary and middle schools has a negative effect on household expenditure, though only the effect of the middle school variable is statistically significant. This result is reasonable if households' expenditure substitutes for the poor quality or the availability of schools in the district. Conversely, the number of secondary schools in the district has a positive effect on expenditure and the effect is often statistically significant. The survey dummy variables show that education expenditure has increased over time, particularly between the last two surveys.[34] This effect is statistically significant for both household groups and in most specifications of the model. The selection terms are significant in the majority of cases and their inclusion reduced the selection bias introduced by using sub-samples of farmers and labourers.

Rainfall variability has the expected negative effect on household education expenditure. The significance of the effect is supported by three of the four specifications of the model in the case of farmers, and by two specifications in the case of agricultural labourers. The size of the effect is larger for the sample of farmers, though this is not apparent from the output of tables 4.14 and 4.15, because coefficient estimates were rounded to the third decimal place. This does not mean that labourers face less uncertainty than farmers, rather it suggests that labour markets are not particularly affected by agricultural output variability or that agricultural labourers are relatively insulated from output fluctuations by other means.

The charts in Figure 4.7 show the shares of education expenditure predicted by the model plotted against household per capita expenditure. The predicted shares were obtained using the coefficient estimates of the semiparametric tobit, because it is the only model that found a highly statistically significant effect of rainfall variability on education expenditure for both farmers and agricultural labourers. The plots were produced nonparametrically by applying LOWESS only to the positive values of the predicted shares.

[34] The increase in education expenditure between the last two surveys may be biased by the overestimation of total household expenditure in the survey of 1999-00 due to changes in the household questionnaire discussed in Section 2.4.3.

Figure 4.7 Predicted education expenditure shares with and without rainfall variability



Source: calculated from NSSO data.

The purpose of the charts is to simulate the effect of eliminating the uncertainty about agricultural output determined by rainfall variability. The dotted lines plot the predicted shares after setting rainfall variability (sigma) to zero, while the solid line is calculated at the average values of all explanatory variables. Note that the two charts are plotted on different scales of shares and per capita expenditure because in the sample of farmers these variables are on average twice the size of those of agricultural labour households. The charts show a small but clear effect of weather variability on education expenditure. The effect is slightly larger in the case of farmers, and increasing with per capita expenditure.

Tables 4.16 and 4.17 show the results of the expenditure model run separately for samples of irrigated and non-irrigated farmers and for samples of agricultural labour households living in irrigated and non-irrigated villages. Only a limited number of coefficient estimates are reported, because most coefficients of the full specifications do not add any new insight. The demographic variables have larger effects among irrigated households, both in the case of farmers and labourers. Expenditure increases with the age group of the child and more so for irrigated households. The skilled wage is never statistically significant, while the unskilled wage is occasionally significant and has the expected negative sign in some of the specifications for non-irrigated farm households.

Rainfall variance is rarely statistically significant, and when it is, its negative sign points to a negative effect on education expenditure of irrigated farmers and labourers. The estimated effect is of the opposite sign of what expected. In other words, farmers and agricultural labourers with access to irrigation seem to spend more on the education of their children and up to higher grades, but they are also more affected by rainfall variability than their non-irrigated counterparts. While the income effect is positive, the effect of the variance is negative.

One possible interpretation of this result is that the strategy adopted to overcome the selectivity problem was not the right one. But assuming that the selection bias was removed from the estimation, an alternative explanation is needed. One possibility is that farmers and labourers with access to irrigation derive a larger fraction of their income from agriculture. As a result, their incomes vary more widely with weather fluctuations though weather fluctuations have in general a lower effect on their agricultural output. Therefore, households with access to irrigation invest more in education than non-irrigated ones due to their higher incomes and not because they are less affected by income uncertainty.

Table 4.16 Education expenditure model: irrigated and non-irrigated farmers

	OLS		Heteroscedastic tobit		Semiparametric tobit		2-part	
	Irrigated	Non-irrigated	Irrigated	Non-irrigated	Irrigated	Non-irrigated	Irrigated	Non-irrigated
Non-school age children (M)	-0.007	-0.024	-0.013	-0.008	-0.014**	-0.041***	-1.192*	-2.098**
Non-school age children (F)	-0.011	-0.030**	-0.017	-0.030*	-0.027***	-0.028**	-1.568**	-2.678**
Primary school age children (M)	0.033**	-0.005	0.059***	0.058**	0.048***	0.018**	0.793	-1.152
Primary school age children (F)	0.018	-0.004	0.040***	0.032*	0.047***	0.020**	1.051*	-0.843
Middle school age children (M)	0.034**	-0.013	0.071***	0.049**	0.057***	0.027**	1.168**	-0.302
Middle school age children (F)	0.034**	-0.012	0.051***	0.041**	0.043***	0.021**	1.703***	-0.304
Secondary school age children (M)	0.023*	0.019	0.035**	0.026	0.031***	0.013	2.151***	0.923
Secondary school age children (F)	0.016	-0.005	0.009	-0.014	0.028***	-0.018*	1.911**	0.283
Skilled wage	0.001	0.001	0.001	0.001	0.001	-0.001	0.001	0.006
Unskilled wage (male)	-0.001	-0.001**	-0.001*	-0.001	-0.001	0.001	-0.005	-0.020**
Rainfall variance	0.001	0.001	0.001**	0.001	-0.001***	0.001	-0.001	-0.001
Selection term	-0.162***	-0.022	-0.034**	-0.015	-0.025	-0.069**	-6.485**	0.392
Observations	3679	1473	3679	1473	1605	677	1953	738
R-square	0.23	0.25			0.08	0.11	0.44	0.47
Log likelihood			2657.3	1094.2	34	17		
P-value (chi2)	0.000	0.000						

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

Table 4.17 Education expenditure model: agricultural labourers of irrigated and non-irrigated villages

	OLS		Heteroscedastic tobit		Semiparametric tobit		2-part	
	Irrigated	Non-irrigated	Irrigated	Non-irrigated	Irrigated	Non-irrigated	Irrigated	Non-irrigated
Non-school age children (M)	-0.004	-0.006	-0.008	-0.020**	-0.013**	-0.020***	-0.819	-1.138**
Non-school age children (F)	-0.009	-0.004	-0.021	-0.020**	-0.013**	-0.025***	-1.495**	-1.046*
Primary school age children (M)	0.014	0.011**	0.064***	0.033***	0.018***	0.012***	0.923	0.684
Primary school age children (F)	0.003	0.007	0.048***	0.024**	0.014**	0.002	-0.030	-0.034
Middle school age children (M)	0.006	0.030**	0.049***	0.044***	0.020***	0.019***	1.034	2.409***
Middle school age children (F)	0.006	0.008	0.034**	0.016	0.017***	0.012***	1.142*	1.316**
Secondary school age children (M)	0.023*	0.011*	0.039**	0.002	0.001	-0.009**	2.676***	2.118**
Secondary school age children (F)	-0.008	-0.001	-0.030**	-0.029**	-0.011*	-0.019***	1.873**	0.875
Skilled wage	0.001	-0.001	-0.001	-0.001	-0.001	0.001	0.004	0.002
Unskilled wage (male)	0.001	-0.001	-0.001	0.001	-0.001	0.001	-0.006	-0.008
Rainfall variance	0.001	0.001	-0.001	0.001	-0.001*	0.001	-0.001*	-0.001
Selection term	-0.021*	-0.012	-0.025	-0.025*	-0.019**	0.007	-0.955	0.906
Observations	2631	3145	2631	3145	1063	1014	1075	1187
R-square	0.13	0.18			0.09	0.11	0.37	0.41
Log likelihood			1613.5	1907.7	24	18		
P-value (chi2)								

Note: *statistical significance at 10%, **statistical significance at 5%, ***statistical significance at 1%.

4.8 Conclusions

This chapter assessed the impact of income uncertainty on schooling decisions made by rural households. A simple model of educational choices was developed for rural households that are poor, liquidity constrained and face substantial income risk. Schooling decisions are modelled as solutions of a utility maximisation problem, which includes costs and benefits of education over the life cycle. The hypothesis that uncertainty of agricultural income forces households to save resources that could be otherwise invested in education is tested using an educational attainment model, set up as a duration model with time covariates. This empirical model includes historical rainfall variability among the explanatory variables as a predictor of perceived income uncertainty. Models of household education expenditure are also estimated in order to assess the effect of uncertainty on different households groups. The results of the empirical analysis are as follows:

- Andhra Pradesh has made considerable progress over the last 50 years in increasing completion rates in primary education, and in reducing disparities in educational attainments between boys and girls, social groups, and regions. Completion rates increased from less than 20% in the 1950s to more than 80% in the 1990s in primary school, from less than 15% to 60% in middle school, and from less than 5% to 50% in lower secondary school. Girls have almost closed the gap with boys in primary and lower secondary schools in the 1980s and 1990s, though differences still remain in middle school. By the early 1990s, children of scheduled caste had almost entirely closed the gap with children of higher castes in primary, middle and lower secondary school. Children of scheduled tribes however are lagging behind and the gap in completion rates has increased over time with the exception of primary school completion rates, where the gap was slightly reduced. No substantial regional differences in completion rates were found across the three macro regions, though there are signs that the Rayalseema region was lagging behind other regions in the late 1980s and early 1990s.

- Socioeconomic parents' background, returns to schooling, and school availability are important determinants of children progress through school.

Income uncertainty, represented by historical rainfall variability, significantly reduces the probability that a child will attain higher education levels. When compared to other determinants of progress through school the effect is found to be modest but not negligible.

- The socioeconomic status of the household and the opportunity cost of schooling are important determinants of expenditure on education by farmer households. Households spend less in the education of girls compared to boys, but the differences were not found to be statistically significant. Expected income variability was found to negatively affect household expenditure on education in three of the four specifications of the model.
- Expenditure on education of agricultural labourers is largely explained by the socioeconomic status of the household. Agricultural labour households spend less in the education of girls compared to boys, but the difference is not statistically significant. Expected income variability was found to negatively affect household expenditure on education in two of the four specifications of the model.
- Determinants of education expenditure, including income variability, do not affect farmers with and without irrigation differently. Agricultural labour households with access to irrigation are negatively affected by income uncertainty in two of the four model specifications.

The empirical analysis found that income uncertainty negatively affects schooling decisions of rural households and expenditure on education of farmers and agricultural labourers. Because output variability was approximated by rainfall variability rather than being observed directly, the size of this risk is not easily quantifiable. In order to measure weather-related risk we would need to know rainfall probabilities at a much more disaggregated level. The lack of extensive historical rainfall data is a main constraint to the analysis conducted in this chapter. In addition, it is not certain that farmers assess future income risk based on the historical record of rainfall patterns (Luseno et al., 2003).

Rainfall risk however can only underestimate the effect of uncertainty on households' decisions, because the idiosyncratic risk faced by Indian households is much higher than the one represented by rainfall variability. Morduch (2004) finds that idiosyncratic risk constitutes between 75% and 95% of total income variability in the ICRISAT villages of south India. Following the influential study by Townsend (1994), most development economists have tended to dismiss the importance of idiosyncratic risk. Townsend (1994), using the ICRISAT data, showed that Indian households were able to insure consumption against idiosyncratic shocks through community exchange of gifts and financial help. Similar results were obtained in other contexts, suggesting that policy makers should focus their attention on region-wide shocks rather than on idiosyncratic ones (Besley, 1995, Fafchamps, 1992, Udry, 1994). However, in a follow-up study Lim and Townsend (1998) found that consumption smoothing in the ICRISAT villages happened through self-insurance, by building up and drawing on grain reserves, rather than through informal community based sharing. This evidence is corroborated by several empirical studies pointing to the inability of rural households to fully insure against idiosyncratic risk (Carter, 1997, Dercon and Krishnan, 2000b, Gertler and Gruber, 2002). The impact of overall uncertainty, resulting from both its idiosyncratic and covariate components, on households' investment in human capital is likely to be much higher than the one detected by this study. The advantage of using rainfall variability as a predictor of uncertainty is that it offers a source of risk variation that is exogenous to household characteristics.

The obvious remedy against the negative effects of uncertainty is the provision of insurance. Some economists believe that the provision of insurance to the poor will be a major milestone in the fight against poverty (Dercon, 2007). The reasons for the absence of insurance in poor rural areas relate to the notions of moral hazard and adverse selection (Ray, 1998). First, once insured, farmers may be less likely to make production decisions that increase their chances of success. In this way, insurance may increase the probability of losses. Second, those more willing to purchase the insurance are also those more inclined to make risky decisions, but insurers cannot know in advance who is most risky. The application of a flat insurance rate has the effect of pushing away safe buyers, while a reduction in the premium would reduce the insurer's profits.

In addition to these information-related problems, there are some practical issues in offering insurance to poor farmers (Dercon, 2004, Morduch, 2006). Contracts with farmers are for small amounts. As a result, the collection of payments and the verification of damage can be difficult to carry out in rural areas and there are no economies of scale to exploit. In other words, the transaction costs of an insurance scheme in rural areas are very high. For these reasons, market-based insurance in rural areas of poor countries is rare, and even when it is supplied or subsidised by governments or external assistance it is difficult to implement (Hazell, 1992, Hazell et al., 1986).

Andhra Pradesh is currently running two innovative insurance programmes. The first is the National Agricultural Insurance Scheme (NAIS) that since 1999 provides seasonal insurance at the area level (*mandal*) for a number of crops including: rice, groundnut, castor, cotton, chillies, and sugarcane. In the agricultural year 1999-00, nearly 16% of farmers were insured in the Kharif season (Vyas and Singh, 2006). Under this scheme, all farmers growing a given crop in a given area, and accessing bank loans, make compulsory insurance payments. Crop failure is evaluated at the *mandal* level in terms of shortfall of yields or rainfall respect to an historical average. Payment is made in terms of part or of the entirety of the loan given. The second insurance programme run by the state is the rainfall insurance scheme implemented by the ICRISAT and the World Bank since 2004 in the Mahaboobnagar and Anantapur districts (Gine et al., 2008). Similarly to the NAIS, this scheme compensates farmers for losses deriving from rains below a minimum historical threshold at the area level. The advantage of these schemes is that they avoid the information-related problems outlined above. By insuring every farmer in a given area and by assessing an average damage, they prevent the occurrence of both adverse selection and moral hazard problems. In addition, the regional coverage of the schemes simplifies the administration task and reduces transaction costs.

The evaluation of these and other insurance programmes is in its infancy, and the evidence available in support of a microinsurance revolution is thin, though the perception of its need is very high (Morduch, 2006). Evaluations of microinsurance programmes have so far focused on issues of take-up and welfare outcomes in terms of expenditure levels and income stabilisation (Cole et al., 2009, Gine et al., 2008, Gine and Yang, 2007). The results presented in this chapter suggest that a proper assessment of insurance schemes should include an evaluation of their impact on long term household decisions and on investments in human capital in particular.

5 Conclusions

This concluding chapter offers some explanations of the persistence of poverty in rural Andhra Pradesh based on the findings of the three previous chapters. The first section illustrates the poverty trap model, while the following sections discuss the findings of the previous chapters in light of the poverty trap model. Each section also discusses a number of possible directions for future research.

5.1 The poverty trap model

Standard neoclassical economic theory is based on models in which countries and individuals face opportunities that reward efforts and savings. Bowles et al. (2006a) call these models the achievement models of income determination. According to these models, over the long term all countries converge to the income level of rich countries (Solow, 1956), while over the generations incomes of poor households converge to those of richer households (Loury, 1981). The empirical evidence available does not conclusively support these models. There is considerable evidence of the lack of convergence in the rates of economic growth across countries (see for example Pritchett, 1997, Ray, 1998), and studies of intergenerational transmission of wealth point to the presence of substantial social immobility both in developed and developing countries (see for example Bourguignon et al., 2007, Mazumder, 2005).

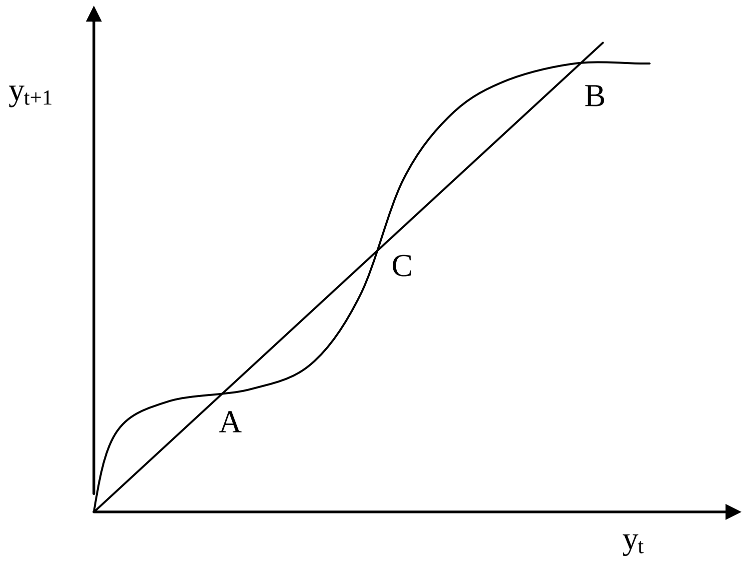
In contrast to the neoclassical model, poverty traps models explain why poverty and inequality persist among countries, generations and individuals. Generally speaking, a poverty trap is any self-reinforcing mechanism that causes poverty to persist (Azariadis and Stachurski, 2005). The idea of a poverty trap is one of the oldest in economics. The population model of Malthus, in which economic growth is constrained by population growth, and the Ricardian trap of diminishing rates of profits that bring the economy to a ‘stationary state’, are early examples of poverty trap models (Hayami and Godo, 2005). Several poverty traps models were formulated in the early stages of development economics. Rosenstain-Rodan (1943) argued that the industrialisation of south-eastern Europe was constrained by the small size of the markets. Singer (1949) described the presence of vicious circles within vicious circles of poverty, whereby low productivity entailed inability to save, from which modest investment and poor productivity would follow. Nurske (1953) maintained that underdeveloped countries were trapped in vicious circles of poverty and capital accumulation, whereby inability to save would lead to low productivity levels and viceversa. Leibstein (1957) illustrated how some economies were trapped in a low level equilibrium, resulting from the predominance of income-depressing forces over income-raising forces at very low levels of income. Myrdal (1957) discussed how processes of cumulative causation determined backwash effects in less developed areas within and across countries. In these processes, capital and labour moved to growing areas, while non-growing area spiralled down into depression.

The fundamental intuition behind early poverty trap models is the importance of increasing returns to scale in the process of economic growth. The idea that investment decisions depend on the size of the market and that the size of the market in turns depend on decisions to invest is already present in the *Wealth of Nations* of Adam Smith (1970). Early development economist further elaborated this idea describing various processes of positive feed-back, vicious circles and cumulative causation as fundamental explanations of the poverty of nations. The theoretical treatment of the topic by early development economists was however informal and making little use of mathematical and graphical notation, which largely contributed to its neglect by mainstream economic literature (Krugman, 1997). The interest in this type of models was recently revived by the formal work of Murphy et al. (1989a), Matsuyama (1991), and Krugman (1991), among others.

Though poverty traps may emerge from endogenous dysfunctional institutions and social interactions (Bowles et al., 2006a), we are concerned here with threshold poverty traps that originate from market imperfections (Azariadis, 2006). Two main market imperfections may generate threshold poverty traps: increasing returns to scale and liquidity constraints. The exposition of the threshold poverty trap model begins with a description of the positive feed-back mechanism that generates multiple equilibria.

The canonical threshold poverty trap model (Azariadis and Stachurski, 2005, Carter and Barrett, 2006) is best illustrated with the use of a phase diagram (Chiang, 1974). Poverty traps are generated by feed-back mechanisms like those produced by increasing returns to scale. If output depends on output scale, income increases over time as income increases, and income in the next period (y_{t+1}) is a function of income in the present period (y_t). This iterative process is represented algebraically by a difference equation of the form $y_{t+1} = f(y_t)$. While linear difference equations can be solved explicitly, nonlinear difference equations cannot. When nonlinearity occurs difference equations can be analysed qualitatively with the use of a phase diagram like the one depicted in Figure 5.1. The curve corresponding to $f(y_t)$ is a phase line and describes the dynamic path of income over time.

Figure 5.1 Phase diagram of the canonical poverty trap model



The 45 degrees line in Figure 5.1 represents all points in which income in the current period and income in the following period are exactly equal. If we start from an initial value of income on the horizontal axis, we can trace the subsequent values of income in the following way. First, find the value of income in the following period on the phase line. Second, transplot this value in the next period on the horizontal line. This can be accomplished by hitting the 45 degrees line from the phase line and then reporting the value on the horizontal axis. Proceeding in this way by iteration, we move along the dynamic path of income over time.

Figure 5.1 was drawn in such a way to show three intertemporal equilibrium points, of which two are stable (*A* and *B*) and one is unstable (*C*). Starting from values of income in the immediate right and left of point *A*, income converges to the equilibrium point *A*. Similarly, starting from values of income to the near left and right of point *B*, income converges to point *B*. By contrast, the equilibrium point *C* is unstable, and any movement from this point brings income either to point *A* or *B*.

The chart in Figure 5.1 depicts two stable equilibrium points, one characterised by a high income level and one characterised by a low income level. Point *A* is a poverty trap: small income increases beyond *A* lead income to converge again to *A*. The intersection of the phase line with the 45 degrees line at point *C* is the poverty trap threshold. Movements of income beyond this point push a country or household permanently from one equilibrium state to the other.

One implication of this model is that temporary events may have permanent consequences (Azariadis, 2006). Random events like wars, natural disasters or wrong policies may compromise long term economic development by pushing a country below the critical threshold. In the same way a massive foreign aid programme may have long term consequences if it is able to push the country beyond point *C*. Similarly, households trapped in the low equilibrium of point *A* need a substantial injection of income in order to go beyond the threshold *C*. Once the threshold is overcome, the household continues along the positive dynamic path that brings to the high equilibrium point *B*. By contrast, a catastrophic health shock may push a household temporally from point *B* to below point *C*. From point *C* the household is then permanently attracted to the poverty trap of point *A*.

Threshold poverty traps arise for several reasons both at the macro and microeconomic levels. An obvious one is the presence of increasing returns to scale related to the characteristics of the production process (Azariadis and Stachurski, 2005) or to the process of urbanisation (Fujita et al., 2001). At the micro level poverty traps arise, for example, because individuals have no access to the best technologies available (Carter and Barrett, 2006), or because do not have the level of income or assets that allow them exploiting economies of scale.

One necessary condition for poverty trap to occur is the presence of liquidity constraints. If households or countries were able to borrow freely, nothing would prevent them from reaching the threshold and adopting the technologies leading to the high level equilibrium path. One important assumption of poverty traps model is therefore that household are unable to borrow. In the presence of liquidity constraints the only strategy countries and households can adopt in order to reach the high return technology is autarchic saving. Household and countries may sacrifice current consumption in order to accumulate assets. However, very poor households may not be able to reduce consumption any further in order to invest in more profitable activities. The decision on whether to adopt such a autarchic saving strategy may depend on the distance from the point where increasing returns can be exploited (Carter and Barrett, 2006). If the distance is very large, then the adoption of an autarchic saving strategy may be undesirable. This issue was explored at length by Zimmerman and Carter (2003), who identified a critical asset threshold below which an autarchic accumulation strategy is not rational nor feasible.

Poverty trap models provide little scope in terms of quantitative restrictions and implications that can be tested using empirical data (Azariadis and Stachurski, 2005). Examples of empirical tests of poverty trap models at the country level include Bloom et al. (2003) and Graham and Temple (2006), while at the microeconomic level, Ravallion and Jalan (2002) find evidence of poverty traps in rural China, though McKenzie and Woodruff (2006) cannot find evidence of poverty traps in rural Mexico. One reason for the difficulty of testing the validity of poverty traps models is the lack of panel datasets that map out the dynamics of income and other key variables over time. In addition, there are a number of identification and estimation problems that to date have found no solution (Carter and Barrett, 2006).

The three chapters of this dissertation do not investigate the presence of dynamic paths leading to permanent poverty, and the simple observation of poverty over time cannot be taken as evidence of poverty traps. However, all three chapters describe mechanisms that maintain households into poverty. These mechanisms are necessary ingredients of poverty trap models. In the following sections, the determinants of persistent poverty identified in each chapter are discussed in more detail within the theoretical poverty trap model described above.

5.2 Chapter 2: The under-consumption trap

Chapter 2 focused on the fundamental determinants of structural change in the eonomy of Andhra Pradesh: technological progress in agriculture and inelastic food demand.

The effects of food demand on the development of the agricultural sector were analysed in comparative statics for a change in technological progress. The derivative 2.21 obtained from the solution of the model described in Section 2.3 shows that the share of employment in agriculture decreases with technological change. Since the model assumes increasing returns to scale in manufacturing, linked to population size in urban areas, migration from agriculture induced by technical change gives rise to a dynamic growth path for the entire economy.

The analysis of the derivative 2.21 also shows that the negative effect of technological change on agricultural employment does not take place under certain conditions. First, if agriculture does not produce enough food to satisfy minimum requirements, any increase in income translates in an increase in food demand that turns to positive the value of the derivative 2.21. When the rural population lives below a minimum subsistence level, an increase in labour productivity results in an increase in labour employment and food consumption. Second, the larger the propensity to consume food out of current income, the smaller is the value of the derivative 2.21, and hence the smaller is the effect of technological change on agricultural employment. Even if living standards are above minimum food requirements, the effect of technical progress on labour demand vanishes if food demand is very high. Third, the analysis conducted in Section 2.6.3 shows that this effect is reinforced by the shape of the income distribution. Higher income equality is associated with higher food demand. The income food elasticity is large when most population lives in rural areas in similar conditions of extreme poverty.

The conclusions to Chapter 2 discussed the validity of the results of this model under the more realistic assumption that the economy is open to international trade rather than closed. It was acknowledged that in the open economy case the conclusions of the model would be different. However, the role of exports in fostering industrialisation is often overplayed (Murphy et al., 1989b). For example, Chenery et al. (1986) conclude that domestic demand is the main determinant of industrial growth. Using a sample of rapidly growing economies over the period from the early 1950s to the early 1970s, they find that domestic demand accounts for more than 70% of the increase in domestic industrial output. The investigation of the effects of technological progress in the open economy case is a potential area of further research that would require data on the trade balance of Andhra Pradesh versus the rest of India and the rest of the world.

Chapter 2 offers an interpretation of the inability of Andhra Pradesh to industrialise which is in line with theoretical expositions of poverty traps found in the macroeconomic literature (Ros, 2000). The extreme poverty of rural population limits the market for industrial products. Since increasing returns to scale are at the origin of the industrialisation process, demand factors prevent the economy to take off until the basic nutritional requirements of the population are satisfied. Similar interpretations of underdevelopment based on limited market size for industrial product can be found in Murpy et al. (1989b) and in Eswaran and Kotwal (1993).

There are two main policy implications to be drawn from this analysis. First, redistributive policies and policies that increase the living standards of the poorest sectors of the population have a positive impact on long term prospects of economic growth. It is suggested that chronic extreme poverty prevents agrarian economies like Andhra Pradesh from taking off. Second, state interventions in agricultural markets may slow the pace of the agricultural transformation. Andhra Pradesh subsidises agricultural producers though the provision of cheap fertiliser, electricity free of charge, and by purchasing unsold production at minimum support price (Gulati and Narayan, 2003, Jha et al., 2007). These subsidies result in an increase in the production of food and labour demand in agriculture which retard the process of industrialisation.

One shortcoming of the model developed in Chapter 2 is that the results are entirely discussed in comparative statics while a dynamic analysis would be preferable. Because different sectors of the economy display different returns, changes in demand composition and in income distribution brought about by technical change in agriculture affect overall economic growth. In turn, changes in income distribution generate changes in demand composition and therefore in factor demands, which may change income distribution again. The interactions between demand composition, factor demand, income distribution and growth need to be analysed dynamically (Ray, 1998). One way to perform this analysis is by using a Computable Genereal Equilibrium model, whose formulation is well beyond the scope of this chapter. This exercise is left for future research.

5.3 Chapter 3: Irrigation and the liquidity trap

Chapter 3 explored the effect of irrigation on consumption and saving decisions of poor farmers and agricultural labourers. The hypothesis was formulated that irrigation stabilises households' consumption by reducing uncertainty of future output. The empirical analysis found that consumption of rural households with access to irrigation is more stable over the seasonal cycle, and that consumption of households without access to irrigation tracks current income more closely. The empirical analysis also found that farmer households without irrigation tend to save more for precautionary reasons. Irrigation, by reducing risk, affects consumption smoothing and precautionary saving.

The inability to smooth consumption and the presence of precautionary savings alone are not evidence of poverty traps. They are however important contributing factors. Precautionary savings are an inefficient form of storing wealth. By stocking savings in cash or jewellery the poor forego investments in more profitable activities (Deaton, 1990). Moreover, production risk reduces the ability to borrow even if credit is available. This happens because if credit contracts are enforced and if the risk of being unable to repay a loan is high, poor household will not be able borrow (Carroll, 2001). Risk has therefore the effect of reinforcing poverty traps. The poor are cut out of the credit market below a given level of wealth because are unable to provide valid collateral (Banerjee and Newman, 1993). In this context, production risk has the effect of further reducing the ability to repay and of increasing the wealth required to access credit.

Credit constraints, risk and inability to borrow would not matter if households were able to save out of current consumption in order to accumulate wealth. One additional assumption of the model of Chapter 3 is therefore that the poor cannot reduce consumption below a minimum level. The idea that poor households are impatient, meaning that they are pressed to spend their income over the short term, dates back to Fisher (1930) and it is a common ingredient of dynamic consumption models (see for example Deaton, 1991, Glewwe, 1999).

Two main policy implications follow from this analysis. First, programmes providing production insurance or other form of income protection might have an impact on investment behaviours. A consensus is now emerging around the idea that the provision of insurance for the poor might represent a milestone in the fight against poverty (Dercon, 2004). The implementation and the study of insurance programmes in developing countries however are only in their infancy. Few evaluations of such programmes exist and those available do not go beyond the analysis of the conditions for insurance uptake by the intended beneficiaries (Cole et al., 2009, Gine et al., 2008, Gine and Yang, 2007). The analysis of the impact of these interventions on consumption and investment behaviour is an interesting area of future research.

The second policy implication is that the provision of credit to the poor is a promising area of intervention. This policy has already been adopted by the government of Andhra Pradesh on a large scale. Over the last 10 years the state has witnessed the explosion of the self help group movement. A number of programmes run in collaboration with the UNICEF, the World Bank and other donors since the late 1990s has provided millions of women with microcredit, capacity building and social empowerment (Galab and Rao, 2003). The data used in this study do not reflect the impact of these programmes because the most recent survey round of NSSO used dates back to 2004-05. However, recent impact evaluations of microcredit programmes in Andhra Pradesh, have found positive effects on consumption, nutritional status and asset accumulation even for the poorest households (Deininger and Liu, 2009).

Finally, it should be emphasised that the conclusions of this theoretical and empirical investigation rest on strong assumption regarding precautionary saving behaviour and time preferences. These assumptions are rarely tested using empirical data. Few studies have tried to assess the size of precautionary savings (Carroll and Kimball, 2007). These studies have found that in rich countries the amount of wealth held in the form of precautionary saving against future uncertainty is rather high (Carroll and Samwick, 1998). The extent of risk faced by rural household in developing countries would suggest that precautionary savings in poor areas should be even higher, but no similar studies have been attempted in developing countries. This represents an exciting area of future research.

The assumption that the poor are impatient makes appeal to common sense, but has been rarely tested. Barry and Packard (2000) found the rich to be more patient in a sample of Chilean pensioner, but Kirby et al. (2002) found no correlation between time preferences and wealth in a sample of Bolivian horticulturalists. The evidence in this area is mixed and more behavioural experiments are needed in developing countries to test fundamental assumptions of models used by development economists (Cardenas and Carpenter, 2008).

5.4 Chapter 4: the educational poverty trap

Chapter 4 investigated the effect of income uncertainty on child schooling, a topic largely neglected by the literature on the economics of education. The hypothesis was made that income uncertainty negatively affects investments in human capital of rural households. The empirical analysis found that output uncertainty increases the risk of school dropouts of rural children, and reduces household expenditure on education, particularly in farmer households. This outcome is a result of liquidity constraints. Poor household cannot access credit in order to invest in early childhood development and successive schooling because a market for this type of loans does not exist (Loury, 1981). In the absence of credit markets the alternative strategy a household can adopt is autarchic saving out of current income. Poor households however can hardly adopt such a strategy because their rates of time preference for immediate consumption are very high (Glewwe, 1999).

Risk makes credit constraints more stringent (Banerjee, 2004). Poor households are more risk averse than richer ones and production risk in rural areas is very high. As a result, poor household underinvest and do not adopt the educational and occupational careers that offer the highest rates of return. This occupational trap is a source of persistent inequality because the existing distribution of income is reproduced across the generations. It is also a source of inefficiency because does not allow people at the lower end of the income distribution to fully exploit their capabilities (Ray, 1998). The interplay of credit constraints and inability to invest in human capital generates economy wide inequality and inefficiency (Galor and Zeira, 1993, Ljungqvist, 1993).

Credit market imperfections and high production risk contribute to polarise the society between skilled workers exploiting high returns to education, and a mass of uneducated wage workers that are too poor to save up the funds required for schooling. This polarising effect is stronger in India than elsewhere because of the characteristics of the recent process of economic growth. Technological innovation and openness to international trade have resulted in an increase in demand for highly skilled occupations and their corresponding wages (Riboud et al., 2007). The review of the literature on returns to education conducted in Section 4.3.3 confirms this. Returns to schooling in India are diminishing from primary to middle school but increasing over higher education levels.

To summarise, uninsured risk contributes to the negative effect of credit market imperfections on investments in human capital. This in turn reduces the overall efficiency of the economy and generates wide and persisting income disparities. Two policies are particularly promising in addressing these issues: the provision of insurance for the poor and the provision of incentives for school attendance. The first type of policies is in its infancy and has been discussed in the previous section. Andhra Pradesh is currently running interesting experiments in the provision of weather insurance (Lilleord et al., 2005). The evaluation of these experiments is a new area of research. The analysis conducted in Chapter 4 suggests that the evaluation of these programmes should be performed over the long term in order to assess to what extent a reduction in risk affects investment decisions, including investments in human capital.

The provision of conditional cash transfers is another policy that might increase enrolment and reduce drop-out rates. Conditional cash transfer programmes, which provide money to poor families contingent on ensuring school attendance, have proved to be successful in a number of Latin American countries (Rawlings and Rubio, 2005). Most evidence on conditional cash transfer programmes however is on their impact on immediate welfare outcomes like consumption, health and school attendance. The evaluation of the effectiveness of these programmes in orienting educational and occupational decisions of the rural poor over the long term is an area of potential future research.

Finally, it should be emphasised that the models discussed in this section and the model of educational choice under production risk described in Section 4.3.1 , are based on the assumption that poor households are risk averse, and more risk averse than richer ones.

This hypothesis has found limited support in empirical tests (Binswanger, 1981, Pender, 1996). Cardenas and Carpenter (2008) in reviewing the experimental literature on risk aversion conclude that there is no support for the idea that poor people in developing countries are more risk averse than rich people in developed countries. This is an important area of research as it provides the micro foundations for our understanding of household decision making in developing countries.

Related to the issue of assessing risk attitudes is the problem of measuring risk. Most studies, including the present one, attempting to analyse the impact of risk and uncertainty on households' economic decisions rely on indicators of rainfall variability (see for example Gurgand, 2003, Jacoby and Skoufias, 1998, Jensen, 2000). Rainfall however is a very imperfect indicator of risk because rainfall data are collected over large areas with high heterogeneity in precipitations. Rainfall data are used as a matter of convenience because are easily available, but new and imaginative ways of measuring risk should be sought in order to shed light on the impact of risk on the welfare of the poor.

Appendix A: Derivation of demand functions

Derivation of homothetic (Cobb-Douglas) demand functions

Consumers maximise the utility function:

$$U = C_a^{\alpha} C_m^{1-\alpha} \quad \text{(A.1)}$$

with respect to the following budget constraint:

$$Y = C_a + C_m P_m \quad \text{(A.2)}$$

Set the Lagrangean function:

$$L = C_a^{\alpha} C_m^{1-\alpha} - \lambda (Y - C_a - C_m P_m) \quad \text{(A.3)}$$

$$\frac{\partial L}{\partial C_a} = \alpha C_a^{\alpha-1} C_m^{1-\alpha} - \lambda \quad , \qquad \lambda = \alpha \frac{U}{C_a} \quad , \qquad C_a = \alpha \frac{U}{\lambda} \quad \text{(A.4)}$$

$$\frac{\partial L}{\partial C_m} = (1-\alpha) C_a C_m^{-\alpha} - \lambda \quad , \quad \lambda = (1-\alpha) \frac{U}{C_m P_m} \quad , C_m P = (1-\alpha) \frac{U}{\lambda}$$

Two solutions are possible. Either substitute the first order conditions in the budget constraint and obtain that $Y = \frac{U}{\lambda}$, then substitute U in the first order conditions to obtain the demand equations:

$$C_a = \alpha Y \quad \text{and} \quad C_m P_m = (1-\alpha) Y \quad \text{(A.5)}$$

or equate the lambdas from the first order condition to obtain:

$$C_a = \frac{\alpha}{1-\alpha} C_m P_m \quad \text{(A.6)}$$

Income elasticities are equal to one:

$$\frac{\partial C_a}{\partial Y} = \alpha \quad , \qquad \frac{\partial C_a}{\partial Y}\frac{Y}{C_a} = \alpha \frac{Y}{C_a} = \frac{\alpha}{\alpha} = 1 \qquad \text{(A.7)}$$

Derivation of quasi-homothetic (LES) demand equations

Consumers maximise the utility function:

$$U = c\ln(C_a - g_a) + (1-c)(C_m - g_m) \qquad \text{(A.8)}$$

with respect to the budget constraint:

$$Y = C_a + C_m P_m \qquad \text{(A.9)}$$

Set the Lagrangean function:

$$L = c\ln(C_a - g_a) + (1-c)(C_m - g_m) + \lambda(Y - C_a - C_m P_m) \qquad \text{(A.10)}$$

The first order conditions are:

$$\frac{\partial L}{\partial C_a} = \frac{c}{C_a - g_a} - \lambda \quad , \quad \lambda = \frac{c}{C_a - g_a} \quad , \quad C_a = \frac{c}{\lambda} + g_a \qquad \text{(A.11)}$$

$$\frac{\partial L}{\partial C_m} = \frac{1-c}{C_m - g_m} - P_m\lambda \quad , \lambda = \frac{1-c}{P_m(C_m - g_m)} \quad , \quad C_m P = \frac{1-c}{\lambda} + g_m P_m$$

By equating the lambdas:

$$\frac{c}{C_a - g_a} = \frac{1-c}{P_m(C_m - g_m)} \qquad \text{(A.12)}$$

$$c(C_m - g_m)P_m = (1-c)(C_a - g_a) \qquad \text{(A.13)}$$

$$C_m P_m = \frac{1-c}{c}(C_a - g_a) + g_m P_m \tag{A.14}$$

Substituting in the budget constraint:

$$C_a = Y - C_m P_m = Y - \frac{1-c}{c}(C_a - g_a) - g_m P_m \tag{A.15}$$

$$C_a + \frac{1-c}{c} C_a = Y + \frac{1-c}{c} g_a - g_m P_m \tag{A.16}$$

$$\frac{C_a}{c} = Y + \frac{1-c}{c} g_a - g_m P_m \tag{A.17}$$

$$C_a = cY + (1-c) g_a - c g_m P_m \tag{A.18}$$

The final demands for the two goods are:

$$C_a = g_a + c(Y - g_a - g_m P_m) \tag{A.19}$$

$$C_m P = g_m P_m + (1-c)(Y - g_a - g_m P_m) \tag{A.20}$$

Appendix B: Solutions of the agricultural transformation model

Homothetic preferences

The basic equations of the model are:

$$A = B_a(\alpha N)^{\alpha} \quad \text{and} \quad M = B_m\left[(1-\alpha)N\right]^{\beta+\lambda} \tag{B.1}$$

$$w_a = \alpha \frac{A}{aN} \quad \text{and} \quad w_m = \beta \frac{M}{(1-\alpha)N} P_m$$

$$C_a = cY \quad \text{and} \quad C_m = (1-c)\frac{Y}{P_m}$$

Notice that in order to simplify the algebra, land (*L*), and capital (*K*) were considered fixed and were not included in the agricultural and manufacturing production functions. The aim of this exercise is deriving expressions for the effect of technological change on agricultural employment, and the inclusion of land and capital in the solution of the model would not alter the main results.

Setting the equality of wages in the two sectors gives an expression for the terms of trade between agriculture and the modern sector:

$$P_m = \frac{\alpha}{\beta}\frac{A}{M}\frac{1-a}{a} \tag{B.2}$$

Substituting the terms of trade in the consumption function: $C_a = \frac{c}{1-c} C_m P_m$, an expression is obtained for consumption of the agricultural good:

$$C_a = \frac{c}{1-c} C_m \frac{\alpha}{\beta}\frac{1-a}{a}\frac{A}{M} \tag{B.3}$$

Setting the equilibrium in the product markets the production functions in the two sectors are substituted for consumption of the two goods in the expression above:

$$A = \frac{c}{1-c} M \frac{\alpha}{\beta} \frac{1-a}{a} \frac{A}{M} \tag{B.4}$$

$$a(1-c)\beta = c\alpha(1-\alpha) \tag{B.5}$$

$$a(\beta - \beta c + \alpha c) = c\alpha \tag{B.6}$$

Therefore the share of employment in agriculture is fixed:

$$a = \frac{c\alpha}{\beta - \beta c + \alpha c} \tag{B.7}$$

Hierarchic preferences

The basic equations of the model are:

$$A = B_a(\alpha N)^{\alpha} \quad \text{and} \quad M = B_m\left[(1-\alpha)N\right]^{\beta+\lambda} \tag{B.8}$$

$$w_a = \alpha \frac{A}{aN} \quad \text{and} \quad w_m = \beta \frac{M}{(1-\alpha)N} P_m$$

$$C_a = Y \quad \text{and} \quad C_m = 0 \quad \text{if} \quad Y \leq g_a \tag{B.9}$$

$$C_a = g \quad \text{and} \quad C_m = y - g_a \quad \text{if} \quad Y > g_a$$

As long as income is less than the minimum food requirement $Y<g_aN$, all labour is employed in agriculture $Ca=A$. The share of employment in agriculture is therefore equal to one, and there is no manufacturing sector. The share of employment in agriculture does not change until income reaches a level that allows the consumption of the manufacturing good. When this level of income is achieved, consumption and production in agriculture remain fixed at the level given by the minimum food requirement: $C_a = g_a N = A$.

Starting from the agricultural production function, by setting the equality in the market for the agricultural good, the production function becomes:

$$B_a a^{\alpha} N^{\alpha} = g_a N \tag{B.10}$$

$$a = g_a^{\frac{1}{\alpha}} N^{\frac{1-\alpha}{\alpha}} B_a^{-\frac{1}{\alpha}} \tag{B.11}$$

and the derivative of a with respect to B_a is:

$$\frac{\partial a}{\partial B_a} = -\frac{1}{\alpha} g_a^{\frac{1}{\alpha}} N^{\frac{1-\alpha}{\alpha}} B_a^{-\frac{1+\alpha}{\alpha}} \tag{B.12}$$

Non-homothetic preferences

The basic equations of the model are:

$$A = B_a(\alpha N)^{\alpha} \quad \text{and} \quad M = B_m\left[(1-\alpha)N\right]^{\beta+\lambda} \tag{B.13}$$

$$w_a = \alpha \frac{A}{aN} \quad \text{and} \quad w_m = \beta \frac{M}{(1-\alpha)N} P_m$$

$$C_a = g_a N + c(Y - gN) \quad \text{and} \quad C_m P_m = (1-c)(Y - gN)$$

Setting the equality of wages in the two sectors an expression is obtained for the terms of trade between agriculture and the modern sector:

$$P_m = \frac{\alpha}{\beta} \frac{A}{M} \frac{1-a}{a} \tag{B.14}$$

Substituting the terms of trade in the consumption function: $C_m P_m = \frac{1-c}{c}(C_a - g_a N)$, we obtain:

$$cC_m P_m \frac{\alpha}{\beta}\frac{A}{M}\frac{1-a}{a} = (1-c)(C_a - g_a N) \tag{B.15}$$

Setting the equality in the product markets:

$$cM\frac{\alpha}{\beta}\frac{A}{M}\frac{1-a}{a} = (1-c)(A - g_a N) \tag{B.16}$$

$$c\frac{\alpha}{\beta}\frac{1-a}{a} = (1-c) - (1-c)\frac{ag_a N}{A} \tag{B.17}$$

$$c\alpha(1-a) = \alpha\beta(1-c) - (1-c)\beta\frac{ag_a N}{A} \tag{B.18}$$

$$c\alpha(1-a)A - a\beta(1-c)A + \beta a g_a N(1-c) = 0 \tag{B.19}$$

$$c\alpha(1-a) - a\beta(1-c) + \frac{\beta a g_a N(1-c)}{B_a (aN)^\alpha} = 0 \tag{B.20}$$

Differentiation of *a* with respect to B_a is not simple, but can be calculated implicitly using:

$$\frac{\partial a}{\partial B_a} = -\frac{F'_{B_a}}{F'_a} \tag{B.21}$$

$$F'_{B_a} = \frac{-\beta a g_a N a^\alpha N^\alpha (1-c)}{(B_a a^\alpha N^\alpha)^2} = \frac{-\beta a g_a N(1-c)}{B_a{}^2 a^\alpha N^\alpha} = -\frac{\beta a g_a N(1-c)}{B_a A} \tag{B.22}$$

$$F'_a = -c\alpha - \beta + \beta c + \frac{B_a a^\alpha N^\alpha \beta g_a N(1-c) - \alpha B_a a^{\alpha-1} N^\alpha \beta a g_a N(1-c)}{(B_a a^\alpha N^\alpha)^2} \tag{B.23}$$

$$F'_a = -c\alpha - \beta + \beta c + \frac{\beta g_a N(1-c)(1-\alpha)}{B_a a^\alpha N^\alpha} \tag{B.24}$$

and therefore:

$$\frac{\partial a}{\partial B_a} = -\frac{-\dfrac{\beta a g_a N(1-c)}{B_a A}}{-c\alpha - \beta + c\beta + \dfrac{\beta g_a N(1-c)(1-\alpha)}{A}} \tag{B.25}$$

$$\frac{\partial a}{\partial B_a} = -\frac{-\dfrac{\beta a g_a N(1-c)}{B_a}}{(\beta c - c\alpha - \beta)A + g_a N\beta(1-c)(1-\alpha)} \tag{B.26}$$

In order for the derivative to be negative, the denominator must be negative. Since the first term in the denominator is negative (because $\beta c < \beta$), and the second term is positive (because all the variables and parameters are positive), the derivative is negative if the first term is larger than the second term in absolute value, that is if:

$$A(\beta + c\alpha - \beta c) > g_a N\beta(1-c)(1-\alpha) \tag{B.27}$$

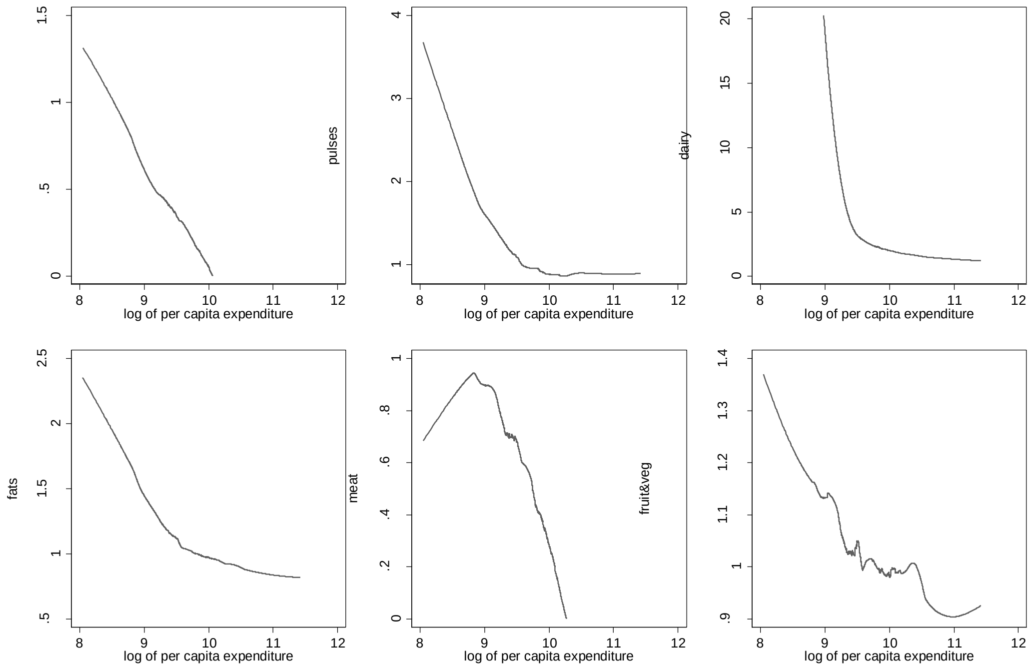
$$\frac{A}{N}[\beta(1-c) + c\alpha] > g_a \beta(1-c)(1-\alpha) \tag{B.28}$$

$$\frac{A}{N} > g_a \frac{\beta(1-c)(1-\alpha)}{\beta(1-c) + c\alpha} \tag{B.29}$$

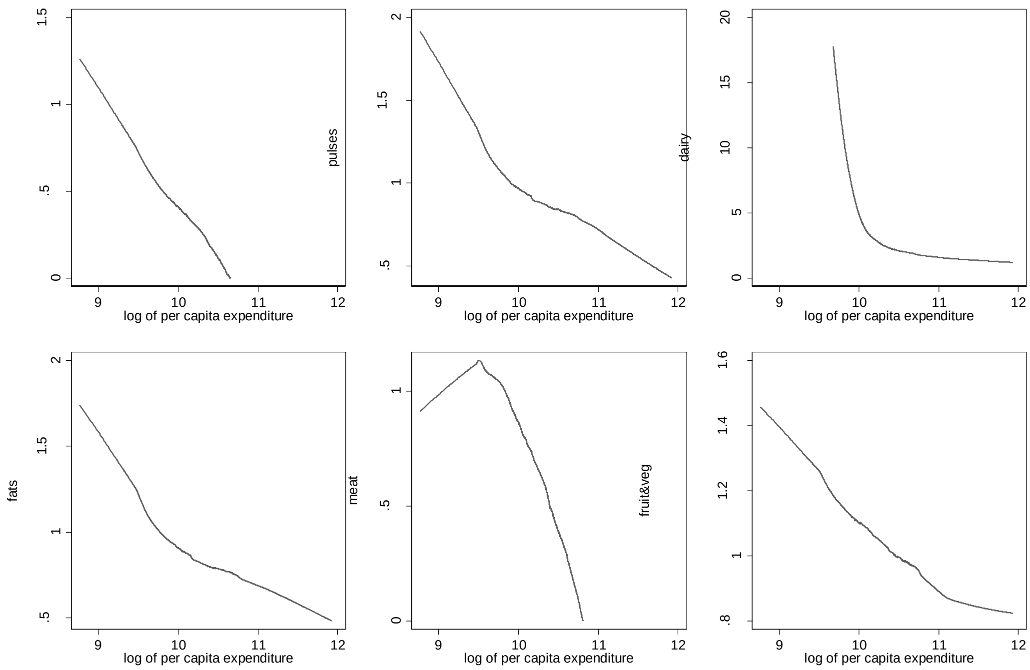
The factor multiplying *g* on the right-hand-side of the inequality is less than one. Therefore, the inequality is assured provided *A/N* is at least as large as *g*.

Appendix C: Semiparametric elasticities by food category

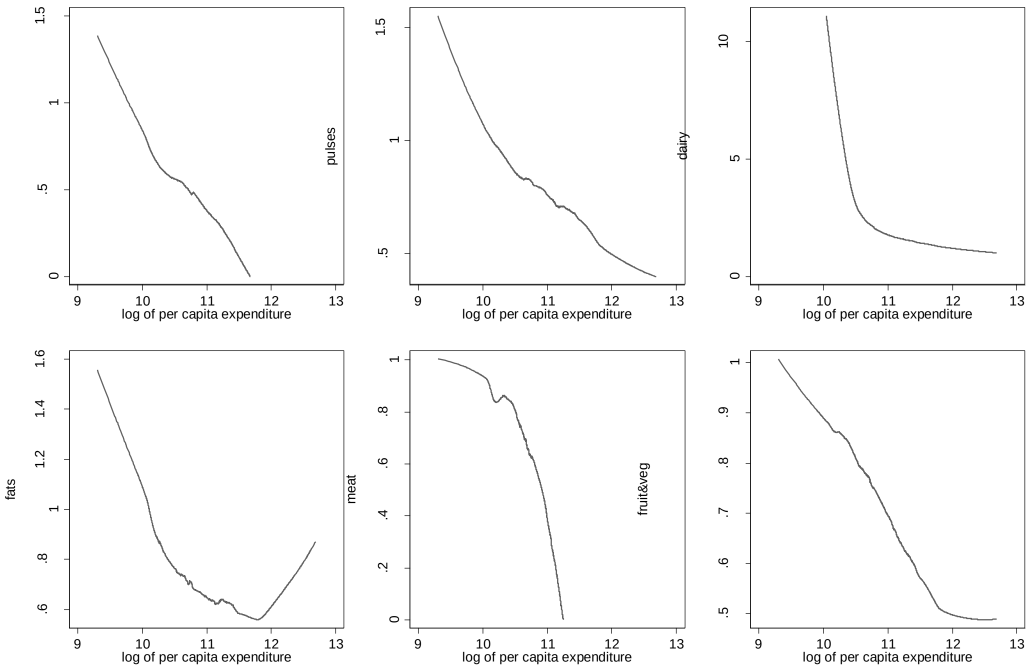
43rd round (1987-88)



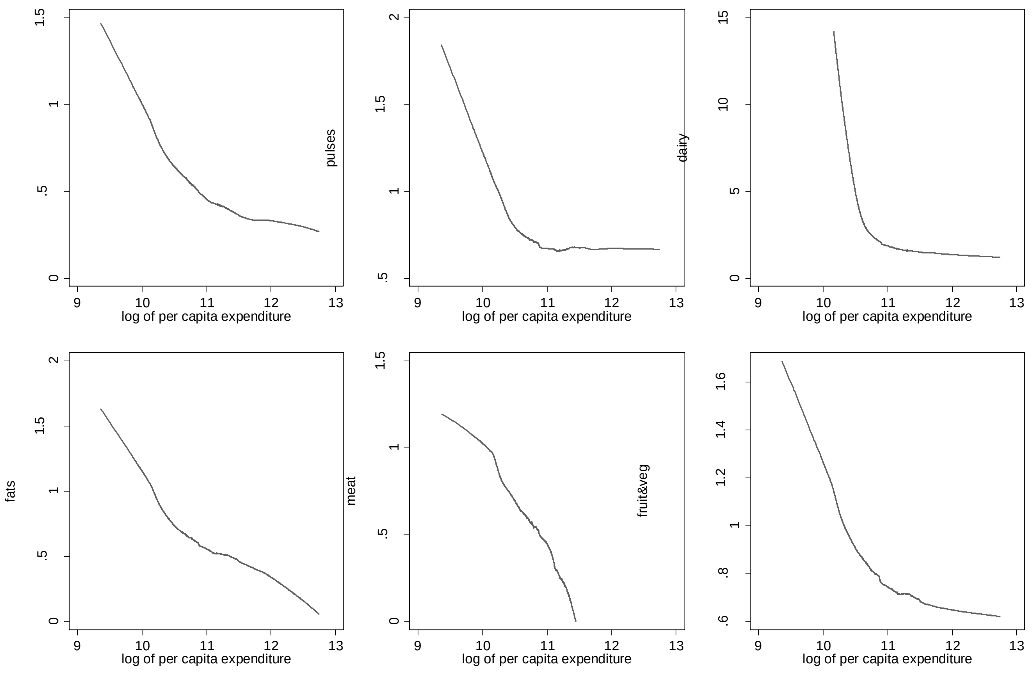
50th round (1993-94)



55th round (1999-00)



61st round (2004-05)



Appendix D: The constant relative risk aversion utility function

The CRRA utility function is:

$$u(c_t) = \frac{c_s^{1-\gamma} - 1}{1-\gamma}, \qquad \text{with} \quad \gamma > 0 \tag{D.1}$$

This utility function incorporates the property of constant relative risk aversion, which makes the cost of risk increasing proportionally with farmers' incomes. The number one in the numerator is required to make the expression defined also in the special case where γ, the coefficient of relative risk aversion, is equal one.

The first derivative of the utility function is positive:

$$u'(c) = c^{-\gamma} \tag{D.2}$$

and utility increases with consumption.

The second derivative is negative:

$$u''(c) = -\gamma c^{-(\gamma+1)} \tag{D.3}$$

and the utility function is concave down, which implies risk aversion on the part of the consumers.

The third derivative, that is the derivative of the marginal utility function, is positive:

$$u'''(c) = \gamma(\gamma+2)c^{-(\gamma+2)} \tag{D.4}$$

which implies prudence.

The absolute coefficient of risk aversion is positive:

$$-u''(c)/u'(c) = \gamma / c \tag{D.5}$$

which means that the consumer prefers a certain level of consumption to an uncertain level of consumption having the same expected value.

Finally, the relative degree of risk aversion is constant:

$$-u''/u' = \gamma \tag{D.6}$$

meaning that the cost of risk is proportional to the level of consumption.

Appendix E: Solution of the seasonal consumption model

This appendix illustrates the step-by-step derivation of the seasonal consumption equations (17) in the text. These equations were obtained using the method of dynamic programming. The notation used is the same already introduced in the main text.

The objective of this dynamic program is to find the level of consumption c_t, at each time over the period considered, that maximises the sum of the net benefits at all times:

$$\max_{c_t} V = \sum_{t=1}^{T} u_t(c_t) \qquad \text{(E.1)}$$

Maximisation occurs subject to the transition equation that characterises the accumulation or decumulation of assets over time:

$$A_{t+1} = h_t(A_t, c_t) \qquad \text{(E.2)}$$

In this particular case the aim is finding the seasonal consumption path c_1, c_2, c_3, that maximises the sum of discounted utilities over three periods (seasons):

$$u(c_s) = \sum_{s=1}^{3} \delta^s \left[\frac{c_s^{\,1-\gamma} - 1}{1-\gamma} \right] \qquad \text{(E.3)}$$

subject to the following budget constraint:

$$A_{s+1} = r(A_s + y_s + c_s) \qquad \text{(E.4)}$$

Where $\delta = \dfrac{1}{1+\rho_{seas}}$ is the time preference seasonal discounting factor, and $r = (1+i_s)$ is the seasonal interest rate.

The following conditions are set: A_1=0 (there are no assets in the first period), A_4=0 (all assets are consumed by the end of the third period), and y_3=0 (no income is produced in the third period).

In the last period the value function to maximise is:

$$V_3 = \max_{c_3} \delta^3 \left[\frac{c_3^{1-\gamma} - 1}{1-\gamma} \right] \tag{E.5}$$

subject to:

$$A_4 = r(A_3 + y_3 - c_3) = r(A_3 - c_3) = 0 \tag{E.6}$$

Consumption in the last period is simply: $\boxed{c_3 = A_3}$ (E.7)

This value is substituted in the value function of the third period:

$$V_3 = \max_{c_3} \delta^3 \left[\frac{A_3^{1-\gamma} - 1}{1-\gamma} \right] \tag{E.8}$$

In the second period the value function is:

$$V_2 = \max_{c_2} \left[\frac{\delta^2 (c_2^{1-\gamma} - 1)}{1-\gamma} + \frac{\delta^3 A_3^{1-\gamma} - 1}{1-\gamma} \right] \tag{E.9}$$

subject to:

$$A_3 = r(A_2 + y_2 - c_2) \tag{E.10}$$

Now A_3 is substituted in the value function:

$$V_2 = \max_{c_2}\left[\frac{\delta^2(c_2{}^{1-\gamma}-1)}{1-\gamma}+\frac{\delta^3 r^{1-\gamma}(A_2+y_2-c_2)^{1-\gamma}-1}{1-\gamma}\right] \tag{E.11}$$

Now differentiate the value function with respect to c_2:

$$\frac{\partial V_2}{\partial c_2}=\frac{(1+\gamma)\delta^2 c_2{}^{-\gamma}}{(1+\gamma)}+\frac{(1-\gamma)\delta^3 r^{1-\gamma}(A_2+y_2-c_2)^{-\gamma}}{(1-\gamma)}(-1) \tag{E.12}$$

$$\frac{\partial V_2}{\partial c_2}=\delta^2 c_2{}^{-\gamma}-\delta^3 r^{1-\gamma}(A_2+y_2-c_2)^{-\gamma} \tag{E.13}$$

Then set the derivative to zero and solve for c_2:

$$c_2{}^{-\gamma}=\delta r^{1-\gamma}(A_2+y_2-c_2)^{-\gamma} \tag{E.14}$$

$$\delta r^{1-\gamma}=(A_2+y_2-c_2)^{\gamma}c_2{}^{-\gamma} \tag{E.15}$$

$$\delta r^{1-\gamma}=\left(\frac{A_2+y_2-c_2}{c_2}\right)^{\gamma} \tag{E.16}$$

$$\left(\delta r^{1-\gamma}\right)^{\frac{1}{\gamma}}=\frac{A_2+y_2}{c_2}-1 \tag{E.17}$$

Set $d=\left(\delta r^{1-\gamma}\right)^{\frac{1}{\gamma}}$ in order to simplify the expression:

$$\frac{A_2+y_2}{c_2}=d+1 \tag{E.18}$$

Consumption in the second period is: $\boxed{c_2=\frac{A_2+y_2}{d+1}}$ (E.19)

The value function of the second period can now be rewritten and simplified in the following way:

$$V_2 = \max_{c_2}\left[\frac{\delta^2\left[\left(\frac{A_2+y_2}{d+1}\right)^{1-\gamma}-1\right]}{1-\gamma}+\frac{\delta^3 A_3^{1-\gamma}-1}{1-\gamma}\right] \tag{E.20}$$

$$V_2 = \max_{c_2}\left[\frac{\delta^2\left(\frac{A_2+y_2}{d+1}\right)^{1-\gamma}}{1-\gamma}+\frac{\delta^3[r(A_2+y_2-c_2)]^{1-\gamma}}{1-\gamma}\right] \tag{E.21}$$

$$V_2 = \max_{c_2}\left[\frac{\delta^2\left(\frac{A_2+y_2}{d+1}\right)^{1-\gamma}}{1-\gamma}+\frac{\delta^3[r(A_2+y_2-\frac{A_2+y_2}{d+1})]^{1-\gamma}}{1-\gamma}\right] \tag{E.22}$$

$$V_2 = \max_{c_2}\left[\frac{\delta^2\left(\frac{A_2+y_2}{d+1}\right)^{1-\gamma}}{1-\gamma}+\frac{\delta^3 r^{1-\gamma}(A_2+y_2)^{1-\gamma}\left(\frac{d}{d+1}\right)^{1-\gamma}}{1-\gamma}\right] \tag{E.23}$$

$$V_2 = \max_{c_2}\left[\frac{\delta^2(A_2+y_2)^{1-\gamma}\left(\frac{1}{d+1}\right)^{1-\gamma}}{1-\gamma}(1+\delta r^{1-\gamma}d^{1-\gamma})\right] \tag{E.24}$$

Now notice that:

$\delta r^{1-\gamma} = d^{\gamma}$ and $1+\delta r^{1-\gamma}d^{1-\gamma} = 1+d$

$$V_2 = \max_{c_2}\left[\frac{\delta^2 (A_2 + y_2)^{1-\gamma}\left(\frac{1}{d+1}\right)^{1-\gamma}}{1-\gamma}(1+d)\right] \tag{E.25}$$

$$V_2 = \max_{c_2}\left[\frac{\delta^2 (A_2 + y_2)^{1-\gamma}(1+d)^{\gamma}}{1-\gamma}\right] \tag{E.26}$$

Now write the value function for the first period:

$$V_1 = \max_{c_1}\left[\frac{\delta(c_1^{1-\gamma} - 1)}{1-\gamma} + \frac{\delta^2 (A_2 + y_2)^{1-\gamma}(1+d)^{\gamma}}{1-\gamma}\right] \tag{E.27}$$

subject to the constraint:

$$A_2 = r(A_1 + y_1 - c_1) = r(y_1 - c_1) \tag{E.28}$$

Substitute A_2 in the value function:

$$V_1 = \max_{c_1}\left[\frac{\delta(c_1^{1-\gamma} - 1)}{1-\gamma} + \frac{\delta^2 (ry_1 - rc_1 + y_2)^{1-\gamma}(1+d)^{\gamma}}{1-\gamma}\right] \tag{E.29}$$

Differentiate the value function with respect to c_1:

$$\frac{\partial V_1}{\partial c_1} = \frac{(1-\gamma)\delta c_1^{-\gamma}}{1-\gamma} + \frac{\delta^2 (1-\gamma)(ry_1 - rc_1 + y_2)^{-\gamma}(1+d)^{\gamma}}{1-\gamma}(-r) \tag{E.30}$$

Set the derivative to zero, simplify the expression, and solve for consumption in the first period:

$$\delta c_1^{-\gamma} = \delta^2 r(ry_1 - rc_1 + y_2)^{-\gamma}(1+d)^{\gamma} \tag{E.31}$$

$$c_1^{-\gamma} = \delta r^{1-\gamma}(y_1 - c_1 + \frac{y_2}{r})^{-\gamma}(1+d)^{\gamma} \qquad \text{(E.32)}$$

$$\delta r^{1-\gamma} = (y_1 - c_1 + \frac{y_2}{r})^{\gamma}(1+d)^{-\gamma} c_1^{-\gamma} \qquad \text{(E.33)}$$

$$\delta r^{1-\gamma} = \left(\frac{y_1 + \frac{y_2}{r} - c_1}{c_1} \right)^{\gamma} (1+d)^{-\gamma} \qquad \text{(E.34)}$$

$$\delta r^{1-\gamma} = \left(\frac{y_1 + \frac{y_2}{r}}{c_1} - 1 \right)^{\gamma} (1+d)^{-\gamma} \qquad \text{(E.35)}$$

$$(\delta r^{1-\gamma})^{\frac{1}{\gamma}} = \left(\frac{y_1 + \frac{y_2}{r}}{c_1} - 1 \right)(1+d)^{-1} \qquad \text{(E.36)}$$

$$d = \left(\frac{y_1 + \frac{y_2}{r}}{c_1} - 1 \right)(1+d)^{-1} \qquad \text{(E.37)}$$

$$d(1+d) = \left(\frac{y_1 + \frac{y_2}{r}}{c_1} - 1 \right) \qquad \text{(E.38)}$$

$$\frac{y_1 + \frac{y_2}{r}}{c_1} = d(1+d) = d + d^2 + 1 \qquad \text{(E.39)}$$

Consumption in the first period is: $\boxed{c_1 = \frac{y_1 + \frac{y_2}{r}}{1 + d + d^2}}$ (E.40)

Bibliography

Ahluwalia (1975) Inequality, Poverty and Development, *Journal of Development Economics* 6 307-42

Alesina, A. and Rodrik, D. (1994) Distributive Politics and Economic-Growth, *Quarterly Journal of Economics* 109 (2): 465-90

Anand, S. and Kanbur, R. (1993) The Kuznets Process and the Inequality Development Relationship, *Journal of Development Economics* 40 (25-52)

Antonelli, G. B. (1886) *Sulla Teoria Matematica Della Economia Politica*, Pisa: Tipografia del Falchetto

Atkinson, A. B., Gomulka, J. and Stern, N. H. (1990) Spending on Alcohol - Evidence from the Family Expenditure Survey 1970-1983, *Economic Journal* 100 (402): 808-27

Azariadis, C. (2006) The Theory of Poverty Traps, in S. Bowles, S. N. Durlauf and K. Hoff (eds), *Poverty Traps*, Princeton: Princeton University Press

Azariadis, C. and Stachurski, J. (2005) Poverty Traps, in P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Dordrecht: North Holland

Bagliano, Fabio C. and Bertola, Giuseppe (2004) *Models for Dynamic Macroeconomics*, Oxford: Oxford University Press

Baland, J. M. and Ray, D. (1991) Why Does Asset Inequality Affect Unemployment - a Study of the Demand Composition Problem, *Journal of Development Economics* 35 (1): 69-92

Banerjee, A. (2004) The Two Poverties, in S. Dercon (ed.) *Insurance against Poverty*, Oxford: Oxford University Press

Banerjee, A. and Iyer, L. (2005) History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India, *American Economic Review* 95 (4): 1190-213

Banerjee, A., Iyer, L. and Somanathan, R. (2005) History, Social Divisions and Public Goods in Rural India, *Journal of the European Economic Association* 3 (2-3): 639-47

Banerjee, A. and Picketty, T. (2005) Top Indian Incomes, 1922-2000, *The World Bank Economic Review* 19 (1): 1-20

Banerjee, A. V. and Newman, A. F. (1993) Occupational Choice and the Process of Development, *Journal of Political Economy* 101 (2): 274-98

Banerjee, B. and Knight, J. B. (1985) Caste Discrimination in the Indian Urban Labor-Market, *Journal of Development Economics* 17 (3): 277-307

Banks, J., Blundell, R. and Lewbel, A. (1997) Quadratic Engel Curves and Consumer Demand, *Review of Economics and Statistics* 79 (4): 527-39

Bardhan, P. and Udry, C. (1999) *Development Microeconomics*, London: Oxford Univerisity Press

Barr, A. and Packard, T. (2000) Revealed and Concealed Preferences in the Chilean Pension System: An Experimental Investigation, *Discussion Paper Series Department of Economics University of Oxford*

Barrett, C. B., Bezuneh, M. and Aboud, A. (2001) Income Diversification, Poverty Traps and Policy Shocks in Cote D'ivoire and Kenya, *Food Policy* 26 (4): 367-84

Baulch, Bob and Hoddinott, John (2000) *Economic Mobility and Poverty Dynamics in Developing Countries*, London: Frank Cass

Becker, Gary Stanley (1975) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, Chicago: University of Chicago Press, 1980

Behrman, J. R. and Deolalikar, A. B. (1989) Agricultural Wages in India: The Role of Health, Nutrition, and Seasonality, in D. E. Sahn (ed.) *Seasonality in Third World Agriculture*, Baltimore and London: The John Hopkins University Press

Berry, R.A. and Cline, W.R. (1979) *Agrarian Structure and Productivity in Developing Countries*, Geneva: International Labour Organisation

Bertola, G. (1993) Factor Shares and Savings in Endogenous Growth, *American Economic Review* 83 (5): 1184-98

Besley, T. (1995) Nonmarket Institutions for Credit and Risk Sharing in Low-Income Countries, *Journal of Economic Perspectives* 9 (3): 115-27

Besley, T. and Burgess, R. (2000) Land Reform, Poverty Reduction, and Growth: Evidence from India, *Quarterly Journal of Economics* 115 (2): 389-430

Besley, T., Burgess, R. and Esteve-Volart, B. (2007) The Policy Origins of Poverty and Growth in Rural India, in t. Besley and L. J. Cord (eds), *Delivering on the Promise of Pro-Poor Growth,* New York and Washington DC: Palgrave Macmillan and the World Bank

Bhalla, G.S. and Singh, G. (2001) *Indian Agriculture: Four Decades of Development*, New Delhi: Sage

Bhalla, S.S. (2002) *Imagine There Is No Country: Poverty, Inequality and Growth in the Era of Globalisation*, Washington D.C.: Institute for International Economics

Bhalotra, S. and Attfield, C. (1998) Intrahousehold Resource Allocation in Rural Pakistan: A Semiparametric Analysis, *Journal of Applied Econometrics* 13 (5): 463-80

Binswanger, H. P. (1981) Attitudes Towards Risk:Theoretical Implications of an Experimental Study in Rural India, *Economic Journal* 91 867-90

Binswanger, H., Deininger, K. and Feder, g. (1995) Power, Distortions, Revolt and Reform in Agricultural Land Relations, in J. R. Behrman and P. V. Srinivasan (eds), *Handbook of Development Economics*, Amsterdam: Elsevier

Binswanger, H., Yang, M. C., Bowers, A. and Mundlak, Y. (1987) On the Determinants of Cross-Country Aggregate Agricultural Supply, *Journal of Econometrics* 36 (1-2): 111-31

Bleany, M. F. (1976) *Under-Consumption Theories*, London: Lawrence and Wishart

Bloom, D. E., Canning, D. and Sevilla, J. (2003) Geography and Poverty Traps, *Journal of Economic Growth* 8 (4): 355-78

Blundell, R. and Stoker, T. M. (2005) Heterogeneity and Aggregation, *Journal of Economic Literature* 43 (2): 347-91

Bourguignon, F. (1990) Growth and Inequality in the Dual Model of Development - the Role of Demand Factors, *Review of Economic Studies* 57 (2): 215-28

Bourguignon, F., Ferreira, F. H. G. and Menendez, M. (2007) Inequality of Opportunity in Brazil, *Review of Income and Wealth* 53 (4): 585-618

Bowles, S., Durlauf, S. N. and Hoff, K. (2006a) Introduction, in S. Bowles, S. N. Durlauf and K. Hoff (eds), *Poverty Traps*, Princeton: Russel Sage Foundation

—— (2006b) *Poverty Traps*, Princeton: Princeton University Press

Buchinsky, M. (1994) Changes in the United-States Wage Structure 1963-1987 - Application of Quantile Regression, *Econometrica* 62 (2): 405-58

Burgess, R. and Pande, R. (2005) Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment, *American Economic Review* 95 (3): 780-95

Caldwell, J. C., Reddy, P. H. and Caldwell, P. (1985) Educational Transition in Rural South-India, *Population and Development Review* 11 (1): 29-51

Cameron, A.C. and Trivedi, P.K. (2005) *Microeconometrics*, Cambridge: Cambridge University Press

Cardenas, J. C. and Carpenter, J. (2008) Behavioural Development Economics: Lessons from Field Labs in the Developing World, *Journal of Development Studies* 44 (3): 311-38

Carneiro, P., Hansen, K. T. and Heckman, J. J. (2003) Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice, *International Economic Review* 44 (2): 361-422

Carroll, C. D. (2001) A Theory of the Consumption Function, with and without Liquidity Constraints, *Journal of Economic Perspectives* 15 (3): 23-45

Carroll, C. D. and Kimball, M. S. (2007) *Precautionary Savings and Precautionary Wealth*, entry in the New Palgrave DIctionary of Economics 2nd edition

Carroll, C. D. and Samwick, A. A. (1998) How Important Is Precautionary Saving?, *Review of Economics and Statistics* 80 (3): 410-9

Carter, M. R. (1984) Identification of the Inverse Relationship between Farm Size and Productivity - an Empirical-Analysis of Peasant Agricultural Production, *Oxford Economic Papers-New Series* 36 (1): 131-45

—— (1997) Environment, Technology, and the Social Articulation of Risk in West African Agriculture, *Economic Development and Cultural Change* 45 (3): 557-90

Carter, M. R. and Barrett, C. B. (2006) The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach, *Journal of Development Studies* 42 (2): 178-99

Carter, M. R. and Maluccio, J. A. (2003) Social Capital and Coping with Economic Shocks: An Analysis of Stunting of South African Children, *World Development* 31 (7): 1147-63

Carter, M. R. and May, J. (2001) One Kind of Freedom: Poverty Dynamics in Post-Apartheid South Africa, *World Development* 29 (12): 1987-2006

Caselli, F. and Coleman, W. J. (2001) The Us Structural Transformation and Regional Convergence: A Reinterpretation, *Journal of Political Economy* 109 (3): 584-616

Chamarbagwala, Rubiana (2008) Regional Returns to Education, Child Labour and Schooling in India, *Journal of Development Studies* 44 (2): 233-57

Chambers, R., Longhurst, R. and Pacey, A. (1981) *Seasonal Dimensions to Rural Poverty*, Exter: SRP Ltd.

Chambers, R.G. and Pope, R.D. (1992) Engel's Law and Linear-in-Moments Aggregation, *American Journal of Agricultural Economics* 74 682-8

Chatterji, A. K. (1992) *Monsoons, Cyclones and Floods in India*, London: Sangam Books

Chenery, H. B., Robinson, S. and Syrquin, M. (1986) *Industrialization and Growth: A Comparative Study*, New York: Oxford University Press

Chenery, H.B. and Syrquin, M. (1975) *Patterns of Development, 1950-1970*, London: Oxford University Press

Chiang, A. (1974) *Fundamental Methods of Mathematical Economics*, Tokyo: Mac-Graw-Hill

Cole, S., Gine, X., Tobacman, J., Topalova, P., Townsend, P. and Vickery, J. (2009) Barriers to Household Risk Management: Evidence from India, *Harvard Business School Working Paper 09-116*

Da Corta, L. and Venkateshwarlu, D. (1999) Unfree Relations and the Feminization of Agricultural Labour in Andhra Pradesh, 1975-1995, in T.J. Byres, K. Kapadia and J. Lerche (eds), *Rural Labour Relations in India*, London: Frank Cass

Datt, G. and Ravallion, M. (1998) Farm Productivity and Rural Poverty in India, *Journal of Development Studies* 34 (4): 62-85

Deaton, A (1990) Saving in Developing Countries: Theory and Review, in The World Bank (ed.) *Proceedings of the World Bank Annual Conference on Development Economics 1989*, Washington DC: The World Bank

—— (1991) Saving and Liquidity Constraints, *Econometrica* 59 (5): 1221-48

—— (1997) *The Analysis of Household Surveys*, Washington DC: The World Bank

—— (2003) Prices and Poverty in India, 1987-2000, *Economic and Political Weekly* January 25 362-8

—— (2008) Price Trends in India and Their Implications for Measuring Poverty *Economic and Political Weekly* February 9 43-9

Deaton, A. and Dreze, J. (2002) Poverty and Inequality in India: A Re-Examination, *Economic and Political Weekly* (September 7): 3729-48

Deaton, A and Dreze, J (2008) Nutrition in India: Facts and Interpretations *Mimeo*

Deaton, A and Muellbauer, John (1980a) *Economics and Consumer Behaviour*, Cambridge: Cambridge University Press

Deaton, A. and Muellbauer, J. (1980b) An Almost Ideal Demand System, *American Economic Review* 70 (3): 312-26

Deaton, A. and Kozel, V. (2005) *The Great Indian Poverty Debate*, New Delhi: Macmillan India

Deaton, A. and Tarozzi, A. (2000) Prices and Poverty in India, Research Program in Development Studies, *mimeo*

Deaton, A. and Zaidi, S. (2002) Guidelines for Constructing Consumption Aggregates for Welfare Analysis, *Living Standard Measurement Study Working Paper* 135 (104)

Deininger, K. and Liu, Y. (2009) Longer Term Impacts of Self-Help Groups in India, *Policy research Working Paper 4886* The World Bank

Deininger, K. and Squire, L. (1996) A New Dataset Mesuring Income Inequality, *World Bank Economic Review* 10 565-91

Dejanvry, A. and Sadoulet, E. (1983) Social Articulation as a Condition for Equitable Growth, *Journal of Development Economics* 13 (3): 275-303

Dela Cruz-Dona, Rena and Martina, Alan (2000) Some Links between Education, Household Well-Being and Credit Markets: Evidence from Rural Philippines *Oxford Development Studies* 28 (3)

Dercon, S. (2004) *Insurance against Poverty*, Oxford: Oxford University Press

—— (2007) *Designing Insurance for the Poor*, IFPRI

Dercon, S. and Krishnan, P. (2000a) In Sickness and in Health: Risk Shaving within Households in Rural Ethiopia, *Journal of Political Economy* 108 (4): 688-727

—— (2000b) Vulnerability, Seasonality and Poverty in Ethiopia, *Journal of Development Studies* 36 (6): 25-53

DES (1975-2005) *Agricultural Wages in India*, New Delhi: DES

—— (2005) *Golden Jubilee of Andhra Pradesh: 50 Years*, Hyderabad: DES

Devereux, S., Vaitla, B. and Houenstein Swan, S. (2008) *Seasons of Hunger*, London Pluto Press

Dhawan, B. D. (1988) *Irrigation in India's Agricultural Development*, New Delhi: Sage

Diewert, W. E. (2004) The Treatment of Seasonal Products, in ILO (ed.) *Consumer Price Index Manual: Theory and Practice,* Geneva: ILO

Dixit, A. (1973) Models of Dual Economies, in J.A. Mirrlees and N.H. Stern (eds), *Models of Economic Growth: Proceedings of a Conference Held by the International Economic Association at Jerusalem,* London: Macmillan

Dreze, J and Sen, A. (1995) *India, Economic Development and Social Opportunity,* Delhi: Oxford University Press

Dreze, J. and Sen, A. (2002) *India: Development and Pariticipation*, Oxford: Oxford University Press

Dreze, Jean and Kingdon, G. G. (1999) School Participation in Rural India, *mimeo*

Duan, Naihua, Manning, Willard G., Morris, Carl N. and Newhouse, Joseph P. (1983) A Comparison of Alternative Models for the Demand for Medical Care, *Journal of Business & Economic Statistics* 1 (2): 115-26

Dubey, A. and Palmer-Jones, R. (2005a) Poverty in India since 1983: New Poverty Counts and Robust Poverty Comparisons, *Artha Vijnana* XLVII (3-4): 287-328

—— (2005b) Price, Price Indexes and Poverty Counts in India During 1980s and 1990s: From Cpis to Poverty Lines?, *Artha Vijnana* XLVII (3-4): 259-86

—— (2005c) Prices, Price Indexes and Poverty Counts in India during 1980s and 1990s: Calculation of Unit Value Consumer Price Indexes, *Artha Vijnana* 2005 (XLVII): 223-58

Duraisamy, P (2002) Changes in the Returns to Education in India, 1983-94: By Gender, Age-Cohort, and Location, *Economics of Education Review* 21 (6): 609-22

Duryea, S. and Arends-Kuenning, M. (2003) School Attendance, Child Labor and Local Labor Market Fluctuations in Urban Brazil, *World Development* 31 (7): 1165-78

Dutta, Bhakaskar and Ramaswami, Bharat (2001) Targeting and Efficiency in the Public Distribution System. The Case of Andhra Pradesh and Maharashtra, *Economic and Political Weekly* XXXVI (18): 1524-32

Dutta, Puja Vasudeva (2006) Returns to Education: New Evidence from India, 1983-1999, *Education Economics* 14 (4): 431-51

Echevarria, C. (1997) Changes in Sectoral Composition Associated with Economic Growth, *International Economic Review* 38 (2): 431-52

Eswaran, M. and Kotwal, A. (1993) A Theory of Real Wage Growth in Ldcs, *Journal of Development Economics* 42 (2): 243-69

Fafchamps, M. (1992) Solidarity Networks in Preindustrial Societies - Rational Peasants with a Moral Economy, *Economic Development and Cultural Change* 41 (1): 147-74

—— (2003) *Rural Poverty, Risk and Development*, Bodmin: MPG Books Ltd

Fafchamps, M., Udry, C. and Czukas, K. (1998) Drought and Saving in West Africa: Are Livestock a Buffer Stock?, *Journal of Development Economics* 55 (2): 273-305

Fan, J. and Gijbels, I. (1996) *Local Polynomial Modelling and Its Applications*, London: Chapman & Hall/CRC

Fan, S. G. and Hazell, P. (2001) Returns to Public Investments in the Less-Favored Areas of India and China, *American Journal of Agricultural Economics* 83 (5): 1217-22

Fisher, Irving (1930) *The Theory of Interest, as Determined by the Impatience to Spend Income and Opportunity to Invest It*, New York: Macmillan

Foster, A. D. and Rosenzweig, M. R. (1996) Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution, *American Economic Review* 86 (4): 931-53

—— (2004) Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms: India, 1970-2000, *Economic Development and Cultural Change* 52 (3): 509-42

Fraser Jackson, L. (1984) Hierarchic Demand and the Engel Curve for Variety, *The Review of Economics and Statistics* 66 (1): 8-15

Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002) Time Discounting and Time Preference: A Critical Review, *Journal of Economic Literature* 40 (2): 351-401

Friedman, M. (1957) *A Theory of the Consumption Function* Princeton University Press

Fujita, M., Krugman, P. and Venables, A.J. (2001) *The Spatial Economy*, Princeton: MIT Press

Gadgil, S., Rao, P. R. Seshagiri and Rao, K. Narahari (2002) Use of Climate Information for Farm-Level Decision-Making: Rainfed Groundnut in Southern India, *Agricultural Systems* 74 431-57

Galab, S. and Rao, N.C. (2003) Women and Self Help Groups: Poverty Alleviation and Empowerment, in C.H. H. Rao and S. M. Dev (eds), *Andhra Pradesh Development: Economic Reforms and Challenges Ahead,* Hyderabad: CESS

Galor, O. and Zeira, J. (1993) Income-Distribution and Macroeconomics, *Review of Economic Studies* 60 (1): 35-52

Gang, I. N., Sen, K. and Yun, M. S. (2008) Poverty in Rural India: Caste and Tribe, *Review of Income and Wealth* 54 (1): 50-70

Garikipati, S. (2006) Feminisation of Agricultural Labour and Women's Domestic Status: Evidence from Labour Households in India, *mimeo*

Gertler, P. and Gruber, J. (2002) Insuring Consumption against Illness, *American Economic Review* 92 (1): 51-70

Gill, G. J. (1991) *Seasonality and Agriculture in the Developing World*, Cambridge: Cambridge University Press

Gine, X., Townsend, R. and Vickery, J. (2008) Patterns of Rainfall Insurance Participation in Rural India, *World Bank Economic Review* 22 (3): 539-66

Gine, X. and Yang, D. (2007) Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi, *Policy Research Working Paper 4425* The World Bank, Washington DC

Glewwe, Paul (1999) *The Economics of School Quality Investments in Developing Countries : An Empirical Study of Ghana*, Basingstoke: Macmillan in association with Centre for the Study of African Economies, University of Oxford

Glewwe, P. and Hall, G. (1998) Are Some Groups More Vulnerable to Macroeconomic Shocks Than Others? Hypothesis Tests Based on Panel Data from Peru, *Journal of Development Economics* 56 (1): 181-206

Gollin, D., Parente, S. and Rogerson, R. (2002) The Role of Agriculture in Development, *American Economic Review* 92 (2): 160-4

Gong, X., van Soest, A. and Zhang, P (2005) The Effects of the Gender of Children on Expenditure Patterns in China: A Semiparametric Analysis, *Journal of Applied Econometrics* 20 (4): 509-27

Gopal, Iyer K. (2004) *Distressed Migrant Labour in India: Key Human Rights Issues*, Edited book

Government of India (2007) *Poverty Estimates for 2004-05*, New Delhi: Press Information Bureau

Govinda, R. (2001) Providing Education for All in India, in R. Govinda (ed.) *India Education Report: A Profile of Basic Education,* New Delhi: Oxford University Press

Graham, B. S. and Temple, J. R. W. (2006) Rich Nations, Poor Nations: How Much Can Multiple Equilibria Explain?, *Journal of Economic Growth* 11 (1): 5-41

Gulati, A. and Narayan, S. (2003) *The Subsidy Syndrome in Indian Agriculture*, Oxford: Oxford University Press

Gurgand, M. (2003) Farmer Education and the Weather: Evidence from Taiwan (1976-1992), *Journal of Development Economics* 71 (1): 51-70

Hanushek, E. A. (1992) The Trade-Off between Child Quantity and Quality, *Journal of Political Economy* 100 (1): 84-117

Hardle, W. (1990) *Applied Non-Parametric Regression*, New York, NY: Cambridge University Press

Hayami, Y. and Godo, Y. (2005) *Development Economics: From the Poverty to the Wealth of Nations*, London: Oxford University Press

Hayami, Y. and Ruttan, V.W. (1971) *Agricultural Development: An International Perspective*, Baltimore, MD: John Hopkins Press

Hazell, P. (1992) The Appropriate Role of Agricultural Insurance in Developing Countries, *Journal of International Development* 4 567-81

Hazell, P., Pomareda, C. and Valdes, A. (1986) *Crop Insurance for Agricultural Development*, Washington DC: IFPRI

Hazell, P. and Ramaswamy, C. (1991) *The Green Revolution Reconsidered: The Impact of High Yielding Varieties in South India*, Baltimore MD: John Hopkins University Press

Hazell, Peter B R (1984) Sources of Increased Instability in Indian and U.S. Cereal Production, *American Journal of Agricultural Economics* 66 (3): 302-11

Heckman, J. J., Lochner, Lance J. and Todd, Petra e. (2003) Fifty Years of Mincer Earnings Regressions, *mimeo*

Himanshu (2007) Recent Trends in Poverty and Inequality: Some Preliminary Results, *Economic and Political Weekly* February 10 497-508

Hoddinott, J. and Kinsey, B. (2001) Child Growth in the Time of Drought, *Oxford Bulletin of Economics and Statistics* 63 (4): 409-38

Hoff, K. and Pandey, P. (2006) Discrimination, Social Identity, and Durable Inequalities, *American Economic Review* 96 (2): 206-11

IEG (2007) *An Impact Evaluation of India's Second and Third Andhra Pradesh Irrigation Project. A Case of Poverty Reduction with Low Economic Returns,* Washington DC: World Bank

ILO (2004) *Consumer Price Index Manual: Theory and Practice*, International Labour Office

Jacoby, H. G. and Skoufias, E. (1997) Risk, Financial Markets, and Human Capital in a Developing Country, *Review of Economic Studies* 64 (3): 311-35

—— (1998) Testing Theories of Consumption Behavior Using Information on Aggregate Shocks: Income Seasonality and Rainfall in Rural India, *American Journal of Agricultural Economics* 80 (1): 1-14

Jairath, J. (2000) Participatory Irrigation Management in Andhra Pradesh, paper presented at South Asia Regional Poverty Monitoring and Evaluation Workshop Conference, New Dheli, June 8-10 2000

Jalan, J. and Murgai, R. (2007) An Effective "Targeting Shortcut"? An Assessment of the 2002 Below-Poverty-Line Census Method, *mimeo*

Jeffrey, C., Jeffery, P. and Jeffery, R. (2005) Reproducing Difference? Schooling, Jobs, and Empowerment in Uttar Pradesh, India, *World Development* 33 (12): 2085-101

Jeffrey, C., Jeffery, R. and Jeffery, P. (2004) Degrees without Freedom: The Impact of Formal Education on Dalit Young Men in North India, *Development and Change* 35 (5): 963-86

Jensen, R. (2000) Agricultural Volatility and Investments in Children, *American Economic Review* 90 (2): 399-404

Jha, R. (2000) Growth, Inequality and Poverty in India: Spatial and Temporal Characteristics, *Economic and Political Weekly* March 11 921-8

Jha, Shikha, Srinivasan, P. V. and Landes, Maurice (2007) *Indian Wheat and Rice Sector Policies and the Implications of Reform*, Economic Research Report Number 41, USDA

Johnston, J. and Dinardo, J. (1997) *Econometric Methods*, New York, NY: McGraw-Hill

Jorgenson, D. W. (1961) The Development of a Dual Economy, *Economic Journal* 71 (282): 309-34

—— (1967) Surplus Agricultural Labour and Development of a Dual Economy, *Oxford Economic Papers-New Series* 19 (3): 288-312

Kalbfleisch, J. D. and Prentice, R. L. (1980) *The Statistical Analysis of Failure Time Data*, New York ; Chichester: Wiley

Kambhapati, U. S. and Pal, S. (2001) Role of Parental Lilteracy in Explaining Gender Difference: Evidence from Child Schooling in India, *The European Journal of Development Research* 13 (2): 99-119

Kaul, V. (2001) Early Childhood Care and Education, in R. Govinda (ed.) *India Education Report: A Profile of Basic Education*, New Delhi: Oxford University Press

Kedir, A. and Girma, S. (2007) Quadratic Engel Curves with Measurement Error: Evidence from a Budget Survey, *Oxford Bulletin of Economics and Statistics* 69 (1): 123-38

Kerr, J. (2002) Watershed Development, Environmental Serivces, and Poverty Alleviation in India, *World Development* 30 (8): 1387-400

Kerr, J., Pangare, J. and Pangare, V. L. (2002) *Watershed Development Projects in India. An Evaluation.*, Research Report 127, Washington DC: IFPRI

Keynes, J.M (1936) *The General Theory of Employment, Interest, and Money*, London: Macmillan

Kimball, M. S. (1990) Precautionary Saving in the Small and in the Large, *Econometrica* 58 (1): 53-73

King, E. M. and Lillard, L. A. (1987) Education Policy and School Attainment in Malaysia and the Philippines, *Economics of Education Review* 6 (2): 167-81

Kingdon, G. G. (1998) Does the Labour Market Explain Lower Female Schooling in India?, *Journal of Development Studies* 35 (1): 39-65

—— (2005) Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure, *Economic Development and Cultural Change* 53 (2): 409-51

—— (2007) The Progress of School Education in India, *Oxford Review of Economic Policy* 23 (2): 168-95

Kingdon, G. G. and Unni, Jeemol (2001) Education and Women's Labour Market Outcomes in India, *Education Economics* 9 (2): 173-95

Kirby, K. N., Godoy, R., Reyes-Garcia, V., Byron, E., Apaza, L., Leonard, W., Perez, E., Vadez, V. and Wilkie, D. (2002) Correlates of Delay-Discount Rates: Evidence from Tsimane' Amerindians of the Bolivian Rain Forest, *Journal of Economic Psychology* 23 (3): 291-316

Kochar, A. (1999) Smoothing Consumption by Smoothing Income: Hours-of-Work Responses to Idiosyncratic Agricultural Shocks in Rural India, *Review of Economics and Statistics* 81 (1): 50-61

—— (2003) Higher Schooling Investments and Poverty in the Indian Economy, *mimeo*

—— (2004) Urban Influences on Rural Schooling in India, *Journal of Development Economics* 74 (1): 113-36

Kodde, D. A. (1986) Uncertainty and the Demand for Education, *Review of Economics and Statistics* 68 (3): 460-7

Kremer, M., Muralidharan, K., Chaudhury, N., Hammer, J. and Rogers, R. H. (2005) Teacher Absence in India: A Snapshot, *Journal of the European Economic Association* 3 (2-3): 658-67

Krishna, A., Kapila, M., Pathak, S., Porwal, M., Singh, K. and Singh, V. (2004) Falling into Poverty in Villages of Andhra Pradesh, *Economic and Political Weekly* July 17 3249-56

Krishnaji, N. (2001) Poverty, Gender and Schooling: A Study of Two Districts in Andhra Pradesh, in A. Vaidyanathan and N. P. R. Gopinathan (eds), *Elementary Education in Rural India,* New Delhi: Sage

Krugman, P. (1991) Increasing Returns and Economic-Geography, *Journal of Political Economy* 99 (3): 483-99

—— (1997) *Development, Geography, and Economic Theory*, Cambridge: MIT Press

Kurian, N.J. (2000) Widening Regional Disparities in India: Some Indicators, *Economic and Political Weekly* February 12 538-50

Kurosaki, T. and Fafchamps, M. (2002) Insurance Market Efficiency and Crop Choices in Pakistan, *Journal of Development Economics* 67 (2): 419-53

Kuznets, S. (1955) Economic Growth and Income Inequality, *American Economic Review* 45 1-18

—— (1957) Quantitative Aspects of the Economic Growth of Nations Ii. Industrial Distribution of National Product and Labor Force, *Economic Development and Cultural Change* 4 (Supplement): 1-110

Laitner, J. (2000) Structural Change and Economic Growth, *Review of Economic Studies* 67 (3): 545-61

Leibenstein, H. (1957) *Economic Backwardness and Economic Growth*, New York: Wiley

Leonard, d. and Van Long, N. (1992) *Optimal Control Theory and Static Optimization in Economics*, Cambridge: Cambridge University Press

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

Levhari, D. and Weiss, Y. (1974) Effect of Risk on Investment in Human Capital, *American Economic Review* 64 (6): 950-63

Levinson, B. A. and Holland, D. C. (1996) The Cultural Production of the Euducated Person: An Introduction, in B A. Levinson, D. E. Foley and D. C. Holland (eds), *The Cultural Production of the Educated Person: Critical Ethnographies of Schooling and Local Practice*, Albany, NY: State University of New York Press

Lewis, W. A. (1954) Economic Development with Unlimited Supply of Labour, *Manchester School* 22 (2): 139-91

Lilleord, H.B., Gine, X., Townsend, P. and Vickery, J. (2005) Weather Insurance in Semi-Arid India, *mimeo*

Lim, Y, . and Townsend, R. M. (1998) General Equilibrium Models of Financial Systems: Theory and Measurement in Village Economies, *Review of Economic Dynamics* 1 (1): 59-118

Lipton, Michael (1989) *New Seeds and Poor People*, With Richard Longhurst Johns Hopkins Studies in Development series. Baltimore: Johns Hopkins University Press

Lipton, Michael, Litchfield, Julie and Faures, Jean-Marc (2003) The Effects of Irrigation on Poverty: A Framework for Analysis, *Water Policy* 5 413-27

Ljungqvist, L. (1993) Economic Underdevelopment - the Case of a Missing Market for Human-Capital, *Journal of Development Economics* 40 (2): 219-39

Loury, G. C. (1981) Intergenerational Transfers and the Distribution of Earnings, *Econometrica* 49 (4): 843-67

Luseno, W. K, McPeak, J. G., Barrett, C. B., Little, P. D. and Gebru, G. (2003) Assessing the Value of Climate Forecast Information for Pastoralist: Evidence from Southern Ethiopia and Northern Kenya, *World Development* 31 (9): 1477-94

Maddala, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press

Matsuyama, K. (1991) Increasing Returns, Industrialization, and Indeterminacy of Equilibrium, *Quarterly Journal of Economics* 106 (2): 617-50

—— (1992) Agricultural Productivity, Comparative Advantage, and Economic Growth, *Journal of Economic Theory* 58 (2): 317-34

Mazumder, B. (2005) The Apple Falls Even Closer to the Tree Than We Thought: New and Revised Estimates of the Intergenerational Inheritance of Earnings, in S.

Bowles, H. Gintis and M. Osborne Groves (eds), *Unequal Chances: Family Background and Economic Success*, Princeton, NJ: Princeton University Press

McCulloch, N. and Baulch, B. (2000) Simulating the Impact of Policy Upon Chronic and Transitory Poverty in Rural Pakistan, *Journal of Development Studies* 36 (6): 100-30

McKenzie, D. J. (2003) How Do Households Cope with Aggregate Shocks? Evidence from the Mexican Peso Crisis, *World Development* 31 (7): 1179-99

McKenzie, D. J. and Woodruff, C. (2006) Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Enterprise, *Economic Development and Cultural Change* 55 (1): 3-42

Miron, J. (1996) *The Economics of Seasonal Cycles*, MIT Press

Modigliani, F. (1966) Life Cycle Hypothesis of Saving, Demand for Wealth and Supply of Capital, *Social Research* 33 (2): 160-217

MOL (several issues) New Delhi: Labour Bureau of the Ministry of Labour, Government of India

Morduch, J. (1990) Risk, Production and Savings: Theory and Evidence from Indian Households, *mimeo*

—— (1994) Poverty and Vulnerability, *American Economic Review* 84 (2): 221-5

—— (2004) Consumption Smoothing across Space: Testing Theories of Risk-Sharing in the Icrisat Study Region of South India, in S. Dercon (ed.) *Insurance against Poverty*, Oxford Oxford University Press

—— (2006) Microinsurance: The Next Revolution?, in A. V. Banerjee, R. Benabou and D. Mookherjee (eds), *Understanding Poverty*, Oxford: Oxford University Press

MRD (1994) *Guidelines for Watershed Development Programs*, New Dehli: Government of India

Mundlak, Y. (2000) *Agriculture and Economic Growth: Theory and Measurement*, Cambridge, MA: Harvard University Press

—— (2005) Economic Growth: Lessons from Two Centuries of American Agriculture, *Journal of Economic Literature* 43 (4): 989-1024

Mundlak, Y., Larson, D. F. and Crego, A. (1997) Agricultural Development: Issues, Evidence, and Consequences, *mimeo*

Muralindharan, M. and Kremer, M. (2006) Public and Private Schools in Rural India, *mimeo*

Murphy, K. M., Shleifer, A. and Vishny, R. (1989a) Income-Distribution, Market-Size, and Industrialization, *Quarterly Journal of Economics* 104 (3): 537-64

Murphy, K. M., Shleifer, A. and Vishny, R. W. (1989b) Industrialization and the Big Push, *Journal of Political Economy* 97 (5): 1003-26

Myrdal, G. (1957) *Economic Theory and Underdeveloped Regions*, London: Duckworth

Nambissan, G. B. (1996) Equity in Education? Schooling of Dality Children in India, *Economic and Political Weekly* 31 (16/17): 1011-19

Nambissan, G. B. and Sedwal, M. (2001) Education for All: The Situation of Dalit Children in India, in R. Govinda (ed.) *India Education Report: A Profile of Basic Education*, New Delhi: Oxford University Press

NCERT (1997) *All India Education Survey*, New Delhi: National Council of Educational Research and Training

Newbery, David M. G. and Stiglitz, Joseph E. (1981) *The Theory of Commodity Price Stabilization : A Study in the Economics of Risk*, Oxford: Clarendon

NSSO (1993) *Instructions to Field Staff*, NSSO

—— (1999) *Instructions to Field Staff*, NSSO

Nurske, R (1953) *Problems of Capital Formation in Underdeveloped Countries*, Oxford: Oxford University Press

Oblitas, K. and Peter, R. (1999) Transferring Irrigation Management to Farmers in Andhra Pradesh, India, *World Bank Technical Paper 449*

Office of the Director of Census Operations (2001) *Census of India: Andhra Pradesh, Village and Town Directory 1991 and 2001, and Urban and Rural Pca 1961-2001*

Ota, M. and Moffatt, P. G. (2007) The within-Household Schooling Decision: A Study of Children in Rural Andhra Pradesh, *Journal of Population Economics* 20 (1): 223-39

Pal, S. (2004) How Much of the Gender Difference in Child School Enrolment Can Be Explained? Evidence from Rural India, *Bulletin of Economic Research* 56 (2): 133-58

Palmer-Jones, R. and Sen, K. (2001) On India's Poverty Puzzles and Statistics of Poverty, *Economic and Political Weekly* January, 20 211-7

Palmer-Jones, Richard and Sen, Kunal (2003) What Has Luck Got to Do with It? A Regional Analysis of Poverty and Agricultural Growth in Rural India, *Journal of Development Studies* 40 (1): 1-31

Patnaik, U. (2007) Neoliberalism and Rural Poverty in India, *Economic and Political Weekly* July 28 3132-50

Paxson, Christina H (1993) Consumption and Income Seasonality in Thailand, *Journal of Political Economy* 101 (1): 39-72

Pender, J. L. (1996) Discount Rates and Credit Markets: Theory and Evidence from Rural India, *Journal of Development Economics* 50 257-96

Perotti, R. (1992) Income-Distribution, Politics, and Growth, *American Economic Review* 82 (2): 311-6

Persson, T. and Tabellini, G. (1994) Is Inequality Harmful for Growth, *American Economic Review* 84 (3): 600-21

Pitt, M. M. and Khandker, S. R. (2002) Credit Programmes for the Poor and Seasonality in Rural Bangladesh, *Journal of Development Studies* 39 (2): 1-24

Pradhan, B. K, R., Salulja M., Roy, P.K. and Shetty, S. L. (2003) *Household Savings and Investment Behaviour in India*, Mumbai: NCAER

Prais, S.J and Houthakker, H.S. (1971) *The Analysis of Family Budgets*, Cambridge: Cambridge University Press

Prasad, K. V. E. (1987) *Wastage, Stagnation and Inequality of Opportunity in Rural Primary Education: A Case Study of Andhra Pradesh*, New Delhi: Ministry of Human Resources and Development

Pritchett, L. (1997) Divergence, Big Time, *Journal of Economic Perspectives* 11 (3): 3-17

PROBE Team (1999) *Public Report on Basic Education in India*, New Delhi: Oxford University Press

Psacharopoulos, G. (1985) Returns to Education - a Further International Update and Implications, *Journal of Human Resources* 20 (4): 583-604

—— (1994) Returns to Investment in Education - a Global Update, *World Development* 22 (9): 1325-43

—— (2002) Returns to Investment in Education: A Further Update, *World Bank Policy Research Working Paper* (2881)

Quizon, Jaime and Binswanger, Hans (1986) Modeling the Impact of Agricultural Growth and Government Policy on Income Distribution in India, *World Bank Economic Review* 1 (1): 103-48

Ranis, G (1988) Analytics of Development: Dualism, in H. Chenery and T. N. Srinivasan (eds), *Handbook of Development Economics,* Amsterdam: Elsevier

Rao, M.G., Shand, R.T. and Kalirajan, K.P (1999) Convergence of Incomes across Indian States: A Divergent View, *Economic and Political Weekly* March 27 769-78

Ravallion, M. (2004) Pro-Poor Growth: A Primer, *World Bank Policy Research Working Paper* 3242

—— (2008) A Global Perspective on Poverty in India, *Economic and Political Weekly* October 25 31-7

Ravallion, M., Chen, S. and Sangraula, P. (2007) New Evidence on the Urbanization of Global Poverty, *Population and Development Review* 33 (4): 667-

Ravallion, M. and Datt, G. (1996) How Important to India's Poor Is the Sectoral Composition of Economic Growth?, *World Bank Economic Review* 10 (1): 1-25

Ravallion, M. and Jalan, J. (2002) Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China, *Journa of Applied Econometrics* 17 (4): 329-46

Ravi, C. and Indrakant, S. (2003) Food Security and Public Distribution System, in *Andhra Pradesh Development Economic Reforms and Challenges Ahead,* Hyderabad: Center for Economic and Social Studies

Rawlings, L. B. and Rubio, G.M. (2005) Evaluating the Impact of Conditional Cash Transfer Programs, *The World Bank Research Observer* 20 (1): 29-55

Ray, D. (1998) *Development Economics*, Princeton: Princeton University Press

Ray, R. and Lancaster, G. (2005) On Setting the Poverty Line Based on Estimated Nutrient Prices, *Economic and Political Weekly* January 1 46-56

Reddy, S.S, Galab, S. and Rao, P. P. (2003a) Environment and Sustainable Development: Status and Strategies, in C.H Hanumantha Rao and S. Mahendra Dev (eds), *Andhra Pradesh Development Economic Reforms and Challenges Ahead*, Hyderabad: Center for Economic and Social Studies

Reddy, S.S., Galab, S. and Rao, P. P. (2003b) Trends and Determinants of Poverty, in C.H Hanumantha Rao and S. Mahendra Dev (eds), *Andhra Pradesh Development Economic Reforms and Challenges Ahead*, Hyderabad: Center for Economic and Social Studies

Reddy, V. Ratna (2003) Irrigation: Development and Reforms, in C.H Hanumantha Rao and S. Mahendra Dev (eds), *Andhra Pradesh Development Economic Reforms and Challenges Ahead*, Hyderabad: Center for Economic and Social Studies

Reddy, V. Ratna and Rao, R. Nageswara (2003) Primary Education: Causes and Constraints, in C.H. Hanumantha Rao and S. Mahendra Dev (eds), *Andhra Pradesh Development*, Hyderabad: CESS

Riboud, M., Savchenko, Y. and Tan, H. (2007) *The Knowledge Economy and Education and Training in South Asia,* Human Development Unit South Asia Region, The World Bank

Robinson, P. M. (1988) Root-N-Consistent Semiparametric Regression, *Econometrica* 56 (4): 931-54

Robinson, S. (1975) A Note on the U-Hypothesis Relating Income Inequality and Economic Development, *American Economic Review* 66 437-40

Romer, D. (2006) *Advanced Macroeconomics*, London: McGraw-Hill/Irwin

Ros, J. (2000) *Development Theory and the Economics of Growth*, Ann Arbor: The University of Michigan Press

Rosenstein-Rodan, P. (1943) Problems of Industrialization of Eastern and Southeastern Europe, *Economic Journal* 53 202-11

Rosenzweig, M. R. and Wolpin, K. I. (1993) Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries - Investments in Bullocks in India, *Journal of Political Economy* 101 (2): 223-44

Sachs, J., Bajpai, N. and Ramiah, A. (2002) Understanding Regional Economic Growth in India, *Asian Economic Papers* 1 (3): 32-62

Sahn, D. E. (1989) Policy Recommendation for Improving Food Secuirity, in D. E. Sahn (ed.) *Seasonality in Third World Agriculture*, Baltimore and London

Santhapparaj, A.S. (1996) Job Search and Earnings of Migrants in Urban Labour Market: A Study of Madurai Metropolis, *Indian Journal of Labour Economics* 39 (2): 269-86

Schady, N. R. (2004) Do Macroeconomic Crises Always Slow Human Capital Accumulation?, *World Bank Economic Review* 18 (2): 131-54

Screpanti, e. and Zamagni, S. (2005) *An Outline of the History of Economic Thought*, Oxford: Oxford University Press

Seale, J., Regmi, A. and Bernstein, J. (2003) International Evidence on Food Consumption Patterns, *USD Technical Bulletin* (1904): 1-67

Sen, A. H. and Himanshu (2004) Poverty and Inequality in India - Ii Widening Disparities during the 1990s, *Economic and Political Weekly* (September 25): 4361-75

Singer, H. (1949) Economic Progress in under-Developed Countries, *Social Research* 1 (March)

Sinha, Sidharth (2004) Agricultural Insurance in India, *Economic and Political Weekly* June 19 2605-12

Sipahimalani, V. (2000) *Financing of Elementary Education in India in the 1990s*, New Delhi: World Bank

Skoufias, E. and Parker, S. W. (2006) Job Loss and Family Adjustments in Work and Schooling during the Mexican Peso Crisis, *Journal of Population Economics* 19 (1): 163-81

Smith, A. (1970) *The Wealth of Nations*, Harmondswhorth: Penguin

Snow, A. and Warren, R. S. (1990) Human-Capital Investment and Labor Supply under Uncertainty, *International Economic Review* 31 (1): 195-206

Solow, R. M. (1956) A Contribution to the Theory of Economic-Growth, *Quarterly Journal of Economics* 70 (1): 65-94

Springate-Baginsky, O., Reddy, V. R., Reddy, M. G. and Galab, S. (2002) *Watershed Development in Andhra Pradesh. Improving Policy-Livelihood Relationships in India*, Policy Review Paper 5, DFID

Stiglitz, J. E. (1969) Effects of Income, Wealth, and Capital Gains Taxation on Risk-Taking, *Quarterly Journal of Economics* 83 (2): 263-83

Subrahmanyam, S. (2003) Regional Disparities: Causes and Remedies, in C.H Hanumantha Rao and S. Mahendra Dev (eds), *Andhra Pradesh Development Economic Reforms and Challenges Ahead*, Hyderabad: Center for Economic and Social Studies

Sujatha, K. (2001) Education among Scheduled Tribes, in R. Govinda (ed.) *India Education Report: A Profile of Basic Education*, New Delhi: Oxford University Press

Sundaram, K. and Tendulkar, S. D. (2003) Poverty among Social and Economic Groups in India in 1990s, *Economic and Political Weekly* December 13 5263-76

Tarozzi, A. (2005) The Indian Public Distribution System as Provider of Food Security: Evidence from Child Nutrition in Andhra Pradesh, *European Economic Review* 49 (5): 1305-30

Temple, J. (2005) Dual Economy Models: A Primer for Growth Economists, *Manchester School* 73 (4): 435-78

Theil, H., Chung, C.F. and Seale, J.L. (1989) *International Evidence on Consumption Patterns*, Greenwich, CT: JAI Press Inc.

Thomas, D., Beegle, K., Frankenberg, E., Sikoki, B., Strauss, J. and Teruel, G. (2004) Education in a Crisis, *Journal of Development Economics* 74 (1): 53-85

Tilak, J. B. G. (1990) Education and Earnings: Gender Differences in India, *International Journal of Development Planning Literature* 5 (4): 131-9

—— (2001) Financing Elementary Education in India, in R. Govinda (ed.) *India Education Report: A Profile of Basic Education*, New Delhi: Oxford University Press

—— (2002) Determinants of Household Expenditure on Education in Rural India, *mimeo*

Timmer, C.P. (1988) The Agricultural Transformation, in H. Chenery and T. N. Srinivasan (eds), *Handbook of Development Economics Vol. 1*, Amsterdam: Elsevier

Tobin, J. (1958) Estimation of Relationships for Limited Dependent-Variables, *Econometrica* 26 (1): 24-36

Townsend, R. M. (1994) Risk and Insurance in Village India, *Econometrica* 62 (3): 539-91

Trivedi, K. (2006) Educational Human Capital and Levels of Income: Evidence from States in India, 1965-92, *Journal of Development Studies* 42 (8): 1350-78

Udry, C. (1994) Risk and Insurance in a Rural Credit Market - an Empirical-Investigation in Northern Nigeria, *Review of Economic Studies* 61 (3): 495-526

UNESCO (2007) Regional Overview: South and West Asia *Education for all global monitoring report* , UNESCO

Unni, Jeemol (1995) Returns to Education by Gender among Wage Employees in Urban India, *mimeo*

Vaidyanathan, A. and Gopinathan, N. P. R. (2001) *Elementary Education in Rural India*, New Delhi: Sage

Varshney, A. (1995) *Democracy, Development, and the Countryside: Urban-Rural Struggles in India*, Cambridge: Cambridge University Press

Vyas, V. S. and Singh, Surjit (2006) Crop Insurance in India: Scope for Improvement, *Economic and Political Weekly* November 4 4585-94

Wade, R. (1988) *Village Republics*, Cambridge University Press

Walker, T. S. and Ryan, James G. (1990) *Village and Household Economies in India's Semi-Arid Tropics*, Baltimore: Johns Hopkins University Press

White, H. and Masset, E. (2003) The Importance of Household Size and Composition in Constructing Poverty Profiles: An Illustration from Vietnam, *Development and Change* 34 (1): 105-26

Wolpin, K. I. (1982) A New Test of the Permanent Income Hypothesis - the Impact of Weather on the Income and Consumption of Farm Households in India, *International Economic Review* 23 (3): 583-94

Working, H. (1943) Statistical Laws of Family Expenditure, *Journal of the American Statistical Association* 38 43-56

World Bank (1990-2005) *World Development Indicators*

—— (2001) *Attacking Poverty*, World Development Report Washington DC: The World Bank

—— (2005) *India's Water Economy: Bracing for a Turbulent Future*, Washington DC: The World Bank

—— (2006) *World Development Report: Equity and Development*, Washington DC: The World Bank

Yadav, M. S., Bhardwaj, M., Sedwal, M. and Gaur, N. (2001) Learning Conditions and Learning Achievement in Primary School, in R. Govinda (ed.) *India Education Report: A Profile of Basic Education*, New Delhi: Oxford University Press

Yatchew, A. (2003) *Semiparametric Regression for the Applied Econometrician*, Cambridge: Cambridge University Press

Zimmerman, F. J. and Carter, M. R. (2003) Asset Smoothing, Consumption Smoothing and the Reproduction of Inequality under Risk and Subsistence Constraints, *Journal of Development Economics* 71 (2): 233-60