Essays on risk, nutrition and poverty dynamics

by

Ingo W. Outes Leon

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

UNIVERSITY OF OXFORD

Oriel College
January 2013

Essays on risk, nutrition and poverty dynamics

by

Ingo W. Outes Leon

Oriel College

Submitted to the Department of Economics
on 29th January 2013, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This collection of essays explores questions on risk, poverty dynamics and early child development. The first essay analyses issues of cognitive development in early childhood. An extensive literature documents linkages between nutrition and cognition, but few studies show correlations to be causal. Building on Glewwe et al (2001) and Alderman et al (2006), we re-examine this relationship and find a negative cognitive effect of early nutritional disinvestments among a sample of Peruvian pre-schooling siblings.

The second essay examines nutritional catch-up growth among poor Ethiopian children and the role household wealth in this process. We find that nutritional catch-up patterns vary across socioeconomic groups: average catch-up growth in height-for-age is near perfect among children in better-off households. However, for children above 5 years of age, household wealth no longer affects height velocity. Our findings suggest that nutritional remediation is effective early on in life, but that this window of opportunity might already be closed by the age of five.

The third essay contributes to the literature on poverty traps and thresholds. In spite of their popularity, little empirical evidence supports their existence. Studies testing poverty traps often are restrictive in their methodology (e.g. Lokshin and Ravallion, 2004) or fail to provide causal estimates (Lybbert et al, 2004). This essay tests and finds evidence of multiple equilibria in income dynamics in rural India using a 30-year long panel. Simulations suggest the presence of two equilibria: a stable high-income equilibrium and a low-level unstable saddle point. To the best of our knowledge, the paper provides first evidence of multiple equilibria in income dynamics.

(262 words)

A mis dos niñas.

Acknowledgements

I wish to express my gratitude to all the people that contributed to this thesis. First and foremost I am indebted to my supervisor, Stefan Dercon, whose support, patience and encouragement at all stage of this process has been invaluable. He contributed to each of the chapters with helpful and insightful comments that have influenced greatly this work as well as my approach to economic research. I can only thank him for being there when needed as an academic advisor and as a friend.

I would also like to express my gratitude to those that contributed with this work either by reading and commenting on earlier versions of my chapters or by discussing specific aspects of them during my presentations at seminars and conferences, especially Adrian Wood, Marcel Fafchamps, Jere Behrman, Sudhir Anand, Javier Escobal, Santiago Cueto Alemayehu Seyoum, Pieter Serneels, Andrew Zeitlin, Michele DiMaio, Niels Kemper and Joachim DeWeerdt. Special thanks to Francis Teal and Martin Browning, who as transfer interviewers influenced the direction of my work. My friends and colleagues Bob Rijkers, Sofya Krutikova, Oswaldo Molina, Bet Caeyers and Martina Kirchberger for fruitful discussions. I am particularly indebted to Catherine Porter and Alan Sanchez for being loyal friends and patient collaborators. Without them the thesis would have been a far less attractive endeavor.

Different institutions helped during these years. I acknowledge financial support from the Herbert & Ilse Frankel Foundation from Oriel College, the Inter-American Development Bank and the University of Oxford Department of Economics. I thank the Young Lives Project and ICRISAT for granting me access to the databases used in this analysis.

I am also grateful to all my dear friends that were always there for us during the DPhil. My close friends Darius, Leslie, Oswaldo, Ursula, Lasse, Steffie, Magalie and Albert. I am also indebted to Jo and Brian for welcoming us as part of their family.

This is dedicated to those I love. To my parents, Betty y Pepe, for their values and everything else they have taught me. To my brother, Brais. To Noa for teaching us who

we want to be. And to Julia for a shared life.

Contents

Word Count: approx. 60,000.

Statement of Authorship

The following document details the authorship of the four chapters submitted by Ingo Outes-Leon towards the fulfilment of the requirements for the DPhil in Economics, at the University of Oxford.

Chapter One: ”Introduction”

Ingo Outes is the sole author of this chapter.

Chapter Two: ”Revisiting the Nutrition and Cognition link”

This paper is a further developed version of work started by a team of the Young Lives project, in practice led by Ingo Outes and Alan Sanchez. Alan collaborated in the development of the underlying ideas, but most of detailed robustness tests and further analysis was the exclusive work of Ingo Outes.

Chapter Three: ”Catching up while there is time”

This paper is joint work between Ingo Outes and Catherine Porter. While Ingo Outes developed the core ideas, Cath collaborated in the analysis and drafting of the paper.

Chapter Four: ”Income Dynamics in Rural India”

The paper is the output of a Research Officer position Ingo Outes held during his period as a doctoral student. The RO position was part of an ESRC-funded project held by Prof Stefan Dercon titled ”Risk, Shocks, Growth and Poverty”. The paper is a co-authored work between Stefan Dercon and Ingo Outes. Some of the initial ideas were jointly developed, but the development of the precise empirical methodology, data analysis as well as the drafting of the paper is exclusively the work of Ingo Outes.

Chapter 3 in this doctoral thesis is forthcoming for publication as: Outes, I. and C. Porter, (forthcoming) ”Catching up from early nutritional deficits? Evidence from rural Ethiopia” Economics and Human Biology.

Chapter 1

Introduction

This collection of essays is motivated by the notion that uninsured shocks can have persistent and permanent effects on the welfare of individuals and households. Poverty trap models suggest the existence of binding dynamic thresholds in income and asset accumulation that result in low mobility and in potentially permanent effects of uninsured income shocks. A similar dynamic pattern can be found in models of early child development and the notion of critical periods, which suggest that investments in human capital might only be effective during particular stages of a child's development.

The first two essays explore issues of early child development and critical periods. The first essay revisits the nutrition and cognition link in early childhood, exploiting shocks during the nutritional critical period to identify causality, while the second examines heterogeneity in catch-up growth and the extent to which nutritional remediation is effective. Finally, the third essay contributes to the literature on poverty traps by testing for the existence of multiple equilibria and dynamic thresholds in an agrarian economy.

1.1 Critical Periods and Early Child Development

One strand of the literature on poverty and poverty traps focuses on human capital and health dynamics. Poor health and low adult BMI might reduce the ability to generate income through labour productivity (e.g. Thomas and Strauss (1997)) as well as increase morbidity and mortality rates (Fogel, 1997; Strauss and Thomas, 1998; Deaton, 2005). To the extent that these mechanisms are sufficiently large, poor health and nutrition might give rise to poverty traps (Dasgupta and Ray, 1986, 1987; Ray, 1993). However, few empirical studies have found conclusive evidence of adult health dynamic thresholds being binding (Strauss and Thomas, 1998).

Although not always recognized as such, evidence in the literature on early child development shows strong features of binding dynamic thresholds. In the medical and pediatrics literature, it is well understood that nutritional inputs during pregnancy and the first two or three years of a child's life are crucial in determining later adult height, health and other significant outcomes (see Doyle et al. (2009), Victora et al. (2010) and Strauss and Thomas (2008)). Temporary economic shocks and periods of short-term deprivation can have persistent effects on a child's long-term wellbeing (Almond and Currie, 2011). Moreover, deficiencies can also have effects on cognitive development and human capital early on in life (Grantham-McGregor et al., 2007; Thompson and Nelson, 2001).

Ultimately these deficiencies can result in poorer health and shorter lives, as well as affect educational attainment, employment and wage income later in life (Hoddinott et al. (2008); Maluccio et al. (2009); Alderman et al. (2006) and Glewwe et al. (2001), and overviews by Glewwe and Miguel (2008) and Strauss and Thomas (2008)). Further, the birth-weight of newborns is closely related to the birth-weight of their mothers, thus indicating that low body weight at birth is persistent across generations (Behrman and Rosenzweig, 2001, 2004). Such inter-generational linkages could be perpetuating low levels of economic mobility and poverty traps (Osmani and Sen, 2003).

Models of human capital formation that incorporate these insights have become influential in the economics discipline (Cunha and Heckman, 2007, 2008; Almond and Currie, 2011). These models acknowledge that human capital incorporates multiple skills - health, cognitive and non-cognitive -, and that skill formation might show features of self-productivity and skills complementarities. In other words, skills beget skills. They also recognize that the efficacy of skill investments might vary over time, and that early life might be defined by 'sensitive' and even 'critical' periods, after which further investments are largely ineffective.

As such these models suggest that the efficiency-equity trade-off does not exist in early childhood - i.e. investment is both efficient and equitable -, prompting Alderman (2010) to argue persuasively that spending on early childhood nutrition should not be viewed as a redistributional tool, but as investments in health, nutritional and cognitive development.

Related to the concept of critical periods is the notion that nutrition and other human capital dimensions might be subject to catch up growth. Understanding the magnitude of catch-up growth is crucial for issues of vulnerability and long-term welfare: if catch up is large, temporary shocks might have relatively mild long-terms effects. A small literature in economics has sought to quantify this phenomenon in nutrition (Deolalikar, 1996; Fedorov and Sahn, 2005; Alderman et al., 2006; Mani, 2008), but while there is agreement that catch-up is highest in early childhood, estimates of catch-up growth differ widely by method and age-range of analysis. A feature of these studies is that they focus on average rates of catch-up across the entire distribution of children, regardless of nutritional status of the child or the economic ability of their families to carry our remedial investments. Yet the medical literature suggests that catch-up growth might be expected among those below the standard growth curve, and only once the original nutritional depravation has been removed (see for example, Boersma and Wit (1997)).

In Chapter Three we seek to address these shortcomings and explore how socio-economic status affects catch-up growth while allowing for rates to vary across the height distribution. Our results are compelling in that we find that catch-up growth is highly non-linear but only among children in wealthier households: catch-up is near perfect among 'wealthy' children in the bottom quartile of the height-for-age distribution. Moreover, while this pattern of nutritional remediation is effective between the ages of one and five, beyond the age of five household wealth no longer affects height growth. We interpret this finding as evidence that the window of opportunity might be closed by the age of five.

Chapter Two exploits the notion of a critical period in nutrition to analyze the effect of early nutritional deficiencies on cognitive development. Glewwe et al. (2001) and Alderman et al. (2006) present evidence of a causal nutrition-cognition link, though the validity of their instruments is contested (Glewwe and Miguel, 2008; Strauss and Thomas, 2008). In a similar setting to theirs, we apply instrumental variable methods to a sibling-difference specification of preschool children in Peru. The challenge of estimating the effect of the health inputs on cognitive development is that of other inputs being missing (Todd and Wolpin, 2003, 2007). By focusing on pre-school children, we reduce the sphere of cognitive influence to the home environment only. A further advantage of our data is that the paired-siblings outcomes are measured roughly at the same age, allowing us to control more effectively for unobserved age-specific investments.

Our instrumentation strategy uses food price changes and household shocks as instruments for height-for-age differences between siblings. Different lengths of exposure to the shocks during the critical developmental period of a child identify our analysis. We find that our instruments are reasonably strong and pass standard validity tests, and even though they capture very different sources of variation in sibling-difference nutrition, both sets of instruments providing similar estimates. The evidence shows that there is a significant and causal impact of early nutrition on cognitive ability. The magnitude of the effects being sizeable especially considering that these children are yet to enroll in formal schooling with deficits likely to further widen further in later years.

1.2 Poverty Dynamics

Over the years, a wide range of poverty trap models have been developed (Dasgupta and Ray (1986); Galor and Zeira (1993); Banerjee and Newman (1993); Mookherjee and Ray (2002) to name but a few). While these models vary greatly in the mechanisms they aim to portrait, they all share their reliance on some type of market failure and a dynamic function with certain non-linear features. The combination of these factors leads to a separating equilibrium whereby agents below a particular threshold converge over time to a lower equilibrium. These type of models predict both low levels of mobility and a strong degree of path dependence (see Benabou (1996) and Azariadis and Stachurski (2005) for reviews of this topic).

In spite of their popularity, little empirical evidence supports their existence. Faced with a multiplicity of mechanisms, empirical studies have sought to test reduced form models of income and asset dynamics. Studies that have applied flexible non-parametric methods have found some evidence of non-linear patterns in asset dynamics (notably Lybbert et al. (2004); Adato et al. (2006); Barrett et al. (2006) and Quisumbing and Baulch (2009)), but their methods do not satisfactorily address the endogeneity of lagged assets. On the other hand, studies that do so, typically applying Arellano-Bond type of methods (Antman and McKenzie, 2007; Jalan and Ravallion, 2004; Lokshin and Ravallion, 2004), find no evidence of multiple equilibria. A number of factors might conspire against these studies. The use of panel data with short duration - typically less than five years Dercon and Shapiro (2007) - may not capture the long dynamics behind poverty persistence. Further, if measurement error is serially correlated, the use of distant lags as instruments, as Arellano-Bond methods do, will overstate mobility Antman and McKenzie (2007).

In Chapter Four, we contribute to this literature by addressing some of these technical challenges. We follow Lokshin and Ravallion (2004) and others in modeling current income as a polynomial function of lagged income while allowing for household heterogeneity. However, we use exogenous instruments to correct for endogeneity problems, obviating the need for Arellano-Bond methods and addressing the critical problem of measurement error in income. Moreover, our dataset - a 30-year long panel from rural India - is ideally suited to capture poverty trap dynamics.

Our empirical analysis shows that the income generating dynamics in rural India follow a quadratic polynomial function with pronounced concavity. Income simulations suggest the presence of two equilibria: a stable high-income equilibrium and a low-level unstable saddle point. While households with sufficiently high fixed effect income follow the high-equilibrium dynamic path, almost half of our sample has too low an income steady state to overcome the dynamic point of divergence. To the best of our knowledge, this constitutes first evidence of multiple equilibria in income dynamics.

References

Adato, M., Carter, M., and May, J. (2006). Exploring poverty traps and social exclusion in south africa using qualitative and quantitative data. *The Journal of Development Studies*, 42(2):226–247.

Alderman, H. (2010). The economic cost of a poor start to life. *Journal of Developmental Origins of Health and Disease*, 1(1):19–25.

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

Almond, D. and Currie, J. (2011). *Human Capital Development before Age Five*, volume 4 of *Handbook of Labor Economics*, chapter 15, pages 1315–1486. Elsevier.

Antman, F. and McKenzie, D. (2007). Poverty traps and nonlinear income dynamics with measurement error and individual heterogeneity. *The Journal of Development Studies*, 43(6):1057–1083.

Azariadis, C. and Stachurski, J. (2005). Poverty traps. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, chapter 5. Elsevier.

Banerjee, A. V. and Newman, A. F. (1993). Occupational choice and the process of development. *Journal of Political Economy*, 101(2):274–98.

Barrett, C., Marenya, P. P., Mcpeak, J., Minten, B., Murithi, F., Oluoch-Kosura, W., Place, F., Randrianarisoa, J. C., Rasambainarivo, J., and Wangila, J. (2006). Welfare dynamics in rural kenya and madagascar. *The Journal of Development Studies*, 42(2):248–277.

Behrman, J. R. and Rosenzweig, M. R. (2001). The returns to increasing body weight. PIER Working Paper.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to birthweight. *The Review of Economics and Statistics*, 86(2):586–601.

Benabou, R. (1996). Equity and efficiency in human capital investment: The local connection. *Review of Economic Studies*, 63(2):237–64.

Boersma, B. and Wit, J. (1997). Catch-up growth. *Endocrine Reviews*, 18(5):646–661.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. *The American Economic Review*, 97(2):31–47.

Cunha, F. and Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *J. Human Resources*, 43(4):738–782.

Dasgupta, P. and Ray, D. (1986). Inequality as a determinant of malnutrition and unemployment: Theory. *Economic Journal*, 96(384):1011–34.

Dasgupta, P. and Ray, D. (1987). Inequality as a determinant of malnutrition and unemployment: Policy. *Economic Journal*, 97(385):177–88.

Deaton, A. (2005). The great escape: A review essay on fogel's 'the escape from hunger and premature death, 1700-2100'. NBER Working Papers 11308, National Bureau of Economic Research, Inc.

Deolalikar, A. B. (1996). Child nutritional status and child growth in kenya: Socioeconomic determinants. *Journal of International Development*, 8(3):375–393.

Dercon, S. and Shapiro, J. (2007). Moving on, staying behind, getting lost: Lessons on poverty mobility from longitudinal data. In Narayan, D. and Petesch, P., editors, *Moving out of Poverty*. World Bank.

Doyle, O., Harmon, C. P., Heckman, J. J., and Tremblay, R. E. (2009). Investing in early human development: Timing and economic efficiency. *Economics & Human Biology*, 7(1):1–6.

Fedorov, L. and Sahn, D. E. (2005). Socioeconomic determinants of children's health in russia: A longitudinal study. *Economic Development and Cultural Change*, 53(2):479–500.

Fogel, R. W. (1997). New findings on secular trends in nutrition and mortality: Some implications for population theory. In Rosenzweig, M. R. and Stark, O., editors, *Handbook of Population and Family Economics*, volume 1 of *Handbook of Population and Family Economics*, chapter 9, pages 433–481. Elsevier.

Galor, O. and Zeira, J. (1993). Income distribution and macroeconomics. *Review of Economic Studies*, 60(1):35–52.

Glewwe, P., Jacoby, H. G., and King, E. M. (2001). Early childhood nutrition and academic achievement: a longitudinal analysis. *Journal of Public Economics*, 81(3):345–368.

Glewwe, P. and Miguel, E. A. (2008). *The Impact of Child Health and Nutrition on Education in Less Developed Countries*, volume 4 of *Handbook of Development Economics*, chapter 56, pages 3561–3606. Elsevier.

Grantham-McGregor, S., Cheung, Y., Cueto, S., Glewwe, P., Richter, L., and Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. *The Lancet*, 369(9555):60 – 70.

Hoddinott, J., Maluccio, J. A., Behrman, J. R., Flores, R., and Martorell, R. (2008). Effect of a nutrition intervention during early childhood on economic productivity in guatemalan adults. *The Lancet*, 371(9610):411–416.

Jalan, J. and Ravallion, M. (2004). Household income dynamics in rural china. In Dercon, S., editor, *Insurance Against Poverty*. Oxford University Press.

Lokshin, M. and Ravallion, M. (2004). Household income dynamics in two transition economies. *Studies in Nonlinear Dynamics & Econometrics*, 8(3):4.

Lybbert, T. J., Barrett, C. B., Desta, S., and Coppock, D. L. (2004). Stochastic wealth dynamics and risk management among a poor population. *Economic Journal*, 114(498):750–777.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. (2009). The impact of improving nutrition during early childhood on education among guatemalan adults. *Economic Journal*, 119(537):734–763.

Mani, S. (2008). Is there complete, partial, or no recovery from childhood malnutrition? empirical evidence from indonesia. Technical report.

Mookherjee, D. and Ray, D. (2002). Contractual structure and wealth accumulation. *American Economic Review*, 92(4):818–849.

Osmani, S. and Sen, A. (2003). The hidden penalties of gender inequality: fetal origins of ill-health. *Economics & Human Biology*, 1(1):105–121.

Quisumbing, A. and Baulch, B. (2009). Assets and poverty traps in rural bangladesh. Working Papers id:2158, eSocialSciences.

Ray, D. (1993). Labor markets, adaptive mechanisms, and nutritional status. In et al, B., editor, *Essays in Honour of K.N.Raj.* Oxford University Press.

Strauss, J. and Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2):766–817.

Strauss, J. and Thomas, D. (2008). *Health over the Life Course*, volume 4 of *Handbook of Development Economics*, chapter 54, pages 3375–3474. Elsevier.

Thomas, D. and Strauss, J. (1997). Health and wages: Evidence on men and women in urban brazil. *Journal of Econometrics*, 77(1):159–185.

Thompson, R. and Nelson, C. (2001). Developmental science and the media: Early brain development. *American Psychologist*, 56(1):5–15.

Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33.

Todd, P. E. and Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, 1(1):91–136.

Victora, C. G., de Onis, M., Hallal, P. C., Blssner, M., and Shrimpton, R. (2010). Worldwide timing of growth faltering: Revisiting implications for interventions. *Pediatrics*, 125(3):e473–e480.

Chapter 2

Revisiting the Nutrition and Cognition link: A Within-Sibling Investigation of Peruvian pre-school children

Chapter 2

Revisiting the Nutrition and Cognition link: A Within-Sibling Investigation of Peruvian pre-school children

Abstract

An extensive literature documents linkages between early nutritional deficiencies and reduced cognitive ability, educational attainment and, ultimately, lower labor market performance. Few of these studies, however, have shown these correlations to be genuinely causal. We reexamine the nutrition and cognition link, applying instrumental variable methods to a sibling-difference specification for a sample of Peruvian pre-school children. We use household shocks and food price changes as instruments. As such our analysis also quantifies the nutritional and cognitive costs of the 2006-08 global food price crisis. We find that there are significant and negative cognitive effects of early childhood nutritional disinvestments: a decrease in Height-for-Age z-score leads to a reduction in the Peabody Picture Vocabulary Test score of 17%-21%. The accumulated deficits are sizeable considering that these children are only 3-6 years old and are yet to enroll in formal schooling, with deficits likely to widen in later years.

2.1 Introduction

A growing literature in economics, nutrition and sociology has built a substantial evidence base on the linkages between early nutritional deficiencies and reduced cognitive ability, educational attainment and ultimately lower market wages later on in life. Few of these studies, however, have been able to show these correlations to be genuinely causal. In this paper we seek to establish the causal relationship between nutritional achievement and cognitive development in a sample of pre-school aged Peruvian children. Concern about the long-term effects of childhood malnutrition has been amplified by the food price crisis that led to a global rise of 40% in food prices during the 2006-2008 period (von Braun, 2010). Whilst not experiencing the highest rates of inflation in the region, Peru showed some of the most rapid increases, accumulating a 20% increase in food prices between 2006 and 2008 (Cuesta and Jaramillo, 2010). Exploiting the variation in nutritional intake resulting from the food price changes as well as other household-specific shocks, we are able to show that nutrition indeed does have a causal impact on cognitive ability. Our results have a further policy meaning, since they are a quantification of the nutritional and subsequent cognitive costs of the global food crisis on pre-school aged Peruvian children.

Understanding causality in the nutrition-cognition nexus is complicated by the endogenous nature of a child's health status. As illustrated by Behrman and Lavy (1994), both a child's health and her cognitive achievement can be understood as the outcomes of a utility-maximization process whereby parents invest in a child's human capital subject to initial conditions - genetic innate abilities -, parental taste for child's quality and budget constraints. Since parental preferences and their ability to turn inputs into outcomes as well as genetic endowments are unobserved, ordinary least squares (OLS) estimations of the cognitive returns to early nutritional investments are likely to be biased.

Grantham-McGregor (1995) and Grantham-McGregor and Baker-Henningham (2005) review evidence from the nutrition literature, and find that school-aged children who were severely malnourished in the early years are more likely to suffer from cognitive deficits, though they stress that whilst the evidence is strong, it is not unequivocal, and that a number of questions remain unanswered.[1] A strand of the literature uses experimental studies of supplementation to address the issue of endogeneity. The INCAP study in Guatemala (Pollitt et al., 1993) showed substantial positive effects of early childhood nutritional supplements on cognitive achievement among teenagers, and later on their adult development (Maluccio et al., 2009). A similar study implemented in Jamaica found that early child stimulation and nutritional supplements were effective in increasing cognitive achievement at the age of 8 (Grantham-McGregor et al.), but by the age of 12, only 'stimulation' children had higher cognitive achievement than the control group

[1] For a policy review of the relevant evidence, see also World Bank (2006).

(Grantham-McGregor et al., 1997). If appropriately randomized, experimental studies are powerful tools for testing causal linkages. However, ethical and budgetary issues limit the replicability of such studies, especially with regard to studying the effects of undernutrition. Using non-experimental data, Alderman et al. (2006) and Glewwe et al. (2001) (hereafter AHK and GJK, respectively) show that early childhood nutritional deficiencies in the form of low height-for-age can be linked to poorer cognitive attainment later in life. In doing so, they exploit within-sibling variations to deal with the endogeneity bias resulting from unobserved household heterogeneity. They also address differential parental investments resulting from child heterogeneity in innate abilities by applying instrumental variable estimation.

Following AHK and GJK, in this paper we combine a sibling-difference specification with instrumental variable methods to study the relationship between nutrition and cognitive achievement during the pre-school period for a sample of Peruvian children. We use data from a novel sample of paired-siblings to estimate a conditional demand function for cognitive achievement in sibling-difference form, in which all investments common to both siblings are removed. To alleviate concerns of differential parental investments between siblings that might drive differences in nutrition and cognitive outcomes, we incorporate additional controls to the sibling-difference specification in dimensions related to birth-order and birth-sex-order. We also control for changes in household and community circumstances that might lead to differential outcomes between siblings.

In our instrumentation strategy, we use two sets of instruments for height-for-age differences between siblings, namely food price changes and household shocks occurring during the critical developmental period of a child. We exploit differences between siblings by looking at food prices prevalent during the first three years of their life as a source of exogenous variation in nutritional inputs experienced by the siblings. Given that our sample comprises paired-siblings born in the periods 2001-2 and 2003-5 respectively, younger siblings were affected during their early years by the food price crisis, which impacted Peruvian households most severely in 2006 and 2007. In addition, we use household-specific short-term shocks that took place between 2000 and 2002 and between 2007 and 2009 as a further set of exogenous instruments. Because of their timing, these shocks can be considered to be child-specific, since they affected the household when one of the siblings was in her critical nutritional period, with the counterfactual sibling either relatively old or yet to be born.

We apply this strategy to the Young Lives Peruvian survey.[2] The data are a novel sample of paired-siblings born in 2001-2 and 2003-5 respectively, for which anthropometric and cognitive measures were collected at roughly the same age-period - mostly between 4 and 6 years of age - at two different points in time, the 2006 and 2009 waves of the Young Lives survey. We use Peabody Picture Vocabulary Test (PPVT) scores as the cognitive outcome measure and contemporaneous height-for-age z-scores (HAZ) as the nutritional measure.

An advantage of our data is that the paired-siblings outcomes are measured roughly at the same age. Cognitive achievement at a particular age can be modeled as a function of a child's innate genetic ability and the cumulative effect of present and past cognitive investments in both the home and school environment (see Todd and Wolpin (2003, 2007)). The challenge of estimating the effect of the health inputs on cognitive development is that of other inputs being missing. By focusing on a sample of children consisting primarily of pre-school age children, we reduce the sphere of cognitive influence mainly to the home environment. As such we contribute to the literature on cognitive development during pre-schooling age (see Paxson and Schady (2007), Berlinski and Galiani (2007), and Behrman et al. (2004) among others, and Schady et al. (2006) for a review). .

Furthermore, our methodology allows us to go beyond previous studies. GJK argue that while instrumentation with birth weight can solve problems of differential parental investments across siblings, it does not deal with unobserved genetic factors affecting both nutritional and cognitive outcomes. Instead they suggest the use of nutritional shocks as instruments, to recreate the identifying conditions of a 'natural' experiment, but data limitations prevented the application of their preferred methodology. Using drought and civil war incidence as instruments of a child's stature, AHK implement the methodology outlined by GJK. However, the validity of their choice of instruments has been contested (Glewwe and Miguel, 2008).

[2] www.younglives.org.uk

Our analysis shows that there is a significant and causal impact of early nutrition on cognitive ability. Diagnostics of the first-stage results indicate that both price-inflation and household-specific shocks are valid and reasonably strong instruments. Even though they capture different sources of variation in nutritional investments, both sets of instruments provide similar point estimates, suggesting that both might identify a single underlying mechanism.

We believe that our proposed set of instruments are not only well placed to meet the stringent conditions set-out in GJK, but are arguably sufficiently short-lived to have little impact on later cognitive achievement other than through their impact on anthropometric status. Indeed, we show that our results remain stable throughout a number of robustness checks on potential violations of the exclusionary restriction. The main concern with the price-shock instrument is that it might lead to persistent household level effects. We present evidence that our findings are robust to potential changes in household time allocation patterns; fertility choices; as well as persistent effects on household assets and consumption. Furthermore, our results are also robust to the inclusion of controls for delayed school and pre-school enrolment, suggesting that our analysis captures a nutrition-cognition parameter beyond the cognitive effects of delayed enrolment.

The effects uncovered appear to be substantial in magnitude; a one standard deviation increase in Height-for-Age would lead to an increase in the PPVT score of 17%-21% of a standard deviation. The magnitude of these effects is significant, considering that the cognitive deficits have been accrued only during the first few years of a child's lifetime. Moreover, they provide a first quantification of the nutritional and subsequent cognitive costs of the food crisis among pre-school age Peruvian children.

The remainder of the paper is organized as follows. Section 2.2 describes our conceptual framework and lays out in detail the empirical strategy to be used. Section 2.3 introduces the key features of the sample, the measurement variables and the instrumental variables chosen for the analysis. Section 2.4 shows our main results while section 2.5 discusses a number of robustness checks. Finally, section 2.6 concludes.

2.2 Methodology

2.2.1 Conceptual framework

Our conceptual model is a two-period characterization of early child development. Consider a framework in which the first part of early childhood (the first 2-3 years) is considered as Period 1, and the remainder of early infancy as Period 2. Denote nutritional status accumulated at the end of Period 1 of child k from household h as $H_{t-1,k,h}$ and pre-school cognitive achievement at the end of Period 2 as $CA_{t,k,h}$. We assume that $H_{t-1,k,h}$ summarizes all the investment made in the child during Period 1. In turn, $H_{t-1,k,h}$ is assumed to be an input for $CA_{t,k,h}$. Both variables are chosen by parents on the basis of preferences, budget constraints and initial conditions. We focus on the following equation,

$$CA_{t,k,h} = \alpha H_{t-1,k,h} + X_{t,k,h}\Pi + \eta_{CA,h} + \mu_{CA,k,h} + \epsilon_{k,h} \tag{2.1}$$

where $X_{t,k,h}$ is a vector that includes Period 2 child and household observable characteristics that have an influence on cognitive achievement; $\mu_{CA,k,h}$ represents child unobservable characteristics; and, $\eta_{CA,h}$ captures unobserved household and environmental characteristics affecting cognitive development. Equation (1) can be interpreted as a conditional demand function for cognitive achievement such that $H_{t-1,k,h}$ is the input of interest and $X_{k,h}$, $\mu_{CA,k,h}$ and $\eta_{CA,h}$ are unobservable determinants of parental cognitive investments. For instance, $\eta_{CA,h}$ reflect aspects such as household intellectual environment, whereas $\mu_{CA,k,h}$ incorporates aspects such as child innate ability (for a similar setup, see Glewwe and Miguel (2008)). As described in Behrman (1996), Behrman and Lavy (1994), GJK and AHK, the main challenge of estimating equation 2.1 arises from the possibility that at least one of the following conditions does not hold:

$$E(H_{t-1,k,h}, \eta_{CA,h}) = 0 \tag{2.2}$$

$$E(H_{t-1,k,h}, \mu_{CA,k,h}) = 0 \tag{2.3}$$

If either 2.2 or 2.3 do not hold, then an OLS estimation of the parameter of interest, α, would be biased. A violation of condition 2.2 could arise if there are unobservable household characteristics that simultaneously explain why some families are more likely to raise both healthy and well-educated children. Specifically, determinants of child health not included already in $X_{k,h}$ might be correlated with household unobservable characteristics that influence cognitive achievement (e.g., parental health knowledge might be correlated with household intellectual environment). Similarly, unobserved community characteristics, if correlated with health status (e.g., communities with better health services are also likely to have better educational services) would also lead to violations of 2.2. In turn, a violation of condition 2.3 could arise if child-specific unobservables are correlated with health status. Two possible mechanisms for this phenomenon have been suggested in the literature. Firstly, parental nutritional investments might be adjusted as a child's innate cognitive abilities are revealed - condition (2.2). Secondly, the health status and the cognitive ability of a child might be correlated through a common unobserved genetic endowment - condition (2.3).

Although in principle an instrumental variable approach should suffice to deal with endogeneity due to infringements of conditions 2.2 and 2.3, finding a valid, strong instrument for pre-school nutrition is challenging. Instead, the standard approach has consisted in following a two-prong strategy whereby household fixed effects and instrumental variable are jointly implemented. In the context of cognitive returns to investments in early nutrition, this was first applied by GJK and AHK, in turn echoing earlier studies (see Rosenzweig and Wolpin (1995) for an example and references to studies that have used kinship data). Specifically, assuming data on cognitive achievement and nutritional status is available for a pair of siblings i and j, one can estimate a sibling-difference version of equation 2.1. As illustrated by 2.6, such a strategy allows us to eliminate any factor that is common across siblings.

$$CA_{t,i,h} = \alpha H_{t-1,i,h} + X_{t,i,h}\Pi + \eta_{CA,h} + \mu_{CA,i,h} + \epsilon_{i,h} \tag{2.4}$$

$$CA_{t,j,h} = \alpha H_{t-1,j,h} + X_{t,j,h}\Pi + \eta_{CA,h} + \mu_{CA,j,h} + \epsilon_{j,h} \tag{2.5}$$

$$\Delta_{i,j}CA_{t,h} = \alpha\Delta_{i,j}H_{t-1,h} + \Delta_{i,j}X_{t,k,h}\Pi + \Delta_{i,j}\mu_{CA,h} + \Delta_{i,j}\epsilon_h \quad (2.6)$$

In effect, this sweeps out any potential bias due to departures from (2.2) but leaves endogeneity due to violation of (2.3) unresolved. Parents might still be allocating investments differently across siblings based on sibling differences unobserved to the researcher. Alternatively, inherently healthier siblings might also be more likely to be smarter by nature. Thus, an instrumental variable approach is still required in 2.6 for the $\Delta_{i,j}H_{t-1,h}$ term. In addition, the use of instrumental variables helps dealing with the increased noise-to-signal ratios that occurs when implementing sibling-difference methods (Ashenfelter and Krueger, 1994). The data collected on siblings (see below) were collected only in the latest round of the survey, and therefore in the empirical application we are effectively constrained to use contemporaneous height-for-age instead of its lag (e.g. H_t not H_{t-1}). However, the long-term nature of height-for-age as an indicator of nutritional investments, especially in the critical period for each child, combined with the choice of instruments reflecting events that took during Period 1 of our conceptual model mitigate the shortcomings of the data. Cunha and Heckman (2007) and Heckman (2007) provide a discussion on the theoretical notion of a 'critical period' for child development in the early years, when there are complementarities between investments in the early period, and those in subsequent periods.[3]

2.2.2 Empirical strategy

Empirically, we estimate equation (2.6) using data on matched-siblings born in 2001-2 and 2003-5, respectively. The data allow us to compare two siblings at a similar nutritional and cognitive developmental stage. More specifically, we relate differences in cognitive achievement and Height-for-Age between siblings when aged approximately 4-6 years but measured at different points in time, 2006 and 2009 for the index and sibling children respectively. Our baseline econometric specification is represented by equation (2.7).

[3]There is limited evidence on what precisely the critical period is (are) for aspects of child development, though a consensus that the foetal period and the first three years are extremely important. Almond and Currie (ming) review the limited evidence. See also Rai for a cohort study.

$$\Delta_{i,j}^{06,09} PPVT_h = \alpha \Delta_{i,j}^{06,09} HAZ_h + \gamma_1 \Delta_{i,j}^{06,09} Age_h + \gamma_2 Demo_h^{06} + \vartheta_c + \varepsilon_h \tag{2.7}$$

We include a range of variables designed to capture changes in the conditions that might have effected cognitive investments across siblings, $\Delta_{i,j} X_{t,k,h}$. Equation 2.7 includes controls for household demographics ($Demo_h^{06}$), changes in community services (ϑ_h), as well as sibling differences in age between 2006 and 2009 ($\Delta_{i,j}^{06,09} Age_h$).

Household demographic controls are designed to capture differences in parental investments that might lead siblings to follow different developmental paths. We address two possible such patterns of differential investments, both linked to parental child preference. First, birth-order might play a role in both the time dedicate to a child and the level of resource competition in the household. The index child might have benefited more for being of a lower birth order relative to her sibling[4] or because the household was smaller at the time. We deal with this by including controls for the birth-order of the index child and for the number of siblings born after the index. Secondly, the gender of the child might be a determinant of investments. While extreme patterns of gender bias, such the 'missing women' cases found in South Asia, are not known in Peru, task allocations - such as chores, child care and other household production activities - are likely to be linked to both gender and birth-order issues. Accordingly, in our model specification, we include dummies for all gender-birth-order combinations. The baseline model also includes community fixed-effects controls, (ϑ_h), designed to capture changes at the community level, as they could also drive differences between siblings according to their date of birth (e.g., changes in access to and quality of preschool programs).[5]

We apply IV estimation methods to equation (2.7) to address the remaining endogeneity related to unobserved child-specific investments. The application of IV methods requires the identification of instruments that meet the following conditions:

[4] There is indeed evidence pointing out towards this possibility (e.g., Behrman (1988); Horton (1988)).

[5] In our empirical analysis, we use $H_{t,k,h}$ as a proxy for $H_{t-1,k,h}$, because we do not observe the latter for both siblings. Although we acknowledge concerns of a possible simultaneous determination of $H_{t,k,h}$ and $CA_{t,k,h}$, our instrumental variable strategy, where we use shocks that took place during the first 3 years of life to identify $d_{i,j} CA_{t,h}$, reconciles our estimation this with our conceptual model.

$$E(\Delta_{i,j}HAZ_h \cdot Z) \neq 0 \tag{2.8}$$

$$E(\Delta_{i,j}\epsilon_h \cdot Z) = 0 \tag{2.9}$$

Condition (2.8) - or the strength condition - states that the vector of instruments (Z) should be correlated with the endogenous variable. 'Weak' IVs can lead to biased estimates and invalid standard errors. Condition (2.9) defines a valid IV. The instrument should not be correlated with the error term in the main equation. In our context, this means that we are looking for events that are exogenous to the determination of a child's PPVT score but are sufficiently strong to affect the stature of the child.

As a potential source of exogenous variation that could meet these requirements, we look at changes in the conditions faced by each of the siblings during their first three years of life. We consider two sets of instruments for within-siblings nutrition. The first set of instruments correspond to changes in a selected group of food items that together represent around 54% of the household consumption basket. Our motivation for this choice of instrument set stems from the 2006-8 food price crisis. In particular, the fact that the older siblings, born between 2001-2, were not affected by the crisis during their critical nutritional period, while their younger siblings, born between 2003-5, were. Thus we compare changes in the prices faced by the siblings between the sixth and the thirty-fifth month of life (the first six months of life being excluded as for most of this period are infants are likely to still rely on breastfeeding).

The second set of instruments correspond to household-level shocks happening during Period 1 of development for the index child or its younger sibling. Specifically, we focus on negative shocks such as frost and illness or death of other household members that took place between 2000 and 2002 and between 2007 and 2009. The 2000-02 period is linked to the critical nutritional period of the index child at a time when its younger siblings would not have been born yet. Similarly, the 2007-09 period is associated to the critical period of the younger siblings at a time where the older siblings had already surpassed its critical nutritional period. Furthermore, since the cognitive scores for the older siblings were collected by early 2007, the possibility that the 2007-9 shocks could have had an effect on the older siblings can ruled out. Consequently, we treat these household shocks as child-level shocks. We present the selected set of instruments in more detail in the next section. Conditional on $\Delta_{i,j}X_{t,k,h}$, the selected instruments are assumed to act only through a nutritional channel. Glewwe et al. (2001) list the ideal requirements for the use of shocks as instruments. Shocks should be (a) of sufficient magnitude and persistence to affect a child's Height-for-Age; (b) sufficiently variable across households; and (c) sufficiently transitory not to affect the sibling. We believe that our proposed set of instruments are well placed to meet these stringent conditions and, in particular, are sufficiently short-lived to have little impact on later cognitive achievement other than through their impact on physical growth.

The conceptual framework places particular emphasis on the necessity to control for changes in household circumstances over time ($\Delta_{i,j}X_{t,k,h}$) as a control for differences in cognitive investments across siblings. As a robustness check on our main econometric specification (2.7), we include information on changes in non-food household real expenditure per capita and household assets measured at period t.[6] The inclusion of these additional controls serves an additional purpose. Shocks that have a persistent and delayed effect on household welfare, through reduced assets and income generating abilities, could be creating a spurious relationship between cognition and nutrition.

Controlling for changes in the household following the occurrence of the shocks, ensures that the exclusionary restriction (2.9) is not violated. A further potential channel for violations of the instrument validity condition is when shocks affect cognitive investments directly. This could be either because the income effect of the shocks reduces the household's ability to invest in their children's cognitive skills; or because shocks might increase household fertility, and thereby increase sibling competition for cognitive investments. We perform a number of robustness checks on specification (2.7), by including additional information on the time allocations of all children in the households, and the activities of the mother or caregiver, as well as information on fertility choices after the index child was born.

[6] Changes in household assets holding are proxied by changes in an estimated wealth index. This index is the average of three sub-indices: a consumer durables index, an access to services index and a housing quality index. We follow a definition of this index equivalent to that used in the Demographic and Health Surveys.

A further set of checks concerns unobserved cognitive investments outside the household. At time of measurement, a substantial number of the children in our sample are enroled in preschool, while a small proportion of older children have even started primary school education. While the sibling-difference model combined with controls for age-differences should capture the variation generated by the timely enrolment of both siblings, differences between siblings in preschool enrolment as well as the age of preschool enrolment could lead to biases in our estimates. In particular, to the extent that parents decide to delay, or bring forward, enrolment of the sibling because she is smaller, or bigger, than the index child at the same age, this will create a positive correlation between the PPVT and Height-for-Age scores unrelated to the nutrition-cognition causal link (Glewwe and Jacoby, 1995).While we do not have information on preschool enrolment of the siblings, in our robustness checks section, we test the sensitivity of our core results to including controls for preschool enrolment and age of enrolment of the index child only.

2.3 Data

In our analysis, we use the Young Lives Peruvian Survey, a longitudinal sample of a cohort of children born in 2000-2001. The baseline sample is cluster stratified, with twenty districts randomly selected across the country (7 in the coast, 10 in the Highlands and 3 in the jungle). Sites were chosen from a list of districts that excluded the 5% richest districts as measured by a district poverty ranking. This was in line with the policy aim of the project of oversampling children living in poor households (Wilson et al., 2006). Within each selected district, around 100 households with at least one child born between 2001 and 2002 were chosen randomly to participate in the project. The panel of children that is being followed is 2,000 (hereafter, the 'index' children).[7] The surveys collect anthropometric data and administer tests on cognitive ability to all index children, alongside background information on their families and their local communities (*centros poblados* or towns).[8] Currently, three survey waves are available: the baseline round in 2002 and two follow-ups in 2006-7 and 2009. During the time of the surveys, the index

[7] For a detailed description of the sampling design, see Flores and Escobal (2008).

[8] In many instances, the districts selected contain many *centro poblados.* The community-level surveys were administered in the 80 *centros poblados* identified within the 20 districts selected.

children were aged 6-20 months, 4-5 and 7-8 years of age, respectively.

In the 2009 wave, for each sampled household, the anthropometric module and cognitive achievement test were also administered to the sibling born immediately after the index child - hereafter, the 'younger sibling' - provided he/she was at least 4 years of age at the moment of the survey. The vast majority of these younger siblings were born between 2003 and 2005, so that they were between 4 and 6 years of age when the data was collected, a very close match to the age-range (4-5 years) of the index children in the 2006-7 survey wave.

A common problem in longitudinal studies arises due to household attrition. The tracking exercise implemented by the Young Lives project ensures that attrition is kept to a minimum. Dercon and Outes-Leon (2009) explore this point in detail for the first two waves of the study and find that attrition rates are very low by international standards with only 3.7% of the children lost or dropped out, leaving a panel sample of 1,963 children.[9] Based on analysis of observable characteristics, they find that attrited households are not systematically different from non-attrited households. While differences in unobservable characteristics can not be ruled out, low attrition rates suggest that potential biases in the results due to attrition are likely negligible. Subsequent analysis of the third wave shows that

2.3.1 Measurement Variables

We use the Peabody Picture Vocabulary Test (PPVT), Spanish version, as the measure for cognitive achievement, while nutritional status is proxied by the height-for-age z-score (HAZ).

The PPVT is a test of receptive vocabulary that can be used to measure preparedness for school. We use the Spanish version of the test (PPVT-R) to ensure that questions are appropriate for our Peruvian sample. [10] A number of studies have used this test

[9]Subsequent analysis over the three waves (Cueto et al., 2011), shows that attrition rates have been even smaller between the second and third wave (0.7%). This drop is consistent with findings in Dercon and Outes-Leon (2009) that show that the main contributor to attrition between the first two rounds is early mortality, a factor likely to small as children grow older

[10]The test was originally developed for English speakers, but the Spanish version of the test was latter

as the basis for investigations into cognitive development in Latin America (e.g. Paxson and Schady (2007)). In the PPVT test children are asked to select between four pictures the one that best represented the meaning of a word presented to them orally by the enumerators. The test score is based on the child's responses to a maximum of 125 such questions. Although all children face the same list of questions ascending in difficulty, the number of questions and their difficulty differ by child: the starting point is set by the child's age while its performance determines how far up she reaches into the difficulty scale. When administering the tests clear protocols were followed (as established in the Examiner's Manual in Dunn et al. (1986)). The tests do not require any assessment on the part of the enumerators, only that the correctly establish the age of the child and recorder its responses. [11]

A common concern when using the PPVT measure of cognitive development is that scores can be censored. This might happen if the test is too difficult and a child receives the minimum score. The histogram of the Raw score presented in Figure 2-1 does indeed shows some degree of censoring. PPVT score is is also strongly skewed distribution. To standardise the raw scores, the Spanish PPVT (PPVT-R) tests provides a normalising sample of Spanish speaking children. However, we find that in our case, some of the variation in the raw score was being lost in the normalization. Because of this and to address the challenge of censoring, we implement an alternative in-sample standardization procedure. We exploit the fact that children of different ages should be referenced differently; we standardize by age cohorts (3, 4, 5, 6 or 7 years of age), to have a mean of zero and a standard deviation of one for each age group and across the entire sample. Figure 2-2 shows that this procedure reduces to some extent the incidence of censoring and produces a less skewed distribution.

However, it should be noted that our main outcome variable is not the level of the PPVT score, but the difference in the standardised score between siblings. Figure 2-3 presents a histogram of this measure. Reassuringly, the sibling-difference PPVT score has a distribution close to normal and suffers from little censoring. However, the

developed and normed. See Dunn et al. (1986)

[11] see Cueto et al. (2009) for further details of the test and its properties in the context of the Young Lives samples. Using data from the 2006-07 round, Cueto et al. (2009) test the internal consistency of the PPVT tests and find the measurements highly reliable.

absence of censoring the sibling-difference measure does not mean that the problem has disappeared. Censoring in the level variable, reduces the variance in the sibling-difference score. This transforms the censoring problem into a measurement error (of outcome variables) problem.[12]

To measure the stock of nutritional achievement of the children we use the height-for-age z-score (HAZ). Z-scores compare the height of a child relative to a reference group of healthy children. [13] Height-for-age z-scores are recommended by the World Health Organisation (WHO) as a measure of child development, in particular as a correlate of the long-run investments in child nutrition.

2.3.2 Analysis Sample

As described above, our analysis relates differences in PPVT scores and their Height-for-Age measure between siblings at a similar age but different points in time. We exclusively use the sample of matched index and sibling children for which anthropometric and cognitive data was collected, which consists of 900 children in 450 households.[14] As shown in Table 2.1 most of the children, index or sibling, were aged between 4-6 years of age at the time of measurement, 2006 and 2009 for the index and sibling children respectively.[15]

Basic descriptive statistics comparing the paired-siblings households to the rest of the Young Lives households are reported in Table 2.2. We see that there are some significant differences between the households used in the siblings analysis, suggesting that our sample of analysis could be a selected sample. This is not surprising given the

[12]Nevertheless, our core results stand regardless of the measure used. Sensitivity analysis shows that using the Raw score, PPVT-provided standardised score, or the within-sample standardised does not affect the basis of our results.

[13]see http://www.who.int/growthref/ and references on the website. We use the standard measures updated in 2007.

[14]To remove implausible observations and alleviate the problems of attenuation bias in our sibling-difference specification, the sample of matched-siblings used in our analysis has been trimmed off outliers: we exclude the top and bottom 2.5% tails of the sibling-difference Height-for-Age distribution.

[15]Even though the sampling frame for the siblings would have ruled out children below 4 years of age, Table 2.1 shows that data was collected for six siblings aged 3 years. To prevent further reductions to an already small sample size, we include these children in our analysis. However, their elimination would be inconsequential to our results. In any case, in our analysis, differences in age between siblings are control for in the form of sibling age-difference dummies.

sampling frame applied in the collection of the siblings. Accordingly, we are careful to treat our results as representative only of relatively young - and poor - families that have at least two children. However, the fact that we use a particularly young sample of households, with active fertile lives, might suggest that concerns regarding endogenous fertility should be considered with particular care. We return to this point later in the paper.

We nevertheless estimate the cross-section for cognitive achievement using OLS and reject the null hypothesis that the cognition effect of nutrition is significantly different between our paired-siblings sample and the excluded sample (not reported). This alleviates concerns that our results could be driven by sample selection.

At this point it is instructive to understand how our sibling sample compares with the wider population in Peru and the Latin American region for our key indicators. The Young Lives sampling frame ensures that ours is a pro-poor sample. Comparisons with the 2005 DHS survey in Peru, a nationally representative sample, confirm the pro-poor bias. While both samples have similar levels of rural households (45%), 32% of 4-5 year olds in the DHS sample are stunted,(DHS) while 45% of our sample of siblings have HAZ below -2 s.d. When assessing the cognitive development of our sample of siblings, we compare them with the study by (Paxson and Schady, 2007). They use data from Ecuadorian children sampled for the purpose of assessing a social protection programme - the sampling also following a pro-poor bias - with surveys administered in 2003-04. However, descriptives indicate that their sample has better educated mothers and children with better nutrition. Using the officially-standardised Spanish PPVT, they report a score of 86.24 for a sample of 3-6 year olds). In comparison, the children (aged 4-5) in our sample score poorly with an average score of 83.

Table 2.3 presents some descriptive statistics for the index children and their younger siblings. It shows that on average the younger siblings are better nourished than their older counter-parts. As is shown in the next sub-section, siblings were exposed to substantial price increases during the 2006-2008 period, prior to their measurement. If these price increases had an effect on nutrition, as would be required for them to be *strong* instruments, this is not apparent from Table 2.3. However, these differences are hardly surprising considering that they were measured three years after their older *index* siblings; their better nutrition is possibly a reflection of being born at a later stage of the household's life-cycle, benefiting from improved economic conditions, or simply the results of a secular trend. They could also be linked to improvements in access to health services and nutritional programs at the community-level. As described above, in our econometric specification we include district dummies and controls for non-food consumption and household assets designed to capture time-varying effects resulting from life-cycle trends.

Before presenting the results of the econometric analysis, it is useful to plot the correlation between nutrition and cognition in our data. The solid line in figure 2-4 presents the kernel density of Height-for-Age sibling-difference, while the dashed line depicts the kernel smoothing estimate of PPVT on Height-for-Age in sibling-difference form. The third line, in dash-dot form, shows the kernel smoothing estimates for the pooled OLS model.[16] We find that the pooled OLS slope is substantially steeper than the nutrition-cognition slope for the sibling-difference model, suggesting that time-invariant household characteristics substantially bias the cognition- nutrition vector. However, as discussed above, the sibling-difference relationship depicted in Figure 2-4, might at the same time mask substantial endogeneity. While the increased attenuation bias will bias the slope downwards, differential cognitive investments across children could be positively or negative correlated with nutrition, implying that *a priori* it is not possible to sign the direction of the remaining bias. The aim of our instrumentation strategy is to establish the direction of that bias. We now turn to the discussion of our instruments.

2.3.3 Instrumental variables

We use two sets of exogenous instruments; food price changes and idiosyncratic shocks that affected the household in the critical nutritional period. Food prices are clearly relevant to nutrition, exogenous to the household and vary sufficiently over time during

[16]Note that the x-axis in Figure 2-4 is to be interpreted very differently depending on whether we are considering the sibling-difference or pooled OLS relationship. While the ΔHAZ mean is approximately around the value of zero and has no nutritional meaning, the levels line depicts the Height-for-Age in our sample.

the period of study. Moreover, in light of the nature of the food price crisis, we argue that the event was grave and sufficiently short-lived around the critical period of one of the siblings cohorts: the younger sibling. Figure 2-5 shows that the most dramatic stage of the crisis took place between 2006 and 2008, coinciding with the critical period of the younger siblings born between 2004 and 2005. In contrast, the older siblings born between 2001 and 2002 had already transitioned their critical nutritional stage before the beginning of the crisis.

For the purpose of estimating the impact of the food crisis on child nutrition, we do not center exclusively on products consumed by the child, because the increase in prices was generalized. Rather, we look at fluctuations in food price categories deemed important for overall household food consumption and study their effect on within-siblings early nutrition. There are likely to be a number of channels through which food prices impact child nutrition, none of which can be ruled out *a priori*. For households that are net consumers, increases in food prices can lead to a reduction in child food intake both in terms of the quantity and quality. This can happen directly, when the specific food item is one consumed by the infant (such as diary or cereals), or indirectly due to reallocations within the household consumption basket. In turn, for households that are net producers, increases in food prices can have a positive effect on nutrition through their positive income effect.

Taking price data from the Peruvian institute of statistics (*Instituto Nacional de Estadística e Información*, INEI), we create a child-specific variable, representing food prices in the first 3 years (excluding the first 6 months of breastfeeding), disaggregated by semester (6-11 months, 12-17 months, 18-23 months, 24-29 months, 30-35 months). This strategy was used by Glewwe and King (2001); food prices were also used by Alderman et al. (2001) with slightly broader age categories. Creating the child-specific variable allows us to introduce as much heterogeneity as possible both across and within households - though recall that the index children are all born within a year of each other, by virtue of the cohort sample design.

The food prices we use were obtained from data reported by INEI, who collect price data on a monthly basis across the main cities of the country in order to construct regional consumer price indices.[17] For our purposes, we impute prices by matching the 20 clusters sampled by Young Lives to the prices prevalent in the associated capital of the Department where the districts are located. We then match these prices to each child according to date of birth and use as IV the siblings-difference in log-prices. Note that, since our estimation controls for cluster fixed effects, the effect of prices is identified by sibling differences in date of birth within each cluster.

Concerned with parsimony in the IV specification,[18] we include three of the most important food price sub-categories. Cereals: Bread and Cereals (comprising subcategories wheat, rice, maize, pasta as well as bread and biscuits); Meat: (comprising chicken, red meat and meat products, and processed meat); Dairy: Milk, Eggs and Cheese. The three categories are important for their calorific and quality of protein content. They also represent just over 50% of the purchased consumption basket in the sampled families.[19] We also include the category of Tubers (comprising potatoes, cassava, yucca, and Andean tubers) as the largest category of home production. All four price categories represent 56% of the purchased consumption basket in the third round, and 64% including own-consumption, showing that many households produce their own tubers.

Our vector of price differences is thereby defined as follows:

$$\begin{aligned}\Delta_{i,j}^{06,09} Prices_{t,h} =& \pi_1 \Delta_{i,j}^{Ch.Sem.} Cereals_c + \pi_2 \Delta_{i,j}^{Ch.Sem.} Milk_c \\ &+ \pi_3 \Delta_{i,j}^{Ch.Sem.} Meat_c + \pi_4 \Delta_{i,j}^{Ch.Sem.} Tubers_c\end{aligned} \tag{2.10}$$

Whereby $\Delta_{i,j}$ computes the difference between prices for Index child (i) minus younger sibling (j), for each distinct $ChildSemester$ critical period: 6-11 months, 12-17 months, 18-23 months, 24-29 months, 30-35 months, for the four selected food items.

Figure 1 shows the evolution of prices between 2001 and 2009 for the main food groups used in the analysis (Cereals, Dairy, Meat and Tubers). We can see that the

[17] The geo-political map divides the country into 25 Departments, in turn disaggregated into provinces and districts; INEI collects information for the capital districts of each Department and for other cities.

[18] INEI reports estimates for the main price indices that conform the consumer price index (8 categories), including the food and beverages price index, in turn disaggregated in 14 sub-categories. Using all the sub-categories would generate 70 IVs; with a dataset of only 450, this would be extremely demanding of the data.

[19] Figures computed using information collected in the third wave of data in 2009. Note that we do not have information on food consumption for individual children, only at the household level.

food price 'crisis' began in late 2005 and prices rose for most of 2006-2008, and how this corresponds to the age groupings of the cohorts in the sample.

We also include idiosyncratic shocks as instruments. While the mechanisms of these shocks is similar to the price changes - that is, reductions in the nutritional intake of a specific child during the critical developmental period -, they complement the price data well because they introduce further household-level heterogeneity as well as a different source of nutritional variation. We include three different such events, one from the 2002 wave and two from the 2009 wave. We do not include shocks from the 2006 wave as they could have affected both children during their critical period.

From the 2002 wave, we use a dummy for whether any - from a list of - negative events affected the household between the index child pregnancy period and the time of the survey, when the index children were aged between 6 and 20 months. By definition, these events could only have directly affected the index child, as the younger sibling was yet to be born. Additionally, from the 2009 round, we include indicator variables for whether the household was affected by an event of sever frost, as well as whether a member of the household other than the child died or suffered an illness between 2006 and 2009. Either of these events could not possibly have affected the nutrition of the index child, given that we only use their 2006 height-for-age score.

By virtue of being self-reported these individual shocks could be affected by biased responses - for example, only vulnerable households whose livelihoods have been severely affected report the incidence of the shocks. Testing the validity of self-reporting is challenging but information on frost shocks can shine some light on to this. To the extent that frost is a local phenomenon, we should expect frost events to be correlated across households living in the same community. Reassuringly, we find a high correlation, 0.64, between individual self-reports and the average frost reporting in each of the 27 districts in our sample, a value significantly higher than the correlation that would arise by chance. [20]

[20]We carry out two types of tests. First, we compute the correlation coefficient between individual frost incidence and the average at the district level (or *ubigeo*) and find it to be as high as [0.64]. This is substantially higher than the correlation arising between a random variable and its community average - in simulations we find this correlation to be around [0.18]. Secondly, using seemingly unrelated biprobit methods, we confirm that these differences are significant. Finally, we confirm the strong correlation

Moreover, a further concern arises if individual-shocks are correlated with household characteristics. In columns (1) to (3) of Table 2.4 we test this possibility for the three household-specific shocks: bad events in 2002, frost events and illness shocks in 2009 respectively. We estimate a linear probability model of the incidence of each of the household shocks, using time-invariant household characteristics[21] and cluster fixed effects as controls. There are few patterns common to all three shocks and few variables that are significant: The age of the caregiver reduces the likelihood of incidence of bad-events, and Spanish speakers are less likely to be affected by an illness shock in the 2006-09 period. While reporting a frost event in your community appears more likely among households with higher consumer durables, in other words, the wealthier households. Yet, it is not clear how such correlations should be interpreted: whether they illustrate self-reporting patterns, perceptions of severity of the shock, or a genuine correlation between household characteristics and incidence and severity of the shocks.

While household shocks are arguably less exogenous to the household than the price data - their incidence and severity potentially correlated with unobserved household characteristics - this concern is reduced when using a sibling-difference specification. Just as a fixed-effect models, the sibling-difference methodology controls for any time-invariant household unobservables. And in turn, there is no reason to presuppose any of the household shocks to be correlated with child-specific unobservables. Nevertheless, in the next section, we return to this point, to test whether the correlations reported in Table 2.4 might be biasing our core results.

Our sample, by construct, includes mostly households at a stage in their lifecycle with active fertility. If shocks affect subsequent fertility choices, our estimates might be biased; our instrument could simply be documenting the impact of fertility choices on early nutritional and cognitive development of the siblings. In columns (4) and (5) in Table 2.4 we explore the extent to which fertility choices are affected by the shocks. We use two measures of fertility: "if any additional siblings were born after the younger sibling" - there are 49 such households - and "how many additional children after the

of frost reporting at the local level, when running a simple OLS determinants regression of the frost incidence on district dummies. We find that the dummies explain 40% of the frost incidence variation, and that they are jointly significant.

[21]We use the same vector of household characteristics as used in the pooled OLS model.

younger sibling" (ranging from 0 to 4). Note that these variables measure fertility after 2002 (the earliest possible date the younger sibling would have been born). To the household controls we add, the three sets of instruments likely to affect fertility after 2002: price shocks, frost and illness shocks.[22] We find that neither frost nor illness shocks affect fertility choices, but price shocks are found to be jointly significant. How these patterns might affect our estimates is explored in the robustness section.

2.4 Results

We report pooled OLS and sibling-difference OLS results in Table 2.5. For the pooled OLS model, column 1, we include a parsimonious set of child and household-level controls[23] as well as community fixed effects. Recall that the cross-sectional pooled OLS, is likely to be biased due to unobserved household heterogeneity. Columns 2 to 4 in turn report results for the sibling-difference model specification following equation 2.6. Column 2 includes controls for differences between siblings in terms of age, sex and birth order, while column 3 adds cluster fixed effects, which are designed to capture community-level changes.[24]

Table 2.5 shows that both pooled OLS and sibling-difference models significant and - as expected - positive impact of nutrition on cognitive development and that sibling-difference estimates are robust to the inclusion of controls for community changes. Consistent with the descriptives statistics, we find that the nutritional effect is smaller when we control for unobserved household heterogeneity. However, with coefficient estimates of 0.099 and 0.083 respectively, comparisons between the pooled OLS and the baseline sibling-difference model suggest that the importance of unobserved household heterogeneity might be more modest than commonly assumed.[25]

[22] That is, we exclude shocks measured in 2002, that correspond to events before the birth of the younger sibling.

[23] These include, child's sex and age, mother's years of schooling, household size, household per capita non-food expenditure expressed in logs and household wealth index

[24] Note that in column 1, robust standard errors are clustered at the household level. All other results throughout the paper, including columns 2 to 4 in table 2.5, report standard errors corrected for clustering at the *index child age-community* level. All further specifications report standard errors corrected for cluster and index age-cohort specific correlations. In other words, our inference testing is robust to unobserved correlation between children of the same age living in the same district or cluster

[25] In this sample, it would appear that there is little time-invariant household heterogeneity - taken

Finally, column 4, re-estimates the core model specification for the sub-sample of paired-siblings aged 4 to 5 years only. Siblings in this reduced sample have an age profile that most resembles the index; more importantly all children are below the schooling age (6 in Peru), implying the effect of unobserved schooling investments can be disregarded in this model. We draw comfort from the fact that, in spite of a substantially reduced sample size, 330 paired-siblings, the coefficient remains significant and virtually unchanged. Even though the core model specification does not control for schooling investments, the evidence suggests that schooling investments in the full sample are uncorrelated with height-for-age. We revisit this issue in section 2.5.

In table 2.6 we report the results for the IV GMM sibling-difference model specification. For comparability, in column 1 we duplicate the results from the sibling-difference OLS with full controls (column 3 in table 2.5). Column 2 reports estimates for the model where we use differences in food prices during the first 36 months of life of a child as IVs, while column 3 presents results when using child-level shocks as excluded instruments. While column 4 combines both sets of instruments. Finally, column 5 includes household time-invariant controls to test for the potential endogeneity of some of our shocks. [26] [27]

Diagnostics of the first-stage regressions, indicate that our set of instruments are reasonably strong. When using price changes only as instruments, we obtain a Kleibergen-Paap Wald rk F statistic of 14.53, which is above the critical value for a maximum IV bias of 10 per cent (Stock and Yogo, 2005). Whilst this is a strong result by itself, we also tried combining these instruments with the child-level economic shocks as this introduces more variation across locations - recall our price-shock data are specific to each child, by virtue of their age, but prices are measured at the community level. The combined set of

care of by the sibling-difference model -, that is not already captured by the household level controls used in the pooled OLS model. In particular, while a bi-variate regression of PPVT on HAZ yields a large coefficient consistent with figure 2-4, the pooled OLS coefficient estimates drop substantially when we control for household assets and non-food consumption. In other words, the dramatic differences in slope in figure 2-4 are mostly accounted by observed measures of household wealth rather than unobserved household heterogeneity.

[26] In all cases, results were obtained from a two-step efficient generalized method of moments (GMM) estimator using STATA routines created by Baum et al. (2010). Due to the clustered nature of our data, a robust Kleibergen-Paap Wald rk F statistic is reported to test for the presence of weak instruments

[27] In column 5, we return to the possibility that the IV estimates are driven by the household correlates reported in Table 2.4. We conclude that this is not the case. Including controls for household time-invariant characteristics has little impact on our estimates.

instruments - column 4 - are even stronger with a maximum IV bias of just above 5 per cent. In column 3, we show the contribution of idiosyncratic shocks alone, but find as instruments they do not pass the 'weak IV' test, implying that second-stage inferences will be invalid and an point estimates are likely to include a relative bias between 10-20%. Additionally, we test for the exogeneity of our instruments and find that all three specifications also pass the overidentification test (Hansen J-test).

Second-stage results uncover a nutritional effect that is large in magnitude and strongly significant. When using food prices, point estimates suggest that one standard deviation increase in height-for-age yields higher PPVT scores by 17% of a standard deviation. When combining the two sets of instruments, point estimates rise to 21% of a standard deviation. Point estimate from using shocks only as IVs are even higher at 24% but the coefficient is much less precisely estimated. Even though the latter estimates contain substantial 'weak IV' bias, it is interesting to note that the point estimates are not very different from the results using the price data. Even though the two sets of instruments exploit a different source of nutritional variation, the similarity in coefficients suggests that both instruments might be capturing a single mechanism.

Other studies applying similar sibling-difference IV strategies (e.g. Alderman et al. (2006)), have also obtained coefficients substantially higher than their OLS estimates. The change in parameter estimates could be attributed to the increased attenuation bias that results from applying differentiae methods. However, comparisons of pooled OLS and the sibling-difference models suggest that in our data measurement error might only be part of the story. While one can only speculate, a plausible explanation is that parents allocated household recourses and investments in a compensatory manner (as suggested by Behrman et al. (1982)). That is, if parental attentions are dedicated to children with poorer health and lower height-for-age, the resulting higher child-specific unobserved cognitive investments will reduce the correlation between cognition and nutrition in the sibling-difference specification. Our instrumentation strategy therefore yields a parameter robust to the mediation of compensatory households.

2.4.1 Discussion of first stage results

We now move on to discuss the effect of the instruments on height-for-age as a proxy for nutritional achievement, a question of intrinsic policy interest. Columns 1 to 3 in Table 2.7 report the first-stage results for the three IV model specifications reported in 2.6. Because the height-for-age difference variable is constructed as the difference between the Index HAZ and the HAZ of the sibling, a shock affecting the former would have a negative effect on ΔHAZ. First-stage regressions for the idiosyncratic shocks are reported in column 2. We find as predicted that adverse shocks to the household that affected the index child only, reduce the difference in height-for-age between siblings, and that illness and death of household members in 2009 increase the nutrition gap. Incidence of frost, a common weather shock in the Andean highlands, has the correct sign but is only imprecisely estimated.

Results for the effect of food prices are reported in column 1. Recall the food prices are split into semesters of early childhood defined for each child. Even though the full set of food price changes are sufficiently strong instruments, only a few items are statistically significant on their own. Cereal prices in the period 30-35 months, and tubers and meat prices seems in the early years reduce height-for-age of the sibling, resulting in a positive coefficient on height-for-age differences. On the other hand, meat prices in the period 30-35 months has a negative impact on the difference between sibling heights. This is somewhat puzzling, and we suspect multicollinearity might be affecting both the sign and significance of some of the price variables.

As a better measure to determine the importance of the price data on nutrition, we use tests of joint significance across groups of variables (see Table 2.8). When testing across time periods for joint significance by food groups, we find that the two strongest food items are cereals and meat. We also tested across food items for joint significance by time period, in order to investigate whether some periods were more important. For example Glewwe and King (2001) searching for a 'critical period' found tentative evidence that the period 18-24 months was most important for a sample of children in the Philippines. The period 24-35 months has been identified as a more sensitive period for nutritional development (e.g. Alderman et al. (2006) on Zimbabwe). We find joint significance of the first and last periods under scrutiny (6-11 months, and 30-35 months). However, one should be cautious when interpreting these results, since in our context, the significance of a period will also be determined by the magnitude of the price changes experienced. Indeed, in terms of the timing for our cohorts, the 30-35 months period for the young siblings corresponds with the years 2006-8, the period in which the largest price increases where experienced.

To understand further the role played by price changes in child nutrition, it is helpful to recognise that a number of factors could be at work here. For *normal* goods we expect price increases to be related to reduced consumption and nutritional intake, however, when households are net-producers, profit effects might lead to increases in consumption. We use this understanding to explore further the mechanisms behind our instruments. We follow two strategies.

First, we explore these channels explicitly by testing the extent to which the price changes affect household consumption and its potential heterogenous effect. Secondly, we seek to uncover similar heterogeneous patterns in the first stage of our sibling-difference instrumentation methodology.

For the 2002 and 2006 rounds we have information on household consumption data. We run household a first-difference specification of the impact of price changes on household food expenditure growth between survey rounds. And include a range of household and community level controls, all measured as changes between rounds. [28] While the model specification and period of analysis are slightly different from the one used in the sibling-difference IV specification, the analysis can be informative of the type of mechanisms that link prices and household consumption. For that purpose, we focus on the potential heterogeneous effects of prices by exploring its interaction with rural/urban and farmer/non-farmer status of households.

Column (1) in table 2.9 presents the base model without interactions, while columns

[28]To run these regressions we construct a different type of price series from the one used in our sibling-difference IV specification. We compare the changes in household consumption between rounds 2002 and 2006 with the price changes recorded between rounds. We use the local price on the date of interview of a household for constructing the household specific price series.

(2) and (3) report estimates for models with interactions for rural households, and farming households respectively. In rural Peru, meat is primarily an imported item, and to some extent so are bread and cereals: little home production would be expected for either of these items. On the other hand, tubers (including potatoes) is a typical Peruvian rural home produce; the excess production often sold locally. Diary products are possibly a mixture: with lama diary products sourced locally, but cattle produce imported from urban areas.

Column (1) shows that cereals have a significant negative impact on household consumption, but other food items are insignificant. While interacting prices with farming status yield little further insight (column (3)), column (2) shows clear heterogenous patterns between rural and urban households. In rural areas, cereals and meat have a negative effect on consumption (though only the former is significant) consistent with being imported products. Inflation in diary products and tubers, on the other hand, lead to higher consumption growth. Moreover, the positive profit effects, are only found in rural areas, with tube inflation reducing consumption in urban areas.

Having established that price changes in our food items can have important, but complex, effects on household consumption, we explore similar type of heterogeneities in our sibling-difference IV specification. We augment equation 2.10 to incorporate interactions between prices and rural/urban status:

$$\Delta_{i,j}^{06,09} Prices_{t,h} = [Rural/Urban] \bigotimes [\pi_1 \Delta_{i,j}^{Ch.Sem.} Cereals_c + \pi_2 \Delta_{i,j}^{Ch.Sem.} Milk_c + \pi_3 \Delta_{i,j}^{Ch.Sem.} Meat_c + \pi_4 \Delta_{i,j}^{Ch.Sem.} Tubers_c] \tag{2.11}$$

Second stage results for this equation are reported in table 2.10. [29] With a vector of 40 instruments, first-stage mechanisms are difficult to uncover. To allow a meaningful interpretation of the effect of price shocks on sibling HAZ, however, we explore the

[29] The table also includes results for the interaction model with household farming and non-farming status. It is interesting to note both of these interaction models yield very strong instruments - well above the critical values required for strong instruments and higher than our core IV model -, yet second-stage point estimates are similar to our earlier findings. This is suggestive that our core IV estimates are relatively free from weak IV bias.

predicted model. We compute $H\hat{A}Z$ using the estimates for individual items of equation 2.11, and create predicted HAZ impact for individual price food items and by location. For example, this approach allows us to compare the HAZ predicted impact of inflation in tubers for rural versus urban areas.

Table 2.11 reports predicted HAZ impacts, alongside the corresponding F-tests of joint significance in the first-stage equation. A positive effect indicates that as prices for the index child (versus the sibling) rise, so will its relative nutritional status. The results are interesting in that they mirror heterogenous patterns reported in Table 2.9 for household consumption. Tubers have a positive effect on sibling nutrition, but in only rural areas, with negative effects among urban households, while cereals and meat behave like normal good.

The evidence in table 2.11 is far from conclusive; confidence intervals reported are not sufficient to assess significance of the reported predicted HAZ impacts - they do not incorporate the uncertainty in estimation; while some coefficients remain counterintuitive. Nevertheless, on balance the evidence is suggestive that our instruments generally affect child nutrition in patterns consistent with economic theory, and similar to how consumption growth is affected.

Overall, results presented in this section indicate that the food price crisis had a significant effect on nutritional outcomes of siblings, which resulted subsequently in lower cognitive achievement. A similar mechanisms is uncovered when using household-specific shocks as instruments. In the next section, we turn to test the robustness of these findings.

2.5 Robustness Checks

An extensive literature on health and nutrition has also explored the foetal and prenatal period. Most of the studies agree that deficiencies during pregnancy can have health effects in the medium- and long-term (Godfrey and Barker, 2000; Almond, 2006), although a few studies have also found short-term effects (Maccini and Yang, 2009). In order to keep the number of instruments manageable, we exclude price changes for the pre-natal and breastfeeding period. While appearing to some extent *ad hoc*, this choice is corroborated by further robustness tests. When we re-estimated our main IV model (column 2 in 2.6) with the two pre-natal and the breastfeeding semester replacing the two semesters in the third year of the critical period, we found that - with a Kleibergen-Paap F-statistics of 6.0 - the revised IV set was sufficiently strong to pass the 'weak IV' test. However, the pre-natal and breastfeeding prices had only limited power in explaining height-for-age differences.[30]

While the exogeneity of the shocks is not suspect, at least for the price data, this does not guarantee meeting the exclusionary restriction (2.9). (Glewwe and Miguel, 2008). A particular concern is the possibility that price and household level shocks could have persistent effects on cognitive achievement other than through nutrition, through a direct effect on household cognitive investments. In particular, shocks or increases in food prices in one period could crowd out expenditure in educational items in the next period through a reduction in savings or if the household had to sell assets as a result of the event. If this was indeed the case, our instrumentation strategy would be invalid, because our instruments would be correlated with unobserved cognitive investments; our IV GMM estimates merely quantifying the effect of such a reduction in investments rather than the nutrition effect. However, it should be noted that our core IV GMM sibling-difference model is robust to certain types of persistence. Indeed, if a shock has not only an immediate effect on household nutritional investments but also a *permanent* effect on household cognitive investments, the within-sibling specification would partly capture this phenomenon. That is, to the extent that a shock is permanent, both siblings should be similarly affected by the reduced household wealth - this would be the case for shocks affecting the index child's critical period, but not that of the sibling (given that we have already measured the index child's cognitive development at an earlier period).

To address this concern, we augment our core model specification to include controls

[30]In particular, we find the semester (-2), semester (-1) and semester (1) have F-Statistics of 9.4, 5.4 and 3.7 respectively, with only the first group statistically significant. Even though estimates might contain substantial 'weak IV' bias, it is interesting to note that the second-stage estimates yield a nutrition coefficient that was positive and significant. Results not reported in tables, but can be requested from the authors.

for changes in household assets and non-food consumption. To the extent that shocks affect cognitive investments through reduced household wealth, the additional controls 'switch-off' the argument of the invalidity of our instruments due to their persistent effect.

Table 2.12 reports estimates for the augmented model specification. Columns 1 and 2 reproduce the core results from the previous section when using prices only and prices and shocks as instruments. Columns 3 and 4 expand the core model to include changes in household assets between 2006 and 2009. Since changes in assets might be an imperfect proxy for changes in cognitive investments, columns 5 and 6 report results that also include changes in non-food consumption. Finally, columns 7 and 8 include also controls for changes in households assets between 2002 and 2006.[31] The inclusion of earlier changes is motivated by the concern that some of the price changes might have affected income wealth before 2006, however, by controlling for these early changes we are also eliminating some of the exogenous variation of the instruments that we wanted to exploit. It is therefore not surprising to find that the strength of the IVs is reduced furthest in columns 7 and 8.

In spite of the importance of controlling for household changes in wealth, as a proxy for changes in cognitive investment, results in table 2.12 show little evidence that they affect PPVT sibling differences significantly; while changes in assets have a large coefficient, these are also very imprecisely estimated. When turning to the nutritional effect we find that our core results remain very robust. On the one hand, we find that the inclusion of the additional controls has only a limited effect on the first-stage strength of the instruments; indeed instruments remain 'strong' in all model specifications.[32] On the other hand, point estimates remain remarkably stable, largely staying in the 17-21% range. We can conclude that if our set of instruments have persistent effects on household welfare, these phenomenon do not affect the nutrition-cognition estimates.

[31] Ideally, we would have also included controls for changes in non-food consumption between 2002 and 2006, but no data on non-food consumption were collected in the 2002 wave.

[32] The model in columns 7 has the lowest Kleibergen Paap F-Statistic, 10.68, marginally below the Stock-Yogo critical value for an IV bias of 10%.

2.5.1 Time Allocations and Endogenous Fertility

A related concern is the possibility that shocks have a contemporaneous effect on cognitive investments. If parents are forced to increase their labour supply due to shocks, this might reduce child-specific cognitive investments. Table 2.13 presents a range of strategies to explore this possibility.

First, we include controls on changes in the main activity of the mother. If shocks are sufficiently severe, households might refocus their activities. We create a variable that documents changes from housework to outside work between 2006 and 2009 (recall that the price shocks take place primarily during the 2006-08 period).

Secondly, rather than changes in activity we would like to control for changes in time allocations of the mother or household head. Regrettably, that information is not available for all rounds.[33] However, we use instead a measure of household time allocations based on the work activities of the children in the household. We create a composite variable of the amount of "hours spent on paid and unpaid work" for all children aged between 4 and 17 (including 'index' and 'young sibling') and how it changes between the 2006 and 2009 rounds. Thirdly, if shocks reduce cognitive investments directly, we might be able to control for this using the PPVT score of the index child in 2009 - when aged 7-8. [34]

Table 2.13 also explores the effect of endogenous fertility. If shocks affect household composition and in particular new births, our IV strategy might be simply documenting household fertility choices. Earlier in section 2.3.3 we reported some evidence that the price shocks might be correlated with birth rates after the younger sibling. We include dummies of the number of additional children born in the household after the younger sibling.

Results in Table 2.13 show our core estimates to be robust to the additional con-

[33]This information was only collected in round 2.

[34]It should be noted that this robustness test will also control for potential spill-over effects between siblings. That is, if a child's cognitive development not only depends on the parental investments but its own siblings, any negative effects on the index cognitive score might also reduce the younger sibling's score. Controlling for the changes in PPVT score of the index child between the 2006 and 2009 rounds will capture any indirect impact of the shocks through the spill-over mechanism.

trols. Their inclusion individually or collectively have no substantial effect either on the strength of the instruments or the significance of the HAZ coefficient.

Changes in mother activity (from housework to outside work) increase the PPVT differences (see columns (1) and (2)), suggesting that this might be negative for cognitive stimulation of the younger sibling. [35] While increasing the number of hours worked collectively by the children in the household affects negatively the index child (more than its sibling), possibly because being older his increase in hours is largest (columns (3) and (4)). It seems that adding controls for time allocations seems to have little impact on the HAZ coefficient. Consistent with this, controlling for the index PPVT score in 2009, does not affect HAZ impact (columns (5) and (6)) either. This is reassuring as it suggests that our instruments are not persistent enough to affect PPVT scores in 2009, either through assets, time allocations or sibling spill over effects. Finally, results in columns (7) and (8) show that endogenous fertility is not driving our IV estimates. While the number of siblings does affect PPVT differences, the HAZ coefficient remains unaffected.

2.5.2 School and Pre-School Investments

A further source of concern in our analysis is with respect to child-specific school investments. On the one hand, some children in our sample are already of schooling age - 6 years being the standard age of school enrolment. On the other hand, even if not of schooling age, a large proportion of children may already be enroled in pre-school. Failing to control for differences in schooling between siblings could be biasing our estimates. The concern is that children that are physically small and possibly frail might have their enrolment delayed, and if so, at the time of testing, these children are likely to have a lower PPVT score. The instrumentation strategy is particularly vulnerable to this critique, since it has been shown that early stunting is commonly associated with delayed school enrolment(Glewwe and Jacoby, 1995; Alderman et al., 2001; Glewwe et al., 2001).

[35] Its inclusion reduces the power of the instrument and the significance of HAZ, but this might be due to smaller sample size. Recall that we have information on time spent at work by the mother, but only for the 2006 round. When we include this 2006 variable, the drop in sample size is not as important: the HAZ coefficient remains positive and significant, around 0.22.

Our IV estimates could therefore be capturing the nutritional effect of delayed enrolment rather than the nutrition-cognition link.

Table 2.14 reports a range of robustness checks on our core IV model specification aimed at addressing the problem of omitted cognitive investments originating from delayed pre-school or school enrolment.[36] The pre-school and school data available for the index child are extensive, but for siblings these data are largely unavailable. On pre-schooling we only have information for the index child, while for primary school, we only know whether a sibling was enroled in 2009 but not their age of enrolment. The data limitations imply that for age of enrolment we can only include controls for the index child; we consider this variable a proxy for sibling differences in age of enrolment. If an index child has a particularly early age of enrolment, it is arguably likely that the sibling might have a later enrolment; in our regressions, we would then expect the *levels* measure of age of enrolment of the index child to be positively correlated with the PPVT score *sibling difference* measure.

Columns 2 to 4 in Table 2.14 report IV estimates when we include pre-school enrolment and age of enrolment for the index child only. On its own pre-school enrolment has a large an positive effect on PPVT differences (see column 1). Recall that the PPVT difference variable is constructed as the difference between the Index PPVT score minus the PPVT score of the sibling. Pre-school enrolment of the index child therefore increases the gap between the index and its sibling. Similarly, and as expected, ΔPPVT is decreasing in the age of preschool enrolment (column 3), although when both enrolment and age are included, delayed enrolment (from 3 to 5 years of age) does not have a significant effect (column 4). Columns 5 and 6 expand our set of controls to include primary school. Similar to pre-school enrolment, we find that ΔPPVT is decreasing in school enrolment of the sibling and the age of school enrolment of the index child. At the same time, we find that across all of alternative specifications, both first-stage and second stage IV estimates remain robust. When we control for age and enrolment in school and pre-school, we obtain a Kleibergen-Paap F-statistic of 19.99 - suggesting an IV bias of just below

[36]Note that for our OLS model we reported estimates for the restricted sample of children aged 4 and 5. This sample excludes all children of schooling age but not children that might be enrolled in pre-school. We do not apply this approach to the instrumental variable model, because the restricted sample of 330 observations is too small to produce reliable 2SLS given our large number of instruments.

5%, and find that a one standard deviation increase in height-for-age improves PPVT by 19.2% of a standard deviation.[37] This is a remarkable finding in that, though both schooling and pre-schooling are strong determinants of sibling cognitive differences, their inclusion has no meaningful effect on our estimates of the nutrition-cognition parameter.

2.6 Conclusions

The importance of a good start in life cannot be overstated, as early child outcomes have strong predictive power for future life chances. In the context of Peru, where many children are malnourished, and affected along with many other countries by the food price crisis of 2006-2008, we have revisited the nutrition-cognition relationship. We have provided evidence on the link between nutrition and cognitive achievement for a group of pre-school age children, by using a within-sibling estimation strategy combined with instrumental variables in order to convincingly determine causality.

We find that an increase in the Height-for-Age z-score of one standard deviation - keeping other factors constant - translates into increases in the Peabody Picture Vocabulary Test (PPVT) score of 17%-21% of a standard deviation.

Our instruments include both covariate and idiosyncratic shocks, and in robustness checks, we show that controlling for household assets, time allocation and endogenous fertility does not affect our estimates. If shocks have a cognitive effect through those mechanisms, they are not affecting our estimates of the nutritional channel. Further, we test whether our results are driven by selective school enrolment. Our results remain robust to these sensitivity tests, although we find significant effects of schooling on cognitive development.

The results are of policy concern, not least because this sample of children are only 4-5 years old and we do not yet know the long-term effects of their nutritional deficits. The findings suggest that early nutritional interventions can have substantial cognitive benefits, more so considering that, unless addressed, early deficits are likely

[37] Table 2.14 includes both prices and shocks as instruments. Results remain unchanged when using prices only. Indeed, when we re-estimate column 5 with prices only, we also obtain a point estimate of 0.19, and a Kleibergen-Paap F-statistic of 16.7.

to be followed by further deficits in human capital accumulation resulting from delayed school enrolment, poor educational progression and early drop-out rates.

References

Alderman, H., Behrman, J. R., Lavy, V., and Menon, R. (2001). Child health and school enrollment: A longitudinal analysis. *The Journal of Human Resources*, 36(1):185–205.

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

Almond, D. (2006). Is the 1918 influenza pandemic over? long-term effects of ¡i¿in utero¡/i¿ influenza exposure in the post-1940 u.s. population. *Journal of Political Economy*, 114(4):672–712.

Almond, D. and Currie, J. (forthcoming). *Human Capital Development before Age Five*, volume 4 of *Handbook of Labor Economics*. North Holland.

Ashenfelter, O. and Krueger, A. B. (1994). Estimates of the economic returns to schooling from a new sample of twins. *American Economic Review*, 84(5):1157–73.

Baum, C., Schaffer, M., and Stillman, S. (2010). Ivreg2: Stata module for extended instrumental variables/2sls and gmm estimation. Statistical Software Components, Boston College Department of Economics.

Behrman, J. R. (1988). Nutrition, health, birth order and seasonality : Intrahousehold allocation among children in rural india. *Journal of Development Economics*, 28(1):43 – 62.

Behrman, J. R. (1996). The impact of health and nutrition on education. *The World Bank Research Observer*, 11(1):23–37.

Behrman, J. R., Cheng, Y., and Todd, P. E. (2004). Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *The Review of Economics and Statistics*, 86(1):108–132.

Behrman, J. R. and Lavy, V. (1994). Children's health and achievement in school. Living Standards Measurement Survey Study Working Paper.

Behrman, J. R., Pollak, R. A., and Taubman, P. (1982). Parental preferences and provision for progeny. *The Journal of Political Economy*, 90(1):52–73.

Berlinski, S. and Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, 14(3):665–680.

Cuesta, J. and Jaramillo, F. (2010). Taxonomy of causes, impacts and policy responses to the food price crisis in the andean region. Research department working paper.

Cueto, S., Escobal, J., Penny, M., and Ames, P. (2011). *Tracking Disparities: Who Gets Left Behind? Initial Findings from Peru Round 3 Survey*. Young Lives Country Report.

Cueto, S., Leon, J., Guerrero, G., and Munoz, I. (2009). Psychometric characteristics of cognitive development and achievement instruments in round 2 of young lives. Young Lives Working Paper Series 15, University of Oxford.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. *The American Economic Review*, 97(2):31–47.

Cunha, F., Heckman, J., and Lochner, L. (2006). *Interpreting the Evidence on Life Cycle Skill Formation*, volume 1 of *Handbook of the Economics of Education*, chapter 12, pages 697–812. Elsevier.

Cunha, F. and Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *J. Human Resources*, 43(4):738–782.

Dercon, S. and Outes-Leon, I. (2009). Survey attrition and attrition bias in young lives. Young Lives Technical Note 5, University of Oxford.

DHS. *Peru, 2004-06 - Final Report Continuous (2004-2006) (Spanish)*.

Dunn, L., Padilla, E., Lugo, D., and Dunn, L. (1986). *Manual del Examinador para el Test de Vocabulario en Imgenes Peabody (Peabody Picture Vocabulary Test) Adaptacin Hispanoamericana (Hispanic-American Adaptation)*. Minnesota: AGS.

Flores, E. and Escobal, J. (2008). An assessment of the young lives sampling approach in peru. Open Access publications from University of Oxford http://economics.ouls.ox., University of Oxford.

Glewwe, P. and Jacoby, H. G. (1995). An economic analysis of delayed primary school enrollment in a low income country: The role of early childhood nutrition. *The Review of Economics and Statistics*, 77(1):156–169.

Glewwe, P., Jacoby, H. G., and King, E. M. (2001). Early childhood nutrition and academic achievement: a longitudinal analysis. *Journal of Public Economics*, 81(3):345–368.

Glewwe, P. and King, E. M. (2001). The Impact of Early Childhood Nutritional Status on Cognitive Development: Does the Timing of Malnutrition Matter? *World Bank Econ Rev*, 15(1):81–113.

Glewwe, P. and Miguel, E. A. (2008). *The Impact of Child Health and Nutrition on Education in Less Developed Countries*, volume 4 of *Handbook of Development Economics*, chapter 56, pages 3561–3606. Elsevier.

Godfrey, K. M. and Barker, D. J. (2000). Fetal nutrition and adult disease. *Am J Clin Nutr*, 71(5):1344S–1352.

Grantham-McGregor, S. (1995). A review of studies of the effects of severe malnutrition on mental development. *Journal of Nutrition*, 125(8 Suppl):2233S – 2238S.

Grantham-McGregor, S. and Baker-Henningham, H. (2005). Review of the evidence linking protein and energy to mental development. *Public Health and Nutrition*, 8:1191 – 1201.

Grantham-McGregor, S., Powell, C., Walker, S., and Himes, J.

Grantham-McGregor, S., Walker, S., Chang, S., and Powell, C. (1997). Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *Am J Clin Nutr*, 66(2):247–253.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33):13250–13255.

Horton, S. (1988). Birth order and child nutritional status: Evidence from the philippines. *Economic Development and Cultural Change*, 36(2):341–354.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–26.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. (2009). The impact of nutrition during early childhood on education among guatemalan adults. *Economic Journal*, 119:734–763.

Paxson, C. and Schady, N. (2007). Cognitive development among young children in ecuador: The roles of wealth, health, and parenting. *Journal of Human Resources*, 42(1).

Pollitt, E., Gorman, K. S., Engle, P. L., Martorell, R., Rivera, J., Wachs, T. D., and Scrimshaw, N. S. (1993). Early supplementary feeding and cognition: Effects over two decades. *Monographs of the Society for Research in Child Development*, 58(7):i–118.

Rosenzweig, M. R. and Wolpin, K. I. (1995). Sisters, siblings, and mothers: The effect of teen-age childbearing on birth outcomes in a dynamic family context. *Econometrica*, 63(2):303–326.

Schady, N., Galiani, S., and Souza, A. P. (2006). Early childhood development in latin america and the caribbean [with comments]. *Economía*, 6(2):185–225.

Stock, J. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression.*, pages 80–108. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenber. Cambridge University Press.

Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33.

Todd, P. E. and Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, 1(1):91–136.

von Braun, J. (2010). Rising food prices- what should be done? Ifpri policy brief.

Wilson, I., Huttly, S., and Fenn, B. (2006). A case study of sample design for longitudinal research: Young lives. *Int. J. Social Research Methodology*, 9(3):361–65.

World Bank (2006). Repositioning nutrition as central to development: A strategy for large-scale action. Directions in development.

Figure 2-1: Histogram: PPVT Raw Score, Index Child, 2006

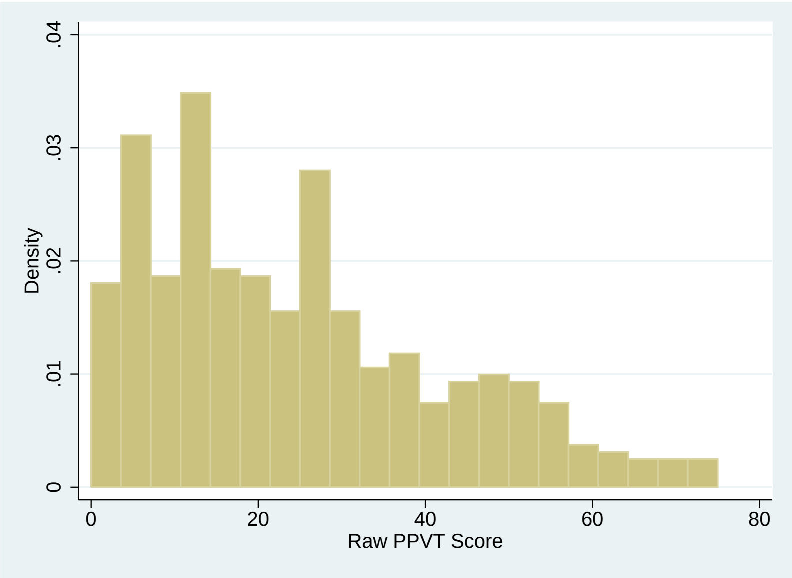


Figure 2-2: Histogram: Standarised PPVT Score, Index Child, 2006

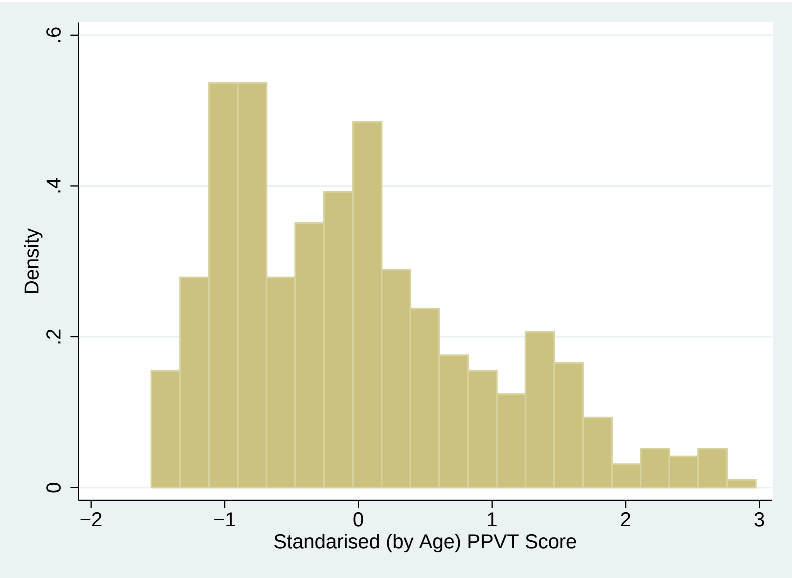


Figure 2-3: Histogram: Sibling Difference, Standarised PPVT Score, Index Child (2006) - Sibling(2009)

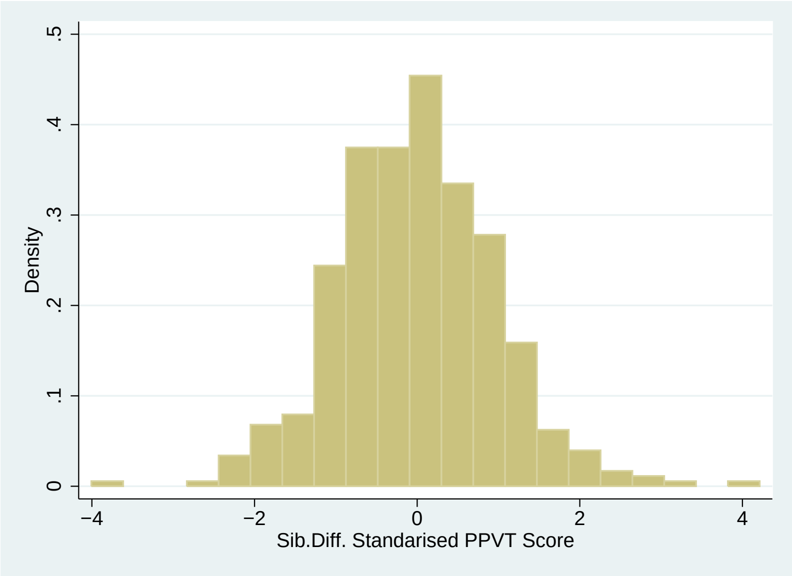


Figure 2-4: Non-parametric nutrition-cognition relationship: pooled sample vs siblings-difference sample

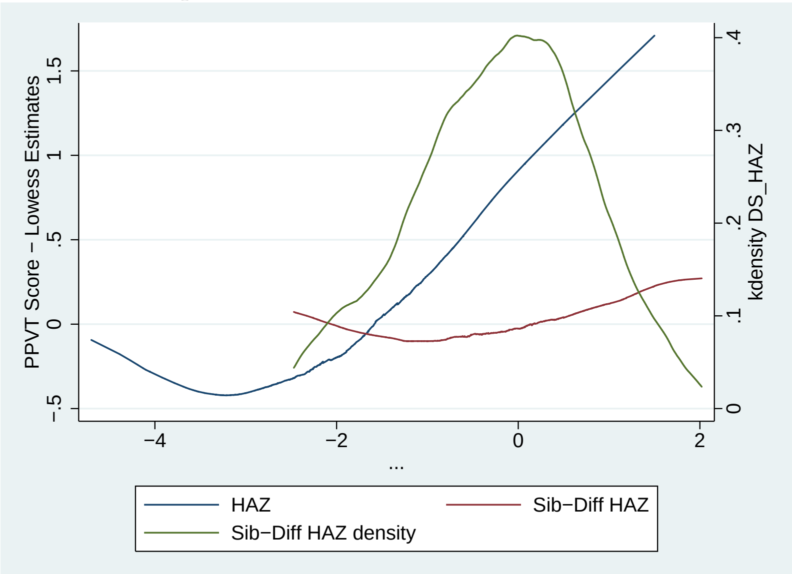


Figure 2-5: Evolution of food prices in Peru: 2000-2009

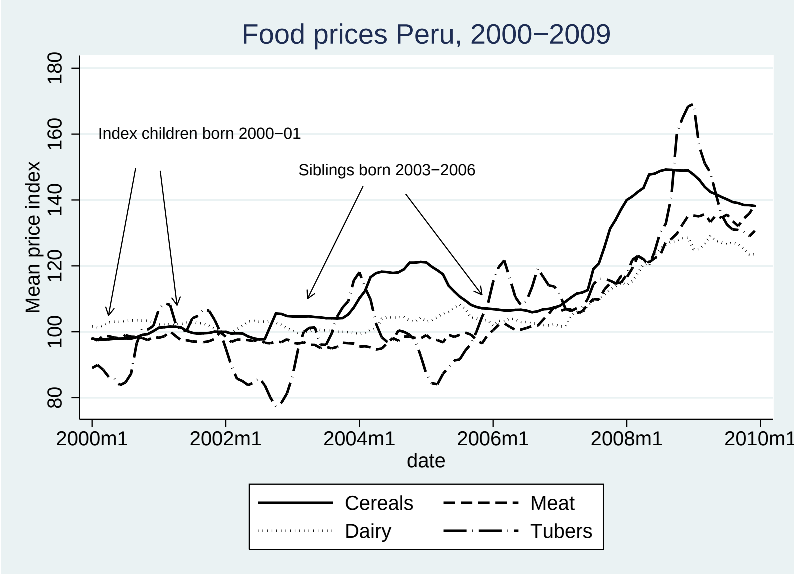


Table 2.1: Age Distribution: Index versus Younger Siblings, 2006 and 2009 YL Waves

	Age of Younger Sibling, 2009					
	3y	4y	5y	6y	7y	Total
Age of Index, 2006						
4y	1	40	60	17	0	118
5y	7	95	135	77	6	320
6y	0	3	4	5	0	12
Total	8	138	199	99	6	450

Table 2.2: Descriptive Statistics: Full versus Paired-Siblings Sample

		Paired-Sib. sample	Rest of YL sample	Diff.
Age of mother in 2001	Mean	25.41	27.49	***
	Std.Err.	0.287	0.180	
Mother's years of schooling	Mean	6.57	8.15	***
	Std.Err.	0.216	0.115	
Height-for-age of index child	Mean	-1.808	-1.848	
	Std.Err.	0.050	0.088	
Raw PPVT score of index child	Mean	24.706	31.102	***
	Std.Err.	0.821	0.475	
PPVT score of index child, Official Standarization	Mean	83.049	86.540	**
	Std.Err.	19.98	33.772	
Nr Observations		450	1514	

Table 2.3: Descriptive Statistics: Index versus Younger Siblings

		Index Children, 2006	Younger Siblings, 2009	Diff.
Height-for-age	Mean	-1.808	-1.613	***
	Std.Err.	.050	.050	
Standarized PPVT score*	Mean	-.022	-.0026	
	Std.Err.	.046	.048	
Age (in years)	Mean	4.764	4.907	***
	Std.Err.	.023	.038	
% of male	Mean	.499	.434	*
	Std.Err.	.024	.023	
Nr Observations		450	450	

*Raw PPVT scores standardized to have mean/var 0/1 across age-groups

Table 2.4: IV Correlates and Fertility Choices, Linear Probability Model

	IV Correlates: HH Characteristics			Fertility Choices	
	Any Bad Event, 2002	Illness Shock, 2007-2009	Frost Event 2007-2009	Any Additional Siblings, 2006-09	Number Additional Siblings2006-09
	LPM	LPM	LPM	OLS	LPM
	(1)	(2)	(3)	(4)	(5)
Illness Shock, 2007-2009				-0.0267	-0.0650
				(0.041)	(0.063)
Frost Event, 2007-2009				0.0076	0.0441
				(0.055)	(0.097)
Non-Food Consumption (per capita), in logs	-0.0115	0.0078	-0.0172	-0.0267**	-0.0189
	(0.018)	(0.015)	(0.015)	(0.011)	(0.015)
Food Consumption (per capita), in logs	0.0540	0.0046	0.0196	-0.0990***	-0.2001***
	(0.046)	(0.040)	(0.043)	(0.037)	(0.062)
Maternal Education Level	-0.0028	-0.0027	-0.0092	-0.0010	0.0057
	(0.009)	(0.005)	(0.006)	(0.005)	(0.008)
Age Caregiver	-0.0087**	0.0003	-0.0033	0.0072**	0.0109**
	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)
Consumer Durables Index, 0-1	0.0046	-0.1034	0.1575*	0.0565	0.0554
	(0.137)	(0.090)	(0.094)	(0.106)	(0.147)
Services Index, 0-1	-0.0200	0.0500	-0.0361	0.0308	0.0869
	(0.123)	(0.080)	(0.056)	(0.066)	(0.093)
Housing Quality, 0-1	0.1867	0.0795	0.0077	0.0768	0.2923
	(0.133)	(0.091)	(0.067)	(0.103)	(0.184)
Spanish Speaker	0.0315	-0.1000**	0.0583	0.0652	0.0679
	(0.074)	(0.048)	(0.051)	(0.044)	(0.076)
Cluster fixed effects	Yes	Yes	Yes	Yes	Yes
Price Inflation Shocks	No	No	No	Yes	Yes
F-Test: Inflation Shocks	-	-	-	43.84	60.24
R-Squared	0.0242	0.0122	0.0175	0.0804	0.0737
Nr Observations	449	449	449	445	445

Notes: All models use a Linear Probability Model, except for column (4). Robust standard errors, clustered at the *index age-cluster* level. *, **, *** denote significance at 10%, 5% and 1% levels.

Table 2.5: Pooled OLS and Sibling-Difference OLS Models. Dependent variable: PPVT age-standardized score

	Pooled OLS	Within-siblings OLS		
	(1)	(2)	(3)	(4)
Height-for-age	0.099			
	(0.029)***			
Siblings-difference height-for-age		0.090	0.083	0.081
		(0.039)**	(0.038)**	(0.044)*
Obs.	898	450	450	330
R^2	0.423	0.076	0.138	0.142
Siblings-difference	No	Yes	Yes	Yes
Child-level controls	Yes	Yes	Yes	Yes
Cluster fixed effects	Yes	No	Yes	Yes
Household level controls	Yes	No	No	No
Age-group	All	All	All	4-5y

Notes: robust standard errors, clustered at the household level in Column (1) and at the *index age-cluster* level in columns (2) to (4); *, **, *** denote significance at 10%, 5% and 1% levels.

Table 2.6: IV GMM Sibling-Difference Model. Dependent variable: PPVT age-standardized score

	Full sample				
	OLS	Instruments: Changes in food prices (a)	Instruments: Household shocks (b)	Instruments (a) + (b)	With HH Controls IV: (a) + (b)
	(1)	(2)	(3)	(4)	(5)
Height-for-age	0.083	0.172	0.239	0.207	0.223
	(0.038)**	(0.079)**	(0.318)	(0.059)***	(0.066)***
Weak identification test:					
Kleibergen-Paap Wald rk F stat	-	14.53	8.38	19.88	19.30
Stock-Yogo weak ID test critical values:					
5% maximal IV relative bias	-	21.38	13.91	21.41	21.41
10% maximal IV relative bias	-	11.45	9.08	11.41	11.41
Overidentification test:					
Hansen J statistic	-	22.348	0.280	23.622	23.490
p-value	-	0.2673	0.8694	0.3674	0.3745
Obs.	450	450	450	450	449
R^2	0.138	-0.001	-0.017	-0.008	-0.0147
Nr. Excluded Instruments	-	20	3	23	23
Siblings-difference	Yes	Yes	Yes	Yes	Yes
Child-level controls	Yes	Yes	Yes	Yes	Yes
Cluster fixed effects	Yes	Yes	Yes	Yes	Yes
Time-Invariant Household controls	No	No	No	No	Yes
Age-group	All	All	All	All	All

Notes: robust standard errors, clustered at the *index age-region* level; *, **, *** denote significance at 10%, 5% and 1% levels.

Table 2.7: First stage results: Height-for-age

	Model 1		Model 2		Model 3	
	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.
Cereal price: 6-11 mths	0.983	0.739			0.936	0.731
Cereal price: 12-17 mths	0.111	0.785			0.128	0.747
Cereal price: 18-23 mths	-0.231	0.944			-0.182	0.905
Cereal price: 24-29 mths	0.257	0.879			0.170	0.868
Cereal price: 30-36 mths	2.666***	0.808			2.746***	0.796
Meat price: 6-11 mths	2.978**	1.381			2.805**	1.388
Meat price: 12-17 mths	2.751**	1.246			2.877**	1.210
Meat price: 18-23 mths	-1.620	1.135			-1.909	1.231
Meat price: 24-29 mths	1.259	1.839			0.897	1.849
Meat price: 30-36 mths	-4.138***	1.545			-4.100***	1.475
Dairy price: 6-11 mths	0.722	1.371			1.290	1.347
Dairy price: 12-17 mths	-1.638	1.393			-1.319	1.339
Dairy price: 18-23 mths	-1.030	1.573			-1.041	1.604
Dairy price: 24-29 mths	2.316	1.552			2.839*	1.474
Dairy price: 30-36 mths	-1.114	1.265			-1.288	1.285
Tubers price: 6-11 mths	0.791*	0.452			0.743*	0.447
Tubers price: 12-17 mths	-0.323	0.428			-0.319	0.441
Tubers price: 18-23 mths	0.032	0.446			0.038	0.442
Tubers price: 24-29 mths	0.008	0.495			0.019	0.516
Tubers price: 30-35 mths	-0.693	0.536			-0.686	0.544
Adverse shocks:2000-02			-0.181**	0.078	-0.181**	0.079
Frosts 2007-09			0.162	0.139	0.171	0.129
Illness (others) 2007-09			0.234***	0.091	0.247**	0.099
Constant	-0.118	0.319	-0.230	0.250	0.145	0.349
R-squared	0.148		0.100		0.165	
N	450		450		450	

Notes: within-household fixed effects estimates. Robust standard errors, clustered at the *index age-region* level; *, **, *** denote significance at 10%, 5% and 1% levels. Food prices groups as used by INEA: (1) Cereals: Bread and Cereals (comprising subcategories wheat, rice, maize, pasta as well as bread and biscuits); (2) Meat (comprising chicken, red meat and meat products, and processed meat); (3) Dairy: Milk, Eggs and Cheese. (4) Tubers: comprising potatoes, cassava, yucca, and andean tubers. Other controls included but not reported: age, sex, birth order, non-food expenditure (2006-2002), change in wealth index (2009-2006), community fixed-effects.

Table 2.8: Joint significance of Food Price, by Rural/Urban, Farming/Non-Farming HH.

	Total	By location		By activity	
		Urban	Rural	Farming	Non-Farming
	(1)	(2)	(3)	(4)	(5)
By item (all 5 semesters)					
Bread and Cereals	19.570	22.050	9.990	31.590	15.450
	(0.00)	(0.00)	(0.08)	(0.00)	(0.01)
Meat	30.440	21.160	16.670	18.450	9.450
	(0.00)	(0.00)	(0.01)	(0.00)	(0.09)
Diary	8.640	7.000	11.030	17.460	6.590
	(0.12)	(0.22)	(0.05)	(0.00)	(0.25)
Tubers	4.910	12.570	2.530	4.010	8.670
	(0.43)	(0.03)	(0.77)	(0.55)	(0.12)
By semester (all 4 groups)					
6-11 mth	17.970	62.890	5.220	14.020	12.010
	(0.00)	(0.00)	(0.27)	(0.01)	(0.02)
12-17 mth	6.480	2.820	14.030	3.710	4.620
	(0.17)	(0.59)	(0.01)	(0.45)	(0.33)
18-23 mth	3.290	9.360	4.960	11.400	6.140
	(0.51)	(0.05)	(0.29)	(0.02)	(0.19)
24-29 mth	7.340	5.860	4.420	13.430	0.390
	(0.12)	(0.21)	(0.35)	(0.01)	(0.98)
30-35 mth	22.200	17.150	12.840	10.440	21.900
	(0.00)	(0.00)	(0.01)	(0.03)	(0.00)
Total (all groups/semesters)					
All	321.650	208.500	283.120	172.500	137.090
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: Food prices groups as above in table 2.7. F-tests of joint significance conducted after estimation of model (3) and the interaction models. p-values reported in brackets.

Table 2.9: Exploring Price Channels: Food Consumption and Price Inflation, 2006-2002

	Base Model	Interacted Model: Rural	Interacted Model: Farming
	(1)	(2)	(3)
Δ Cereals price, 2006-2002	-0.0053*	-0.0031	-0.0066*
	(0.003)	(0.004)	(0.003)
Δ Meat price, 2006-2002	-0.0009	0.0003	0.0012
	(0.006)	(0.006)	(0.005)
Δ Dairy price, 2006-2002	-0.0013	-0.0051	-0.0095
	(0.004)	(0.005)	(0.006)
Δ Tubers price, 2006-2002	-0.0025	-0.0043**	-0.0022
	(0.002)	(0.002)	(0.002)
Δ Cereal price x Rural		-0.0100**	
		(0.004)	
Δ Meat price x Rural		-0.0174	
		(0.011)	
Δ Dairy price x Rural		0.0282***	
		(0.008)	
Δ Tuber price x Rural		0.0074***	
		(0.002)	
Δ Cereal price x Farming			-0.0018
			(0.006)
Δ Meat price x Farming			-0.0066
			(0.013)
Δ Dairy price x Farming			0.0159
			(0.014)
Δ Tuber price x Farming			0.0005
			(0.003)
R-Squared	0.0100	0.0293	0.0157
Nr Observations	449	449	449

Notes: Variables are the change between 2006-2002 for all categories. Rural is a dummy for living in rural area. Agriculture dummy for main occupation of household head. Significance levels as above.

Table 2.10: Interacted Models, by Rural-Urban and Farming-Non-Farming Households

	Interacted: Rural/Urban	Interacted: Farming/Non-Farming
	(1)	(2)
Height-for-age	0.2285***	0.2238***
	(0.003)	(0.010)
Weak identification test:		
Kleibergen-Paap Wald rk F stat	1976.18	3484.01
Stock-Yogo weak ID test critical values:		
5% maximal IV relative bias	21.37	21.37
10% maximal IV relative bias	11.21	11.21
5% maximal IV size	110.04	110.04
10% maximal IV size	56.80	56.80
Overidentification test:		
Hansen J statistic	36.126	41.762
p-value	0.6017	0.3517
Obs.	450	450
R^2	-0.0138	-0.0125
Nr. Excluded Instruments	40	40

Notes: robust standard errors, clustered at the *index age-region* level; *, **, *** denote significance at 10%, 5% and 1% levels.

Table 2.11: Predicted HAZ Impact of Price Changes

	Predicted Model			F-Tests of Joint Significance		
	\hatHAZ	Min	Max	F-Test	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)
Urban x Cereals	0.0843	0.060447	0.108143	10.88	0.0539	*
Urban x Meat	-0.1169	-0.13879	-0.09493	16.08	0.0066	***
Urban x Diary	0.0076	-0.00306	0.018212	9.54	0.0893	*
Urban x Tubers	-0.0625	-0.0736	-0.05134	2.55	0.7691	
Rural x Cereals	-0.0121	-0.04281	0.018515	22.99	0.0003	***
Rural x Meat	-0.3558	-0.40685	-0.30481	20.75	0.0009	***
Rural x Diary	0.0518	0.034777	0.068839	6.68	0.2455	
Rural x Tubers	0.0408	0.027468	0.054216	13.05	0.0229	**

Table 2.12: Robustness Checks: Controlling for Changes in Assets and Consumption - IV GMM Regressions

	IV GMM Core Regressions		Robustness Checks: + Δ Assets, 2006-2009		Robustness Checks: + Δ Non-Food Consumption, 2006-2009		Robustness Checks: + Δ Assets, 2002-2006	
	IV: Prices	IV: Prices & Shocks	IV: Prices	IV: Prices & Shocks	IV: Prices	IV: Prices & Shocks	IV: Prices	IV: Prices & Shocks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Height-for-age	0.1717**	0.2071***	0.1704**	0.2083***	0.1891**	0.2233***	0.1686**	0.2129***
	(0.079)	(0.059)	(0.079)	(0.057)	(0.077)	(0.056)	(0.083)	(0.059)
Δ Assets, 2006-09			0.4390	0.3866	0.4701*	0.4219	0.3048	0.3096
			(0.268)	(0.260)	(0.270)	(0.261)	(0.302)	(0.293)
Δ Non-Food Cons., 2006-09					-0.0339	-0.0242	-0.0322	-0.0225
					(0.021)	(0.016)	(0.020)	(0.016)
Δ Assets, 2002-06							-0.2225	-0.1550
							(0.199)	(0.180)
Weak identification test:								
Kleibergen-Paap F-Test	14.53	19.88	14.66	19.03	15.03	18.57	10.68	13.08
Overidentification test:								
Hansen J statistic	22.35	23.62	21.99	23.13	22.46	23.46	21.68	22.91
p-value	0.27	0.37	0.28	0.39	0.26	0.38	0.30	0.41
Obs.	450	450	450	450	450	450	450	450
R^2	-0.001	-0.008	0.002	-0.006	-0.001	-0.009	0.001	-0.007
Siblings-difference	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Changes in HH Assets	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Changes in HH Consumption	No	No	No	No	Yes	Yes	Yes	Yes
Age-group	All	All	All	All	All	All	All	All

Notes: robust standard errors, clustered at the *index age-region* level; *, **, *** denote significance at 10%, 5% and 1% levels.

Table 2.13: Robustness Checks: Time Allocations, Sibling Spillovers and Fertility Choices

	Robustness Checks: Core Model plus Additional Controls									
	+ Δ Mother Activity, 2006-2009		+ Δ Hours Work, by All Children, 2006-2009		+ PPVT Score, Index Child, 2009		+ Nr Younger Additional Siblings		+ All Controls	
	IV: Prices	IV: Prices & Shocks	IV: Prices	IV: Prices & Shocks	IV: Prices	IV: Prices & Shocks	IV: Prices	IV: Prices & Shocks	IV: Prices	IV: Prices & Shocks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Height-for-age	0.159**	0.170***	0.211**	0.269***	0.175**	0.187***	0.146*	0.197***	0.207**	0.251***
	(0.08)	(0.06)	(0.09)	(0.06)	(0.08)	(0.06)	(0.08)	(0.06)	(0.09)	(0.07)
Δ Mother Activity	0.061*	0.049								
	(0.04)	(0.03)								
Δ Hours Working			-0.292***	-0.287***					-0.308***	-0.303***
			(0.06)	(0.06)					(0.06)	(0.05)
PPVT Score 2009, Index					-0.140***	-0.146***			-0.138***	-0.132***
					(0.04)	(0.04)			(0.04)	(0.03)
One Younger Add.Sib.							0.199**	0.215**	0.225**	0.231***
							(0.09)	(0.08)	(0.09)	(0.08)
Two Younger Add.Sib.							-0.059	-0.145	-0.027	-0.019
							(0.14)	(0.09)	(0.12)	(0.09)
Three or More Add.Sib.							-0.531***	-0.493***	-0.431***	-0.444***
							(0.15)	(0.14)	(0.17)	(0.15)
Weak identification test:										
Kleibergen-Paap F-Test	8.5032	10.9888	13.3206	22.2707	13.2115	16.4234	16.3942	20.5724	13.7615	19.0326
Overidentification test:										
Hansen J statistic	21.3089	21.9045	21.1110	21.7886	20.9927	21.9818	21.3831	22.9777	18.2672	18.8038
p-value	0.3200	0.4656	0.3307	0.4726	0.3372	0.4610	0.3160	0.4030	0.5047	0.6574
Obs.	436	436	450	450	447	447	450	450	447	447
R^2	0.0024	0.0007	0.0103	-0.0067	-0.0014	-0.0042	0.0068	-0.0017	0.0188	0.0076
Siblings-difference	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Core Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: robust standard errors, clustered at the *index age-region* level; *, **, *** denote significance at 10%, 5% and 1% levels.

Table 2.14: Robustness Checks: Controlling for School and Preschool Enrolment - IV GMM Regressions

	Core	Robustness Checks:				
	(1)	(2)	(3)	(4)	(5)	(6)
		coef/se	coef/se	coef/se	coef/se	coef/se
Height-for-age	0,1717**	0.1983***	0.3269***	0.3239***	0.2927***	0.1916***
	(0.079)	(0.064)	(0.061)	(0.061)	(0.063)	(0.064)
Index child attended pre-school		0.1791***		0.2501***	0.3600***	0.3837***
		(0.067)		(0.092)	(0.090)	(0.085)
Index child started preschool at 3yrs.			0.2644***			
			(0.086)			
Index child started preschool at 4yrs.			0.2106***	-0.0483	-0.0534	-0.0320
			(0.075)	(0.052)	(0.051)	(0.042)
Index child started preschool at 5yrs.			0.1520**	-0.1034	-0.0971	-0.0839
			(0.076)	(0.088)	(0.092)	(0.074)
Index was school enroled when interviewed					0.5906*	0.4240
					(0.352)	(0.303)
Siblings was school enroled when interviewed					-0.4403***	-0.4108***
					(0.082)	(0.085)
Age of school enrolment 6yrs.						-0.2629**
						(0.128)
Age of school enrolment 7yrs.						-0.1312
						(0.136)
Age of school enrolment 8yrs.						-0.0067
						(0.117)
R-Squared	-0.001	0.0016	-0.0488	-0.0480	0.0003	0.0407
Nr Observations	450	450	439	439	439	432
Nr Clusters	43	43	43	43	43	43
Hansen J (Overid) test	22.348	23.7786	24.9622	25.0061	24.7080	21.8820
p-value	0.2673	0.3589	0.2989	0.2968	0.3112	0.4670
Kleibergen-Paap F-Test	14.5300	21.8973	28.6230	28.2979	24.3540	19.9923

Notes: robust standard errors, clustered at the *index age-region* level; *, **, *** denote significance at 10%, 5% and 1% levels.

Chapter 3

Catching up while there is time - Evidence of nutritional catch-up and household wealth in Ethiopia

Chapter 3

Catching up while there is time – Evidence of nutritional catch-up and household wealth in Ethiopia

Abstract

We examine the nutritional status of a cohort of poor Ethiopian children and their patterns of catch-up growth in height-for-age between three key development stages: ages one, five and eight. During the earliest period, we find that nutritional catch-up patterns vary substantially across socioeconomic groups: average catch-up growth in height-for-age is almost perfect among children in relatively better-off households, while among the poorer children, relative height is more persistent. Between five and eight years of age, however, we find near-perfect persistence and no evidence of heterogeneity in catch-up growth. Our findings suggest that nutritional remediation can be effective early on in life, but that this window of opportunity might already be closed by the age of five. Calculations suggest that the height-for-age deficits accumulated by the bottom half of the distribution at the age of five might translate into 1.4% of loss annual income during their adult lifetimes.

3.1 Introduction

Childhood undernutrition is of the most pressing issues in current development policy (Grantham McGregor et al., 2007). Evidence from medical and paediatric research indicates that anthropometric and cognitive development of adults is largely determined during gestation and early childhood, and might be subject to 'critical periods'. In particular, nutritional inputs during pregnancy and the first two or three years of a child's life have been documented as crucial in determining later adult height, health and other significant outcomes (Doyle et al., 2009; Victora et al., 2010). Similarly, shocks during these critical developmental periods can have irreversible effects on boys' and girls' long-term cognitive abilities, anthropometric status and health. Models of human capital formation that incorporate these insights have become influential in the economics discipline (Cunha et. al, 2006; Cunha and Heckman, 2007). A feature of these models is that an efficiency-equity trade-off does not exist in the early years (i.e. investment is both efficient and equitable), prompting Alderman (2010) to argue persuasively that spending on early childhood nutrition should not be viewed as a redistributional tool but as investments in health, nutritional and cognitive development.

Related to the concept of critical periods is the notion that nutrition and other human capital dimensions might be subject to catch-up growth. The medical literature documents that higher than normal height velocity can be achieved following a period of retardation, such that previously lagging individuals might return to the statistically normal growth curve. Catch-up growth is characterized by an improvement in the percentile position in the distribution (Boersma and Wit, 1997). Evidence from animal studies shows that near-perfect catch-up is common for mildly undernourished subjects, but stunting might be permanent when nutritional deficits begin early and are prolonged. These studies also show that nutritional remediation is effective, and catch-up growth is achieved only while the critical period of growth remains open (see Wilson and Osbourn, 1960 for an early review). Evidence on human catch-up from nutritional and health literature shows some similar results: Martorell et al. (1994) survey evidence from medical literature and find evidence of catch-up growth when living conditions are improved, especially for younger children. Schroeder et al. (1995) find that nutritional supplementation has a significant impact on growth for under 3 year olds in Guatemala. Adair (1999) finds almost total catch-up for children aged 2-12 in the Philippines.

A similar notion is found in the historic anthropology literature. Socio-economic class and racial status have long been associated with differences in historic height (Eveleth and Tanner, 1976). But if the relationship between adult height and income is nonlinear at the individual level, then the relationship at the aggregate level depends on the income distribution (Steckel, 1995). Accordingly, aggregated inequality has been found to strongly correlate with lower average heights and that height is particularly sensitive to income at low income levels (Steckel, 1983, 1995).

A small literature in economics seeks to quantify catch-up growth in child nutrition (Deolalikar, 1996; Fedorov and Sahn, 2005; Alderman et al., 2006; Mani, 2008). While there is agreement that catch-up is highest in early childhood, estimates of catch-up growth differ widely by age and methods of analysis. In this paper, we contribute to this literature by analysing nutritional catch-up growth in a sample of Ethiopian children. We differ from previous studies in that the primary focus of our analysis is on assessing how household socioeconomic status can influence height velocity. Assuming households have homogenous preferences, children living in relatively rich, less credit-constrained households might be able to achieve higher rates of nutritional catch-up than their poorer counterparts. For this purpose, we use the Ethiopian Young Lives sample, a longitudinal dataset of relatively poor children, that provides three waves of anthropometric and socioeconomic data measured at key stages of development: at the ages of one, five and eight respectively. The study covers the critical stage of infancy up to mid-childhood, but stops before the onset of adolescence and its related complex patterns of teenage growth spurts. Exploiting this structure, we can test whether returns to nutritional investments vary between early childhood and mid-childhood (5-8 years).

A handful of studies have examined socioeconomic differences in nutritional status, mostly in a developed country context. To our knowledge, our paper is the first to investigate differential patterns of catch-up growth for different age groups, and across the socioeconomic distribution. Doyle et al. (2009) review evidence showing that low birth weight in developed countries has clear long-term consequences for human capital, and further find significant differences between high and low socioeconomic groups. Batty et al. (2009) also review several studies that find significant impact of various measures of socioeconomic status on height. Finch and Beck (2011) find significant socioeconomic gradients in height-for-age of a cross section of US children aged 2-6 years. Evidence for developing countries is quite scarce. Rona et al. (2003) find cross-sectional socioeconomic differences in height in Trinidad and Tobago, and Webb et al. (2008) provide evidence in several countries of Eastern Europe that adult height is associated with childhood socioeconomic circumstances. Ruel et al. (1995) expand on nutritional effects of the Guatemala study outlined above (Schroeder et al., 1995) and find that girls from poorer backgrounds benefited more from the supplementation than girls from wealthier families.

In our empirical analysis, we estimate a dynamic model of nutritional status and explore whether the gradient of catch-up growth – that is, the coefficient on lagged height-for-age z-scores – varies by socioeconomic status. By documenting the link between early childhood nutritional status and the socioeconomic gradient of catch-up, our results go beyond the standard findings of the literature. In section 4, we present non-parametric estimates of catch-up growth and find that, between the ages of one and five, undernourished children from better off households experience higher height velocity than their poorer but equally undernourished counterparts. At the same time, we find little evidence of differential catch-up growth across wealth groups for the same children in the period between five and eight years of age, suggesting that nutritional remediation after the age of five might be ineffective.

In section 5, we revisit these findings in a parametric context, including applying instrumental variables (IV) methods to address endogeneity concerns. We find ordinary least squares and IV methods broadly consistent with the non-parametric analysis. Higher height velocity among relatively rich children implies that catch-up growth during early childhood for these children is near total, and the differences between catch-up across quartiles of the wealth distribution are statistically significant. To summarise the contribution of the paper, we find non-linear patterns of heterogeneous catch-up growth, with height velocity highest amongst undernourished children in relatively wealthy households. This evidence is consistent with compensatory investments under imperfect credit markets, but we cannot rule out that differences are driven by household preferences. Our findings indicate that nutritional interventions for the poorest households, even in very poor settings, could have considerable impact on future human capital and that such interventions would only be effective if targeted at infant children.

The remainder of the paper proceeds as follows. In section 3.2 we discuss the theoretical background and methodology. In section 3.3 we describe our data and provide some descriptive statistics, while sections 3.4 and 3.5 present the core of the analysis. We conclude in section 3.6.

3.2 Methodology

Strauss and Thomas (2008) provide a comprehensive survey of recent developments in applied microeconomic analysis regarding health over the life course, and Almond and Currie (2010) summarise recent empirical evidence on the persistent impact of early life conditions on future outcomes including health, cognitive ability and earnings. A theoretical framework that echoes the nutrition literature and has a focus on "critical period programming" has become influential in economics (Cunha et al., 2006). The technology of skill (or human capital) formation determines complementarity or substitutability of investments in different time periods over time, which crucially underpins the possibility for catch-up growth and nutritional remediation. In the extreme case of perfect complementarity, investments in period two (or later) *cannot compensate* for the lack of investment in period one. In sum, early childhood investments must be distinguished from late childhood investments, and an equity-efficiency trade-off may exist for "late" investment, but not for "early" investment.

Related to critical periods is the notion that nutrition and other human capital dimensions might be subject to catch-up growth. The medical literature referenced in the introduction provides evidence that higher than normal height-for-age growth can be achieved following a period of retardation, leading to catch-up growth.[1] Similarly, Almond and Currie (2010) discuss the concept of remediation, or the extent to which a shock experienced in an early

[1] Boersama and Wit (1997) note that catch-up growth may also take the form of an extended period of growth – for example, extending the adolescent growth period.

developmental period can subsequently be mitigated. The authors offer evidence that shocks cause more long-term damage amongst poorer families, even when facing the same shock as wealthier families.

Whether or not remediation is possible depends on several factors: the total productivity of all combined investments through childhood; the timing and combination of shocks and subsequent investments; and the extent of substitutability/complementarity that may exist at different stages. For example, higher nutritional status in an early period may lead to increased absorption of nutrition in later periods. Almond and Currie (2010) note that it is not necessary to observe all investments and estimate the substitutability coefficient for health production; we may simply observe how a shock or nutritional intervention in the first period affects outcomes in later periods. The reduced form estimation of the health effect of such a shock will not only include the pure biological technological effect, but also the effect of the household's responses. Whether parental investments are compensatory or reinforcing will depend on the degree of inter-temporal substitution in the health production function, but also on the functional form that household preferences take. This remains a question open to empirical investigation.

Empirically, height-for-age z-score ($HAZ_{i\,t-1}$), for child (i) at time (t), can be modelled as a function of lagged height-for-age z-score at time (t-1), which proxies for previous nutritional investments, and a vector of child and household characteristics (X_{t-1}), which proxy for contemporaneous nutritional investments. Height-for-age might also be affected by unobserved child and household characteristics (μ_i), as well as community unobservables (μ_v).

(1) $$HAZ_{it} = X_{t-1}\beta + HAZ_{it-1}\gamma + \mu_v + \mu_i + v_{it}$$

Several earlier studies noted in the introduction have used variations of equation (1) to obtain estimates of catch-up growth.[2] When consistently estimated, equation (1) provides parameter estimates for the degree of persistence in height-for-age between period t and t-1. Under perfect catch-up the coefficient on lagged HAZ (γ) would be close to zero, while a coefficient close to unity would be consistent with perfect persistence.

However, while the coefficient on lagged HAZ can provide an estimate of average persistence, it is not clear that this model is sufficiently flexible to capture the complex patterns underlying catch-up growth. In particular, it assumes that nutritional investments do not respond to past nutritional status and imposes linearity in the lagged HAZ coefficient. There are two possible reasons why the relationship may be non-linear; first, behavioural responses – parents may compensate or reinforce nutritional deficits that they observe and; second, biological factors may lead to increased growth velocity for children at the lower end of the growth distribution.

The remediation model discussed above and in Almond and Currie (2010) would suggest that household nutritional investments will respond to early realizations of nutritional status, resulting in either compensatory or reinforcing actions of parents. Which occurs will depend on technology, preferences and resources. We are unable to directly observe whether parents are able to make compensatory investments but we can hypothesise that for poorer families, this may be more difficult. Behrman et al. (1982) show in a model with more than one child, that even if parents would prefer to accumulate the human capital of

---

[2] Several previous studies have specified this model as a growth equation instead, namely $(HAZ_{it} - HAZ_{it-1}) = HAZ_{it-1}\gamma^{*} + \text{cov}\,ariates$. If the relationship between current and lagged height is linear, then the two equations are equivalent, and in comparison with equation (1), $\gamma^{*} = \gamma - 1$. See Federov and Sahn (2005) for further discussion of this point.

the less well-off sibling, it still may not happen if inter-temporal substitution is difficult due to imperfect credit markets.

On the point of biological factors, the paediatric and medical literature suggest that higher velocity is likely to happen among children in the most vulnerable positions, and only in so far as nutritional *remediation* takes place. This would suggest that catch-up growth should show patterns of heterogeneity across different levels of nutritional investment on the one hand, and non-linearity across the distribution of height on the other. Failing to take these features into account is likely to lead to biases in catch-up estimates, and in particular imply that the functional form linking HAZ and lagged HAZ in equation (1) might be mis-specified.

In this paper, we use non-parametric and parametric methods to address the shortcomings of earlier studies. First, we present kernel smoothing estimates of height-for-age z-scores on lagged height-for-age z-scores across wealth quartiles. This analysis is sufficiently flexible to allow for non-linearities in the HAZ relationship and across the socioeconomic distribution, and yields the core of our empirical results. Secondly, we apply an array of parametric methods to corroborate the findings of the graphical analysis. We modify equation (1) to test whether catch-up growth differs across households from different parts of the wealth distribution by interacting the lagged HAZ coefficient with household wealth.

(2) $$HAZ_{it} = X_{t-1}\beta + \sum_{j=1}^{4} w_{ij} * HAZ_{it-1}\gamma + \mu_v + \mu_i + v_{it}$$

(2') $$Stunting_{it} = X_{t-1}\beta + \sum_{j=1}^{4} w_{ij} * Stunting_{it-1}\gamma + \mu_v + \mu_i + v_{it}$$

Where *j=1,2,3,4* represents four quartiles of the distribution of a composite wealth index. We also test for non-linearities in this relationship by examining the persistence of stunting (a binary variable) as expressed in equation (2').

Our panel dataset contains measurements for each child in three time periods: as a new born (aged 0-1), in early childhood (aged 4-5) and in mid-childhood (aged 7-8). This allows us to test whether the relationship in equation (2) is stable over time. The findings of previous studies suggest that from the age of 3 to 5, catch-up growth is substantially lower. We expect a similar pattern regarding the effectiveness of nutritional investments.

Empirical estimates of equation (1) and (2) and, specifically, the relation between nutritional status in two periods, are likely to suffer from endogeneity concerns (Hoddinott and Kinsey, 2001; Fedorov and Sahn, 2005; Strauss and Thomas, 2008). Unobserved parental investments, measurement error and genetic potential are among a number of factors that could lead to biased and inconsistent estimates. We saturate the model with a number of child and household characteristics to reduce potential biases from heterogeneity in the error term. In particular, the vector of controls, (X_{t-1}), includes information on pregnancy and child birth experiences, past illness of the child, child characteristics (age, birth order and gender), household demographics and household composition variables, as well as parental education. We also include maternal height as a proxy for the genetic potential of the child. Finally, we include a full set of cluster fixed effects to capture unobserved heterogeneity in the patterns of catch-up growth across the sample communities.[3]

We address any remaining endogeneity concerns by applying instrumental variable (IV) techniques. We estimate determinants of the lagged height-for-age z-scores (HAZ) using instruments (z_{it-1}) orthogonal to the error term in the main equation. We incorporate the socioeconomic gradient analysis into the IV technique by estimating the IV model for different quartiles of the distribution,

$$(3) \qquad HAZ_{it} = X_{t-1}\beta + HAZ_{it-1}\gamma + \mu_v + \mu_i + v_{it} \qquad \text{if } w_{ij} \le \tilde{w} \qquad \text{(Main-Equation)}$$

$$(4) \qquad HAZ_{it-1} = X_{t-1}\xi + Z_{it-1}\phi + u_{it-1} \qquad \text{if } w_{ij} \le \tilde{w} \qquad \text{(First-Stage)}$$

whereby $\tilde{w}$ equals 2 and 4. In other words, we obtain IV estimates for the full sample and the bottom half of the wealth distribution separately.

Our choice of instruments is motivated by the demographics and medical literature on the impact of seasonality on birth weight and long-term health. Medical studies document patterns of seasonality in foetal development and birthweight (Van Hanswijck de Jonge et al, 2003, McGrath et al, 2005, Chodick et al, 2007, Torche and Corvalan, 2010,), with seasonality varying across countries in different latitudes and level of development (Chodick et al, 2009, Strand et al, 2011). Studies with neonatals also show that seasonality in the availability of nutrients (Doblhammer and Vaupel (2001)) and risk of infection (Costa and Lahey (2005)) in early life, affect the long-term prospects of children. Ethiopian studies have also uncovered substantial seasonality. Dercon and Krishnan (2000) report

---

[3] In the results section, we also report evidence for equations (1) and (2) where the X_{t-1} vector is omitted and only include cluster dummies as controls. We report this model for comparison purposes. The evidence confirms our original suspicion that omission of covariates might bias estimates of catch-up and socioeconomic gradients.

seasonal variations in the nutritional intake of adults, while Ferro-Luzzi et al (2001) report strong correlations between diarrheal child morbidity and seasonal patterns of retroviral infections.

We use the quarter of birth as an instrument for height-for-age at the age of 6 to 18 months. Results for the first-stage regressions indicate that children born in the second quarter of the year and again towards the end of the year have significantly lower HAZ z-scores one year later. (See Table A1 in the Annex). We interpret this as indication of nutritional deficiencies suffered by the newborns during the critical period when babies are no longer being breast fed. Children born towards the end of the year will typically stop breast feeding during the main Ethiopian rainy season or *Kremt* (from June to September) – which coincides with the critical food insecure period among rural households, as well as the period of maximal exposure to diarrheal infections (Ferro-Luzzi et al.,2001). Similarly, children born during the second quarter might be exposed to the second shorter season, or Belg, which occurs in March and April.

The credibility of the IV technique relies on the 'validity' and 'strength' of our instruments. Formally, we require the following conditions:

(3) $$Cov(v_{it}, z_{it-1}) = 0$$ (Exclusionary Restriction)

(4) $$Cov(HAZ_{it-1}, z_{it-1}) \neq 0$$ (Strength)

Although the structure of the YL sample does not allow us to apply the Mother Fixed Effects IV estimator as used in Alderman et al. (2006) – we apply Cluster Fixed Effects IV methods instead –, we believe that our set of instruments are truly orthogonal to the error term in equation (2). For the exclusionary restriction to be met, condition (3), seasonality should not be correlated with any unobserved nutritional investments.

One potential violation of instrument validity might happen if households plan the season of birth of their children. This has been suggested in a sub-Saharan African context in Artadi, 2005). Households that do so might also differ in other ways, such that the timing of birth might be correlated with unobserved nutritional investments and parental preference. We explore the possibility of the endogeneity of the instrument by analysing the correlates of season of birth in Table A2 (see Annex). Column (2) estimates a probit model with season of birth dummy as outcome variable, and a vector of controls (X_{t-1}) and village fixed effects as controls. Reassuringly, we find little correlation between key socio-economic indicators – such as wealth, education or height of the mother -, and quarter of birth. However, some of the household demographic variables were indeed significant. In all our IV specifications, we include the full vector of controls (X_{t-1}), to ensure our IV estimates are not affected by these correlations.[4]

An additional concern in applying IV methods relates to how informative, or strong, the instruments are. Finite-sample theory suggests that IV estimation with weak instruments can lead to very misleading results. Weak IVs provide not only biased point estimates – towards OLS –, but standard errors are understood to be biased downwards, increasing the possibility of accepting significance of a substantially biased estimate (Stock and Yogo, 2002; Murray, 2006). We use the Stock-Yogo (2002) critical values to asses whether the IVs in the first-stage regressions are sufficiently strong. To ensure that any biases are

[4] Another possibility regarding violations of the exclusionary restriction is that quarter of birth as such determines not only nutritional investments early on in life, but subsequent investments through the lifetime of a child. If this was the case, exclusionary restriction (4) would be violated and our IV estimates would be biased, simply documenting the process of subsequent investments. We explore this possibility by estimating our main equation (3) but adding quarter of birth as one of the confounders. Because in equation (3) we control already for lagged HAZ, any effect of quarter of birth on directly would suggest that investments between (t) and (t-1) are indeed correlated with the season of birth. We perform this test for both periods: 2009 on 2006, and 2006 on 2002. Since our analysis focuses strongly on the wealth distribution, we perform the test for each half of the wealth distribution. The findings are conclusive: quarter of birth does not affected HAZ at age 4-5 or 7-8 other than through its effect on lagged HAZ. (Results not reported).

minimal, we report Fuller IV estimates. The Fuller estimator belongs to a family of weak-IV robust estimators, shown to perform better than standard IV methods – such as 2SLS or GMM methods (Murray, 2006). Though the season of birth is highly significant in equation (3), we find that our instruments are marginally weak, but the point estimates are robust.

Given the weak IV issues with the estimation, some uncertainty around the magnitude of the true estimate remains. Recent studies have proposed Anderson-Rubin and CLR Moreira confidence intervals, which have been shown to be *fully* robust to the presence of weak-IVs (Yogo, 2004 and Andrews, Moreira and Stock, 2006). These statistics provide a robust indication – regardless of the exact coefficient magnitude – of whether the effects studied are non-spurious. The tests support our suggestion that in spite of relatively weak instruments, the catch-up estimates are sufficiently robust.

Finally, note that in our parametric analysis we control the error term for potential correlation across members of the same community. In particular, shocks common at the level of the village as well as seasonal shocks might affect standard inference methods. Throughout the paper we report cluster corrected standard errors at the level of the community and the quarter of birth.

3.3 Data and descriptive statistics

The data are obtained from the first three rounds of the Ethiopia Young Lives survey, an ongoing longitudinal study of child poverty. The baseline year was 2002, and the study is planned to continue over a period of fifteen years.[5] Ethical concerns informed all stages of

[5] See www.younglives.org.uk for an overview of the Young Lives project, which also operates in India, Peru, and Vietnam. See Outes-Leon and Sanchez (2008) for an assessment of the sample used.

the data collection process.[6] 100 children aged 6-18 months were randomly sampled in each of the 20 purposely chosen sites, yielding a cohort of just under 2,000 children (1,999 to be precise), which have been followed in subsequent follow up rounds in 2006 and 2009.[7]

Ethiopia is a country that is known for pervasive malnutrition and persistent hunger (Alderman and Christiensen, 2001). Major famines in 1974, 1984 and in the past ten years[8] make the study of nutrition and child development a pertinent issue. The Ethiopian economy has experienced growth in the past two decades, but seasonal hunger continues to be an endemic feature of life in many rural areas.

We calculate the height-for-age z-score for children in the sample (HAZ). HAZ score is a measure of child development that has been shown to correlate with long-run investments in child nutrition (i.e. the 'stock' of health). It shows the height of the child relative to an international reference group of healthy children. We use the latest version of the height distributions, known as the WHO Reference 2007.[9] The international standard allows our study to be compared with other studies, and also allows an analysis of children of different ages, compared to the norm for their exact age in months. Stunting is defined as a HAZ of -2 or less, and severe stunting as HAZ below -3.

We focus our analysis on children located in the rural sentinel sites (just over half the children). Around 85% of the population of Ethiopia live in rural areas where

[6] For a detailed discussion of the research ethics, methods and training of the research team, including issues arising over the course of the longitudinal research, and ongoing informed consent, see Morrow (2009).

[7] Wilson and Huttley (2004) present a justification of this sampling procedure. The five regions selected (Addis Ababa, Amhara, Oromia, SNNP and Tigray) account for 96% of the population of the country. Note though, that this coverage excludes pastoralist communities of Afar and Somaliland.

[8] See the many useful references in Harvey (2009)

[9] Raw data at http://www.who.int/growthref/ last accessed Feb 23rd 2010. Also see de Onis et al. (2004, 2007) and Garza and de Onis (1999) for a discussion of the latest standards for child growth.

undernutrition is more prominent than in urban areas. Further, our IV strategy outlined in section 3, above, uses seasonal variation in nutrition, which is more pertinent for the rural sample.

Attrition in the Young Lives sample is low in international comparison with other longitudinal studies (Outes and Dercon, 2008). In the second round, attrition was 4.5% – 61 children had died, and 29 were untraceable or did not wish to participate – leading to a loss of just 48 children from the rural sample. Outes and Dercon (2008) study attrition bias in round two and conclude that while sample attrition – mainly mortality – in Ethiopia was linked to poor health and nutrition, the magnitude of these effects would not lead to estimation biases in a health equation similar to the one estimated here. Attrition in round three was even lower, at just over one per cent, or thirteen children for the rural sample, which suggests that attrition bias is of little concern in our analysis.[10]

Table 1a presents the mean and standard deviation for the variables in our analysis. Common to other studies, we find that the anthropometric status of children – as measured by height-for-age – deteriorates between the time of birth and 3 years of age (e.g. de Onis et al., 2007). At the age of 6-18 months in round one (2002), the average YL child had a HAZ z-score of -1.13.[11] Four years later the average HAZ z-score was -1.6. By 2009, it recovered slightly to -1.38. 34 per cent of the children were stunted in 2002, falling to 24 per cent in 2009.

We measure wealth by using an index constructed by the Young Lives team. The Wealth Index is a composite variable of housing quality, local services and durable assets

---

[10] We perform similar attrition bias tests as in Outes and Dercon (2008) for round 3 and find no evidence of non-random attrition. Results are not reported here.

[11] Mekonnen et al. (2005) provide a fuller descriptive analysis of the nutritional status of the cross-section of this group of children in 2002, and find strong correlations between wealth and stunting, amongst other things.

standardised across the YL sample that arguably captures the socio-economic status of the households.[12] Housing Quality comprises an average of four dummies equal to one if the walls, roof, floor are of good durability, or if the house has more than one room; Consumer Durables include 11 items such as radio, fridge, bicycle, table. The Services Index comprises averaged dummies for if the household has access to electricity, (clean, piped) water, sanitation (pit latrine or flush toilet) and cooking fuel (not wood or dung). Each sub-index is scored, and then the three are averaged to calculate the final index. The range of the wealth index is then from 0 to 1, and we divide households into quartiles of the wealth distribution across the sample.

We provide a disaggregation of the wealth index into the three components of services, consumer durables and housing in table 1b. We show for each quartile the percentage of households with no items, one or more items, and two or more items in each sub index. It is immediately apparent that this is an extremely poor sample. In the bottom quartile (column 1), households generally speaking have none of the items in any sub-index. In terms of housing, the number of improved housing materials increases steadily across quartiles, though no household in the sample has a floor made of anything other than earth. In the top quartile, more than half have two or more items such as improved roof, walls; or more than one room. There is not a great deal of variation across the sample in terms of consumer durables owned, as most households are poor and own very few. In the top quartile, just under a third own one or two items. In terms of services, the bottom two quartiles have none. In the third quartile, 17 per cent of households have one item (predominantly a pit latrine), and nobody has more than one service. In the top quartile, 42 per cent have one service, and 16 per cent have two services (mainly pit latrine plus electricity).

[12] See Ethiopia Preliminary country report round one on www.younglives.org.uk for further details.

Figure A1 (annex) depicts the kernel density function of the wealth index variable. While this is a substantially skewed distribution, we find that a quartile analysis captures the variation in wealth well. This is not the case for the sub-indices, which in some cases only take a limited number of values, consequently in later analysis, where we decompose the effect of the wealth index into its sub-indices, we use the categorical variables reported in Table 1b instead.

3.4 Catch-up growth and household wealth

In Figure 1 we graphically depict the relation between height-for-age z-scores at age 6-18 months and aged 4-5 years. The graph shows bivariate lowess kernel smoothing estimates with round 1 HAZ in the x-axis and round 2 HAZ in the y-axis. For most of the values of HAZ at the age of 0-1, the relation is linear. Moreover, the slope of the relation suggests a substantial degree of catch-up (as the slope is flatter than the 45-degree line shown on the graph). There is also a marked non-linear effect at the lower end-tail of the HAZ distribution. Early measures of HAZ below -2 (stunted) are linked to higher levels of catch-up growth – in other words, children with particularly low levels of HAZ appear to experience faster catch-up growth than other children. This is consistent with the definition of catch-up in Boersma and Wit (1997), discussed above. Figure 2 incorporates the relation between height at 4-5 years and 7-8 years. The relationship is much closer to the 45 degree line, indicating that children's position in the height distribution appears to be more 'fixed' and that – unlike in the earlier period – height at 4-5 years is a strong predictor of height three years later.[13]

Figure 3 further explores whether catch-up growth differs across wealth categories. Again we present lowess kernel smoothing estimates and now split the original sample between children living in households in the bottom, top and middle two quartiles of the round 1 wealth index (see above for description, panel A depicts the early childhood relationship and panel B the later period). Panel A is striking in two respects. First, the locus of round 1 HAZ and round 2 HAZ is vertically shifted when moving from lower to higher wealth index quartiles. This suggests that children from poorer households will accumulate a growth deficit vis-à-vis wealthier children *with a similar initial level of nutrition*.

Secondly, the original non-linear relationship between early and later HAZ appears to be present only among the relatively wealthy households. While children in households from the bottom quartile of the wealth index experience catch-up rates that remain constant throughout the entire nutritional distribution, better-off children experience increased catch-up growth if their HAZ measure in round 1 was low.

The graph in Figure 3(A) provides compelling evidence that – to the extent that catch-up growth is taking place between the ages of 0-1 and 4-5, it is children in relatively well-off families who are benefiting most. The comparison with the later period (between age 4-5 and 7-8) is also striking. Figure 3(B) shows no clear differences between any of the wealth categories, and again the slopes of the lines converge at a much steeper gradient. In the early years, household wealth appears to enable nutritional catch up, however, by the age of

---

[13] Figure A1 in the annex shows the kernel density estimates of the three rounds of data. Clearly the early years (round 1) are more spread out, which we would expect at this age group, but there is a striking similarity between rounds 2 and 3.

five the window of opportunity for effective nutritional remediation appears to be closed, though it may open again in the adolescent years.

The patterns presented in Panel A of Figure 3 identify two distinct definitions of catch-up, not always separated in the literature. The first being “mean catch-up” rate as identified by the slope of the curves presented. This is the traditional definition used in the economic literature which focuses on the coefficient of lagged HAZ in equation (3): the closer to unity – or the steeper the curve - the lower the rate of catch-up. However, Figure 3 also suggested a second type of catch-up in line with the medical literature, what we call “distributional catch-up”. The notion that catch-up will be expected among children with a height deficit, but only when remediation is applied. In Figure 3, this is represented by the heterogenous patterns of catch-up rates – across wealth groups – present only in the lower part of the HAZ distribution. We believe that while related, “mean” and “distributional” catch-up are distinct phenomena.

Using non-parametric analysis, or OLS for that matter, cannot establish whether catch-up is genuine or driven by endogeneity bias. Measurement error, unobserved child ability and household heterogeneity are well understood causes of statistical bias that will affect catch-up growth. Measurement error and mean reversal generate a negative correlation between HAZ and its lag, creating the impression of higher rates of mobility. Selective mortality creates a similar phenomenon of reduced persistence: if surviving children have better innate health, possibly due to a better genetic pool, rates of recovery will be higher than if we could observe all children. All these factors bias catch-up estimates upwards, or coefficient on lagged HAZ biased towards zero in equation (3). In the next section we present parametric analysis using IV methods that seek to address these issues.

Accordingly, OLS estimates and non-parametric analysis should be interpreted with caution. Indeed, high “mean” catch-up rates suggested by the flat slopes presented in Figure 3 for young children are particularly suspect: mortality rates among infants are non-negligible, and measuring height among infants is notoriously prone to error, as such we expect substantial biases in Panel A. The steeper curves for the older period, see Panel B, suggests both the closing of the critical period, but also the fact that endogeneity bias might not be as high.

However, while endogeneity concerns affect both, “mean” and “distributional” catch-up, we believe that the latter might be less sensitive to these concerns, at least in their conventional formulation. In what follows, we explore the extent to which “distributional” catch-up found in Panel A of Figure 3, could be generated by endogeneity biases.

Measuring the height of small children can be imprecise, especially when babies are too small to stand alone. Our sample includes children aged 6 to 17 months, so we might expect the noise-to-signal ratio to be larger among the youngest in the cohort. We do find evidence that HAZ in the first wave has a higher variance among younger children (aged 6 to 11 months). However, those younger children have on average higher z-scores than their slightly older counterparts (aged 12 to 17 months). In other words, younger children are not concentrated in the lower section of the HAZ distribution. Nevertheless, even if they were, it is not clear why attenuation bias should exclusively affect richer children – as is implied by the convexity of the top and middle quartiles, but not the bottom wealth quartile in Figure 3.

Panel attrition could potentially be generating heterogenous catch-up. If selective mortality is concentrated among children with poor health and nutrition, surviving children in the lower section of the HAZ distribution are likely to have higher innate health. Low HAZ has indeed been found to be correlated with attrition in the second wave, however, the small incidence of attrition implies that attrition bias can be expected to be minimal (Outes and Dercon, 2008). However, this phenomenon cannot explain why catch-up is largest among the richer children. If anything, we would expect selective mortality bias to be highest among the poorest households – those where undernourishment is most likely to lead to mortality.

Finally, children in richer families might differ in their innate health; wealthier families over generations might have cumulated a better genetic pool (for a discussion see Deaton, 2007). Accordingly, wealthier families with possibly taller mothers might have children with higher height potential (Bhalotra and Rawlings, 2011). In this case, children from richer backgrounds who are temporarily stunted in the early period, may simply be reaching their higher underlying potential height, rather than being more heavily invested in. In our subsequent parametric analysis we include information on maternal height and maternal education to control for this mechanism.

3.5 Parametric Analysis

In this section we report an array of econometric methods to substantiate the findings of the non-parametric analysis. Table 2 shows community fixed-effects (OLS) estimates of HAZ on lagged HAZ for early childhood and mid-childhood separately. Columns (1) and (3) are the naïve specifications with only community fixed effects as controls, whilst columns (2) and (4) include full controls; household characteristics that are likely to influence future nutritional status (household composition, assets, mother's height, literacy of the mother). The top panel of the table assumes homogeneity in the lagged HAZ coefficient. The first two columns in this panel show a strong correlation between early nutritional status at the age of 0-1 and later height attainment at the age of 4-5 years of age, with a significant and positive coefficient on the early child height. The point estimate for the lagged dependent variable of 0.23 during this period shows substantial, but only partial, catch up. If persistence in height-for-age were perfect, we would expect a coefficient that is close to one. Columns (3) and (4) use the same specifications, but move on one period; the relation between height-for-age at the age of 4-5 and 7-8 years is assessed. The coefficients are now much higher, closer to 0.7, replicating the high degree of correlation that was shown in panel B of Figure 3.

In terms of the full set of control variables (reported in Annex table A1), we found that the wealth index was significant in determining HAZ in the early period, but not later. Maternal height (our control for genetic factors), literacy of the mother, the number of household members and the health of the child during early years were also significant- though less so for the later period. Beyond the reported controls the model specification also includes information on a range of household, mother and child characteristics, as well as a set of community, household ethnicity and month of birth dummies. In a robustness check we trimmed the top and bottom 5% HAZ scores, removing potential outliers, and results remained largely unchanged.

In the lower panel of table 2, we explore the interaction of wealth and lagged height to examine whether the relationship between early child height and subsequent height differs significantly depending on the wealth of the household. Columns (1) and (2) report results for the early period, and columns (3) and (4) for the later period. We include individual wealth quartile dummies and interact the HAZ in the previous period with each of these quartiles, omitting the poorest quartile. Taking the early period first, the interaction terms are highly significant.[14] We find that persistence of nutritional status is significantly lower for the top quartile than for the bottom quartile, while point estimates indicate that the magnitude of the catch-up growth increases with wealth. This evidence suggests that catch-up growth is stronger for children from relatively wealthier households. Belonging to the top quartile of wealth reduces the persistence coefficient by 0.13 points. In contrast, there are no significant differences for the later period (columns 3 and 4), indicating that the wealth differential in the early period has now disappeared.

In table 3 we further explore the possible nonlinear relationship between HAZ in the two earlier periods by fitting a linear probability model of the likelihood of being stunted (HAZ<-2.0)[15] and severely stunted (HAZ<-3.0) in both the early and later periods as a function of our earlier set of household, maternal and child controls, as well as stunting in the previous period. The likelihood of stunting persisting into the next period is higher in later childhood than earlier on. Further, the results showed a strong correlation between stunting across the early periods for the poorer children compared to the richest. Whilst 36% of poor children remain stunted from age 6-18 months through to age 4-5, only 19% of the previously stunted wealthy children do so. The pattern is similar for severe stunting. Consistent with the previous results, such wealth differentials disappear in the later period. This evidence supports the suggestive patterns of catch-up depicted in Figure 3, namely that in the early period, relatively wealthy children with early nutritional deficiencies appear to achieve higher catch-up growth than equally stunted but poorer children. In the later period, no significant differences emerge.

---

[14] HAZ levels in the current period do not appear to increase with wealth quartiles and are not reported to keep the tables parsimonious – possibly because any level effects are already being captured by HAZ status in round 1.

[15] As noted in the data section, WHO standard (5cm and 9cm below the mean for a one and five year old boy respectively). Note also that we ran a probit model on this variable which showed very similar results, but we report here the linear probability model, as we use a linear model to fit the IV later on in the paper.

3.5.1 Instrumental variables estimation[16]

Despite the careful selection of control variables, we are unable to allay the concern that unobserved factors might determine both nutritional attainment in the early stage of life, and subsequent nutritional development in the second stage of life. OLS estimates of the relation between lagged HAZ and current HAZ could be driven by these unobserved factors and therefore be spurious. Note, however, that as discussed in the previous section, there is no clear reason why endogeneity would cause a spurious disparity in catch-up across the wealth distribution.

IV Fuller estimates for height-for-age z-scores in the early childhood period are presented in table 4 following the system of equations (3) and (4). Column (1) reports IV results for the full sample, while columns (2) and (3) show estimates for the bottom half of the wealth distribution, using wealth index and maternal height respectively as the wealth indicator. For the full sample, we find that Fuller estimates yield a parameter estimate on lagged HAZ similar to OLS estimates (0.25). Estimates for the bottom half of the wealth distribution are substantially higher than the full sample coefficients. For the bottom half of the wealth index, the persistence parameter is 0.42, with a similarly high coefficient for the bottom half of the maternal height distribution. The bottom panel in Table 4 also reports IV estimates on stunting status. Estimates show a substantial increase in the coefficient on lagged height, with parameter estimates of 0.68 for the full sample, and 0.90 – effectively suggesting full persistence – among the poorest households. This may be due to a higher level of measurement error in the binary model – but it could be due to the poorer fit of the first-stage.

---

[16] As there is little variation in the height-for-age between the two later periods of mid-childhood, we present the IV analysis for the first two periods only.

In all models, the coefficient for the poorer group is more precisely estimated than for the full group, which goes some way to dispel the concern that the wealth differences in catch-up were driven by unobservables causing an upward bias; it appears that if anything, the opposite may be true. IV estimates support (and strengthen) our earlier OLS findings: catch-up rates are larger among wealthier children and this effect primarily takes place among undernourished children.

Table 4 also provides information on the strength of our instruments. We compare the F-Statistics of the first-stage explanatory power of the excluded instruments reported at the bottom of the table with the Stock and Yogo (2005) weak-IV test critical values. The Kleibergen-Paap F-Statistic for our IV estimates ranges from 7 to 10.5, marginally passing the weak-IV rule-of-thumb of a value of 10 suggested by Staiger and Stock (1997). Stock and Yogo critical values suggest that our Fuller estimates may contain a 30% bias or more. The remaining bias is understood to be towards the OLS estimates; IV Fuller estimates therefore provide a lower bound to the true point estimates.[17] However, instrument weakness also affects inference testing, rendering standard errors invalid. Table 4 reports the implied p-values from fully robust Anderson-Rubin 95% confidence intervals.[18]

We find that among the poorest households, the lagged HAZ coefficient is significantly different from zero at the 10% level (with a robust weak-IV p-value of 6%). However, the estimated coefficient on persistence for the full sample appears to be insignificant.[19]

---

[17] When we use a GMM/LIML estimator, which arguably carries a smaller OLS bias, but is less robust to the effect of weak instruments, we find indeed coefficient estimates are higher, supporting the theory that the reported Fuller estimates represent a lower bound to the true parameter.

[18] We also compute confidence intervals applying Conditional Likelihood Ratio Moreira (2003) procedure and find results similar to the Anderson-Rubin procedure.

[19] Note that instrument weakness is higher for the full sample than for the bottom half sample, implying that the difference in catch-up rates found in the IV estimates could be driven by differences in the weak-IV bias. Results from the AR

Consequently, even with some uncertainty around average catch-up parameters, we can be confident to conclude that the bottom and top half households in the wealth distribution experience significantly different levels of catch-up growth, with the poorer half experiencing only partial catch-up while the richer half appear to catch-up fully.

Estimates of the degree of catch-up growth in height among children vary significantly in the literature. Using experimental data, Ruel et al. (1995) obtain estimates of persistence in height of 0.75 and 0.61 for boys and girls respectively for the period between 3 years of age and adolescence. When using IV methods, Hoddinott and Kinsey (2001) report point estimates of 0.56 on lagged height for children 12-36 months old, but a coefficient as small as 0.19 for the same age group when applying mother fixed effect methods. Fedorov and Sahn (2005), using yet again a different method – IV Arellano-Bond GMM methods – obtain estimates of 0.20 for children from 0 to 76 months of age. Our full sample IV estimates contribute to this literature. The persistence parameter estimate of 0.25 for 4-5 year olds is closest to Fedorov and Sahn's study, though we can not totally rule out the possibility that catch-up is complete.

However, we show that average catch-up rates can mask substantial heterogeneity affected by nutritional remediation. Differences in catch-up rates between the bottom and top halves of the wealth distribution can be substantial, 0.42 versus an implied value of 0.08 respectively.[20] Moreover, allowing for the non-linear nature of health – by analysing stunting status – further increases the disparity in catch-up rates: a gap of 0.47 between the bottom (0.91) and top (0.441) halves of wealth. This evidence suggests that targeted

---

confidence intervals dispel this possibility and indicate that difference persistent when the effect of any weak instruments is removed.

[20] The implied coefficient for the top half is calculated using the point estimates for the average and those of the bottom half of the sample.

nutritional remediation in early childhood can be very effective, and in the absence of such investments, height is likely to experience only limited catch up. In fact, stunted children in relatively poor households have little chance of catching up with their healthy peers.

The IV and OLS results on wealth differentials in the early years are robust to different methodologies. Indeed we find that OLS estimates underestimate the socioeconomic gradient of catch-up growth. Similarly, we consider that the significant difference between the early and mid-childhood OLS estimates – in particular the fact that the economic gradient disappears – offers convincing evidence that the opportunities for catch-up growth are better in the earlier years, and that nutritional remediation after the age of five might be too late.

The costs of the catch-up deficits accumulated early on in life cannot be understated. Poorer nutrition early on in life will affect other forms of human capital accumulation (such as cognitive development and educational attainment), adult health as well as adult earnings. Recent research on the returns to height in rural Ethiopia (Dercon and Porter, 2012), allow us to derive estimates of the impact of lower catch-up rates documented here on loss of lifetime adult earnings. Using a long running longitudinal study of rural Ethiopian households (the ERHS), Dercon and Porter, 2012, estimate the mincerian returns to height and find that 1 cm lower adult height reduces annual income by 1%. From Table 4 we can obtain a conservative measure of differential catch-up between the bottom and top half of the distribution, 0.17 points of height-for-age z-score. Assuming that these differences remain constant from the age of five until adulthood, as suggested by our evidence and the wider nutritional critical period literature, this deficit in nutritional catch-up would translate into 1.4% lower annual income over their adult lifetime. [21] This is a remarkably large loss of income considering that it only captures the impact of height, but not other confounders, such education.[22] It also illustrates how nutritional deficits among poor children will have a substantial welfare effect over their lifetime.

3.5.2 Channels of Nutritional Remediation

We now investigate the wealth differentials in more detail. Our composite wealth variable outlined in section 3 comprises three sub-indices: *housing quality*, *services index* (water, sanitation, electricity and cooking fuel), and *consumer durables*. To further investigate the channels through which wealth impacts on catch-up in the early period, we re-estimate equation (2) for each sub-index independently. Because sub-indeces can be very discrete in their distribution, we use indicator variables to capture their variation instead of the quartile analysis. Table 1b shows a detailed descriptive breakdown of the wealth index and the indicator variables.[23]

Table 5 reports a summary of our findings. Among the three components, *access to services* appears to be the most important. Having one of the four services increases catch-up growth by 7 per cent and having two or more services by 13 per cent (column 3). In terms of stunting, one (two or more) service(s) reduces the probability by 17 (35) per cent. We interpret these results as evidence that services that improve the child's environment have

---

[21] We take a conservative approach, by assuming that catch-up rates equal the full sample estimates reported in column (1) of Table 4. The 0.17 coefficient is thereby the difference between the coefficients reported in column (2) and (1). With a standard deviation of 1.08 in the height-for-age distribution at the age of 4-5 years, 0.17 represents 0.16% of a standard deviation.

[22] Dercon and Porter (2012) report a 1% loss in annual income for every 1cm lost in adult height, the latter representing a drop in 11.6% of a standard deviation in adult height. In their equation they do control for education, gender and household composition, but not for adult health. Accordingly, the height effect should be interpreted as the impact of height on income through a range of channels, such as health, physical ability, status, etc, but excluding other confounding effects such as education.

[23] Using the quartile analysis for the sub-indices yields similar results to the analysis discussed here, though substantially less precise.

complementary (and possibly separate) impacts on nutritional intake in terms of ability to catch-up from nutritional shocks at an early age, for example through reduced infections and illnesses (Burger and Esray, 1995; Alderman et al., 2003; Merchant et al., 2003). Increased *consumer durables* and *housing quality* do not have significant impacts on their own. However, as we would expect - and just as they do for the services sub-index - coefficient estimates are negative and increase in magnitude with the number of items (see the columns (4) and (5) in particular), suggesting that housing quality and consumer durables capture distinct nutritional effects.

So far we have shown that being born in a wealthier household ensures that early nutritional deficiencies can be recovered, in some cases even fully. However, this notion relies on dynamic equations that, as discussed earlier, might be subject to problems of endogeneity. Our instrumental variable approach aims are solving those concerns, however, we now seek to explore the question of differential catch-up in a different fashion. We have shown that season of birth has strong effect on early nutritional attainment, and that birth is exogenous to households and individual children. We analyse here the direct effect of the season of birth on height-for-age. Firstly, whether its effect is persistent over time, and secondly, whether richer households are able to reduce, or even erase, its impact.

Table 6 presents results for a range of model specification. We estimate reduced form equations of the impact of season of birth on HAZ and Stunting status at ages 0-1, 4-5 and 7-8. All regressions include a full set of controls and community fixed effects, but no controls for lagged HAZ. The top panel stratifies the regressions by wealth index: top and bottom half of the distribution. Meanwhile the bottom panel splits the sample using sub-indices of wealth index: households with low housing quality (N=744), low consumer durables (N=483), and low services (N=577).[24]

The table provides evidence consistent with other findings in the paper. First, quarter of birth has a significant effect on infant HAZ (0-1 years) only for children in the bottom distribution of wealth. Richer children are not affected by the timing of the birth. Only the poor are vulnerable to the negative effects of the quarter of birth.[25] Secondly, as the children grow older the impact of the season of birth is diminished, suggesting that there is some catch-up taking place even among the poorest children, which is consistent with the relatively high coefficients we find for lagged HAZ and the non-parametric analysis for HAZ at age 4-5 and 0-1 years.

Column (3) shows that the impact of season disappears by the age of 7-8. While this might be true at the mean of the height-for-age distribution, analysis of stunting suggests otherwise. Columns (4) to (6) indicate that season of birth is a significant and persistent determinant of stunting, even at the age of 7-8 years. As a poor child, even by the age of 7-8 years old, having been born in the second or fourth quarter of the year increases by 8% the likelihood of being stunted.

We are also interested in whether early life nutrition has different consequences for boys and girls – differences may be driven by behavioural factors – such as a pro-boy bias in the intra-household allocation of resources – or biological factors (or both). Studies of animals (see summaries in Boersma and Wit, 1997) have shown that females have higher potential

---

[24] We define a household as scoring 'low' on a specific sub-index, when it has no item in the list that conform each sub-index (See table 1b). For example, low housing quality means that a household had neither of the following: wall, roof, floor of good durability, or if the house has more than one room. Estimates for those scoring 'high' on the sub-indices are not shown.

[25] The same is found when the effect of season is explored among households with 'high' scores in the sub-index. Results not shown in the table.

for catch-up growth. Additionally, in a well-known Guatemalan nutrition study Ruel et al. (1995) found that a nutritional supplement had a higher impact for girls of lower socio-economic status at age 3, which persisted into adolescence. Deolalikar (1996) in a Kenyan study found higher rates of catch-up for girls in terms of weight (though did not test across socioeconomic status), especially at younger ages.

The four panels in Figure 4 provide lowess estimates similar to Figure 3 for each gender separately. It becomes apparent that nutritional remediation differs across genders. Girls benefit more from living in wealthier households than their equally undernourished male counterparts. Table A5 in the Annex reports OLS parametric estimates split by gender. Whilst there are not many significant differences, girls have significantly higher rates of catch-up in the top wealth quartile (0.16 lower than the average coefficient of 0.27). Indeed for boys, nutritional remediation appears to be ineffective. Consistent with earlier findings, we also find that in mid-childhood nutritional investments are no longer effective for either gender. Nevertheless, with our data we are not able to test whether these gender patterns are because in early childhood girls' nutrition is more sensitive to other inputs such as sanitation, or that in credit-constrained or poorer households, girls are allocated less food.

3.6 Conclusion

Examining a group of poor rural Ethiopian children, we find a clear relationship between nutritional status measured at age 6-18 months and nutritional status four years later. We find that malnourished children in the richer households do experience significantly higher catch-up rates than malnourished children in the bottom quartile of the wealth distribution. We find evidence that catch-up among wealthier girls is substantially higher than among poorer girls, whereas the socioeconomic status of the household does not appear to affect catch-up growth of boys. Considering that our sample is pro-poor from a very poor country, the results are quite alarming – even in this context, being poorer seems to indicate fewer possibilities to catch-up from negative shocks in early life. Examining the same children during the 5-8 years period, HAZ position in the distribution is very persistent, and does not vary by household wealth, nor by gender.

We cannot distinguish between several competing explanations in terms of why there is differential catch-up according to wealth in the early period. Both richer and poorer households may wish to compensate for poor endowments, but richer households may be more successful as they have more resources available. Our findings indicate that it is through access to essential services such as sanitation and electricity that wealth has a strong effect on nutritional catch up.

We cannot fully rule out the influence of unobservable characteristics of households. Significant differences across the wealth distribution could possibly arise because richer and poorer households have different preferences (for example, poorer households simply prefer to reinforce sibling differences in endowments). Rosenzweig and Wolpin (1980) show that there is an under identification problem in terms of distinguishing between parameters of the child nutrition "production function" and the preferences of the household (though in the context of cognitive ability). Another possibility that we cannot exclude is that richer households may have unobserved ability (beyond their education which is factored into our model) that allows them to compensate and achieve catch-up growth, additional to their extra wealth (e.g. noticing malnutrition sooner; or knowing that if the child is small then feeding more/improving nutritional intake will help the child; and which foods are better suited to this purpose). Or, richer households may have access to better investment `technology', that is they are more effective in improving nutritional outcomes for a given amount of expenditure over and above improved water and sanitation, which are included in our wealth measure (for example, purchasing food with a higher nutritional content).

A priori, it seems to us unlikely that poorer households would have completely different preferences, and rather more likely that the marginal benefit of investing in an undernourished child has higher opportunity cost in terms of other households members' nutrition. Whether or not there is more than one channel of impact, the outcome is clear: that early life chances are considerably lowered for under nourished children from poorer households. Further, if this is not addressed before the age of five, then the window of opportunity for catch-up may have closed. Moreover, other evidence in the literature shows that even in developed countries, health differentials that manifest in the early years appear to be exacerbated as children get older (Case et al., 2002).

This suggests that in the absence of other interventions, future surveys of Young Lives children in Ethiopia may find increasing disparities in nutritional outcomes and other measures of child development as the children age. Our own calculations suggest that height-for-age deficits accumulated by the bottom half of the distribution at the age of five, might translate in 1.4% of loss annual income during their adult lifetimes. It is sobering to realise that children as young as five, might already be faced with a lifetime burden in the form of reduced health, education and earnings.

References

Adair, L.S., 1999. Filipino children exhibit catch-up growth from age 2 to 12 years. Journal of Nutrition 129: 1140-1148

Alderman, H., 2010. The economic cost of a poor start to life. Journal of Developmental Origins of Health and Disease 1, 19-25.

Alderman, H. and L. Christiaensen, 2001, Child Malnutrition in Ethiopia: Can Maternal Knowledge Augment The Role of Income? . Africa Region Working Paper Series 22.

Alderman, H., Hentschel, J., & Sabates, R. (2003). With the help of one's neighbors: externalities in the production of nutrition in Peru. Social Science & Medicine, 56(10), 2019-2031.

Alderman, H., Hoddinott, J., Kinsey, B., 2006. Long term consequences of early childhood malnutrition. Oxford Economic Papers 58, 450

Almond D., Currie, J., 2010. Human Capital Development Before Age Five. NBER Working Paper No. 15827 March 2010, Revised January 2011

Andrews, D.W. K.; Moreira, M. J.; Stock., J. H., 2006. Optimal Two-Sided Invariant Similar Tests for Instrumental Variables Regression. Econometrica 74, 715-752.

Artadi, E. 2005. Going into labor: Earnings vs infant survival in rural Africa. Mimeograph, Bocconi University.

Barker, D.J.P., 1992, Fetal and infant origins of adult disease. British Medical Journal, xv, 343.

Batty, G. D., Shipley, M. J., Gunnell, D., Huxley, R., Kivimaki, M., Woodward, M., Man Ying Lee, C., Smith, G. D., 2009. Height, wealth, and health: An overview with new data from three longitudinal studies. Economics & Human Biology, Volume 7, Issue 2, July 2009, 137-152

Behrman, J.R., Pollak, R.A. and Taubman, P., 1982. Parental Preferences and Provision for Progeny. Journal of Political Economy, 90, 52-73.

Behrman, J. R., Rosenzweig, M. R. 2004. Returns to Birthweight. Review of Economics & Statistics 86, 586-601

Bhalotra, S. and Rawlings, S.B., 2011. Intergenerational persistence in health in developing countries: The penalty of gender inequality? Journal of Public Economics, 95, 286-299.

Boersma, B., Wit, J.M., 1997. Catch-up Growth. Endocrine Reviews. October 1, 1997 vol. 18 no. 5, 646-661

Burger, S. E.,. Esrey, S.A., 1995. Water and sanitation: health and nutrition benefits to children. In: Pinstrup-Andersen, P., D. Pelletier, H.Alderman (eds.) Child growth and nutrition in developing countries : priorities for action. Ithaca; London : Cornell University Press.

Case, A., Lubotsky, D., Paxson, C., 2002. Economic Status and Health in Childhood: The Origins of the Gradient. American Economic Review 92, 1308-1334.

Chodick G, V. Shalev, I. Goren and P.D. Inskip, 2007, Seasonality in birth weight in Israel: new evidence suggests several global patterns and different etiologies. Annals of Epidemiology. Vol. 17, 440–446.

Chodick G, S. Flash, Y. Deoitch and V. Shalev, 2009, Seasonality in birth weight: review of global patterns and potential causes. Human Biology, Vol. 81, 463–477.

Costa, D. L. and J. N. Lahey. 2005. Predicting older age mortality trends. Journal ofthe European Economic Association, 3(2-3):487{493, April-May

Cunha, F., Heckman, J., Lochner, L., Masterov, D. 2006. Interpreting the Evidence on Life Cycle Skill Formation. In: Hanushek, E., Welch, F. (eds.) Handbook of the Economics of Education, chapter 12, pages 697-812. Elsevier

Cunha, F, and Heckman, J., 2007. The Technology of Skill Formation.American Economic Review 97(2):31–47.

Currie, J. Moretti, E., 2007. Biology as Destiny? Short and Long Run Determinants of Intergenerational Transmission of Birth Weight. Journal of Labor Economics 25, 231-264.

Deolalikar, A. B., 1996. Child nutritional status and child growth in Kenya: Socioeconomic determinants. Journal of International Development 8(3): 375-393.

de Onis M, Garza C, Victora CG, Bhan MK, Norum K.R., 2004. The WHO Multicentre Growth Reference Study (MGRS): Rationale, planning and implementation. Food and Nutrition Bulletin, Volume 25, Number 1, Supplement 1, March 2004

de Onis, M., O. A.W., et al., 2007. Development of a WHO growth reference for school-aged children and adolescents. Bulletin of the World Health Organisation 85(9): 649-732

Deaton, A., 2007. Height, health and development. Proceedings of the National Academy of Sciences of the United States of America vol:104 (33):p13232.

Dercon, S., Krishnan, P., 2000. In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia. Journal of Political Economy 108, 688

Dercon, Stefan and C. Porter, 2012. Live aid revisited: long-term impacts of the 1984 Ethiopian famine on children. CEPR Discussion Papers 9033, C.E.P.R. Discussion Papers

Doblhammer, G and J. W. Vaupel. 2001. Lifespan depends on month of birth. Proceedings of the National Academy of Sciences, 98(5):2934{2939.

Doyle, O., Harmon, C. P., Heckman, J. J., Tremblay, R. E., 2009. Investing in early human development: Timing and economic efficiency. Economics & Human Biology, 7(1), 1-6

Elbers, C., Gunning, J. W., Kinsey, B. H., 2002. Convergence, Shocks and Poverty. Discussion Paper 2002–035/2, Tinbergen Institute, Amsterdam

Eveleth, P.B. and J.M. Tanner, 1976. Worldwide Variation in Human Growth. Cambridge: Cambridge University Press.

Fedorov, L., Sahn, D. E., 2005. Socioeconomic Determinants of Children's Health in Russia: A Longitudinal Study. Economic Development and Cultural Change, Vol. 53, No. 2, pp. 479-500, January 2005

Ferro-Luzzi, A., Morris, S. Taffesse, S. Demissie T., D'Amato M., 2001. Seasonal Undernutrition in Rural Ethiopia. IFPRI Research Report 118.

Finch, B.K. and Beck, A.N., 2011. Socio-economic status and z-score standardized height-for-age of U.S.-born children (ages 2-6). Economics & Human Biology, 9, 272-276.

Garza C, de Onis M. A new international growth reference for young children. American Journal of Clinical Nutrition, 1999, 70:169S-172

Grantham-McGregor, S. et al. 2007. Developmental potential in the first 5 years for children in developing countries. Lancet 2007; 369: 60–70.

Harvey, P (ed) (2009) Disasters special issue on Ethiopia available online. http://www.wiley.com/bw/vi.asp?ref=0361-3666#383 (accessed April 14, 2011)

Hoddinott, J.; Kinsey, B., 2001. Child Growth in the Time of Drought. Oxford Bulletin of Economics and Statistics, vol. 63(4), 409-36, September.

Mani, S., 2008. Is there Complete, Partial, or No Recovery from Childhood Malnutrition? Empirical Evidence from Indonesia, Fordham Economics Discussion Paper Series dp2008-19, Fordham University, Department of Economics.

Martorell, R., Khan, L. K., Schroeder, D. G. Reversibility of stunting: epidemiological findings in children from developing countries. European journal of clinical nutrition 1994;48 Suppl 1:S45-57.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell and A. R. Quisumbing, 2006, The impact of an experimental nutritional intervention in childhood on education among Guatemalan adults, FCND Briefs (International Food Policy Research Institute).

McGrath, J.J., D. Keeping, S. Saha, D.C. Chant, D.E. Lieberman, and M.J. O'Callaghan 2005. The association between birth weight, season of birth and latitude. Ann. Hum. Biol. 32, 547–559.

Mekonnen, A., B. Tefera, T. Woldehanna, N. Jones, J. Seager, T. Alemu, and G. Asgedom, 2005, Child Nutritional Status in Poor Ethiopian Households: The Role of Gender, Assets and Location Young Lives Working Paper 26.

Merchant, A. T., Jones, C., Kiure, A., Kupka, R., Fitzmaurice, G., Herrera, M. G., 2003. Water and sanitation associated with improved child growth. European Journal of Clinical Nutrition, 57(12), 1562-1568.

Moreira, M.J., 2003. A Conditional Likelihood Ratio Test for Structural Models. Econometrica, 71 (4), 1027-1048.

Morrow, V., 2009. The Ethics of Social Research with Children and Families in Young Lives: Practical Experiences, Working Paper 53, Oxford: Young Lives

Murray, M. P., 2006. Avoiding Invalid Instruments and Coping with Weak Instruments. Journal of Economic Perspectives 20(4): 111-132.

Outes-Leon, I. and Dercon, S., 2008. Survey Attrition and Attrition Bias in Young Lives, Young Lives Technical Note 5, Oxford: Young Lives

Outes-Leon, I. and Sanchez, A., 2008. An Assessment of the Young Lives Sampling Approach in Ethiopia, Young Lives Technical Note 1, Oxford: Young Lives

Rona, R.J., Mahabir, D., Rocke, B., Chinn, S. and Gulliford, M.C., 2003. Social inequalities and children's height in Trinidad and Tobago. European Journal of Clinical Nutrition, 57, 143-150.

Rosenzweig, M. R.; Wolpin, K. I., 1980. "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," Econometrica, vol. 48(1), 227-40, January.

Ruel, M.T., Rivera, J., Habicht, J.P. and Martorell, R., 1995. Differential Response to Early Nutrition Supplementation: Long-Term Effects on Height at Adolescence. International Journal of Epidemiology, 24, 404-412.

Schroeder D.G., Martorell R., Rivera J.A., Ruel, M.T., Habicht J-P.,1995. Age differences in the impact of nutritional supplementation on growth. Journal of Nutrition 1995;125(4 suppl):1051S1051

Staiger, D. and Stock, J.H.,1997. Instrumental Variables Regression with Weak Instruments. Econometrica, 65, 557-586.

Steckel, R., 1983 “Height and per Capita Income.” Historical Methods 16 (1983): 1-7.

Steckel, R., 1995. “Stature and the Standard of Living.” Journal of Economic Literature 33 (1995): 1903-40.

Stock, J. H., Yogo, M.,2002. Testing for Weak Instruments in Linear IV Regression. NBER Technical Working Papers. I. National Bureau of Economic Research.

Strand, L.B, A.G. Barnett and S. Tong, 2001, The influence of season and ambient temperature on birth outcomes: A review of the epidemiological literature, Environmental Research, Vol. 111, Issue 3, April 2011, Pages 451–462

Strauss, J., Thomas, D., 2008. Health over the life course. Handbook of Development Economics, Volume IV. T. P. Schulz and D. Thomas (eds), Elsevier Press.

Todd, P., Wolpin, K., 2003. On the Specification and Estimating of the Production Function for Cognitive Achievement, Economic Journal, Feb, 2003, pp F3-F33.

Van Hanswijck de Jonge, L., G. Waller and N. Stettler, 2003. Ethnicity modifies seasonal variations in birth weight and weight gain of infants. J. Nutr. 133, 1415–1418.

Victora, C. G., de Onis, M., Hallal, P. C., Blossner, M., & Shrimpton, R., 2010. Worldwide Timing of Growth Faltering: Revisiting Implications for Interventions. Pediatrics, 125 (3), e473-480.

Walker, S. et al., 2007. Child development: risk factors for adverse outcomes in developing countries. Lancet 2007; 369: 145–57

Webb, E., Kuh, D., Peasey, A., Pajak, A., Malyutina, S., Kubinova, R., Topor-Madry, R., Denisova, D., Capkova, N., Marmot, M. et al., 2008. Childhood socioeconomic circumstances and adult height and leg length in central and eastern Europe. Journal of Epidemiology and Community Health, 62, 351-357.

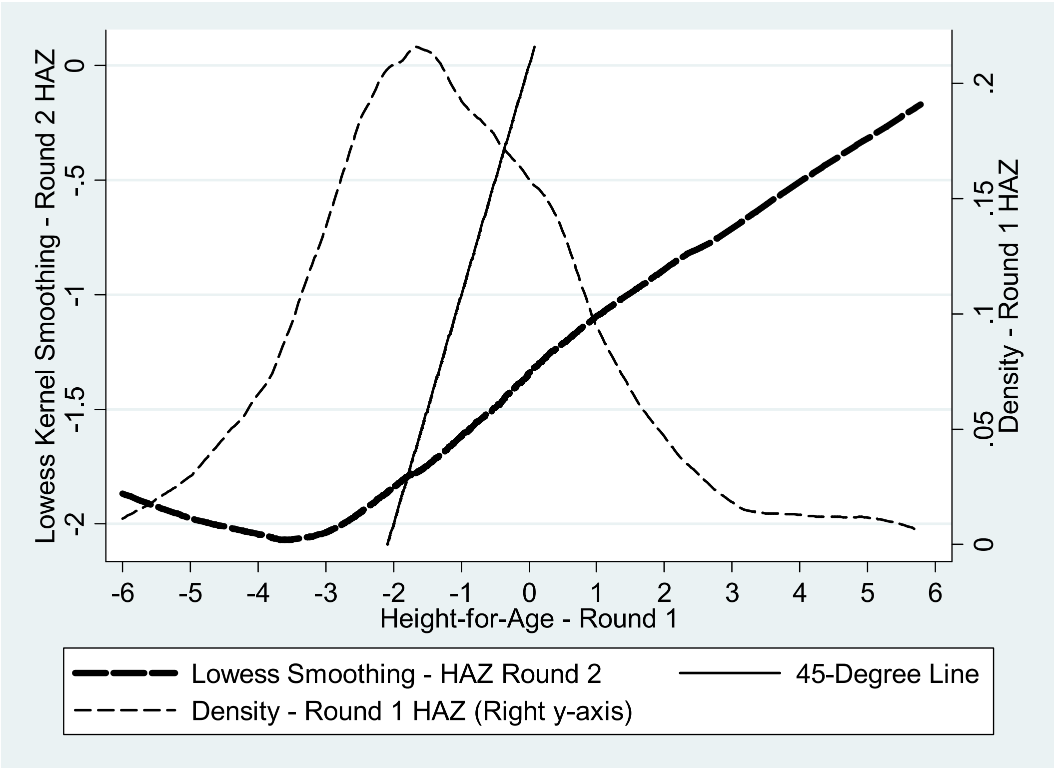
Wilson, I, Huttly S.R.A., 2004. Young Lives: A Case Study of Sample Design for Longitudinal Research, Oxford: Young Lives, Working Paper 10, 2004.

Wilson PN, Osbourn DF 1960 Compensatory growth after undernutrition in mammals and birds. Biol Rev 35:324–363

Yogo, M, 1997. Estimating the Elasticity of Intertemporal Substitution when instruments are weak. The Review of Economics and Statistics, August 2004, 86(3): 797–810.

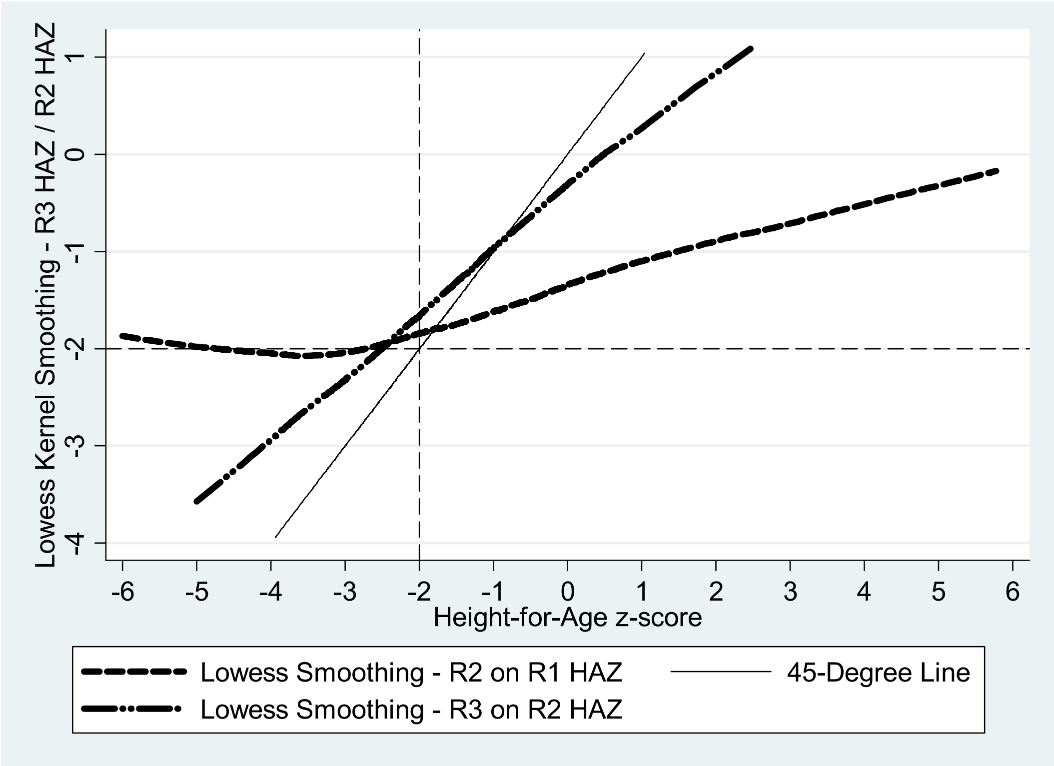
Figures

Figure 1: Average catch-up. Nonparametric estimation of height-for-age in Round 1 (2002) and Round 2 (2006)



Notes: Dashed line depicts the kernel density function of height-for-age z-scores in round 1 (2002) when children were aged 6-18 months. Thick long dashed line is the lowess kernel smoothing estimate of height-for-age z-score in round 2 (2006) against the height-for-age in round 1. The figure includes data for 913 children. For reference the solid line indicates the 45-degree line.

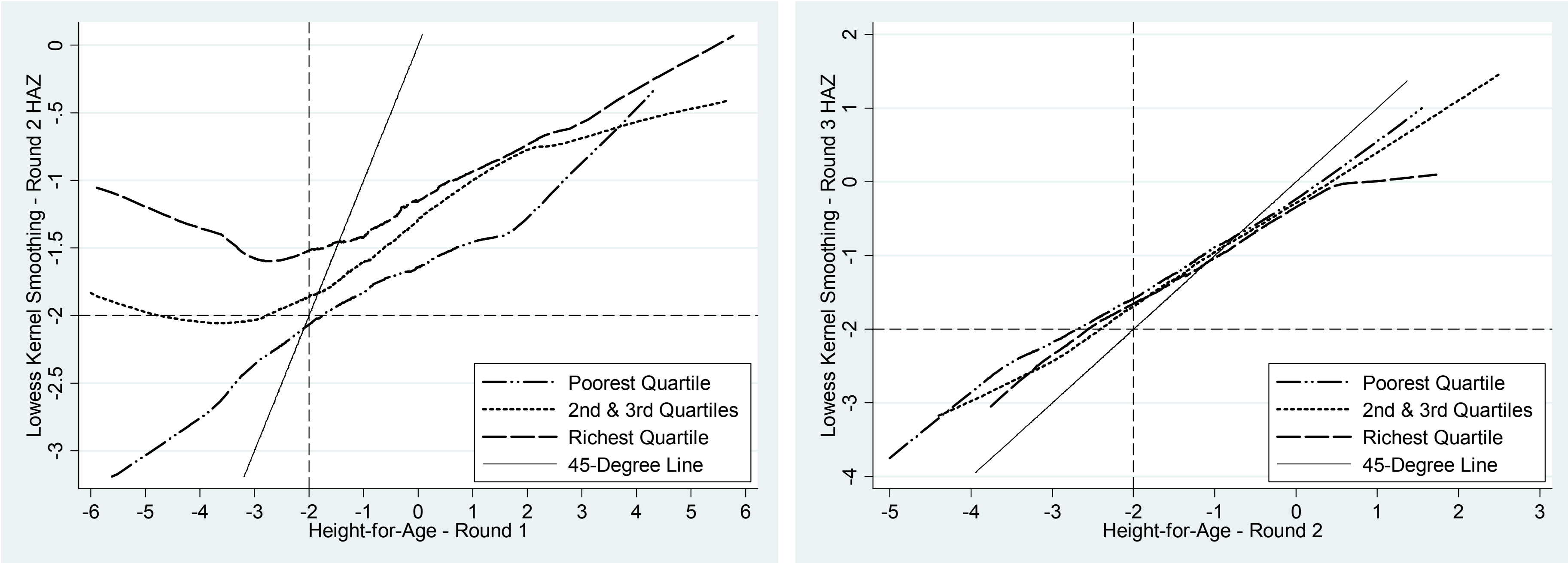
Figure 2: Average catch-up. Non-parametric relation between HAZ and HAZ lagged for early and mid-childhood periods



Notes: Dashed line depicts the kernel density function of height-for-age z-scores in round 1 (2002) when children were aged 6-18 months. Bold curves are lowess kernel smoothing estimates of height-for-age z-score and height-for-age z-score lagged. The 'longdash' line represents the relation between HAZ in round 2 (4-5 years) and HAZ in round 1 (0-1 years), while the 'longdash-dot-dot' line represents the relation between HAZ in round 3 (7-8 years) and HAZ in round 2 (4-5 years). Each kernel curve includes data for 913 and 903 children respectively. For reference the solid line indicates the 45-degree line.

Figure 3: Non-parametric estimates of catch-up by Wealth Quartiles

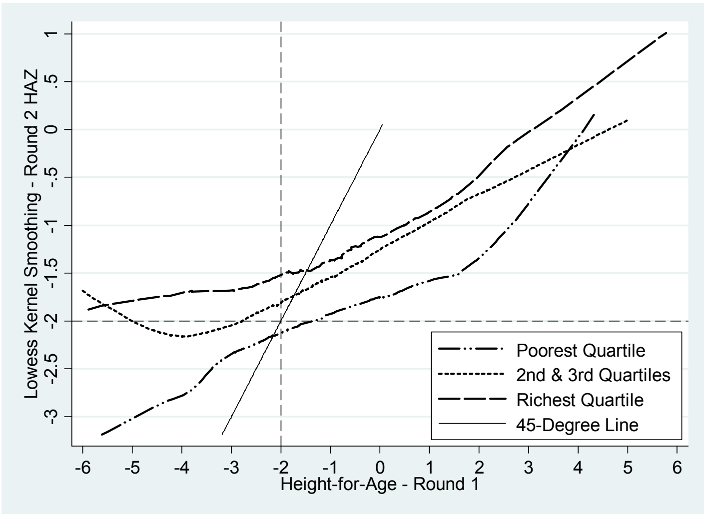
Panel A - Early childhood: 0-1 to 4-5 years of age **Panel B - Mid-Childhood: 4-5 to 7-8 years of age**



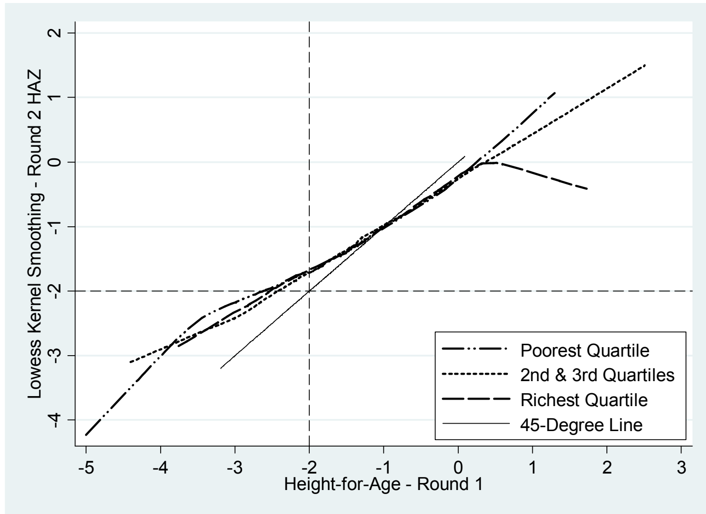
Notes: Curves are lowess kernel smoothing estimates of height-for-age z-score and height-for-age lagged for different wealth quartiles. The 'longdash' line represents the top quartile of wealth, the 'shortdash' line combines the second and third quartiles, while the 'longdash-dot-dot' line represents the bottom wealth quartile. Panel A reports the relation between round 1 (0-1 years) HAZ and round 2 (4-5 years) HAZ, while panel B reports the relation between round 2 (4-5 years) and round 3 (7-8 years) HAZ. The measure of wealth is an index comprised of three items: housing quality index, local services index and consumer durables index, all measured in round 1 (2002). Panels A and B include data for 913 and 903 children respectively across all wealth quartiles. For reference the solid line indicates the 45-degree line.

Figure 4: Non-parametric estimates of catch-up by gender and wealth

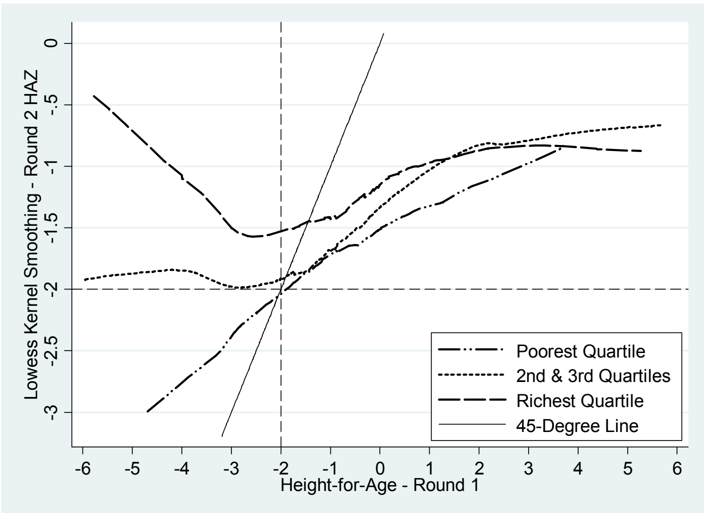
Panel A – Boys – Early childhood: 0-1 to 4-5 years of age



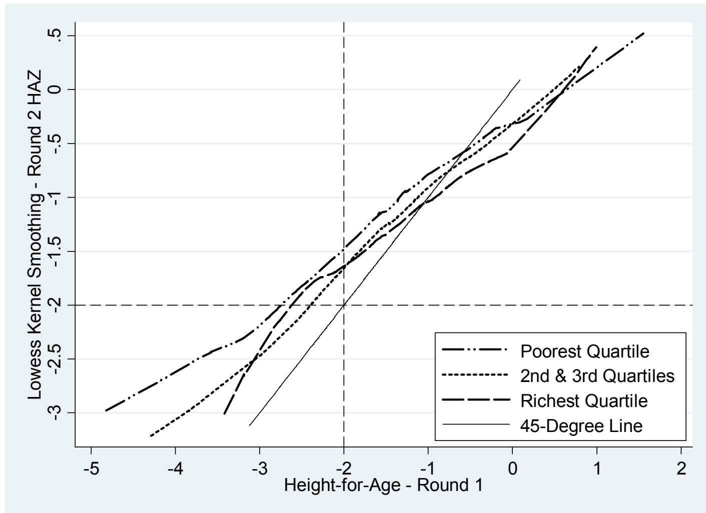
Panel B – Boys – Mid-Childhood: 4-5 to 7-8 years of age



Panel C – Girls – Early childhood: 0-1 to 4-5 years of age



Panel D – Girls – Mid-Childhood: 4-5 to 7-8 years of age



Notes: The ‘longdash’ line represents the top quartile of wealth, the ‘shortdash’ line combines the second and third quartiles, while the ‘longdash-dot-dot’ line represents the bottom wealth quartile. Panels A and C include data for 483 and 430 boys and girls respectively, while panels B and D include data for 478 and 425 boys and girls respectively. Solid line indicates the 45-degree line.

Tables

Table 1a: Summary of descriptive statistics

	Mean	Std. Dev.	Nr Observ.
Height-for-Age z-score, 2009, Age 7-8yr	-1.375	1.0309	903
Height-for-Age z-score, 2006, Age 4-5yr	-1.603	1.0794	913
Height-for-Age z-score, 2002, Age 0-1yr	-1.134	2.0268	913
Stunted, HAZ<-2.0, 2009, Age 7-8yr	0.244	0.4295	903
Stunted, HAZ<-2.0, 2006, Age 4-5yr	0.364	0.4813	913
Stunted, HAZ<-2.0, 2002, Age 0-1yr	0.341	0.4742	913
Stunted, HAZ<-3.0, 2009, Age 7-8yr	0.060	0.2372	903
Stunted, HAZ<-3.0, 2006, Age 4-5yr	0.096	0.2953	913
Stunted, HAZ<-3.0, 2002, Age 0-1yr	0.147	0.3541	913
Sex of the Child, 2002, Female=1	0.471	0.4994	913
HH Head Sex, 2002, Female=1	0.080	0.2714	913
HH Size, 2002	5.715	2.0706	913
Nr HH Adults, 2002	2.346	0.8249	913
Nr Male members, 2002	2.892	1.4735	913
Nr Brothers, 2002	1.317	1.3152	913
HH Wealth Index, 2002	0.087	0.0913	913
Caregiver reads, 2002, Easily=1	0.147	0.3541	913
Caregiver reads, 2002, With difficulty=1	0.112	0.3152	913
Any Schooling?, 2002, HH Head, Yes=1	1.817	0.3868	913
Age of the Mother, 2002	27.467	6.1651	913
Maternal Height, 2002, in CM	158.504	5.7379	913

Table 1b: Composition of the wealth index

Sub-Indices Items		*Wealth Index*			
		First Quartile	*Second Quartile*	*Third Quartile*	*Fourth Quartile*
Housing Quality	None	99.1%	80.3%	24.6%	5.7%
	One	0.4%	19.7%	56.1%	35.5%
	Two or More	0.4%	0.0%	19.3%	58.8%
Consumer Durables	None	100.0%	46.5%	80.7%	24.6%
	One	0.0%	52.6%	6.1%	44.3%
	Two or More	0.0%	0.9%	13.2%	31.1%
Services	None	100.0%	100.0%	82.5%	41.2%
	One	0.0%	0.0%	17.5%	42.1%
	Two or More	0.0%	0.0%	0.0%	16.7%
Nr Observations		229	228	228	228

Notes: Cells refer to the percentage of households in each quartile (first being the bottom quartile, and fourth the top) who own items as indicated. Housing Quality items: wall, roof, floor of good durability, or if the house has more than one room; Consumer Durables include 11 items such as radio, fridge, bicycle, table. Services Index comprises access to electricity, (clean, piped) water, sanitation (pit latrine or flush toilet) and cooking fuel (not wood or dung).

Table 2: Catch-up growth by Age Period - HAZ at Age 4-5 and 7-8 years, OLS Estimates

OLS Estimates	*Dependent variable: HAZ (t), Age 4-5 yr*		*Dependent variable: HAZ (t), Age 7-8 yr*	
	^ (t-1) = Age 0-1		*^ (t-1) = Age 4-5*	
	No Controls	*Full Controls*	*No Controls*	*Full Controls*
	(1)	(2)	(3)	(4)
Average Catch Up Model				
Height-for-Age (t-1) ^	0.236***	0.239***	0.711***	0.693***
	(0.019)	(0.020)	(0.028)	(0.024)
Full Controls	No	Yes	No	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes
R-Squared	0.295	0.384	0.540	0.574
Nr Observations	913	913	903	903
Wealth Interaction Model				
Height-for-Age (t-1) ^	0.295***	0.295***	0.693***	0.667***
	(0.034)	(0.032)	(0.045)	(0.046)
(HAZ (t-1)) x (Quartile 2, Wealth Index)	-0.042	-0.035	0.045	0.074
	(0.053)	(0.050)	(0.050)	(0.057)
(HAZ (t-1)) x (Quartile 3, Wealth Index)	-0.070	-0.069	0.040	0.039
	(0.052)	(0.043)	(0.052)	(0.056)
(HAZ (t-1))x (Quartile 4, Wealth Index)	-0.128**	-0.101**	-0.010	-0.009
	(0.052)	(0.049)	(0.061)	(0.063)
Wealth Quartile Dummies	Yes	Yes	Yes	Yes
Full Controls	No	Yes	No	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes
R-Squared	0.306	0.386	0.542	0.577
Nr Observations	913	913	903	903

Notes: * p<0.10, ** p<0.05, *** p<0.01. ^ For each model, HAZ (t-1) refers to the height-for-age z-score of the child in the previous period.
All estimates are Ordinary Least Squares with community fixed effects. Full controls include information on household characteristics (including caregiver gender, height, health, household wealth and demographics), mother and child information, as well as dummies for age in months, ethnic group and birth order. The lower panel of the table shows the interaction of the lagged HAZ with the wealth quartile (bottom quartile omitted) measured in round one. When interacting by wealth quartiles household wealth is excluded from the core controls. Standard errors are clustered at the community and quarter of birth level.

Table 3: Linear Probability Model, Stunting at 4-5 and 7-8 years of age

Linear Probability Model	*Dependent variable: Stunting (t), Age 4-5 yr* ^ *(t-1) = Age 0-1*		*Dependent variable: Stunting (t), Age 7-8 yr* ^ *(t-1) = Age 4-5*	
	Stunted (HAZ<-2.0)	*Stunted (HAZ<-3.0)*	*Stunted (HAZ<-2.0)*	*Stunted (HAZ<-3.0)*
	(1)	(2)	(3)	(4)
Stunted (t-1) ^	0.363***	0.196**	0.471***	0.236***
	(0.052)	(0.088)	(0.047)	(0.062)
(Stunted (t-1)) x (Quartile 2, Wealth Index)	0.028	-0.094	0.023	0.193*
	(0.085)	(0.089)	(0.064)	(0.099)
(Stunted (t-1)) x (Quartile 3, Wealth Index)	-0.026	-0.130	0.029	-0.128
	(0.090)	(0.095)	(0.065)	(0.123)
(Stunted (t-1)) x (Quartile 4, Wealth Index)	-0.173**	-0.217**	0.005	0.039
	(0.085)	(0.085)	(0.078)	(0.149)
Wealth Quartile Dummies	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes
R-Squared	0.106	0.025	0.283	0.138
Nr Observations	913	913	903	903

Notes: * p<0.10, ** p<0.05, *** p<0.01. Dependent variable is binary (0,1) which is positive if child is stunted, defined as HAZ<-2.0 (columns 1 and 3) or HAZ <-3.0 (columns 2 and 4). ^ For each model, Stunted (t-1) refers to stunting status of the child in the previous period. All estimates are Linear Probability Model with community fixed effects and full controls. Full controls include information on household characteristics (including caregiver gender, height, health and demographics), mother and child information, as well as dummies for age by month, ethnic group and birth order. Columns (1) and (2) estimate the relation between stunting at the age of 0-1 and 4-5 years, columns (3) and (4) explore the relation for the later period from 4-5 to 7-8 years of age. Standard errors are clustered at the community and quarter of birth level.

Table 4: Catch-Up Effects - Summary of IV Fuller and OLS Estimates

	Full Sample	*Bottom Half Wealth Index*	*Bottom Half Maternal Height*
	(1)	(2)	(3)
Dependent Variable: Height-for-Age (t), 4-5 years of age			
OLS - HAZ (t-1), Age 0-1 years	0.239***	0.270***	0.223***
	(0.020)	(0.031)	(0.026)
IV Fuller - HAZ (t-1), Age 0-1 years	0.250**	0.419***	0.368***
	(0.122)	(0.149)	(0.125)
First-Stage: IV F-Statistic	7.158	10.496	7.032
Anderson-Rubin p-value	0.280	0.059	0.032
Nr Observations	913	457	457
Instrument	season	season	season
Dependent Variable: Stunted (t), (HAZ<-2.0), 4-5 years of age			
OLS - Stunting (t-1), Age 0-1 years	0.326***	0.363***	0.339***
	(0.034)	(0.047)	(0.044)
IV Fuller - Stunting (t-1), Age 0-1 years	0.676**	0.911***	0.906***
	(0.277)	(0.315)	(0.271)
First-Stage: IV F-Statistic	5.253	8.304	4.044
Anderson-Rubin p-value	0.024	0.009	0.001
Nr Observations	913	457	457
Instrument	season	season	season

Notes: * p<0.10, ** p<0.05, *** p<0.01. All estimates include full controls and community fixed-effects. Full controls include information on household characteristics (including caregiver gender, height, health and demographics), mother and child information, as well as dummies for age by month, ethnic group and birth order. Instrumental variable ‘season’ takes a value of one for children born in the second and fourth quarter. IV estimates obtained using weak-IV robust k-class Fuller estimator. First-stage F-Statistic reports results for Kleibergen-Paap test of weak identification. Standard errors are clustered at the community and quarter of birth level.

Table 5: Disaggregated Wealth Index- Housing, consumer durables, services

OLS Estimates	*Dependent variable: HAZ, Aged 4-5*			*Dependent variable: Stunted (HAZ<-2.0), Aged 4-5*		
	Nr of Items by Wealth Sub-Index			*Nr of Items by Wealth Sub-Index*		
	Housing Quality Sub-Index	*Consumer Durables Sub-Index*	*Services Sub-Index*	*Housing Quality Sub-Index*	*Consumer Durables Sub-Index*	*Services Sub-Index*
	(1)	(2)	(3)	(4)	(5)	(6)
Height-for-Age (t-1), 0-1 years	0.264***	0.251***	0.257***			
	(0.027)	(0.022)	(0.020)			
Stunted (t-1) (HAZ<-2.0), 0-1 years				0.375***	0.344***	0.350***
				(0.048)	(0.034)	(0.036)
(HAZ / Stunted) x (Sub-Index, Items: One)	-0.040	-0.044	-0.074*	-0.069	-0.043	-0.168**
	(0.038)	(0.032)	(0.042)	(0.079)	(0.060)	(0.070)
(HAZ / Stunted) x (Sub-Index, Items: Two or More)	-0.055	0.014	-0.130**	-0.163	-0.073	-0.359***
	(0.047)	(0.037)	(0.058)	(0.104)	(0.108)	(0.093)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for Nr of Items by Wealth Sub-Indices	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.197	0.200	0.198	0.106	0.100	0.104
Nr Observations	913	913	913	913	913	913

Notes: * p<0.10, ** p<0.05, *** p<0.01. All OLS estimates include full controls and community fixed-effects. Full controls include information on household characteristics (including caregiver gender, height, health and demographics), mother and child information, as well as dummies for age by month, ethnic group and birth order, dummies for the wealth index quartiles and its components. Sub-indices for the wealth index are explained and decomposed in table [1b]. Stunting models are estimated using a Linear Probability Model. Standard errors are clustered at the community and quarter of birth level.

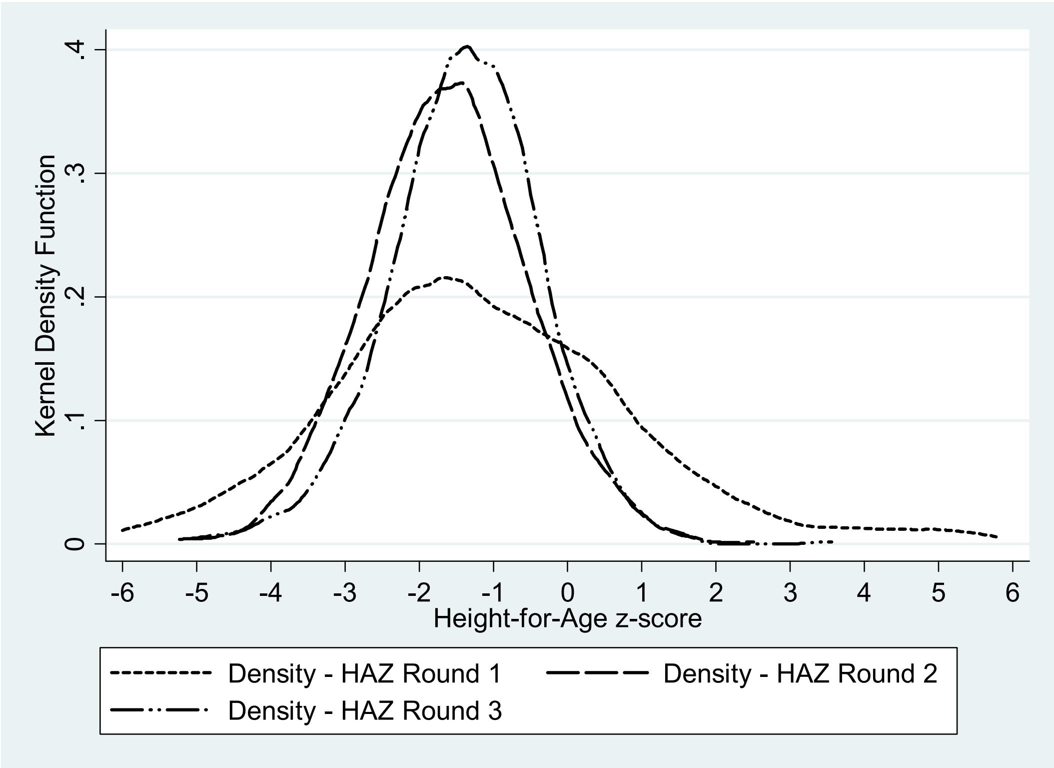
Table 6: Reduced Form – Impact of Quarter of Birth, on HAZ and Stunting, by Wealth Index and Sub-Indices

OLS Estimates	Height-fo-Age			Stunted HAZ		
	HAZ, 0-1 years	*HAZ, 4-5 years*	*HAZ, 7-8 years*	*HAZ, 0-1 years*	*HAZ, 4-5 years*	*HAZ, 7-8 years*
	(1)	(2)	(3)	(4)	(5)	(6)
Bottom Half - Wealth Distribution, N=457						
Quarter of Birth	-0.6121***	-0.3349**	-0.178	0.1454***	0.2157***	0.0772*
	(0.175)	(0.161)	(0.129)	(0.047)	(0.074)	(0.042)
Top Half - Wealth Distribution, N=457						
Quarter of Birth	-0.272	-0.004	-0.153	0.074	0.007	0.011
	(0.213)	(0.113)	(0.098)	(0.062)	(0.047)	(0.053)
Low Housing Quality, N=744						
Quarter of Birth	-0.4910**	-0.3417**	-0.2600*	0.1032*	0.2086***	0.1030**
	(0.209)	(0.157)	(0.133)	(0.055)	(0.066)	(0.051)
Low Consumer Durables, N=483						
Quarter of Birth	-0.4582**	-0.2413**	-0.3187***	(0.067)	0.0990**	0.0789**
	(0.183)	(0.105)	(0.117)	(0.048)	(0.050)	(0.039)
Low Services, N=577						
Quarter of Birth	-0.4202***	(0.156)	-0.1975**	0.0911**	0.1057**	(0.056)
	(0.163)	(0.115)	(0.096)	(0.045)	(0.043)	(0.034)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * p<0.10, ** p<0.05, *** p<0.01. Estimates include full controls and community fixed-effects. Full controls include information on household characteristics (including caregiver gender, height, health and demographics), mother and child information, as well as dummies for age by month, ethnic group and birth order. Standard errors are clustered at the community and quarter of birth level. Stunting models are estimated using a Linear Probability Model. Low Sub-Index samples correspond to households with a single item of the sub-index. See table [1b] for the decomposition of the sub-indices.

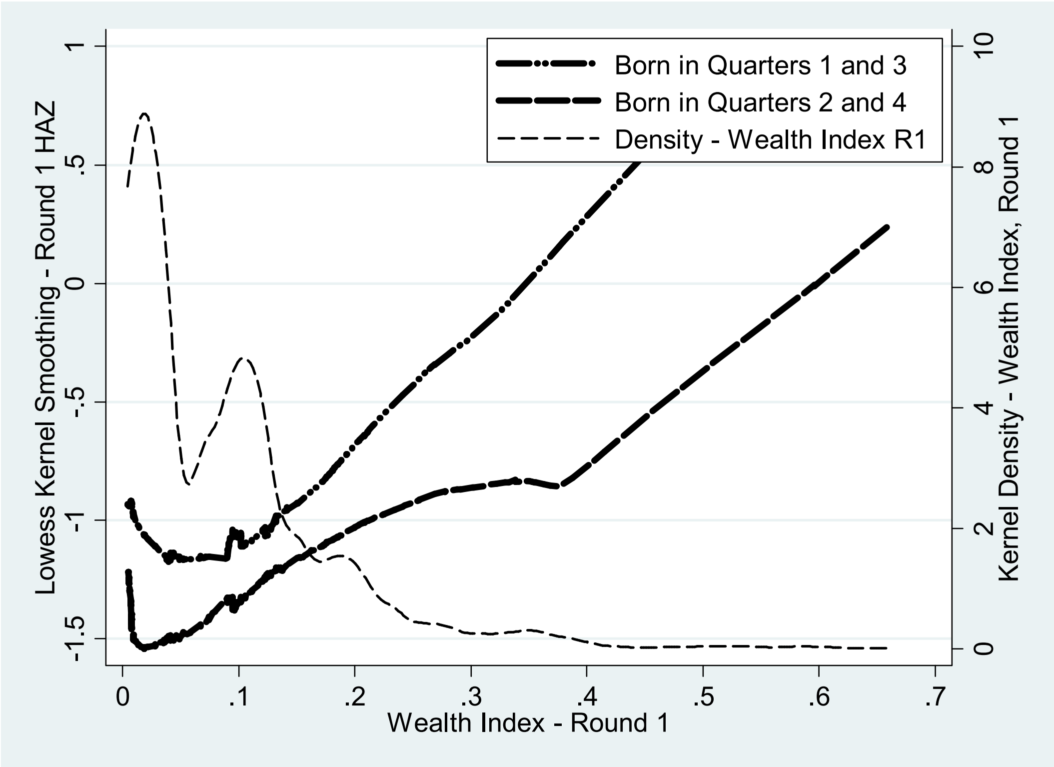
Annex

Figure A1: Height-for-age z-score density functions by round



Notes: Curves depict kernel density functions of height-for-age z-scores for different rounds. The 'shortdash' line stands for the kernel density for round 1 (0-1 years), the 'longdash' line for round 2 (4-5 years) while the 'longdash-dot-dot' line for round 3 (7-8 years).

Figure A2: Height-for-age z-score and wealth by quarter of birth



Notes: Dashed line depicts the kernel density function of height-for-age z-scores in round 1 (2002) when children were aged 6-18 months. Bold curves are lowess kernel smoothing estimates of height-for-age z-score and the wealth index. The measure of wealth is an index comprised of three items: housing quality index, local services index and consumer durables index, all measured in round 1 (2002). The 'longdash' line depicts the HAZ - wealth index relation for those children born in the first and third quarter of the year (393 children), while the 'longdash-dot-dot' line represents the relation for those born in the second and fourth quarter of the year (520 children).

Table A1: Detailed Baseline Estimates, Catch-up growth in early and mid-childhood

OLS Estimates	*Dependent variable: HAZ (t), Age 4-5 yr*		*Dependent variable: HAZ (t), Age 7-8 yr*	
	Average Catch Up	*Weatlh Interactions*	*Average Catch Up*	*Weatlh Interactions*
	(1)	(2)	(3)	(4)
Height-for-Age (t-1), Age 0-1 years	0.239***	0.295***		
	(0.020)	(0.032)		
Height-for-Age (t-1), Age 4-5 years			0.693***	0.667***
			(0.024)	(0.046)
(HAZ (t-1)) x (Quartile 2, Wealth Index)		-0.035		0.074
		(0.050)		(0.057)
(HAZ (t-1)) x (Quartile 3, Wealth Index)		-0.069		0.039
		(0.043)		(0.056)
(HAZ (t-1))x (Quartile 4, Wealth Index)		-0.101**		-0.0090
		(0.049)		(0.063)
Quartile 2, Weath Index		0.016		0.075
		(0.119)		(0.143)
Quartile 3, Weath Index		0.031		0.119
		(0.101)		(0.133)
Quartile 4, Weath Index		0.030		-0.074
		(0.126)		(0.150)
Sex, Female	0.150	0.157	0.177*	0.188*
	(0.118)	(0.116)	(0.095)	(0.096)
HH Head Sex, Female	-0.125	-0.123	0.086	0.077
	(0.158)	(0.156)	(0.094)	(0.093)
HH Wealth Index	0.987**		0.0210	
	(0.452)		(0.416)	
HH Size	0.034	0.035	-0.036	-0.041
	(0.039)	(0.038)	(0.033)	(0.034)
Nr HH Adults	-0.050	-0.044	0.062	0.067
	(0.051)	(0.049)	(0.044)	(0.044)
Nr Male HH members	0.084	0.088	0.085	0.090
	(0.088)	(0.088)	(0.067)	(0.067)
Nr Brothers	-0.068	-0.071	-0.064	-0.068
	(0.087)	(0.085)	(0.056)	(0.056)
Age of Mother	-0.005	-0.005	0.004	0.004
	(0.007)	(0.007)	(0.006)	(0.006)
Maternal Height, in CM	0.031***	0.031***	0.009***	0.009***
	(0.006)	(0.006)	(0.004)	(0.004)
Caregiver can read (with difficulty)	0.221**	0.205*	-0.143*	-0.145*
	(0.107)	(0.107)	(0.083)	(0.083)
Caregiver can read (easily)	0.201**	0.231**	0.0228	0.0415
	(0.093)	(0.092)	(0.069)	(0.069)
Any Schooling?, HH Head	-0.039	-0.068	-0.155*	-0.161
	(0.088)	(0.085)	(0.081)	(0.079)
Constant	-5.714***	-5.488***	-2.012***	-2.061***
	(1.249)	(1.243)	(0.693)	(0.691)
Full Controls	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes
R-Squared	0.384	0.386	0.574	0.577
Nr Observations	913	913	903	903

Notes: * p<0.10, ** p<0.05, *** p<0.01. ^ For each model, HAZ (t-1) refers to the height-for-age z-score of the child in the previous period. All estimates are Ordinary Least Squares with community fixed effects. Full controls include further information on household characteristics, as well as dummies for age in months, ethnic group and birth order. Stunting models are estimated using a Linear Probability Model. Standard errors are clustered at the community and quarter of birth level.

Table A2: First-Stage IV Model and Determinants of IV, Age 4-5 years

	First-Stage Model	*Determinants of IV*
	Dependent variable: HAZ, Age 4-5	*Dependent variable: Season of Birth*
	(1)	(2)
Season of Birth, 2 and 4 Quarter	-0.4212***	
	(0.151)	
Sex, Female	0.201	-0.032
	(0.183)	(0.032)
HH Head Sex, Female	0.002	0.034
	(0.313)	(0.041)
HH Wealth Index	2.2711***	-0.027
	(0.845)	(0.122)
HH Size	0.056	0.0316**
	(0.106)	(0.013)
Nr HH Adults	-0.095	-0.0348**
	(0.132)	(0.016)
Nr Male HH members	-0.129	-0.022
	(0.163)	(0.025)
Nr Brothers	0.2789*	0.030
	(0.160)	(0.023)
Age of Mother	0.009	-0.0056*
	(0.014)	(0.003)
Maternal Height, in CM	0.0350***	-0.001
	(0.010)	(0.001)
Caregiver can read (with difficulty)	0.3830*	0.000
	(0.199)	(0.031)
Caregiver can read (easily)	-0.111	-0.047
	(0.190)	(0.035)
Any Schooling?, HH Head	-0.344	0.009
	(0.212)	(0.027)
Constant	-6.1035***	1.5842***
	(1.902)	(0.301)
Full Controls	Yes	Yes
Community Fixed Effects	Yes	Yes
R-Squared	0.291	0.628
Nr Observations	913	913

Notes: * p<0.10, ** p<0.05, *** p<0.01. All estimates include full controls and community fixed-effects. Full controls include information on household characteristics (including caregiver gender, height, health and demographics), mother and child information, as well as dummies for age by month, ethnic group and birth order. Instrumental variable 'season' takes a value of one for children born in the second and fourth quarter. Column (1) reports the results from the first-stage IV regressions. Column (2) explores the correlates of the instrument. Standard errors are clustered at the community and quarter of birth level.

Table A3: Examining catch-up growth, OLS Estimates detail

OLS Estimates	*Dependent variable: HAZ, Age 4-5 yr*			*Dependent variable: Stunted (HAZ<-2.0), Age 4-5 yr*		
	Full Sample	*Bottom Half Wealth Index*	*Bottom Half Maternal Height*	*Full Sample*	*Bottom Half Wealth Index*	*Bottom Half Maternal Height*
	(1)	(2)	(3)	(4)	(5)	(6)
Height-for-Age, Age 0-1 years	0.2386***	0.2703***	0.2235***			
	(0.020)	(0.031)	(0.026)			
Stunted (HAZ<-2.0), Age 0-1 years				0.3264***	0.3633***	0.3391***
				(0.034)	(0.047)	(0.044)
Sex, Female	0.150	-0.025	0.051	-0.057	0.012	-0.017
	(0.118)	(0.129)	(0.137)	(0.057)	(0.070)	(0.075)
HH Head Sex, Female	-0.125	-0.129	-0.255	0.039	0.021	0.080
	(0.158)	(0.246)	(0.197)	(0.073)	(0.129)	(0.095)
HH Wealth Index	0.9874**	5.316	0.150	-0.158	-0.399	0.014
	(0.452)	(3.248)	(0.641)	(0.174)	(1.513)	(0.327)
HH Size	0.034	0.077	0.012	-0.0437*	-0.027	-0.038
	(0.039)	(0.047)	(0.050)	(0.023)	(0.034)	(0.029)
Nr HH Adults	-0.050	-0.005	0.006	0.0760***	0.023	0.065
	(0.051)	(0.076)	(0.074)	(0.028)	(0.042)	(0.042)
Nr Male HH members	0.084	-0.090	-0.025	-0.015	0.009	0.036
	(0.088)	(0.114)	(0.107)	(0.045)	(0.057)	(0.057)
Nr Brothers	-0.068	0.090	0.066	-0.016	-0.040	-0.085
	(0.087)	(0.113)	(0.101)	(0.046)	(0.059)	(0.056)
Age of Mother	-0.005	-0.015	-0.002	0.004	0.004	0.006
	(0.007)	(0.013)	(0.010)	(0.004)	(0.007)	(0.006)
Maternal Height, in CM	0.0314***	0.0298***	0.0320***	-0.0102***	-0.0135***	-0.002
	(0.006)	(0.008)	(0.011)	(0.003)	(0.004)	(0.006)
Caregiver can read (with difficulty)	0.2214**	0.2401*	0.158	-0.047	-0.1230*	-0.095
	(0.107)	(0.143)	(0.110)	(0.045)	(0.069)	(0.059)
Caregiver can read (easily)	0.2006**	0.032	0.3766***	-0.1079***	-0.060	-0.084
	(0.093)	(0.142)	(0.118)	(0.037)	(0.067)	(0.055)
Any Schooling?, HH Head	-0.039	0.112	0.044	0.036	-0.031	0.000
	(0.088)	(0.147)	(0.107)	(0.031)	(0.060)	(0.045)
Constant	-5.7144***	-4.6167***	-5.3939***	1.6384***	1.7615**	0.029
	(1.249)	(1.595)	(1.847)	(0.571)	(0.884)	(0.965)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.384	0.344	0.340	0.245	0.242	0.284
Nr Observations	913	457	457	913	457	457

Notes: * p<0.10, ** p<0.05, *** p<0.01. All estimates include full controls and community fixed-effects. Full controls include information on household characteristics (including caregiver gender, height, health and demographics), mother and child information, as well as dummies for age by month, ethnic group and birth order. Standard errors are clustered at the community and quarter of birth level.

Table A4: Examining catch-up growth, IV Fuller detail

IV Fuller Estimates	Dependent variable: HAZ, Age 4-5 yr			Dependent variable: Stunted (HAZ<-2.0), Age 4-5 yr		
	Full Sample	Bottom Half Wealth Index	Bottom Half Maternal Height	Full Sample	Bottom Half Wealth Index	Bottom Half Maternal Height
	(1)	(2)	(3)	(4)	(5)	(6)
Height-for-Age, Age 0-1 years	0.250265**	0.418738***	0.368461***			
	(0.122)	(0.149)	(0.125)			
Stunted (HAZ<-2.0), Age 0-1 years				0.676446**	0.910998***	0.906345***
				(0.277)	(0.315)	(0.271)
Sex, Female	0.147	-0.079	0.048	-0.053	0.038	0.013
	(0.119)	(0.133)	(0.136)	(0.063)	(0.081)	(0.098)
HH Head Sex, Female	-0.125	-0.073	-0.247	0.016	-0.036	0.029
	(0.160)	(0.242)	(0.211)	(0.073)	(0.117)	(0.095)
HH Wealth Index	0.960908*	5.183	-0.345	0.078	0.083	0.676
	(0.542)	(3.283)	(0.817)	(0.291)	(1.534)	(0.509)
HH Size	0.033	0.090484*	0.009	-0.035	-0.023	-0.027
	(0.041)	(0.050)	(0.052)	(0.025)	(0.031)	(0.030)
Nr HH Adults	-0.049	0.006	0.031	0.077695**	0.012	0.053
	(0.053)	(0.077)	(0.090)	(0.030)	(0.046)	(0.048)
Nr Male HH members	0.085	-0.107	0.005	-0.036	-0.012	0.016
	(0.090)	(0.114)	(0.109)	(0.044)	(0.063)	(0.059)
Nr Brothers	-0.071	0.089	0.002	0.013	-0.017	-0.040
	(0.095)	(0.107)	(0.116)	(0.045)	(0.060)	(0.058)
Age of Mother	-0.005	-0.017	-0.004	0.004	0.006	0.006
	(0.007)	(0.013)	(0.010)	(0.004)	(0.008)	(0.007)
Maternal Height, in CM	0.031030***	0.023342***	0.033467***	-0.006564*	-0.006	0.004
	(0.006)	(0.008)	(0.013)	(0.003)	(0.005)	(0.006)
Caregiver can read (with difficulty)	0.216940*	0.137	0.028	-0.009	-0.057	0.037
	(0.112)	(0.183)	(0.145)	(0.064)	(0.097)	(0.097)
Caregiver can read (easily)	0.201697**	0.074	0.412453***	-0.102349**	-0.076	-0.045
	(0.093)	(0.155)	(0.143)	(0.042)	(0.087)	(0.077)
Any Schooling?, HH Head	-0.035	0.109	0.085	0.027	0.011	0.001
	(0.098)	(0.158)	(0.115)	(0.030)	(0.087)	(0.055)
Constant	-5.6e+00***	-3.200	-5.4e+00***	0.964	0.047	-1.200
	(1.353)	(2.020)	(2.016)	(0.746)	(1.385)	(1.176)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.383	0.295	0.287	0.153	0.030	0.059
Nr Observations	913	457	457	913	457	457
First-Stage: IV F-Statistic	7.1581	10.4961	7.0319	5.2531	8.3041	4.0439
Instrument	season	season	season	season	season	season

Notes: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. All estimates include full controls and community fixed-effects. Full controls include information on household characteristics (including caregiver gender, height, health and demographics), mother and child information, as well as dummies for age by month, ethnic group and birth order. Instrumental variable 'season' takes a value of one for children born in the second and fourth quarter. IV estimates obtained using weak-IV robust k-class Fuller estimator. First-stage F-Statistic reports results for Kleibergen-Paap test of weak identification. Standard errors are clustered at the community and quarter of birth level.

Table A5: Catch-up growth by Gender - OLS Estimates

OLS Estimates	*Dependent Variables: Anthropometrics (t), Age 4-5 yr*			*Dependent Variables: Anthropometrics (t), Age 7-8 yr*		
	Full Sample	*Male*	*Female*	*Full Sample*	*Male*	*Female*
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Height-for-Age (t)						
Height-for-Age (t-1), 0-1 years	0.2945***	0.3185***	0.2706***			
	(0.032)	(0.031)	(0.051)			
Height-for-Age (t-1), 4-5 years				0.6674***	0.6822***	0.6866***
				(0.046)	(0.070)	(0.059)
(HAZ (t-1)) x (Quartile 2, Wealth Index)	-0.035	-0.061	0.005	0.074	0.086	0.033
	(0.050)	(0.063)	(0.069)	(0.057)	(0.083)	(0.082)
(HAZ (t-1)) x (Quartile 3, Wealth Index)	-0.069	-0.043	-0.1099*	0.039	-0.006	0.073
	(0.043)	(0.042)	(0.063)	(0.056)	(0.081)	(0.089)
(HAZ (t-1))x (Quartile 4, Wealth Index)	-0.1010**	-0.037	-0.1688**	-0.009	-0.054	0.084
	(0.049)	(0.048)	(0.070)	(0.063)	(0.093)	(0.082)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.200	0.235	0.174	0.487	0.463	0.539
Nr Observations	913	483	430	903	478	425
Dependent Variable: Stunted (t) (HAZ<-2.0)						
Stunted (t-1) (HAZ<-2.0), 0-1 years	0.3634***	0.3264***	0.3598***			
	(0.052)	(0.074)	(0.080)			
Stunted (t-1) (HAZ<-2.0), 4-5 years				0.4705***	0.3771***	0.5430***
				(0.047)	(0.072)	(0.074)
(Stunted (t-1)) x (Quartile 2, Wealth Index)	0.028	0.092	0.025	0.023	0.142	-0.068
	(0.085)	(0.101)	(0.156)	(0.064)	(0.101)	(0.097)
(Stunted (t-1)) x (Quartile 3, Wealth Index)	-0.026	-0.013	-0.001	0.029	0.126	0.063
	(0.090)	(0.120)	(0.130)	(0.065)	(0.106)	(0.120)
(Stunted (t-1)) x (Quartile 4, Wealth Index)	-0.1725**	-0.131	-0.144	0.005	0.166	-0.140
	(0.085)	(0.128)	(0.136)	(0.078)	(0.119)	(0.094)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.106	0.100	0.126	0.283	0.273	0.324
Nr Observations	913	483	430	903	478	425

Notes: * p<0.10, ** p<0.05, *** p<0.01. For each model, HAZ (t-1) refers to the height-for-age z-score of the child in the previous period. All estimates are Ordinary Least Squares with community fixed effects. Full controls include information on household characteristics (including caregiver gender, height, health, household wealth and demographics), mother and child information, as well as dummies for age in months, ethnic group and birth order. The lower panel of the table shows the interaction of the lagged HAZ with the wealth quartile (bottom quartile omitted) measured in round one. When interacting by wealth quartiles household wealth is excluded from the core controls. Standard errors are clustered at the community and quarter of birth level.

Chapter 4

Income Dynamics in Rural India: Testing for Poverty Traps and Multiple Equilibria

Chapter 4

Income Dynamics in Rural India: Testing for Poverty Traps and Multiple Equilibria

Abstract

Do serious climatic shocks lead to processes of persistent poverty and poverty traps? We have access to a panel data set spread across 30 years, building on the ICRISAT data on six villages in the semi arid tropics to investigate this question. We identify the dynamic income process showing the transition dynamics in response to rainfall shocks, using a fixed effects dynamic model allowing for multiple equilibria. We show that there is serious persistence and evidence of multiple equilibria in the data: in the data period analysed, many households were initially in a precarious (unstable) equilibrium that could lead to a downward cycle into destitution, but also take-off if appropriate circumstances presented themselves. Many managed to escape in the period 1984-2004 towards higher and stable equilibria, leading to considerably better living conditions. For the median income houseold, the higher equilibrium is at roughly 210 US dollars per year per adult at 1975 prices, and the unstable equilibrium below which a downward spiral would emerge is about 55 US dollars. By investigating the fixed effects, we find that those with higher assets, especially in the form of initial levels of education in the family in the 1970s, higher land holdings and/or high physical capital were faced with a much lower level of income at which a downward spiral could have followed. Those with few assets in these different forms could experience the downward spiral at much higher levels of income: their livelihood was far more precarious. For them, climatic shocks, even at reasonable levels of incomes in preceding years could lead to destitution.

Methodologically, we use Lokshin and Ravallion's estimation method but using rainfall as instruments rather than black box dynamic identification methods as in Arrellano-Bond estimators. The instruments are reasonably strong, and the link between climate shocks and destitution appears to be very strong. We find that rainfall-induced lower income does not only have a simple contemporaneous effect, as would be the case if rainfall caused the error in an income process that would be described as independently distributed errors. Instead, we find that rainfall-induced income levels have a persistent impact, possibly causing destitution.

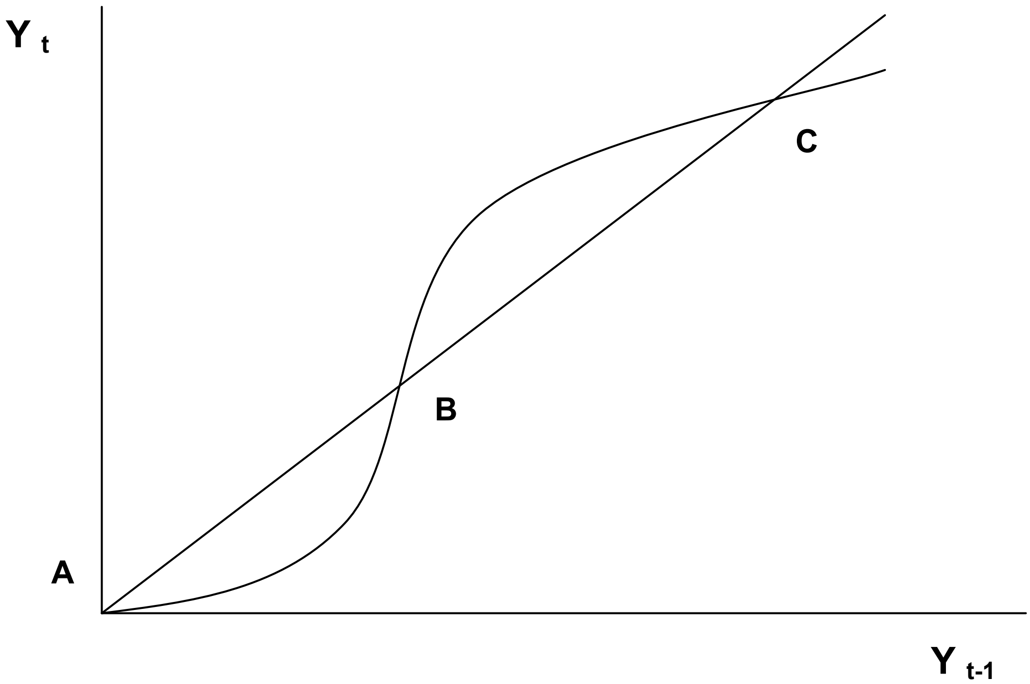
4.1 Introduction

With about 2.5 billion people receiving less than two dollars per day in income (Chen and Ravallion, 2008) the issue of whether and why the destitute escape poverty constitutes a central question in economic research. Theories of poverty traps explain why living in poverty at some time causes a person to remain poor in the future, or why a country's poverty causes the country to remain in future poverty. These theories imply stark conclusions: a positive income shock could prevent a person from living in poverty for the indefinite future, while a sufficiently grave negative shock to income could prevent a person from ever escaping poverty. Some such theories assume that a person requires a fixed and indivisible investment to purchase a good, like education or credit (Banerjee and Newman, 1993, and Galor and Zeira, 1993); others assume increasing returns to income via nutrition or another means (Dasgupta and Ray, 1993); still others show how leaving the poor without bargaining power can cause the poor not to save (Mookherjee and Ray, 2002) or how poverty itself can stifle individual aspirations (Ray, 2006) and investments in their children (Mookherjee, Napel and Ray, 2010). In this paper, we offer a test for the existence of poverty traps using a long-term panel from rural India.

Admittedly imperfect tests of these elegant models have offered little empirical support, leaving Dasgupta (1997) to describe that they reside 'awkwardly' in development thinking. A model of nutrition poverty traps has received empirical criticism from several studies (Bliss and Stern, 1982; Swamy, 1997; Rosenzweig, 1988), though Dasgupta (1997) argues that they use flawed tests. A theory of fixed costs to entering businesses has received similarly little support (McKenzie and Woodruff, 2006).

Several recent studies have proposed that a poverty trap could arise through a combination of mechanisms, or through some unstudied mechanism. These studies essentially examine whether a regression of some welfare measure – income, consumption, or assets – on its lag has a shape that could indicate the presence of a poverty trap, as described in Figure 1, linking the welfare measure Y at t and $t-1$ in some non-linear way. A 45 degree line is drawn in to identify equilibrium points, i.e. where Y_t equals Y_{t-1}. As is well-known, the shape shown offers multiple equilibria, whereby A and C are stable (low and high) equilibria, while B constitutes an unstable state; as once removed from B, a household would drift towards A or C according to the dynamic relationship shown in figure 1, below.

Figure 1: Poverty trap and multiple equilibria



Nonparametric kernel regressions of current on lagged assets using small samples from Kenya, Ethiopia, Madagascar, and South Africa show unstable equilibria over some low values of income that suggest the presence of a possible poverty trap (Adato et al., 2006;

Barrett et al., 2006; Lybbert et al., 2004, Naschold (2009) and Quisumbing and Baulch, 2009). However, these studies ignore the endogeneity of lagged income and assets in dynamic panel models. Furthermore, the potential bias in data obtained from many-year recall questions, limited generalizability of sample sizes under 200 individuals, and bias of bivariate kernel regressions at discontinuities (Fan, 1992), give their conclusions limited scope. In higher income areas, studies applying methods with corrections for various econometric challenges in estimating income dynamics to data from Eastern Europe and Urban Mexico have found evidence for some stable low-level equilibria but no evidence of a poverty trap; evidence from China shows similar findings (Antman and McKenzie, 2007, Jalan and Ravallion, 2003, and Lokshin and Ravallion, 2004).

The econometric challenges involved in testing for the presence of poverty traps are not trivial, and most create bias towards failing to reject the hypothesis that poverty does not entrap people. Hence one could reasonably conclude that the existing literature fails to establish whether poverty traps actually do not exist or whether available data and methods have inadequate power to detect them as there are econometric problems abound. Panel data with short duration – typically less than five years (Dercon and Shapiro, 2007) – may not capture the dynamics that ensure poverty's persistence. The nature of a dynamic panel model ensures that regression of income on its one-period lagged value will inflate the effect of lagged income on current income. Measurement error in income creates a mirage of income mobility, so a person whose true income remains constant over time may appear to first enter, and then escape, poverty. Similarly the stochastic nature of income might result in equally high levels of mobility (Carter and Barrett, 2006).

Existing studies address some, but not all of these concerns. Jalan and Ravallion (2003) and Lokshin and Ravallion (2004) apply Arellano-Bond GMM (Arellano and Bond, 1991) estimators to identify the association of a cubic polynomial of lagged income with current income. But if measurement error has serial correlation, as at least one U.S. comparison of survey-reported income with independent income reports suggests (Bound and Krueger, 1991), then using distant lags of income as instruments for once-lagged income, as Arellano-Bond methods do, will overstate mobility. Antman and McKenzie (2007) for this and other reasons condemn the possibility of using panels for identifying nonlinear income dynamics, and propose instead the use of pseudo-panels to average out measurement error across individuals.

The present study shows how panel methods can address these econometric criticisms and consistently test for the presence of a poverty trap. We test whether a poverty trap characterizes the income dynamics of individuals in an unusually long 30-year panel from six villages in India's semi-arid tropics, based on a recent extension of the ICRISAT Village Level Studies. We follow Lokshin and Ravallion (2004) and others in estimating a dynamic equation where income is modelled as a cubic polynomial of lagged income that allows for unobserved income heterogeneity. However, unlike these studies, we use exogenous instruments to correct for endogeneity problems. We interact rainfall shocks with household characteristics to provide valid and informative instruments for a polynomial function of lagged income, obviating the need for Arellano-Bond methods and addressing the critical problem of measurement error in income.

Further, the first-differences econometric specification allows for household heterogeneity in the income generating process. We can retrieve these individual effects and explore their correlates with starting period household characteristics, therefore uncovering those household assets that could have led to sustained increases in the trajectory of incomes.

The downside of using long panels is that survey attrition might be high. Indeed, of the original 1998 individuals found in the ICRISAT sample in 1975-1984, only 654 (or 33%) where included in the 2001 survey. Ignoring the possibility that attriting individuals may have substantially different characteristics from individuals remaining in the panel may overstate or understate mobility. We correct for attrition in observables applying inverse probability weighting (IPW) methods.

What may look like an econometric solution to statistical problems, the method we use – the estimation by IV methods of a lagged polynomial of income using rainfall shocks as genuinely exogenous instruments – also has a clear conceptual meaning. Theoretical models of poverty traps typically imply that only a 'shock' can move people between equilibria, as shown in Figure 1. In semi-arid India, from where the data are derived, the key is rainfall. Using rainfall as an instrument, we aim to make a direct causal link between rainfall as a cause of lower or higher income, affecting whether the household experiences a shock high enough to move between equilibria and the speed by which it moves to a new equilibrium.

Furthermore, the household fixed effects will allow households to have different underlying equilibria, reflecting for example different assets and human capital levels, offering a further interpretation of the meaning of the precariousness and potential vulnerability in their livelihoods.

Our empirical analysis shows that the income generating dynamics in rural India follow a quadratic polynomial function with pronounced concavity. We find that the uncovered income dynamics remain robust to a series of robustness checks. Firstly, although precision is reduced, results remain unchanged when we correct for non-random attrition. Secondly, we find that the uncovered dynamics are stable between the two panel periods of our sample, 1975-1983 and 2001-2004. Thirdly, while the presence of ‘weak’ IVs in our baseline results can not be ruled out, we find that a series of robustness checks designed to address problems of ‘weak IV’ finite-sample bias have no substantive effect on our estimates. Moreover, when estimating the dynamic model on the shorter 1975-1983 panel, we find instruments to be strong by standard measures (Stock and Yogo, 2005), while still predicting a quadratic polynomial function with substantial concavity.

Income simulations based on the estimated parameters suggest the presence of two equilibria: a stable high-income equilibrium and a low-level unstable saddle point. While households with sufficiently high fixed effect income follow the high-equilibrium dynamic path, almost half of our sample has too low an income steady state to overcome the dynamic point of divergence. Unpacking the household-specific fixed effects, we find that household assets in the mid-1970s, – essentially education, land holdings and physical assets – are positively associated with higher levels of steady state income.

Our analysis suggests that over the past 30 years many households have managed to escape towards higher and stable equilibria leading to considerably better living conditions. However, those with few assets to start with could experience the downward spiral at much higher levels of income: their livelihood was far more precarious. For them, climatic shocks, even at reasonable levels of incomes in preceding years, could lead to destitution.

The paper proceeds as follows. Section 2 outlines the econometric obstacles inherent in testing those models of poverty traps which only consider income dynamics. Section 3 describes the 30-year panel data set. Section 4 presents the main results, and section 5 concludes.

4.2 Econometric Methodology

The consistent estimation of a dynamic model of income is a challenging exercise and requires addressing a number of statistical problems: the endogeneity of lagged income in a dynamic model; measurement error in income; unobserved individual heterogeneity as well as non-random attrition. We discuss each in turn, and the econometric methodologies we propose to solve them.

4.2.1 Dynamic panels and measurement error

We estimate an AR(1) model where the income y_{it} of person i at time t depends on a linear function of a polynomial of person i's lagged income,[1] and a composite error term with time-invariant and idiosyncratic components ρ_i and v_{it}:

$$y_{it} = \beta g(y_{i,t-1}) + \rho_i + v_{it} \tag{1}$$

$$y_{it} = \beta_1(y_{i,t-1}) + \beta_2(y_{i,t-1})^2 + \beta_3(y_{i,t-1})^3 + \rho_i + v_{it} \tag{1'}$$

Where the $g(.)$-function represents the polynomial function which in equation (1') is assumed to constitute a cubic polynomial function. Since lagged income correlates positively with the composite error term $\rho_i + v_{it}$, estimating equations (1) or (1') by OLS generate inconsistent estimates of β. First-differencing the equations eliminates the individual unobserved effect ρ_i:

[1] Some studies describe such an estimate as a test of ``non-linear income dynamics'' (Antman and McKenzie, 2007; Jalan and Ravallion, 2003). While specifying lagged income as a higher-order polynomial allows current income to vary nonlinearly with lagged income, the regression function itself is linear in the higher-order terms of lagged income.

$$\Delta y_{it} = \beta \Delta g(y_{i,t-1}) + \Delta v_{it} \tag{2}$$

$$\Delta y_{it} = \beta_1 \Delta(y_{i,t-1}) + \beta_2 \Delta(y_{i,t-1})^2 + \beta_3 \Delta(y_{i,t-1})^3 + \Delta v_{it} \tag{2'}$$

where $\Delta y_t = y_t - y_{t-1}$. However, the dynamic nature of the income generating process implies that OLS estimation of β in equation (2) will still produce inconsistent estimates, resulting from lagged errors ($v_{i,t-1}$) affecting both Δv_{it} and $\Delta y_{i,t-1}$ (Nickell, 1981).

Arellano-Bond IV General Method of Moments (GMM) estimators use lags of the endogeneous variable as instruments for $\Delta y_{i,t-1}$. If the error term Δv_{it} in equation (2) lacks second-order serial correlation, and if equation (1) is dynamically complete, then the second lag and any further lags of income can serve valid instruments of $\Delta y_{i,t-1}$ (see Arellano and Bond (1991) for the Arellano-Bond *dif* version of the estimator). Under these assumptions, the general method of moments provides consistent and efficient estimates of β, the parameter of interest (Anderson and Hsiao, 1982; Arellano and Bond, 1991; Blundell et al. 2000; Bond, 2002).

Several existing studies use Arellano-Bond type estimators with short panels to test whether a poverty trap characterizes income. One of these studies shows overidentification tests of the instruments' validity and a test that the residuals have second-order autocorrelation (Jalan and Ravallion, 2003); the other studies do not mention these tests (Lokshin and Ravallion, 2004 and Antman and McKenzie, 2007) and none of the three studies evaluate the weakness of the instruments, an important problem when a dynamic panel has a near-unit root (Stock et al., 2002; Blundell et al, 2000), and a problem which large samples do not eliminate (Bound et al., 1995).

Measurement error also creates a more substantial problem in these papers, since some U.S. data suggest that measurement error in an individual's income has positive autocorrelation across waves of a panel (Bound and Krueger, 1991). Antman and McKenzie (2005) show that in the presence of such measurement error, GMM estimators provide inconsistent estimates of the parameters in equation (2'), and that the severity of the problem persists upon specifying lagged income as a higher-order polynomial.

When applying Arellano-Bond methods, measurement error, the stochastic nature of income, as well as poor instrument validity of higher-lags of income will all conspire in creating the mirage of an income generating process with high levels of mobility. In this paper, instead, we propose the use of exogenous instruments in the form of rainfall shocks to estimate the parameters of a cubic polynomial as defined in equation (2').

Rainfall provides a useful instrument for income in poor agricultural areas.[2] Since the economies of agricultural villages heavily depend on weather, flood and drought sharply affect the income of most households in the village: agricultural households have lower yield in seasons of extreme weather, households that earn income from agricultural labour find less work in times of extreme weather, and most individuals in these communities depend on good weather for strong income.

However, while rainfall shocks are common to all households in a village, we are interested in modelling household-specific dynamics. To capture the heterogeneous effects of the shocks, in our instrument set, we interact rainfall shocks with two household-specific variables: access to land and household demographics.

[2] See, for example, Paxson (1992), Miguel (2005), and the review in Rosenzweig and Wolpin (2000).

Positive and negative rainfall shocks might have different effects across households with varying reliance on agricultural income. The benefits to a particularly good rainfall year are likely to rise with area of cultivation. At the extreme, landless workers might accrue little (indirectly perhaps through higher wages) to no benefits. On the other hand, in economies of household production with imperfect labour markets such as rural India, labour-scarce households will not be able to make the most of a particularly good rainy season.

Our three instruments of choice are defined as follows. First, we use the signed square of rainfall shocks, whereby the shock is defined as the deviation in a year's precipitation relative to its historic mean. For this we use monthly village-level rainfall measures from several Mandal-level collection stations. The higher moment of rainfall deviations is used to increase the power of the instrument when predicting the second and third moments of lagged income.

We interact the rainfall shock variable with the land area (in hectares) operated by the household in that year. Size of land usage makes intuitive sense as an instrument, since households with larger plots will receive more benefit from years with good rainfall and more harm from years with drought. We use our third instrument, the interaction between rainfall shocks and the number of children aged 0-8, as a measure of the labour-scarcity of the household. Households with more inactive members too young to contribute and requiring care will be less flexible to insure their income against rainfall shocks or to profit from a positive shock.

For each rainfall instrument $Z_{i,t-1}$ and endogenous variable, we require two conditions to be met, which for the case of the first moment of lagged income can be written as:

$$cov(Z_{i,t-1}, \Delta y_{i,t-1}) \neq 0 \tag{3}$$

$$cov(Z_{i,t-1}, \Delta v_{i,t}) = 0 \tag{4}$$

Condition (3) requires that the instrument strongly correlates with lagged income – the 'strong' IV condition –, while condition (4) requires the instrument to be orthogonal to measurement error in lagged income or with other unobserved components of the structural equation – the so-called 'validity' condition. One could interpret our critique of Arellano-Bond estimates of equation (2) as violations of the latter condition.

While we expect our rainfall instruments to be 'valid' in the sense that lagged rainfall shocks are likely to be orthogonal to current household income, the set of interactions might be suspect to capture unobserved household heterogeneity. With that in mind, we avoid the worse effects of such endogeneity by including only the interaction term between the household indicator and the rainfall shock.[3]

Further, as a choice for rainfall interactions, land and household demographics are well suited because they remain relatively stable over time. The inclusion of liquid assets in the instruments sets might create a dynamic bias that would render our instruments invalid.[4] Household demographics are unlikely to be flexible enough to respond effectively to shocks. In rural India, access to land remains a strong determinant of household wealth and other welfare related variables. Other studies on ICRISAT villages have used land as a good indicator of physical wealth and one that might be relatively stable over time (see Townsend (1994) or Ligon (1998)).

---

[3] Nevertheless, it is not possible to rule out entirely the possibility that these interactions might actually capture unobserved time-varying heterogeneity across households. To test this empirically we perform a type of placebo test. We test whether out core results are driven by the unobserved effects of the household indicators. We carry out a placebo regression whereby rainfall instruments are replaced by land in operation and number of children in the household. As we expected, this model yields very different results from our base model. We interpret this as evidence that our instrument set is not being driven by any unobserved heterogenous effects correlated with land area or child demographics. (Results not reported).

[4] For example, more liquid assets such as livestock might be affected by past shocks. Including such an asset as an interaction for our instrument set might create a dynamic bias between past shocks and current income.

The concern over the dynamic properties of our instrument set is a serious one in our set-up. In particular, shocks might affect the accumulation of assets and productive investment that might affect future income. If this was the case, the AR(1) model assumed in equation (1) might be misspecified. We return to this point in the Annex, where we test the dynamic properties of the income series and explore whether the dynamic impact of shocks might generate an income process closer to an AR(2) model.

Violations of assumption (3) might also pose a challenge to our results. Problems arising from 'weak' IVs are particularly severe when one considers the finite-sample properties of IV estimators. Point estimates are rendered biased and inconsistent, while standard errors are invalid (see Staiger and Stock, 1997; Hahn and Hausman, 2005; and Murray, 2006a, among others). As a robustness check on the IV GMM estimates, we apply 'weak' IV-robust estimators. Fuller k-class estimators and limited information maximum likelihood (LIML) estimators are understood to perform better under 'weak' IVs.[5] Furthermore, in a world of weak IVs, the potential IV bias can be reduced when using a parsimonious set of instruments (see Stock and Yogo, 2005). Accordingly, we re-estimate our preferred model specification with a reduced set of instruments.

4.2.2 Individual heterogeneity

Fixed individual factors – education, geographic location, and others – may affect the trajectory of an individual's income. Since these fixed factors may correlate with income and hence bias regression estimates, equation (2') uses first-differencing to eliminate these fixed

[5] Both LIML and Fuller methods are examples of k-class methods which are asymptotically equivalent to the GMM estimator. They differ from GMM in the weighting placed on instruments. Studies have shown that the Fuller k-class of methods dominates over other 'weak' IV-robust estimators. However, LIML methods are also often used for its nesting properties: when the model is exactly identified GMM and LIML are identical, and Fuller estimates are mean-square-error corrected versions of LIML. See Hahn, Hausman, and Kuersteiner (2003), Anderson, Kunitomo, and Matsushita (2005) and the excellent review article by Murray (2006b).

effects.

However, the effects themselves have economic interest and recovering these parameters allows us to observe the correlation between observable individual characteristics and the part of an individual's income trajectory which does not depend on short-term income dynamics. It will give us insight into any factors that may cause households to achieve higher levels of income. Furthermore, as it will be shown further below, the fixed effects will affect the location of the equilibrium, while at the same time allowing us to identify the type of households – in terms of their early assets – that have higher equilibria compared to others. Jalan and Ravallion (2005) and Antman and McKenzie (2005) estimate models with household fixed effects, but fail to examine further their correlates, even though they both highlight their relevance and how they shift a person’s trajectory.

Since the idiosyncratic errors have zero mean across the population, we estimate the individual effect by the deviation of an individual's mean outcome from the predicted mean (Antman and McKenzie, 2007):

$$\hat{\rho}_i = \overline{Y}_i - \hat{\beta}_1 \overline{Y}_{i,t-1} - \hat{\beta}_2 \overline{Y}^2_{i,t-1} - \hat{\beta}_3 \overline{Y}^3_{i,t-1}$$

where we average the dependent and independent variables across the years in which they would appear if we had not first-differenced the model. We then investigate the correlates of these fixed effects by regressing them on a vector Z_i of fixed individual characteristics:

$$\hat{\rho}_i = \phi_0 + Z_i \phi_1 + \varepsilon_i$$

The parameters ϕ_1 show the correlation of individual characteristics with the fixed effects. A positive association $\phi_j > 0$ for some element j of the vector Z_i implies that ϕ_j gives an individual permanently higher income regardless of shocks.

4.2.3 Panel Duration and Stability of Income Dynamics

The dataset we use has the advantage of an unusually long duration: 30 years, a length paralleled by only a small handful of existing datasets (Dercon and Shapiro, 2007). Unfortunately, the panel has a gap of about fifteen years. Households were surveyed yearly between 1975 and 1983, which we name the VLS1 panel, and again between 2001 and 2005 – the VLS2 panel.[6] In the effort to examine the long-run factors that influence poverty and welfare, such long-term panel duration provides critical information on income dynamics. But given our focus on income dynamics, ignoring this gap in the middle and treating 1983 as if it preceded 2001 would yield problematic estimates for later years.

To address the 1984-2000 gap, we use one-year lags of variables for all years. Taking the first difference of an income model with one-period lags as a repressor, drops two years in each of the two period panels. Hence, in our model specification we use the first difference of current income from nine waves of the panel (1977, 1978, 1979, 1980, 1981, 1982, 1983, 2003, 2004), while we use the first difference of the independent variables (lagged income) from a different set of nine waves (1976, 1977, 1978, 1979, 1980, 1981, 1982, 2002, 2003). Since we use lagged rainfall as an instrument rather than the many lags of income used in Arellano-Bond estimators, the 1984-2001 gap creates no other obstacles in estimating the dynamic panel model.

The 15-year gap between the two annual survey panels raises a further concern. Our model

[6] Income data in year 1984 included only a small subset of individuals. A 1992 round of income data included few individuals

specification assumes that a single set of polynomial parameter values underlie the income generating process. Similarly, the model assumes that the household-specific fixed effects are stable over time. However, should the income generating process in rural India have changed sufficiently during the 15-year gap between VLS1 and VLS2, our model specification might be grossly mis-specified resulting in uninformative parameter estimates. We address this issue by assessing the parameter stability of the income polynomial function across the two panel periods.

4.2.4 Non-Random Attrition

Requiring a very long panel comes at least at one cost: attrition. Over a 30-year period, a considerable number of households were lost, partly due to the well-documented rules of tracking the ICRISAT panel (Foster and Rozenzweig, 2001), which in principle did not track anyone leaving the household. In the VLS2 rounds, split-offs that remained in the study villages were largely included, but migration out of the villages is not irrelevant, and constitutes the main cause of attrition (Badiani et al., 2007). Moreover, even among households interviewed throughout, not all individual members might be included in all surveys. Our sample of analysis is therefore a sample of individuals living in the study villages – throughout the length of the panel –, directly related to the original ICRISAT households included in the first wave of the study in 1975. This might affect external validity, but as long as coefficients are interpreted exactly as relevant for this sample, it is not problematic, and obviously, given the relative low mobility in rural India (Munshi and Rosenzweig, 2009), this is not an irrelevant sub-population.

If attrition randomly removed observations from each wave of a survey, then attrition would

and had different methodology than other years, while 2005 data are still being processed. See the following section for more only decrease the precision of estimated parameters. But attrition may occur for non-random reasons: individuals leave villages due to fixed and time-variant characteristics like shocks and job opportunities that cause a person or household to move. Since attrition may correlate with observed and unobserved characteristics which influence income, estimating equation (2) by any method without addressing attrition can produce an inconsistent estimate of β.

The problem has similarities to selection models where an econometrician observes a response variable for only a subset of a cross-sectional survey, and indeed the first selection models explicitly discussed their potential for addressing panel attrition (Heckman, 1979). Lokshin and Ravallion, 2004, for example, simultaneously estimate a Arellano-Bond GMM regression with an equation where baseline household composition, education, and location variables serve as instruments for selection in a regression of income on its lag.

But it is difficult to argue that these or any variables affect selection but not income, as one requires for consistent estimates of the parameters in equation (2). Other authors suggest more detailed procedures which estimate selection models for each time period, but these too require exclusion restrictions (see Wooldridge, 2002b, pp. 581-590). Any selection model requires observation of factors which vary across individuals, affect the probability of disappearing from the panel, and are independent of income. Fitzgerald, Gottschalk and Moffitt (1998) propose a cost-benefit model wherein individuals consider the net value of participating in a survey, and interview duration or interview payments affect the value of the survey. ICRISAT offers no such variation of interview payments across respondents. Furthermore, ICRISAT attrition occurs rarely due to refusal and more often due to migration and death. We conclude that no variables from available data can credibly satisfy the required exclusion restriction.

---

details on the data.

Weighted least squares (WLS), sometimes called inverse probability weighting, eliminates the need for exclusion restrictions, though WLS does require the model to satisfy nontrivial identification assumptions (Fitzgerald, Gottschalk and Moffitt, 1998, Wooldridge, 2002a, and Wooldridge, 2002b). In some cross-sectional surveys where individuals refuse to participate, surveyors construct weights to represent an individual's probability of participating in the survey, and inference using these surveys weights responses by the inverse of these probabilities.[7] WLS in the present context plays a similar role.

The attrition-corrected results use a random population sample at time $t = 1$ and define the selection variable s so an observation appears in a wave if and only if $s_{it} = 1$. We treat attrition as an absorbing state, in that an individual who attrits from the sample at time t does not reappear, so $s_{it} = 1 \rightarrow s_{rt} = 1 \quad \forall\, r < t$. Although such an approach forces us to drop observations that vanish for one or more rounds then reappear, this loss of precision and information allows us to use a potentially consistent estimator of regression parameters even in the face of substantial attrition.

For this correction to provide a consistent estimator, we must assume that a set of baseline covariates z_{i1} has enough predictive power that outcomes and covariates at any future time are independent of selection:

$$P(s_{it} = 1 \mid y_{it}, x_{it}, z_{i1}) = P(s_{it} = 1 \mid z_{i1}) \qquad (5)$$

Writers generally describe assumption (5) as selection on observables or ignorability of selection. To consistently estimate equation (2) while assuming selection on observables, for

[7]A separate reason for constructing and using weights arises when survey design, and not respondent refusal or absence, causes individuals to have unequal probabilities of appearing in the data.

each time period, we estimate a probit of s_{it} on z_{i1} using all observations that appear in the baseline survey. We obtain estimated probabilities $\hat{p}_{it}$ for each time period and individual, then weight the regression by the inverse of these fitted probabilities, equivalent to minimizing the following function:

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{s_{it}}{\hat{p}_{it}}\left(\Delta y_{it}-\alpha\Delta y_{i,t-1}\right)\right)^{2} \tag{6}$$

An analogy argument can show that, under assumption (5), equation (6) produces a consistent estimator which has a probability limit identical to an unweighted regression if the data had no attrition (Wooldridge, 2002a; and Wooldridge, 2002b) .

4.3 Data: the 30-year ICRISAT Panel

The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) near Hyderabad, India, collected annual surveys between 1975 and 1984 (VLS1), then for the same households in the period 2001-2005 (VLS2). The core data included 60 households each from six villages (240 in total in 1975) in India's semi-arid topics: the villages of Aurepalle and Dokur in the Mahbubnagar District of the Indian state of Andhra Pradesh; the villages of Shirapur and Kalman in the Sholapur District of the state of Maharashtra, and the villages of Kanzara and Kinkheda in the Akola District of Maharashtra. Villagers generally work in dryland farming, with limited irrigation (Badiani et al. 2008).

For the early data collection, interviewers lived in the villages and interviewed households every 3-4 weeks to obtain income information. The more recent data (VLS2) use one interview per year for 2001-2003 and two per year for 2004. Detailed checks on comparability between these years and with the VLS1 are reported in Badiani et al. (2008). A tracking survey allowed follow-up of individuals interviewed in the 1975-84 rounds. Additionally, the 2001-2005 study re-surveyed new households to compensate for the reduced sample sizes due to attrition. Walker and Ryan (1990) provide detailed description of the early survey rounds and research stemming from them, while Badiani et al. (2008) provide an appendix with further detail on the recent data collection. A key point to notice is that this paper provides an assessment of the comparability of different indicators, as the frequency of data collection is different in the VLS1 and VLS2. Badiani et al. (2008) show that trends in those variables that were collected with the same frequency in both surveys and those that were not showed remarkable similarity, suggesting that comparability may not be negatively affected.

We define our measure of annual income as net real income (excluding asset sales). This is a measure consistent with the income series analysed by Badiani et al. (2008) and includes income from crops, livestock income, transfers, income from trade, migration income and labour income. To allow comparison across households of different sizes and composition we re-weight the income series using adult equivalence scales to obtain a measure of income per adult equivalent. [8] Finally, the income series is deflated to rupees at 1975 prices.

Table (1) provides descriptive statistics of the general structure of the ICRISAT panel. As one would expect of a 30-year long panel, household attrition is substantial. Of the total 1998 individuals found in the VLS 1 sample, only 654 were included in the first year of the VLS 2 panel, amounting to a 77% attrition rate at the individual level. While part of this attrition was due to death, patterns of migration analysed in Badiani et al (2008) suggest that non-migrants

[8] Sensitivity analysis on the baseline model shows that changes in the adult equivalence scales have no substantive effect on our results. Similarly, our findings remain broadly unchanged when we model income dynamics with income per capita or

constitute a non-random sample of the original VLS1 sample.

With such high rates of attrition and the prospect of non-random migration, it is crucial to correct for attrition bias. To apply inverse probability weighting methods, we require the data to have a particular structure (Wooldridge, 2002a; and Wooldridge, 2002b). Firstly, while income is measured at the household level, we model attrition at the individual level. Accordingly, we structure our sample as a panel of individuals.[9] Secondly, we require attrition to be an absorbing state, that is, starting from all individuals present in the first period of the panel – year 1975 – we drop any individuals missing in at least one subsequent wave, excluding the 1984-2001 gap. This not only will exclude individuals that have permanently left the villages – because of death or migration – but also individuals that might have temporarily left the sample. Similarly, new additions post-1975 to VLS1 households, as well as members of the newly surveyed VLS2 households, are not included in the analysis. As a result of these data restrictions – and as shown in Table 2 –, our analysis sample is a balanced panel of individuals that starts with 1333 individuals in year 1975 and concludes with 325 in year 2004.

Finally, before moving on to discussing our results, we note that throughout the paper we apply standard error corrections for heteroskedasticity and autocorrelation clustering at the level of the household. Moreover, given the pattern of split-off households in the VLS2 panel, we define the clustering option at the level of the 'original' household – i.e. VLS1 households

simply income with no correction for household demographics. Other sensitivity tests performed show that excluding transfer income or using crop income only yield results consistent with the evidence presented here.

[9] For individuals that might have split off from the original household, we compute their income series as a composite incomes series across households and time, whereby their annual income is attributed on the basis of the matching between the individual and the household IDs in the annual rosters. Specifically, an individual found in a new household in the VLS2 sample, will be inputed , the incomes series of the new household during the VLS2 years, while during the earlier years it will take the income series of the original household.

– capturing any correlation in the error tem across family dynasties.[10]

4.4 Results

4.4.1 Income Trends

Table 2 reports descriptive statistics for the income measure – income per adult equivalent per year in rupees in 1975 prices – both at the individual and household level. The individual income series corresponds with the data used in our analysis. As discussed earlier, we treat attrition as an absorbing state, such that only individuals present in all waves of the panel since 1975 are included in the series. The household level income series includes VLS1 households and their later split-offs, and includes the income values that are matched to the individual income series. In that respect, the two series are equivalent and divergences are only due to differences in the weighting across households resulting from different household sizes.

During the early panel, income shows a clear upward trend, while income in the later panel, VLS2, is substantially higher, most notably due to the 2001 and 2002 high income years, a finding that coheres with the results of Badiani et al (2008). Graph 1 provides an even more striking representation of the income increases benefiting the households in the ICRISAT sample. The graph plots kernel densities of income for individual years 1975, 1983, 2001 and 2004. We observe that not only has mean income increased over this period, but the spread of the income distribution has also grown dramatically. As we will argue later, the evidence presented in the paper goes some way towards explaining the observed expansion in income

[10] Additionally, given the structure of our data, individuals belonging to the same household will have identical income values. The clustering function will also yield standard errors corrected for the repeated nature of the data.

inequality.

Before moving on to the discussion of the results of our parametric analysis, it is instructive to plot the raw income data. The econometric model assumes that the income generating process follows a polynomial function. Graph 2 uses locally weighted (Lowess) methods to obtain a non-parametric estimate of income lagged ($y_{i,t-1}$) on current income ($y_{i,t}$). The income patterns implied by the graph suggest substantial convergence and few non-linearities in the income generating process. In the next section, we assess whether this remains an accurate description of the true income generating process once issues of measurement error, income stochastic patterns and unobserved heterogeneity have been addressed.

4.4.2 Regression results

As discussed in section 3 above, our econometric methodology estimates the parameter of a polynomial function of lagged income in first-differenced form – equation (2') – using exogenous instruments. Tables 3 and 4 report IV GMM first and second-stage estimates, for three alternative functions: linear, quadratic and cubic polynomial. For each of these models, we estimate one specification with year dummies and another without.

Columns (1) and (2) in Table 3 show that our instruments significantly affect changes in income. When no year dummies are included, we find rainfall deviations significantly increase household income. The significance of the signed square of the rainfall deviations indicates a complex nonlinear relation between precipitation and income. Similarly, the interaction effects between rainfall shocks and household characteristics are significant and take signs as expected. We find that households operating larger plots and households with fewer kids among their members appear to benefit (lose) the most when rainfall is abundant (poor).

While estimates for interaction instruments remain largely unchanged when year dummies are included, rainfall shocks lose their significance. Given that the pattern of rainfall in the Indian sub-continent is mostly determined by the timing and profusion of the monsoon season, we would expect rainfall shocks in a given year to be highly correlated across villages. It is therefore not surprising to find rainfall shocks not to vary sufficiently across villages to be identified beyond the annual effect.

Recall that our first-difference specification removes any time-invariant unobservables, such as household heterogeneity and village fixed effects. However, time-varying effects such as time trends, or village specific trends are not controlled for. Our aim is to model household income dynamics, not so much to document how aggregate income or precipitation patterns change over time. It is in this context that year dummies are important in ensuring that our instrumentation strategy captures household specific processes. Their inclusion ensures that the income effect of rainfall shocks is exclusively identified by the variation in village precipitation, while at the same time controlling for time-varying trends common to all villages. The fact that the power of our instruments in explaining income dynamics is highest when including year dummies merely confirms this understanding. It is on this model specification that we focus throughout the paper.

Evidence in columns (3) to (6) in Table 3 shows that both the signed square of rainfall deviations as well as the interaction instruments are not only good predictors of the first moment, but also of the second and third moments of income. Estimated coefficients both take signs that make economic sense and are significant at standard levels of confidence.

However, when we test for their joint-significance we obtain relatively low F-Statistic, 7.9, 8.1 and 7.1 for the first, second and third moments of income respectively.

Even though instruments are correctly signed and are significant in explaining different moments of the changes in lagged income, we cannot ignore inference problems arising from 'weak' instruments. Indeed, the implied Cragg-Donald statistics suggest that our IV GMM estimates might contain substantial finite-sample biases. We follow two alternative strategies when addressing the issue of weak IVs. First, we estimate the model with alternative estimators – such as LIML and Fuller estimators – which are more robust to the presence of 'weak' IVs. Secondly, we re-estimate the quadratic polynomial model with a parsimonious set of instruments, which improves the joint explanatory power of the instruments and reduces the remaining bias in the second-stage estimates. While we discuss each of these strategies in detail in the appendix, we find reassuring that estimates from these alternative methods yield results remarkably similar to the baseline model.

In Table 4 we present our main results. The table reports second-stage estimates for the three different polynomial specifications. When including year dummies, we uncover substantial annual shifts in income growth – most notably a drop in growth between 2002 and 2003 and the subsequent rebound in 2004. This is consistent with the incidence of failed monsoon rains that affected interior areas of Andrah Pradesh and Maharashtra during the 2003, 2004 and 2005 Kharif season. Failing to control for such large year-specific growth patterns – possibly related to aggregate shocks affecting all villages – might hinder the identification of individual income dynamics. Indeed, we find that the inclusion of year dummies dramatically improves the precision in the estimates of the parameters of the polynomial.

As expected, when fitting a linear model we find a positive correlation between lagged income and current income (see column 2). The coefficient is statistically significant but modest in magnitude suggesting a relatively high degree of income mobility. In columns (3) to (6) we report the results for the quadratic and cubic model specifications.

Estimates for the quadratic model provide some striking results. First, we find that the income generating process might not be linear in nature. When we include the second-moment of lagged income we find it to be significant at the 5% level – see column (4). Secondly, the point estimates for the first-moment of lagged income are significant and very large in magnitude, suggesting convergence might be relatively slow or even unachievable. Indeed, point estimates show that the income generating process might follow a concave pattern, opening the possibility for the existence of multiple equilibria.

When allowing for a third-moment in the income generating process (column 6), point estimates change wildly and significance is lost. We attribute this to multicollinearity and poor identification power of the IV GMM estimator. We show in the appendix that alternative estimators – such as IV Fuller – perform better and yield estimates similar to the quadratic model specification.[11]

As suggested earlier, ignoring attrition in a long panel such as ICRISAT might lead to substantial biases in parameter estimates. Table 5 reproduces results for the IV GMM estimates where we apply inverse probability weighting (IPW) methods. The weights are computed as the inverse of the probability of attrition predicted by a vector of individual and household characteristics measured at the beginning of the survey period.[12] Table 5 shows that point estimates are remarkably similar in magnitude to our original estimates. In other

[11]The correlation coefficient between the second and third moment of lagged income in these regressions is very high, reaching 0.96 when estimating the model in column (6). Other estimators seem to perform better under these conditions, the cubic models for the IPW IV GMM and IV Fuller estimates provide point estimates similar to the quadratic model parameters. See Appendix for Fuller estimates.

words, while attrition in the ICRISAT sample is large and almost certainly non-random, patterns of attrition appear to be relatively orthogonal to income dynamics in the period of study. This is consistent with Badiani et al (2008), who find that correcting for attrition had relatively little impact on the determinants of consumption growth over the same period. At the same time, we find that significance of the estimated parameters is lost. This can be attributed to the loss in efficiency from applying the IPW procedure (Wooldridge, 2002b) and is corroborated by the increase in standard errors observed in Table 5 relative to the IV GMM standard errors (see Table 4).[13]

In the appendix, we discuss a number of further checks for robustness of the core results. The finding that income dynamics in the ICRISAT villages appear to be driven by a quadratic polynomial with pronounced concavity remains robust to changing the period of analysis, between the VLS1 and VLS2 panels, as well as to applying alternative estimation methods more robust to the presence of 'weak' IVs.

4.4.3 Multiple Equilibria and Household Heterogeneity

Showing that the current income follows a polynomial of lagged income with a concave pattern does not constitute proof of the existence of multiple equilibria. For that we need to show that the derivative of the resulting polynomial is larger than unity when $y_{i,t} = g(y_{i,t-1})$ – where the g-function represents the dynamic polynomial function – and that this condition is met within the range of values of the income distribution. Furthermore, in our model

[12] Attrition probit estimates are not reported, but results can be requested from the authors.

[13] As a further robustness check of the impact of attrition on our estimates, we re-estimate our baseline model only with individuals present in every year of the panel. This reduces the sample size substantially – from 8279 to 3301 observations – as well as the external validity, but it ensures that the parameter estimates will accurately reflect, free from attrition bias, the income generating process for the sub-sample of non-attiring individuals. In spite of the substantial attrition and the reduction in sample size, we find that baseline results remain largely unchanged. If anything, parameters become larger in magnitude and more precisely estimated. Results are not reported in the paper.

specification, household heterogeneity will shift the dynamic patterns and therefore any potential multiple equilibria will be household specific (see also Antman and McKenzie, 2005).

The easiest way to illustrate this is graphically. Graph 3 plots income simulations based on the parameter estimates of the polynomial from the quadratic model specification reported in Table 4.[14] [15]

The results presented in Graph 3 are truly striking. First, when considering the dynamic path for the median household, we find two equilibria in the range of reasonable values of income: namely a high stable equilibrium, and a low unstable saddle point. The high stable equilibrium is approximately 1900 Rupees per adult per year in 1976 prices, approximately 210 US dollars at the exchange rate of that time. The lower unstable equilibrium is about Rp 500 or 55 US dollars. The derived dynamic path implies that households with fixed effect values close to the median will face a divergence point or threshold whereby the dynamic paths separate. Namely, households lying above the saddle point will converge over time towards the high equilibrium, while households below the saddle point will inevitably suffer further losses in their future income. Lying below the threshold or being pushed over it by shocks such as rainfall would appear to put households on a path towards 'perdition'.

A second aspect to note from Graph 3 is that not all households are exposed to the risk of destitution. Households with sufficiently high individual fixed effects, as represented by the top decile in Graph 3, have a single dynamic equilibrium. For these types of household, we

[14] We report in the Appendix the full set of simulations based on all six of the specifications estimated in Table [4]. After recovering the household fixed effects implied by the model estimates, we plot the dynamic simulations for different percentiles of the fixed effects distribution. Specifically Graph 3 reports the simulated dynamic path for the 10th, 50th and 90th percentiles of the individual fixed effects.

expect their income levels to converge – although the rate of convergence suggested by the concave polynomial would appear to be very slow, even for relatively high values of lagged income. Furthermore, the existence of a single equilibrium for this group of households cannot be understated. Even when faced with large shocks, these households would appear to face little risk of being put on a path of structural divergence. It is as if their livelihood faces no vulnerability: even if they occasionally have low incomes, will not permanently remain in this state.

Thirdly, among households with low fixed effects, we find that the dynamic path also appears to experience multiple equilibria. However, for this group vulnerability does have a qualitatively different meaning than for other households. For low levels of fixed effect income, the stable equilibrium and the divergence threshold are close to each other. In other words, households that have already reached their steady state equilibrium could be pushed over the divergence threshold by a relatively small shock. Furthermore, it should be noted that the lower the fixed effects, the higher lies the saddle point. In other words, households with low fixed effects could be set on a structural divergence path even though they might have values of current income higher than households with higher fixed effects that are on a convergence path towards their high stable equilibrium.

Two further points are worth noting. Analysing the median household dynamic path we see that the threshold for this group lies on a relatively low value of income. It is therefore plausible that most households might have current income above the threshold, resulting in a relatively low risk of structural destitution. If this applies to the median, it is possible that the risk of being set on a divergence path is only real for a relatively small sub-sample of

[15] As one would expect, robustness checks on simulated dynamics provide similar results to those for model estimates. We find that using 'weak' IV robust estimators as well as attrition correction methods does not change the simulated dynamics. See Graph A2 in the Appendix.

households. However, Graph 3 provides some evidence against this possibility. Comparing the median versus the bottom decile households' curves, we see that there is only a small vertical distance between the two. In other words, half of our sample faces a dynamic threshold that lies between 500 and 1200 Rupees – the unstable equilibrium for the bottom decile.

Additionally, although clear from the graph, it is instructive to recognise the fact that households on a convergence path will ultimately converge to their own steady state equilibrium. Their eventual steady state income level – and arguably their level of long-term welfare – will therefore depend on their own individual fixed effects. This feature combined with the high inequality in the individual fixed effects in the top of the distribution – as represented by the large vertical distance between the median and the top decile of fixed effects – predicts substantial structural growth for this group of individuals. This prediction would appear to be consistent with the dramatic increases observed in income inequality over the period of analysis depicted in Graph 1.

Before moving on to the analysis of individual fixed effects, it is worth noting, that although our evidence suggests that income dynamics in rural India follow a concave process, the archetypal poverty trap model would predict a process following a complex polynomial with at least three moments. We think that these two views are not incompatible. When the dynamic path in poverty trap models is sufficiently low, the lower equilibrium will be defined by the so-called subsistence income. Failure to meet this minimum level of consumption cannot be prolonged for long periods without severe health consequences. With consumption and income bounded from below, the dynamic process is likely to be kinked at the lower end of the income distribution.

Figure 2: Poverty Trap with Kink

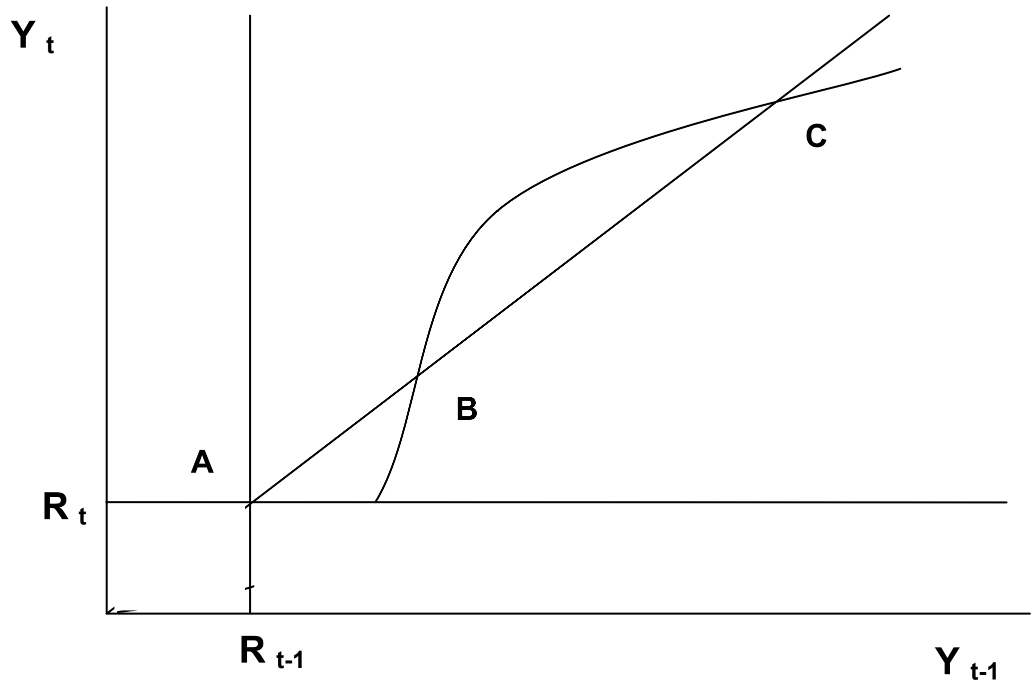


Figure 2 represents this type of process. Subsistence income is the true “road to perdition”; once this low steady state is reached it is near impossible to escape it. Yet, our methodology is ill equipped to identify its existence. On the one hand, parametric methods are too inflexible to identify kinks in the dynamic process – non-parametric are better suited. On the other hand, the nature of the kink implies that most households will not be found at the steady state (point A) rather than along the R(t)-line. Estimating reliably a third moment of the dynamic process is thereby made difficult by the fact that few households can be found outside the low equilibrium.

4.4.4 Unpacking Individual Fixed Effects

The importance of the individual fixed effects can hardly be overstated. Not only do they determine the steady state income households will eventually reach, but individuals with fixed effects approximately below the median are faced with the real possibility of suffering a shock that might put them on a dynamic path towards destitution. It is with this in mind that we now move towards understanding what lies behind these fixed effects.

To retrieve some idea of the correlates of income fixed effects, we regress them on a set of starting period household characteristics as measured at the beginning of our panel in 1975. Columns (1) and (2) in Table 6 report the correlates of individual fixed effects from the quadratic models (3) and (4) in Table 4. These columns include time-invariant household characteristics only as determinants of household fixed effects, while columns (3) and (4) include 'land area owned' and 'value of household assets' as additional time-varying correlates.

We find results in Table 6 are robust across all polynomial models. For our preferred model, the quadratic polynomial with year dummies, we find that beyond village dummies, education of the household head appear to be significantly correlated with fixed effect income. When we add land ownership and value of assets to education, we find all three to be strongly correlated with the fixed effects. Despite their geographic proximity, these villages have substantial heterogeneity in soil and other characteristics (Walker and Ryan, 1990). Correspondingly, individuals in different villages have different income trajectories: compared to the village of Aurepalle in the state of Andhra Pradesh, the default category, the villages of Shirapur, Kalman and Kinkheda in the state of Maharashtra have substantially lower fixed effects.

We interpret these results as suggestive that while changes in India over the past 30 years have increasingly created opportunities for substantial welfare improvements, not all households have been in the position to benefit. Human capital and physical assets appear crucial in ensuring that households are well enough equipped to take up these new opportunities.

4.5 Conclusions

A variety of theories suggest why a person who becomes poor at any time will remain poor indefinitely. Most such theories focus on a technology with increasing returns to scale which arises from a particular social mechanism – nutrition, education, fixed costs to entering a business, or another. The ideas of poverty traps that arise from these theories constitute a central theory of development economics at both the micro and macro levels. Yet these theories have received extremely little empirical support, possibly due to econometric pitfalls in the methods underlying the relevant empirical studies, as Dasgupta (1997) argues occurs for tests of the nutrition-efficiency wage theory, or possibly because no poverty trap in fact exists.

The large number of people in extreme penury constitutes only one reason underpinning the importance of understanding whether and why the destitute escape poverty. The presence of poverty traps would also imply a startling policy conclusion: a small transfer to a poor individual or household could change that person from low- to high-level equilibrium and permanently remove a person from poverty.

Since most existing theories of poverty traps assume some form of fixed investment cost, or increasing returns to assets or income, we examine whether income dynamics give evidence of increasing returns. A variety of econometric problems arise in this analysis: lagged income is inherently endogenous in a dynamic panel model; measurement error in income will cause OLS or GMM estimates to understate income's persistence; individual heterogeneity may disguise the fact that some individuals face a poverty trap even though the average individual does not; and short panel duration may give inadequate time to observe sufficient movement in income.

The bivariate kernel regressions or Arellano-Bond methods that existing papers use address some, but not all, of these pitfall. We apply IV GMM methods in a dynamic income equation that addresses issues of measurement error and endogeneity of lagged income, while allowing household heterogeneity in an income steady state. Unlike similar studies (Jalan and Ravallion, 2003 and Lokshin and Ravallion, 2004), we use exogenous instruments in exploiting deviations in annual precipitation to explain future income. Indeed, first-stage estimates reveal rainfall deviations to be a strong predictor of year-on-year changes. In particular, we interact rainfall with household land operated and household composition variables to obtain valid and relatively strong instruments for a polynomial function of contemporaneous income.

Our analysis shows that income generating dynamics in rural India follow a quadratic polynomial function with pronounced concavity. We find these results to be robust to changes in the period of analysis and corrections for non-random attrition. Further, parameter estimates obtained by applying alternative estimation methods more robust to the presence of 'weak' IVs remain consistent with our original results.

A recent study applying non-parametric methods to the same dataset as ours reports income mobility to be static, but provides no evidence of the existence of multiple equilibria (Nashold, 2009). When allowing for individual heterogeneity and using exogenous instruments to address income endogeneity, our analysis provides a different conclusion. Income simulations based on the estimated parameters suggest the presence of two distinct equilibria: a stable high-income equilibrium and a low-level unstable saddle point. While households with sufficiently high fixed effect income follow the high-equilibrium dynamic path, almost half of our sample has too low an income steady state to overcome the dynamic point of divergence. Analysis of income fixed effects shows that schooling and other assets at the beginning of the sample period are linked to high levels of steady state income.

We interpret our results as suggesting that changes in India over the past 30 years increasingly provide opportunities for substantial welfare improvements, but not all households are well placed to benefit. Education appears crucial in ensuring these opportunities are being taken. Those with higher assets have an income process with a much lower low-level unstable equlbrium than those with fewer assets: the latter's lives are far more precarious and even at higher income levels they risk sliding down dramatically. For some with high assets, this low unstable equilibrium would correspond to large negative current income positions. While in an agricultural setting occassional negative incomes are possible (and indeed observed in the data), it suggests that only for a rather negative and almost improbable income draw, they would face such fate.

Appendix

A.1 Estimation with Weak Instruments

As discussed in section 2, consistent IV GMM estimates require two instrumental variable conditions to hold, the 'validity' condition and the 'strong IV' condition. The validity condition states that the set of instruments should not be correlated with any unobserved determinant of income. The exogenous nature of the rainfall shocks and its potential heterogeneous effects suggest that our set of instruments is unlikely to be invalid. Indeed, Hansen J Over-identification statistic reported in Table 4 indicates that we cannot reject the null hypothesis that all excluded instruments are exogenous.

However, Cragg-Donald F-statistics related to Table 4 suggest that our set of instruments might not be sufficiently strong. Although our excluded instruments are good predictors of individual moments of lagged income, the Cragg-Donald statistics test the null hypothesis that the set of excluded instruments is jointly sufficiently strongly correlated with the set of endogenous variables. We find that in spite of first-stage F-Statistics of 7.90, 8.10 and 7.10 for the three-moments of changes in lagged income, the Kleibergen-Paap rank corrected Cragg-Donald statistics amounts to values of 3.5 and .0.0028 for the quadratic and cubic models with year dummies, respectively. When compared with Stock-Yogo critical values, these Cragg-Donald statistics suggest the presence of 'weak' instruments resulting in IV GMM estimates containing absolute biases approximately exceeding 30%.

Such magnitude of bias casts doubts on the reliability of the results presented earlier. Here we present two alternative approaches designed to provide further evidence of the robustness of our earlier results. First, we apply alternative estimators that are more robust to 'weak' instruments. We use LIML and Fuller estimators to obtain alternative point estimates for our main TSLS results. Secondly, in a world of weak IVs, the size of the resulting bias is understood to increase with the number of instruments (Stock and Yogo, 2005). As a robustness check we re-estimate our quadratic model, using only the interaction effects of the rainfall shocks. While removing the rainfall shock itself from the instrument set might reduce the bias, this comes at little cost, as rainfall shocks themselves are not significant when the model is estimated with year dummies.

Table A1 in the Appendix reproduces Table 4 for the alternative LIML and Fuller estimators. Comparing these results with the IV GMM estimates, we find that both point estimates and significance remain robust to estimation by LIML and Fuller.[16] Although still suspect of suffering from weak IV bias, we draw some comfort from the fact that these alternative estimators provide point estimates consistent with the IV GMM estimates.

Additionally, Table A2 reports results for our 'parsimonious IV' estimates of the quadratic model with year dummies. Using only the rainfall shocks interacted with land area and with the number of kids in the household, we improve the strength of our instrument. First-stage F-Stats increase to 10.10 and 9.70 for first and second moments respectively, and the Cragg-Donald statistic reaches a value of 4.22. Although the latter is not sufficiently high for weak IVs to be ruled out, the revised estimates would be expected to contain a smaller bias. Specifically, Stock-Yogo (2005) critical values for the IV GMM/LIML model suggest an absolute bias less than 20%.

Results reported in Table A2 are broadly consistent with our earlier results. It is also interesting to note that relative to LIML and Fuller estimates with three excluded instruments in Table A1, as well as to IV GMM estimates in Table 4, point estimates in the parsimonious IV model increase in magnitude for both the first and second moment. We interpret this change as an indication that the remaining 'weak' IV bias might be biasing downward the estimated parameters. In other words, the estimated dynamic patterns might be understating the degree of persistence and concavity of the true income generating process. Table A2 further shows that our 'parsimonious IV' results are also robust to correcting for attrition in observables.

The weak IV bias could also have implications for the dynamic simulations. Graph A2 in the Appendix reproduces the simulation dynamics for the Fuller regressions using the 'parsimonious IV' set reported in Table A2. We find that the estimated dynamics remain consistent with our original results. This is also the case when we correction for attrition in observables.

It should be noted that it is unclear whether the Cragg-Donald statistic is the appropriate diagnostic in our case. Our model specification is a particular case of a system of equations 'nonlinear in the endogenous variables' – to use the terminology in Wooldridge (2002b). By virtue of the second and third moments of lagged income being nonlinear functions of lagged income, the identifying conditions (rank and order) for the system are the same as in the case of the simple linear lagged income case (Fisher, 1965). In effect, any additional nonlinear term arrives with its own instruments – namely the first and higher moments of the instruments used in predicting the linear endogenous variable. For the assessment of the rank condition, Fisher (1965) proposes to treat the nonlinear terms as independent additional

[16] Note that the LIML estimation method is identical to the IV GMM procedure when the model is just identified, which is the case in our analysis when we estimate the cubic polynomial model.

variables, but these should be excluded from the rank condition tests.

The Cragg-Donald F-Statistic in its original form and the Kleibergen-Paap rank corrected version treat each endogenous variable as an independent equation; an assumption that would appear inappropriate in our case. Indeed, the performance of the Cragg-Donald statistic in our 'parsimonious IV' model casts further doubts on the validity of the test. While individually the first and second moments appear to be strong instruments following the Staiger and Stock rule of thumb – with F-Statistics of 10.10 and 9.70 respectively – the Cragg-Donald statistic yields a value (4.22) that would nevertheless predict a bias close to 20%.

In the case of a system of equations nonlinear in endogenous variables, Wooldridge (2002b) suggests an alternative estimation approach: instead of treating the nonlinear terms as independent variables, the nonlinear relation between the different moments can be explicitly exploited. Wooldridge proposes a procedure that uses $(\hat{y}_{t-1})^2$ as a single instrument for ${y_{t-1}}^2$, where the former is the square of the predicted first-moment of lagged income using the set of excluded instruments (Wooldridge, 2002b, p. 237). While more parsimonious, when we apply this alternative estimation methodology to our quadratic polynomial model, we find that – with an F-Statistic of 1.66 – the predictive power of $(\hat{y}_{t-1})^2$ on ${y_{t-1}}^2$ is in fact lower than the predictive power of the excluded instruments. If we nevertheless estimate the parameters of the quadratic polynomial following Wooldridge (2002b), we obtain estimates $y_t = 0.51 * y_{t-1} - 0.000127 * (y_{t-1})^2$. These estimates suggest a concave dynamic function with a single stable equilibrium, but low speed of convergence.

A.2 Parameter Stability and Panel Structure

Our model specification assumes that a single set of polynomial parameter values underlie the income generating process. However, considering the 15-year gap between VLS1 and VLS2, it is plausible that income dynamics have dramatically changed during that time. We address this issue by testing the parameter stability of the income polynomial function across the two panel periods. Table A3 reports results from estimating the quadratic model specification for the VLS1 and VLS2 panel samples separately. We find that VLS1 estimates have strong IVs – according to the Cragg-Donald test with the Kleibergen-Paap rank correction – and report parameter estimates broadly consistent with the full panel.

On the other hand, estimates for the VLS2 panel indicate the presence of severely weak instruments. Point estimates are also very imprecisely estimated. Both of these findings should not be surprising as we effectively estimate our differencing model using a two-period panel (2003 and 2004) only. Moreover, the dramatic income increases experienced during 2001 and 2002 – see Table 2 –, imply that the VLS2 sample might be a particularly noisy period during which to estimate income dynamics with a short time component. It is therefore especially encouraging to nevertheless find that point estimates are remarkably similar to estimates for the VLS1 sample only.

Accordingly, dynamic simulations presented in Graph A3 show very similar patterns between the two periods. We note that while the point of divergence appears stable between the periods, the higher stable equilibrium is substantially higher in the later panel. At the same time, estimated fixed effects in VLS2 appear to have a higher spread. Each of these findings is consistent with the rise in income variance shown in Graph 1 and suggest that the increased opportunities that modern India has to offer appear to be benefiting only a fraction of households.

More formally we test for the parameter stability between the two periods applying a pooling test or Chow test. We augment the quadratic dynamic model as characterised in equation (1) with interaction effects for the VLS1 and VLS2 periods with each of the dynamic terms. The Chow test analyses the null hypothesis that the dynamic parameters are identical between the two periods. Unable to appropriately estimate the augmented model by two-stage least squares – faced with the impossible task of instrumenting six endogenous variables –, we apply the test instead to the OLS model. The results of the Chow test are in line with earlier evidence; we find that the stability of the quadratic polynomial parameters estimated by OLS can not be rejected.

A.3 Dynamic Properties of the Income series

A concern through the paper is with regards to the appropriate specification of our underlying model, and in particular whether the income series might be an AR(2) process rather than the assumed AR(1) with polynomial. If this was the case, not only would our base model be misspecified, but the validity of the rainfall instruments would be put into question. Formally, an AR(2) process would be found if the error term, v(t), in equation (1) follows an AR(1) process, as depicted in equation (6) below:

$$y_{it} = \beta g(y_{i,t-1}) + \rho_i + v_{it} \qquad (1)$$

$$v_{it} = Shock_t + Shock_{t-1} \qquad (6)$$

There are a number of mechanisms through which previous rainfall shocks might affect current income. Perhaps rainfall deviations are correlated over time. However, more likely is that past shocks might have a persistent effect on income by affecting household asset accumulation and investment choices. Rural households are expert agents at coping with shocks; to do so they might sell assets or livestock, change crop choices, or modify their fertiliser practices. Many of these measures, however, can have dynamic effects on their ability to generate future income. As such, equation (6) might be plausible depiction of income dynamics.

To test this possibility we explore the auto-correlation properties of first-differenced equation (2), which removes the unobserved and perfectly auto-correlated household unobserved term, ρ_i. In a first-differences specification, by construct the error term will follow an AR(1) process. To test for persistent dynamic effects of shocks in the levels equation, we ought to test for the presence of AR(2) processes in the first-differenced error term, Δv_{it}. Note that this condition is similar to the requirements of dynamic GMM models such as Arellano-Bond estimators.

Table A4 reports the results of testing for AR(1) and AR(2) in the first-differenced equation for different model specifications of equation (2). We consider linear, quadratic and cubic models as well as models with and without year dummies. As expected we find evidence throughout of AR(1) differenced residuals, but in spite of our concerns, there is no evidence of AR(2) dynamics patterns in the differenced income series. The results are reassuring, and suggest that our dynamic model is correctly specified, and that rainfall shocks do not have a persistent dynamic effect on income.

References

Adato, Michelle, Michael. Carter, and Julian May. Exploring Poverty Traps and Social Exclusion in South Africa Using Qualitative and Quantitative Data. *Journal of Development Studies* 42(2):26–247, 2006

Anderson, T.W. and Cheng Hsiao. Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics*, 18(1):47-82, 1982.

Antman, Francesca and David McKenzie. Poverty traps and Nonlinear Income Dynamics with Measurement Error and Individual Heterogeneity. *Journal of Development Studies* 43 (6) 1057-1083, 2007.

Arellano, Manuel and Stephen Bond. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 48(2):277-297, 1991.

Badiani, Reena, Stefan Dercon, Pramila Krishnan and K.P.C. Rao. Changes in Living Standards in Villages in India 1975-2004: Revisiting the ICRISAT village level studies. CPRC Working Paper Series, no. 85, 2008.

Banerjee, Abhijit and Andrew Newman. Occupational Choice and the Process of Development. *Journal of Political Economy*, 101(2):274-298, 1993.

Barrett, C.B., P.P. Marenya, J.G. McPeak, B. Minten, F. Murithi, W. Oluoch-Kosura, F. Place, J.C. Randrianarisoa, J. Rasambainarivo, and J. Wangila. Welfare Dynamics in Rural Kenya and Madagascar. *Journal of Development Studies* 42:248–277, 2006

Bliss, Christopher J. and Nicholas H. Stern. *Palanpur: The economy of an Indian village*. Oxford University Press, Oxford, 1982.

Blundell, Richard and Stephen Bond and Frank Windmeijer. Estimation in dynamic panel data models: improving on the performance of the standard GMM estimators. 2000.

Bond, Stephen R. Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, 1(2):141-162, 2002.

Bound, John and David A. Jaeger and Regina M. Baker. Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak. *Journal of the American Statistical Association*, 90(430):443-450, 1995.

Bound, John and Alan B. Krueger. The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?. *Journal of Labor Economics*, 9(1):1-24, 1991.

Carter, Michael and Chris Barrett. The economics of poverty traps and persistent poverty: An asset based approach. *Journal of Development Studies*, 42, 178–199, 2006

Chen, Shaohua and Martin Ravallion. “The Developing World Is Poorer Than We Thought, But No Less Successful in the Fight against Poverty”, Policy Research Working Paper 4703, The World Bank. 2008.

Dasgupta, Partha and Debraj Ray. Inequality as a Determinant of Malnutrition and Unemployment: Theory. *The Economic Journal*, 96(384), 1986.

Dasgupta, Partha, “An Enquiry into Well-Being and Destitution”, Oxford University Press, 1993.

Dercon, Stefan and Joseph Shapiro. “Moving On, Staying Behind, Getting Lost: Lessons on poverty mobility from longitudinal data”, chapter in Deepa Narayan and Patti Petesch, *Moving out of Poverty*, World Bank 2007.

Fan, Jianqing. Design-adaptive Nonparametric Regression. *Journal of the American Statistical Association*, 87(420), 1992.

Fisher, Franklin M. Identifiability Criteria in Nonlinear Systems: A Futher Note. *Econometrica*, 33, 197-205, 1965

Fitzgerald, John and Peter Gottschalk and Robert Moffitt. An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. *Journal of Human Resources*, 33(2):251-299, 1998.

Galor, Oded and Joseph Zeira. Income Distribution and Macroeconomics. *Review of Economic Studies*, 60(1):35-52, 1993.

Heckman, James J. Formulation and Estimation of Dynamic Models Using Panel Data. *Econometrica*, 47(1):153-162, 1979.

Jacoby, Hanan and Emmanuel Skoufias. Estimating the Return to Schooling: Progress on Some Persisitent Econometric Problems. *Review of Economic Studies*, 64, 1997.

Jalan, Jyotsna and Martin Ravallion. *Insurance Against Poverty*, chapter Household Income Dynamics in Rural China. Oxford University Press, 2003.

Ligon, Ethan, 1998. "Risk Sharing and Information in Village Economics," Review of Economic Studies, Wiley Blackwell, vol. 65(4), pages 847-64, October.

Lokshin, Michael and Martin Ravallion. Household Income Dynamics in Two Transition Economies. *Studies in Nonlinear Dynamics & Econometrics*, 8(3), 2004.

McKenzie, David and C Woodruff. Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises *Economic Development and Cultural Change*, Vol. 55, pages 3-42, 2006.

Mookherjee, Dilip and Debraj Ray. Contractual Structure and Wealth Accumulation. *American Economic Review*, 92(4):818-849, 2002.

Mookherjee, Dilip and Debraj Ray. Aspirations, Segregation and Occupational Choice. *Journal of the European Economic* Association, Vol. 8, No. 1, Pages 139-168, 2010

Munshi, Kaivan and Mark Rosenzweig. Why is Mobility in India so Low? Social Insurance, Inequality, and Growth. NBER Working Papers 14850, 2009,

Murray, Michael. Avoiding Invalid Instruments and Coping with Weak Instruments. *Journal of Economic Perspectives*. American Economic Association, vol. 20(4), pages 111-132, 2006a.

Murray, Michael. The Bad, the Weak, and the Ugly: Avoiding the Pitfalls of Instrumental Variables Estimation, 2006b. Available at SSRN: http://ssrn.com/abstract=843185.

Naschold, F. Poor stays poor - Household Asset Poverty Traps in Rural Semi-arid India. Mimeo. 2009.

Nickell, Stephen. Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417-1426, 1981.

Quisumbing, Agnes and Bob Baulch. Assets and Poverty Traps in Rural Bangladesh. IFPRI Working Papers, July 2009, No. 143. 2009

Ray, Debraj. Aspirations, Poverty, and Economic Change in *Understanding Poverty*, Oxford University Press, 2006

Rosenzweig, Mark R. Labour markets in low-income countries in *Handbook of Development Economics* . North-Holland, Amsterdam, 1988.

Stock, James H. and Jonathan H. Wright and Motohiro Yogo. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics*, 20(4):518-529, 2002.

Stock, J. H., and M. Yogo. Testing for Weak Instruments in Linear IV Regression. In Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, ed. D.W. Andrews and J. H. Stock, 80–108, 2005.

Swamy, Anand V. A simple test of the nutrition-based efficiency wage model. *Journal of Development Economics*, 53:85-98, 1997.

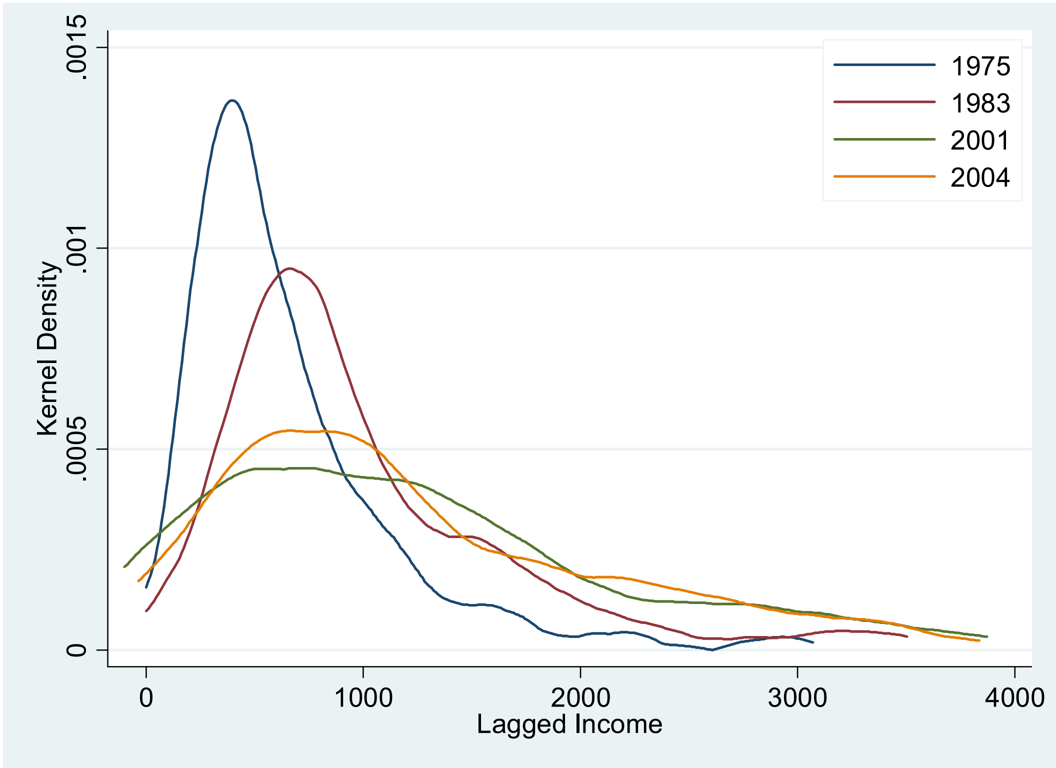
Townsend, Robert M, 1994. "Risk and Insurance in Village India," Econometrica, Econometric Society, vol. 62(3), pages 539-91, May.

Walker, Thomas S. and James G. Ryan. *Village and Household Economies in India's Semi-Arid Tropics*. Johns Hopkins Press, Baltimore, 1990.

Wooldridge, Jeffrey M. Inverse probability weighted M-estimators for sample selection, attrition, and stratification. Institute for Fiscal Studies Working Paper CWP11/02, 2002a.

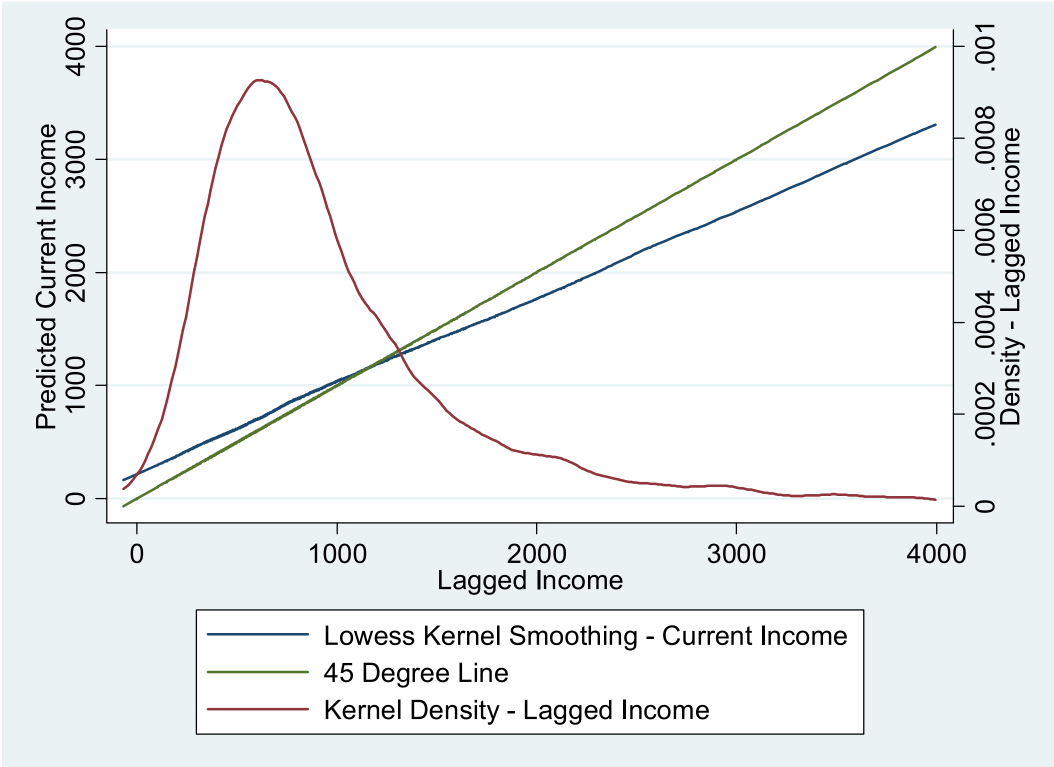
Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2002b.

Graph 1: Evolution of Household Income Distribution Over 30-Year Period
Kernel Densities of Income for Years 1975, 1983, 2001 and 2004



Note: Kernel smoothing densities for specific years. For ease of presentation incomes plotted are restricted to values -100 and 4000 Rupees in 1975 prices. All VLS1 households and their later split-off households are included, but new VLS2 households have been excluded.

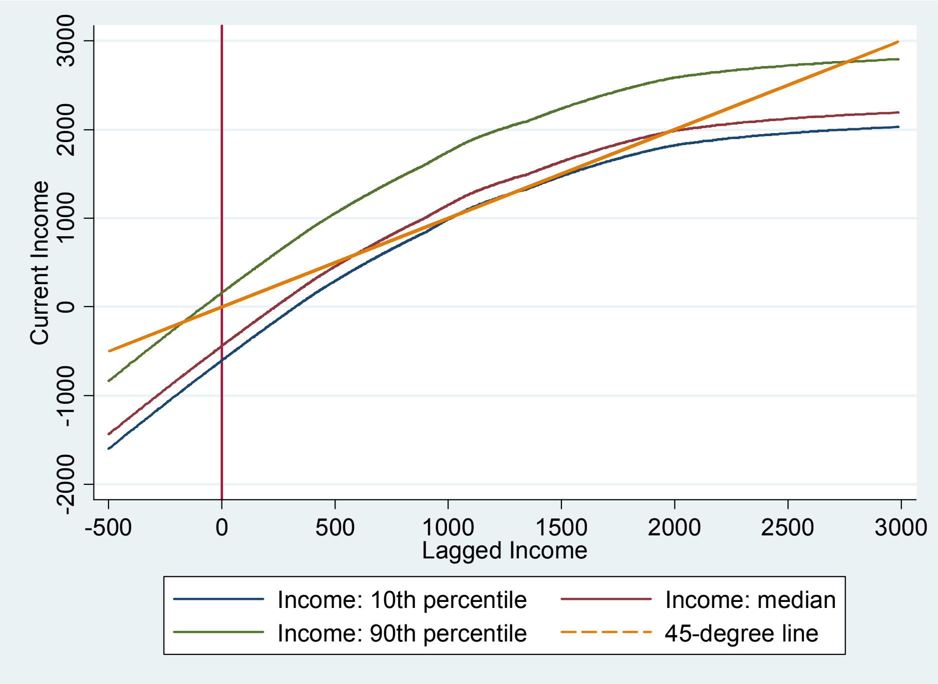
Graph 2: Bivariate Lowess Estimates of Current Household Income on Lagged Household Income, and Kernel Densities of Lagged Income



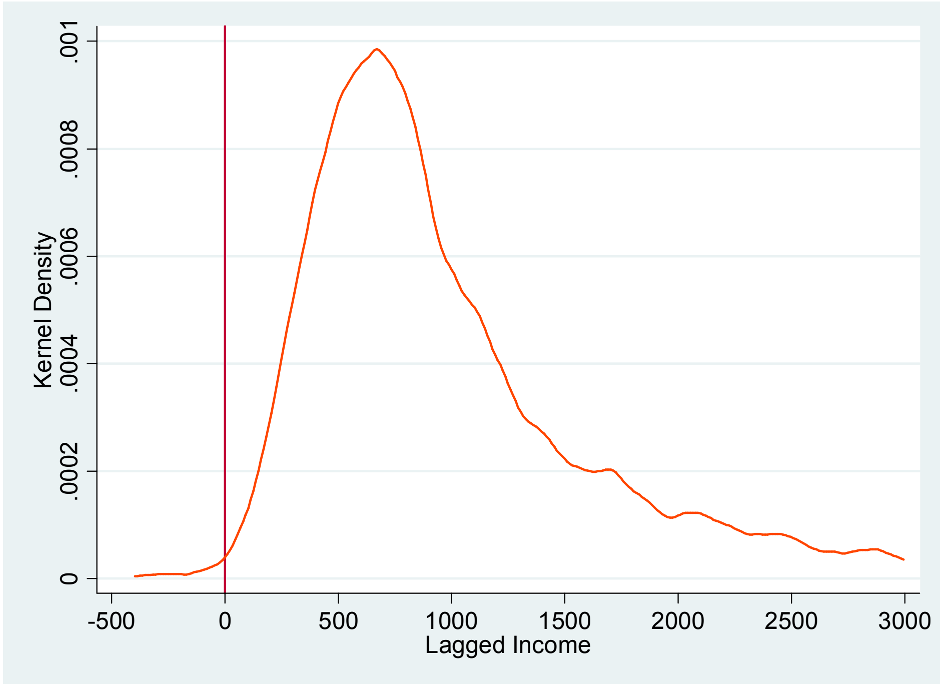
Note: 45-degree line indicates locus where current income equals lagged income. Data used for the graphs is restricted to the years used in the parametric analysis, namely 1977-1983, and 2003-2004. For ease of presentation, incomes plotted are restricted to values -100 and 4000 Rupees in 1975 prices. Parametric analysis in the paper uses all income values. All VLS1 households and their later split-off households are included, but new VLS2 households have been excluded.

Graph 3: Simulated Income Dynamics for Median, 10th percentile and 90th percentile of fixed effects
Model Specification: TSLS GMM Estimates of Quadratic Model with Year Dummies,

Panel A – Simulated Dynamics

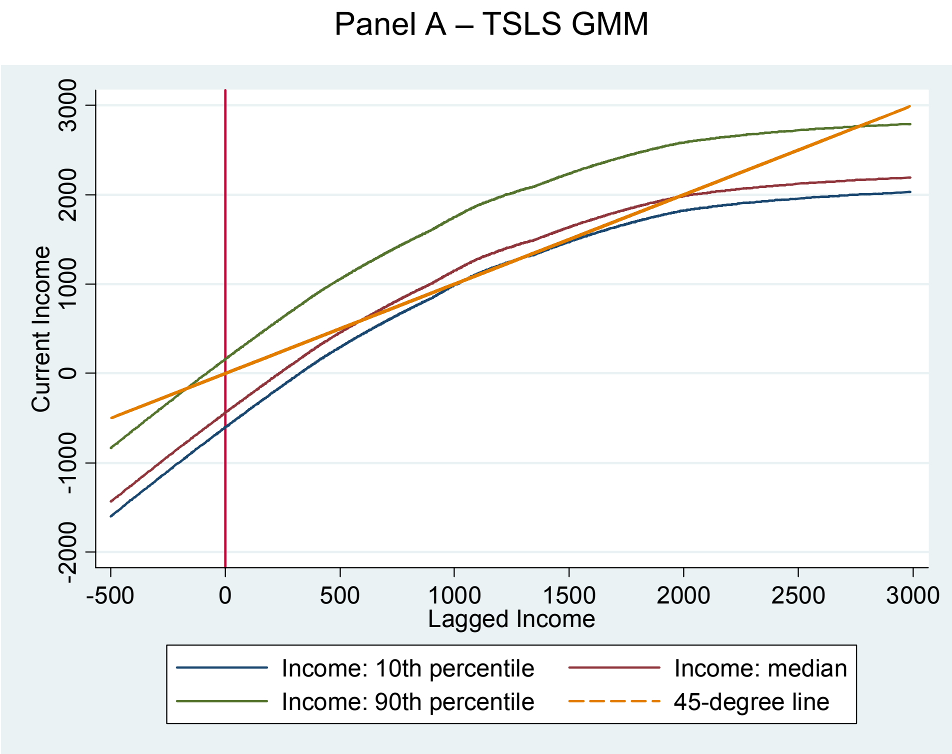


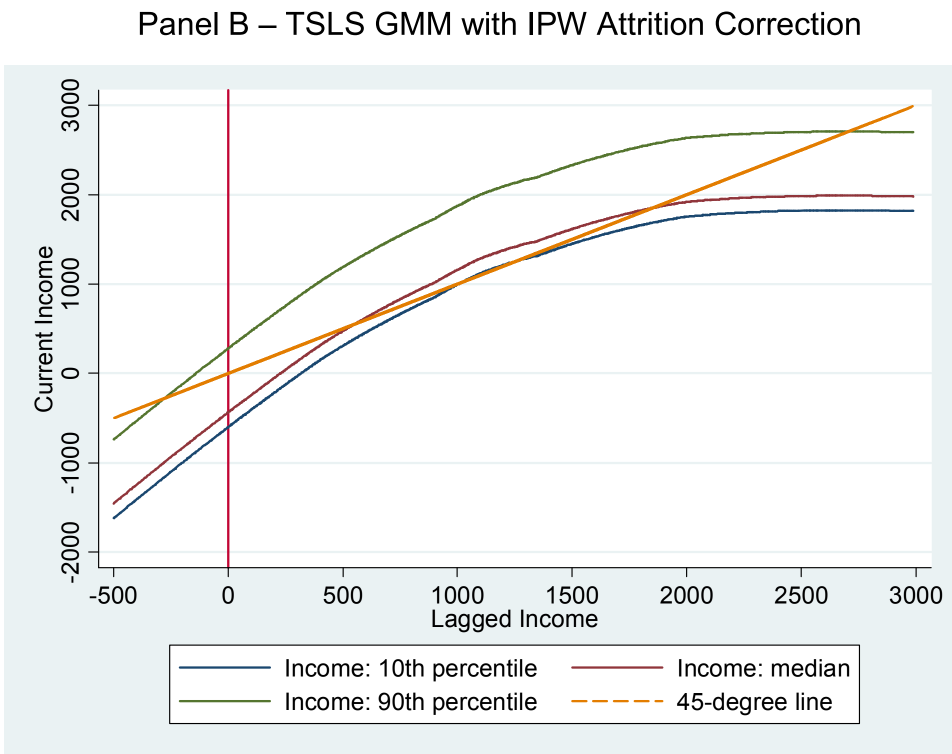
Panel B – Kernel Density of Lagged Income



Note: Simulations of income dynamics have been computed using the estimated parameter of the polynomial of lagged income. Plotted simulation based on TSLS GMM estimates for model specification with quadratic polynomial and year dummies. Computed fixed effects act as shifters of the locus of the polynomial. Simulation curves are depicted for a realistic range of lagged-income. Intersections between the dynamic trajectories and the 45-degree line indicate a potential equilibrium. Vertical line indicates zero income. Panel B kernel densities of lagged income does not show 104 observations with values of lagged income above 3000 Rupees, as well as 10 observations with values below -500. In both cases values not shown are otherwise included in all regressions.

Graph 4: Attrition Correction – Simulated Income Dynamics for Median, 10th percentile and 90th percentile of fixed effects Model Specification: TSLS GMM Estimates of Quadratic Model with Year Dummies, with and without Attrition Correction





Note: Simulations of income dynamics have been computed using the estimated parameter of the polynomial of lagged income. Plotted simulations based on IV GMM estimates for model specification with quadratic polynomial and year dummies (Panel A) as well as with Inverse Probability Weighting (IPW) attrition correction (Panel B). Computed fixed effects act as shifters of the locus of polynomial. Simulation curves are depicted for a realistic range of lagged-income. Intersections between the dynamic trajectories and the 45-degree line indicate a potential equilibrium. Vertical line indicates zero income.

Table 1: Tracking and attrition in the 2001-2004 survey

Status by 2004-05	Full sample of individuals included in 1975-1984 (VLS1) with tracking information in 2005	Of which: Included in the 2001 survey, i.e. in the village and in the sample in 2001	Of which: Not included in the VLS2, 2001 survey
Dead in 2005?	432	24	408
Migrated in 2005?	675	45	630
In village in 2005?	857	581	276
No information in 2005?	34	4	30
Total	1998	654	1344

Note: based on attempts to track 1998 individuals included at some point between 1975-84 in the original households of the 1975-84 sample. Not including servants. See Badiani et al (2008)

Table 2. Mean and Standard Deviation of Income Series at Household and Individual Level, by Survey Year

	VLS 1									VLS 2				*Total*
	1975	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*	*1983*	*2001*	*2002*	*2003*	*2004*	
Income (HH Level)	668.3	863.5	1006.8	1017.4	1032.6	971.7	1017.7	1180.2	1049.8	1415.4	1647.0	1165.6	1300.9	1076.3
	(553.8)	(713.0)	(783.3)	(741.1)	(826.7)	(792.9)	(857.9)	(981.5)	(826.3)	(1,605.3)	(1,633.6)	(1,250.5)	(1,487.4)	(1,034.7)
Nr of Households	238	240	240	238	238	237	236	236	237	173	171	171	170	2825
Income (Indiv. Level)	663.8	871.4	1038.9	1045.8	1064.8	1011.7	1066.6	1220.1	1085.7	1547.8	1809.2	1212.7	1434.2	1055.8
	(544.7)	(714.8)	(801.8)	(760.8)	(827.7)	(800.3)	(851.4)	(962.9)	(849.7)	(1,674.4)	(1,697.2)	(1,376.9)	(1,375.9)	(942.8)
Nr of Individuals	1333	1250	1193	1127	1074	1020	1000	980	964	360	353	338	325	11317

Note: Income measured in 1975 rupees per adult equivalent per year. Standard deviations appear in parentheses beneath mean values of continuous variables. Household series includes all VLS1 households and their later split-off households, but does not include new VLS2 households. The Individual income series corresponds with the panel of individuals used in our analysis. Attrition is treated as an absorbing state, such that only individuals present in all waves since 1975 (excluding the 1984-2000 gap) are included.

Table 3. First Stage Results – Impact of Rainfall on Lagged Income

Dependent variable:	ΔY_{t-1}	ΔY_{t-1}	ΔY_{t-1}^{2}	ΔY_{t-1}^{2}	ΔY_{t-1}^{3}	ΔY_{t-1}^{3}
	(1)	(2)	(3)	(4)	(5)	(6)
Square Rainfall shock, t-1	0.0007**	0.0004	3.5253**	2.8780*	1.9e+04*	1.4e+04*
	(0.000)	(0.000)	(1.71E+00)	(1.51E+00)	(1.20E+04)	(7.70E+03)
Rainfall shock * land operated (ha), t-1	28.7496***	13.5687	1.4e+05***	1.0e+05*	6.0e+08**	5.2e+08*
	(9.486)	(10.479)	(5.00E+04)	(5.40E+04)	(2.50E+08)	(2.70E+08)
Rainfall shock * children 0-8 in HH, t-1	-9.1588**	-27.8874***	-6.5e+04***	-1.1e+05***	-3.7e+08***	-4.8e+08***
	(4.529)	(6.485)	(2.00E+04)	(2.40E+04)	(1.20E+08)	(1.10E+08)
ΔYear: 1977		228.1362***		5.9e+05***		1.9e+09***
		(35.671)		(1.40E+05)		(6.00E+08)
ΔYear: 1978		194.7497***		5.0e+05**		1.30E+09
		(51.615)		(2.30E+05)		(9.80E+08)
ΔYear: 1979		14.1897		-150,000.00		-9.60E+08
		(41.736)		(1.70E+05)		(7.20E+08)
ΔYear: 1980		42.9072		160,000.00		8.90E+08
		(40.541)		(1.90E+05)		(9.40E+08)
ΔYear: 1981		-49.6454		-170,000.00		-900000000
		(50.153)		(2.20E+05)		(1.10E+09)
ΔYear: 1982		62.6628		99,000.00		5.60E+08
		(52.514)		(2.10E+05)		(9.90E+08)
ΔYear: 1983		195.2823***		7.0e+05***		2.9e+09***
		(36.451)		(1.80E+05)		(9.90E+08)
ΔYear: 2003		342.6323**		1,300,000.00		8.60E+09
		(145.482)		(9.30E+05)		(7.20E+09)
ΔYear: 2004		-7.7e+02***		-3.3e+06***		-1.9e+10*
		(161.081)		(1.30E+06)		(1.00E+10)
Observations	8279	8279	8279	8279	8279	8279
R-squared	0.0071	0.0776	0.0050	0.0388	0.0024	0.0224
F Test for joint significance of IV	6.3751	7.9327	4.8832	8.0919	3.8767	7.0878

Note: Standard errors robust to heteroskedasticity and 'original' household-level clustering appear in parentheses. Income measured as real Rupees per adult equivalent. 'Rainfall shock' measured as the ratio between Kharif precipitation (rainy season) and the Kharif historic precipitation as observed across all waves of the panel. 'Rainfall shock' is interacted with the size of land operated in hectares as well as the number of children aged 0-8 in the household. Variable 'Square Rainfall shock' is the 'signed' square of the Kharif precipitation difference in a wave and the historic mean.

Table 4. Modelling Income Dynamics – TSLS GMM Estimates

Dependent variable: ΔYt	Linear Polynomial		Quadratic Polynomial		Cubic Polynomial	
	(1)	(2)	(3)	(4)	(5)	(6)
ΔY_{t-1}	0.209664	0.457039***	0.074571	2.037840**	3.854833	-1.200000
	(0.257)	(0.155)	(0.652)	(0.886)	(4.348)	(23.849)
ΔY_{t-1}^{2}			0.000033	-0.000407**	-0.002223	0.003986
			(0.000)	(0.000)	(0.002)	(0.032)
ΔY_{t-1}^{3}					0.000000	-0.000001
					(0.000)	(0.000)
ΔYear: 1977		1.1e+02**		-17.0000		-360.0000
		(43.768)		(94.034)		(2,457.988)
ΔYear: 1978		-1.0e+02**		-1.6e+02**		-660.0000
		(52.173)		(74.585)		(3,650.417)
ΔYear: 1979		18.0000		-21.0000		-70.0000
		(36.261)		(53.182)		(375.374)
ΔYear: 1980		-8.9e+01*		-60.0000		80.0000
		(50.787)		(62.560)		(1,009.924)
ΔYear: 1981		9.5e+01*		120.0000		-14.0000
		(48.974)		(85.581)		(975.295)
ΔYear: 1982		1.3e+02***		1.3e+02*		330.0000
		(35.737)		(70.826)		(1,471.271)
ΔYear: 1983		-2.2e+02***		-2.3e+02***		-400.0000
		(44.455)		(51.340)		(1,280.467)
ΔYear: 2003		-7.5e+02***		-8.6e+02***		1400.0000
		(174.674)		(252.021)		(17,000.000)
ΔYear: 2004		6.8e+02***		5.8e+02*		-2000.0000
		(224.472)		(325.147)		(20,000.000)
Observations	8279	8279	8279	8279	8279	8279
R-squared	-0.208	-0.4627	-0.2362	-1.7644	-18.0796	-159.6338
First-Stage Diagnostics						
Number of Instruments	3	3	3	3	3	3
Cragg-Donald F-Statistic	19.8126	15.7872	1.4762	3.515	0.3546	0.0028
Second-Stage Diagnostics						
Anderson-Rubin F stat	2.756	5.092	2.756	5.092	2.756	5.092
Prob > F	0.043	0.002	0.043	0.002	0.043	0.002
Hansen J Statistic (overidentification)	6.3701	4.311	6.6131	0.3628	-	-
Prob > chi-squared	0.0414	0.1158	0.0101	0.547	-	-

Note: Standard errors robust to heteroskedasticity and ‘original’ household-level clustering appear in parentheses. Income measured in 1975 Rupees per adult equivalent. Anderson Rubin F statistic tests the null hypothesis that the endogenous regressors are jointly insignificant in structural equation. Hansen J Statistic, reported for overidentified models tests null hypothesis that the excluded instruments are uncorrelated with the structural equation error. Cragg-Donald F-Statistics reported includes Kleibergen-Paap rank correction. Cragg-Donald 'Weak IV' statistic tests for the null hypothesis that instruments are strongly correlated with the set of endogenous variables.

Table 5. Attrition Corrected Income Dynamics – TSLS GMM Estimates with Inverse Probability Weighting (IPW), Model estimates with Year Dummies

Dependent variable: ΔYt	Linear Polynomial	Quadratic Polynomial	Cubic Polynomial
	(1)	(2)	(3)
ΔY_{t-1}	0.416034**	2.091244	2.645707
	(0.194)	(1.325)	(11.219)
ΔY_{t-1}^{2}		-0.000451	-0.000867
		(0.000)	(0.008)
ΔY_{t-1}^{3}			0.000000
			(0.000)
ΔYear: 1977	1.1e+02**	-13.0000	17.0000
	(52.097)	(106.990)	(580.139)
ΔYear: 1978	-9.2e+01*	-1.5e+02*	-110.0000
	(55.746)	(76.522)	(786.512)
ΔYear: 1979	4.9864	-34.0000	-46.0000
	(37.127)	(64.493)	(265.456)
ΔYear: 1980	-9.4e+01*	-71.0000	-63.0000
	(56.946)	(72.290)	(180.953)
ΔYear: 1981	120.0000	180.0000	210.0000
	(73.779)	(155.866)	(571.674)
ΔYear: 1982	1.2e+02***	73.0000	44.0000
	(40.841)	(122.490)	(591.942)
ΔYear: 1983	-2.1e+02***	-2.2e+02***	-180.0000
	(46.890)	(68.703)	(795.031)
ΔYear: 2003	-7.8e+02***	-9.4e+02**	-1100.0000
	(197.158)	(433.624)	(2,542.320)
ΔYear: 2004	6.6e+02***	150.0000	29.0000
	(238.309)	(649.886)	(2,837.888)
Observations	8279	8279	8279
R-squared	-0.4441	-3.2669	-5.0413
Anderson-Rubin F stat	3.439	3.439	3.439
Prob > F	0.018	0.018	0.018
Hansen J Statistic (overidentification)	3.4689	0.0037	-
Prob > chi-squared	0.1765	0.9513	-

Note: TSLS GMM model estimated with inverse probability weights. Weights are computed as the inverse of the predicted probability of attrition in a given wave. Individual, household and household head characteristics are used as predictors of wave attrition. Standard errors robust to heteroskedasticity and household-level clustering appear in parentheses. Income measured in 1975 Rupees per adult equivalent. Anderson Rubin F statistic tests the null hypothesis that the endogenous regressors are jointly insignficant in structural equation. Hansen J Statistic, reported for overidentified models tests null hypothesis that the excluded instruments are uncorrelated with the structural equation error.

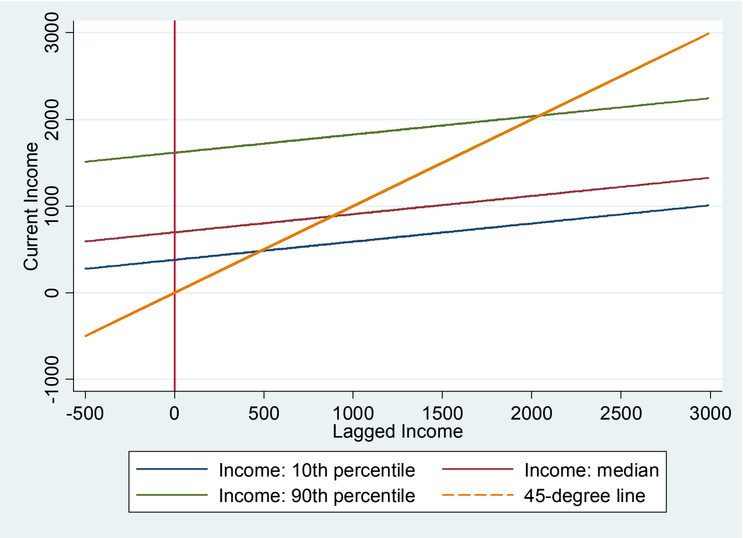
Table 6. Unpacking Individual Fixed Effects - Correlates of HH Steady State Income

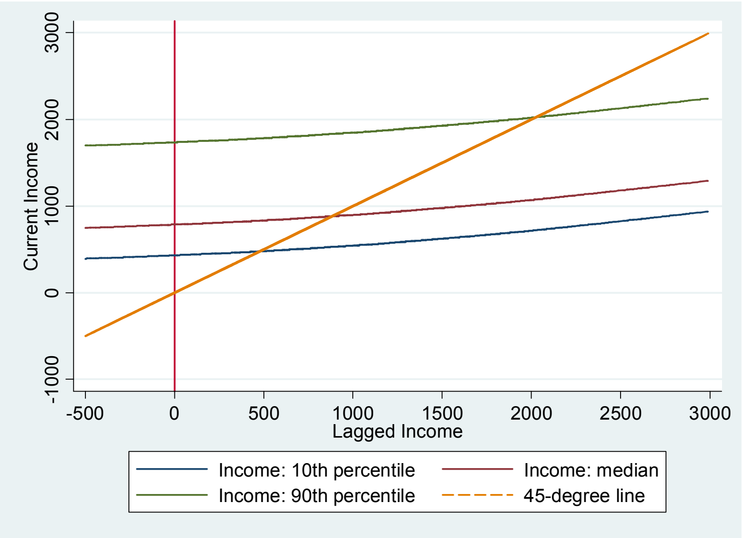
Dependent variable: Individual Fixed Effects	Quadratic Polynomial			
	No Year Dummies	Year Dummies	No Year Dummies	Year Dummies
	(1)	(2)	(3)	(4)
HH Head Education, 1975	1.7e+02***	1.6e+02***	7.9e+01**	1.1e+02***
	(38.57)	(39.38)	(32.16)	(32.29)
HH Head Age, 1975	2.61	2.31	1.59	1.56
	(3.03)	(2.57)	(2.40)	(2.32)
HH Head Sex, 1975	-59.00	-95.00	-30.00	-81.00
	(154.75)	(71.80)	(128.14)	(73.92)
HH Size, 1975	22.00	-1.50	4.92	-14.00
	(24.14)	(19.04)	(18.81)	(18.54)
Nr of Members ages 9-14, 1975	-14.00	-50.00	6.41	-34.00
	(50.81)	(33.57)	(38.18)	(30.87)
Nr of Members ages 0-8, 1975	-8.1e+01*	-29.00	-37.00	-0.93
	(41.09)	(33.14)	(31.72)	(29.23)
Area Owned by HH (in Ha), 1975			2.1e+02***	1.2e+02**
			(34.08)	(47.48)
Value of HH Assets (in Ru), 1975			0.007989**	0.006248**
			(0.00)	(0.00)
Low Caste Dummy	-3.2e+02***	-5.90	-1.5e+02***	1.1e+02**
	(64.43)	(47.36)	(56.74)	(45.48)
Village Dokur	-67.00	-1.8e+02*	83.00	-85.00
	(122.92)	(105.29)	(101.51)	(100.44)
Village Shirapur	-2.2e+02**	-2.9e+02***	-62.00	-1.8e+02**
	(104.55)	(92.59)	(91.84)	(90.64)
Village Kalman	-100.00	-2.6e+02**	-76.00	-2.5e+02**
	(118.46)	(103.28)	(108.08)	(103.96)
Village Kanzara	-5.60	-150.00	110.00	-77.00
	(132.56)	(149.20)	(115.85)	(147.98)
Village Kinkheda	-60.00	-3.0e+02**	100.00	-190.00
	(158.02)	(128.65)	(134.27)	(123.15)
Constant	6.5e+02**	-3.3e+02*	4.7e+02**	-4.3e+02**
	(274.94)	(187.33)	(235.25)	(183.62)
Observations	1176	1176	1176	1176
R-squared	0.235	0.235	0.1681	0.2638

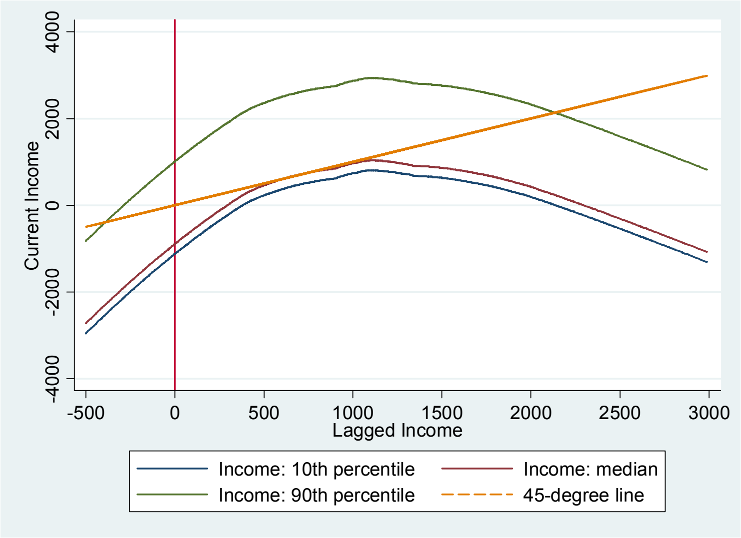
Note: Standard errors robust to heteroskedasticity and household-level clustering appear in parentheses. We restrict the sample size to the number of individuals for which we can compute income fixed effects. Fixed effects used in columns (1) to (4) obtained using parameter estimates corresponding to model specifications reported in columns (3) and (4) in Table 3.

Graph A1: Simulated Income Dynamics for Median, 10th percentile and 90th percentile of fixed effects
All Model Specifications: TSLS GMM Estimates - Linear, Quadratic and Cubic

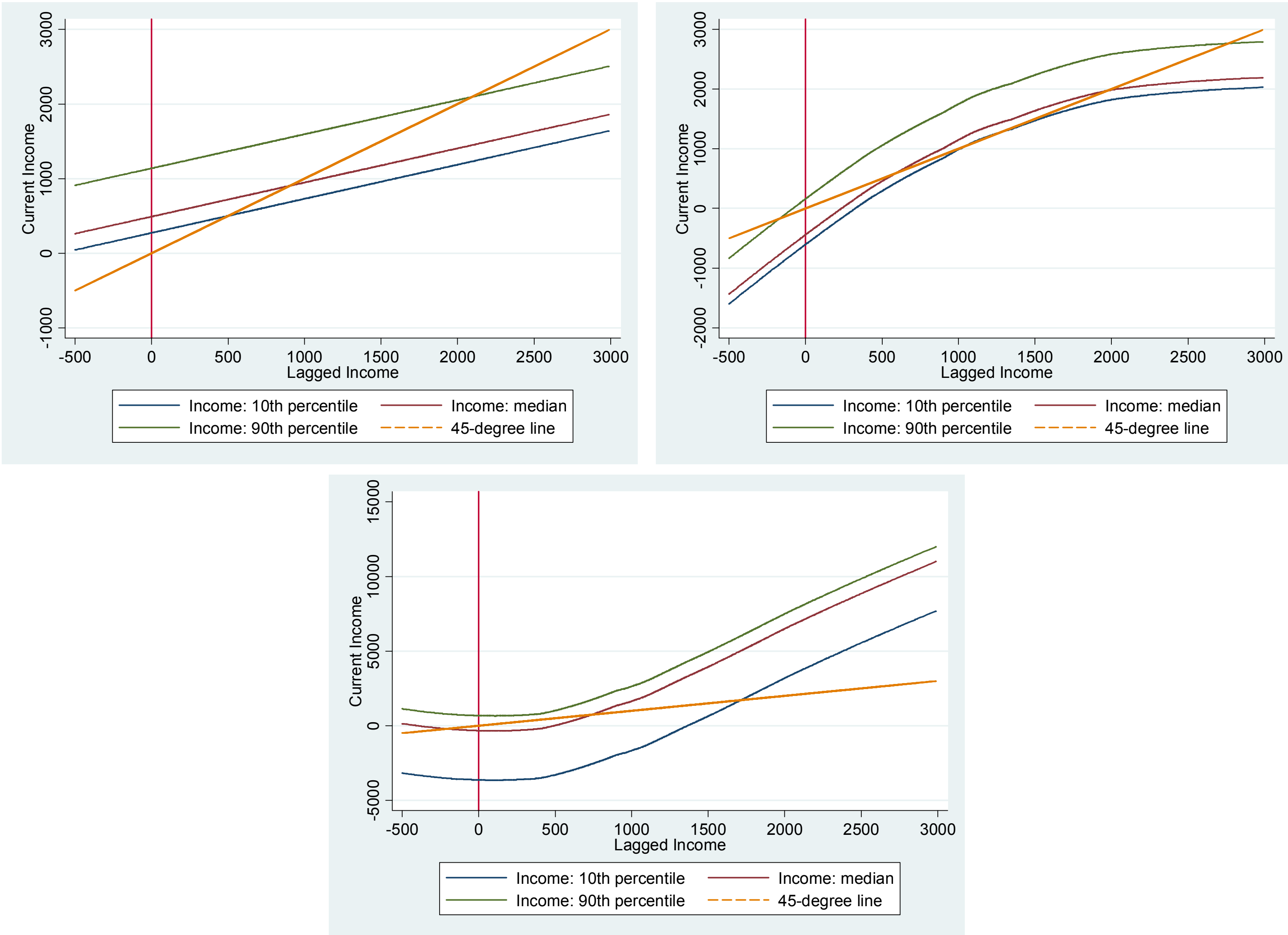
Panel A – No Year Dummies – Linear, Quadratic and Cubic







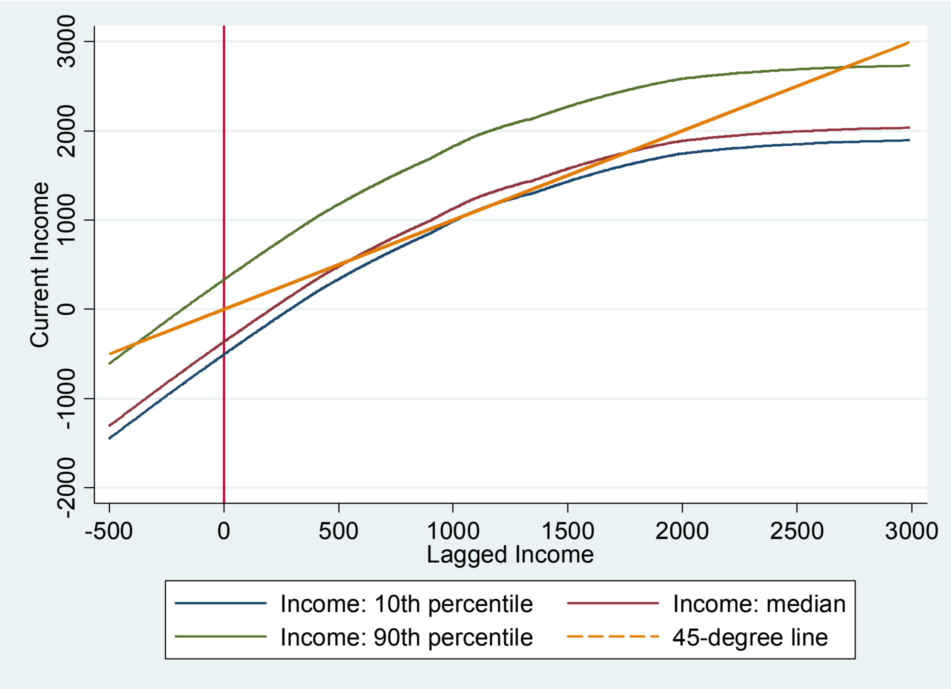
Panel B – With Year Dummies – Linear, Quadratic and Cubic with Year Dummies



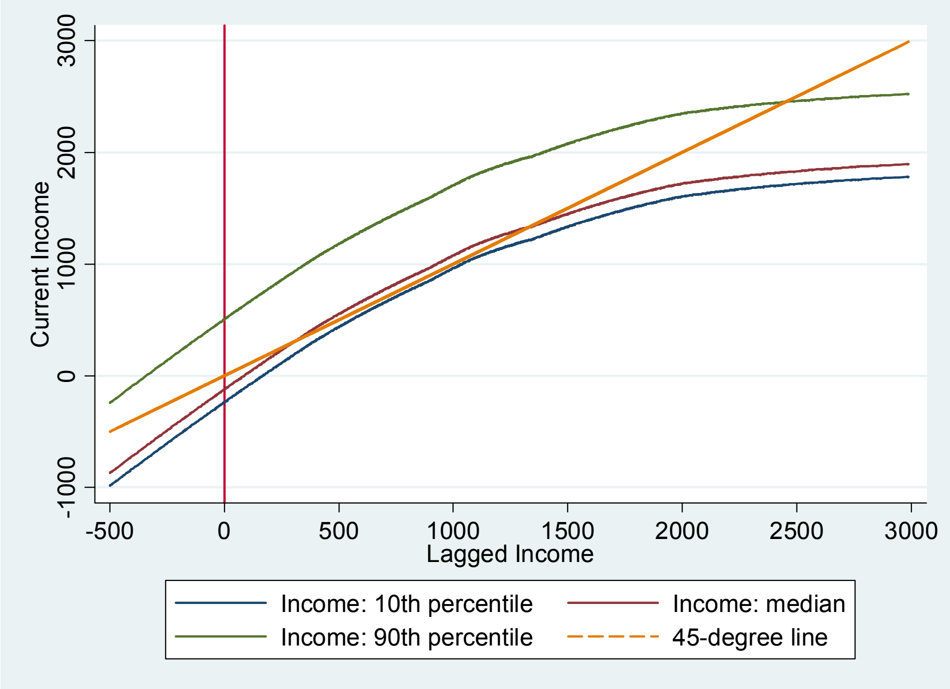
Note: Simulations of income dynamics have been computed using the estimated parameter of the polynomial of lagged income. Computed fixed effects act as shifters of the locus of polynomial. Simulation curves are depicted for a realistic range of lagged-income. Intersections between the dynamic trajectories and the 45-degree line indicate a potential equilibrium. Vertical line indicates zero income.

Graph A2: Parsimonious IV Model – Simulated Income Dynamics with Median, 10th percentile and 90th percentile of fixed effects Model Specification: Fuller Estimates with Two Instruments [(Rainfall x Land) and (Rainfall x Kids)] – Quadratic Polynomial Model with Year Dummies

Panel A – Fuller Estimates



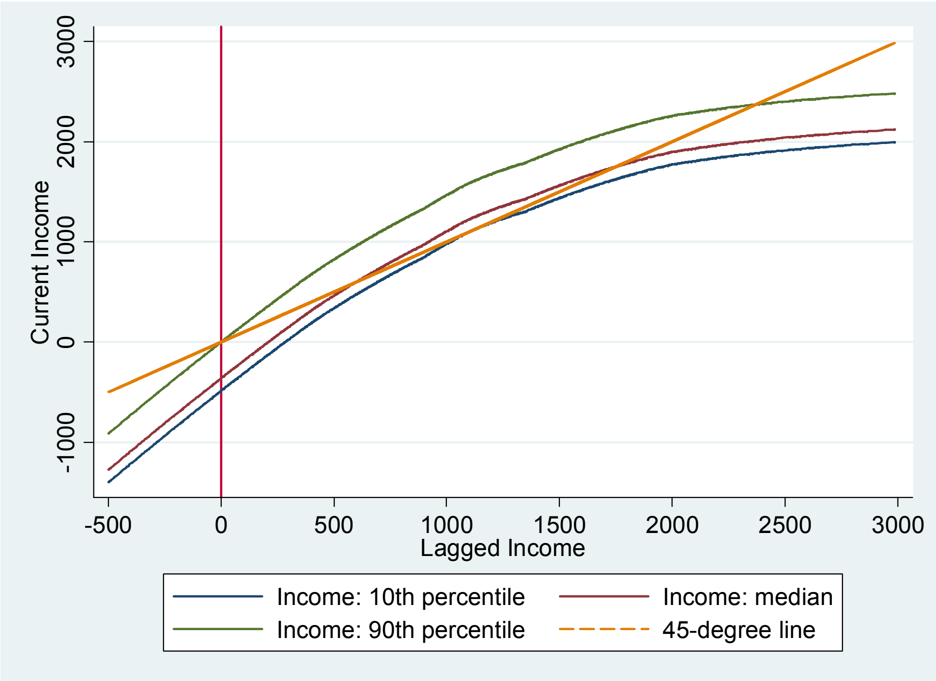
Panel B – Fuller Estimates with IPW Attrition Correction



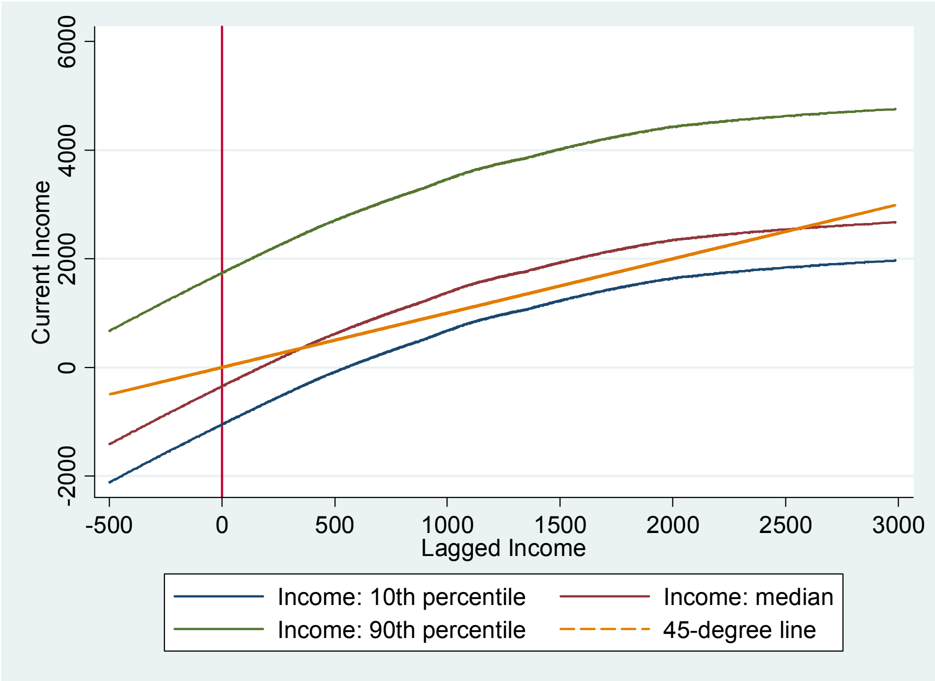
Note: Panel A reports Fuller estimates for the quadratic polynomial model with year dummies. Panel B reports Fuller estimates with Inverse Probability Weights correction. IPW weights are computed as the inverse of the predicted probability of attrition in a given wave. Individual, household and household head characteristics are used as predictors of wave attrition. Simulations of income dynamics have been computed using the estimated parameter of the polynomial of lagged income. Computed fixed effects act as shifters of the locus of polynomial. Simulation curves are depicted for a realistic range of lagged-income. Intersections between the dynamic trajectories and the 45-degree line indicate a potential equilibrium. Vertical line indicates zero income.

Graph A3: VLS1 and VLS2 Dynamics – Simulated Income Dynamics with Median, 10th percentile and 90th percentile of fixed effectsModel Specification: Fuller Estimates – Quadratic Polynomial Model with Year Dummies by Panel Period

Panel A – VLS1 (1975-1983)



Panel B – VLS2 (2001-2004)



Note: Simulations of income dynamics have been computed using the estimated parameter of the polynomial of lagged income. Computed fixed effects act as shifters of the locus of polynomial. Simulation curves are depicted for a realistic range of lagged-income. Intersections between the dynamic trajectories and the 45-degree line indicate a potential equilibrium. Vertical line indicates zero income. Estimation methodology drops two years from each panel period, 1975-76 and 2001-2002 in VLS1 and VLS2 respectively.

Table A1. Robustness Checks: 'Weak' IV Robust Estimators – TSLS LIML and TSLS Fuller

Dependent variable: ΔYt	IV LIML Estimates			IV Fuller Estimates		
	Linear Polynomial	Quadratic Polynomial	Cubic Polynomial	Linear Polynomial	Quadratic Polynomial	Cubic Polynomial
	(1)	(2)	(3)	(4)	(5)	(6)
ΔY_{t-1}	0.502424*	2.034827**	-1.200000	0.477534*	1.867273**	1.886554**
	(0.271)	(0.954)	(23.849)	(0.261)	(0.839)	(0.871)
ΔY_{t-1}^{2}		-0.000404*	0.003986		-0.000369*	-0.000458
		(0.000)	(0.032)		(0.000)	(0.000)
ΔY_{t-1}^{3}			-0.000001			0.000000
			(0.000)			(0.000)
ΔYear: 1977	60	-21.0000	-360.0000	65.0000	-7.4000	5.0278
	(-57.666)	(98.245)	(2,457.988)	(55.944)	(90.013)	(78.072)
ΔYear: 1978	-93	-1.6e+02**	-660.0000	-89.0000	-1.5e+02**	-1.4e+02**
	(-71.381)	(78.186)	(3,650.417)	(69.865)	(72.388)	(65.317)
ΔYear: 1979	8.126988	-22.0000	-70.0000	7.9915	-19.0000	-18.0000
	(-42.394)	(54.435)	(375.374)	(41.799)	(51.933)	(48.933)
ΔYear: 1980	-85	-59.0000	80.0000	-84.0000	-61.0000	-65.0000
	(-54.861)	(63.530)	(1,009.924)	(54.422)	(61.575)	(59.156)
ΔYear: 1981	88	120.0000	-14.0000	86.0000	110.0000	110.0000
	(-60.889)	(88.267)	(975.295)	(59.837)	(83.611)	(84.074)
ΔYear: 1982	1.4e+02***	1.3e+02*	330.0000	1.4e+02***	1.3e+02*	1.3e+02*
	(-43.105)	(73.068)	(1,471.271)	(42.286)	(68.516)	(69.459)
ΔYear: 1983	-2.2e+02***	-2.3e+02***	-400.0000	-2.2e+02***	-2.3e+02***	-2.2e+02***
	(-57.338)	(52.672)	(1,280.467)	(56.269)	(50.150)	(48.464)
ΔYear: 2003	-8.5e+02***	-8.5e+02***	1400.0000	-8.4e+02***	-8.4e+02***	-8.8e+02***
	(-208.418)	(259.896)	(17,000.000)	(204.824)	(245.629)	(249.309)
ΔYear: 2004	7.2e+02**	550.0000	-2000.0000	7.0e+02**	5.4e+02*	5.9e+02**
	(-303.245)	(341.497)	(20,000.000)	(295.831)	(321.720)	(286.611)
Observations	2260	2260	2260	2260	2260	2260
R-squared	-0.5354	-1.7527	-159.6338	-0.4945	-1.4715	-1.2559
First-Stage Diagnostics						
Number of Instruments	3	3	3	3	3	3
Cragg-Donald F-Statistic	15.7872	3.515	0.0028	15.7872	3.515	0.0028
F-Statistic - ΔY_{t-1}	7.9327	7.9327	7.9327	7.9327	7.9327	7.9327
F-Statistic - ΔY_{t-1}^{2}		8.0919	8.0919		8.0919	8.0919
F-Statistic - ΔY_{t-1}^{3}			7.0878			7.0878
Second-Stage Diagnostics						
Anderson-Rubin F stat	5.092	5.092	5.092	5.092	5.092	5.092
Prob > F	0.002	0.002	0.002	0.002	0.002	0.002
Hansen J Statistic (overidentification)	3.446	0.348	-	3.535	0.379	-
Prob > chi-squared	0.179	0.555	-	0.171	0.538	-

Note: Standard errors robust to heteroskedasticity and 'original' household-level clustering appear in parentheses. Income measured in 1975 Rupees per adult equivalent. Anderson Rubin F statistic tests the null hypothesis that the endogenous regressors are jointly insignificant in structural equation. Hansen J Statistic, reported for overidentified models tests null hypothesis that the excluded instruments are uncorrelated with the structural equation error. Cragg-Donald F-Statistics reported includes Kleibergen-Paap rank correction. Cragg-Donald 'Weak IV' statistic tests for the null hypothesis that instruments are strongly correlated with the set of endogenous variables.

Table A2. Robustness Checks – Parsimonious IV Set: Interactions (Rainfall x Land) and (Rainfall x Kids)

Dependent variable: ΔYt	Quadratic Polynomial Model		With Attrition Correction	
	GMM or LIML Estimators	Fuller Estimator	GMM or LIML Estimators	Fuller Estimator
	(1)	(2)	(3)	(4)
ΔY_{t-1}	2.120567**	1.926574**	2.155082	1.531708*
	(0.908)	(0.788)	(1.748)	(0.798)
ΔY_{t-1}^{2}	-0.000439**	-0.000395**	-0.000474	-0.000301
	(0.000)	(0.000)	(0.001)	(0.000)
ΔYear: 1977	-19.0000	-4.9000	-14.0000	17.0000
	(94.943)	(86.906)	(109.716)	(87.593)
ΔYear: 1978	-1.6e+02**	-1.5e+02**	-1.5e+02*	-1.2e+02*
	(77.627)	(71.703)	(77.972)	(63.890)
ΔYear: 1979	-25.0000	-22.0000	-36.0000	-20.0000
	(54.682)	(51.970)	(74.304)	(50.860)
ΔYear: 1980	-56.0000	-59.0000	-69.0000	-79.0000
	(64.502)	(62.088)	(79.734)	(64.260)
ΔYear: 1981	110.0000	110.0000	190.0000	160.0000
	(90.439)	(85.096)	(163.178)	(126.450)
ΔYear: 1982	1.3e+02*	1.3e+02*	71.0000	89.0000
	(75.813)	(70.433)	(129.007)	(92.980)
ΔYear: 1983	-2.3e+02***	-2.2e+02***	-2.1e+02***	-2.1e+02***
	(54.213)	(51.392)	(77.258)	(59.359)
ΔYear: 2003	-8.3e+02***	-8.2e+02***	-9.4e+02**	-9.0e+02***
	(272.697)	(254.679)	(446.218)	(328.488)
ΔYear: 2004	500.0000	490.0000	110.0000	320.0000
	(385.892)	(357.504)	(999.787)	(519.470)
Observations	8279	8279	8279	8279
R-squared	-1.9732	-1.6124	-3.5570	-1.6163
First-Stage Diagnostics				
Number of Instruments	2	2	2	2
Cragg-Donald F-Statistic	4.2201	4.2201	0.7536	0.7536
F-Statistic - ΔY_{t-1}	10.138	10.138	5.6812	5.6812
F-Statistic - ΔY_{t-1}^{2}	9.6768	9.6768	4.2166	4.2166
Second-Stage Diagnostics				
Anderson-Rubin F stat	7.216	7.216	4.784	4.784
Prob > F	0.001	0.001	0.009	0.009

Note: Parsimonious instruments are as follows: (Rainfall x Land Operated (Ha)) and (Rainfall x Nr Kids 0-8). Columns (3) and (4) report estimates with IPW attrition correction. Inverse probability weights are computed as the inverse of the predicted probability of attrition in a given wave. Individual, household and household head characteristics are used as predictors of wave attrition. Standard errors robust to heteroskedasticity and 'original' household-level clustering appear in parentheses. Income measured in 1975 Rupees per adult equivalent. Anderson Rubin F statistic tests the null hypothesis that the endogenous regressors are jointly insignificant in structural equation. Hansen J Statistic not defined since both models are only just-identified Cragg-Donald F-Statistics reported includes Kleibergen-Paap rank correction. Cragg-Donald 'Weak IV' statistic tests for the null hypothesis that instruments are strongly correlated with the set of endogenous variables.

Table A3. Robustness Checks: Parameter Stability between VLS1 and VLS2

Dependent variable: ΔYt	VLS1 - Quadratic Polynomial			VLS2 - Quadratic Polynomial		
	GMM	LIML	Fuller	GMM	LIML	Fuller
	(1)	(2)	(3)	(4)	(5)	(6)
ΔY_{t-1}	1.993613**	1.950957**	1.869939**	3.205948	3.340816	2.181551
	(0.885)	(0.941)	(0.886)	(3.820)	(3.985)	(2.086)
ΔY_{t-1}^{2}	-0.000393*	-0.000382	-0.000365*	-0.000574	-0.000587	-0.000412
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
ΔYear: 1977	-17.0000	-17.0000	-9.6000			
	(-87.513)	(90.805)	(86.885)			
ΔYear: 1978	-1.6e+02**	-1.6e+02**	-1.5e+02**			
	(-73.134)	(76.020)	(73.194)			
ΔYear: 1979	-21.0000	-20.0000	-19.0000			
	(-52.997)	(53.914)	(52.680)			
ΔYear: 1980	-62.0000	-61.0000	-62.0000			
	(-62.485)	(63.173)	(62.233)			
ΔYear: 1981	120.0000	110.0000	110.0000			
	(-84.251)	(86.130)	(83.896)			
ΔYear: 1982	1.3e+02*	1.3e+02*	1.3e+02*			
	(-69.139)	(70.708)	(68.526)			
ΔYear: 1983	-2.3e+02***	-2.3e+02***	-2.3e+02***			
	(-50.614)	(51.600)	(50.515)			
ΔYear: 2003				-1000.0000	-1000.0000	-8.9e+02**
				(-645.111)	(-673.000)	(-422.235)
ΔYear: 2004				930.0000	970.0000	640.0000
				(-1,005.532)	(-1,044.257)	(-608.382)
Observations	7717	7717	7717	562	562	562
R-squared	-1.4339	-1.3844	-1.2836	-4.5706	-4.8245	-2.2488
First-Stage Diagnostics						
Number of Instruments	3	3	3	3	3	3
Cragg-Donald F-Statistic	8.741	8.741	8.741	0.816	0.816	0.816
F-Statistic - ΔY_{t-1}	7.8407	7.8407	7.8407	1.382	1.382	1.382
F-Statistic - ΔY_{t-1}^{2}	8.435	8.435	8.435	1.7235	1.7235	1.7235
Second-Stage Diagnostics						
Anderson-Rubin F stat	6.078	6.078	6.078	2.034	2.034	2.034
Prob > F	0.001	0.001	0.001	0.112	0.112	0.112
Hansen J Statistic (overidentification)	0.332	0.332	0.336	0.0422	0.0411	0.0611
Prob > chi-squared	0.565	0.565	0.562	0.8373	0.8394	0.8048

Note: Standard errors robust to heteroskedasticity and 'original' household-level clustering appear in parentheses. Income measured in 1975 Rupees per adult equivalent. Anderson Rubin F statistic tests the null hypothesis that the endogenous regressors are jointly insignificant in structural equation. Hansen J Statistic, reported for overidentified models tests null hypothesis that the excluded instruments are uncorrelated with the structural equation error. Cragg-Donald 'Weak IV' statistic tests for the null hypothesis that instruments are strongly correlated with the set of endogenous variables. Estimation methodology drops two years from each panel period, 1975-76 and 2001-2002 in VLS1 and VLS2 respectively.

Table A4. Tests of Auto-Correlation in First-Differenced Income Series

		No Year Dummies			With Year Dummies		
		Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
		1	2	3	4	5	6
AR(1) Test	z-score	-11.84	-13.45	-14.44	-9.15	-11.95	-12.65
	p-value	0.00 ***	0.00 ***	0.00 ***	0.00 ***	0.00 ***	0.00 ***
AR(2) Test	z-score	-1.34	-1.33	-0.97	-1.42	-1.39	-1.05
	p-value	0.18	0.18	0.33	0.16	0.17	0.29

Note: Tests of auto-correlation are carried out on the household-level income series. We test for evidence of auto-correlation in the first-difference equation. AR(1) processes in the first-differenced residuals are to be expected but evidence of AR(2) processes, would suggest an AR(1) process in the levels equation.