Title
Essays in Schooling, Nutrition and Public Policy

Author
Ramachandran, Maithili

Publication Date
2011

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Essays on Schooling, Nutrition and Public Policy

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Economics

by

Maithili Ramachandran

August 2011

Dissertation Committee:

Professor Anil B. Deolalikar, Co-Chairperson
Professor Jorge M. Agüero, Co-Chairperson
Professor Aman Ullah
Professor David Malueg

The Dissertation of Maithili Ramachandran is approved:

Co-Chairperson

Co-Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisors for their patience, instruction and encouragement throughout my doctoral studies. My first debt of gratitude is to Professor Anil Deolalikar. Anil has always provided the perspective and balance I needed to be a good research scholar. His advice, empathy and efforts took me through the most difficult points of my graduate career. I learnt from him the importance of striving for excellence without getting derailed by the pursuit of perfection.

I would like to thank my co-chair, Jorge Agüero for teaching me everything I know about the craft of research in economics. This dissertation would not have taken shape without his guidance and infinite commitment. I thank Professor Aman Ullah for patiently answering all my questions on nonparametric econometrics. His intuitive explanations greatly facilitated my programming and choice of regressions. Thank you to Professor David Malueg for reading through my work so carefully and helping me improve the presentation.

The genesis of my penultimate chapter lies in the research project Jorge introduced me to, in the Summer of 2008. This chapter is co-authored with him and included in this dissertation with his permission. I thank the personnel of the Central Statistical Office of Zimbabwe for making a 10% sample of the 2002 Population Census available to us. Special thanks to Craig Richardson for sharing his time-series data on rainfall in Zimbabwe. I am grateful to Todd Sorensen, Steven Helfand, Mindy Marks, Chris Barrett, and Prashant Bharadwaj for their extensive comments on the paper. Discussions at many invited presentations contributed to the evolution of this paper - I would like to thank seminar participants at the Applied Economics colloquium in UCR, the Cornell-IBEPP conference, PACDEV, CSAE-Oxford, the PAA's Economic Demography Workshop, the inaugural Applied Microeconomics conference at Claremont McKenna College and the NBER Children's Program meeting for their thoughtful reactions. Likewise, my work on the calorie-income relationship in Maharashtra would not have been possible without the National Sample Survey data that Anil provided me so generously.

I would like to thank my friends, Asha and Brad for their love and kindness. For the last five years, they have been my family away from home, taking care of me and helping me through countless practical difficulties. My final, silent gratitude is to god and my family.

To my parents and sister.

ABSTRACT OF THE DISSERTATION

Essays on Schooling, Nutrition and Public Policy

by

Maithili Ramachandran

Doctor of Philosophy, Graduate Program in Economics
University of California, Riverside, August 2011
Professor Anil B. Deolalikar, Co-Chairperson
Professor Jorge M. Agüero, Co-Chairperson

This thesis examines human capital outcomes in developing countries. It is especially concerned with evaluating the policies meant to raise the rate of accumulation of such capital in poor households.

Chapter 3 considers this problem in the context of calorie deprivation in India. Rapid economic growth has been advocated as the instrument of choice in tackling undernutrition in India. Yet, while India's annual economic growth has never dipped below four percent in two decades, calorie consumption has been falling across the income distribution. This poses a disturbing trend against the background of widespread malnutrition. Chapter 2 details how the literature has typically defined nutrition, and attempted to derive the causal impact of income on it. The next essay investigates the calorie-income puzzle using a random sample of poor households in rural Maharashtra from the 2004 National Sample Survey. The nonparametric estimate of the expenditure elasticities of calories reveals that there is a gradual fall from 0.28 to zero over the income distribution, which roughly

translates to a halving of the elasticities from 1983. Controlling for household and district level characteristics does not alter the basic estimate. It closes by showing that food price inflation could not have been the cause of the calorie decline.

Chapter 4 examines the intergenerational transmission of schooling in Zimbabwe. After Independence in 1980, Zimbabwe implemented a substantial reform of the racially-segregated education system it inherited from colonial times. A key element of the reform was the elimination of restrictions governing progress from primary to secondary school. Consequently, primary school graduates of 1980 entered secondary school at a rate three times higher than the class of 1979. Exploiting the fuzzy discontinuity implicit in this natural experiment, I find that an additional year of schooling acquired by a woman increased her child's by about five percent of a standard deviation. Estimates of schooling transmitted from fathers to children were thrice as large, significant and robust. Descriptive evidence suggests that public investment in the quality of schooling would have enhanced the intergenerational benefits of the reform.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Human capital, as Strauss and Thomas put it, is multidimensional, spanning a wide array of investments. This collection of essays focuses on schooling and nutrition, which are but two indicators of human capital. Schooling and nutrition have long dominated public policy discussions, especially in developing countries. This is not surprising, given their ability to profoundly affect the well-being of nations. They are powerful drivers of economic change, shifting many countries from a low to a high growth cycle. In a study based on a panel of about one hundred countries observed from 1965 to 1995, Barro (2001) shows that human capital is strongly related to national growth. In broad strokes, higher schooling attainment not only increases complementarity of the labor input with technology, it stimulates growth by inducing lower fertility. Further, the quality of schooling has an equal, if not greater impact than the level, on outcomes like adult earnings, morbidity, as well as on child mortality, anthropometrics and schooling. Thus, human capital creation has a tremendous policy appeal.

Education and health are probably the two most important concerns for policy-makers in poor countries. In recent times, a substantial amount of public funds have been devoted to creating the right incentives and programs to nudge households in the direction of greater human capital accumulation. Conditional cash transfers (CCTs) have become popular in the education sphere; the well-known instances include Mexico's PROGRESA-Oportunidades, Bolsa Escola in Brazil, and the Noon Meal Scheme in India. Malawi and Morocco are the latest entrants to this style of fighting poverty; pilot CCTs were implemented in the late 2000s with the goal of reducing the ultra-poverty rate, increasing school enrolment and establishing longer-term employment opportunities for poor households.

The fundamental expectation from anti-poverty programs is that in the presence of the right incentives, households will accelerate human capital investment. This in turn would better their economic prospects and help them break the intergenerational transmission of poverty. Development economists have repeatedly looked for evidence bearing out this expectation. Their findings have been positive but at times tinged with uncertainty, mostly due to the difficulty of obtaining a robust causal estimate. Program evaluation is a concern of the essays presented here but not as an end in itself. Programs are used as the setting against which I can study why households make certain choices in the level and quality of their human capital investment, as well as the kinds of outcomes - immediate and long-term - that emerge for them. While the outcomes help me evaluate whether a particular program was effective, the evaluations themselves do not reveal *why* they may have failed. Thus, one of the questions I attempt to answer in each essay, is what mechanisms of development were in play and if the reason a policy worked was because it triggered the relevant mechanism.

In Chapter 3, the objective is to understand if the nutrition of poor households can be improved by increasing their incomes. The setting is rural Maharashtra, a state in India where the policy thinking for the better over two decades has been that hunger is the consequence of insufficient incomes; hence, economic growth is the channel that can best cure undernutrition. India is a country where the proportion of underweight children below five years of age has stayed above 40 percent despite three decades of employment generation programs, mother-child health schemes and most recently, nutrition supplementation efforts. The rates of anemia among adult women and general caloric-undernourishment have also remained persistently high – above 20 percent. This has contrasted strongly with the rate of economic growth, which has averaged over four percent annually since the Nineties. In effect, although the cross-sectional correlation between household income and calorie consumption is positive, the correlation over time for any single observation tends to be negative.

To test if the policy emphasis on economic growth has therefore had a positive impact on nutrition, Chapter 3 uses household survey data on consumption and expenditure habits of households over a period of twenty years and estimates the income elasticity of calorie consumption for each point in the per capita household expenditure distribution. The data delivers an income elasticity of calories in 2004 that is only half its magnitude in 1983 and not even significantly different from zero for the top quartile of households.

Over the same period, the income elasticity of calorie prices are more or less unchanged. The poorer half of the sample shows a decided trend towards greater fat and protein consumption – a substitution pattern motivated by a preference for variety. There is some evidence to suggest that calorie *requirements* could have declined in the twenty year period, given health improvements, medical advances and falling physical demands of work. The study recommends against relying on income policies to raise calorie intake, unless it is by targeting the households in the lower tail of the income distribution.

Chapter 4 is an exploration of the intergenerational accumulation of schooling. Here the setting is not so much a CCT as a government reform of the education system in a poor country. This is the education reform of Zimbabwe following its Independence in 1980 from apartheid. The reform provides a wonderful natural experiment to study the effect of parent education on child human capital.

The study uses a ten percent random sample of the 2002 Zimbabwe Population Census. Two generations are considered: parents (restricted to the heads of households and their spouses in the sample) in the age group of 28 to 44 years and children matched to these parents, selected from the ages of six to fifteen years in 2002. The literature on estimating the effects of parental schooling on children's outcomes is fraught with the difficulty of accounting for omitted variable bias and measurement errors. The main threat to validity that this Chapter 4 must tackle is, if the effect on the child human capital attributed to parental education is simply the reflection of a collection of unobserved influences correlated with parental schooling. The study navigates these tricky waters by taking advantage of the regression discontinuity design (RDD) embedded in the reform. The fuzzy RDD provides the opportunity to estimate with great precision, the effect of parental schooling on different measures of child education for the sample of observations in the neighborhood of the cutoff. It sacrifices some external validity for the estimate obtained. Nevertheless, the study succeeds on several levels: it is able to show that merely the level of parental schooling has important effects for the rate of accumulation of child schooling. The estimates are stronger for fathers than for mothers – the reasons for this are yet to be fully understood – it is possible that the effects of higher maternal schooling affects children through multiple channels. The quality of schooling is not explicitly accounted for in the 2SLS estimates. However, it can be inferred that the estimates are a lower bound, largely because various indicators of schooling quality record a steep decline over the period of the reform.

Seen in terms of the human capital investments it engineered, the reforms realized much of their transformational power: Zimbabwe achieved universal primary school access and literacy in a very brief period of time. Economic growth increased and poverty fell steadily until government spending on education reversed from an annual three percent of GDP in the Eighties to less than two percent in the Nineties. This makes the intergenerational gains in human capital as of 2002 even more impressive. The study concludes by considering in brief, the pathways through which the schooling investments occurred. It finds that exposure to the reforms for women is strongly correlated with reduced fertility, delayed marrying-age and assortative matching.

Despite the common view that educating women is a viable mechanism for economic development, few studies actually explore the degree of human capital accumulation in subsequent generations. Most such studies have also documented the results for developed countries. Given how crucial education has been shown to breaking the inter-generational transmission of poverty, Chapter 4 makes a vital contribution by presenting robust evidence in favor of the such policies.

The dissertation is laid out as follows: Chapter 2 provides a literature review of the measurement of income-calorie elasticities. It is followed by Chapter 3 which presents this issue in Maharashtra and evaluates the policy wisdom behind encouraging household nutrition investments. Chapter 4 moves to the other aspect of human capital highlighted here, schooling. It discusses the intergenerational transmission of this investment against the background of the education reform in Zimbabwe. Chapter 5 summarizes the contributions and concludes.

Chapter 2

A Review of the Literature on Calorie-Income Elasticities

2.1 Nutrition, Nutrients and Income

Nutrition has been a topic of long-standing interest to development economists. Several papers have been devoted to understanding the determinants of nutrition and as many have examined the impact of higher or better nutrition on numerous outcomes, including but not limited to infant and child mortality, schooling, adult health, labor productivity and lifetime income.

The theme of the subsequent chapter centers on the extent to which nutrition, as represented by caloric intake, is affected by the economic status of a household, in turn represented by the size of household expenditure. For this reason, I review mostly the literature that speaks directly to this relationship. I am concerned with describing the typical definitions chosen for nutrition in this literature, the econometric measurement of its relationship to income, and interpreting the estimates obtained from various poor economies and over time.

I begin with a general framework linking nutrition to a its principal determinants and outcomes. This framework is borrowed from Schiff and Valdes (1990). Suppose N stands for nutrition. Nutrition is distinct from nutrient intake, which will be used to refer to the quantity of a nutrient consumed by an individual or household. Nutrient intake is denoted by n. Usually, most studies think of calories when they speak of nutrient intake, although the term can encompass macronutrients like proteins and fats as well as micronutrients (such as vitamins and trace minerals). Thus, nutrition is the output of nutrient intake, and other inputs; so the equation below is essentially a nutrition production function:

$$N = N(n, q, p, k, H, sex, age, I\{rural\}) \tag{2.1}$$

where the input vectors following n are $q =$ the size of non-nutrient attributes of food, $p =$ privately-provided inputs, $k =$ publicly-provided inputs, and $H =$ the health status of the individual in the past and current periods.

To fix ideas, we could think of nutrition as the height, weight or body-mass-index of an individual. It could also be an indicator variable, denoting if an individual is obese, anemic, or whether a household has any chronically malnourished children under age five. Clearly, these "outputs" are different from health, which is a much broader term. Health could be measured as the rate of incidence of a major disease in a household over a period of time, as the absence of stress-related ailments, as a composite of similar such measures. Health is different from nutrition in that it may be influenced by the quantity and quality of medical care a household seeks and is able to afford from time to time. Medicine is not quite an input into nutrition though. Schiff and Valdes (1990) propose an analogue of the nutrition production function for health. They consider all of the above inputs in the nutrition production function as influencing health, in addition to which, they include m for the medical services availed by a household.

For simplicity and ease of estimation, several studies equate nutrition to calorie intake. This decision skips over estimating the effect of other covariates on nutrition. For instance, non-nutrient attributes of food have turned out to be very important in the preferences of households. Indian data[1] shows that a larger income does not promote always higher nutrient intake, and even households falling below a calorie-poverty line trade-off higher nutrient consumption for variety in diet. These choices can be read as hurting the household's nutrition, but this might be true only of the most impoverished households. If, following an increase in income, a household elects to spend a larger fraction on non-nutrient attributes, this need not lead to a decline in long-term nutrition. Non-nutrient attributes, such as food quality and freshness can affect nutrition positively, by supplying more nutrients per unit of food than do foods of a lower grade.

Nutrient intake is likely to dominate all others in raising nutrition mainly when a household is severely undernourished. At very low levels of n, the marginal effect of all

[1] Indian data referred to are the consumption and expenditure rounds of the National Sample Surveys; see Behrman and Deolalikar (1987) for a discussion on the shift in household expenditure patterns from nutrients to other aspects of food.

other inputs on nutrition would be negligible. At the same time, nutrient intake does affect nutrition positively and significantly, but this marginal effect diminishes as nutrient intake rises. As Schiff and Valdes suggest,

$$\frac{\partial^2 N}{\partial n^2} \quad < \quad 0$$

$$\frac{\partial N}{\partial x} \quad \rightarrow \quad 0 \text{ as } n \rightarrow 0 \text{ where } x = q, p, k.$$

From the nutrition production function, the estimate of interest is the elasticity of nutrition with respect to income. From the preceding discussion, the elasticity of nutrition can be equated to the elasticity of nutrient intake only at very low levels of calorie or protein consumption. Otherwise, these two elasticities are far from being identical. Their relationship is mediated by a whole set of income elasticities, viz. the income elasticity of food quality, the income elasticity of public goods like sewerage, potable water, and electricity, and the elasticity of nutrition with respect to all these variables.

Using ϵ_{xz} to denote the elasticity of a variable x with respect to another variable z, the income elasticity of nutrition can be stated as:

$$\epsilon_{NY} = \eta_{Nn}\epsilon_{nY} + \eta_{Nq}\epsilon_{qY} + \eta_{Np}\epsilon_{pY} + \eta_{Nk}\epsilon_{kY} + \eta_{NH}\epsilon_{HY} \tag{2.2}$$

The η represent the elasticities of nutrition with respect to various inputs in the production function. As Schiff and Valdes point out, the income elasticity of nutrition is probably positive since all the terms in the equation 2.2 are positive. However, it is an empirical matter whether ϵ_{NY} is greater than one. Equation 2.2 also highlights how inputs besides the nutrient intake affect ϵ_{NY}. As income increases, a household is likely to invest more in nutrition-enhancing private goods such as refrigeration, appliances that facilitate faster everyday preparation of meals (input p). Greater income is also likely to bring better public goods to a neighborhood – richer households get better water supply, relatively uninterrupted power, regular waste collection and sewerage services from their local governments (input k). With more income, households may seek to purchase foods of better quality, but this is by no means a guarantee. Households may choose "unwisely" – from a nutrition standpoint, due to either a lack of information or simply because their preferences run that way. This last point suggests that *a priori*, there is no way to sign ϵ_{qY}, even if η_{Nq} is expected to be positive.

It is clear however, that the income elasticity of nutrition will more often than not, diverge from the income elasticity of calories. One implication of this is that it would be a major simplification to regard the calorie content of a food or a food group as unchanging with respect to income. Suppose as a household's income rises, it purchases a higher quality of the same food but one that has lower calorie per unit. A household might also switch between foods in the same broad group, e.g. from coarse to finer cereals; coarse cereals typically hold higher calories per unit compared to the finer cereals like rice, but the latter are perceived as socially superior goods, owing to their widespread consumption among the richer classes. These substitution behaviors contradict the assumption of a constant food-to-calorie conversion. The income elasticity of calories is often smaller than the income elasticity of food expenditure, a result of the substitution of non-nutrient characteristics for calories. Under these circumstances, it is even more misleading to treat the income elasticity of nutrition synonymously with the income elasticity of calorie consumption. Nutrition need not suffer because of the substitution behaviors discussed above. This is especially true if we believe that households are rational and will not trade off adequate food intake simply for variety, flavor or status in food purchases.

As incomes expand, households have the opportunity to choose to spend a larger fraction of income on non-nutrient characteristics of food. Likewise, they have the opportunity to increase their expenditure on aspects of human capital, which would contribute to their long-term health. It is possible that households are willing to make short-term trade-offs for long-term benefits: higher calorie intakes may be sacrificed in order to afford other goods or make certain investments that households believe will lead to a permanently higher income and food consumption. The policies advocated by the World Bank ((Behrman and Deolalikar 1989, p.666)) routinely emphasized the importance of raising farm sector incomes to reduce the prevalence and depth of undernutrition. Hunger was seen as a manifestation of inadequate nutrient intake, brought on by low incomes. Policies to educate households on the importance of a balanced diet, or of maintaining minimum daily allowances were secondary to the policy of promoting employment opportunities and directly subsidizing the cost of staple foods for poor families. In a number of studies, it emerged however, that such policies were unlikely to succeed with raising nutrition. It was not just the fact that

Engel's law prediction of a shrinking food budget was borne out in the median and richer households in a sample, but that most households did not spend the marginal dollar entirely on calories, when they experienced an increase in income.

2.2 Estimation of the Income Elasticity of Calories

The income elasticity of calories have been estimated by two different methods in the literature. The first is known as an *indirect method*, whereby the income elasticity of food expenditure is estimated by aggregating the expenditures on several individual commodities and the calorie-income elasticity is derived subsequently from calorie conversion factors applied to the various food quantities. The second method is a *direct* method, where the income elasticity of calories is obtained from a reduced form demand function for calories, with income placed on the right-hand-side of that regression.

Authors who adopted the indirect system of obtaining nutrient elasticities include Strauss (1982) and Pitt (1983). Many more have approached the direct method of estimating the income elasticity. With no attempt to be exhaustive, the list includes Timmer and Alderman (1979), Behrman and Wolfe (1984), Pitt and Rosenzweig (1985), Bouis and Haddad (1992), Ravallion (1990), Subramanian and Deaton (1996), Gibson and Rozelle (2002), Skoufias (2003), Kochar (2005) and Jha, Gaiha, and Sharma (2009b). Table 2.1 provides a summary of the recent papers estimating the gradient of the income-calorie relationship. Behrman and Deolalikar (1987) estimate the elasticity using each method in turn and argue that the difference in the estimates obtained is largely owed to households substituting away from nutrients towards other aspects of food. More recently, a few authors have tried to study the intra-household allocation of food and the response of calories to change in the resources shifting from husband to wife or from an increase in a single member's earnings. These authors estimate the elasticity of calorie intake, not with respect to household resources, but with respect to the calorie consumption per member (see Mangyo (2008) and Shimokawa (2010)).

In general, the elasticity estimates drawn from the indirect method tend to be somewhat larger than the estimates yielded by the reduced-form calorie demand functions. Further, data and measurement of the principal variables – nutrients and household resources – have affected the size of the elasticity estimate. Broadly speaking, the data/measurement issue falls into three parts. These are:

1. Measurement of the dependent variable, i.e. calorie intake.
2. Measurement of economic status – income, expenditure, or an index of household wealth.
3. Measurement errors in the data on calories and household resources.

Measurement of the Dependent Variable

Datasets vary according to whether they permit the computation of calorie availability or intake. Often, surveys of consumption expenditures collect information on the purchases of various foods, the dollar amounts and the physical quantities. The calorie data are later constructed by applying calorie conversion factors to the quantities of different foods purchased by the household. The surveys separately record data on the number of meals consumed on various occasions - within and outside the household, including the number of meals served by a household to outsiders. This requires the calorie data to be adjusted for the calorie content of the meals obtained by a household from others and the calorie content of meals served by the household to others. Even with the adjustment, the calorie data are not regarded as intake. Thus, calorie availability regressed on income generates the income elasticity of these availabilities – its magnitude is usually greater than the income elasticity of intake (Strauss and Thomas 1995, p.1899).

The other method by which surveys have collected calorie data is by measuring the quantity of prepared meals as well as wastage in households every day for a continuous period. The Brazilian ENDEF survey collected data in seven daily visits to a sample of households where the enumerators not only noted the meals consumed in the last 24 hours, but also the number of household members and number of guests. The Indonesian budget surveys gathered details of food consumption over a recall period of seven days with a similar purpose. Yet, this manner of data collection is not only costly but is still subject to recall error; respondents are likely to overstate the food consumption over a shorter period than over a month. Their calculation of the calorie, protein and fat content of their meals may also be flawed. If the measurement of calorie intake is the objective of a survey, it would be superior to collect such data from perhaps the quantity and quality of the ingredients used in the preparation of meals, rather than details of meal portions. Two instances of surveys undertaking this challenge are the ICRISAT (this is a panel survey that began in the late seventies, interviewing 240 households spread across six villages in semi-arid south-central India) and the Bukidnon surveys in the Philippines. The recall periods in these surveys were the last twenty-four hours and intakes are computed for each individual within the household. Intake data have usually given small, sometimes statistically insignificant elasticity estimates. This can be seen from Table 2.1.

With regard to the variable used to represent the economic status of households, some surveys have measured only total expenditure to elicit the most reliable responses and minimize measurement error. The National Sample Surveys in India have done this traditionally. Other surveys have inferred household income, by imputing a market value to farm production and the productive inputs of the household such as the NCAER's surveys (Jha, Gaiha, and Sharma 2009a, p.985).

Measurement Errors

The problem of common measurement errors between the dependent variable (nutrients) and the main regressor (household income or expenditure) has been tackled in several ways in the literature. The problem arises from the fact that food purchase reports are the basis on which all the major and minor nutrient intakes are computed. Food expenditure is naturally correlated with total expenditure as well as with income. In the presence of positively correlated measurement errors, the marginal effect of income on nutrient consumption is exaggerated and the income elasticity of calories can be overestimated.

To avoid upward bias in estimation of the elasticities, the literature has resorted to the use of either instrumental variables or a panel data where fixed effects can at least wipe out the time-invariant source of the bias. Of these, the instrumental technique is probably preferable but finding an instrument that varies closely with income or total expenditure but not subject to the common measurement error problem is not easy. Subramanian and Deaton use the nonfood component of total expenditure both with and without household size as a control in their calorie-income regressions. They show that nonfood expenditure is able to provide a lower bound to the elasticity estimate and reduce the bias in the OLS estimate. Behrman and Deolalikar (1987) and later, Jha, Gaiha, and Sharma (2009b) use a number of instruments for income (or expenditure) such as the farm size, the square of farm size, the percentage of farm area under deep (high-fertility) soil, family size, proportions of the household that are adult males and females, age and schooling of the household head, total annual rainfall in the village of residence and various village and year dummies.

Instruments are preferred to fixed effect estimation because the latter remove variables that affect nutrient intakes but do not vary over time within the units of observation. In the context of the nutrition production function, variables that fixed effects estimation could drop potentially include publicly-provided inputs. In the nutrient demand functions, community endowments, locational and even some demographic characteristics of households can be erased from consideration in a differenced regression equation. Further, fixed effects only eliminate a part of the bias. In the most well-known instance of a study combining both instruments and fixed effects to estimate calorie elasticities, Behrman and Deolalikar find that the elasticity is statistically insignificant. However, one of the limitations of their estimate lay in the small sample size from which they procure it. As Subramanian and Deaton observe, precision is crucial because the standard error of Behrman and Deolalikar's (1987) estimate is 0.37. A larger sample could afford the possibility of estimating the elasticity, at each quartile of the income distribution, if not for each unit of observation in the sample.

2.3 Nonparametric Estimation Methods

Theoretically, there is a reason to expect the elasticity of calories to vary by income level. Engel's law predicts that the share of food in the household budget declines as the households grow richer. Physically, there is an upper limit to the amount of food consumption an individual requires and desires. In fact, we see households substitute taste, appearance, variety and status for calorie content. Calorie needs decline as households shift to more sedentary jobs, which they do when they tend to have more education and thus, more income. For all of these reasons, the calorie response to an increase in income is unlikely to be uniform across the income distribution. If the income elasticity of calorie intake is itself a function of income, then the simplest regression specification capable of capturing this relationship is perhaps a quadratic term in household resources. However, in the early attempts to model nonlinearities in this relationship - such as regressing logarithm of calories on the inverse of the logarithm of total expenditure and its quadratic - did not always succeed. Nonparametric methods were all the more important for being able to offer insight into the shape of the calorie-income curves.

Poleman (1981) and Lipton (1983) believed that the calorie-income curve could actually be elbow-shaped. They go further to suggest that the budget share of food may increase initially when there is an increase in the income of the poorest households. In this phase, the calorie response is unlikely to be a diminishing since nutrient intakes would override all other wants of the "ultra-poor", as Lipton refers to them. In the last two decades, investigations into the income elasticity of calories have taken a rigorous non-parametric approach. (Subramanian and Deaton 1996) demonstrated the power of this technique with the 1983 NSS data for India (specifically, the state of Maharashtra). They use a nearest-neighborhood estimator based on a rectangular kernel function and graph the nonparametric regressions of the logarithms of calorie availability and calorie price against log monthly household outlay. They find that the elasticity estimates are significant and decline gradually from about 0.5 for the poorest decile to about 0.3 for the richest. Following them, several others have implemented nonparametric regressions.

Roy (2001) uses the local linear regressions and finds that the calorie-income curve is upward-sloping for the most part, implying a non-zero elasticity. Her most sophisticated formulation involves a semiparametric regression model with random effects which estimates a mean elasticity of 0.11 and ranges from 0.14 in the lower tail of the income distribution to - 0.03 for the richest individuals. Roy also studies the elasticities by gender and reports that females have a higher income elasticity of calories than males. Gibson and Rozelle also use nonparametric LOWESS graphs of the calorie-income relationship in urban Papua New

Guinea. They find an unconditional calorie elasticity estimate of approximately 0.6 for the poorest half of the population, most of whom cannot access the recommended daily calories per capita. Gibson and Rozelle approximate the nonlinearities in the calorie-expenditure relationship by using expenditure splines when they extend the regression analysis to include other covariates of nutrient intake. Nonparametric estimation was also employed by Abdulai and Aubert (2004) for Tanzania. They find a strongly significant and positive relationship between household expenditure and calorie demand[2], which they interpret in favor of conventional wisdom that income growth alleviates hunger. They also find a negative and significant relationship between food prices and calorie demand, which they regard as sufficient basis for targeted food subsidies.

2.4 Household size and Composition

While nonparametric techniques are better able to describe the curvature of the calorie-income regression, they are hindered by the massive demands they make on data size and computing power. Nevertheless, one of the most important sources of variation in elasticities is household size and composition – a variable whose impact can be gleaned even from the bivariate nonparametric calorie regressions.

Subramanian and Deaton illustrate the gradient of the calorie-income relationships drawn after grouping households by size. They find that household size reduces their elasticity estimate to 0.35 from 0.40. This is explained by the negative correlation between

[2]The basic estimate is about 0.60 and this declines to 0.40 when lagged expenditures instrument for current outlay combined with either first-differences or fixed effects in the estimation process

household size and per capita outlay, so that when household size is an omitted variable in the regression of per capita calories on per capita outlay, the slope coefficient is biased upwards. In general, the calorie Engel curve for the smaller households lie above those of larger households.

Strauss and Thomas (1995) claim that in the Brazilian ENDEF data, introducing demographics into the nonparametric calorie-expenditure regression produces more or less the same curve that did not control for any household-level variables. Of course, controlling for demographics is not straightforward. It can be argued that household composition is endogenous. Calorie needs vary by age, sex and occupation and to the extent that households align calorie intakes with needs, the omission of demographics would bias the elasticities.[3]

2.5 Broader Concerns: Norms and Nutrition Education

As mentioned before, substitution away from nutrient intake is inevitable as households grow richer over time. The problem with this behavior – from a normative standpoint – is when it happens "too early", i.e. households that are still below the nutrition-poverty line spend a larger fraction of an income increase on non-nutritive aspects of food. However, before we question the wisdom of rational households, a number of other issues present

[3]Supposing as is often the case, that higher incomes are negatively correlated with household size but positively correlated with per capita calorie consumption. Then among households of the same income level, the larger ones will have lower per capita calorie intake because the proportion of children in them is greater and the calorie consumption of children is much less than that of adults. Consequently, regressing per capita calories directly on expenditure but excluding household size would inflate the income elasticity of calories. In a long panel, this omission would lead to a much larger upward bias, as fertility rates decline in response to a rising per capita income.

themselves. Among them is the nutrition poverty line itself. The recommended amount of daily calories per capita has not evolved with changing food quality, diet patterns, work habits and health advances in most countries. There is also some disagreement over both the actual calorie requirements per person as well as debate over the capability of some indicators to truly capture the extent of undernutrition in a society. As such, some innovations in the way we measure undernutrition may be in order.

Gopalan (1983) favored that a yardstick that measured undernutrition using official calorie norms but warned that sole reliance on that one measure would be misleading. He suggested supplementing that indicator with diet surveys, clinical and anthropometric assessments, selective biochemical evaluations such as hemoglobin levels and a monitoring of improvement in environmental standards. Deaton and Drèze (2009) echo the same sentiment, believing that no there is no evidence of a tight link - in the Indian context - between incomes and calorie consumption, and between calorie consumption (n) and nutrition, (N), or health. They too, suggest closer nutrition monitoring to ensure that data from different agencies are not giving inconsistent reports on the undernutrition rates.

More recently, Jensen and Miller (2010b) proposed a method based on revealed preferences to estimate the rate of undernutrition in a society. They argue that the share of staple calories in a person's diet would decline sharply when the subsistence (minimum nutrition) threshold is passed. This threshold would very likely vary by the individual's age, sex, location, occupation, as well as health circumstances (s)he may have been exposed to. Yet, it could still be discovered by identifying the level of household expenditure at which the share of staple calories declined dramatically. Jensen and Miller attempt to do exactly this. They construct a model of a utility-maximizing consumer and apply it to panel data from the China Health and Nutrition Survey where they believe the economic gains in nutrition are better captured by their method than by traditional poverty lines. Concluding, they make their point as follows:

> Policy makers should perhaps not care about whether someone meets a calorie threshold, which can't be determined precisely anyway, but instead getting the consumer to the point where the marginal utility of additional calories is revealed to be low, suggesting they are not a priority for the consumer, and thus should not be for the policy maker (and since policies promoting increased caloric will not be very effective at that point anyway).

The dilemma of policymakers between advocating good nutrition and getting the poorest of households at least subsistence nutrition is an old one. Certainly, Jensen and Miller make a very valid point given the non-monotonic response of calories to wealth and income. It is efficient to devote resources where they achieve the maximum good for society. Hence, it is useful to know which households actually do care about increasing their nutrient intake foremost before targeting them with employment programs or food subsidies. On the matter of how best the marginal dollars can be spent, households are likely to make a better decision (i.e. a decision that benefits its most vulnerable members or increases the average health level in the household) when they have more information. The information could pertain to the cheapest ways to secure the minimum nutrient intakes, the advantages of a balanced diet, fresh foods, and the timeliness of nutrition – specifically, how good nutrition in infancy and early childhood help proximate outcomes like schooling attainment, but also elevate long-term health prospects.

Table 2.1: **Elasticities of Calorie Demand - Recent Previous Literature**

Author(s) (Year)	Country	Unit of Study	Measure of Calories	Household Resources	Estimation Method	Preferred Elasticity Estimate
I. Elasticities of Calorie Demand with Respect to Income or Expenditure						
Behrman and Deolalikar (1987)	India	Individual	Intake	X	2SLS	0.17*
Behrman and Deolalikar (1987)	India	Individual	Intake	X	2SLS, FD	0.37*
Behrman and Deolalikar (1990)	India	Individual	Intake (24-hr recall)	Y	OLS, FE	0 to 0.07*
Bouis (1994)	Kenya	Family	Intake (24-hr recall)	X	2SLS	0.14 to 0.17*
Bouis (1994)	Philippines	Household	Availability (unadjusted)	X	2SLS	0.52
Jha, Gaiha, and Sharma (2009b)	India	Households	Intake	X	2SLS	0.07
Subramanian and Deaton (1996)	India	Household	Availability	X	2SLS, NP	0.30 to 0.50
Dawson and Tiffin (1998)	India	Country	Intake	Y	FD, ML	0.34
Tiffin and Dawson (2002)	Zimbabwe	Country	Intake	Y	FD, ML	0.31
Gibson and Rozelle (2002)	Papua New Guinea	Individual	Availability	X	2SLS, NP	0.60
Skoufias (2003)	Indonesia	Household	Availability	Y	FE, NP	0.32 to 0.45
Aromolaran (2004)	Nigeria	Individual	Intake (48-hr recall)	X	2SLS	0.03 to 0.04*
Kochar (2005)	India	Household	Intake (30-day recall)	X	2SLS	0.24
Abdulai and Aubert (2004)	Tanzania	Household	Availability	X	NP, 2SLS, FE	0.49 to 0.62
II. Elasticities of Calorie Demand with Respect to Calorie per Household Member						
Mangyo (2008)	China	Individual	Intake	n/a	FD, 2SLS	0.7 to 1.1 †
Shimokawa (2010)	China	Individual	Intake	n/a	FD, 2SLS	– †

Notes: * indicates a statistically *insignificant* estimate at the 5% level. Under the column, household resources, the abbreviations X and Y stand for expenditure and income of the unit of study respectively. Under estimation method, the abbreviations expand as follows: 2SLS: two-stage least-squares, FE = fixed effects, FD = first-differences, ML = maximum likelihood. Both Behrman and Deolalikar (1990) and Dawson and Tiffin (1998) also calculate elasticities of calorie demand with respect to food prices in their work while Kochar (2005) reports an elasticity of calories with respect to food price subsidies of about 0.06. The estimate from Gibson and Rozelle pertains to the poorer half of the urban households in their sample. The estimate from Behrman and Deolalikar (1990) is taken from the row reporting the implied household average elasticity. The estimates of Skoufias (2003) are calculated at the 25th percentile of the household total outlay distribution from 1996, and the lower estimate is for urban, the higher for rural. † The elasticities vary by demographic. There is no single preferred estimate.

Chapter 3

The Income-Calorie Relationship in Rural Maharashtra

3.1 Introduction

Undernutrition has been a long-standing concern in poor countries, with economists and policy-makers looking for ways to reduce its severity and prevalence. For decades, developing countries have pursued single-minded policies of higher economic growth, trusting to the income-response of calories to solve undernutrition[1]. However, India has proven a paradox to this strategy. Over the better part of two decades, increasing rates of economic growth have been accompanied by a gradual downward drift in the calorie Engel curve.

The puzzle of declining calorie-intakes in India has been examined frequently in the development literature; recent discussions include Deaton and Drèze (2009), Patnaik

[1] Although insufficient calorie intake cannot be equated to undernutrition, the idea is that it is easiest to manipulate nutritional status through calorie intake.

(2004, 2007), Meenakshi and Viswanathan (2005) and Palmer-Jones and Sen (2001). Much of the research on calorie consumption has adopted the route of estimating the income (or expenditure) elasticity of calories, in order to determine if higher incomes led to greater food consumption. Belief in a positive and significant income-elasticity was the basis of many employment and nutrition programs in India. However, investigating this belief with consumption survey data has laid out mixed evidence. Some authors such as Subramanian and Deaton (1996) and Jha, Gaiha, and Sharma (2009b) find a significant and positive elasticity, while others, mainly Behrman and Deolalikar (1987, 1989) do not. This paper follows the practice of estimating income elasticities for total calories and the price paid to obtain them, from a cross-section of households in rural Maharashtra. In Maharashtra, rapid growth has failed to lower poverty and undernutrition commensurately, as in most other states of India. Incidentally, it is also a state where past nutrition studies have occurred, so that a new set of results would afford interesting comparisons.

Maharashtra has been one of the best-performing states in India, growing at a rate of 4.3% since 1960-61 until 1990-91, and after the liberalization reforms went into effect, at a rate of 6.5% in the nineties (Reserve Bank of India 1999). Yet, according to the report of the Planning Commission (1993, 2004), the poverty rate in the entire state of Maharashtra was no less than a quarter of the total population as of 2004. In fact, the decline in poverty has been slow and small, falling from 58% in 1973 to 45% in 1983 to about 25% in 2004-05, a rate still higher than the national average. The poverty rate varies with the geography of the state: western and coastal areas are less poor than the interior eastern parts.

In untangling the income-calorie nexus, one of the difficulties of establishing a causal relationship arises in the form of the efficiency-wage hypothesis (Stiglitz 1976). It is always possible that causation might travel from nutrition to income. Richer and better-nourished households would possess higher marginal products of labor that lead to greater incomes, more so than food-deprived, poorer households are capable of. Thus, poor households may be caught in a poverty trap, not because their marginal propensity to spend on calories is low but because they are unable to generate a large enough income to demonstrate a significant income elasticity of calories. The efficiency-wage hypothesis is difficult to rule out: a good instrument for calories is not easily available and simultaneous equations techniques are not reliable.

However, a look at the rural daily wage of the agricultural labor around the time of the survey suggests that nutrition is unlikely to have held back the earning capacity of households. The rural daily wage in 2004 for casual agricultural labor (i.e. manual work in cultivation or other agricultural activities) in Maharashtra was anywhere between 35 and 40 Rupees for women and between 46 and 54 Rupees for men, (about 9 to 15 Rupees in 1983 prices). These wages are at least 7 times as much as the price of a thousand calories, so the cost of obtaining 1810 calories per day[2] would require at most 21% of the rural daily wage. Although this is a larger percentage of the wage than reported for rural Maharashtra of 1983 by Subramanian and Deaton (1996), it indicates that minimum nutrition standards are well within the reach of the poorest households. To that extent, it is more likely that nutrition is constrained by income than the converse.

[2] 1810 calories per day per capita is the FAO norm for India. More background on the norm is provided in the Section 3.2.

This paper finds that the calorie Engel curve in Maharashtra has drifted downward since 1983, and the income response of calories has shrunk considerably, falling by roughly half throughout the expenditure distribution. Where Subramanian and Deaton find that the outlay elasticity of calories ranged between 0.5 and 0.3 in 1983, in 2004, the elasticity is only 0.12 at the mean, although the difference in the calorie elasticities between the poorest and richest deciles is about the same as in 1983. Through nonparametric LOWESS algorithms and local linear regressions, it emerges that the outlay elasticity for calorie purchases in the poorest decile is 0.2 and it ia significant. The richest decile, however, produces an elasticity not statistically different from zero. The elasticity of the price paid per calorie has fallen since 1983 too, but the decline is of smaller magnitude. It ranges from 0.4 at the lowest decile to -0.1 at the top end of the expenditure distribution. The bivariate relationship between calories and income is explored further with spline regressions which find that the elasticity estimates are robust to the introduction of controls, such as the size and composition of households as well as district fixed effects.

3.2 Data

The data in this study comes from the sixty-first round of the National Sample Survey in India, conducted from July 2004 to June 2005. The survey follows a multi-stage, stratified sampling design where the first stage units in rural areas are census villages and the ultimate-stage units are the households. The data for rural Maharashtra is a sample of 504 villages spread across 35 districts (the relevant number of districts is actually 34 since the district of Mumbai is considered to be fully urbanized and thus does not figure in the rural sample). Ten households are interviewed from each hamlet-group within a census village. The number of hamlet groups is determined by the population size of the village. Therefore, heavily populated villages contribute more households to the survey data. The survey typically gathers detailed information on the purchases of various goods - from both food and non-food categories - over a well-defined recall period. The food data in the sixty-first round was collected on the basis of 30-day recalls, a reference period used regularly in past Consumption Expenditure surveys of the NSSO. As food items tend to be purchased frequently, this recall period is preferred to a 7-day period as it is likely to yield less-exaggerated estimates of the quantities and expenditure values of different foods. In addition, the survey records the demographic characteristics of households, as well as the educational attainment and occupational status (primary livelihood) of its members.

The data of primary interest is the food consumption of households. Foods may be procured by households in one or more of five ways: by direct purchase, from own stock of production, as gifts or exchanges in kind, as free collection and in the form of ready-made meals consumed outside the home. Household members might receive meals at work, as part or full compensation for labor services rendered; they may also receive meals on special occasions (like local festivities) or as guests at another household. Of course, most households also serve meals to non-members (i.e. employees or guests). The National Sample Survey records the number of such meals received and given by each household. This meal adjustment to the total food purchases of a household is crucial since the outcome of interest in calorie availability, if not intake.

The survey breaks down the food data into about 148 individual items. The quantities of these foods is converted into calories using the conversion factors suggested in National Sample Survey Organisation (2007, p.18). The total calories arrived at by summing the calorie content of all foods purchased by the household is the *unadjusted* calorie availability. I next adjust for the calorie content of meals received and given by the household over the same period. Four different types of meal-variables are constructed in the data: m_1 representing the number of meals consumed at home, m_2 representing the number of meals received as gifts, exchanges or for free, m_3 representing the number of meals received from employers, and m_4 representing the number of meals given to non-household members. The meal calorie adjustment is made following the procedure recommended in National Sample Survey Organisation (2007):

$$\text{Meal Calories} = [m_1 + m_2 + m_3]/[m_1 + m_4] * 1200$$

where 1200 is assumed to be the mean calorie content of prepared meals; the ratio preceding 1200 in the formula above is typically greater than one for poor households as they tend to net recipients of meals but smaller than one for the relatively better-off.

Thus, adjusted total calories = unadjusted calories + meal calories.

It turns out that for about 34% of households in the data, this type of meal adjustment lowers their total calorie availability while the remainder registers an increase in calorie availability, having received more meals than they gave to others.

The calorie data in this paper will be regarded as calorie availability rather than intake. In general, intake differs from availability due to wastage, the number of meals served to guests or employees of a household and consumption-smoothing choices. To begin with, households need not consume all the food that they actually purchase. Wastage can drive a wedge between availability and intake, although households that already suffer from an energy deficiency are less likely than rich households to incur large wastage. Intake could differ from availability also because a fraction of a household's food purchases is set aside for catering to employees, servants or guests. Both wastage as well as the practice of serving meals to non-household members are more frequent and typical of rich households. Hence, the gap between the calorie intake and availability is expected to be large for rich households, but relatively small for poor households.

Calorie availability (and intake) are both likely to vary substantially by household size and composition. Although the data available do not permit an analysis of the intra-household allocation of calories and perhaps, age- or sex-based discrimination, an attempt is made to control for the extent to which the average person in a household meets the requisite daily norm. By assigning an adult male in the age group of 20-39 years, one "consumer unit", a measurement designed to represent his daily energy requirement, the requirements of all other individuals - male and female, of other age-brackets - are computed. Thus, summed consumer units within a household stand for the number of equivalent adult males, 20-39 years of age, present in that household and help control for the diverse energy requirements of children and adults.

Calories, Prices and Food Shares

The total sample for rural Maharashtra contains 5023 individual households. However, some of the values reported for per capita daily calorie consumption are improbable, so I choose to keep values in the distribution that range from a minimum of 1000 to a maximum of 4200 calories. In addition, I choose to keep households with a monthly per capita expenditure in the range of 25 to 7500 Rupees. This reduces the sample to 4840 households.

Table 3.1 provides a first look at expenditures and calorie availabilities post-meal adjustment by each decile of monthly per capita expenditure. Columns [2] and [3] list the real per capita expenditure of households and the budget share of food respectively. The relationship between these variables is straightforward: the budget share of food falls as real expenditures rise. The poorest of households spend as much as four-fifths of their total budget on food. This fraction declines to 58 percent in the median class and is as high as 45 percent in the eighth decile. Only the richest decile devotes less than a third of its budget to food. The food expenditure data used to calculate the budget share of food includes the expense on meals eaten outside the home but excludes the approximate cost of meals prepared by the household and served to guests or employees.

The last three columns tabulate the calorie availability adjusted for meals by expenditure decile and then, the percentage of two different calorie norms the availability data satisfy. The first calorie norm is an average daily intake of 2400 calories per rural person, which has been in use in most of the literature since being framed by the Indian Planning

Commission in its Report of 1979.[3] Although this norm has not been updated officially by the Government of India, different organizations have been using alternative benchmarks for some time now. For instance, the FAO placed the calorie norm for South Asia at 2110 daily calories per head and for India in particular, at 1810 per capita daily calories (Bajpai, Sachs, and Volavka 2005). In Column [6] of Table 3.1, the availability data are compared against the norm of 1810 calories.

The lowered norms are a recognition of lifestyle and public health changes since the seventies: work in rural India today is less demanding in terms of the manual effort expended compared to thirty years ago; further, disease burdens have declined and transport facilities have improved. Therefore, imposing the daily calorie requirement of 2400 per capita could imply a higher rate of energy deficiency than is sensible. In fact, in Table 3.1, no decile meets the norm of 2400 calories per capita per day and the calorie availability of the poorest decile is only 70 percent of the requirement. On the other hand, all but the bottom-most decile satisfy the FAO norm. It should also be noted that the mean calorie availability is only 79 percent of the official norm, so the average rural household in Maharashtra is more likely to be energy deficient.

Table 3.2 lists the summary statistics for the main variables of interest. The mean per capita daily calorie availability is 1990 calories, which is not very different from the

[3]In the late 1970s in India, the Planning Commission constituted a Task Force on Projection of Minimum Needs and Effective Consumption Demand which, on the basis of a systematic study of nutritional requirements, recommended a rural norm of 2,435 calories daily per capita and an urban minimum of 2,095. Rural areas were thought to need more calories because it was believed that a greater proportion of the population would be engaged in manual labor. The calorie norms were built by weighting the age, sex and activity-specific calorie allowances by the population proportions of each age-sex-occupation category. The weights were estimated using demographic data from the 1971 Census and participation rates across occupations from the Employment Survey of the 27th NSS Round (Planning Commission 1979).

unadjusted mean calorie availability per person per day. The mean monthly per capita expenditure on all goods is 642 Rupees; in 1983 terms, this translates to 165 Rupees. The mean monthly food expenditure is 325 Rupees. Food forms about 56% of the monthly budget on an average. In addition, the mean price paid for 1000 calories is a little over 5 Rupees, or about 1.38 Rupees in 1983 values.The average household has nearly 5 persons but the mean number of consumer units is 3.9, i.e. the calorie requirement of the mean household is approximately that of a four-person, adult-male household. Bear in mind that fertility rates have been falling over time, so that the proportion of adults in a given size of household in 2004 would be greater than the proportion found in a household of the same size in 1983. Other things being equal, lower fertility must imply that for any given household size, calorie requirements in 2004 are higher than twenty years ago. It may be preferable to measure calorie availabilities per adult equivalent rather than per capita to capture the rate of calorie deficiency and impoverishment. Still, working with calorie availabilities per capita affords easier comparison to other studies, even if such data tends to understate shortfalls from the calorie norm, than availability per adult equivalent will do.

Returning to Table 3.1, it is worth noting the differences between the bottom and top deciles: while the poorest 10% of households incur a total monthly expenditure of 66 Rupees, the top 10% spends five-and-a-half times that amount. Yet food is only 30 percent of the total budget in the top decile, while the poorest households allot 83 percent. Finally, the top and bottom deciles are only 322 calories apart in per capita daily availabilities. This suggests a substantial difference in the price each decile pays for its calories, but it turns out that the poorest and richest households are almost the same on average. Thus, the actual price per thousand calories is Rupees 6.32 among the richest and Rupees 4.16 among the poorest. This difference is large enough to be important later when the downward drift of the Engel calorie curves are analyzed.[4]

Thus, one of the main differences between the poorest and richest households in Maharashtra in 2004 lies in their budget shares of food. The other lies in their calorie availabilities: measured against the norm, the shortfall of the poorest decile is considerably larger than that of the top decile. Part of the shortfall can be attributed to the downward drift of the calorie Engel curve since twenty years ago. Reasons for the drift are discussed in a later section, and it is likely that a combination of variables are at play, including improvements in public health, in the quality of infrastructure, the degree of physical demands of current-day rural labor, and households substituting non-nutrient characteristics of foods (like taste or "status") for the calorie-content.

Table 3.3 examines the behavior of the poorest and richest deciles with regard to the manner in which they dispense their food budget across broadly defined commodity groups. Of these, cereals and cereal substitutes are the main source of calories, with oils and fats being the other important source. The remainder of the food groups contribute proportionately less to total calories and more to protein and essential micronutrients. The figures appearing in the column titled "Mean" under "Expenditure Shares" were calculated by averaging the expenditure share of a food group over all households in the sample. The corresponding figures in columns called "Bottom 10%" and "Top 10%" were calculated

[4] In their study of Maharashtra using NSS data from 1983, Subramanian and Deaton (1996) state that the bottom decile paid about 88 paise (.88 Rupees) for a thousand calories and the top paid Rupees 1.50. Adjusting for inflation, the prices calculated in 2004 translate to Rupees 1.07 and Rupees 1.62 respectively.

similarly, but using households in the first and last deciles respectively. Calorie shares of the food groups are likewise averages over households in the appropriate samples. The price paid per calorie is stated in Rupees per one thousand calories. The price data for 2004/05 have been deflated using the Consumer Price Index for Agricultural Labourers (CPI-AL), so all the prices in Table 3.3 are in 1983 values and can be compared readily against the values in Table 1 of Subramanian and Deaton (1996).

The cereals food group lumps staple cereals such as rice, wheat, their derived products, with coarser cereals, like millet, sorghum, and maize, with cereal substitutes like tapioca and jackfruit. This group accounts for 30% of the total outlay on food at the mean, and for 32% and 26% respectively in the bottom and top deciles. Ignoring the "other foods" category, which is an assortment of goods such as salt, spices, beverages and intoxicants, fruits and vegetables claim the next largest share in the mean budget, with oils and fats, then dairy following. This pattern is repeated for the bottom and top deciles, although for the richest households, dairy switches position with oils and fats in the list of largest budget shares.

Compared to the pattern of disposition in 1983 [5], oils and fats now take precedence over pulses in expenditures shares. Pulses have diminished in expenditure shares across deciles. Partly, this trend is owed to the change in the calorie price of these food groups. Calories from oils and fats as well as calories from dairy cost less in 2004/05 than in 1983. The top decile pays a marginally higher price for dairy in 2004 than in 1983, but not enough to reduce the share of dairy in total expenditure. In contrast, pulses have turned out to be

[5] See columns 1 through 3 in Table 1, p.140 in Subramanian and Deaton (1996).

more expensive in 2004, eliciting substitution away towards cheaper food categories.

The next three columns in Table 3.3 show the calorie shares of different food groups. Calorie shares are fairly similar across deciles for most food groups. Cereals are the only group where the difference between the richest and poorest households is larger than 10%. While cereals still contribute 57% of total calorie consumption in the top decile (no change since 1983), they make up 70% of the bottom decile's calorie purchases (a drop from 77% in 1983). On an average, cereals generate 64% of the mean household's total calories, a fall from 71% in 1983. This gap is filled by the expanded shares of oils and fats, dairy and fruits and vegetables. Compared over time, oils and fats show the largest increase in calorie shares, almost doubling for the poorest households.

The average cost of a thousand calories was 1.14 Rupees in 1983 and it rose to 1.30 Rupees in 2004. However, this masks the widely different changes in the cost of group calories. The final three columns of Table 3.3 set out the price per calorie by commodity group and year. Numbers in the first row represent the cost of one thousand calories from each commodity group in 2004; numbers in the second row show the corresponding costs in 1983. It is clear that for any given commodity group, the price per calorie in 2004 does not vary much between the top and bottom deciles: the latter uniformly pays a little less in all categories except sugar. However, two important observations emerge regarding the price per calorie in 2004 versus 1983: one, oils and fats are the only food group for which the real cost of calories has *declined* over the period; two, cereals calories are four to five times more expensive in 2004 than in 1983. A thousand cereal calories cost 3.77 Rupees at the top decile, falling to 3.46 Rupees for the bottom decile. The average cost was 3.59 Rupees in 2004, compared to 64 paise in 1983.[6]

In the next section, the elasticity estimates for both calorie availability and the price per calorie are presented and discussed. Preliminary to that, it is helpful to invoke the expressions presented by Subramanian and Deaton (1996). If the calorie content of each good is denoted as k_{Gi} and the quantity of good i that is bought is called q_{Gi} then the total calorie availability, ignoring meals, is

$$c = \sum_{G} \left[\sum_{i \in G} q_{Gi} k_{Gi} \right] \tag{3.1}$$

The expenditure elasticity of total calories can be derived as a weighted average of expenditure elasticities of expenditures on different foods with the calorie shares of the food in the group acting as weights. Suppose x denotes total monthly per capita expenditure.

$$\epsilon_{cx} = \sum_{G} \eta_G \sigma_G \left[\sum_{i \in G} \eta_{Gi} \omega_{Gi} \right], \tag{3.2}$$

where η_G is the total expenditure elasticity of the expenditure on food group G, σ_G is the calorie share of group G in c, η_{Gi} is the elasticity of expenditure on good i in group G with respect to the expenditure on that food group G, and ω_{Gi} is the calorie share of good i in group G.

The expenditure elasticity of calories can also be related to the the expenditure elasticity of the price paid per calorie. Suppose the expenditure elasticity of the calorie price of a food group is denoted by ξ_G. Then the expenditure elasticity of calories is simply the

[6] 100 paise is equal to 1 Rupee.

difference between the expenditure elasticities of total food expenditure and group prices per calorie. Summed over all groups, the calorie elasticity is again a weighted average:

$$\epsilon_{cx} = \sum_G \sigma_G \Big[\eta_G - \xi_G \Big] \tag{3.3}$$

The total expenditure elasiticity of food expenditure is

$$\epsilon_{xx} = \sum_G \omega_G \eta_G, \tag{3.4}$$

where ω_G stands for the budget share of group G and η_G is again the expenditure elasticity of the expenditure on food group G.

Note that the richest households in the sample were found to pay slightly more per calorie than was paid by the poorest households for nearly all calorie sources (the exception was sugar). It would appear that the expenditure elasticity of price per calorie is positive. If so, the expenditure elasticity of calories would be smaller than the expenditure elasticity of food spending. If a higher price per calorie occurs in richer households, then the expenditure elasticity of the price per calorie would drive a wedge between the elasticities of total calories and total food. Both inter- and intra-group substitution of foods would contribute to this result.

Suppose for now that intra-group substitution is ignored. Substitution between commodity groups still implies a positive expenditure elasticity of price if the groups whose calorie shares are now increasing in the household's choice set are (uniformly) more expensive than the groups declining in calorie shares. It is not necessary to have positive expenditure elasticities for the group calorie price. It is possible for this elasticity to be zero or even negative. A zero price elasticity is likely in the absence of intra-group substitution. A negative price elasticity could arise if richer households actually paid less than the poorer ones consistently for all food groups.

Intra-group substitution is assumed to be relatively small in this study. Substitution between food groups is examined further below. Between 1983 and 2004, the calorie share of cereals declined nonlinearly through the real expenditure distribution. The shares of oils and fats rose the most (from 4.8 to 9.1 percent in the bottom decile and from 7.6 to 12 percent at the top), followed by dairy (2.8 to 4.6 percent for the mean household) and fruits and vegetables (from 3.5 to 5.7 percent at the mean). Households have generally diversified their consumption of calories, substituting fats and foods not particularly rich in calories for cereals. The inter-group substitution may have been caused in part by the change in unit prices of the broad food groups since 1983. Although the real price per calorie is not remarkably different in the cross-section of households in 2004, every category of food has grown more expensive, with the exception of oils and fats. Among major calories sources, cereals are the second-most expensive after meat and eggs. Both the bottom and top deciles pay more for cereal calories cost than for fat calories, and about thrice as much for meat, vegetable and fruit calories as for cereals.

One category where the unit price rose but calorie and budget shares either declined or remained unchanged is meat. At any given expenditure class, meat does form a very small percentage of the total budget and calorie availability. This does not undo the general intergroup substitution towards cheaper calories; households may well trade-off a higher price for some diversity in their diet. Although the factors motivating inter-group substitution may be varied, it is a useful simplification to regard it as the only significant type of substitution in the food choices of households. If so, based on the previous formulas for the calorie elasticity, the estimates for ϵ_{cx} and ϵ_{xx} are 0.118 and 0.332. The next section looks into the elasticities more rigorously, using nonparametric methods to characterize the regression function and the gradient at different levels of monthly per capita expenditures.

3.3 Expenditure Elasticities of Calories and Prices

As past studies have shown (see Subramanian and Deaton (1996), Roy (2001), Gibson and Rozelle (2002) and Skoufias (2003)), the elasticities of calorie intake and prices are not usually identical across the expenditure distribution. Poor households behave differently from rich households, having to make food choices dictated more by survival considerations and less by taste or status. Thus, not only is the level and pattern of food expenditure different but so is the marginal calorie response to increased income.

I begin by exploring the nonparametric joint density of the logarithm of per capita daily calories and the logarithm of monthly per capita expenditure (LMPCE) in Figure 3.1. This joint density was estimated using a bandwidth selected by the maximum likelihood cross-validation procedure with a Gaussian kernel over the sample of 4840 observations. The contours of the joint distribution in Figure 3.2 suggest that it is approximately normal. Figure 3.3 shows the unconditional density of the logarithm of per capita daily calories. There is a slight suggestion of skewness to the right in the diagram, which is to be expected given that the bulk of the households in rural Maharashtra do in fact consume less than the median. Most also purchase fewer calories than the norm of 2400 calories per head per day (the calorie norm falls at about 7.78 on the log scale).

I ran a parametric regression of the logarithm of per capita daily calories on the logarithm of expenditures to compare how well it approximated the true relationship as revealed by a nonparametric model. I favor a quadratic specification in the logarithm of monthly per capita total outlay for two reasons: first, the logarithms of both calories and expenditures substantially reduces the nonlinearity from the relationship; secondly, the quadratic specification affords some flexibility with respect to evaluating the elasticities at each point in the household expenditure distribution. Figure 3.4 draws the fitted regressions of the quadratic model in red broken lines and the nonparametric model in the solid blue curve.

The two models produce similar fit, the quadratic largely agreeing with the nonparametric regression. The difference occurs at the higher levels of expenditure, where the nonparametric fit is somewhat flatter. The resemblance between the fitted models cannot be due to oversmoothing in the local linear regression. Lower bandwidths were experimented with, including window-widths half as small as the one which produced Figure 3.4 (where the optimal window-width was 0.336) and while the basic resemblance persisted, the variance increased. It is worth noting that even though both calories and expenditures were cast in logarithms in the regressions, the predicted regression lines are rather nonlinear. A look at the gradients associated with the two models suggests more palpable differences. The gradients are shown in Figure 3.5. The slopes from the parametric model clearly overstate the elasticity for poorer households and understate them for richer households. It predicts an average elasticity of almost 0.15 at the mean, when the level of real household expenditure is 146 Rupees per month per person. The most probable explanation for the difference in the gradients of the two models, despite the likeness of the calorie-expenditure relationships estimated before, is the nonlinearity detected in Figure 3.4. A similar behavior occurs in Gibson and Rozelle (2002, p.34) in their study of the income-calorie relationship in urban Papua New Guinea.

Figure 3.6 and Figure 3.7 present the 95% confidence bands for the nonparametric local linear regression regression and the gradients therefrom. Figure 3.6 presents the local linear fit from a bandwidth of 0.336 with a second-order Gaussian kernel with standard errors bootstrapped to construct the 95% confidence interval. The bootstrap procedure does not take into account cluster effects; however, they do correct for possible heteroscedasticity in the errors. Cluster effects have usually been found to be rather small. The bootstrapping procedure was performed with 50, 100 and 200 replications using alternatively all and half the total observations in the sample. The confidence bands from the standard errors obtained in all these procedures were near-identical (Efron and Tibshirani (1993) advise no more than 200 replications to estimate standard errors, unless one is bootstrapping the confidence intervals themselves). The bands are fairly tight around the middle of the regression and only widen at the extreme values. Overall, the local linear regression estimates the relationship with relative precision.

Figure 3.7 graphs the expenditure elasticities of calories with bootstrapped standard errors. The confidence interval is approximately $[0.05, 0.25]$ and the mean elasticity estimate is 0.15. The gradients are not estimated precisely except perhaps in the lowest and top five percentiles. The elasticities are no greater than 0.20 in the poorest decile and fall gently to about 0.05 at the topmost decile. This is far removed from the rural Maharashtra of 1983, when the elasticities were as high as 0.65, falling across the expenditure distribution to 0.40. Indeed, not only have levels of calorie availability diminished over the two decades since then, but the strength of the calorie response has declined too.

Figure 3.8 is the counterpart to Figure 3.4 for the price paid per calorie. The vertical axis plots the logarithm of the calorie price (in Rupees per thousand calories) and the horizontal axis shows the logarithm of the per capita monthly outlay. The local linear regression is the solid blue curve, and it overlaps with the quadratic model (broken red lines) in the middle of the expenditure distribution. Again, the most noticeable difference lies at the extremes of the expenditure distribution; near the top, for instance, the slope of the nonparametric model becomes negative. The local linear fit also appears flatter than the quadratic regression at these extremes. The confidence bands for the local linear regression are shown in Figure 3.9. As with the calorie regression, the bands are very tight in the median of the expenditure distribution. The price per calorie regressions are drawn based on the assumption that intra-group substitution is a minor response to changes in household income, and the bulk of the substitution response occurs between food groups.

The elasticities produced by the nonparametric regression are in Figure 3.10 along with the confidence intervals. There is a steeper decline in the elasticity of price, starting at 0.40 for the poorest households and falling to -0.10 for the richest. The negative elasticity at the top of the expenditure distribution was hinted in the nonparametric regression line from Figure 3.8, although the elasticity estimate is less precise than in the middle of the distribution. At the mean monthly expenditure - about 4.98 on the logarithmic scale - the expenditure elasticity of price per calorie is around 0.23, which agrees with the gap between the outlay elasticities of calories (0.118) and of food expenditure (0.332). Both the steep drop in the price elasticity across the distribution of households and the negative elasticity at the very top are in sharp contrast to the results of Subramanian and Deaton (1996) for rural Maharashtra in 1983.

The next section begins to explore how the robust nonparametric elasticities estimated so far are. By introducing selected covariates of calorie availability, it examines if the bivariate relationships described hold up or change in the presence of new variables which affect the calorie consumption of households.

3.4 Regressions of Calorie Availability

The nonparametric regressions ignored important determinants of the amount of calorie consumption in a household, such as household size. As argued in Section 3.2, household size and composition play an important role in the total quantity and variety of food purchased by a household. The regressions will use the number of consumer units, which is a computed measure of the calorie requirement in a household, based on the age and sex of its various members, in an attempt to distinguish the expenditure elasticity of calories between households of similar size but different composition. Additional regressors include the age and sex of the head of the household, the social class and religious group of the household and the main source of livelihood.

Figure 3.11 summarizes the results of the calorie-expenditure relationship. It shows a lowess smoothing based on Cleveland (1979). This is the LOWESS algorithm whereby a series of weighted regression functions are estimated over grids of evenly spaced points in the data and the weights are assigned according to the distance of an observation from a selected sample point in the grid; the weight diminishes with distance from an observation in the grid. The lowess regression is the solid blue curve in the graph and the fitted quadratic model is shown in broken red lines. The vertical bars in green represent the first and third quartiles in the expenditure distribution. The diagram indicates a possible approximation to the nonparametric model, i.e. a spline function in MPCE in a linear regression of log calories on log expenditures. Although visual inspection suggests that the slope of the nonparametric regression is not particularly different between the second and third quartiles, a spline for each quartile of expenditure is preferred in the regressions.

Thus, Table 3.4 begins with the results of two alternative parametric models with controls for the demographic and occupational characteristics of households. Columns [1] through [4] contain the regression results from using linear splines for each of the four quartiles of the expenditure distribution. The first column controls only for household size and consumer units. The second column introduces demographic controls, including the age and sex of the household head, as well as the religion and caste group. The third column brings in dummy variables for the primary income-generating occupation of the household. There are four labor dummies for occupation: self-employment in agriculture, self-employment in non-agriculture, supplying market labor in the agricultural sector and supplying market labor outside agriculture. The base category in occupation is simply other types of labor.

In all the six regressions, the covariates that were statistically significant at least once were the household size, the consumer units associated with a household, the dummy variables for belonging to a scheduled caste, to a scheduled tribe, and to a specific religious group and an indicator for self-employed agricultural households. Scheduled caste and tribe households have a lower elasticity than the average household. Producer-consumer households in the agricultural sector purchase more calories than the mean household. Household size is inversely related to calorie availability. Calorie availability increases in households with greater consumer units, so households of the same size but containing a greater proportion of adults and males, purchase more calories.

As expected from the nonparametric gradients, the regressions in Table 3.4 predict that elasticities fall with higher levels of income. Columns [1] through [4] show that the elasticity declines from 0.18 in the first quartile to 0.06 for the median class and to nearly zero for the richest quarter of the population. It is true however, that it is only the spline for the fourth quartile that is significant. District fixed effects which are included in the fourth regression were mostly significant and they have produce smaller elasticity estimates, which fall through the distribution from 0.13 in the first quartile to 0.07 in the median to 0.02 in the fourth quartile.

Columns [5] and [6] show the regression results with a quadratic in the logarithm of monthly per capita expenditures. Both quadratic regressions predict a mean expenditure elasticity for calories of 0.10. The addition of demographic covariates and controls for the occupations do not alter this point estimate. The inclusion of district fixed effects in the quadratic regression does not affect the gradient (results not shown in the Table 3.4). However, the exclusion of the district fixed effects produced a significant F-statistic of 17.55. This might be a reflection of the broad similarities that households in the same district might share such as the price of food, community and social infrastructure although the fixed effects are not serious enough to change the elasticities.

3.5 Conclusion

The elasticity estimates discovered for rural Maharashtra in the consumption survey data of 2004/05 are in what would be considered the low range. The estimates are even smaller than the values reported by Behrman and Deolalikar (1987) with panel data from the International Crop Research Institute for the Semi-Arid Tropics - Village Level Studies (ICRISAT-VLS). For rural Maharashtra in 1976-1978, Behrman and Deolalikar (1987) favor a calorie elasticity no higher than 0.37 at the mean, which was incidentally a statistically insignificant estimate. Examining rural Maharashtra with the 1983 NSS data, Subramanian and Deaton (1996) argue that the true mean estimate was in the range of 0.28 to 0.38, and statistically different from zero.

The ICRISAT data differs from the National Sample Survey data in certain important ways: it collects intake, rather than availability data, it uses a recall period of 24 hours instead of a 30-day period, it collects income data, while the NSS does not, and it tracks only 240 households in six selected villages in the states of Maharashtra and Andhra Pradesh, sampling about five percent of the total number of households interviewed in Maharashtra alone in the National Sample Surveys. Intake data typically generate a lower estimate of elasticities than availability data, since the latter tend to be computed by converting quantities of foods using fixed calorie contents. The aggregation process to arrive at total calories tends to inflate the calorie consumption of households. Adjustments based on the calorie content of meals mitigates some of the inflation in total calorie values, yet as Strauss and Thomas point out intake data have often returned low elasticities in the literature (See Table 34.1 in Strauss and Thomas (1995)).

Shorter recall periods generally increase the reported rates of monthly expenditure, and would bias estimates of the expenditure elasticity of calories downward. If the larger sample size in the NSS data guarantees a higher precision in the estimates and the elasticity reported by Subramanian and Deaton (1996) are preferred as a point of comparison, then the present study finds that the elasticities have diminished considerably over the period 1983-2004. Between the 38th round (1983) and the 61st round (2004/05), the NSS has not changed the design of its questionnaires very significantly. It continues to use the recall period of 30 days for all frequently purchased goods and by and large, the list of food items has remained intact.

One important influence examined in the regressions following the bivariate relationships presented in the graphs in Section 3.3 is household size and composition. The mean estimate of about 0.15 drops at the most, by a third in the presence of these covariates. The basic estimate does not suffer excessively when district fixed effects are introduced.

Calorie elasticities reported here are comparable to the estimate of Bouis and Haddad (1992) for Philippine farm households where OLS, two-stage least-squares and fixed effects estimates all converge at elasticities between 0.08 to 0.12. In more recent studies, Aromolaran (2004) finds that the calorie elasticity in rural South-West Nigeria is close to zero and no higher than 0.04 and that redistribution from men to women in household resources reduces per capita intakes. In contrast, Gibson and Rozelle (2002) report slightly higher estimates of the calorie elasticity for Urban Papua New Guinea. They determine that it lies in the range of 0.42, when they run a least-squares regression with no controls to 0.18, when they instrument for household total expenditures with non-food expenditures and control for household demographics, schooling and occupational characteristics as well as use cluster fixed effects.

Although the feedback from income to calories is not "high" in rural Maharashtra of 2004, the estimates are still significant. This is important to policy framing since income supplementation and employment programmes may show some success in securing higher calorie consumption to the poor. What may have contributed to the low elasticities, however, is that some of the undernourished households do not have as much of an incentive to increase their calorie consumption today, compared to about twenty years ago.

Since the elasticity of price paid for calories is about the same magnitude and relatively unchanged since 1983, it suggests that households could not have been affected by higher food prices in trying to increase their calorie consumption. India went through two decades of relatively low food inflation beginning in the eighties. Statistics published annually by India's central bank reveal that food grain prices declined over the period 2000-05 with the inflation rate in rural areas coming down from 8.1 per cent in 1993-2000 to 1.90 per cent in 2000-05. Official poverty estimates in the early 2000s attributed the decline in poverty to low food inflation. It is true that since 2006, food prices have been rising -

inflation even hit 14 per cent in 2010 - but the data used here come from 2004/05. Further, given that real per capita incomes grew at an average of four percent each year between 1999 and 2004/05, households could not have been held back by low food inflation if they were trying to increase calorie consumption.

Food production, especially of cereals, declined over the period 1990-91 to 2004-05 in Maharashtra. Mishra and Panda (2006) report that the area under cultivation and the output of all cereals declined but one coarse cereal, jowar (sorghum) recorded a decline of 1.5% per annum between 1990 and 2005. Agriculture has fallen in terms of its contribution to the state domestic product in Maharashtra from 28% in 1980 to 12% in 2004. The growth rates have also declined since the end of the eighties, from an annual rate of over 5% to less than 2% since. Somewhat puzzlingly, the decline in prosperity has not driven a whole number of households out of agriculture. The population earning its livelihood in this sector has only declined from 62% to 55% over the period 1980 to 2004. As work burdens fall, the energy needs of households that remain in the primary sector fall too, so that these households could make do with fewer calories to sustain themselves.

There is some evidence to this notion that households have experienced lower calorie requirements since 1983. The nature of work in the agricultural sector has undergone a large change, and relative mechanization has lowered the need for much manual labor today. Yet, the percentage of workers engaged in agriculture has only dropped from 62% in 1983 to 55% in 2004/05 - not exactly a large change. Of course, a large decline in the agriculture-dependent population would not be sufficient to conclude that they were engaged in physically less demanding occupations by 2004. Employment in urban construction, for instance, would not necessarily lower calorie needs. The type of occupation is not a complete indicator of calorie requirements. The phenomenon of underemployment also gets in the way of using occupations to guess calorie requirements for an agricultural household. A worker who is less than fully-employed will not be distinguishable from a full-time employee unless there is more precise information, such as the number of days in a year worked and the duration of a work day.

Indirect evidence for lowered calorie requirements might be sought in the kind of assets rural households possess today: even cheap means of transport, like scooters or bicycles could reduce the time and energy taken to commute to town-centers, local markets, to fetch drinking water, etc. Even television sets at home might reduce the calories spent per capita per day, but the saving from this activity is not likely to be large. Deaton and Drèze argue that improvements in the health environment could have also contributed to lower calorie needs than two decades ago.

The decline in rural Maharashtra's calorie consumption and elasticity cannot be ascribed to falling real incomes, since both per capita incomes and total real expenditures have increased over the period 1983-2004. Likewise, there has been a moderate increase in the real price paid for all calories, but not of the order of magnitude that could cause declining calorie consumption or a shrinking elasticity. The best explanation for the decline lies in the idea of diminished calorie needs. Given that the cost of obtaining the daily calorie minimum is within the budget of the poorest rural workers in Maharashtra, an undernutrition trap is unlikely to be at work.

The trend of declining calorie consumption in India has raised concern largely because solving hunger has been a policy focus for more than fifty years. Decidedly, calories are not the best indicator of nutritional status. Calorie needs vary by an individual's sex, age, activity levels, stress levels, basal metabolism, periods and frequency of illness, etc. These variations are not reflected in the household-level calorie norms, which themselves have been recognized as outdated or inaccurate. When the norms are not reliable, considering calorie shortfalls alone to assess nutritional well-being can be misleading. A nutrition production function, as pointed out by Schiff and Valdes (1990) would include several non-nutrient food determinants, many as crucial as caloric intake. For instance, publicly-provided inputs like piped water, electricity, nutrition education, and sewage systems help raise the nutritional status of a population. As the quality of these inputs increases, the nutritional status improves for any given calorie intake. Thus, as Deaton and Drèze (2009) suggest, it may be necessary to look to more than just calories when seeking to assess the nutritional status of a society. Calorie deficiency is not unimportant, and other micronutrient deficiencies may well overlap with it but relying on calories exclusively may lead to an overestimation of undernutrition.

Thus, while not completely abandoning calories as a signal of how well-nourished a population might be, Jensen and Miller (2010b) offer an interesting strategy to determine the level of subsistence nutrition. Their approach - one that would suit the Indian context well - centers on using the share of staple calories in diets to locate the calorie threshold at which households begin diversifying their food consumption, by substituting away from calories. Thus, they identify *individual*-specific thresholds where calories cease to be the main priority and its income elasticity diminishes. Thresholds like these would enable policy-makers to target income-augmenting programs to households that still exhibit a "high" income-calorie elasticity. This revealed preference approach to gauging calorie sufficiency is an attractive alternative to relying on un-updated calorie norms to measure undernutrition in a developing country. If current norms do not indicate the extent of calorie deprivation well, then an analysis along similar lines for rural Maharashtra would go a long way in two respects: for one, it would lend new perspective on the extent of undernourishment in this region, and ascertain if there is cause for anxiety. For another, it could provide the basis for improved nutrition monitoring of poor households, by providing a standard that can be checked and updated relatively easily over time.

Table 3.1: **Expenditures and Calorie Availability: Rural Maharashtra, 2004/05**

Decile of MPCE (1)	Real monthly per capita expenditure (2)	Budget share of food (3)	Per capita meal-adjusted daily calorie availability (4)	Percentage of calorie norm (5)	Percentage of alternative calorie norm† (6)
Bottom	65.88	83.3	1673	69.7	92.4
Second	86.25	71.6	1811	75.5	100.0
Third	99.44	66.7	1863	77.6	102.9
Fourth	112.09	61.4	1830	76.3	101.1
Fifth	124.37	58.0	1900	79.2	105.0
Sixth	138.97	55.7	1958	81.6	108.2
Seventh	156.48	50.7	1945	81.0	107.5
Eighth	181.04	45.9	1977	82.4	109.2
Ninth	223.93	39.3	1997	83.2	110.3
Top	362.24	29.5	1995	83.1	110.2
Mean	154.99	56.2	1895	79.0	104.7

Notes: MPCE stands for monthly per capita expenditure. The per capita daily calories are adjusted for the number of meals eaten at home, at work, received for free and provided to non-household members. The budget share of food is also computed from the meals-adjusted expenditure on food.

† The alternative calorie norm is 1810 calories, which was recommended by the FAO. Column [5] uses the norm of 2400 calories per capita per day, which dates back to the recommendation of the 1979 Planning Commission of the Indian Government.

Table 3.2: **Summary Statistics of the Main Variables**

Variable	Mean	SD	Min	Max	Obs.
Per Capita Daily Calorie Availability					
Unadjusted	1982.3	488.7	986.5	4109.6	4840
Adjusted	1990.8	490.1	1002.3	4124.2	4840
On the log scale	7.6	0.2	6.9	8.3	4840
Monthly Per Capita Expenditures (MPCE)					
On all goods	641.5	355.3	131.5	2581.3	4840
On food	324.9	143.2	72.2	2166.5	4840
Log real MPCE	5.0	0.5	3.5	6.5	4840
Rupees Paid Per 1000 Calories					
Nominal price	5.4	1.5	1.5	26.3	4840
Log real price	0.3	0.3	-0.9	1.9	4840
Other Variables					
Household size	4.9	2.4	1.0	26.0	4840
Consumer units	3.9	1.9	0.5	22.0	4840

Notes: All figures are calculated from the 61st round of the National Sample Survey for a sample of 4840 households from rural Maharashtra. The topmost panel reports calorie availability, rather than calorie intake, because the survey gathers data on the quantities of different foods *purchased* by a household, which may differ from the quantities consumed. Adjustment in the calorie data is to do with the number of meals consumed at home, work and elsewhere; see text for more details. The recommended daily calorie norm for rural agricultural laborers is 2400 calories. Monthly expenditures are divided into food and non-food categories. Real variables appearing under expenditures and price per calorie are in 1983 Rupees. Consumer units refer to the number of "equivalent adults" in a household; the calorie requirement of a given individual is expressed as a multiple of a young, adult male's minimum calorie needs.

Table 3.3: **Expenditure Patterns, Calorie Shares and Price Per Calorie: Rural Maharashtra, 2004/05**

Commodity Group	Expenditure Shares (%)			Calorie Shares (%)			Price Per Calorie (Rupees Per 1000 Calories)		
	Mean	Bottom 10%	Top 10%	Mean	Bottom 10%	Top 10%	Mean	Bottom 10%	Top 10%
Cereals	29.70	31.90	25.80	64.30	69.70	57.00	3.59	3.46	3.77
							0.64	0.51	0.79
Pulses	7.80	9.00	6.70	5.30	5.40	5.40	1.93	1.83	1.98
							1.51	1.44	1.60
Dairy	10.00	8.00	13.20	4.60	3.40	6.90	3.36	3.10	4.00
							3.69	3.59	3.92
Oils and Fats	11.20	11.60	11.00	10.30	9.10	11.90	1.48	1.43	1.56
							1.74	1.67	1.81
Meat and Eggs	4.30	4.30	4.80	0.50	0.40	0.60	14.77	14.68	14.82
							11.70	11.00	12.20
Fruits and Vegetables	15.10	14.00	16.30	5.70	4.40	7.20	11.30	11.14	12.45
							3.90	3.83	3.85
Sugar	6.20	6.10	5.30	7.00	5.70	7.50	1.17	1.21	1.14
							1.01	0.94	1.09
Other Food	15.60	15.00	16.80	2.30	1.90	3.50	133.62	106.03	186.16
							17.40	16.80	15.90

Notes: Expenditure shares refer to the share of a commodity group in the household's monthly budget. Calorie shares refer to the share of a commodity group in the unadjusted total calorie consumption of the household. Both are calculated for individual households and then averaged over the sample. The price per calorie is the number of rupees needed to procure 1000 calories from a commodity group. The prices have been deflated to their 1983 values - using the Consumer Price Index for Agricultural Labourers (CPI-AL) - for easier comparison with the second row of numbers under each food group. These numbers come from Table 1 in Subramanian and Deaton (1996, p.140) for rural Maharashtra based on 1983 NSS data. The columns titled "Mean" pertain to the mean household; the columns called "Bottom 10%" refer to the households in the bottom decile of the nominal per capita monthly household expenditure distribution while the columns called "Top 10%" refer to the households in the top decile.

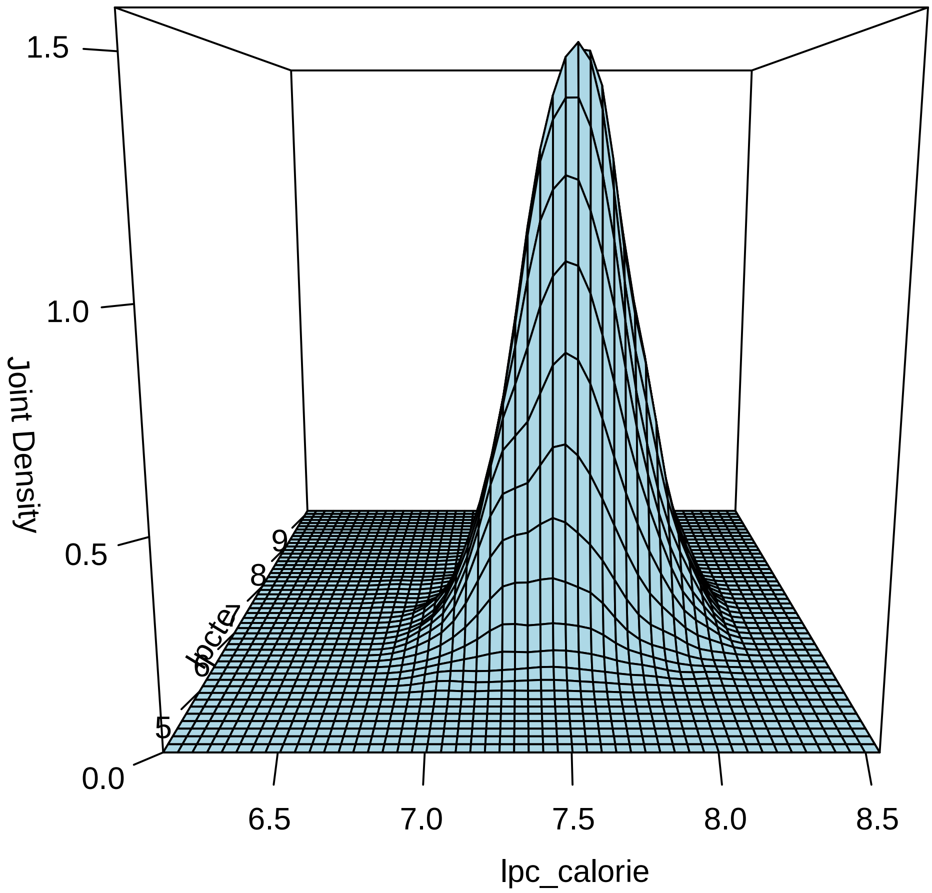


Figure 3.1: Joint Density of Calories and Expenditure

Notes: The figure shows the estimated joint density of the logarithms of per capita calories and monthly expenditures. See text for more details.

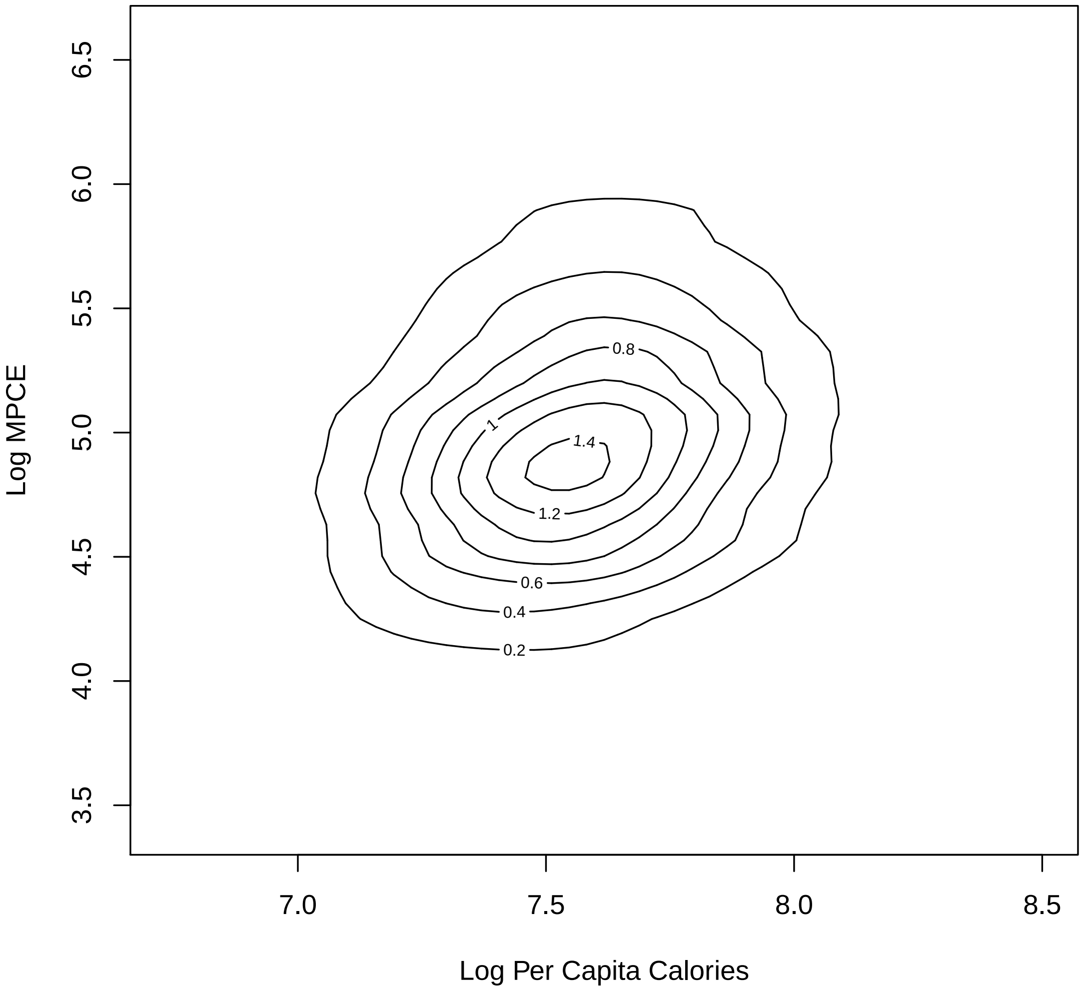


Figure 3.2: Contours of the Joint Density of Calories and Expenditure

Notes: This figure gives the contours corresponding to the joint density in figure 3.1. Although the contours are more circular than elliptical, the regularity of the shapes suggests that the joint density of the logarithmic transformation approaches normal.

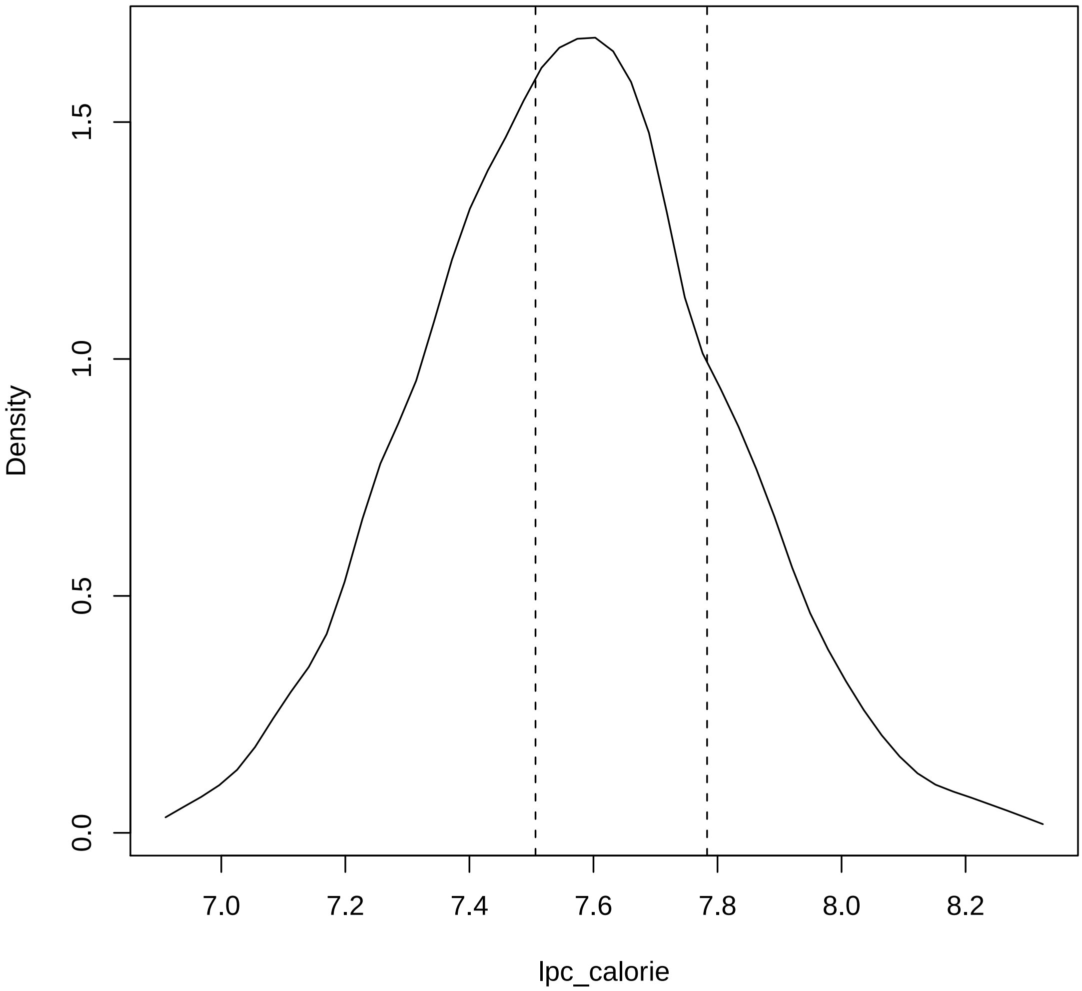


Figure 3.3: Unconditional density of Log per capita daily Calorie availability

Notes: The unconditional density was obtained in **R** by opting for a bandwidth chosen through cross-validation using the maximum likelihood procedure. This bandwidth = 0.035. The total number of observations in the sample is 4840. The vertical lines indicate the daily per capita calorie thresholds of 2400 and 1810 respectively.

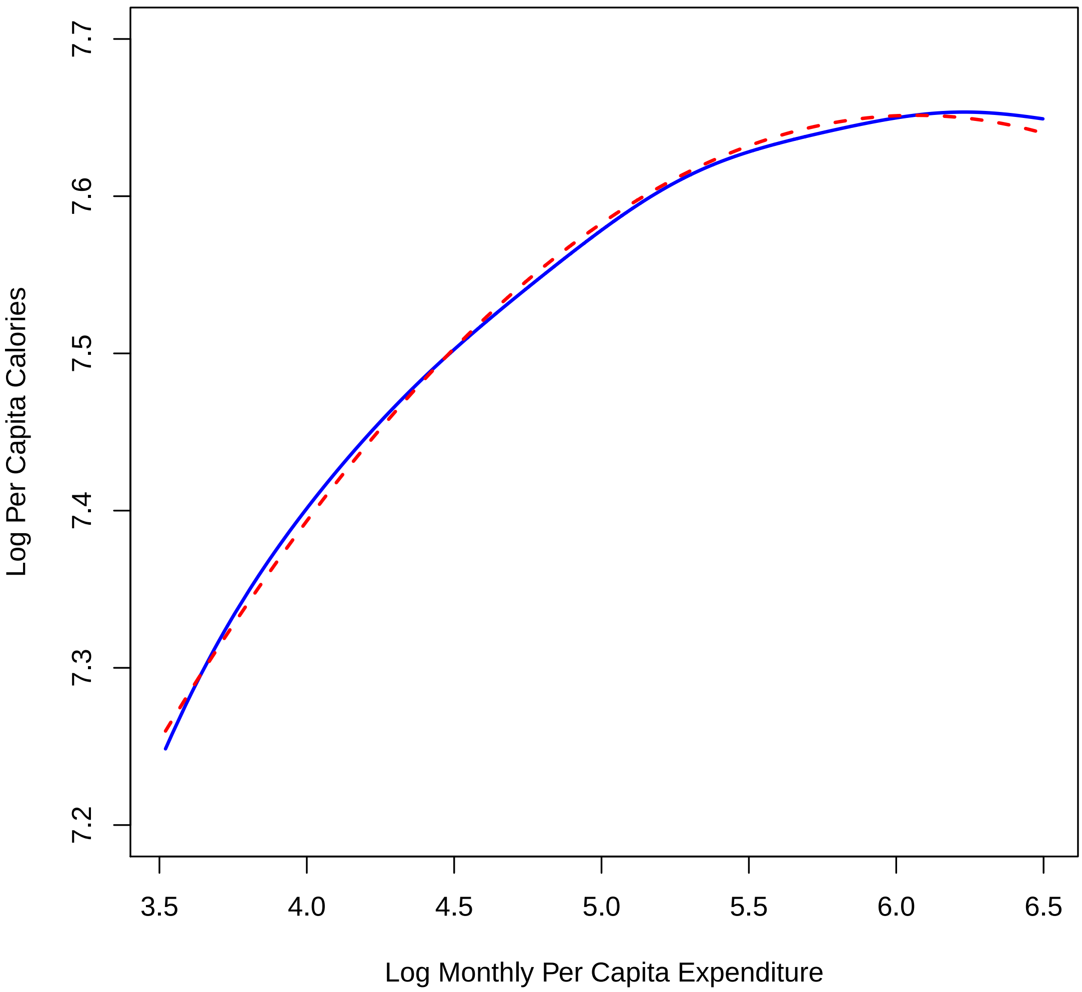


Figure 3.4: Quadratic and Non-Parametric estimation of the Calorie-Expenditure relationship

Notes: The figure shows two different regressions of log per capita daily calories on log monthly per capita expenditures. The quadratic fit is shown by the broken red curve and the nonparametric local-linear regression fit, by the solid blue curve. The smoothing parameter employed in the non-parametric specification was chosen through cross-validation by minimizing the Akaike Information Criterion. The bandwidth produced was 0.336. The total number of observations in the sample is 4840.

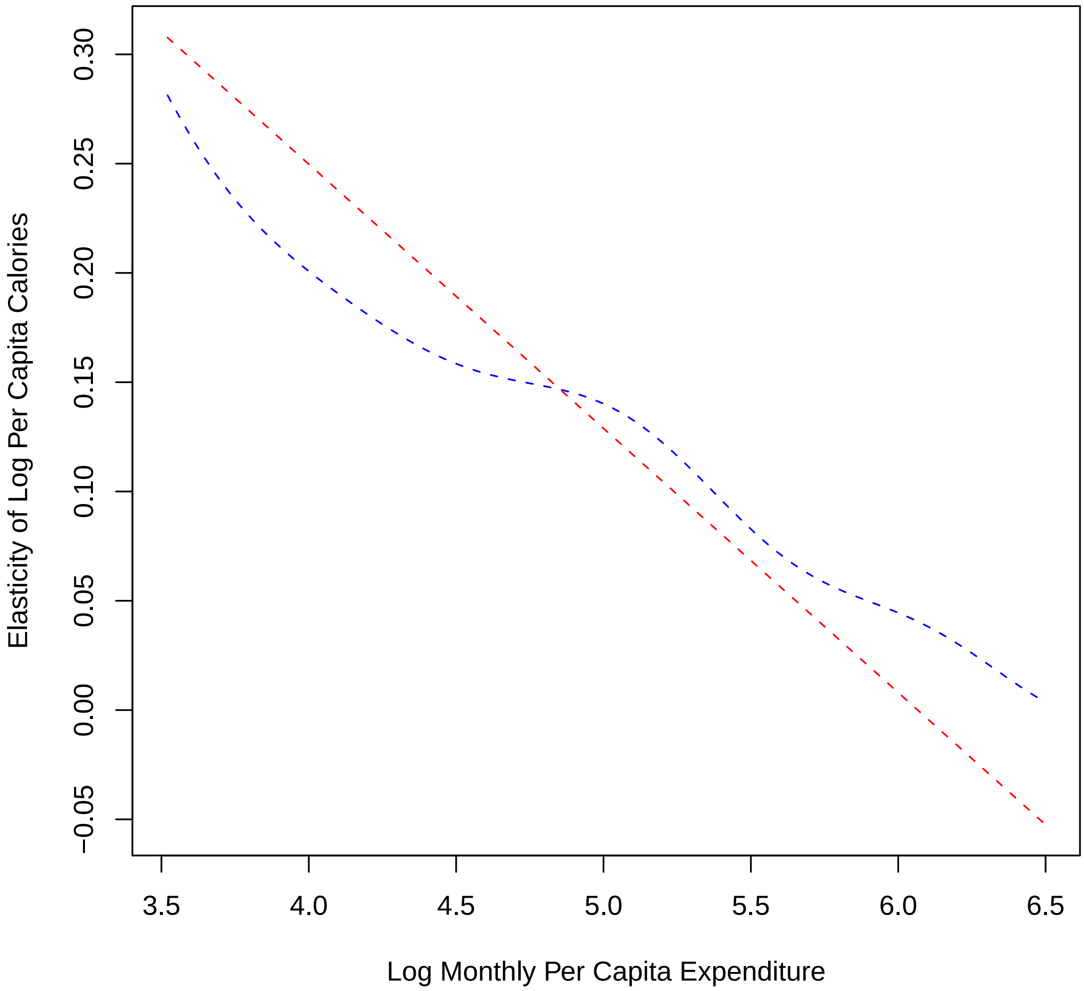


Figure 3.5: Expenditure Elasticities of per capita daily Calories

Notes: The figure shows the gradients obtained from the two different regressions of log per capita daily calories on log monthly per capita expenditures. The gradient associated with the quadratic model is shown by the broken red curve and the elasticity from the nonparametric local-linear regression fit, by the broken blue curve. A fixed bandwidth of 0.336 chosen through cross-validation was used in the non-parametric specification. The total number of observations in the sample is 4840.

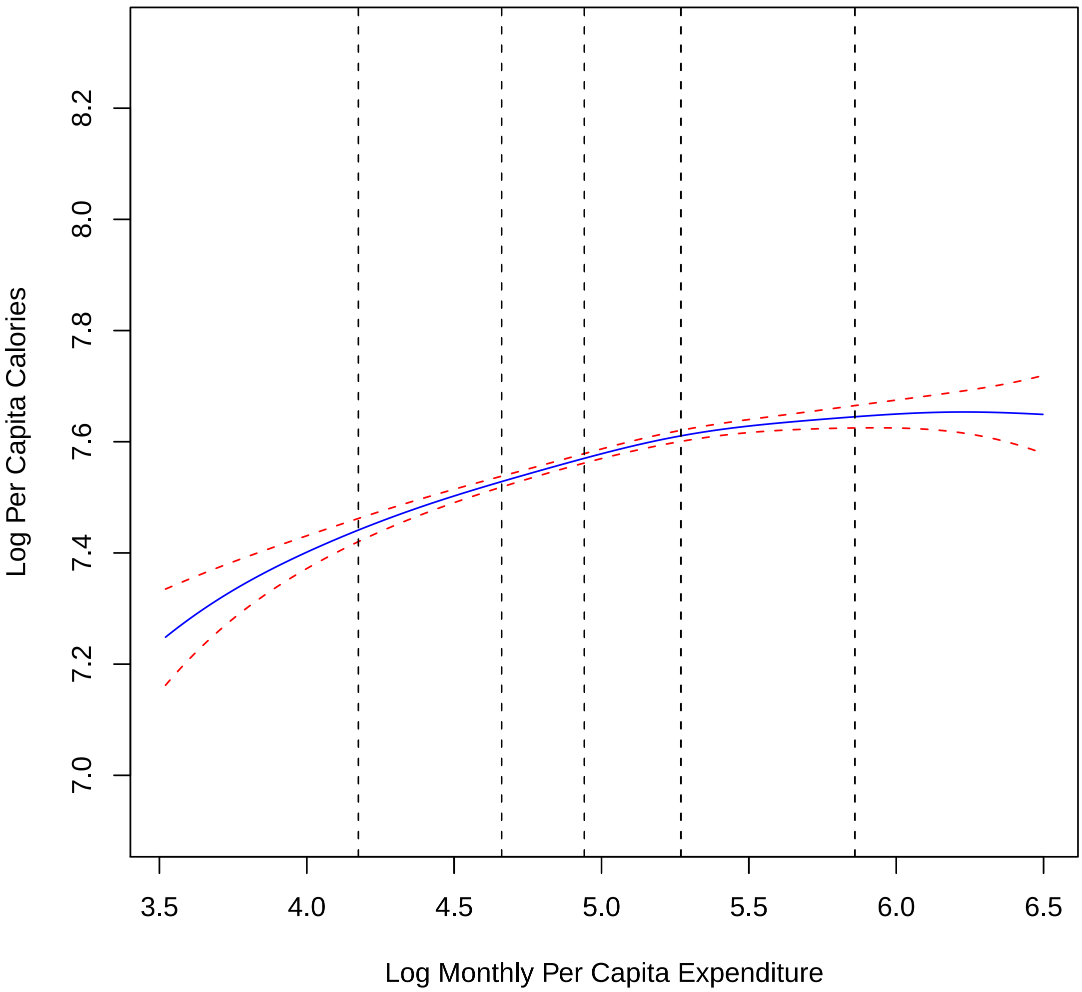


Figure 3.6: Local-linear Regression of Log per capita daily Calories: Rural Maharashtra, 2004/05

Notes: The figure shows the nonparametric local-linear regression of log per capita daily calories on log monthly per capita expenditures. The broken red bands are the 95% confidence intervals obtained by bootstrapping the standard errors for the regression model. The black vertical lines mark successively, the bottom decile, the first quartile, the median, the third quartile and top decile in the expenditure distribution. A fixed bandwidth of 0.336 chosen through cross-validation was used in the non-parametric estimation. The total number of observations in the sample is 4840.

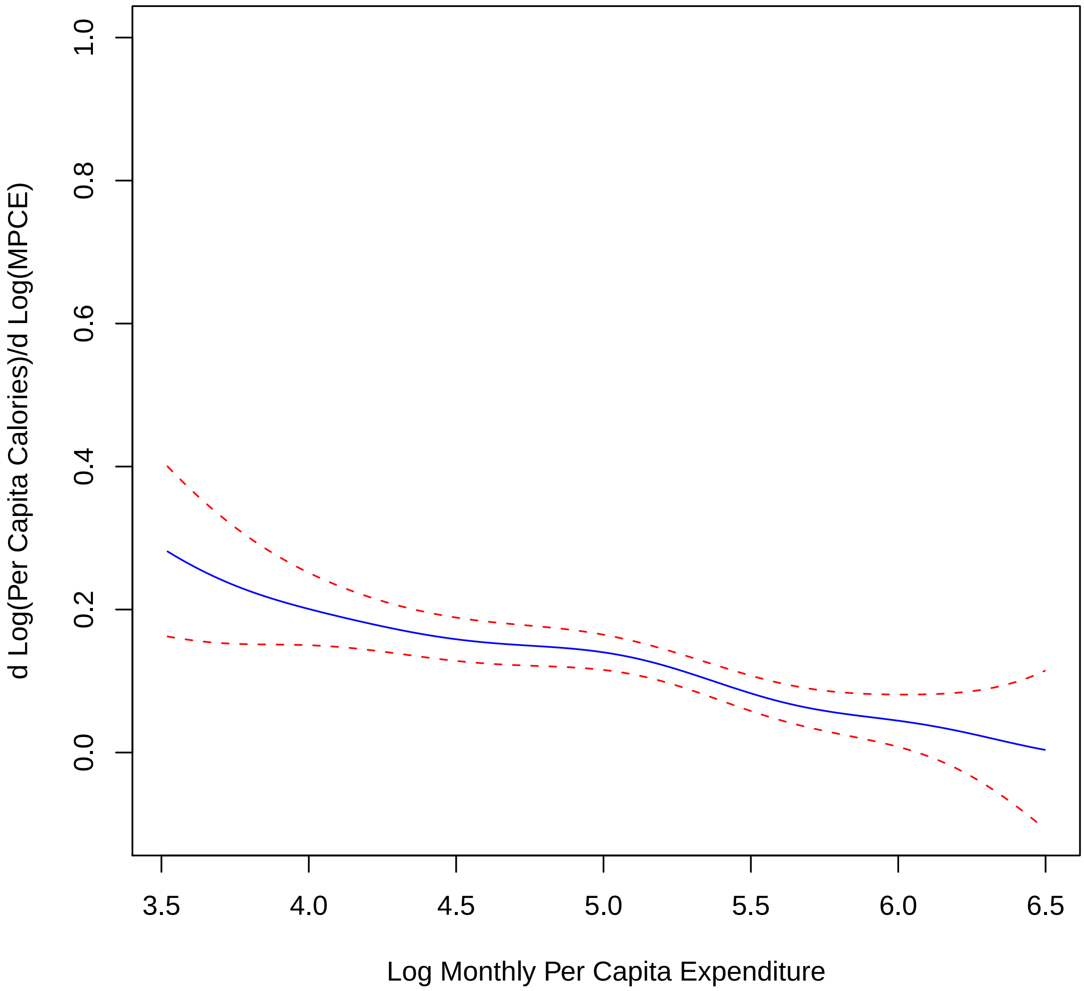


Figure 3.7: Local-linear Expenditure Elasticity of Calories: Rural Maharashtra, 2004/05

Notes: The figure shows the gradients obtained from the local-linear regression of log per capita calories. The gradient appears as the solid blue curve and the broken red bands represent the 95% confidence intervals constructed by bootstrapping the standard errors of the gradients. A fixed bandwidth of 0.336 chosen through cross-validation was used in the non-parametric specification to derive the gradients. The total number of observations in the sample is 4840.

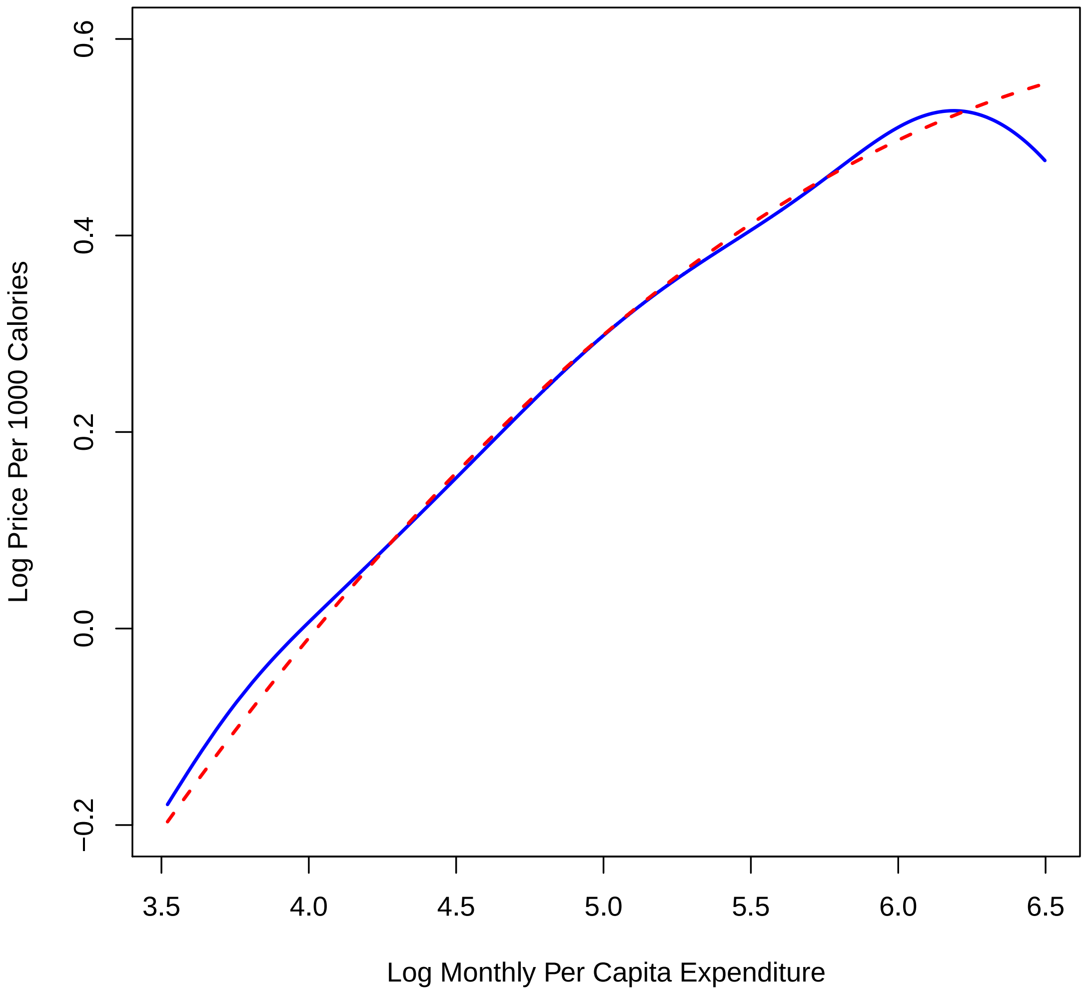


Figure 3.8: Quadratic and Non-Parametric estimation of the Calorie Price-Expenditure relationship

Notes: The figure shows two different regressions of the log price per 1000 calories on log monthly per capita expenditures. The quadratic fit is shown by the broken red curve and the nonparametric local-linear regression fit, by the blue curve. The smoothing parameter employed in the non-parametric specification was chosen through cross-validation by minimizing the Akaike Information Criterion. The bandwidth produced was 0.251. The total number of observations in the sample is 4840.

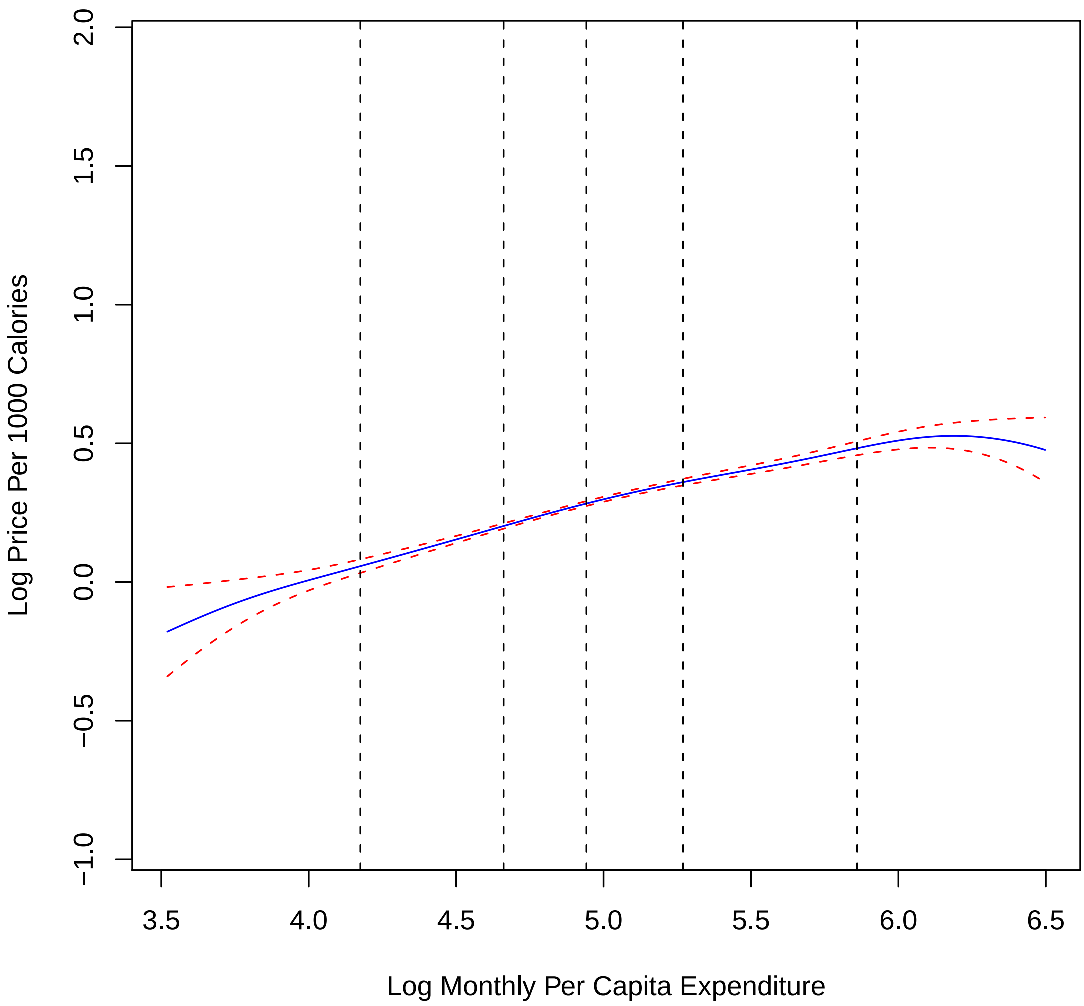


Figure 3.9: Local-linear regression of Log daily Calorie Price, Rural Maharashtra, 2004/05

Notes: The figure shows nonparametric local-linear regression of log price per 1000 calories on log monthly per capita expenditures. The broken red bands signify the 95% confidence intervals obtained by bootstrapping the standard errors for the regression model. A fixed bandwidth of 0.251 chosen through cross-validation was used in the non-parametric estimation. The total number of observations in the sample is 4840. The vertical black lines mark successively, the bottom decile, the first quartile, the median, the third quartile and top decile in the expenditure distribution.

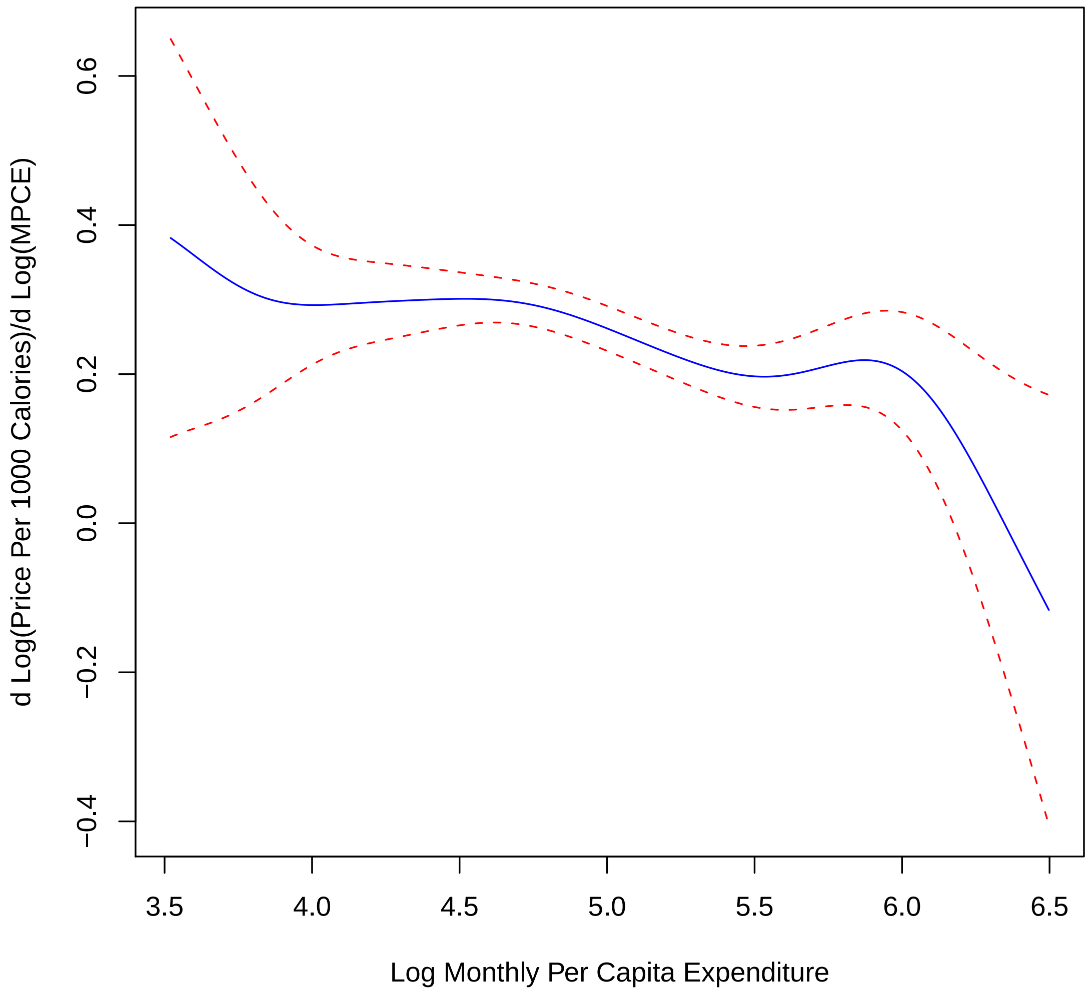


Figure 3.10: Local-linear expenditure elasticity of Log Price per Calorie, Rural Maharashtra, 2004/05

Notes: The figure shows the gradients obtained from the local-linear regression of log price per 1000 calories on log monthly per capita expenditures. The gradient appears as the solid blue curve and the broken red bands represent the 95% confidence intervals constructed by bootstrapping the standard errors of the gradients. A fixed bandwidth of 0.251 chosen through cross-validation was used in the non-parametric specification. The total number of observations in the sample is 4840.

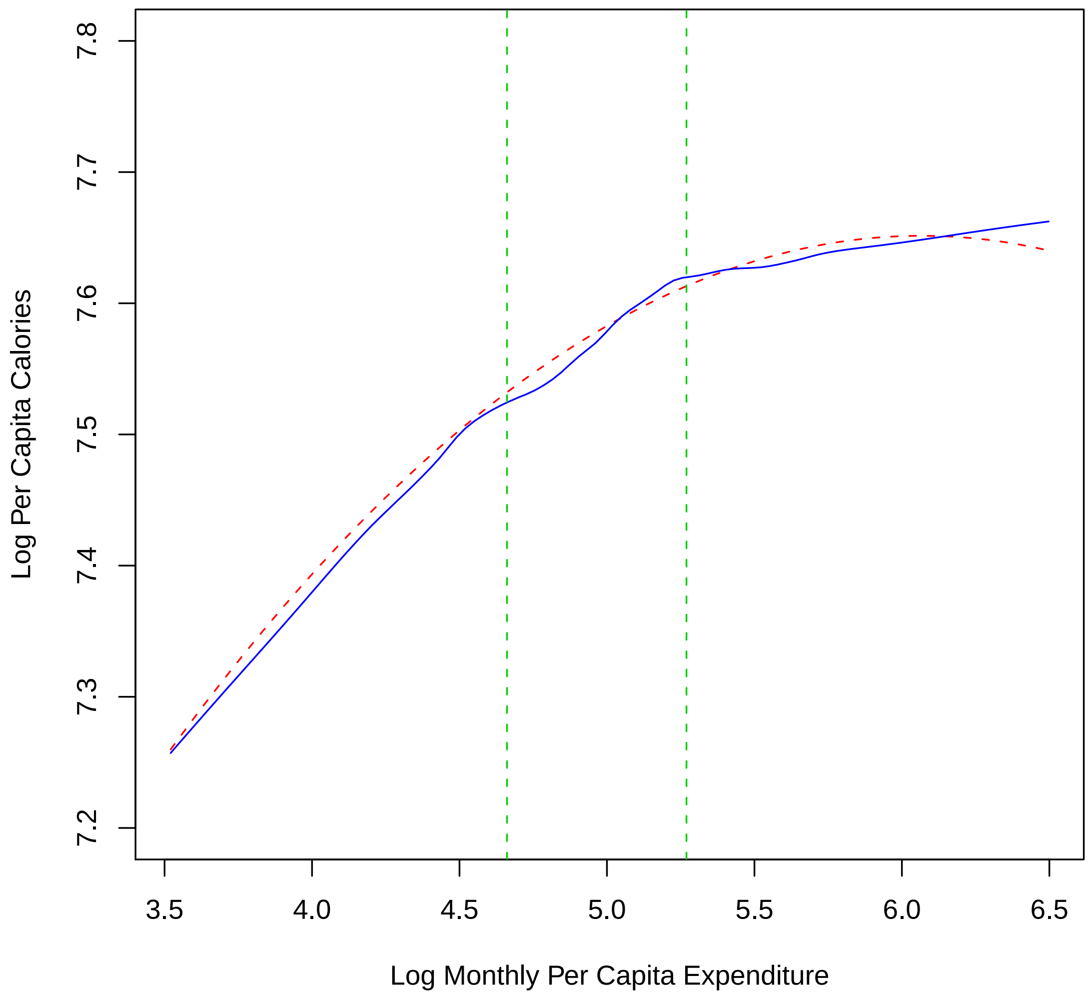


Figure 3.11: Lowess Smoothing of Log Calories

Notes: The figure shows the quadratic regression of the log calories on log monthly per capita expenditures and the lowess smoothed regression. The green vertical lines demarcate the first and third quartile of the expenditure distribution. The total number of observations in the sample is 4840.

Table 3.4: **Regressions of Calorie Availability: Rural Maharashtra, 2004/05**

Dependent Variable: Log Per Capita Calorie Availability						
	[1]	[2]	[3]	[4]	[5]	[6]
Log MPCE	0.184	0.157	0.158	0.13	0.832	0.684
	[0.051]***	[0.050]***	[0.050]***	[0.043]***	[0.158]***	[0.162]***
Log MPCE * I{Q_2}	-0.128	-0.126	-0.122	-0.064		
	[0.119]	[0.118]	[0.118]	[0.103]		
Log MPCE * I{Q_3}	-0.072	-0.066	-0.063	-0.026		
	[0.111]	[0.111]	[0.111]	[0.100]		
Log MPCE * I{Q_4}	-0.183	-0.158	-0.148	-0.114		
	[0.061]***	[0.060]***	[0.061]**	[0.056]**		
I{$Q_2 = 1$}	0.603	0.592	0.573	0.299		
	[0.559]	[0.554]	[0.554]	[0.485]		
I{$Q_3 = 1$}	0.348	0.322	0.305	0.122		
	[0.543]	[0.546]	[0.544]	[0.489]		
I{$Q_4 = 1$}	0.921	0.799	0.748	0.574		
	[0.293]***	[0.289]***	[0.294]**	[0.270]**		
Log MPCE Squared					-0.073	-0.059
					[0.016]***	[0.016]***
Log household size	-0.307	-0.302	-0.302	-0.321	-0.306	-0.302
	[0.023]***	[0.023]***	[0.023]***	[0.021]***	[0.023]***	[0.023]***
Consumer units	0.035	0.033	0.033	0.037	0.035	0.033
	[0.006]***	[0.006]***	[0.006]***	[0.005]***	[0.006]***	[0.006]***
Scheduled tribe		-0.061	-0.058	-0.027		-0.057
		[0.016]***	[0.016]***	[0.016]*		[0.016]***
Scheduled caste		-0.028	-0.023	-0.037		-0.023
		[0.014]**	[0.014]*	[0.013]***		[0.014]*
Self employed in agriculture			0.034	0.041		0.035
			[0.015]**	[0.015]***		[0.015]**
Observations	4840	4840	4840	4840	4840	4840

Notes: Robust s.e. in brackets. * is for significance at 10%, ** for 5% and *** for 1%. Regressions in columns [1] through [4] use a (linear) spline in expenditure; column [2] introduces controls for the age, sex, caste and religion of the household head; column [3] brings in dummies for occupations and column [4] adds district fixed effects, the omitted district being Thane. (Mumbai is fully urbanized and does not figure in the regression.) Columns [5] and [6] employ a quadratic in log expenditure, with column [6] controlling additionally for household demographics. MPCE is short for (real) monthly per capita expenditures. The omitted category under religion is non-Hindu; the omitted caste is others. Four occupation dummies were used for self-employed/market labor in agriculture. Columns report only the significant covariates.

Chapter 4

The Intergenerational Transmission of Parental Schooling: Evidence from Zimbabwe's Education Reform

[1]

4.1 Introduction

The argument that education provides a route to escape poverty is long-standing in the social sciences. Education is envisioned as a form of investment that lifts families out

[1] This chapter is co-authored with Professor Jorge Agüero. It is included in this dissertation with his permission. All errors are my own.

of constrained living circumstances and sets them on a path of higher income and better health. From time to time, governments have designed programs aimed at encouraging initial enrolment and promoting continued attendance through secondary school. Often, economic incentives are offered, such as free primary education, cash transfers conditional on children attending school, tax-breaks or cash transfers for families sending girls to school, subsidized schoolbooks, and free uniforms. The literature evaluating these programs provides reassuring evidence of their effectiveness in raising school attainment[2].

However, their effects on the *intergenerational* transmission of schooling are less clear. Well-educated parents often have well-educated children. Part of this correlation is driven by selection: highly-educated parents usually earn large incomes, and so they are able to afford more schooling for their children. Yet, education could also change the choices parents make in their children's human-capital formation. A few well-designed studies have explored how and by how much the education of parents affects the human capital of the children but these exist in the context of developed countries rather than developing countries, where policies meant to break the intergenerational transmission of poverty are vital (e.g. Behrman and Rosenzweig 2002, Black, Devereux, and Salvanes 2005). In this paper we describe a unique natural experiment that helps estimate the transmission of human capital from parents to children in Zimbabwe.

The transmission of human capital across generations is also a reference to the causality problem: how do we know that increased schooling in one generation by itself leads to greater schooling in the next? Some studies have used data on twins to attempt an

[2]See Glewwe (2002) for a review of the recent literature.

answer. Behrman and Rosenzweig (2002) use a sample of parents who are one of a pair of identical (monozygotic) twins. They find a significant and positive relationship between the schooling levels of fathers and children but a *negative* and significant relationship between mothers and children, after they control for assortative mating, female earnings and the mother's child-rearing endowments. Antonovics and Goldberger (2005) contest the results, believing them sensitive to the coding of data. The intergenerational theme is revisited by Currie and Moretti (2003) and Chen and Li (2009), who try to identify the impact of increased mother's schooling on children's health. While Currie and Moretti (2003) rely on instrumental variables, Chen and Li (2009) use data on adopted and biological children in China.

Plug (2004) and de Walque (2009) use data on adoptees in the United States and Rwanda respectively. When adoption occurs randomly, children are matched to adopting families that are unlikely to share the same genetic characteristics. Thus, adoption filters out the nature component in child-schooling, helping to identify the link between parental nurturing and children's schooling. Plug finds a positive marginal effect on children's schooling with respect to fathers but no significant impact from mothers. de Walque finds that additional schooling acquired by either parent in Rwanda benefitted adoptee-orphans. The drawback, as pointed out by Black, Devereux, and Salvanes (2005), is that in practice, children are not placed randomly with parents; so children's education could bear the results of unobserved parental characteristics, including patience and ability (p.438)[3].

[3]Sacerdote (2002) also uses adoptees but his paper focuses on the effect of family socioeconomic background instead of the causal effect of parental education.

Lastly, Oreopoulos, Page, and Stevens (2003) and Black, Devereux, and Salvanes (2005) use time and spatial variation in compulsory schooling laws to identify the effects of parental education on children's education. Thus, Oreopoulos, Page, and Stevens find that increasing parental education decreases the probability that a child would repeat a grade with U.S. census data but Black, Devereux, and Salvanes find that only fathers positively affect the education of the next generation. A potential confound in relying on regional variations in a reform is the possibility of selective migration into states or municipalities that implemented the reform early.

Aside from de Walque (2009) and Chen and Li (2009), the literature on the intergenerational transmission of human capital is largely drawn from developed countries. Duflo (2001), and Duflo (2004) provide an in-depth look at the consequences of Indonesia's 1973 primary school construction program, and finds that an additional school per 1000 children increased male education by 0.12 to 0.19 years and male wages by 3 to 5.4%. A preliminary study on Pakistan by Andrabi, Das, and Khwaja (2009) finds significant, positive effects of mother's schooling on children's test-scores, despite seeing no impact on the extensive margin of her time spent performing household chores or paid work.

After Independence in 1980, the first black government of Zimbabwe embarked on a large reform of its education system. A key feature of the reorganization involved the elimination of restrictions to progress to secondary education. Because of apartheid-style rules in Rhodesia, a 14 year-old graduating from primary school in 1979 had about a twenty percent probability of advancing to secondary school. Post-reform, a child who was 14 years old in *1980* had almost an eighty percent chance of attending secondary school.

The reform essentially created a *fuzzy* discontinuity in the probability of treatment (rather than in the treatment itself); think of the treatment as the event of attending secondary school. We use this discontinuity as a source of exogenous variation in the education of mothers and fathers in our data. Thus, by 2002, a woman aged 14 in 1980 had accumulated 0.8 more years of education on average, compared to one who was 15 in 1980. We further find that a child born to the former mother had 0.13 more years of education (or 5 percent of a standard deviation) than a child born to the latter (even after controlling for the child's age.) Examining fathers, we find that the size of the transmission of schooling to children is almost three times as much as from mothers. We provide an explanation for the relative strengths of the two transmissions and show that our results are robust to several confounding factors.

The remainder of the paper is organized as follows. The post-Independence reform in Zimbabwe's education system is laid out in section 4.2, followed by a description of the data in section 4.3. Section 4.4 presents the identification strategy employing fuzzy regression discontinuity and instrumental variables. Section 4.5.1 shows the main results of the paper followed by a set of robustness tests that confirm our findings (section 4.5.3). Section 4.6 summarizes and closes the paper.

4.2 Post-Independence Schooling Reform in Zimbabwe

Present-day education in Zimbabwe commences with basic and primary education. Primary school is a seven-year cycle which children enter at the age of seven years [4]. Barring

[4]In 1998, the official age of entry to primary school was lowered to six years. However, the EFA 2000 Assessment Report for Zimbabwe notes that according to the 1998 annual school census, most newly-admitted

grade repetitions, students phase into secondary school at around age fourteen. Secondary school consists initially of four grades, and Form IV leads to the O-level examinations. Provided their O-level performance is adequate, students may gain admission to an advanced two-year secondary school program (Lower and Upper Form VI), which culminates in the A-level examinations. The last stage of formal education is attending college and university. (See Nhundu 1992, Nherera 2000, Kanyongo 2005).

Apartheid had severely thwarted schooling opportunities for blacks in Rhodesia [5]. According to Riddell (1980, cited by Nhundu), at least 25% of black school-aged children failed to enter primary school due to a lack of places. Opportunities were further restricted when the Education Plan of 1966 allowed only 50 percent of black primary school graduates to enter secondary school. By 1976, for every 1,000 black school-aged children, 37 reached Form IV and less than 3 reached lower Form VI (Nhundu 1992, p.79). [6]

However in April 1980, the Republic of Zimbabwe came into existence. The first black government had campaigned with the goal of "establishing free and compulsory primary and secondary education for all Zimbabwean children regardless of their race, sex or class" (Nhundu 1992, p.78). The ensuing reform has been widely documented in the literature (See for example Edwards 1995, Edwards and Tisdell 1990, Dorsey 1989). Four key initiatives were undertaken by the new government: (1) the introduction of free and com-

first-graders were over the official admission age. Further, over-age boys were more likely to be admitted to first grade than over-age girls.

[5] Zimbabwe was known as Southern Rhodesia until political Independence in 1980. For a history of Apartheid-Rhodesia's education system and the policies dictating the quantity and quality of schooling Africans received, see (Atkinson 1972) or (O'Callaghan and Austin 1977). Nherera (2000) provides a discussion in the context of present-day globalization and livelihoods in Zimbabwe.

[6] Policies were also calculated to stifle supply: from 1961 to 1972, African attendance at academic secondary schools grew 1.5 times from 15,640 to 23,602, but legislation permitted the addition of just one government African secondary school (Zvobgo 1981).

pulsory primary education; (2) the removal of age restrictions to allow over-age children to enter school; (3) community support for education and; (4) automatic grade progression, in particular from primary to secondary school (Nhundu 1992, p.80). [7] These initiatives were hugely successful, in terms of increasing both access to secondary education and completed schooling levels among blacks.

As Figure 4.1 shows, gross enrolment in secondary schools climbed from 66,215 in 1979 to 482,000 in 1985 (an increase of 628 percent), peaking at a little over 700,000 in 1991 (Nhundu 1992, p.82). As secondary enrolment surged, the transition rate from Grade 7 to Form I rose from 27 percent in 1979 to 87 percent in 1980. Figure 4.2 shows that the transition rate jumped from an average of 30% in the seventies to 70% in the next decade. This dramatic rise in the probability of attending secondary school is the basis of our identification strategy: if the treatment under the reform is attending secondary school, then cohorts graduating from Grade 7 in the late seventies and early eighties clearly have very different probabilities of being treated. The treatment probability is discontinuous in age at the time of the reform, so that a fourteen year old in 1980 is thrice as likely as fourteen year old in 1979 to enrol in Form I. Thus, for reasons exogenous to their choice, younger individuals from the eighties tend to have higher levels of schooling; by instrumenting their completed schooling with their age in 1980, it becomes possible to extract a causal estimate of the human capital transmission to the next generation.

[7] We do not attempt to untangle the relative effects of the different initiatives, though Nhundu (1992) suggests that over-age enrolment and free primary education alone evoked "the nation's largest-ever Grade 1 intake in 1980". Our interest is in estimating the intergenerational transmission of schooling, so we prefer to view the reform as a single variable affecting schooling attainment. In the next section, we shall show that this variable is exogenous to several factors typically determining schooling.

4.3 Data

The main data in this study comes from a 10 percent random sample of the Zimbabwe Population Census of 2002 [8]. The sample of interest is composed of children whose parents are either heads of households or the spouses of household heads. In order to capture the appropriate "treated" and comparison cohorts, we choose children aged six to fifteen years in Census 2002, with at least one living parent in the age-group of 28 to 44 years (these parents would have been in the age-group of 6-22 years in 1980). This leaves us with 187,743 child records to be matched to one or both parents. Table 4.1 shows the descriptive statistics from three different samples. The panel titled mother's (father's) sample refers to the dataset where the children have been matched with the mothers (fathers).

The average child is 10.4 years old. Up to 95 percent of all children had attended school at one time, and 96 percent were in school at the time of the census. About 68 percent of all children were below grade level for their age, a consequence of over-age enrolment as much as grade repetition. Since ever-attendance and current school enrolment are so high, we do not expect to observe a significant marginal effect of increased schooling in the parent generation on either of these outcomes. Any intergenerational transmission of schooling should be more likely to show up in the child's (standardized) years of schooling or "delay", which is measured as the shortfall from schooling attainment for a given age.

The average mother in the sample is a little over 35 years of age, has 7.9 years of schooling and gone through nearly five pregnancies. The average mother also had her first birth at just 19.5 years of age. The average father is about 37 years old and has completed

[8] To the best of our knowledge, our paper is the first causal analysis based on this census.

nearly 9.5 years of schooling. Thus the average parent in the sample has had at least some education beyond primary school.

4.4 Identification Strategy

Apartheid had made secondary education a sphere of deep polarization between whites and blacks. In the early years of the reform, the greatest resources were spent opening new secondary schools, especially in rural areas (Dorsey 1989). If the reform had had any impact, the proportion of blacks completing secondary education should undoubtedly reflect it. Figure 4.2 is promising: between the pre- and post-reform periods, there is a marked jump in the transition rates. Following Imbens and Lemieux (2008), we plot the conditional expectation of the observed outcome for mothers in Figure 4.3 and for father in Figure 4.4. The outcome is completed schooling detrended for age in 1980 [9]. A clear discontinuity occurs in both graphs: for mothers, completed schooling jumps by nearly two years between fourteen and fifteen years of age in 1980 [10]; for fathers, the jump is about one-and-a-half extra years of schooling and is timed somewhat later.

These discontinuities are *fuzzy*, rather than sharp because the *probability* of attending secondary school could not have moved from zero to one for all individuals meeting the entry-age for Form I at the time of the reform. It is not improbable that a girl aged fifteen or sixteen in 1980 would have enrolled in Form I after the reform - particularly be-

[9] Negative correlations between schooling and age have been noted before. Using household data from South Africa, Nigeria, Côte d'Ivoire, Kenya, Burkina Faso and Ghana, Schultz (2004) describes the progress in female attainment over time in these countries as slow but continuous. By contrast, Zimbabwe underwent a very different experience because of its reform. Detrending schooling conveys the impact of that reform that much more strongly.

[10] This result is eight times as large as the effect the Indonesian school construction program created on male education during the 1970's Duflo (2001) and Duflo (2004).

cause entry-age rules were not strictly enforced - but it is much more probable that girls enrolling in Form I after the reform were just fourteen years old in 1980 or younger. The same argument applies to boys with the modifier that the discontinuity is less suggestive at fourteen years of age in 1980 than fifteen - we believe this is evidence that enrolment at a grade level tended to be more overage than age-exact for boys. Inasmuch as this rings as an assumption, the first-stage we estimate for fathers with fifteen years as the cut-off turns out to be much stronger than it is with fourteen. Thus, we choose to treat fourteen years of age in 1980 as the cut-off at which treatment probability becomes discontinuous for women and, fifteen as the equivalent for men.

The next step is where we break the parent generation in our data into two groups, the younger of which would have accumulated more schooling due to the timing of their births. This timing provides the source of exogenous variation in parental schooling which will later help establish the transmission of education to the next generation. The treatment group in mothers is chosen as the age-group of six to fourteen years in 1980 while the control consists of mothers who were fifteen to twenty-two years old at the same time. For fathers, the only difference is that the sample is broken at the cut-off age of fifteen years in 1980. Then our estimation strategy is to compare, for instance, the schooling attainment of children born to mothers aged fourteen or younger in 1980 to the schooling attainment of children born to the older mothers. As in van der Klaauw (2002), the indicator for the age cut-off serves as an excluded instrument in our Two-Stage Least-Squares (2SLS) regressions.

To account for differences in schooling attainment due to age, we standardize years of schooling of the child as follows: let s_{ic} be the years of schooling of child i that belongs to cohort c. Then,

$$\tilde{s}_{ic} = \frac{s_{ic} - \mu_c}{\sigma_c}$$

where μ_c and σ_c are the mean and standard deviation of schooling for children in cohort c. Thus, $\tilde{s}_{ic}$ represents the number of standard deviations of child schooling. The summary statistics for the standardized schooling attainment of children is reported in the bottom two panels of Table 4.1.

To obtain causal estimates of parental education on child schooling, consider the following relation between the mother and the child:

$$CS_i = \alpha + \beta MS_i + x_i'\theta + e_i \tag{4.1}$$

where CS_i is a measure of the human capital of child i. We consider four measures of CS_i: the standardized years of schooling, the delay in schooling attainment experienced by the child, whether the child ever attended school prior to the date of the census and child enrolment at the time of the census. The variable MS_i stands for the years of schooling of the mother of child i. β is the parameter of interest. Equation (4.1) includes a vector of child characteristics (x_i') such as age and sex. OLS estimates of β may be biased due to endogeneity in parental schooling and omitted variable bias.

More educated parents tend to have higher earnings, which typically increases the chances of the child being enrolled in school and going through more schooling. If the relationship between parental schooling and earnings is not modeled, then $\hat{\beta}_{OLS}$ is likely to be biased upward. Similar biases would occur from the omission of other (un)observable parent characteristics that could affect the schooling attainment of the child, such as academic aptitude, health, household wealth, and home or community endowments. Formally, Equation (4.1) could be modified as follows:

$$CS_i = \alpha_1 + \beta\, E[MS_j | A_j, x_j] + f(A_j) + \theta\, x_i' + e_i \tag{4.2}$$

$$E[MS_j | A_j, x_j] = \alpha_2 + \delta\, 1\{A_j \leq \overline{A}\} + g(A_j) + \gamma\, x_j \tag{4.3}$$

where A_j is the age in 1980 of mother j, $1\{\cdot\}$ is an indicator function for when the mother's age in 1980 is less than or equal to the cutoff $\overline{A} = 14$, and functions $f(\cdot)$ and $g(\cdot)$ are flexible polynomial in the mother's age in 1980.

Based on Hahn, Todd, and Van der Klaauw (2001), a consistent estimation of β by 2SLS will rely on two assumptions: first, that maternal schooling is discontinuous at the cut-off (this is testable); second, that $f(\cdot)$ and $g(\cdot)$ are locally continuous at the age cut-off. When the functions $f(\cdot)$ and $g(\cdot)$ are correctly specified, they will capture all potential effects of age on the mother and children's education in the neighborhood of the cut-off. The cut-off indicator can then be used as an excluded instrument to achieve a consistent estimate of β. Following van der Klaauw (2002), $f(\cdot)$ and $g(\cdot)$ will be represented by piece-wise linear approximations [11]. Thus,

$$f(A_j) = \psi_1\, A_j + \psi_2\, (\overline{A} - A_j)\, 1\{A_j \leq \overline{A}\}$$

[11] See Ferraz and Finan (2009) for an example in other contexts.

A similar expression holds for $g(\cdot)$. A more flexible approximation with $K > 1$ can be written as

$$f(A_j) = \sum_{k=1}^{K} \psi_{1k}\, {A_j}^k + \sum_{k=1}^{K} \psi_{2k}\, (\overline{A} - A_j)^k\, 1\{A_j \leq \overline{A}\}$$

Our basic results were unaffected by approximations where $K > 1$ in $f(\cdot)$ and $g(\cdot)$.

4.5 Results

4.5.1 Intergenerational Transmission of Education

OLS Results

In Table 4.2, columns marked OLS show the estimates for β as described in Equation (4.1). Four child outcome variables are regressed against the education of the mother and father separately. The OLS results show a positive association between parental schooling and all four child-schooling outcomes. In the first cell of the table, an additional year of education gained by the mother is associated with an increase in the child's education up to 9% of a standard deviation above the mean. This is equivalent to an increase of 0.1 ($= 0.093 \times 0.962$) years of schooling for the average child. An additional year of mother's schooling is also associated with a decline of 0.1 years in delay. The OLS estimates of the marginal effects of father's schooling are as great as the effects of mother's schooling. Still, these effects are strictly positive associations between parental and child schooling, notwithstanding their high statistical significance.

Two-Stage Least-Squares Estimates

The 2SLS estimates of β as described in Equation (4.2) appear in the even columns of Table 4.2. Each of the four different child schooling outcomes are regressed against the child's age, the child's sex, the parent's education and a piece-wise linear approximation on the parent's age in 1980. The cut-off point at age fourteen (fifteen) in 1980 serves as the instrument for mother's (father's) education. To show the intuition behind this approach, consider the Wald estimator for this effect using the raw data as follows:

$$\hat{\beta}_{\text{Wald}} = \frac{\overline{CS}(z=1) - \overline{CS}(z=0)}{\overline{MS}(z=1) - \overline{MS}(z=0)} = \frac{0.110 - 0.067}{8.4 - 7.6} = 0.05375$$

where $\overline{CS}$ and $\overline{MS}$ denote the mean schooling of children and mothers respectively, and $z = 1$ if the mother's age in 1980 was either thirteen or fourteen and $z = 0$ if it was fifteen or sixteen years instead. The Wald estimate suggests a positive effect, where an additional year of education earned by the mother increases the standardized schooling of the child by about 5 percent of a standard deviation, or approximately 0.049 years.

This magnitude is confirmed in the first TSLS result reported in Table 4.2: the coefficient of maternal schooling is significant and equal to 0.043 implying a marginal effect of $0.043 \times 0.926 = 0.04$ additional child years of schooling. The coefficient is smaller than the OLS estimate, suggesting that the observed covariance between mother and child schooling included the effect of unobservables affecting *both* education variables positively. Thus, a part of the intergenerational transmission from maternal education is a transmission of academic ability, among others. An additional year of mother's education continues to impact delay negatively and significantly, while unsurprisingly, the effects on ever-attendance and current enrolment disappear.

2SLS estimates of the transmission of father's schooling are presented in the lower panel of Table 4.2. These are larger than the corresponding OLS estimates, highly significant and predict that an additional year of father's schooling would raise the child's by 0.14 years (14% of a standard deviation $\times 0.99$). The effect on school enrolment is not significant after instrumenting for father's education but ever-attendance still shows a significant rise of 3 percentage points and delay in child schooling a decline of 0.13 years.

4.5.2 Impact of the Reform on Parental Schooling

The first-stage results for the intergenerational transmissions discussed above are in Panel [1] of Table 4.3. The estimate in each cell corresponds to δ in Equation (4.3). The full sample employs parents in the age-group of six to twenty-two years in 1980, while the shorter age-span restricts this interval to ten to twenty years.

At the point of discontinuity, mothers just younger than fourteen in 1980 had about 0.8 more years of education than those who were fifteen or older. This coefficient is statistically significant at 1% and the F-statistic strongly rejects the possibility that age is a weak instrument for maternal schooling. In the fathers sample, the cut-off predicts a difference of 0.65 years of schooling between the treated and control groups. Adding rainfall to the regressions produces very little change in the point estimates or their significance. Similarly, slicing the data by the sex of the child yields very comparable estimates of the instrument's effect on parental schooling. In the columns under the title of shorter age-span, the instrument is checked for weakness in a more restricted neighborhood around the cut-off. The coefficient is only slightly smaller, ranging from 0.54 to 0.60, but highly significant and the F-statistics still reject the null at 1%.

4.5.3 Robustness Checks

Whites and Foreign-born Blacks

The validity of an RD research design rests partially on the extent to which it excluded a subpopulation it did not mean to treat. In each of Figures 4.6a, 4.6b, 4.7a and 4.7b, the schooling levels of white and foreign parents are fairly smooth across the cut-off points. The schooling attainment of white women in particular (Figure 4.6b) suggests that younger white mothers completed fewer years of schooling, on an average than before the reform. Less of a trend is apparent in the case of white men or foreign-born parents.

Table 4.4 bears out these basics: the instrument is significant and negative in the full sample of white mothers but collapses to zero in the restricted sample. The F-statistics reassure us that the cut-off carries no importance in the context of white schooling levels generally, both prior to and after the reform. Likewise, Table 4.5 presents the first-stage regressions for foreign-born black parents. In the full sample, the instrument does not detect any significant difference between the schooling levels of younger and older foreign-born parents. Where the shorter age-span was used, the instrument bears a significant but negative coefficient, so even if the reform affected the schooling of the younger cohorts, it had the opposite effect of lowering schooling attainment.

Rainfall

Our next robustness check against the 2SLS estimates is to examine if there exists an unmodeled cohort effect correlated with parental schooling at the cut-off age in 1980. For example, many in the younger cohort may have been born in "better" years, thereby inheriting better health, wealth or any other characteristic helpful in acquiring more human capital. If so, the instrument may be confounding the effect of parental schooling with other parental attributes that matter to the child's human capital accumulation.

It is possible that being born in a relatively good rainfall year increases the probability of receiving more schooling. Richardson (2007) documents the strong association between Zimbabwe's rainfall and the growth of its GDP per capita during the period, 1959-2001. Alderman, Hoddinott, and Kinsey (2006) analyze the effect of successive droughts in 1982 and '83 on child nutrition in Zimbabwe, and show that long-term consequences include lower education, greater delay in enrolment, and poorer adult health. Hoddinott (2006) shows that droughts also have an adverse impact on household assets.

Figure 4.5 shows rainfall between 1959 and 1985 from an average of 38 stations across the country. The data is displayed as deviations from the mean, standardized by the standard deviation over the whole period (both moments were calculated over 1959-2001.) Women aged fourteen in 1980 would have been born in 1966, a year with average rainfall. However, all years but three in the periods 1959-'65 and 1967-'72 received rainfall noticeably below-average. Thus most women, whether younger or older than 14 in 1980, were likely to have seen below-average rainfall in their years of birth. The first-stage is re-estimated with rainfall in the parent's year of birth in panel [2] of Table 4.3. It leaves the instrument unaltered in significance and magnitude while rainfall itself is insignificant. The first-stage relationships between the timing of the reform and parental schooling are still very strong. Table 4.6 shows the results of re-estimating Equation (4.2) in the shorter age-span controlling for rainfall at the time of birth. The estimates for the transmission of parental schooling are robust for sons but the mother-daughter transmissions vanish.

District Fixed Effects

Within our census sample, there are 137 districts in Zimbabwe, classified into urban, rural and town councils. We employ a fixed effect for each district. The district fixed effects should filter out any non-randomness in the spread of the reform (e.g. intensity of school construction) from parental schooling. In Table 4.7, the results of including these fixed effects are shown in the first three panels for mothers and in the next three for fathers. Once again, the transmission of maternal schooling loses precision and significance whereas the 2SLS estimates of the transmission of father's schooling are robust to the introduction of fixed effects.

Cohort Effects

The FRD design must respect the exclusion restriction in order to be internally valid: i.e. parents a few years older and younger than the cut-off age in 1980 must not differ significantly in observable or unobservable characteristics that could affect child schooling outcomes. While it is never possible to exhaust the set of observable (even less unobservable) characteristics, we show below that our results are unlikely to be driven by several cohort effects. We start by restricting the two samples of parents closer to their respective cut-off ages in 1980. In the first panel of Table 4.8, we choose mothers aged twelve and thirteen in 1980 for our treated group and consider mothers aged fifteen to sixteen years for the control. All the first-stages are weak, as evidenced by the F-statistic, which in turn drives the imprecise estimates of the maternal transmission of schooling. The second panel performs a similar exercise for fathers: fathers in the interval of $[13, 17]$ years of age in 1980 are used but fathers at the cut-off age of fifteen in 1980 are omitted. In this sub-sample, only the child's Z-score for schooling experiences a small but significant, positive impact from increased father-schooling. In further contrast to panel [1], all the first-stages for fathers are strong.

The last two panels of Table 4.8 also compare units just to the left and right of the cut-off. The age-group of 12-13 in 1980 is set up as the treated group, 16-17 year-olds in 1980 are chosen as the control, panel [3] reports the results from this restriction for mothers and panel [4] does the same for fathers. It turns out that this alternate restriction produces the same type of results on the intergenerational transmission of mother and father's schooling as already seen in panels [1] and [2].

Polynomials in Age Spline

The 2SLS estimates in Table 4.2 used a linear spline in the parent's age in 1980. In Table 4.9 and Table 4.10, we test if those estimates are sensitive to the functional form specified for the age-spline. These tables present the 2SLS results of parent-to-child transmission of schooling when the spline is alternately a cubic, a quadratic and a linear function in age. Columns [1] through [6] in Table 4.10 show that the transmission of father's schooling is robust to these various spline forms. The point estimates tend to be somewhat larger when more flexible spline specifications are used but they are strongly significant. The mother-child transmission of schooling fades in the presence of the nonlinear splines and it is telling that in columns [3] and [6], the linear spline is the only other significant regressor, besides the instrumented variable. Overall, the transmissions of maternal schooling have not proved very precise or significant to the robustness checks so far.

Afrobarometer

Young men and women in 1980 were not only exposed to the change in education policies in 1980 but to a new political and social environment. Independence brought the abolition of apartheid laws and raised the political power of black Zimbabweans. Exposure to this new environment could have had different effects on pre-teenagers than older individuals [12]. Thus, young pupils in 1980 were probably exposed to a "combined treatment": expanded schooling opportunities as well as a different (and possibly *higher*) sense of citizenship. If this were true, the 2SLS estimates in Table 4.2 contain an upward bias.

The Population Census does not collect information that would help rule out this confound. However, in the past ten years, a new set of nationally representative surveys, the Afrobarometer [13] has been engaged in gathering data on individual values and attitudes

[12] Marx, Ko, and Friedman (2009) show that the White-Black disparity in a verbal exam found during the summer of 2008 vanishes for those taking the exam right after President Obama's nomination acceptance speech and just after his election victory.

[13] The Afrobarometer is a non-partisan and independent project conducted in different African nations, where it seeks to gauge the social, economic and political atmosphere of the countries. The Afrobarometer Network is led jointly by the Institute for Democracy in South Africa (IDASA), the Centre for Democracy and Development in Ghana (CDD-Ghana), and Michigan State University (MSU). The 1999 Zimbabwean survey was conducted a few months before political turmoil mounted over the national parliamentary elections of

towards democracy, economic life, the quality of governance, engagement in civil society, and citizenship in several African countries. Zimbabwe has figured in all four rounds so far (surveys were conducted in 1999, 2003, 2005 and 2009). We use data from the 1999 and 2003 rounds. Each survey is administered to a nationally representative sample of 1200 individuals from the universe of over-18 and eligible-to-vote adults. We extracted data pertaining to political values, community participation and questions of identity for Zimbabwean men and women in the ages of six to 22 years in 1980.

Figure 4.8 presents for a pooled sample of parents from the 1999 and 2003 rounds, various indicators of citizenship and political involvement plotted against the parent's age in 1980. The five (binary) indicators are: (a) strongly agreeing with the statement "identify with being/proud to be Zimbabwean" (b) never attend a demonstration or protest march (c) voting in the 1996 elections (d) being close to a political party and (e) close to the ruling party (ZANU-PF) conditional on being close to some party. The graphs show that the covariates are smooth across the cut-off for the most part; minor but insignificant discontinuities are noticeable in Figure 4.8a and Figure 4.8e. Tables 4.11 and 4.12 confirm this basic inference in separate regressions for mothers and fathers respectively. Each covariate is regressed against the cut-off and a piece-wise linear function in age. Column [4] in Table 4.12 shows that the cut-off is just significant at the 10% level for men considering themselves close to a political party. The cut-off is insignificant in all the other regressions, with the F-statistic vouching for its weakness in general. Therefore, it is unlikely that the intergenerational transmission of schooling is driven by cohort-differences in political views

2000.

and values.

4.5.4 Intergenerational Transmissions by Children's Sex

Table 4.13 presents the transmission of schooling from parents to children after slicing the sample by sex in each generation. Columns [1] through [4] show results from the full sample of parents and columns [5] through [8] list the 2SLS estimates from the restricted sample. The top two panels display the marginal effect of an additional year of maternal schooling on the education of sons and daughters respectively. It is clear that only the standardized schooling and delay variables bear any kind of effect and the estimate are qualitatively similar in the full sample. The effect vanishes for daughters in the shorter sample, and doubles for sons; however, as the various robustness checks demonstrated, the transmission of maternal schooling is not robust.

The point estimates for the father's transmission to the children are close to the baseline estimate of 14% of a standard deviation in Table 4.2. They are also highly significant. In the shorter age-span, the estimates grow larger for sons and somewhat smaller for daughters; throughout, the first-stage relationship is strong. The conclusion that emerges is that increased schooling, in the younger school-going cohorts of the eighties, did lead to a significant though small increase in the schooling outcomes of the next generation. In the relationship between fathers' and children's schooling, this effect is both precise and especially robust. However, increased schooling in the younger cohort of mothers has had less direct impact on the human capital accumulation of the next generation.

4.5.5 Mechanisms

Quantity and quality of education

So far we have examined the effect of the reform in terms of just the quantity of education. Issues of quality are obviously important too, in determining why or how much the reform affected education in Zimbabwe. To begin with, quality is relevant simply as a confounding variable in the intergenerational transmission. Suppose the reform increased not merely the schooling opportunities for blacks, but the quality as well (in terms of curriculum, class-size, teacher-pupil ratio, racial integration even). Then omitting quality from the regressions biases the estimate of the quantity of schooling transmission upwards.

Several studies suggest that quality declined after Independence even as the reform increased the number of schools. Edwards (1995) remarks that 1984 was the last year with "good" quality outcomes in the educational system - at least in primary education. Nhundu (1992) cites documents from the Ministry of Education on the decline in quality after 1984: school enrolment during the 1980s occurred "faster than classrooms and teacher's houses could be built" (p.87). Further, a significant number of secondary schools had been built as extensions of extant primary schools. Schools also adopted "hot seating" (i.e. double- or multiple-shifts), where the length of the school day is reduced to accommodate a larger number of students (p.88). Dorsey (1989) and Nhundu (1992) state that after Independence, growth in enrolment was not matched by growth in teacher-staff. Student-teacher ratios rose and the proportion of qualified teachers declined. In 1980, 36.1% of teachers were untrained, almost all of whom were employed at primary schools. In 1984, 41.9% of teachers were untrained and 83.2% of them were working at primary schools. By 1988, 27.8% of untrained teachers were handling secondary level education.

An indirect measure of lowered quality is student performance. As mentioned in Nhundu (1992) and Mackenzie (1988), the reform guaranteed to all students entering Grade 1, eleven years of education. At the end of the eleventh year (Form IV), students sit for the "O"-level exams. Admission to lower Form VI is conditional on passing five or more 'O'-level courses (English language included). Assuming student ability did not differ significantly across cohorts, results on these exams can be used to infer the decline in the quality of instruction. In 1981, exam-takers had started secondary school before Independence, and the pass-rate was 70.8% and 2% of the failing group did not pass a single course. In 1984, when the first post-Independence cohort took the exams, only 22.2% passed all five subjects. Among those who failed, 38.1% failed in all subjects. In 1988, the last year analyzed by Nhundu (1992), only 10.2% passed the 'O'-level exams and 42.6% of the failing group failed all subjects. Revealingly, the number of 'O'-level takers grew by 2,253% between 1981 and 1988, and the failure rate increased by 7,220% (p. 88) during the same period. These figures narrate a sharp decline in the quality of education in Zimbabwe. Our findings of a significant, positive but small transmission of education from parents to children must be interpreted as a lower-bound: it occurs *despite* the low quality of education received by younger parents in our sample.

Pathways for the transmission of parental schooling

What are the mechanisms behind the intergenerational transmission of education from mothers to children? Possible avenues explored in the literature include female employment, fertility decisions and the marriage market. Women with more education could increase their labor force participation and their hours of work as the opportunity cost of their time rises with higher wages. The Population Census used in this paper contains information on women's employment at the time of the census as well as their marital status, age-at-first-birth, total pregnancies, the number of children born and living with the parent. The education level of the women's partners or husbands can also be obtained. We select three variables - age-at-first-birth, children ever born and partner's completed schooling - to test if they serve as mechanisms for the transmission of maternal schooling to children.

Table 4.14 reports two-stage least-squares regression results in six columns, three using the full sample and three with the restricted sample. Column [1] shows that an additional year of mother's schooling increases her age-at-first-birth by 0.32 years, a coefficient significant at the 1% level (the mean age at which mothers had their first birth was 19.5 years). The estimate increases slightly to 0.36 (column [2]) when the sample is restricted to mothers in the age-group of 10-20 in 1980. Columns [3] and [4] follow up with the effect of increased female schooling on the number of children ever born to a mother. The mean is about 5, so a coefficient of -0.116 to -0.064 predicts a 1.3% to 2% fall in fertility among younger women. Thus, a mother completing eleven years of education is likely to have one child fewer than a mother who only went to one year of primary school. Columns [5] and [6] discuss the association between women and their husband's schooling. Positive assortative mating is clearly in evidence here: one additional year of female schooling causes women to marry husbands with at least an extra half-year of education. Since the average woman has a partner with education at least up to Form II, younger women are marrying men with at least ten years of education. Thus, one pathway through which mother's schooling might have increased schooling of the children is positive assortative mating.

4.6 Conclusions

Despite the common view that educating women is a viable mechanism for economic development, there are not many studies that explore the impact of increased female schooling in one generation on the human capital accumulation of the next. Most existing studies of this kind are based on developed countries. This paper uses the dramatic changes in the educational policies of a poor country to understand the intergenerational transmission of schooling. Notable changes emanating from the 1980 reform in Zimbabwe include an impressive rate of school construction in secondary education, the elimination of racial barriers and primary school fees and legislation committing the government to providing at least eleven years of education for all. The reform evoked a massive increase in enrolment rates at all levels. In addition, the primary-to-secondary transition rate increased three-fold between the Grade 7 graduating classes of 1979 and 1980. Thus, girls (boys) aged fourteen (fifteen) or less in 1980 had a much higher probability of enroling in secondary school compared to those over fourteen (fifteen) years in 1980. This discontinuity in treatment probability is at the heart of the FRD design that allows us to measure the causal impact of higher parental schooling on child schooling.

Age in 1980 serves as the excluded instrument in our 2SLS regressions of child schooling. Using data from the 2002 census, we find that children born to more educated mothers and fathers tend to have more schooling. The effects are small but positive; and robust where father-to-child transmissions are concerned. Although maternal transmissions are relatively imprecise, they do not mean that mother's education fails to affect the child's human capital; rather, the pathways through which maternal schooling benefits children are varied and interacting.

These results are important in the light of policy discussions on what breaks the intergenerational transmission of *poverty*. In particular, they recommend the conditional cash transfer (CCT) approach to fostering human capital in poor countries. In countries with such programs, the transfers are typically greater if parents send girls to school. Given the gender disparity in the parent-to-child transmissions here, a strong case can be made for CCTs in Zimbabwe. A conditional cash transfer also promises returns over more than one generation of children. To truly complete the story of intergenerational transmissions of schooling, it will be necessary to consider how other dimensions of human capital come into play as well. It will be intriguing to see, for instance, if the Behrman, Murphy, Quisumbing, and Yount (2009) Guatemala finding of the positive effects of higher maternal cognitive ability on children holds in Zimbabwe.

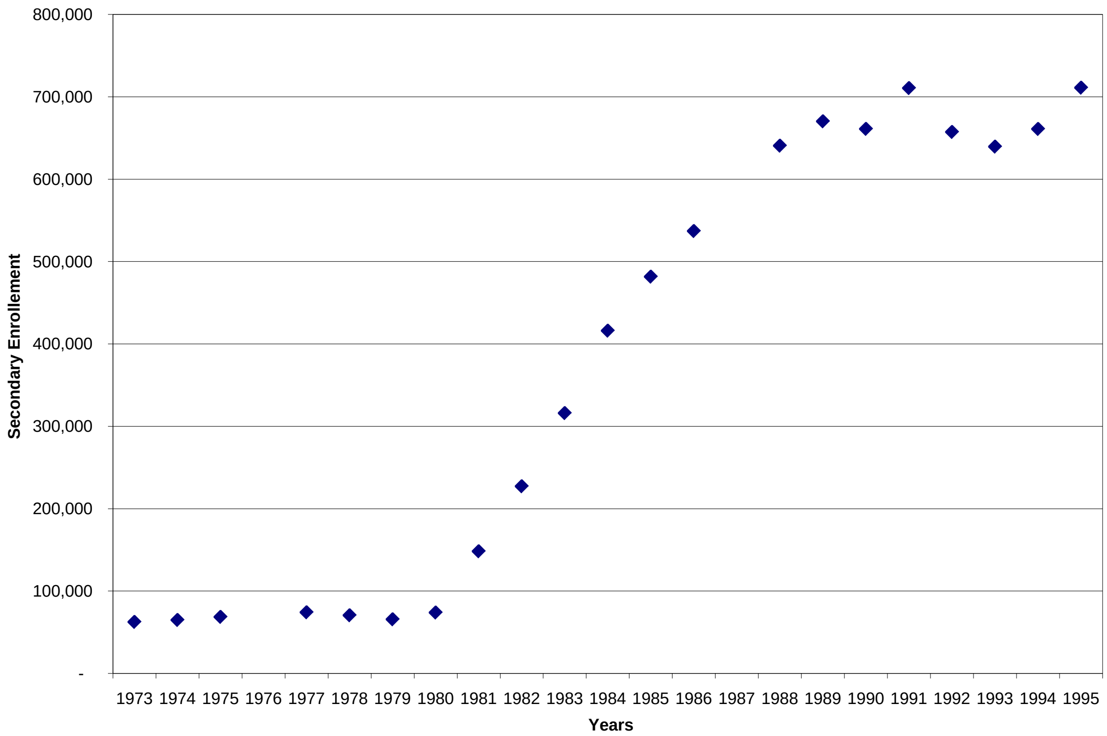


Figure 4.1: Secondary School Enrolments in Zimbabwe: 1973-1995

Source: United Nations, *Statistical Yearbook*, 1975, 1980, 1982, 1984, 1985-1989, 1992, 1994, 1995, and 1997.

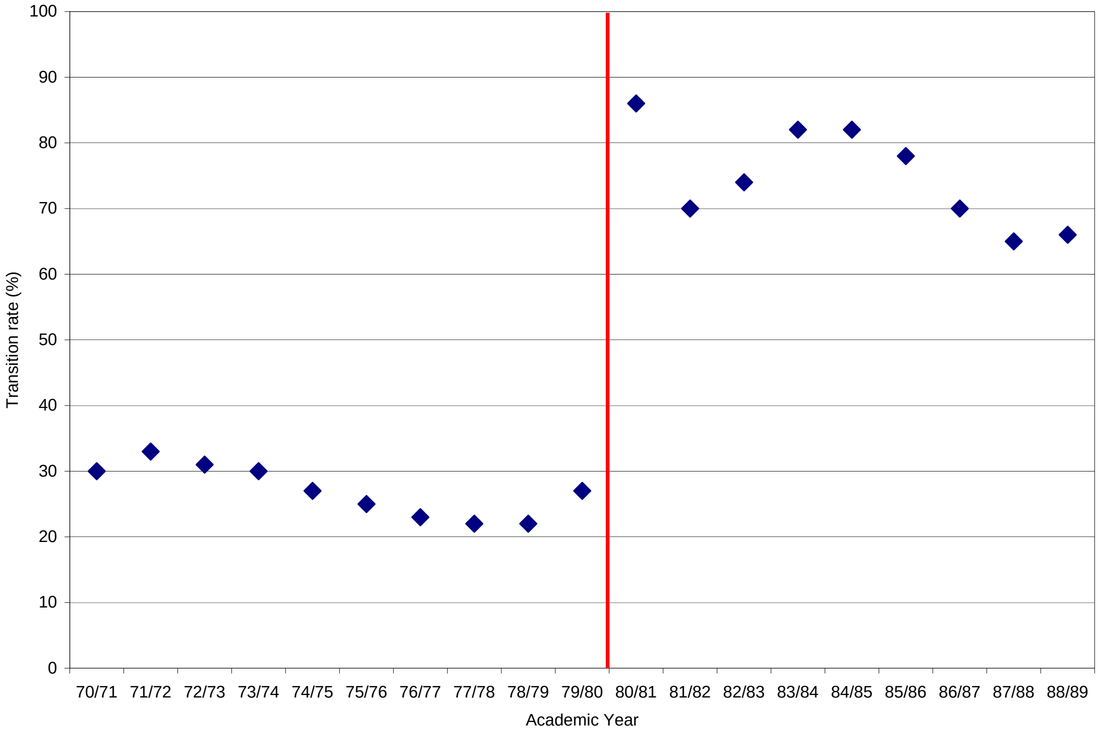


Figure 4.2: Grade 7 to Form 1 Transition Rates: 1970/71-1988/89

Source: Riddell and Nyagura (1991) Table 1.1. Notes: Grade 7 is the last year of primary education and Form I is the first year of secondary education.

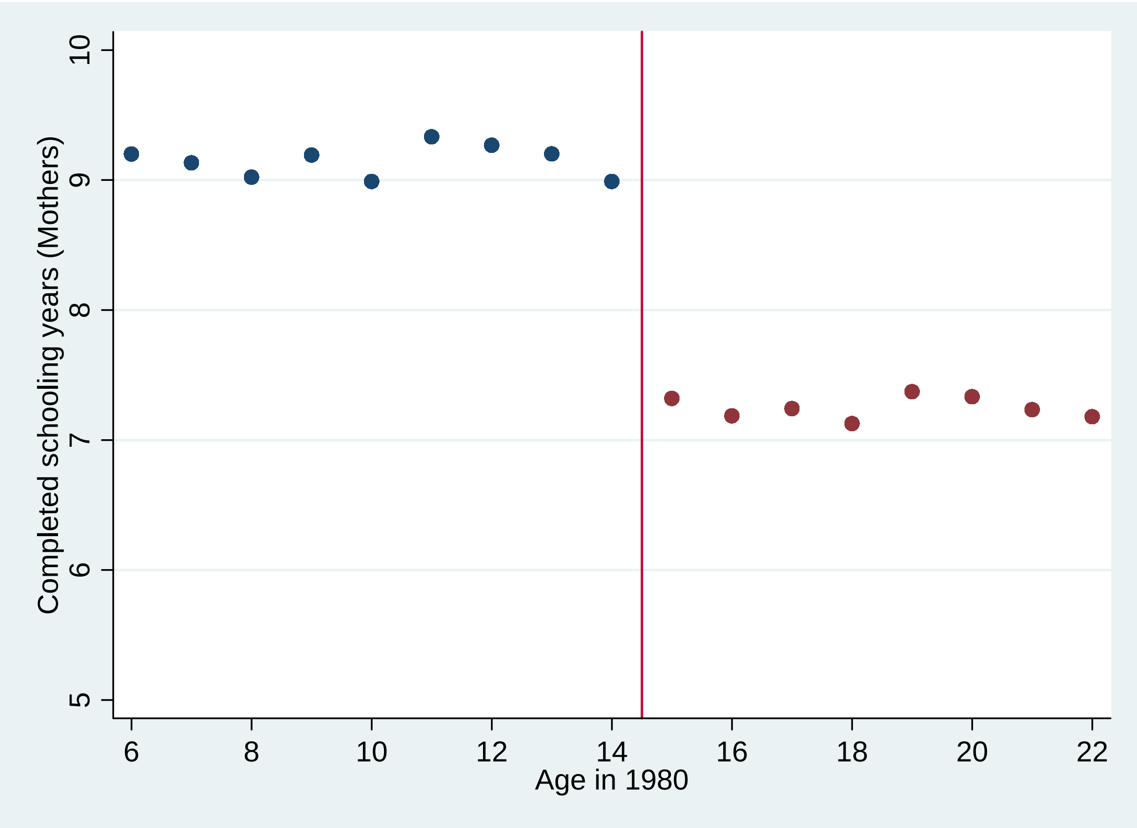


Figure 4.3: Mother's Years of Schooling by Age in 1980

Notes: Circles show mean years of detrended schooling for women who were six to 22 years of age in 1980 and whose children are aged six through 15 years in Census 2002. The vertical line represents the cut-off for treatment.

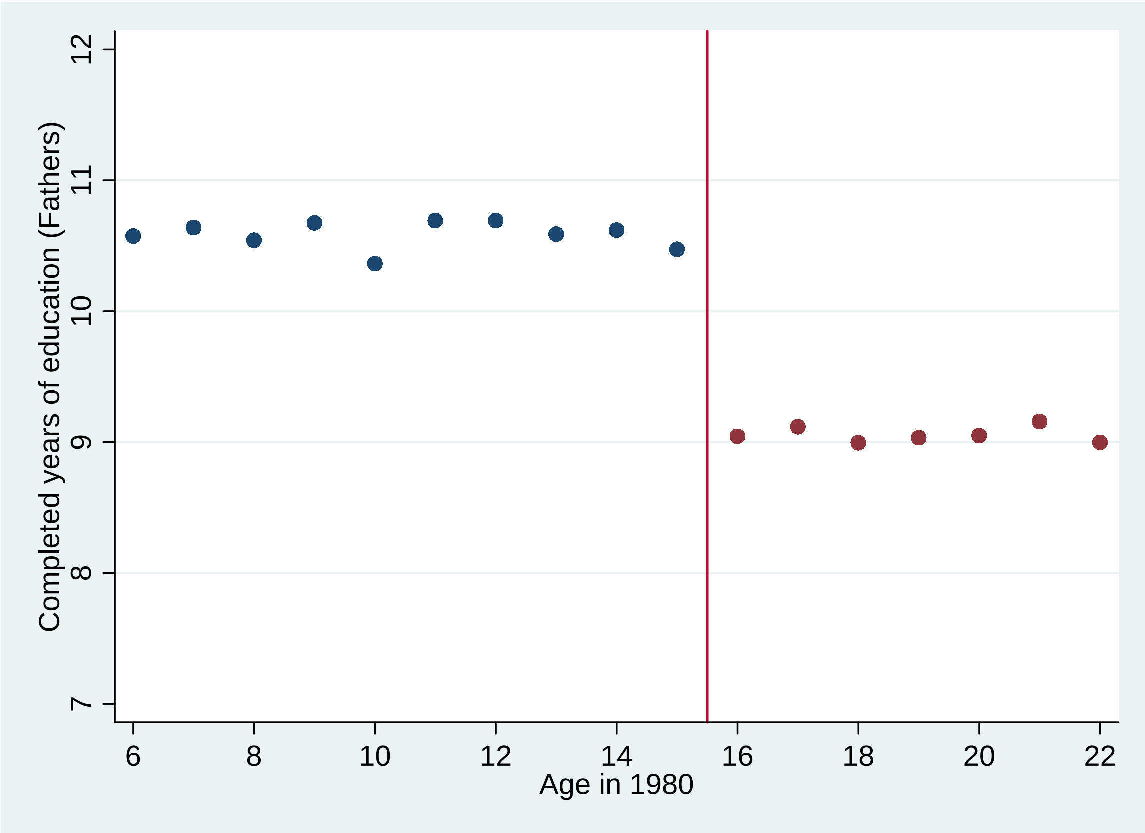


Figure 4.4: Father's Years of Schooling by Age in 1980

Notes: Circles show mean years of detrended schooling for men who were six to 22 years of age in 1980 and have children aged six through 15 years in Census 2002. The vertical line represents the cut-off for treatment.

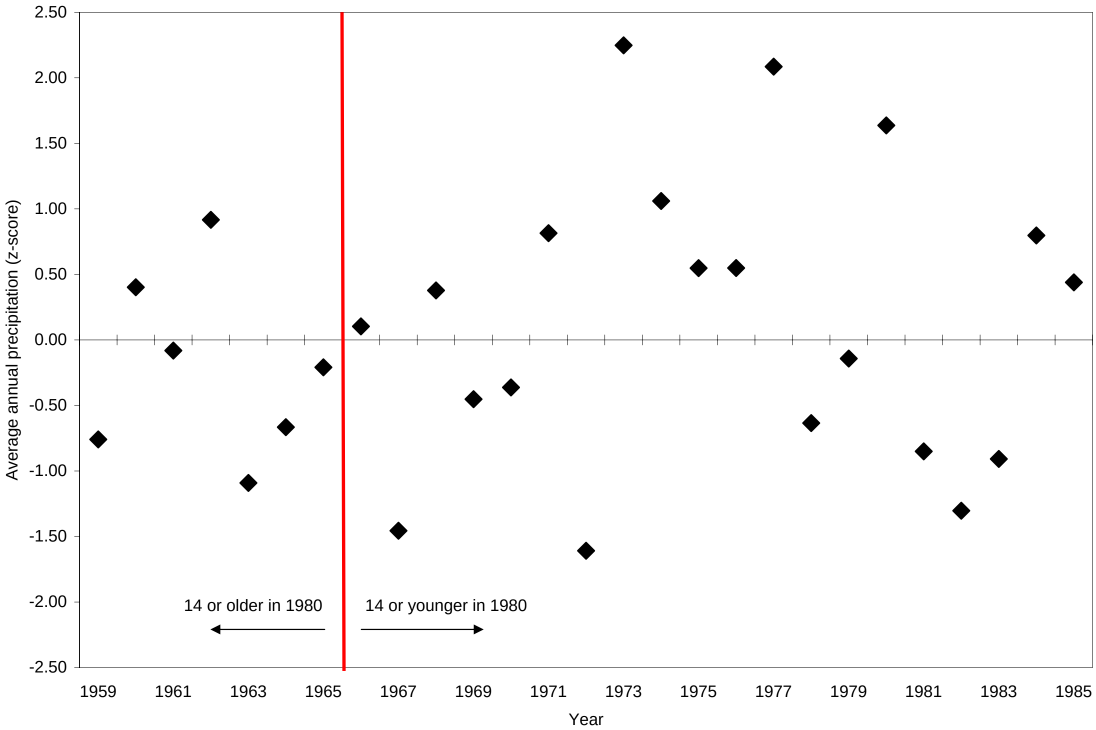


Figure 4.5: Annual rainfall: 1959-1985

Source: Zimbabwe Meteorological Service Department. Notes: Annual rainfall comes from a sample of 38 rainfall stations across the country; data was provided by Craig Richardson. A given year, such as 1970, refers to the 1970-1971 crop-year. It is measured as deviations from the mean relative to the standard deviation. Both moments were obtained from the 1959-2001 series.

Table 4.1: **Summary Statistics of Children, Mothers and Fathers in the Sample**

Variable	Units	Mean	SD	Min	Max	Observations
		Children's Sample				
Girl	Binary	0.493	0.500	0.0	1.0	187743
Age	Years	10.384	2.853	6.0	15.0	187743
Ever attended school	Binary	0.947	0.224	0.0	1.0	187743
In school†	Binary	0.959	0.199	0.0	1.0	177774
Delay	Binary	0.681	0.466	0.0	1.0	163618
		Mother's Sample				
Age	Years	35.435	4.650	28.0	44.0	100463
Schooling	Years	7.921	3.166	0.0	16.0	100463
Age at first birth	Years	19.505	3.129	12.0	39.0	100463
Children ever born	Number of children	4.816	2.128	1.0	15.0	100463
Child's Schooling	Z-score	0.118	0.962	-4.893	2.758	95810
		Father's Sample				
Age	Years	37.100	4.432	28.0	44.0	58430
Schooling	Years	9.471	3.305	0.0	16.0	58430
Child's Schooling	Z-score	0.092	0.990	-4.893	2.758	54802

Notes: The sample in the topmost panel is composed of children from six to 15 years of age, with at least one parent born before the reform and alive in Census 2002. The mothers and fathers in the samples are aged from six to 22 years in 1980.
† Conditional on attending school in the past.

Table 4.2: **Intergenerational Transmission of Schooling**

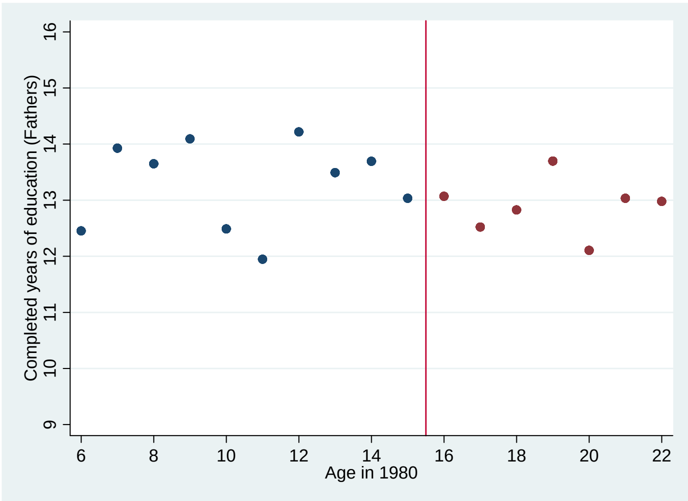
Dependent	Child's Schooling							
Variables:	Years		Delay		In school		Ever Attended	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
Mother-All	0.093	0.043	-0.108	-0.038	0.005	-0.001	0.006	0.002
	[0.001]***	[0.015]***	[0.001]***	[0.018]**	[0.000]***	[0.003]	[0.000]***	[0.003]
Observations	95810	95810	88325	88325	96224	96224	100463	100463
F-test		338.92		306.44		340.30		337.06
Father-All	0.093	0.136	-0.107	-0.130	0.004	0.005	0.008	0.029
	[0.001]***	[0.026]***	[0.001]***	[0.027]***	[0.000]***	[0.004]	[0.000]***	[0.006]***
Observations	54802	54802	48619	48619	55058	55058	58430	58430
F-test		127.42		124.45		127.49		142.00

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. The full sample is composed of children aged six to fifteen years in 2002 with either one or both parents alive in 2002 and belonging to the age-group of six to twenty-two years in 1980. All regressions control for the child's age and sex, and use a linear spline in the parent's age in 1980. The instrument in the two-stage least-squares regressions is the point of discontinuity in treatment-probability, which occurs at the 1980 age of fourteen years for mothers and fifteen for fathers. The reported F-statistics refer to these excluded instruments.

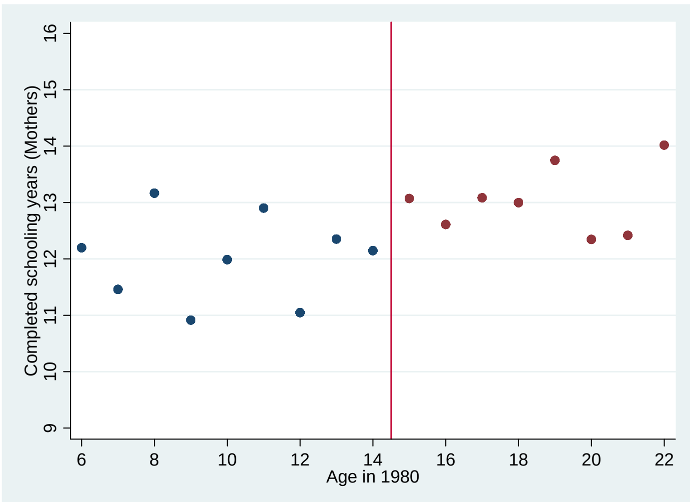
Table 4.3: **First-Stage**

Dependent Variable	Parent's Schooling			
	Full Sample		Shorter Age-span	
	Mothers	Fathers	Mothers	Fathers
All children	0.782	0.645	0.583	0.565
	[0.043]***	[0.056]***	[0.053]***	[0.067]***
Observations	95810	54802	65594	41408
F test	326.96	134.55	122.74	71.23
All children	0.749	0.690	0.581	0.587
	[0.046]***	[0.059]***	[0.054]***	[0.069]***
Rainfall	0.000	-0.042	0.003	-0.033
	[0.010]	[0.017]**	[0.021]	[0.027]
Observations	92186	51108	65594	41408
F test	269.63	135.18	116.83	71.85
Sons	0.766	0.655	0.566	0.553
	[0.061]***	[0.076]***	[0.074]***	[0.092]***
Observations	47902	29648	32865	22167
F test	157.77	74.75	58.42	36.25
Daughters	0.797	0.654	0.600	0.544
	[0.061]***	[0.077]***	[0.075]***	[0.093]***
Observations	47908	28782	32729	21722
F test	169.29	71.46	64.43	34.53

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. All first-stage estimates are reported from parent-child samples with non-missing values for years of child schooling. The full sample holds parents in the age-group of six to 22 years in 1980. The shorter age-span refers to samples where the parent's age in 1980 is restricted to ten to twenty years. Each coefficient is a different estimate of δ in Equation (4.3). The F-statistics correspond to the null, δ=0. Rainfall is expressed in number of standard deviations from the mean for the period, 1959-2001.



(a) Years of Schooling by Age in 1980 - *White Men*



(b) Years of Schooling by Age in 1980 - *White Women*

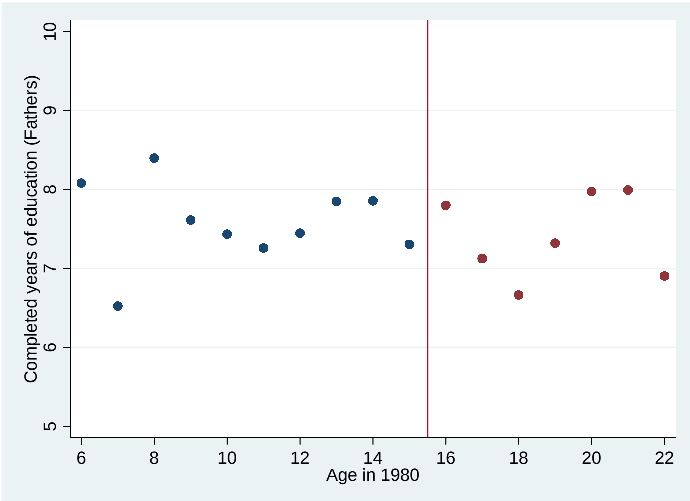
Figure 4.6: Mean Schooling of White Zimbabweans

Notes: In Figures 4.6a and 4.6b, circles show the sample mean years of schooling for white men and women, respectively. These men and women were aged six to 22 years in 1980 with children in the ages of six through 15 in Census 2002. The vertical lines represent the cut-off for treatment; for women, this cut-off is approximately between ages 14 and 15 in 1980, while for men, it occurs between 15 and 16 years of age in 1980. Whites in Zimbabwe must not have been affected *positively* by the educational reforms.

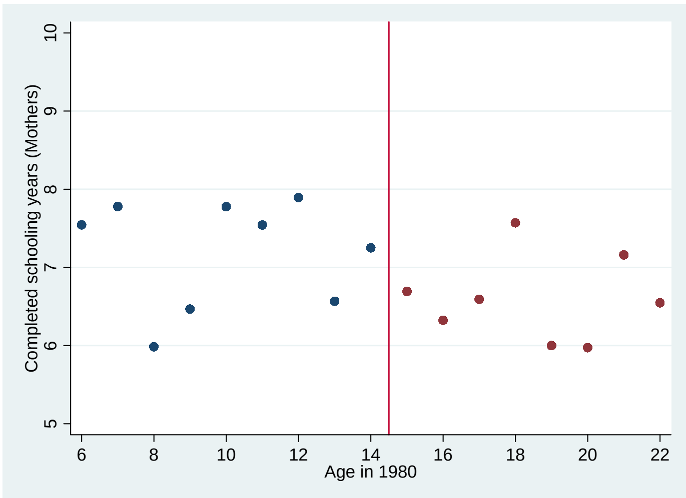
Table 4.4: **First-Stage Results for White Zimbabweans**

	Parent's Schooling			
Dependent	Mothers		Fathers	
Variables:	Full Sample	Shorter Age-span	Full Sample	Shorter Age-span
$1\{x \leq \overline{A}\}$	-1.774	0.025	1.006	1.119
	[0.680]***	[0.792]	[0.791]	[1.021]
x	0.008	0.238	0.051	0.020
	[0.142]	[0.213]	[0.150]	[0.285]
$(\overline{A} - x) * 1\{x \leq \overline{A}\}$	-0.055	-0.565	0.025	-0.119
	[0.169]	[0.332]*	[0.181]	[0.355]
Observations	263	215	195	137
F test: cut-off=0	6.81	0.00	1.62	1.20
Prob > F	0.01	0.97	0.20	0.28

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. All first-stage results are reported for the sample of children with non-missing values for years of schooling. The full sample contains parents aged 6 to 22 years in 1980 and the shorter age-span restricts them to the ages of 10 to 20 years in 1980. and x refers to age in 1980 while $\overline{A}$ in $1\{x \leq \overline{A}\}$ refers to the cut-off age at which the reform could have induced a discontinuity in schooling; this age is 14 and 15 years in 1980 for women and men respectively. The F-statistics come from the test for the null that the coefficient of $1\{x \leq \overline{A}\}$=0.



(a) Years of Schooling by Age in 1980 - *foreign-born Black Men*



(b) Years of Schooling by Age in 1980 - *foreign-born Black Women*

Figure 4.7: Mean Schooling of Foreign-born black Zimbabweans

Notes: In Figures 4.7a and 4.7b respectively, circles show the sample mean years of schooling for foreign-born black men and women in Zimbabwe, aged from six to 22 years in 1980 and having children in the ages of six through 15 in Census 2002. The vertical lines represent the cut-off for treatment; for women, this is between ages 14 and 15 in 1980, and for men, it is between the ages of 15 and 16 years in 1980.

Table 4.5: **First-Stage Results for Foreign-born black Zimbabweans**

	Parent's Schooling			
Dependent	Mothers		Fathers	
Variables:	Full Sample	Shorter Age-span	Full Sample	Shorter Age-span
$1\{x \leq \overline{A}\}$	-0.466	-1.393	-0.569	-1.324
	[0.476]	[0.549]**	[0.422]	[0.520]**
x	-0.181	-0.300	-0.132	-0.322
	[0.081]**	[0.119]**	[0.071]*	[0.132]**
$(\overline{A} - x) * 1\{x \leq \overline{A}\}$	-0.209	0.059	-0.158	-0.179
	[0.100]**	[0.180]	[0.097]	[0.170]
Observations	1086	760	1086	837
F test: cut-off=0	0.96	6.44	1.82	6.49
Prob > F	0.33	0.01	0.18	0.01

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. All first-stage results are reported for the sample of children with non-missing values for years of schooling. The full sample contains parents aged 6 to 22 years in 1980 and the shorter age-span restricts them to the ages of 10 to 20 years in 1980. and x refers to age in 1980 while $\overline{A}$ in $1\{x \leq \overline{A}\}$ refers to the cut-off age at which the reform could have induced a discontinuity in schooling; this age is 14 and 15 years in 1980 for women and men respectively. The F-statistics come from the test of the null that the coefficient of $1\{x \leq \overline{A}\}$=0.

Table 4.6: **Robustness Check: Rainfall**

Dependent Variables:	Standardized Years [1]	Delay [2]	In School [3]	Ever Attended [4]
Mothers-All	0.067	-0.072	0.001	-0.001
	[0.026]***	[0.029]**	[0.005]	[0.005]
Observations	65594	61268	65877	68398
F-test	116.87	112.03	115.35	110.95
Mothers-Sons	0.096	-0.090	0.001	0.002
	[0.036]***	[0.039]**	[0.006]	[0.007]
Observations	32865	30739	32986	34273
F-test	59.80	61.96	59.29	56.52
Mothers-Daughters	0.038	-0.050	0.000	-0.004
	[0.037]	[0.043]	[0.007]	[0.007]
Observations	32729	30529	32891	34125
F-test	57.12	50.35	56.12	54.47
Fathers-All	0.153	-0.157	0.007	0.012
	[0.035]***	[0.037]***	[0.005]	[0.008]
Observations	41408	37086	41604	43889
F-test	68.90	66.82	68.92	69.58
Fathers-Sons	0.190	-0.192	0.009	0.011
	[0.049]***	[0.052]***	[0.007]	[0.011]
Observations	20851	18703	20947	22167
F-test	38.12	35.46	37.85	36.08
Fathers-Daughters	0.111	-0.118	0.005	0.013
	[0.052]**	[0.052]**	[0.008]	[0.011]
Observations	20557	18383	20657	21722
F-test	30.78	31.28	31.04	33.34

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. The estimate reported in each cell is the coefficient of parent years of schooling. Each estimate comes from a different regression. All regressions control for rainfall, measured in number of standard deviations from the mean over the period, 1959-2001. Samples contain parents aged ten to twenty years in 1980 with children aged six through fifteen years in 2002.

Table 4.7: **Robustness Check: Birth-District Fixed Effects**

Dependent Variables:	Standardized Years [1]	Delay [2]	In School [3]	Ever Attended [4]
Mothers-All	0.030	-0.021	-0.001	0.002
	[0.016]*	[0.019]	[0.003]	[0.003]
Observations	95810	88325	96224	100463
Mothers-Sons	0.030	-0.010	0.003	0.002
	[0.023]	[0.028]	[0.004]	[0.005]
Observations	47902	44244	48090	50272
Mothers-Daughters	0.029	-0.029	-0.006	0.002
	[0.023]	[0.027]	[0.004]	[0.004]
Observations	47908	44081	48134	50191
Fathers-All	0.130	-0.123	0.005	0.031
	[0.027]***	[0.028]***	[0.004]	[0.006]***
Observations	54802	48619	55058	58430
Fathers-Sons	0.154	-0.152	0.010	0.031
	[0.036]***	[0.040]***	[0.005]*	[0.008]***
Observations	27712	24600	27835	29648
Fathers-Daughters	0.099	-0.088	-0.001	0.029
	[0.040]**	[0.040]**	[0.007]	[0.009]***
Observations	27090	24019	27223	28782

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%; ** at 5% and *** at 1%. The estimate reported in each cell is the coefficient of parent years of schooling from a 2SLS regression with fixed effects for the parent's district of birth. Samples contain parents aged ten to twenty years in 1980 with children aged six through fifteen years in 2002.

Table 4.8: **Treatment Effect in the Immediate Neighborhood of the Cut-off**

Dependent Variables:	Standardized Years [1]	Delay [2]	In School [3]	Ever Attended [4]
Panel [1]: Mothers, age 12-13 in 1980 vs. 15-16				
	0.052	-0.052	-0.005	-0.025
	[0.162]	[0.159]	[0.031]	[0.039]
Observations	24204	22578	24323	25228
F test	2.95	3.78	2.82	2.47
Panel [2]: Fathers, age 13-14 in 1980 vs. 16-17				
	0.155	-0.148	0.003	0.014
	[0.089]*	[0.100]	[0.012]	[0.017]
Observations	15464	13884	15540	16355
F test	11.18	9.00	12.40	13.77
Panel [3]: Mothers, age 12-13 in 1980 vs. 16-17				
	0.143	-0.229	-0.007	-0.012
	[0.135]	[0.168]	[0.023]	[0.027]
Observations	24723	23104	24847	25773
F test	4.37	3.95	4.78	4.48
Panel [4]: Fathers, age 12-13 in 1980 vs. 16-17				
	0.163	-0.236	0.017	-0.02
	[0.097]*	[0.112]**	[0.014]	[0.027]
Observations	15482	13771	15553	16393
F test	10.62	8.64	11.12	8.46

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%; ** at 5% and *** at 1%. Each entry is an instrumental variable estimate of the effect of an additional year of parental schooling on a specific education outcome of the child. The estimation sample is in the title of each panel, the younger age-group always representing the treated set. The treatment cut-off is at fourteen years of age in 1980 for women and fifteen for men. The F-statistics are for the null hypothesis that the cut-off is not a weak instrument for completed parental schooling.

Table 4.9: **Robustness Check: Polynomials in Age Spline - Mothers**

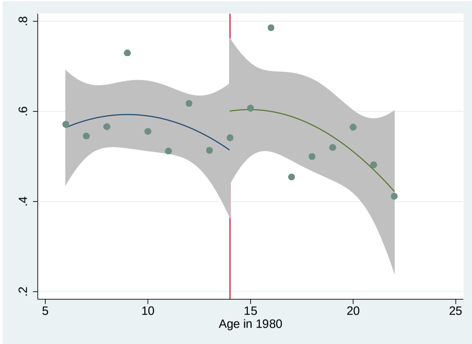
Dependent	Standardized Years			Delay			In School			Ever Attended		
Variables:	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Nature of Spline:	*Cubic*	*Quadratic*	*Linear*	*Cubic*	*Quadratic*	*Linear*	*Cubic*	*Quadratic*	*Linear*	*Cubic*	*Quadratic*	*Linear*
Schooling	-0.019	-0.807	0.057	0.028	1.002	-0.058	0.007	-0.012	0.000	0.003	-0.258	-0.001
	[0.100]	[1.803]	[0.025]**	[0.128]	[2.421]	[0.029]**	[0.018]	[0.102]	[0.005]	[0.021]	[1.240]	[0.005]
Age in 1980	0.094	-1.238	0.009	-0.163	1.520	-0.010	0.178	-0.021	0.000	0.132	-0.344	-0.001
	[1.151]	[2.565]	[0.008]	[1.449]	[3.465]	[0.009]	[0.196]	[0.144]	[0.001]	[0.246]	[1.670]	[0.002]
Linear spline	-0.017	0.049	-0.014	0.021	-0.055	0.018	-0.002	-0.001	0.000	0.000	0.028	0.000
	[0.052]	[0.119]	[0.006]**	[0.049]	[0.142]	[0.007]**	[0.009]	[0.007]	[0.001]	[0.011]	[0.106]	[0.001]
Square age	-0.010	0.030		0.014	-0.037		-0.010	0.001		-0.008	0.008	
	[0.170]	[0.062]		[0.217]	[0.084]		[0.029]	[0.003]		[0.036]	[0.040]	
Quadratic spline	0.004	-0.129		-0.007	0.159		0.006	-0.002		0.007	-0.038	
	[0.034]	[0.260]		[0.029]	[0.350]		[0.006]	[0.015]		[0.009]	[0.177]	
Cube age	0.000			0.000			0.000			0.000		
	[0.003]			[0.004]			[0.001]			[0.001]		
Cubic spline	-0.003			0.004			0.000			-0.001		
	[0.015]			[0.016]			[0.003]			[0.004]		
Observations	65594	65594	65594	61268	61268	61268	65877	65877	65877	68398	68398	68398

Notes: Robust standard errors in brackets. * refers to statistical significance at 10%; ** significance at 5% and *** significance at 1%. All regressions instrument for mother's schooling using the cut-off of fourteen years for her age in 1980, and control for the child's age and sex. The full sample is composed of children aged six to 15 years in 2002 with either one or both parents alive in 2002. Mothers are in the ages of ten to twenty years in 1980.

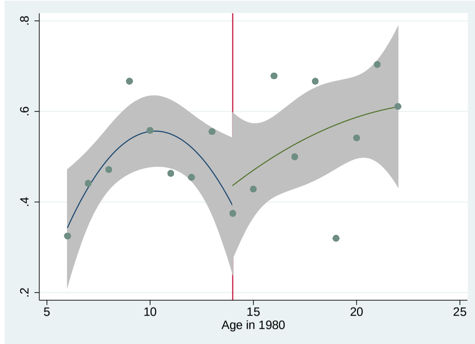
Table 4.10: **Robustness Check: Polynomials in Age Spline - Fathers**

Dependent Variables:	Standardized Years [1]	[2]	[3]	Delay [4]	[5]	[6]	In School [7]	[8]	[9]	Ever Attended [10]	[11]	[12]
Nature of Spline:	*Cubic*	*Quadratic*	*Linear*	*Cubic*	*Quadratic*	*Linear*	*Cubic*	*Quadratic*	*Linear*	*Cubic*	*Quadratic*	*Linear*
Schooling	0.193	0.252	0.150	-0.206	-0.271	-0.147	0.008	0.002	0.005	0.012	-0.005	0.011
	[0.084]**	[0.119]**	[0.035]***	[0.089]**	[0.133]**	[0.036]***	[0.012]	[0.016]	[0.005]	[0.016]	[0.023]	[0.007]
Age in 1980	-0.487	0.214	0.042	0.469	-0.265	-0.043	0.164	-0.006	0.001	0.117	-0.035	0.001
	[1.154]	[0.173]	[0.013]***	[1.171]	[0.212]	[0.014]***	[0.159]	[0.024]	[0.002]	[0.221]	[0.033]	[0.003]
Linear spline	0.001	0.032	0.010	0.000	-0.036	-0.007	0.005	-0.001	0.000	-0.013	-0.007	0.000
	[0.104]	[0.030]	[0.015]	[0.150]	[0.036]	[0.016]	[0.013]	[0.004]	[0.002]	[0.030]	[0.006]	[0.003]
Square age	0.034	-0.004		-0.034	0.005		-0.009	0.000		-0.007	0.001	
	[0.145]	[0.004]		[0.160]	[0.005]		[0.020]	[0.001]		[0.029]	[0.001]	
Quadratic spline	0.007	0.012		-0.010	-0.015		0.000	0.000		0.006	-0.002	
	[0.025]	[0.012]		[0.019]	[0.014]		[0.004]	[0.002]		[0.004]*	[0.002]	
Cube age	-0.001			0.001			0.000			0.000		
	[0.003]			[0.003]			[0.000]			[0.001]		
Cubic spline	-0.002			0.002			0.000			-0.001		
	[0.002]			[0.003]			[0.000]			[0.001]		
Observations	41408	41408	41408	37086	37086	37086	41604	41604	41604	43889	43889	43889

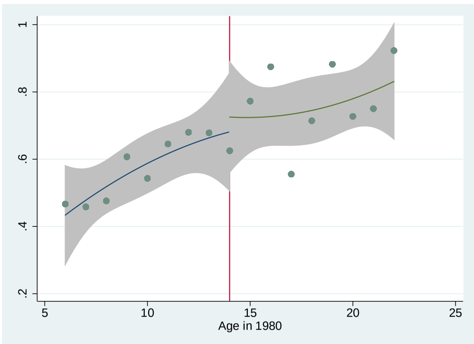
Notes: Robust standard errors in brackets. * refers to statistical significance at 10%; ** significance at 5% and *** significance at 1%. All regressions instrument for father's schooling using the cut-off of 15 years for his age in 1980 and control for the child's age and sex. The full sample is composed of children aged six to 15 years in 2002 with either one or both parents alive in 2002. Fathers are in the ages of ten to twenty years in 1980.



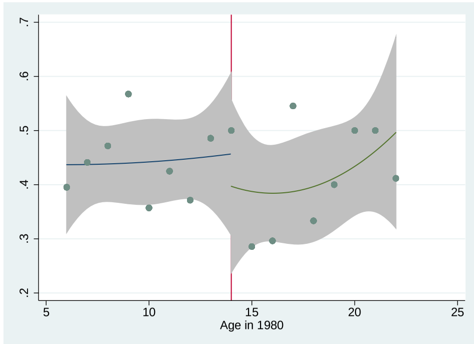
(a) Proud to be Zimbabwean



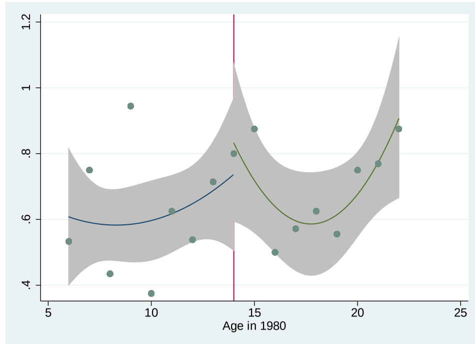
(b) Would never attend a demonstration



(c) Voted in last elections



(d) Close to a political party



(e) Close to ZANU-PF

Figure 4.8: Values and Political Affiliations by Age in 1980

Source: 1999 and 2003 Zimbabwean Afrobarometer. Notes: In each figure, circles show the average values for black Zimbabweans by their age in 1980. A pair of quadratic regressions on either side of the cutoff and 95% confidence intervals also appear. The cutoff point - chosen as age fourteen in 1980 - is represented by the vertical lines in the figures . All variables are binary and figure (a) refers to those who strongly agree with the statement “Proud to be Zimbabwean” in 1999 and “Identify with being Zimbabwean” in 2003.

Table 4.11: **Women's Citizenship and Political Involvement**

Dependent Variables:	Proud to be Zimbabwean [1]	Voted in 1996 Elections [2]	Would Never join a Protest march [3]	Close to a political Party	
				Any [4]	ZANU-PF [5]
$1\{x \leq 14\}$	-0.018	0.117	0.147	-0.050	-0.057
	[0.139]	[0.145]	[0.137]	[0.138]	[0.208]
x	-0.018	0.036	0.029	0.009	-0.004
	[0.022]	[0.023]	[0.021]	[0.022]	[0.031]
$(14-x)*1\{x \leq 14\}$	-0.021	0.019	0.030	0.029	-0.006
	[0.028]	[0.029]	[0.027]	[0.027]	[0.040]
Observations	260	210	264	263	104
F test	0.02	0.65	1.16	0.13	0.08
Prob > F	0.90	0.42	0.28	0.72	0.78

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. The running variable x refers to the woman's age in 1980. In column (1), the dependent variable takes the value one, for those who strongly agree with the statement "Proud to be Zimbabwean" in the 1999 survey and "Identify most with being Zimbabwean, rather than any single ethnic group" in 2003. The sample is composed of women aged between six and 22 in 1980 from the 1999 and 2003 rounds of the Zimbabwe Afrobarometer Survey.

Table 4.12: **Men's Citizenship and Political Involvement**

Dependent Variables:	Proud to be Zimbabwean	Voted in 1996 Elections	Would Never join a Protest march	Close to a political Party: Any	Close to a political Party: ZANU-PF†
	[1]	[2]	[3]	[4]	[5]
$1\{x \leq 15\}$	-0.215	-0.097	-0.156	0.227	0.181
	[0.134]	[0.139]	[0.133]	[0.135]*	[0.455]
x	-0.061	-0.024	-0.020	0.046	-1.373
	[0.028]**	[0.028]	[0.027]	[0.028]	[1.074]
$(15 - x) * 1\{x \leq 15\}$	-0.056	-0.068	-0.033	0.042	-0.269
	[0.030]*	[0.031]**	[0.030]	[0.030]	[0.247]
Observations	280	215	278	279	106
F test	2.58	0.49	1.39	2.84	0.16
Prob > F	0.11	0.49	0.24	0.09	0.69

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. The running variable x refers to the man's age in 1980. In column (1), the dependent variable takes the value one, for those who strongly agree with the statement "Proud to be Zimbabwean" in the 1999 survey and "Identify most with being Zimbabwean" in 2003. The sample is composed of men aged between six and 22 in 1980 from the 1999 and 2003 rounds of the Zimbabwe Afrobarometer Survey.
† A quadratic spline in father's age was used in the regression of this covariate.

Table 4.13: **Transmission of Parental Schooling by Sex of the Child**

	Child's Schooling							
Dependent	Full Sample				Shorter Age-span			
Variables:	Years	Delay	In school	Ever Attended	Years	Delay	In school	Ever Attended
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Mothers-Sons	0.045	-0.030	0.003	0.002	0.088	-0.080	0.001	0.001
	[0.022]**	[0.026]	[0.004]	[0.004]	[0.036]**	[0.040]*	[0.006]	[0.007]
N	47902	44244	48090	50272	32865	30879	32986	34273
F-test	164.91	151.92	165.09	162.30	58.63	60.39	57.79	54.62
Mothers-Daughters	0.042	-0.046	-0.005	0.003	0.028	-0.035	-0.001	-0.002
	[0.021]**	[0.025]**	[0.004]	[0.004]	[0.035]	[0.041]	[0.007]	[0.007]
N	47908	44081	48134	50191	32729	30529	32891	34125
F-test	174.08	154.50	175.29	174.86	64.23	56.07	63.01	61.32
Fathers-Sons	0.159	-0.159	0.011	0.029	0.188	-0.186	0.008	0.012
	[0.036]***	[0.039]***	[0.005]**	[0.008]***	[0.049]***	[0.052]***	[0.007]	[0.011]
N	27712	24600	27835	29648	20851	18703	20947	22167
F-test	60.79	52.83	60.08	64.23	30.93	27.45	30.74	29.14
Fathers-Daughters	0.112	-0.101	-0.001	0.029	0.109	-0.105	0.002	0.011
	[0.037]***	[0.036]***	[0.006]	[0.008]***	[0.051]**	[0.051]**	[0.008]	[0.011]
N	27090	24019	27223	28782	20557	18383	20657	21722
F-test	56.55	57.6	55.69	64.75	28.27	27.4	27.63	30.74

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. Children in all regressions are between six and 15 years of age. The full sample contains parents aged six to 22 years in 1980 while the columns headed "shorter age-span" include parents aged from ten to twenty in 1980. All regressions use a linear spline in the parent's age and control for the child's age. The reported F-statistics refer to the point of discontinuity in treatment probability with respect to parent's age.

Table 4.14: **Mechanisms for Transmission of Mother's Schooling**

Dependent Variables:	Age at First Birth [1]	[2]	Number of Children [3]	[4]	Husband's Schooling [5]	[6]
Woman's Schooling	0.318	0.361	-0.116	-0.064	0.605	0.450
	[0.052]***	[0.087]***	[0.032]***	[0.054]*	[0.069]***	[0.170]**
Shorter Age-span	N	Y	N	Y	N	Y
Observations	100463	68398	100463	68398	40905	20090
Mean of Dependent Variable	19.42	19.42	5.01	5.01	9.63	9.63

Notes: Robust standard errors in brackets. * indicates statistical significance at 10%, ** at 5% and *** at 1%. The full sample contains mothers aged between six and 22 years in 1980 and the shorter age-span refers to women aged ten to twenty years in 1980.

Chapter 5

Conclusions

Education and health are two of the most vital components of human capital. Across countries, research has revealed that the quantity of human capital affects the level of per capita income and the rate at which total factor productivity increases. Within a country, human capital investment frequently improves the well-being of households: the poorest of families tend to experience longer life expectancies, reduced mortality and a chance to escape poverty. The policies that design these opportunities are often founded on a combination of observational analysis and received wisdom. This dissertation explores the links between human capital and development at the household level.

The Income-Calorie Relationship in Rural Maharashtra

Chapter 2 begins by describing a nutrition production function. It points out that a major strand of the development economics literature has regarded nutrient intake as synonymous with nutrition, a not unreasonable approach in extremely poor societies. In such societies, public policy has long sought ways to raise calorie consumption. Increasing per capita income was believed to be a particularly effective tool – several papers have since examined if this tool has really worked. Many of these are reviewed in Chapter 2, especially those relying on nonparametric estimations of the income (or expenditure) elasticity of calories.

In Chapter 3, I look at the level and pattern of nutrition in Maharashtra, an Indian state where, as in the rest of the country, a paradox of rising incomes and falling calorie consumption has posed a puzzle for the better part of two decades. Some recent research has suggested that the best explanation for this puzzle lies in declining calorie *requirements*. However, there has been no direct verification of this hypothesis yet. I examine if the evidence from rural Maharashtra is capable of supporting that hypothesis, as at least one line of thinking.

Using data from the National Sample Survey, I document the changes in the calorie availability in rural households, their pattern of food expenditure and the price paid for food between 1983 and 2004. I find a downward drift of the calorie Engel curve as well as a steady flattening of the relationship. The local linear regression of log calories on the logarithm of household expenditure appears more concave in 2004 and lies almost entirely below the curve of 1983. At the same time, the nonparametric expenditure elasticities of calories show a decline from a median elasticity of 0.5 in 1983 to about 0.15 in 2004. In addition, I find that although the price paid for calories rose in real terms, its expenditure elasticity is relatively unchanged since 1983. The expenditure elasticity of food itself is roughly half its magnitude of 1983: about 0.33.

I rule out reverse causality in the calorie-income relationship, based on the fact that the recommended daily calorie intake costs no more than a third of the rural daily wage in 2004. The second set of robustness checks are performed using parametric regressions approximating the bivariate nonparametric model. In these regressions, controlling for household and district-level characteristics does not alter the basic elasticity estimates of 2004.

Inflation in food prices in Maharashtra was moderate over the study period, averaging about 5% through the Nineties and less than 3% in the early 2000s. In the absence of a strong upward trend in the relative price of food, substitution effects in calorie consumption are relatively small, so that most of the calorie decline cannot be ascribed to price changes. While a preference for variety accounts for a part of the decline in the lowest decile of expenditure, lower calorie requirements remains a persuasive hypothesis.

From a policy perspective, there is a two-fold implication of falling calorie needs over time: policies seeking to end hunger and deprivation through employment guarantees and income increases are less likely to succeed than in the past. Yet, Maharashtra, like the rest of India, records a poverty rate of about a quarter, and even higher rates of child undernutrition. Higher calorie consumption is a necessary goal only at the lower tail of the calorie distribution. An approach that is consistent with this realization and the evidence of a declining income elasticity needs to target only those households that still exhibit a positive, significant income elasticity for calorie consumption. A rethinking of calorie norms based on more systematic nutrition monitoring may also be required.

The Intergenerational Transmission of Schooling in Zimbabwe

Chapter 4 is an exploration of the intergenerational accumulation of schooling. It seeks to uncover a causal impact of parental schooling on the educational attainment of the child. Although the schooling levels of parents and children are strongly correlated, it is a matter of empirical investigation whether increased schooling of a mother in itself promotes the schooling of her children. Jorge and I examine this question in Zimbabwe, where a fortunate natural experiment presents itself in the education reforms implemented thirty years ago.

Zimbabwe was one of the very last African nations to win freedom from apartheid. As Rhodesia, its education system had been racially-segregated and blacks were systematically sidelined from education opportunities beyond primary school. In April 1980, the first democratically-elected Black government of Zimbabwe enacted sweeping reforms, to fulfil their promise of increasing black admission to secondary schools. A significant move was the elimination of an exam that determined transition from primary to secondary school. As a result, the reform dramatically increased the schooling opportunities of the cohort graduating from primary school in the early Eighties. This *experimental* cohort went on to attain at least a year more schooling on an average, than the cohort that finished primary school before 1980 (the control cohort). We observe both age-groups in the Census of 2002, where several in them are parents. We then establish that the reform was a relevant and valid source of exogenous variation in the schooling levels of these parents. This allows us to compare the schooling outcomes of their children and we find that the children of the experimental cohort out-perform the children of the control, i.e. the parents who were relatively unexposed to the reforms.

The analysis is based on a 10% sample of the 2002 Zimbabwe Population Census[1]. To summarize, the work in Chapter 4

1. evaluates the local average treatment effect of the educational reforms for the experimental cohort versus the control group,

2. measures the intergenerational effects of increased parental schooling by investigating the schooling outcomes of the children of the two cohorts exposed to the reforms, and

3. examines the mechanisms through which the intergenerational spillovers took hold in Zimbabwe.

After the reforms, black primary school graduates in 1980 entered secondary school at a rate four times higher than those in 1979. This *fuzzy* discontinuity in the probability of attending secondary school is exploited to test for parent-to-child transmission of education. It also suggest an instrument for the parent's schooling levels: their age in 1980. We consider four outcomes for children sampled in the ages of six to fifteen years[2]: completed years of schooling in relation to peers, delay (due to grade repetition or late entry), the probability of ever attending school and the probability of being currently enrolled in school.

The size of the human capital transmission is largest in the first two outcomes. We find that a one year increase in the mother's education causes an increase in the children's schooling by about 5 percent of a standard deviation. More specifically, an extra year of maternal schooling obtains a marginal increase of 8% in the sons' standardized schooling

[1]To the best of our knowledge, this paper is the first to provide a causal analysis of the schooling reforms, and to use this data set to do so.

[2]This is because six is the entry-age for primary school, and thirteen for secondary school

while only bringing about a 3% increase for daughters. Interestingly, father's education raises the human capital of sons and daughters more strongly than mother's schooling: an additional year of schooling for the father moves up the son's standardized score for years of schooling by 15 to 18% while raising a daughter's by 9% of a standard deviation. Although transmission by birth order cannot be investigated in our data set, these are robust estimates for the average child. We also establish that these findings are unlikely to be driven by confounding factors, as our instrument (age in completed years in 1980) does not predict events or choices correlated with the schooling level of the parents.

This paper also provides the lower bound estimate on the size of human capital transmission in Zimbabwe. We believe this is the case because supply was outmatched by the demand for secondary education, leading to quality deterioration throughout the late Eighties and the Nineties, and so dissipating the full potential impact of the reforms.

Despite the common view that educating women is a viable mechanism for economic development, few studies actually explore the degree of human capital accumulation in subsequent generations. Most such studies have also documented the results for developed countries. Given how crucial education has been shown to breaking the inter-generational transmission of poverty, our paper makes vital contributions on several levels: one, it exploits the natural experiment setting of the reforms to produce good, causal estimates on the level and rate of accumulation of education. Two, it validates conventional policy wisdom on the value of educating women. When we briefly examine the channels of human capital transmission, we find that exposure to the reforms for women is strongly correlated with reduced fertility, delayed marrying-age and assortative matching on spouses. Thus, education enhances the well-being of women, as much as her children's. Finally, the paper is also a program evaluation of the reforms, and the results affirm belief in the potential of conditional cash transfers (CCTs) as an excellent incentive to accelerate investment in human capital.

Bibliography

Abdulai, A., and D. Aubert (2004): "Nonparametric and Parametric Analysis of Calorie Consumption in Tanzania," *Food Policy*, 29(2), 113 – 129.

Alderman, H., J. Hoddinott, and B. Kinsey (2006): "Long term consequences of early childhood malnutrition," *Oxford Economic Papers*, 58(3), 450–474.

Andrabi, T., J. Das, and A. I. Khwaja (2009): "What Did You Do All Day? Maternal Education and Child Outcomes," *SSRN eLibrary*.

Antonovics, K. L., and A. S. Goldberger (2005): "Does Increasing Women's Schooling Raise the Schooling of the Next Generation? Comment," *American Economic Review*, 95(5), 1738–1744.

Aromolaran, A. B. (2004): "Household income, Women's Income Share and Food Calorie Intake in South Western Nigeria," *Food Policy*, 29(5), 507–530.

Atkinson, N. (1972): *Teaching Rhodesians: A History of Educational Policy in Rhodesia.* Longman, London.

Bajpai, N., J. Sachs, and N. Volavka (2005): "India's Challenge to Meet the Millennium Development Goals," CGSD Working Papers 24, The Earth Institute at Columbia University.

Barro, R. J. (2001): "Human Capital and Growth," *The American Economic Review*, 91(2), pp. 12–17.

Behrman, J. R., and A. B. Deolalikar (1987): "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India," *Journal of Political Economy*, 95(3), 492–507.

——— (1989): "Is Variety the Spice of Life? Implications for Calorie Intake," *The Review of Economics and Statistics*, 71(4), 666–72.

Behrman, J. R., and A. B. Deolalikar (1990): "The Intrahousehold Demand for Nutrients in Rural South India: Individual Estimates, Fixed Effects, and Permanent Income," *The Journal of Human Resources*, 25(4), pp. 665–696.

Behrman, J. R., A. B. Deolalikar, and B. L. Wolfe (1988): "Nutrients: Impacts and Determinants," *The World Bank Economic Review*, 2(3), 299–320.

Behrman, J. R., A. D. Foster, M. R. Rosenzweig, and P. Vashishtha (1999): "Women's Schooling, Home Teaching, and Economic Growth," *Journal of Political Economy*, 107(4), 682–714.

Behrman, J. R., A. Murphy, A. R. Quisumbing, and K. Yount (2009): "Are returns to mothers' human capital realized in the next generation?: The impact of mothers' intellectual human capital and long-run nutritional status on children's human capital in Guatemala," Discussion paper.

Behrman, J. R., and M. R. Rosenzweig (2002): "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?," *American Economic Review*, 92(1), 323–334.

Behrman, J. R., and B. L. Wolfe (1984): "More Evidence on Nutrition Demand : Income seems Overrated and Women'S Schooling Underemphasized," *Journal of Development Economics*, 14(1), 105 – 128.

——— (1987): "How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation?," *Journal of Econometrics*, 36(1-2), 185–204.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005): "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, 95(1), 437–449.

Bouis, H. E. (1994): "The Effect of Income on Demand for Food in Poor Countries: Are our Food Consumption Databases giving us Reliable Estimates?," *Journal of Development Economics*, 44(1), 199 – 226.

Bouis, H. E., and L. J. Haddad (1992): "Are Estimates of Calorie-Income Elasticities too high? A Recalibration of the Plausible Range," *Journal of Development Economics*, 39(2), 333–364.

Breierova, L., and E. Duflo (2004): "The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less Than Mothers?," NBER Working Papers 10513, National Bureau of Economic Research, Inc.

Chen, Y., and H. Li (2009): "Mother's education and child health: Is there a nurturing effect?," *Journal of Health Economics*, 28(2), 413 – 426.

Cleveland, W. S. (1979): "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74(368), pp. 829–836.

Currie, J., and E. Moretti (2003): "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *The Quarterly Journal of Economics*, 118(4), 1495–1532.

Dawson, P. J., and R. Tiffin (1998): “Estimating the Demand for Calories in India,” *American Journal of Agricultural Economics*, 80(3), pp. 474–481.

de Walque, D. (2009): “Parental Education and Children’s Schooling Outcomes: Evidence from Recomposed Families in Rwanda,” *Economic Development and Cultural Change*, 57(4), 723–746.

Deaton, A., and J. Drèze (2009): “Nutrition in India: Facts and Interpretations,” *Food Policy*, 44(7), 42–65.

Dorsey, B. J. (1989): “Educational Development and Reform in Zimbabwe.,” *Comparative Education Review*, 33(1), 40–58.

Duflo, E. (2001): “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment,” *American Economic Review*, 91(4), 795–813.

——— (2004): “The medium run effects of educational expansion: evidence from a large school construction program in Indonesia,” *Journal of Development Economics*, 74(1), 163–197.

Edwards, G. (1995): “Rapid educational expansion and primary school efficiency in Zimbabwe, 1983-9,” *Development Southern Africa*, 12(1), 87–95.

Edwards, G., and C. Tisdell (1989): “The Educational System of Zimbabwe Compared with Those of Selected African and Advanced Countries: Costs, Efficiency and Other Characteristics,” *Comparative Education*, 25(1), 57–76.

——— (1990): “Post-independence Trends in Education in Zimbabwe,” *South African Journal of Economics*, 58(4), 298–307.

——— (1992): “The Internal Efficiency of the Primary and Secondary Education Systems in Zimbabwe,” *Eastern Africa Economic Review*, 8(1), 12–25.

Efron, B., and R. J. Tibshirani (1993): *An Introduction to the Bootstrap*. Chapman and Hall, London.

Ferraz, C., and F. Finan (2009): “Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance,” NBER Working Papers 14906, National Bureau of Economic Research, Inc.

Gibson, J., and S. Rozelle (2002): “How Elastic is Calorie Demand? Parametric, Nonparametric and Semiparametric results for Urban Papua New Guinea,” *The Journal of Development Studies*, 38(6), 23–46.

Glewwe, P. (2002): “Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes,” *Journal of Economic Literature*, 40(2), 436–482.

GOPALAN, C. (1983): "'Measurement of Undernutrition': Biological Considerations," *Economic and Political Weekly*, 18(15), pp. 591–595.

HAHN, J., P. TODD, AND W. VAN DER KLAAUW (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–09.

HODDINOTT, J. (2006): "Shocks and their consequences across and within households in Rural Zimbabwe," *The Journal of Development Studies*, 42(2), 301–321.

IMBENS, G. W., AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615–635.

JENSEN, R. T., AND N. H. MILLER (2010a): "Do Consumer Price Subsidies Really Improve Nutrition?," Working Paper 16102, National Bureau of Economic Research.

——— (2010b): "A Revealed Preference Approach to Measuring Hunger and Undernutrition," Working Paper 16555, National Bureau of Economic Research.

——— (2010c): "Using Consumption Behavior to reveal Subsistence Nutrition," Incomplete Draft.

JHA, R., R. GAIHA, AND A. SHARMA (2009a): "Calorie and Micronutrient Deprivation and Poverty Nutrition Traps in Rural India," *World Development*, 37(5), 982 – 991.

——— (2009b): "Modelling Variety in Consumption Expenditure on Food in India," *International Review of Applied Economics*, 23(4), 503–519.

KANYONGO, G. Y. (2005): "Zimbabwes Public Education System Reforms: Successes and Challenges.," *International Education Journal*, 6(1), 65–74.

KOCHAR, A. (2005): "Can Targeted Food Programs Improve Nutrition? An Empirical Analysis of India's Public Distribution System," *Economic Development and Cultural Change*, 54(1), 203–35.

LIPTON, M. (1983): "Poverty, Undernutrition and Hunger," Discussion paper, Staff Working Papers, No. 597.

MACKENZIE, C. G. (1988): "Zimbabwe's educational miracle and the problems it has created," *International Review of Education*, 34(3), 337–353.

MANGYO, E. (2008): "Who benefits more from higher Household Consumption? The Intrahousehold Allocation of Nutrients in China," *Journal of Development Economics*, 86(2), 296–312.

MARX, D. M., S. J. KO, AND R. A. FRIEDMAN (2009): "The "Obama Effect": How a salient role model reduces race-based performance differences," *Journal of Experimental Social Psychology*, 45(4), 953–956.

MEENAKSHI, J., AND B. VISWANATHAN (2005): "Calorie Deprivation in Rural India between 1983 and 1999-2000," in *The Great Indian Poverty Debate*, ed. by A. Deaton, and V. Kozel, pp. 570–583. Macmillan, New Delhi.

MENON, P., A. DEOLALIKAR, AND A. BHASKAR (2009): *The India State Hunger Index: Comparisons Of Hunger Across States.* International Food Policy Research Institute, New Delhi.

MISHRA, S., AND M. PANDA (2006): "Growth and Poverty in Maharashtra," IGIDR working papers, Indira Gandhi Institute of Development Research, Mumbai.

NATIONAL SAMPLE SURVEY ORGANISATION (2007): "Nutritional Intake in India 2004-05, Report 513," Discussion paper, Ministry of Statistics and Programme Implementation, Government of India, New Delhi, Mimeo.

NHERERA, C. M. (2000): "Globalisation, Qualifications and Livelihoods: the case of Zimbabwe.," *Assessment in Education: Principles, Policy and Practice*, 7(3), 335–362.

NHUNDU, T. J. (1992): "A Decade of Educational Expansion in Zimbabwe: Causes, Consequences, and Policy Contradictions.," *Journal of Negro Education*, 61(1), 78–98.

O'CALLAGHAN, M., AND R. AUSTIN (1977): *Southern Rhodesia: The Effects of a Conquest Society on Education, Culture and Information.* Syndenhams Printers, Dorset.

OREOPOULOS, P., AND M. E. PAGE (2006): "The Intergenerational Effects of Compulsory Schooling," *Journal of Labor Economics*, 24(4), 729–760.

OREOPOULOS, P., M. E. PAGE, AND A. H. STEVENS (2003): "Does Human Capital Transfer from Parent to Child? The Intergenerational Effects of Compulsory Schooling," NBER Working Papers 10164, National Bureau of Economic Research, Inc.

PALMER-JONES, R., AND K. SEN (2001): "On Indian Poverty Puzzles and Statistics of Poverty," *Economic & Political Weekly*, 36(3), 211–17.

PATNAIK, U. (2004): "The Republic of Hunger," *Social Scientist*, 32(9/10), pp. 9–35.

——— (2007): "Neoliberalism and Rural Poverty in India," *Economic & Political Weekly*, 42(30), 3132–50.

PITT, M. M. (1983): "Food Preferences and Nutrition in Rural Bangladesh," *The Review of Economics and Statistics*, 65(1), pp. 105–114.

PITT, M. M., AND M. R. ROSENZWEIG (1985): "Health and Nutrient Consumption Across and Within Farm Households," *The Review of Economics and Statistics*, 67(2), pp. 212–223.

PLANNING COMMISSION (1979): "Report of the Task Force on Projection of Minimum Needs and Effective Demand," Discussion paper, Perspective Planning Division, Government of India, New Delhi.

——— (1993): "Report of the Expert Group on Estimation of Proportion and Number of Poor," Discussion paper, Perspective Planning Division, Government of India, New Delhi.

——— (2004): "Poverty Estimates for 2004-05," Discussion paper, Perspective Planning Division, Government of India, New Delhi, Press Note.

Plug, E. (2004): "Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees," *American Economic Review*, 94(1), 358–368.

Poleman, T. T. (1981): "Quantifying the Nutrition Situation in Developing Countries," *Food Research Institute Studies*, 18(1), pp. 1–58.

Ravallion, M. (1990): "Income Effects on Undernutrition," *Economic Development and Cultural Change*, 38(3), pp. 489–515.

Reserve Bank of India (1999): "Handbook of Statistics on the Indian economy," Discussion paper, Mumbai.

Richardson, C. J. (2007): "How much did droughts matter? Linking rainfall and GDP growth in Zimbabwe," *African Affairs (London)*, 106(424), 463–478.

Riddell, A. R., and L. M. Nyagura (1991): "What causes differences in achievement in Zimbabwe's secondary schools?," Policy Research Working Paper Series 705, The World Bank.

Riddell, R. (1980): *Education for employment.* Mambo Press, Gweru, Zimbabwe.

Roy, N. (2001): "A Semiparametric Analysis of Calorie Response to Income Change across Income Groups and Gender," *Journal of International Trade & Economic Development*, 10(1), 93–109.

Sacerdote, B. (2002): "The Nature and Nurture of Economic Outcomes," *American Economic Review*, 92(2), 344–348.

Schiff, M., and A. Valdes (1990): "Nutrition: Alternative Definitions and Policy Implications," *Economic Development and Cultural Change*, 38(2), 281–92.

Schultz, T. P. (2004): "Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education," *Journal of African Economies*, 13(02), ii95–ii14.

Shimokawa, S. (2010): "Asymmetric Intrahousehold Allocation Of Calories In China," *American Journal of Agricultural Economics*, 92(3), 873 – 888.

Skoufias, E. (2003): "Is the Calorie-Income Elasticity Sensitive to Price Changes? Evidence from Indonesia," *World Development*, 31(7), 1291 – 1307, Economic Crises, Natural Disasters, and Poverty.

STIGLITZ, J. E. (1976): "The Efficiency Wage Hypothesis, Surplus Labour, and the Distribution of Income in L.D.C.s," *Oxford Economic Papers*, 28(2), 185–207.

STRAUSS, J. (1982): "Determinants of Food Consumption in rural Sierra Leone : Application of the Quadratic Expenditure System to the Consumption-Leisure component of a Household-firm Model," *Journal of Development Economics*, 11(3), 327 – 353.

STRAUSS, J., AND D. THOMAS (1995): "Human resources: Empirical Modeling of Household and Family Decisions," in *Handbook of Development Economics*, ed. by H. Chenery, and T. Srinivasan, vol. 3 of *Handbook of Development Economics*, chap. 34, pp. 1883–2023. Elsevier.

SUBRAMANIAN, S., AND A. DEATON (1996): "The Demand for Food and Calories," *Journal of Political Economy*, 104(1), 133–62.

TIFFIN, R., AND P. J. DAWSON (2002): "The Demand for Calories: Some Further Estimates from Zimbabwe," *Journal of Agricultural Economics*, 53(2), 221–232.

TIMMER, C. P., AND H. ALDERMAN (1979): "Estimating Consumption Parameters for Food Policy Analysis," *American Journal of Agricultural Economics*, 61(5), pp. 982–987.

VAN DER KLAAUW, W. (2002): "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 43(4), 1249–1287.

ZVOBGO, C. J. M. (1981): "African Education in Zimbabwe: The Colonial Inheritance of the New State, 1899-1979," *Issue: A Journal of Opinion*, 11(3/4), 13–16.