Essays on Education and Work Choices in Developing and Developed Economies

Emla Olivia Anne Fitzsimons

A Dissertation submitted to the Department of Economics
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy
University College London

London

20th April 2004

Abstract

This thesis focuses on education and work choices of individuals in both developing and developed economies. Such economic settings are distinguishable partly in terms of the features and constraints underlying individual decision-making, which impact on the disparity in education levels between low- and high-income economies. In particular, low-income settings are characterised by a large degree of poverty, risk and uncertainty in daily life, the effects of which are often exacerbated by a lack of risk diversification mechanisms due to thin insurance markets and high borrowing constraints. These features have important implications for investment in education.

The higher levels of investment in education in developed economies are largely due to higher levels of wealth, opportunities and well-functioning institutions including markets for insurance and credit, along with the higher availability and quality of educational institutions. Even so, the severity of constraints right throughout the education lifecycle varies extensively across the income spectrum within developed economies. This is reflected in the lower levels of educational attainment of low-income individuals compared to individuals from richer backgrounds.

To begin with, I examine the effects of living in an environment that is inherently risky, on human capital accumulation in Indonesia. I then move on to an analysis of work and education decisions of children in rural Mexico, specifically considering potential interactions between the opportunity costs of education and the sibling composition of the child. Finally, I assess the impact of the Education Maintenance Allowance, a conditional education subsidy, on post-compulsory education and work choices in the UK.

Contents

List of Tables

Acknowledgements

I extend my sincere thanks and appreciation to Orazio Attanasio and Costas Meghir, my supervisors. When inspiration was lacking - as at times it was - their interest and optimism stirred new depths within me. My discussions with them were explorations, led primarily by their insight and expertise. Their advice was always useful and added immeasurably to this work. And as anyone who knows them will attest to, they were a lively and entertaining pair in the manner in which they engaged with each other in cultivating their ideas.

I thank everyone at the IFS for their help and support throughout my four years there. My friends and colleagues in the 'EEE' sector were great. Alissa and Lorraine were supportive and understanding, and would always involve me in discussions and projects on topics related to my work. Erich taught me much in the way of rigorous and methodical research practices and I thank him for his approachability and willingness to help. And things would not have been the same without Barbara, who was always full of chat and laughter, characteristics to savour when you're in the throes of academic pursuit.

The conversations and laughter and happy meals that I shared with Antonio and Renata will always have a special place in my thoughts of my PhD. When the pressure was on, and my spirits were low, they put up with me with such good humour that it was impossible not to be infected by it. I thank Eric, who was always there, no matter what, to listen, to advise and to help me, right throughout the four years. I thank Sandra for her friendship and support.

I am deeply indebted to Elena. Nothing, whether work-related or otherwise, was ever too much to ask of her and her patience and understanding are unmatchable. Her friendship is something to treasure.

My sisters and brothers must accept some responsibility for this work!

It is impossible to overstate how much I have learned from Cora and Lana, who have always led the way for me. They are the rocks that I depend on, the truest of friends. I thank Lana also for perceptive stylistic suggestions on various parts of the text. Lloyd and Ken have always been generous with their time and offer a solid and quiet support of a kind that is always welcome in this noisy world. I thank John for his interest and support, and for his words of advice and encouragement, offered with such kindness, at many critical moments. To Ka and Marina, who are always so caring and kind-hearted, I extend my warm thanks. When I was sitting at my London desk, late of a Sunday night, in the thick of econometrics, thoughts of each and every one of them would bring me home.

I am extremely grateful to the ESRC for the financial support it provided. The analysis in chapter two was also supported by the IFLS Small Grants Program, which is made possible by funding provided to RAND by the U.S. National Institute of Child Health and Human Development.

Most of all, I thank Mum and Dad. The day was not complete without Dad's telephone calls - nightly without fail - which, apart from keeping me up to date on all the goings on in Irish political and social life, were a great source of comfort and strength to me. I thank both of them from the bottom of my heart for all they have done and continue to do for me. I dedicate this work to them, with all my love.

Declaration

1. No part of this thesis has been presented to any University for any degree.

2. Chapter four was undertaken as joint work with Orazio Attanasio and Costas Meghir.

Emla Fitzsimons

Chapter 1

Introduction

In the forthcoming chapters, education and work choices are considered, both in less developed and developed economies. For reasons to be explained, the emphasis is on the former. The theoretical predictions of the standard pure investment model for education are applicable to both: in the absence of credit constraints, individuals accumulate the level of education that equates the marginal return from that level to its marginal cost, and household resources are extraneous to the decision.[1] However, the complete lack of credit constraints is clearly implausible, more so for less developed countries (LDCs), and it is important to consider the more realistic case of limited availability to borrow to invest in education. One implication of this is that in the absence of being able to borrow to augment human capital, reliance for funding is shifted to current income and other household resources. Varying degrees of capital market imperfections, coupled with varying levels of current income and resources across developed and less developed economies, lie behind much of the observed disparity in the education levels of individuals across such economies.[2] However, whilst sub-optimal educational investments are primarily attributed to capital market imperfections, and unequal levels of investment in developed and less developed countries

[1]Note that human capital is multi-dimensional and throughout, I focus on the education component.

[2]It is worth noting that even in the absence of credit constraints, current income matters if education holds a direct consumption value and/or if there is uncertainty of future earnings.

are largely explained by the varying severities of such imperfections, important insights into education choices can be gained by considering other factors that exhibit variation between economies. Differences in environmental constraints, for example, are apparent from the different gradations of availability and quality of schooling, the occurrence of and susceptibility to income shocks, the functioning of insurance markets and the availability of other less formal risk diversification mechanisms. Clearly, observed education choices are also determined by the perceived and actual returns to education, which also differ across economies. An extensive body of empirical research points to the high private and social returns to education in both types of economic environment.[3] However, the realisation of such returns is contingent on education being valued in the future occupation of the child, which is further dependent on such factors as the extent and availability of agricultural versus market work and gender-biased work opportunities. This point is closely related to deep-rooted cultural differences across economies that also have important implications for education and work choices, and again the propensity to invest in the education of children is partly predicated by such norms.

Therefore despite the common theme of education and work choices throughout the ensuing analyses, the focus of each chapter is influenced by the important underlying characteristics of households and the environments within which they dwell. Specifically, in developing economies I stress the importance of accounting for the inherent and persistent riskiness of the village-level environment, the availability of credit and insurance markets, the widespread availability of work as a (partial) substitute for schooling, along with household-level factors such as income variability and sibling composition. In the analysis of choices in the UK on the other hand, factors such as persistent risk and child labour are much less relevant in the decision-making process, and I consider decisions of low-income individuals concerning participation in post-compulsory education, along with subse-

[3]See for example Blundell et al (1999) for evidence on the private returns to education, Sianesi and Van Reenen (2003) for a review of the macro literature and Psacharopoulos (1985, 1994) for evidence on the returns to education in LDCs.

quent enrolment in university. Indeed this represents another key difference in the analyses, with the focus for LDCs on choices from primary school onwards, compared to the focus in the UK on decisions from post-compulsory school onwards. This is indicative of earlier declines in participation in education in LDCs compared to developed economies.[4]

A distinctive feature of LDCs that features to a large extent in the subsequent chapters, is the prevalence of child labour. The ability of children to generate earnings and to contribute to household income from an early age renders them a highly productive economic resource for indigent parents in the impoverished environments of LDCs, which manifests itself in the overall levels of child labour.[5] This has important implications for understanding decisions concerning investment in human capital. Child work activities encompass a variety of work of varying degrees of intensity, from part-time informal household chores and family work, to full-time formal work outside the household and - in the extreme - to hazardous and bonded forms of labour.[6] The long-term effects on child welfare and human capital accumulation are by no means uniform across the wide spectrum of activities. There is no clear evidence on just how detrimental it is to human capital accumulation and on the subsequent effects on the welfare of the child as an adult. This would require a long-term analysis of the effects of

[4]This is borne out in policies that are designed to promote participation in education via conditional subsidies. Post-compulsory participation in education along with increased university enrolment are the key targets of the Education Maintenance Allowance programme in the UK. The Progresa programme in Mexico is instead aimed at fostering an increase in the transition of poor rural youth into junior secondary school and at the further continuation of children in secondary education.

[5]I abstract from all definitional complexities surrounding the term 'child labour'. It should be noted however, that this elusive term must be refined to deal with issues such as the age of a 'child' and what exactly constitutes 'labour'. The International Labour Organisation, defining a child as a person between the ages of five and fifteen, and labour as full-time and part-time work, estimates the number of child labourers at 250 million worldwide, practically all of whom are situate in developing economies.

[6]There is very little empirical work on hazardous forms of child labour, or on the activities of homeless children on the streets, as data on such activities are extremely sensitive and difficult to collect.

working as a child, which are difficult to quantify.[7] However, it is generally agreed that even less severe forms of child labour such as working on the family farm, have, on the whole, adverse long-term effects on human capital accumulation.[8] Consequently there are compelling economic arguments advocating the implementation of policies to reduce child labour, including the positive private returns to education and the spillovers to society from increased human capital accumulation.[9] These are quite apart from humanitarian concern for the health and welfare of children. A brief description of each chapter follows.

In chapter two I study the effects of living in a risky environment on education and child labour in LDCs. The central idea is that households that face more uncertainty, and with limited or no access to formal insurance, have a higher motive for self-insurance and this may, under a set of plausible assumptions, have adverse effects on levels of investment in education. The model predictions are tested using Indonesian data. A negative effect of risk on education would constitute some evidence of children being used as insurance tools to smooth consumption. On the other hand, whilst a negligible effect of risk may indicate that formal insurance markets are well-functioning, it might also reflect the fact that households are using a wide range of other self-insurance mechanisms instead. A key contribution of the chapter is the decomposition of risk into aggregate (village-level) and idiosyncratic components using a unique measure of risk based on five years of wage data on the main earner of the household. Results indicate that in small rural villages where one might expect formal insurance markets to be thin or lacking, idiosyncratic risk has no significant effect on the child's education. There is evidence however, that aggregate village risk affects ed-

[7]See Basu (1999) for an analysis of the dynamic effects of child labour.

[8]However casual types of labour such as working on the family farm during harvest time, may simply lengthen the process of human capital accumulation, without having an adverse effect on the end stock. Indeed work may have adverse effects on schooling, but without work, many children may not be able to attend school at all. However, the evidence points towards child work showing a high degree of persistence, making transition back to school problematic. See Freije and Lopez-Calva (2001) and Guarcello et al (2003).

[9]See for example Weale (1992).

ucation adversely in these villages. These findings are in line with a range of literature which shows that aggregate risk is more difficult to diversify than idiosyncratic risk.[10] This suggests that policy should be carefully crafted in order to provide insurance for households against pervasive income risk, whilst at the same time ensuring that household-level informal insurance mechanisms are not crowded out.

In chapter three I investigate a number of aspects of education and work choices across children within households in rural Mexico. To begin with, I examine the associations between observed activities and the structure of the sibling set of children, taking household size and sibling structure as exogenous. In addition, I also incorporate an analysis of the effects of the child wage on such choices and allow for the (opportunity) costs of schooling to depend on the number of siblings by interacting the wage with sibling composition. In line with previous literature in this area, I find that larger households are associated with less schooling, particularly so for large numbers of very young siblings and/or older brothers.[11] The analysis of schooling costs yields some important new insights, with evidence that the responsiveness of children to wages increases in accordance with the number of the child's siblings. This is particularly so the older the child.

In chapter four I specify and estimate a dynamic discrete-choice multiple-state model for young adults in the UK, some of whom are entitled to a conditional post-compulsory education subsidy, the Education Maintenance Allowance (EMA). I model choices amongst education, full-time work, part-time work and unemployment. The estimation methodology controls for educational selectivity through time along with an extensive range of long-term family and parental background characteristics, thus allowing for heterogeneity and diversity in the EMA population in both observed and unob-

[10] Townsend (1994) presents evidence that whilst agents are successful in insuring against non-covariant (idiosyncratic) forms of risk, pervasive uncertainty is more difficult to insure against. Rosenzweig and Binswanger (1993) similarly find evidence that common shocks appear to have substantially greater consequences for consumption than does idiosyncratic risk, with comparable findings by Udry (1994).

[11] See for example Parish and Willis (1993) and Patrinos and Psacharopoulos (1997).

served dimensions. I also model attrition from the panel through time. This is important as attrition will introduce selection bias if there is a correlation between unobserved characteristics that affect attrition and unobservables that affect the individual's choices. Therefore the effectiveness of the subsidy is not confounded with either dynamic selection bias or attrition bias. Simulations from the estimated model are used examine the sensitivity of schooling choices to education costs. The overall effect on education in response to the subsidy is positive. The proportions in full-time education with a job decrease, but there is a relatively higher increase in participation in full-time education without a job. Increasing the generosity of the subsidy is likely to increase participation in education even further. There is, however, almost no effect of the programme on subsequent university participation.

Chapter 2

The Effects of Risk on Education and Child Labour in Indonesia

2.1 Introduction

In light of the widely documented disparities in the levels of child labour between high and low income economies, it is natural to suppose that observed divergent decisions on the use of a child's time are largely the result of the incongruent economic settings underlying economic choices in both types of economy. In particular, low-income settings are characterised by a large degree of risk and uncertainty in everyday life, and it thus seems reasonable to expect income risk to play an important role in shaping household economic choices. The ability of households to deal with such risk and to smooth consumption across time may be constrained by thin insurance markets for income and higher borrowing constraints, thus cutting off important risk diversification channels. These distinguishing features of low-income settings create the need for households and villages to form alternative ways of coping with uncertainty.[1]

[1] There is an extensive literature that examines the importance of the family unit in coping with uncertainty, and the incorporation of risk into the economic choices and behaviour of households. See for example Rosenzweig (1988), Rosenzweig and Stark (1989),

Within such environments, it is widely accepted that child labour is strongly associated with household poverty. Indeed the relative importance of a variety of observable individual, household and area characteristics, all of which measure poverty to some degree, for child education and work decisions, is widely documented.[2] However, this association is very general, masking as it does the underlying (often unobserved) channels through which poverty translates into lower investment in human capital and higher levels of child labour. Indeed whilst it captures a contemporaneous association between child labour and household resource constraints, it does little to aid an understanding as to the precise reasons why poverty matters. Risk and poverty are strongly correlated, as households with fewer resources are less likely to be able to cope with adverse shocks. However, the effects of risk have been less widely examined, due to the difficulty in quantifying risk. A growing literature examines the specific role of a child as an ex-post mechanism for smoothing out income shocks. Much of this literature finds that unanticipated shocks have positive effects on child labour.[3] Whilst this goes some way towards isolating the key channels through which poverty affects child labour, it is however more informative on the extent to which households face liquidity constraints such that the lack of borrowing opportunities in the event of a shock, may force them to send the child to work instead.

Paxson (1992) Rosenzweig and Binswanger (1993) and Kochar (1995).

[2]A number of characteristics such as parental education and work status, household income, family size, the gender of the household head, ownership of land and school availability, has consistently emerged across studies. See the empirical analyses of the effects of current indicators of household and individual welfare on child labour in Grootaert and Kanbur (1995), Jensen and Nielsen (1997), Patrinos and Psacharopoulos (1997), Duraisamy (2000), Ravallion and Wodon (2000), Ray (2000), Freije and Lopez-Calva (2001), Ejrnaes and Pörtner (2002) and Emerson and Souza (2002).

[3]See Jacoby and Skoufias (1997), Pörtner (2001), Ranjan (2001), Sawada and Lokshin (2001) and Beegle et al (2003). The role of a child as an insurance tool against unforeseen circumstances was proposed by Cain (1982), and work by Grootaert and Kanbur (1995) discusses how child labour may be part of a strategy to minimise the risk of interruption of a household's income stream.

In this chapter, I consider child labour more specifically as part of an intra-household strategy to diversify risk in an uncertain environment.[4] Rather than viewing children as an ex-post mechanism for smoothing consumption, I consider the cumulative effects of merely living in a (perceived) risky environment, on child human capital. The idea is that the riskier the environment, the greater is the incentive of economic agents to build up a buffer stock to cushion against unforeseen adverse events. Insofar as the child is viewed as a liquid economic resource with immediate earnings potential, (s)he can work and contribute to the stock of precautionary savings of the household. The motive for building up a buffer stock will be higher, the less well-functioning are formal mechanisms, such as insurance and credit markets, for protecting against unanticipated shocks.[5] Thus evidence that risk affects child labour in this framework, is indicative of poorly functioning formal insurance markets. To my knowledge, this analysis is among the first to explicitly examine the extent to which living in an intrinsically risky environment affects child education and work choices. Whilst Sawada and Lokshin (2001) show the theoretically negative effect of income instability on investment in education of children, via a precautionary savings motive, their empirical analysis deals with the occurrence of actual household shocks.

I use Indonesian data to test whether children are used to build up buffer stocks in risky environments. I distinguish between risk that is specific to the household (idiosyncratic risk) and risk that is pervasive within a village (aggregate village risk).[6] I find important evidence that children in households facing higher village-level risk do indeed have lower educational

[4]Of course household diversification mechanisms interact with each other, with child labour just one possible strategy amongst a host of others. Ideally one would like to take into account the whole range of possible risk-reducing mechanisms of the household, clearly an onerous task.

[5]As Morduch (1995) discusses, the general consensus regarding insurance markets is that even if household income is partly insurable, as is most likely the case, full insurance is highly unlikely in LDCs.

[6]The distinction between the two is important. A wide body of empirical research underlines the importance of decomposing risk into idiosyncratic and aggregate components, given the range of evidence on the differing responses of economic agents to both types of risk. I refer the reader to footnote (10) in chapter one.

attainment than their counterparts in low-risk environments. To the extent that labour is a substitute for schooling, this translates into higher child labour in these households. The effect is strongest for 10 to 12 year olds, which is the age range that is likely to be most sensitive to severe economic constraints. I find no evidence of idiosyncratic risk affecting children's education. These findings are indicative of pervasive village risk being more difficult to insure against than idiosyncratic risk, and provide some insight into the functioning of insurance markets in these villages. In particular, whilst household-level risk is being diversified away, whether through formal or informal mechanisms, without resorting to children, evidence that aggregate village risk affects education, propels an argument for favouring intervention in the market for insurance against such risk.

The remainder of the chapter is structured as follows: in section 2.2 I outline a simple two-period model of investment in human capital in a risky environment. I show that under certain plausible conditions, investment in human capital is negatively related to the degree of earnings risk facing the financier of the child's schooling (the parent). The theory does not however, rule out possible offsetting positive effects of risk on education, and these are also discussed. In section 2.3 the Indonesian data used in the empirical analysis is described. In this section I define risk more clearly and show how I measure both idiosyncratic and aggregate village risk. Section 2.4 describes the main results and section 2.5 concludes.

2.2 Theoretical Framework

The theoretical effects of a volatile parental income stream on child education are considered within a two-period model. In period 1, a household consists of one parent and one child.[7] This period corresponds to the po-

[7]This enables us to abstract from issues concerning (a) intra-household bargaining amongst parents concerning the child's activity (see Galasso (1999) and Basu (2001)) and (b) compensating and reinforcing human capital decisions amongst siblings (see Becker (1991)). Their incorporation into a formal model and their empirical importance, are topics for future research.

tential schooling phase of the child, throughout which the parent works and earns an exogenous amount of income, y_1^p. Decisions on the use of the one unit of child's time in period 1 are made by the parent. A child may work, go to school, or engage in some combination of the two. This decision is made jointly with the household choice of consumption. In period 2, the child has become an adult and has formed his own household (thus the term 'adult' refers to the offspring in period 2). He earns income, y_2^a, which is an increasing function of human capital. The parent earns an exogenous amount of income, y_2^p.

Within this framework, I lay out one set of plausible conditions that is consistent with risk adversely affecting education. The first assumption is that human capital investments are motivated by parental concern for their offspring. This is modelled by making the adult's utility an argument of the parent's utility function. I assume that there are no transfers from the parent to its adult offspring in the second period, and that the parent leaves no bequests to the child. This rules out the parent compensating for any lack of investment in human capital of the child.[8] Transfers from the offspring to the parent in the second period are also ruled out, in order to capture a realistic intergenerational commitment problem.[9] This may be viewed as a strong assumption, in the sense that one important reason for investing in human capital in low-income countries, is an old age security motive of the parents. This assumption could be relaxed to allow for *ad hoc* transfers from the adult to the parent in the second period, but the crucial assumption is that these transfers are not an enforceable (by the parent) repayment for the parental investment in education.[10] Below, I consider the effects of allowing for transfers from the adult to the parent. Finally, borrowing to invest in

[8]Whilst this is clearly an extreme assumption (see Rosenzweig (1988) for evidence on the importance of intergenerational transfers in developing countries, and Altonji et al (1997) for the US), the simplification is made in order to focus on altruism that acts only through human capital investment.

[9]Baland and Robinson (2000) highlight the importance of this inter-generational problem.

[10]To further justify this assumption, I present evidence in section 2.4 that intergenerational transfers in Indonesia are not increasing in the education level of the donor.

human capital is not permitted, which means that parents must bear all of the costs of their child's schooling, both direct and opportunity in the form of foregone child earnings.[11]

A central feature of the model is uncertainty about future labour income. It is the only source of uncertainty considered throughout and it may arise in two possible ways. In the first, the parent makes its choices about investment in education and household consumption at the beginning of period 1, without knowing its income stream for the following period, y_2^p. However, the parent knows its average income over time, μ, and the variability of its income stream, σ^2, which is general in the sense that it incorporates all past shocks to income, both at aggregate village and idiosyncratic levels.[12] The more volatile its past income, the more uncertain it is about next period's income, y_2^p, and the higher the motive for self-insurance. Households are heterogeneous in the sense that even if their current observed income levels are the same, some have had very volatile past incomes whilst others have had very smooth income draws. Such heterogeneity reflects different risk profiles of households. A second source of potential uncertainty is adult earnings in period 2, y_2^a, which are a function of the fraction of the child's time spent in school in period 1. For simplicity however, I assume that conditional on education, parents anticipate adults' expected earnings with certainty. This assumption is made in order to focus on the effects of uncertain future parental income on child education. The theoretical effects of relaxing this assumption are considered briefly below.

Parental utility in period 1 is a function of household consumption only, and in period 2 is a function of both own consumption and the utility of their offspring, subject to an altruistic parameter. The household chooses consumption and education in period 1 so as to maximise lifetime utility subject to the constraint that the present value of consumption is equal to the present value of income. Assuming that preferences are intertemporally

[11]See Ranjan (1999) who uses a simple two-period model to show how the non-existence of markets for loans against the future earnings of children may give rise to child labour.

[12]The empirical work distinguishes between the two forms of risk.

additive, the parent's problem is to

$$\max_{c_1^{hh}, D_1} U(c_1^{hh}) + \beta E_1 U(c_2^p) + \beta\gamma U(c_2^a) \tag{2.1}$$

where c_i refers to consumption in that period, $i = 1, 2$, the superscripts hh, p and a refer to the household, parent and adult respectively, D_1 is the fraction of child time spent in school in period 1, $\beta = \frac{1}{1+r}$ is the parental discount rate, r is the interest rate, $0 < \gamma < 1$ measures the weight the parent places on the adult's utility and converts the utility of the adult into that of the parent, the expectations operator, E_1, reflects uncertainty (as at time 1) about y_2^p, and the individual period subutility functions are increasing and concave in their single arguments. The concavity assumption is equivalent to the assumption that the parent is risk averse. From (2.1) it is clear that parents derive no utility from education *per se* and care about the end stock of human capital only insofar as it contributes positively to the consumption of their offspring in period 2, thus motivating parental investment in human capital.

The human capital of the child is produced according to the following increasing and concave production function

$$H_1 = g(D_1), \quad \text{where} \quad g'(D_1) > 0, \quad g''(D_1) \leq 0 \tag{2.2}$$

where H_1 is the child human capital stock at the end of the schooling period.

Period 2 earnings of the adult are an increasing and concave function of their end stock of human capital

$$y_2^a = f(H_1), \quad \text{where} \quad f'(H_1) > 0, \quad f''(H_1) \leq 0 \tag{2.3}$$

The life-cycle budget constraint of the parent is

$$Y_L = c_1^{hh} + c_2^p + (p^D + w_1^c) D_1 \tag{2.4}$$

where Y_L is the present value of the lifetime income of the parent, assuming that the labour market earnings of the child when young are pooled with parental resources. Thus Y_L may be written as $Y_L = y_1^p + y_2^p + w_1^c$, where y_i^p equals the exogenous parental earnings in period i, w_1^c is full child income in period 1[13], and p^D is the direct cost of schooling.

Rewriting (2.4) in terms of c_2^p and substituting into (2.1), the parent's problem is

$$\max_{c_1^{hh}, D_1} U(c_1^{hh}) + \beta E_1 U\Big([Y_L - c_1^{hh} - (p^D + w_1^c)D_1]\Big) + \beta\gamma U(c_2^a) \tag{2.5}$$

2.2.1 Education Choice

From (2.5) the first order condition for D_1 is

$$(p^D + w_1^c)E_1[U'(c_2^p)] = \gamma[U'(c_2^a)f'(g(D_1))] \tag{2.6}$$

Equation (2.6) shows that the utility-weighted expected marginal cost of schooling to the parent equals the utility-weighted marginal benefit of additional earnings to the adult in period 2, as a result of schooling. From this first order condition, it can be seen that the risk in second period parental income will affect the education decision in the first period insofar as it affects the expected second period marginal utility of the parent.

Equation (2.6) is used to solve for D_1^*, the optimal level of education in period 1. The problem is complicated by the presence of the expectations operator over the parent's marginal utility of consumption in period 2. There is in general no closed form solution for this problem when labour income is uncertain, with the solution for optimal education choice largely depending on the form of the utility function.

Parental Income Risk

In order to isolate the effect of future parental income risk on c_1^{hh} and D_1, I follow Sandmo (1970) who defines a pure increase in dispersion as a stretching of the distribution of income around a constant mean. This is a combination of additive and multiplicative shifts in the distribution of parental

[13]Note that similar to previous authors (see for example, Jacoby and Skoufias (1997)), I assume that the child wage is not a function of human capital.

income: the additive shift increases the mean whilst holding all other moments constant, and the multiplicative shift stretches the distribution on the right side of zero (assuming income is non-negative).[14] We can thus think of the expected parental income in period 2 as

$$E\left[\delta y_2^p + \theta\right] \tag{2.7}$$

where $\delta > 1$ is the multiplicative shift which increases y_2^p, and θ is an additive shift.

In order for the increase in risk to be mean-preserving, it must be the case that the change in the expected value of future parental income is 0, i.e.

$$\begin{aligned} dE[\delta y_2^p + \theta] &= E[y_2^p d\delta + d\theta] = 0 \\ &\Rightarrow \; d\theta/d\delta = -E[y_2^p] = -\xi \end{aligned} \tag{2.8}$$

The effects of such future income uncertainty on savings and consumption choices have been examined extensively in the literature.[15] The key findings in this extensive literature show that under certain restrictions on household preferences, uncertainty about future income decreases current consumption through increasing expected future marginal utility relative to current marginal utility. In line with this research, it can be shown here that under the assumption of decreasing absolute risk aversion (DARA) in c_2^p, $\frac{\partial c_1^{hh}}{\partial \delta}|_{d\theta/d\delta=-\xi} < 0$. However, the effects on education choice have not been formalised in this literature. Taking expectations of (2.6), the first order condition for education is

$$(p^D + w_1^c)\int_\delta U^{'}(c_2^p)f(y_2^p)d\delta = \gamma[U^{'}(c_2^a)f^{'}(g(D_1))] \tag{2.9}$$

As shown in the appendix, the effect of a mean-preserving increase in risk on investment in education depends on the sign of

$$\left.\frac{\partial D_1}{\partial \delta}\right|_{\frac{d\theta}{d\delta}=-\xi} = \beta(p^D + w_1^c)U^{''}(c_1^{hh})E_1[U^{''}(c_2^p)(y_2^p - \xi)] \tag{2.10}$$

[14]In all of what follows, derivations follow closely on Sandmo (1970) and are detailed in the appendix to this chapter.

[15]See the seminal papers by Leland (1968) and Sandmo (1970) and more recently, Deaton (1992).

Under the assumption of decreasing absolute risk aversion, the above term is negative for all values of y_2^p, which implies that $\partial D_1/\partial\delta|_{d\theta/d\delta=-\xi} < 0$.

It is clear from (2.10) that the more risk averse the parent, the more adverse is the effect of risk on education choice; the same is true the higher the costs of schooling. Note that in the extreme case where $(p^D + w_1^c) = 0$, risk has no effect on education. Clearly, changing education in this case would have no effect on savings as education is costless anyway.

To summarise, under the above conditions, pure risk will never have a positive effect on education, D_1. Uncertainty leads the parent to increase precautionary savings in period 1 in order to maintain a smooth marginal utility of consumption profile across the two periods. Given the assumption that the adult's future environment conditional on education is deterministic, there is no reason for the parent to increase D_1 in response to increased risk of y_2^p, as D_1 does not cushion against y_2^p shortfalls.

Apart from parental prudence and not being able (or not wanting) to borrow to finance investment in education, this unambiguous negative effect of risk on investment in education is driven by two important factors. The first is the assumption of no transfers from the adult to the parent in period 2. If such transfers were allowed, it is plausible that education might increase in response to increased parental income uncertainty. This is because in the event of the parent receiving an unexpectedly low income draw in the second period, the adult could make transfers to the parent which are an increasing function of adult income and therefore of education.

The second reason for this unambiguously negative effect, is the assumption of a deterministic environment for the adult in period 2. Clearly this is unrealistic in the sense that there is no reason to believe that future labour market conditions for the adult, conditional on education, would be known with certainty, whilst the parent would face a stochastic future environment. It is more reasonable to expect such uncertainty to also affect future adult earnings, differentially across education levels, and therefore the education choice. Below, the theoretical effects of relaxing these assumptions are discussed.[16]

Reverse Altruism

If we assume that the adult makes transfers to the parent in period 2 (and that the parent anticipates this), the effects of income risk on child education become ambiguous. There are now two competing effects of risk: on the one hand, parents may use child labour as a buffer against short-term income shortfalls, whilst on the other, if transfers are an increasing function of education, they will invest more in their child's education in order to receive more transfers in the event of a possible future income shock. Allowing $U(c_2^a) = V(c_2^a) + \lambda U(c_2^p)$, where $0 < \lambda < 1$ represents adult altruism towards the parent in period 2, the parent's problem in period 1 is now to

$$\max_{c_1^{hh}, D_1} U(c_1^{hh}) + \beta(1+\gamma\lambda)E_1 U\Big([Y_L - c_1^{hh} - (p^D + w_1^c)D_1 + T(D_1)]\Big) \\ + \beta\gamma E_1[V(c_2^a) + \lambda U(c_2^p)] \tag{2.11}$$

where $T(D_1)$ are transfers from the adult to the parent in period 2, and are increasing in D_1.

The first order condition for education is

$$(1+\gamma\lambda)[(p^D + w_1^c) - T'(D_1)]E_1[U'(c_2^p)] = \gamma[V'(c_2^a)f'(g(D_1))] \tag{2.12}$$

and the effect of a mean-preserving increase in risk on education choice depends on

$$\left.\frac{\partial D_1}{\partial \delta}\right|_{\frac{d\theta}{d\delta}=-\xi} = \beta(1+\gamma\lambda)[(p^D + w_1^c) - T'(D_1)]. \\ U''(c_1^{hh})E[U''(c_2^p)(y_2^p - \xi)] \tag{2.13}$$

[16] It is worth pointing out that whilst the focus here is on education, the model is very general in the sense that it can be applied to any type of investment good that may be used to buffer consumption, and the effects of risk on the accumulation of such goods can similarly be considered. The focus on education, however, underlines how the lack of adequate insurance mechanisms may perpetuate long-term poverty.

the sign of which depends on the relative magnitude of education costs, $p^D + w_1^c$, and the effect of education on transfers, $T'(D_1)$, and is ambiguous overall. As one would expect, if $p^D + w_1^c > T'(D_1)$, it is negative.[17]

Adult Income Risk

The assumption of deterministic earnings for the adult in period 2, conditional on education, may also be relaxed. The effects of own future income uncertainty on education choice have been examined by various authors, and centre around the important works of Levhari and Weiss (1974), Eaton and Rosen (1980) and Kodde (1986). I allow adult income to be a function of education and the stochastic variable that captures the *a priori* unknown state of the world, $\delta : y_2^a = f(H_1; \delta)$. In the first order condition for education, this additional source of risk is reflected in the right hand side expectation in (2.14)

$$(p^D + w_1^c)E_1[U'(c_2^p)] = \gamma E_1[U'(c_2^a)f'(g(D_1); \delta)] \tag{2.14}$$

It is in general not possible to theoretically determine the response of education to an increase in future income risk (see Kodde (1986)). The effects depend on the way in which risk is incorporated into the earnings function. Whilst the (negative) effects on education of special cases of additive and multiplicative forms of risk may be determined, the randomness of human capital returns may be modelled within more general frameworks, rendering it possible that investment in education may increase in response to risk. Thus risk may be either increasing or decreasing in education level and extending the model to allow for a stochastic element in adult income provides a possible mitigating effect of the negative precautionary savings effect, on education.

Evidence as to the plausibility of these two assumptions in the empirical setting, is presented in section 2.4.

[17]In the above, it is assumed for simplicity that transfers are given exogenously by the adult in period 2. See Raut and Tran (1998) and Baland and Robinson (2000) for a more complete analysis of intergenerational transfers and the effect on child labour.

Thus the model suggests that if we observe two households at a particular point in time with the same average income levels, what is also important to observe, is the variability of the income stream across time. This is because if one household's past earnings stream is very stable relative to the others, this is likely to have differential impacts on their economic choices. The household with the more variable income stream will have a higher precautionary savings motive than the former, and as outlined in the model, this may have adverse effects on household education choices.

2.3 Data

The data used in the analysis is the first wave of the Indonesian Family Life Survey (IFLS) data, conducted in 1993. The IFLS is a collaborative effort of *Lembaga Demografi* of the University of Indonesia and RAND. It is an ongoing multi-purpose longitudinal survey providing a broad array of data at the individual and household level on fertility, health, education, migration and employment. It encompasses over 30,000 individuals in 7,224 households, spread across 13 provinces in Indonesia. The individual and household data is accompanied by extensive community data that can be linked to individual households. This includes detailed information on transportation, communications, agriculture and industry, credit opportunities, community development activities, and the availability of schools and health facilities. This is extremely advantageous given the well-documented importance of village characteristics on education choices (which is indeed borne out in the analysis below). There is data on a total of 321 communities. However, these regions are all quite diverse, encompassing a large array of individuals and households, and varying greatly in size, with some even resembling large urban sprawls and others resembling close-knit communities.[18] In order to examine the effects of risk on investment in education, ideally one would like to be able to distinguish between villages in which formal insurance mechanisms exist and those in which they do not. The sample could then be split on this basis, on the assumption that in the latter, households have more

[18]There are on average over 3,000 family heads in urban areas compared to just under 1,000 in rural areas.

of a need to self-insure. However, unfortunately comprehensive information on the availability of insurance within villages is not observed. Instead, the analysis is focused throughout on small rural villages, which are chosen as those rural areas with less than 1,000 households, and represent just under 29% of the overall IFLS sample. The emphasis on these areas is an attempt to restrict the analysis to areas where formal insurance is most likely to be thin, and where self-insurance mechanisms are more likely to be reflected in observed behaviour.[19] In addition, the focus on rural households will most likely capture those agents living in more intrinsically risky environments.

Selecting the sample of interest is not wholly innocuous. In the first instance, the sample is chosen on the basis of the age range corresponding to primary education: children aged 7 to 12. This is because after age 12, leaving school is relatively common, and it thus seems reasonable to expect education choices amongst the 7 to 12 year old range to be more closely driven by severe household economic constraints, with that for the 13-plus range to be more culturally acceptable.[20] In the second case, I define the sample according to their likelihood of working. Surveys by Asra (1993, 1996) suggests that the work of children under 10 years old is very small in Indonesia compared to those aged 10 plus. Indeed in the IFLS, data on work activity in the week before the interview is only asked of individuals aged 10 plus.[21] Amongst 10 to 14 year olds, those observed to be working

[19]See Besley (1995a, 1995b) for evidence that formal rural credit markets still remain highly imperfect in low-income economies.

[20]Primary education in Indonesia is free, compulsory, and almost universal. School enrolment drops for both males and females at the end of primary education (around the age of 12). This feature of low continuation from primary to secondary school, with the largest loss occurring after the completion of primary school, has been noted before as an important failure of the Indonesian education system (see Manning (2000)). Relatively high dropout rates from primary school have also been observed, with 20% dropping out before completion of grade 6. Efforts to increase the availability of secondary education have been significant in recent years, with a current secondary school enrolment rate of just under 50%.

[21]If the week before the interview is a school holiday, individuals are asked about their main activity during term time. This can be work, look for work, housework, school or other.

are indeed full-time child labourers, with only 4.6% of them also attending school. Therefore I also carry out the analysis on individuals aged 10-14 years, who are more likely to be viewed as direct income-generating assets to the household, compared to younger persons. The upper cutoff of age 14 in this instance, is consistent with the ILO definition. In addition in Indonesia, the minimum legal working age is 15 years.[22]

Table 2.1 displays characteristics of households and individuals in small rural villages vis-á-vis all other regions (characteristics for households with at least one 7-12 year old are very similar and are therefore not reported). Notable differences in small rural villages include a substantially higher proportion of households reporting farm ownership and a lower proportion of business owners. In addition, banks are less likely to be present in small rural areas. This is interesting as it may in some sense represent a proxy for borrowing and insurance opportunities, and is reassuring in the sense that the analysis is being focused on the areas of interest - where formal credit and insurance markets are likely to be thin.

I observe whether the individual is currently attending school, for 1,366 7 to 12 year olds across 79 villages. For 1,136 10 to 14 year olds, their work/schooling status is observed, again across 79 small rural communities. Summary statistics in table 2.2 show that approximately 81.2% of this sample is enrolled in school, with almost 12% either working or looking for work. The corresponding school enrolment rate amongst 7-12 year olds is 80.8%.

However, it is likely that risk affects not only the *current* schooling status of the child, but also his accumulated educational attainment. By focusing only on current educational status, account is not necessarily taken of all past temporary interruptions to schooling, which are likely to be an important reaction to uncertainty. For this reason, I also focus the empirical work on the child's current years of education, which represents the accumulated effect of previous education decisions and the current stock of

[22]The relevant ILO Convention was ratified by Indonesia in 1999.

human capital of the child. This outcome will capture both those who have been permanently withdrawn from education, and those whose progression through school has been affected by temporary withdrawals, even if they are currently observed to be attending school. The average number of years of education for 7 to 12 year olds is 2.7 years, compared to 4.3 years for 10 to 14 year olds.

2.3.1 Measuring Uncertainty: Idiosyncratic and Aggregate Risk

Whilst the theoretical predictions of the model point towards a possible channel through which uncertainty will adversely affect investment in education, risk in this model is quite broadly defined. A key distinction in the empirical measurement of risk, is its decomposition into idiosyncratic and aggregate (village-level) components.

Estimating Idiosyncratic Wage Volatility

A key contribution of this chapter is to use past earnings volatility to proxy the risk profiles of households and villages. This approach is intuitively appealing, as it is based on the assumption that households use past earnings volatility to predict future volatility, which I believe to be a reasonable starting point (ideally, one would like to observe the uninsured portion of the unanticipated components of earnings variability to obtain an accurate representation of the household's exposure to risk). The data used to measure household and village level risk, are retrospective earnings and labour supply data which are observed for key household members over the previous five years, 1988 to 1992.[23]

An immediate issue arises as to whose earnings to measure the variability of. If one were to examine the variability of total household earnings,

[23]From the point of view of isolating unanticipated income changes, consumption data would be preferable on two fronts. First, expenditures are likely to provide a more accurate picture of economic well-being over the longer term than current income, and second, they are believed to be less prone to measurement error than wealth or income. However, only a one-year measure of consumption is observed in the 1993 wave of the IFLS.

a key concern is that of confounding labour supply responses. The conventional household earnings measure for non self-employed households is $Y^{hh} = w_m L_m + w_f L_f + w_c L_c$ where Y^{hh} is total household income, and w_i and L_i are wages and labour supply for males, females and children respectively ($i = m, f, c$). If a household has anticipated a bad draw of income, this expectation will be reflected in its labour supply, and household income will subsequently comprise earnings that are the result of behaviour that has been taken to minimise exposure to, or to reduce the effects of, risk. Therefore the volatility of Y^{hh} across time would under-estimate the ex-ante uninsurable component of risk. In an attempt to better capture true risk, I measure the hourly wage variability of the household head only, on the assumption that the head of the household is the main earner and will always want to work, and thus their observed work status at the extensive margin is not a response to uncertainty. Hourly wage variability is estimated, due to the fact that annual earnings would be contaminated by labour supply responses to risk.[24]

However for self-employed household heads, and particularly for households with a family enterprise, net profit from the enterprise is reported by the head, $Y^h = \Pi$. Income from the enterprise is attributed to one family member (the head), but it is likely to comprise the work contributions of family labour, thus contaminating the net profit measure by ex-post labour supply adjustments through the inclusion of the opportunity costs of family labour. In particular, an observation that a particular household has a relatively smooth income stream, whilst another's is volatile, may actually be

[24]This wage measure is constructed from self-reported retrospective labour supply data. Monthly earnings are converted to an hourly wage using data on the number of weeks worked per year and the number of hours worked per week. If the head is observed to have both a primary and a secondary job, only their primary wage is used in the construction of wage variability. This is because secondary jobs are often an important risk reduction mechanism. The sample of household heads is restricted to those who are currently between the ages of 25 and 65, in an attempt to capture those individuals who have most likely been the main household earner for each of the past 5 years. Mean individual, household, and community characteristics are very similar in this restricted sample, to those in table 2.1.

due to the availability and use of family labour in the former but not in the latter. Indeed, it is plausible that ex-ante risk is the same in both households.

Ideally, if one knew the marginal product of each family member, it would be possible to calculate their wage and to net out labour supply responses to risk from Y^h. Unfortunately data limitations preclude this option. However, I observe to what extent family versus hired labour is used on the enterprise - in particular whether the head is (a) self-employed without help (b) self-employed with the help of householders/temporary workers or (c) self-employed with the help of regular workers. I exclude individuals whose status is reported as (b) from the construction of volatility. The reported net profit of the remaining sample of self-employed persons is therefore not inclusive of household labour supply, and is also net of wages of employed workers. As this sample selection is by no means innocuous, I return to a discussion of its importance below.

I observe a maximum of 5 years of retrospective income data across 822 households that have at least 1 10-14 year old, across 79 small rural villages.[25] To account for the fact that some component of earnings changes are predictable, the hourly wage of the household head is first regressed on a set of permanent and predictable indicators of earnings, X_{hvt}, including age, age squared and years of education.

$$\ln w_{hvt} = \beta_v X_{hvt} + \beta_{vt} + \epsilon_{hvt} \tag{2.15}$$
$$\text{for} \quad h = 1, \cdots, H; t = 1, \cdots, 5; v = 1, \cdots, 79$$

where $\ln w_{hvt}$ is the log of the hourly primary wage of household head h in region v at time t, β_{vt} is a village time dummy and ϵ_{hvt} includes both unobserved and unanticipated individual and village characteristics that affect

[25] I omit villages in which the total number of wage observations across the entire 5 years is less than 25. This is because the resulting volatility measure is likely to be unreliable, due to it being based on such a small number of observations. This leads to the loss of 12 villages. A comparison of the main characteristics of these regions with the remaining sample shows that they are marginally more remote in terms of distance to the nearest bank and schools. However all other characteristics are very similar across both samples.

the wage of the head.

By relaxing the constraint that all villages have the same parameter vector, a 79-equation seemingly unrelated regression model is obtained. The parameters β_v and β_{vt} are thus estimated separately by village using OLS on the data pooled over individuals and years.

From the estimates of the residual, $\widehat{\epsilon_{hvt}}$ in (2.15), the coefficient of variation of $\exp(\widehat{\epsilon_{hvt}})$ across years for each household is constructed

$$\widehat{cv}_{hv} = sd_{\exp\widehat{\epsilon_{hvt}}} / \mu_{\exp\widehat{\epsilon_{hvt}}} \tag{2.16}$$

where $\widehat{cv}_{hv}$ is the coefficient of variation for household head (and thus household) h in village v, $sd_{\exp\widehat{\epsilon_{hvt}}}$ is the standard deviation of $\exp\widehat{\epsilon_{hvt}}$ and $\mu_{\exp\widehat{\epsilon_{hvt}}}$ its mean, for each household

$$\mu_{\exp\widehat{\epsilon_{hvt}}} = \frac{\exp\widehat{\epsilon_{hv1}} + \exp\widehat{\epsilon_{hv2}} + \exp\widehat{\epsilon_{hv3}} + \exp\widehat{\epsilon_{hv4}} + \exp\widehat{\epsilon_{hv5}}}{5}$$

for $h = 1, \cdots, H$, $v = 1, \cdots, V$.[26]

However, household expectations of their earnings are based on more information than is observed by the analyst. As a result, $\widehat{\epsilon_{hvt}}$ will include both unpredictable components of the wage (true risk), along with anticipated but unobserved components and measurement error. The coefficient of variation as estimated in (2.16) will comprise these factors and will thus over-estimate wage variability that is due to true risk. In order to smooth out this idiosyncratic risk measure, $\widehat{cv}_{hv}$ is regressed on a vector of observable household and village characteristics that I assume affect risk

$$\widehat{cv}_{hv} = \lambda_v + \gamma Z_{hvt} + u_{hvt} \tag{2.17}$$

where λ_v is a vector of time-invariant village dummies, included in order to pick up permanent and unobserved village characteristics that persistently

[26]In practice, the coefficient of variation is multiplied by $\frac{1}{\sqrt{m_h+1}}$, where m_h is the number of years of missing wage data for household head h. In addition, the top and bottom 1% of wage observations are excluded from the calculation of volatility.

affect idiosyncratic risk and γ reflects the effect of various household and village characteristics, Z_{hvt}, on idiosyncratic risk. The predicted $\lambda_v + \gamma Z_{hvt}$ from (2.17) is the smoothed component of the unexplained wage volatility and is used as the measure of true idiosyncratic risk in the estimations below.

In order to include this measure of risk in the human capital equation, the key identifying variables in (2.17) are the number of household shocks for each of the years 1988 to 1992. Intuitively, I argue that the only effect of household shocks on education, is through the variability they induce on earnings. Almost 17% of households reporting at least one form of adverse shock in 1992. Data on shocks is collected retrospectively - yearly back to 1988, and include both household specific shocks such as serious illness, death or unemployment of a householder and economic events such as a fall in household income due to falling prices, crop loss or other business loss.[27] Also included in Z_{hvt} are the education level of the head, the occupation type of the head, and interactions of enterprise ownership with the size of the enterprise. The output from this regression is displayed in table 2.4. The key instrument, the number of household shocks from 1988 through 1992, is positively and significantly associated with the estimated measure of idiosyncratic risk. An interesting finding is that farm ownership, regardless of farm size, is associated with significantly lower idiosyncratic risk compared to non-ownership of either a farm or a business. This suggests that it is individuals who are employed as labourers on another farm or enterprise, who face the highest idiosyncratic risk levels.

However, the validity of the instrument rests on the assumption that conditional on the variability of labour earnings (idiosyncratic risk) and on the current income of the household, past shocks have no independent effect on human capital accumulation. In order for this to be a valid identification assumption, it is important to control for any other variables that may be correlated with past shocks and that are likely to affect education. In particular, to the extent that non-labour income adjustments are used

[27]The retrospective data on shocks displays some recall bias, with the number of reported shocks being higher, the closer it is to the survey year.

(ex-post) by the household as a means of buffering consumption, the current non-labour income of the household is likely to be a function of past shocks. As a result, failure to control for the current non-labour income of the household in the schooling regressions could lead to biased estimates of the effects of idiosyncratic risk on education: non-labour income, itself likely to be a function of past shocks, would comprise part of the error term, thus rendering the identification assumption invalid and underlining the importance of controlling for current household non-labour income in the human capital equations.[28]

As discussed above, the exclusion of household heads who are self-employed and who use family labour, on the basis that their reported wage is inclusive of labour contributions of family members, is by no means innocuous. They constitute just under 40% of the overall sample. An alternative to this strategy would be to measure the volatility of total household income as opposed to the hourly wage of the head. In this way, whilst the earnings measure for the whole sample would include labour contributions of the entire household and would be comparable across households in this regard, it would however under-estimate the raw exposure of households to risk. Indeed any differences in this measure of risk across households would largely reflect the differing abilities of households to cope with risk (for example, as reflected in household composition, such as the number of working-age males for example or the number of children who can potentially work). Thus it seems more reasonable to measure the variability of hourly wages of the head and to exclude the sample of self-employed heads with family labour. The extent to which this group differs from the remaining sample of self-employed heads is clearly of concern, due to possible selection bias. A

[28]The sign of the bias would depend on both the impact of non-labour income on education and on the correlation between idiosyncratic risk and the error term. If current non-labour income is favourable for education, and if there is a positive correlation between idiosyncratic risk and current non-labour income (as is likely to be the case if households have not depleted their stocks ex-post to buffer consumption), the effects of household-level risk on education would be upward-biased. On the other hand, a negative correlation between non-labour income and idiosyncratic risk would lead to a downward bias of the effects of idiosyncratic risk on education.

comparison of mean household characteristics of the two groups in table 2.3 shows that mean characteristics are indeed very similar across both samples, with the obvious exception of business and farm ownership. Finally, whilst this group is excluded from the wage equation (2.15), in (2.17) I predict idiosyncratic risk for the entire sample, including self-employed household heads who use family labour (for whom Z_{hvt} is observed). Thus children who are from self-employed households with family labour, are included in estimated regressions for measuring the effect of risk on schooling. This further alleviates any concerns about sample selection bias.[29]

Estimating Aggregate Village Wage Volatility

Allowing the village-specific time dummies to vary across villages separates out the aggregate component of income, and the coefficient of variation of $\exp \widehat{\beta_{vt}}$ from (2.15) across 1988 to 1992 is computed for each of the 79 villages as follows

$$\widehat{cv}_v = sd_{\exp \widehat{\beta}_{vt}} / \mu_{\exp \widehat{\beta}_{vt}} \tag{2.18}$$

where $\widehat{cv}_v$ is the estimated coefficient of variation of village v and is the measure of aggregate risk, $sd_{\exp \widehat{\beta}_{vt}}$ is the standard deviation of $\exp \widehat{\beta}_{vt}$ across the 5 years for village v, and $\mu_{\exp \widehat{\beta}_{vt}}$ is its mean.

Before turning to the results, the extent to which this estimated aggregate risk measure reflects pervasive risk within a particular village, is important to consider. This is particularly so as it is constructed using retrospective labour supply and earnings data, and thus the reliability of the reported data must be considered. To begin with, if there are concerns about the recall nature of the data, it is likely that even if the precise details on past labour supply and income are misreported, it seems reasonable to expect households to recall whether they have had smooth or volatile past incomes, and to report such fluctuations accordingly. This volatility will be captured in the data. In addition, if the aggregate risk measure

[29]In table 2.10 I present evidence on the sensitivity of results to the exclusion of this sample of heads. I thus re-estimate (2.15) on the entire sample of household heads to see how the results are affected. I return to this below.

is truly capturing village-wide uncertainty, it is informative to examine its correlation with salient village characteristics.[30] Table 2.5 presents results from an OLS regression of the estimate of village risk on a range of village characteristics, and provides some indication as to the correlations between aggregate risk and the characteristics listed in the table. It is reassuring to note that the observed correlations largely conform with expectations. In particular, even after controlling for village wealth, aggregate risk is higher on average in villages in which there is no formal access to credit. This is consistent with these villages being less well developed and therefore likely to be more susceptible to pervasive shocks. In addition, as expected, aggregate risk is positively correlated with the proportion of farming households in the village. Given that farming households tend to have more volatile income streams, this again conforms to priors. Further, the coefficient on the number of adverse shocks in the village over the past 5 years, and the amount of rainfall in 1991 / 1992, are of the expected sign, although not statistically different from zero at conventional levels. Risk is lower in villages with hilly rather than flat land, which is somewhat counter-intuitive (it is not significant at the 5% level however). However in general, this evidence is reassuring in terms of being largely consistent with expectations regarding aggregate risk and village characteristics.

2.4 Results

The equation that forms the basis of the regressions for estimating the effect of risk on investment in education is

$$S_{ivt} = \alpha_0 + \alpha_1 \widehat{cv}_{hv} + \alpha_2 \widehat{cv}_v + \alpha_3 X_{ivt} + \eta_{ivt} \tag{2.19}$$

where S_{ivt} is a measure of human capital of child i in village v at time t (1993), $\widehat{cv}_{hv}$ is the estimated idiosyncratic risk of the household in which person i lives, $\widehat{cv}_v$ is the aggregate risk of the village in which individual i resides, X_{ivt} includes individual, household and village characteristics that affect the schooling of the child, and η_{ivt} includes unobserved individual

[30] I thank Timothy Besley for this suggestion.

and village characteristics that affect the measure of human capital of individual i in period t and that I assume are uncorrelated with the regressors.

In equation (2.19), α_1 represents the effect of idiosyncratic risk on child schooling, and α_2 yields the effect of aggregate village risk on education. As discussed in section 2.3, I consider two different measures of S_{ivt}: the current schooling status of the child and their accumulated years of education. In each of tables 2.6 to 2.10, equation (2.19) is estimated across different sub-samples.[31] In each of the specifications, the following variables are also controlled for: age and gender of the child, gender of the household head, religion of the household, missing parent, unschooled parents, household size, log of household expenditure, mean income of the head, log value of liquid assets, farm ownership, business ownership, number of primary, junior and senior high schools in the area, distance to the nearest school, presence of bank in the area, log of village expenditure, village size, and average village level wages for males, females and children. The effects of these characteristics on education choices are remarkably consistent across samples. To briefly summarise them, I find that the most notable factors having an adverse effect on education (whether on the probability of attending school or on accumulated years of education) include having a higher number of younger siblings, unschooled parents, living farther from a bank and living in a village with a relatively lower number of junior high schools. The log of household expenditure has a strong positive effect in all specifications.[32]

2.4.1 Effects of Risk on Human Capital

I now turn to the effects of aggregate and idiosyncratic risk variables on investment in education. To begin with, marginal effects from a probit estimation as to whether the child is currently in school, are presented in table

[31] Note that the p-value for the joint significance of the key instruments in the first-stage regression is less than 0.04 across all sub-samples, indicating that the instruments have predictive power for the idiosyncratic measure of risk. Note also that standard errors on the risk coefficients have been adjusted for the first-stage prediction.

[32] These are in line with previous findings in this literature (see for example Grootaert (1999)) and are available upon request.

2.6, separately for 7 to 12 and 10 to 14 year olds. This first set of results shows that for both 7-12 and 10-14 year olds, idiosyncratic risk has a negative but insignificant effect on the probability that the child is currently enrolled in school. The effect of aggregate risk in both samples is also insignificant (especially so for the older age group). However, as discussed in section 2.3, the dependent variable in this case is not wholly informative as to the current stock of human capital of the child, or as to whether the child has been working in the past. In particular, the current activity of the child gives no insight as to exactly *when* the child was withdrawn from school, and thus to what extent human capital accumulation has been affected. A more direct measure of investment in human capital of the child is gleaned by examining accumulated years of education. This captures past temporary (and permanent) interruptions to schooling.[33] For this reason, from hereon I focus on the effects of risk on this measure of human capital.

In table 2.7, the results for both the 7-12 and 10-14 year old samples provide strong evidence that aggregate risk has a negative and significant effect on the accumulated years of education of the child. On the other hand, the effects of idiosyncratic risk are not statistically different from zero for either sub-sample. The results are consistent with children being used as insurance in response to aggregate village risk, and with idiosyncratic risk being diversified away without the use of children, in line with previous findings that aggregate risk is more difficult to insure against than idiosyncratic risk. This is a key finding, and its implications are extremely important for understanding factors in education - and possibly child labour - choices in LDCs. Below, I probe this result further to assess its robustness across different subsamples.

[33] If the individual is observed to have repeated a year of school, this is not counted as an extra year of education. Thus the years of education variable captures any delays in the education process through having to repeat, through late enrolment, or through withdrawal for one year (or more). It does not directly pick up seasonal interruptions to schooling, except through seasonal interruptions leading to the child having to repeat a year of schooling.

Whilst defining the samples on the basis of primary school age and (child) working age is intuitively appealing, it is also of interest to examine the effects of risk across other refined age ranges. For example, one might expect the effects of risk to be lower for the < 10 age range, due to the fact that work is relatively uncommon at this age. By a similar argument, it is also reasonable to expect risk to have less of an effect on 15-17 year olds, amongst whom work is relatively more common. I examine the effects of risk on years of education separately for age groups 7-9, 10-12, 13-14 and 15-17. Results in table 2.8 show that most of the adverse effect of aggregate risk is for the 10-12 year old range. This is in line with expectations. After age 12, it is much more common for individuals to be withdrawn from school; before age 10, it is extremely rare for children to work. Thus the 10-12 year old group comprises individuals who are likely to be most sensitive to economic constraints, and this is indeed borne out in the analysis in table 2.8. Throughout, the effects of idiosyncratic risk are not statistically different from zero.

In an attempt to capture those villages in which consumption smoothing is limited by borrowing constraints, the sample is further restricted to small rural villages that do not have access to credit. Credit represents access to some type of formal credit, whether it is for consumption or investment purposes. From table 2.9 we see that in villages without any access to credit, the pattern is the same as for the overall sample, with idiosyncratic volatility having a negative but insignificant effect on years of education. Despite the decrease in sample sizes, a comparison of tables 2.7 and 2.9 is interesting, as it shows that the effect of aggregate risk is negative and even stronger when the sample of villages is restricted to those without formal credit. This pattern of results conforms to children serving as a consumption-smoothing device in areas where one might expect household diversification strategies to be most important, i.e. areas without any formal credit.[34]

[34]Note that a number of other robustness checks have been carried out, including the sensitivity of the results to the age cutoff of the household head, and to the precise cutoff for village size. The same pattern of results holds.

Finally, I base the measures of risk on the sample of household heads that is inclusive of the self-employed who use family labour. As discussed already, I expect this to under-estimate the ex-ante risk facing households given that the wages of this group of self-employed will include contributions of family labour. In table 2.10, we see that this measure of aggregate risk has negative effects across the 7-12 and 10-12 year old samples, which are significant at the 10% level. Idiosyncratic risk again has no significant effect. Thus from this evidence, it is unlikely that the exclusion of the self-employed with family labour from the wage equations, is biasing the estimates of the effect of risk on education choices. Whilst this is not the preferred measure of risk, I view this analysis as a pointer towards the salient issues involved in measuring risk in LDCs, and the importance of ideally having access to in-depth labour supply and earnings data for each family member, in order to better proxy the ex-ante risk facing households. It is hoped that future data collection in LDCs may bear this in mind.[35]

To conclude, the results are consistent with children aged 10-plus serving as a consumption smoothing device against aggregate but not idiosyncratic forms of risk. This result emerges most strongly for the 10 to 12 age group. It is in line with previous literature which finds that aggregate village risk is more difficult to insure than idiosyncratic risk. To the extent that this lower schooling is substituted by work, one can infer risk feeding through to child labour, in the form of a possible buffer stock for the parent. However, this is not possible to conclude here. Ideally, one would like to observe the child's activity at a number of points in time in order to draw any conclusions about the direct effect of risk on child labour. For the moment, one can only suggest that the human capital accumulation reductions, with their adverse dynamic effects, may be facilitating the occurrence of child labour.

[35]Indeed there is already a move towards the more direct collection of data on household risk in micro-level surveys.

2.4.2 Robustness to Assumptions in Model

In terms of aggregate village risk, as discussed in section 2.2, whilst its negative effect on education may be the result of inadequate insurance mechanisms, it must also be discussed in the light of two key assumptions in the model: the lack of intergenerational transfers from adults to parents, and deterministic earnings for the adult, conditional on education. In section 2.2 I outlined how relaxing these assumptions can lead to the effects of risk on education becoming ambiguous. It is important to determine to what extent these assumptions are plausible in Indonesia, and therefore how they may be having a bearing on the empirical results.

In the model of section 2.2, if one permits transfers from the adult to the parent in period 2, the pure effect of these future transfers is to increase investment in education. This is traded off against the value of current child earnings to the adult, as a means of buffer stock accumulation, and the net effect of risk is thus ambiguous (see equation (2.13)). The observed response of education to risk may thus be the net of these two effects. For aggregate risk, it may be the case that the buffer stock motive is taking precedence over transfer motives for the parent. However, it may instead be the case that intergenerational transfers are unrelated to education in Indonesia, in which case the assumption of the model holds and the observed negative effect is likely to be untempered by this transfers motive. One test of this assumption is to examine whether children with higher education transfer more to their parents. I estimate a probit in which the dependent variable is equal to one if the parent is a net transfer recipient. Controlling for a range of household characteristics, table 2.12 shows that there are no differential effects of various education levels of the donor on this dependent variable.[36] Thus this provides evidence that this offsetting force is unlikely to be in operation in the sample.

The second testable assumption relates to the fact that if earnings risk

[36] This is in line with findings by Raut and Tran (1998) and Cameron and Cobb-Clark (2001), also using the IFLS data.

is increasing in education, and if the parent assesses the randomness of the future environment for the adult on the basis of their own experiences, investment in education may decrease in response to parental income risk - without having to appeal to an insurance argument. This assumption is difficult to test however, as the response of investment in education to risk depends on the precise way in which risk interacts with the earnings function. With this caveat in mind, one simple way is to test whether the measure of idiosyncratic risk is decreasing or increasing in education level. If it is decreasing, it is more plausible to expect investment in education to be positively affected by risk (in order to minimise the future risk that will face the adult), compared to the case in which it is increasing. Thus in the former case, the observed adverse response of education to aggregate risk may reflect the buffer stock motive outweighing the incorporation of adult income risk into parental education choices on the basis of perceived interactions between education and income risk. The coefficients on education levels from equation (2.17), shown in table 2.13, indicate that idiosyncratic risk is decreasing in education level, being highest for unschooled individuals, and for those with no/primary qualification, relative to more senior qualifications. It is therefore plausible that the adverse effects of aggregate risk on education are being mitigated by a possible increase in education in response to future earnings risk of the adult. If this is indeed happening, it in fact strengthens the empirical conclusions.

The same line of thought can be applied to the finding that education choices are not affected by idiosyncratic risk. However, it is difficult to think why the two assumptions would operate differently on both types of risk. Instead, it seems more plausible to us that the effects are being largely driven by a lack of insurance markets for aggregate risk, with idiosyncratic risk being diversified away through means other than children. In fact, one way of coping with risk is by drawing on precautionary savings. In Indonesia, jewellery is an important means of saving that is used to buffer against adverse events. Table 2.11 provides estimates of the marginal effects of risk on the probability that the household currently owns jewellery. The results are interesting, providing evidence that in those households in small rural villages with at least one 7-12 or 10-14 year old, higher levels of aggregate risk increase the probability that the household owns jewellery, suggesting that households save more in response to pervasive risk. The effect of idiosyncratic risk is not statistically different from zero at conventional levels. However these findings are merely suggestive of a plausible range of mechanisms being used by the household to deal with uncertainty. Nonetheless, I view them as providing a motivation for further research into the myriad of ways that households cope with risk, with important consideration for the role of children in this risk-coping portfolio.

Finally, whether the results are due to thin insurance markets for income in small rural villages, is also assessed by examining the effects of risk in large urban areas. The results in table 2.14 show that none of the risk variables are statistically different from zero at conventional levels in these areas. On the basis that it is reasonable to expect insurance markets to be better-functioning in these areas[37], and controlling for a wide range of area characteristics, this is consistent with the argument that the negative effect of aggregate risk in small rural areas, is largely due to insurance market failures.

2.5 Conclusion

In this chapter, the effects of uncertainty combined with inadequate insurance and credit markets, on investment in education, have been specifically addressed. Parental income risk has been shown, under certain conditions, to have an adverse effect on education, with children fulfilling the possible role of an insurance mechanism for smoothing consumption.

Empirically capturing the degree of risk facing households is a non-trivial task. Households may use many unobserved mechanisms to anticipate fu-

[37] Apart from formal insurance (and subsequent moral hazard problems), it may also be the case that individuals are less susceptible to covariant forms of risk in large urban areas, and therefore informal insurance arrangements amongst individuals are more likely to be successful.

ture risk, and the assumption underlying the empirical estimation is that past earnings volatility is a useful indicator to the household of persistent risk and possible future uncertainty. Using past earnings volatility to construct measures of idiosyncratic and aggregate village risk, is a fundamental contribution of this chapter. From estimates of wage volatility at both idiosyncratic and aggregate levels, the analysis has pointed towards an adverse effect of aggregate uncertainty on child education in small rural villages, with idiosyncratic risk having no statistically significant effect on years of education. This is suggestive of the use of children to fill possible gaps in insurance markets for aggregate risk and is in line with previous findings which indicate that aggregate risk is more difficult to insure against than idiosyncratic risk. The robustness of this result across different data sets, is a topic for future research.

From a policy perspective, it is clear that the interaction of risk and lack of insurance markets, may combine to have detrimental effects on the education of a child. However, the findings indicate that idiosyncratic risk is being diversified by households without having to resort to the labour of their children. Whether this is through formal insurance provision or through self-insurance in the household/social networks, is not possible to gauge from this analysis. Aggregate village risk on the other hand, seems to be more difficult to insure against, and this is likely to have serious consequences on child human capital acquisition. Policy must be carefully crafted in order that more formal insurance provision does not crowd out insurance mechanisms that households already use to deal with idiosyncratic risk. At the same time, the costliness of such self-insurance to the household must also be considered.

2.6 Tables

Table 2.1: Mean Characteristics by Region Type

	(1) **Small Rural Villages** $\leq 1,000$ households	(2) **All Other Regions** Urban (all households) Rural (>1,000 households)
Age of the child	11.9766	11.9681
	(1.4213)	(1.4173)
Male child	0.5102	0.4954
	(0.5001)	(0.5001)
Child is attending school	0.8300	0.8853
	(0.3758)	(0.3187)
Unschooled mother	0.3214	0.2708
	(0.4672)	(0.4444)
Unschooled father	0.2928	0.2687
	(0.4552)	(0.4434)
Household size	6.0996	6.1916
	(2.0114)	(2.1105)
Farm ownership	0.7247	0.2459
	(0.4468)	(0.4307)
Business ownership	0.3243	0.4155
	(0.4683)	(0.4929)

Table 2.1 contd.

	(1)	(2)
Number of households in village	551.04	2658.69
	(249.57)	(3122.44)
Urban area	-	0.7109
		(0.4534)
Number of senior high schools	1.4182	1.6424
	(1.1982)	(1.0414)
Distance to nearest school (km)	3.8046	1.4041
	(3.1814)	(1.4055)
Presence of bank	0.1119	0.49
	(0.3154)	(0.50)
	$N = 1,136$	$N = 2,729$

Notes: Sample refers to households with at least one 10 to 14 year old. N is the number of children. Standard deviations in parentheses.

Table 2.2: Current Primary Activity of 10-14 Year Olds

Small Rural Villages				
Work	**Look for work**	**Housework**	**School**	**Other**
0.0561	0.0578	0.0272	0.8124	0.0465
(0.2302)	(0.2336)	(0.1626)	(0.3905)	(0.2008)

N=1,136

Notes: Other consists of those who are physically invalid, mentally handicapped or suffer other hardships. Standard deviations in parentheses.

Table 2.3: Mean Characteristics of Households with Self-Employed Head

	With Family Labour	**Without Family Labour**
Age of head	46.17	45.19
	(10.65)	(10.96)
Male head	0.8942	0.8612
	(0.3077)	(0.3458)
Years of education of head	4.2478	4.2578
	(3.5740)	(3.5584)
Liquid asset ownership	0.5160	0.5252
	(0.50)	(0.4995)
Farm ownership	0.7882	0.6354
	(0.4087)	(0.4814)
Business ownership	0.3869	0.4708
	(0.4872)	(0.4993)
Fraction of adults working	0.7353	0.6893
	(0.2829)	(0.2860)
Household size	4.9992	4.7839
	(2.0676)	(2.0698)
	N=1,190	N=2,161

Standard deviations in parentheses. N is the number of households.

Table 2.4: Estimates from Equation (2.17)

	Small Rural Villages Idiosyncratic risk
Number of household shocks 1988 to 1992	0.0292*
	(0.0144)
Owns business	0.0152
	(0.0176)
Owns small-sized farm	-0.0589*
	(0.0269)
Owns medium-sized farm	-0.1002**
	(0.0262)
Owns large-sized farm	-0.0679**
	(0.0244)
Age of household head	-0.0023*
	(0.0011)
	$r^2 = 0.2970$
	Number of villages = 79

Notes: Dependent variable is idiosyncratic risk, as estimated in (2.16) from equation (2.15). The farm and business dummies may be interpreted relative to non-ownership of either a farm or a business. Also include the education and occupation of the household head and village dummy variables. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.5: OLS Estimates of Aggregate Risk on Village Characteristics

	Small Rural Villages
Log village expenditure	-0.0210
	(0.0538)
Land type hilly	-0.0687
	(0.0379)
Soil productivity average / high	0.0103
	(0.0330)
Credit availability	-0.0679*
	(0.0320)
Presence of cottage industry	-0.0019
	(0.0361)
Presence of factory	0.0583
	(0.0538)
Proportion of hhs in village with farm	0.2021*
	(0.0817)
Proportion of hhs in village with business	0.0681
	(0.1316)
Average number of village shocks 1988-1992	0.0053
	(0.0052)
Average village rainfall 1991-1992	-0.0386
	(0.0268)
Village has irrigated ricefields	-0.0242
	(0.0449)
	$r^2 = 0.327$
	N=79

I also include controls for village size and province. N is the number of villages. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.6: Probit Estimates - Effects of Risk on School Attendance

	Small Rural Villages	
	Attend school	
	(1)	(2)
	7-12	10-14
Idiosyncratic risk	-0.0299	-0.0791
	(0.0386)	(0.0794)
Aggregate risk	-0.0713	0.0117
	(0.0473)	(0.1013)
	$r^2 = 0.1604$	$r^2 = 0.2240$
	$N = 1,366$	$N = 1,136$
	Number of villages = 79	

Notes: Also include standard household, child and village level controls. Reference category is not attend school. Standard errors corrected for clustering at the village level. N is the number of children. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.7: OLS Estimates - Effects of Risk on Years of Education

	Small Rural Villages	
	Years of education	
	(1)	(2)
	7-12	10-14
Idiosyncratic risk	0.0240	-0.1358
	(0.3055)	(0.5534)
Aggregate risk	-0.8753*	-1.0746*
	(0.3461)	(0.5044)
	$r^2 = 0.5207$	$r^2 = 0.4235$
	$N = 1,212$	$N = 1,026$
	Number of villages = 79	

Notes: Also include standard household, child and village level controls. Standard errors corrected for clustering at the village level. N is the number of children. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.8: OLS Estimates - Effects of Risk on Years of Education, by Age

	Small Rural Villages			
	Years of education			
	(1)	(2)	(3)	(4)
	7-9	10-12	13-14	15-17
Idiosyncratic risk	0.1564	0.0789	-0.1177	-0.4826
	(0.3364)	(0.4628)	(0.8909)	(1.0281)
Aggregate risk	-0.2734	-1.4546**	-0.8214	-0.9136
	(0.4815)	(0.5073)	(0.7508)	(0.9038)
	$r^2 = 0.2652$	$r^2 = 0.2597$	$r^2 = 0.2278$	$r^2 = 0.2649$
	$N = 565$	$N = 647$	$N = 379$	$N = 493$
	Number of villages = 79			

Notes: Also include standard household, child and village level controls. Standard errors corrected for clustering at the village level. N is the number of children. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.9: OLS Estimates - Effects of Risk on Years of Education

	Villages without Credit	
	Years of education	
	(1)	(2)
	7-12	10-14
Idiosyncratic risk	-0.0209	-0.1052
	(0.2596)	(0.5048)
Aggregate risk	-1.4708**	-1.5807*
	(0.5070)	(0.8048)
	$r^2 = 0.5201$	$r^2 = 0.4053$
	$N = 989$	$N = 849$
	Number of villages = 60	

Notes: Also include standard household, child and village level controls. Standard errors corrected for clustering at the village level. N is the number of children. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.10: OLS Estimates - Effects of Risk on Years of Education

Sensitivity Analysis

Small Rural Villages

	Years of education		
	(1)	(2)	(3)
	7-12	10-14	10-12
Idiosyncratic risk	-0.0495	-0.4833	-0.0787
	(0.2321)	(0.3117)	(0.3268)
Aggregate risk	-0.9352**	-0.4602	-0.9526
	(0.3596)	(0.4710)	(0.5113)
	$r^2 = 0.4660$	$r^2 = 0.4054$	$r^2 = 0.2232$
	$N = 1,212$	$N = 1,026$	$N = 647$

Number of villages = 79

Notes: Include wages of household heads who employ family labour in estimating the wage equations in (2.15). Also include standard household, child and village level controls. Standard errors corrected for clustering at the village level. N is the number of children. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.11: Probit Estimates - Effects of Risk on Jewellery Ownership

	Small Rural Villages	
	(1)	(2)
	Households with 7-12 year olds	Households with 10-14 year olds
Idiosyncratic risk	-0.0470	-0.0135
	(0.0963)	(0.1043)
Village risk	0.1779	0.2259*
	(0.1050)	(0.1140)
	$r^2 = 0.0424$	$r^2 = 0.0501$
	$N = 935$	$N = 799$

Notes: Also include standard household, child and village level controls. Standard errors corrected for clustering at the village level. N is the number of households. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.12: Probit Estimates - Effects of Education on Transfers

	Small Rural Regions
	Net transfer recipient
Elementary	0.0245
	(0.0238)
Junior High	-0.0049
	(0.0278)
Senior High	-0.0156
	(0.0268)
College	0.0284
	(0.0470)
	$r^2 = 0.1258$

Notes: Also include standard household, individual and village level controls. Omitted category is no education. Standard errors corrected for clustering at the village level. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.13: OLS Estimates - Effects of Education on Idiosyncratic Risk

	Small Rural Regions
	Idiosyncratic Risk
Highest Education Level	
Unschooled	0.1060**
	(0.0193)
No Qualification	0.0422*
	(0.0159)
Elementary	0.0440**
	(0.0148)
Junior High	0.0291
	(0.0176)
	$r^2 = 0.6358$

Notes: Also include standard household, individual and village level controls. Omitted category is senior high qualification or above. Standard errors corrected for clustering at the village level. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

Table 2.14: OLS Estimates - Effects of Risk on Years of Education

	Large Urban Regions	
	Years of education	
	(1)	(2)
	7-12	10-14
Idiosyncratic risk	-0.5862	-0.3895
	(0.4265)	(0.4826)
Aggregate risk	0.4750	0.5299
	(0.4682)	(0.5053)
	$r^2 = 0.5857$	$r^2 = 0.5724$
	$N = 1,147$	$N = 989$
	Number of villages = 118	

Notes: Also include standard household, child and village level controls. Standard errors corrected for clustering at the village level. N is the number of children. ** statistically significant at the 1-percent level; * statistically significant at the 5-percent level.

2.7 Appendix

The first order conditions for consumption and education are respectively

$$\Psi_{c_1^{hh}} : U'(c_1^{hh}) - \beta E_1 U'(c_2^p) = 0 \tag{2.20}$$

$$\Psi_{D_1} : \beta\gamma[U'(c_2^a) f'(g(D_1))] - \beta(p^D + w_1^c) E_1 U'(c_2^p) = 0 \tag{2.21}$$

Effect of Parental Income Risk on D_1

$$\frac{\partial D_1}{\partial \delta}\Big|_{\frac{d\theta}{d\delta}=-\xi} = \frac{\begin{vmatrix} \Psi_{c_1^{hh}c_1^{hh}} & -\Psi_{c_1^{hh}\delta} \\ \Psi_{D_1 c_1^{hh}} & -\Psi_{D_1\delta} \end{vmatrix}}{|H|}$$

where H denotes the Hessian, $|H| > 0$ due to the second order condition for a maximum and

$$\Psi_{c_1^{hh}c_1^{hh}} : \left[U''(c_1^{hh}) + \beta E_1 U''(c_2^p)\right] dc_1^{hh}$$

$$\Psi_{D_1 c_1^{hh}} : \left[\beta(p^D + w_1^c) E_1 U''(c_2^p)\right] dc_1^{hh}$$

$$\Psi_{c_1^{hh}\delta} : -\beta E_1 \left[U''(c_2^p)(y_2^p - \xi)\right] d\delta$$

$$\Psi_{D_1\delta} : -\beta(p^D + w_1^c) E_1 \left[U''(c_2^p)(y_2^p - \xi)\right] d\delta$$

where I have substituted $\delta y_2^p + \theta$ for parental period 2 income and am evaluating the derivative keeping the mean of parental period 2 income constant.

$$\frac{\partial D_1}{\partial \delta}\Big|_{\frac{d\theta}{d\delta}=-\xi} = \frac{\begin{vmatrix} U''(c_1^{hh}) + \beta E_1 U''(c_2^p) & \beta E_1[U''(c_2^p)(y_2^p - \xi)] \\ \beta(p^D + w_1^c) E_1 U''(c_2^p) & \beta(p^D + w_1^c) E_1[U''(c_2^p)(y_2^p - \xi)] \end{vmatrix}}{|H|}$$

$$= \beta(p^D + w_1^c) U''(c_1^{hh}) E_1[U''(c_2^p)(y_2^p - \xi)] \tag{2.22}$$

Risk aversion implies that $U''(c_1^{hh}) < 0$. The sign of $E_1[U''(c_2^p)(y_2^p - \xi)]$ is determined below for all values of y_2^p.

$\boxed{y_2^p \geq \xi}$

Under the assumption that risk aversion $-\frac{U''(c_2^p)}{U'(c_2^p)}$ is decreasing in c_2^p

$$-\frac{U''(c_2^p)}{U'(c_2^p)} \leq \left(-\frac{U''(c_2^p)}{U'(c_2^p)}\right)_\xi \quad \text{if} \quad y_2^p \geq \xi \tag{2.23}$$

$$U'(c_2^p)(y_2^p - \xi) \geq 0 \quad \text{if} \quad y_2^p \geq \xi \tag{2.24}$$

Multiply (2.23) by (2.24)

$$\Rightarrow U''(c_2^p)(y_2^p - \xi) \geq -\left(-\frac{U''(c_2^p)}{U'(c_2^p)}\right)_\xi U'(c_2^p)(y_2^p - \xi)$$

Take expected values on both sides

$$\Rightarrow E_1[U''(c_2^p)(y_2^p - \xi)] \geq -\left(-\frac{U''(c_2^p)}{U'(c_2^p)}\right)_\xi E_1[U'(c_2^p)(y_2^p - \xi)] \tag{2.25}$$

To prove that LHS ≥ 0, it is sufficient to show that RHS ≥ 0. This amounts to showing that $E_1[U'(c_2^p)(y_2^p - \xi)] \leq 0$

Since $U''(c_2^p) < 0$,

$$U'(c_2^p) \leq \left(U'(c_2^p)\right)_\xi \text{ if } y_2^p \geq \xi$$

Also,

$$y_2^p - \xi \geq 0 \text{ if } y_2^p \geq \xi$$

$$\Rightarrow U'(c_2^p)(y_2^p - \xi) \leq (U'(c_2^p))_\xi (y_2^p - \xi)$$

Take expected values

$$\Rightarrow E_1[U'(c_2^p)(y_2^p - \xi)] \leq (U'(c_2^p))_\xi E_1(y_2^p - \xi) = 0$$
$$\Rightarrow E_1[U''(c_2^p)(y_2^p - \xi)] \geq 0 \quad \text{from} \quad (2.25)$$
$$\Rightarrow (p^D + w_1^c)U''(c_1^{hh})E_1[U''(c_2^p)(y_2^p - \xi)] \leq 0$$

$\boxed{y_2^p \leq \xi}$

Because risk aversion is decreasing in c_2^p, it must be that

$$-\frac{U''(c_2^p)}{U'(c_2^p)} \geq \left(-\frac{U''(c_2^p)}{U'(c_2^p)}\right)_\xi \text{ if } y_2^p \leq \xi \tag{2.26}$$

Also

$$U'(c_2^p)(y_2^p - \xi) \leq 0 \text{ if } y_2^p \leq \xi \tag{2.27}$$

Multiply (2.26) by (2.27)

$$\Rightarrow U''(c_2^p)(y_2^p - \xi) \geq \left(\frac{U''(c_2^p)}{U'(c_2^p)}\right)_\xi U'(c_2^p)(y_2^p - \xi) \tag{2.28}$$

$$\Rightarrow E_1[U''(c_2^p)(y_2^p - \xi)] \geq \left(\frac{U''(c_2^p)}{U'(c_2^p)}\right)_\xi E_1[U'(c_2^p)(y_2^p - \xi)]$$

To prove that LHS ≥ 0, it is sufficient to show that RHS ≥ 0. So it must be shown that $E_1[U'(c_2^p)(y_2^p - \xi)] \leq 0$

$$U'(c_2^p) \geq \left(U'(c_2^p)\right)_\xi \text{ if } y_2^p \leq \xi$$

Also

$$y_2^p - \xi \leq 0 \text{ if } y_2^p \leq \xi$$

$$\Rightarrow U'(c_2^p)(y_2^p - \xi) \leq (U'(c_2^p))_\xi (y_2^p - \xi)$$

Take expected values

$$\Rightarrow E_1[U'(c_2^p)(y_2^p - \xi)] \leq (U'(c_2^p))_\xi E_1(y_2^p - \xi) = 0$$
$$\Rightarrow E_1[U''(c_2^p)(y_2^p - \xi)] \geq 0 \quad \text{from} \quad (2.28)$$
$$\Rightarrow \beta(p^D + w_1^c)U''(c_1^{hh})E_1[U''(c_2^p)(y_2^p - \xi)] \leq 0$$

$$\Rightarrow \partial D_1/\partial\delta \leq 0 \text{ in } (2.22)$$

Chapter 3

Sibling Composition, Education Costs and Child Activity in Mexico

3.1 Introduction

Whilst the previous chapter explicitly focused on the effects of risk and uncertainty on education and work choices for children, the analysis did not consider within-household interactions in the allocation of resources across children. However, the widely observed overall differential levels of investment in the education of males and females in less developed economies, clearly indicate unequal investment into the education of children within the household and are the end result of a highly complex intra-household decision-making process.[1] For a number of possible reasons, the *structure* of the family is relevant to such decisions. Even if parents are equally altruistic towards their children, if they seek to maximise family income, they will make investments in each child efficiently, investing more in the education

[1]See Schultz (1993) for evidence that females are less well educated than males in developing economies; indeed of children not in school in less developed economies in 2003, females comprise two-thirds - see www.worldrevolution.org. Of course the within-household inequalities are most likely not confined to between-gender. One can also think of birth-order differences for example, although gender inequalities are most easily detectable in macro-level statistics.

of children yielding the highest expected returns.[2] However, the presence of resource and credit constraints is the most likely explanation for differential investment within the family. This is because when the household is not able to perfectly smooth inter-temporal consumption, or to borrow against the future human capital of children, current income matters for education and parents can no longer invest optimally in the education of their children, being forced instead to ration resources between them. This may push siblings into competition with each other for scarce resources and as family size increases, current household resources are diluted even further amongst children.[3] Sibling structure thus becomes an important factor in choices and even conditional on the sibling set, birth order will be relevant in the presence of credit constraints. This is because the variation in family income across the lifecycle will occur at different stages in the education cycles of children and will be more favourable for some children than for others.

The empirical evidence on the relationship between the number of children in the household and human capital investment points, with few exceptions, towards a negative association.[4] *A priori* however, the relationship between household size and education is ambiguous. Whilst larger households face a greater division of resources, the productive capacity of some children may actually be beneficial for the education of others in the household. The likelihood of a particular child being used by the parent to alleviate resource constraints depends on the age and gender of that child, along with those of his siblings. For example, females tend to have a comparative advantage in home production, with males more likely to provide help on the household farm or to engage in market work.[5] Very young children, on the other hand, are likely to represent a pure drain on resources. Older children may thus free up resources for younger children, or may actually add to total household income, thus improving opportunities for later-born children. Apart from this, different children have different perceived and

[2]Equity can then be achieved through transfers within the family.

[3]Strauss and Thomas (1995) refer to this as 'resource crowding'.

[4]See for example Rosenzweig and Wolpin (1980), Parish and Willis (1993) and Ahn et al (1998). Kaestner (1997), however, finds no effect.

[5]See Strauss and Thomas (1995) and Edmonds (2002).

actual costs of and benefits from education, which will thus affect their levels of education. Therefore apart from actual household size, the gender and the relative age composition of the family are critical factors in the education decision-making process. Parish and Willis (1993) find that whilst having more siblings is associated with less education for both males and females in Taiwan, the association varies depending on the composition of siblings. They find that more younger siblings of the same gender adversely affect the educational attainment of both males and females, whilst opposite-sex siblings are generally neutral to one's educational attainment. On the other hand, older brothers reduce the educational attainment only of males, whilst older sisters are associated with more education for both males and females, thus suggesting that older sisters may serve to alleviate the resource constraints on the family. Patrinos and Psacharopoulos (1997) also provide some evidence that older children may work to provide for their younger siblings.[6] The evidence on birth-order effects on education is inconclusive, with some studies finding that last-born children receive more education than earlier-born siblings, consistent with older children freeing up resources for younger siblings, or with parental income being higher at later stages in the lifecycle.[7] Gomes (1984) and Behrman and Taubman (1986), however, point to the favouring of the eldest child with educational resources, regardless of family size, consistent with the parent wishing to obtain income returns at the earliest possible date.

In this chapter, I examine associations between sibling structure, as categorised by age and gender, and discrete indicators of activities of children, using data on rural Mexico. Throughout, I take the number of siblings as exogenous. Apart from controlling for the composition of siblings, I also investigate the effect of the child wage on the economic activity and whether its effect varies by sibling composition. Despite the fact that this is the

[6]The effect of siblings on other human capital outcomes apart from education, such as child health, has also been widely examined. Garg and Morduch (1998), for example, find that sibling sex composition matters importantly in explaining child health outcomes in poor economies.

[7]See for example Ejrnaes and Pörtner (2002) and Emerson and Souza (2002).

largest opportunity cost of education in LDCs, there is a noticeable gap in the literature as to its empirical importance for observed choices.[8] Some exceptions include Rosenzweig and Evenson (1977), Rosenzweig (1982) and Jacoby and Skoufias (1997), all of whom find that increases in child wages negatively affect school attendance. It is plausible that the responsiveness of a particular child to changes in the child wage, is a function of the structure of the sibling set. The main channel through which this may operate is via the demands placed by siblings on household resources, which are a function of such factors as whether or not they work (or are of - subjectively defined - 'working-age'), their education status and the propensity of parents to invest in their human capital. However, it is not clear *a priori* how the wage elasticity of a particular child will be affected, and this will largely be determined by the influence of the above factors. For example, the higher are the demands that other siblings place on resources (whether for pure consumption or for education), the more sensitive one might expect a given child to be to the wage, whilst if siblings in fact contribute to household resources, a given child's wage elasticity may be lower. To my knowledge, this is one of the first studies to explicitly examine whether sibling structure does indeed have a bearing on the responsiveness of the child to the wage. Another feature of the chapter is to allow for an unobserved household-specific component (representing parental preferences for example) to decisions across children within the same household, which is seldom considered in the empirical work in this area, which generally condition on observable characteristics only.[9]

To briefly preview the main results, I find that having a large number of siblings is associated with lower educational attainment of children. Females, especially of relatively young ages, generally tend to stay at home the more siblings they have, whilst males are equally likely to work or to look for

[8]Apart from poor quality wage data, the limited analysis of wage effects is also due to the endogeneity of wages, being only observed for children who choose to work. Indeed a recent analysis by Hazarika and Bedi (2003) on the effect of school costs on child schooling in rural Pakistan, ignores the opportunity costs of school.

[9]A recent exception is Deb and Rosati (2002).

work. Very young siblings and older brothers are the most detrimental for education. In many instances, there are important wage interactions, with larger changes in participation in various activities in response to changes in the wage, the greater the number of siblings.

The chapter proceeds as follows. In section 3.2, I outline a simple static discrete-choice model to show how the effect of the child wage on investment in education may vary by sibling composition and in section 3.3, I discuss the methodological application of this model. Section 3.4 describes the data used in the analysis, which spans seven states in marginalised villages in rural Mexico. Section 3.5 presents the results of the empirical analysis. Section 3.6 concludes and considers some future research topics in this area, along with discussing the possible effectiveness of a variety of policies relating to the encouragement of participation in education and the eventual eradication of child labour in LDCs.

3.2 Theoretical Framework

In order to show how the wage elasticity may vary by household composition, I consider a simple static discrete-choice model in which the choice set for children consists of school (S) and work (W). I consider two households, which vary only in the number of children present. I show how allowing for current parental income to affect choices implies different effects of the wage on choices in each household.[10]

To set up the model, I assume that if child i works, (s)he contributes w_i (the child wage) to household income; the direct cost of schooling is p_i and the discrete gain in human capital as a result of schooling is H_i, which equals 0 if in work and 1 if in school (I represent these values by H_i^W and H_i^S respectively). Household utility depends on both household consumption and the final stock of human capital of the child(ren). I assume

[10]As discussed already, current income may matter due to the parent not being able or not wanting to borrow to invest in human capital, due to the fact that human capital is poor collateral or because the parent is unable to enforce future transfers from the child.

that the parent is the decision-maker, who is rational in the sense of making choices that maximise perceived household utility subject to the relevant constraints.[11] Due to errors in the maximisation as a result of imperfect perception and optimisation, as well as the inability of the parent to precisely measure all of the relevant variables, utility is assumed to be a random function. The two households that I consider are one in which there is only one child, the other in which there are two children.

3.2.1 Households with one child

Consider first the case in which there is only one child in the household. U_j^* is an underlying latent variable denoting the level of utility for household h associated with the child being in activity j.[12]

$$U_j^* \quad = \quad V_j(C_j, H_1^j) + \epsilon_j \quad \text{for} \quad j = S, W \tag{3.1}$$

where C_j is household consumption, which varies across alternatives due to the fact that if the child works he contributes w_1 to income and therefore consumption, whilst if he goes to school the cost is p_1; H_1^j is the human capital level of the child in activity j and ϵ_j is a residual that captures unobserved variations in tastes and in the attributes of alternatives, along with errors in the perception and optimisation by the parent.

I approximate V_j in (3.1) using a form that is linear in the parameters

$$V_j = \alpha_j C_j + \alpha_1 H_1^j \quad \text{for} \quad j = S, W \tag{3.2}$$

where

$$C_j = Y - 1.(j = S)p_1 + 1.(j = W)w_1 \tag{3.3}$$

[11] As in chapter 2, this assumes the existence of a household utility function rather than a weighted average of utility functions of the two parents, and enables me to abstract from bargaining between parents that may arise due to different preferences across parents.

[12] As the household (parent) is the decision-making unit, implicitly all variables are subscripted by h. This indexing is omitted for ease of notation. I also suppress the conditioning on a range of observable household, individual and village characteristics and introduce it further below.

where Y denotes the exogenous earnings of the parent, whom I assume always works.

The probability of observing a particular choice, j, for the child is denoted by P_j^0 (the superscript denotes the number of siblings of the child). This choice will be observed if it yields the maximum utility to the parent across all possible choices

$$\begin{aligned} P_j^0 &= Pr(U_j^* > U_{j'}^*) \\ &= Pr(\epsilon_{j'} < V_j - V_{j'} + \epsilon_j) \\ &= \frac{e^{V_j}}{\sum_{j'} e^{V_{j'}}} \quad \forall \quad j, j' \end{aligned}$$

under the assumption that the residuals are independently and identically distributed with type 1 extreme value distribution.[13] From this, the probability of observing the child in school, P_S^0, may be written as

$$Pr(j = S) = \frac{e^{V_S}}{e^{V_S} + e^{V_W}} \tag{3.4}$$

where from (3.2) and (3.3)

$$\begin{aligned} V_S &= \alpha_S(Y - p_1) + \alpha_1 H_1^S \\ V_W &= \alpha_W(Y + w_1) + \alpha_1 H_1^W \end{aligned}$$

Therefore (3.4) can be written more fully as

$$\begin{aligned} Pr(j = S) &= \frac{e^{\alpha_S(Y-p_1)+\alpha_1 H_1^S}}{e^{\alpha_S(Y-p_1)+\alpha_1 H_1^S} + e^{\alpha_W(Y+w_1)+\alpha_1 H_1^W}} \\ &= \frac{1}{1 + e^{\alpha_W(Y+w_1)+\alpha_1 H_1^W - \alpha_S(Y-p_1) - \alpha_1 H_1^S}} \qquad (3.5) \end{aligned}$$

The marginal effect of the own child wage on the probability of school attendance is obtained by differentiating (3.5) with respect to w_1

[13] See McFadden (1974).

$$\frac{\partial Pr(j=S)}{\partial w_1} = \frac{-\alpha_W\left(e^{\alpha_W(Y+w_1)+\alpha_1 H_1^W-\alpha_S(Y-p_1)-\alpha_1 H_1^S}\right)}{\left(1+e^{\alpha_W(Y+w_1)+\alpha_1 H_1^W-\alpha_S(Y-p_1)-\alpha_1 H_1^S}\right)^2}$$

$$= -\alpha_W P_S^0(1-P_S^0) \quad (3.6)$$

which is always negative.[14]

3.2.2 Households with two children

Consider now a household with two children. In this case, one may think of the *pair* of activities of the two children as the choice variable. There are four such pairs. U_{jk}^* is an underlying latent variable denoting the level of utility for household h associated with child 1 being in activity j and child 2 in activity k

$$U_{jk}^* = V_{jk}(C_{jk}, H_1^j, H_2^k) + \zeta_{jk} \quad \text{for} \quad j = S, W, \quad k = S, W \quad (3.7)$$

where C_{jk} is household consumption, whose value is again a function of the set of activities of the two children; H_1^j, H_2^k are the human capital levels of child 1 and 2 respectively and ζ_{jk} is a residual that captures unobserved variations in tastes and in the characteristics of alternatives and errors in the optimisation by the parent.

I approximate V_{jk} in (3.7) using a form that is linear in the parameters

$$V_{jk} = \beta_{jk}C_{jk} + \beta_1 H_1^j + \beta_2 H_2^k \quad \text{for} \quad j,k = S, W \quad (3.8)$$

where

$$C_{jk} = Y - 1.(j=S)p_1 + 1.(j=W)w_1 - 1.(k=S)p_2 + 1.(k=W)w_2 \quad (3.9)$$

where Y denotes the exogenous earnings of the parent.

[14]Note also that the derivative is largest when $P_S^0 = 1 - P_S^0$, which occurs when $P_S^0 = \frac{1}{2}$, and becomes smaller as P_S^0 approaches zero or one.

The probability of observing a particular set of choices $\{j, k\}$ for the two children is denoted by P_{jk}. This pair of choices will be observed if it yields the maximum utility to the parent across all possible pairs of choices

$$\begin{aligned} P_{jk} &= Pr(U^*_{jk} > U^*_{j'k'}) \\ &= Pr(\zeta_{j'k'} < V_{jk} - V_{j'k'} + \zeta_{jk}) \\ &= \frac{e^{V_{jk}}}{\sum_{j',k'} e^{V_{j'k'}}} \quad \forall \quad j, j', k, k' \end{aligned}$$

again under the assumption that the residuals are independently and identically distributed with type 1 extreme value distribution. From this, the probability of observing child 1 in school, P^1_S (again, the superscript denotes the number of siblings of the child), may be written as

$$Pr(j = S) = P_{SS} + P_{SW}$$

$$= \frac{e^{V_{SS}} + e^{V_{SW}}}{e^{V_{SS}} + e^{V_{WS}} + e^{V_{SW}} + e^{V_{WW}}} \tag{3.10}$$

where

$$\begin{aligned} V_{SS} &= \beta_{SS}(Y - p_1 - p_2) + \beta_1 H^S_1 + \beta_2 H^S_2 \\ V_{SW} &= \beta_{SW}(Y - p_1 + w_2) + \beta_1 H^S_1 + \beta_2 H^W_2 \\ V_{WS} &= \beta_{WS}(Y + w_1 - p_2) + \beta_1 H^W_1 + \beta_2 H^S_2 \\ V_{WW} &= \beta_{WW}(Y + w_1 + w_2) + \beta_1 H^W_1 + \beta_2 H^W_2 \end{aligned}$$

Therefore (3.10) can be written more fully as

$$\begin{aligned} Pr(j = S) =& \left(e^{\beta_{SS}(Y-p_1-p_2)+\beta_1 H^S_1+\beta_2 H^S_2} + e^{\beta_{SW}(Y-p_1+w_2)+\beta_1 H^S_1+\beta_2 H^W_2}\right) / \\ & \left(e^{\beta_{SS}(Y-p_1-p_2)+\beta_1 H^S_1+\beta_2 H^S_2} + e^{\beta_{SW}(Y-p_1+w_2)+\beta_1 H^S_1+\beta_2 H^W_2} + \right. \\ & \left. e^{\beta_{WS}(Y+w_1-p_2)+\beta_1 H^W_1+\beta_2 H^S_2} + e^{\beta_{WW}(Y+w_1+w_2)+\beta_1 H^W_1+\beta_2 H^W_2}\right) \\ =& N/D \end{aligned} \tag{3.11}$$

The marginal effect of the own child wage on the probability of school attendance is

$$\frac{\partial Pr(j=S)}{\partial w_1} = -N.\big[\beta_{WS}\big(e^{\beta_{WS}(Y+w_1-p_2)+\beta_1 H_1^W+\beta_2 H_2^S}\big)+ \beta_{WW}\big(e^{\beta_{WW}(Y+w_1+w_2)+\beta_1 H_1^W+\beta_2 H_2^W}\big)\big]/D^2$$

$$= -P_S^1\left[\beta_{WS}P_{WS} + \beta_{WW}P_{WW}\right] \tag{3.12}$$

In this model, allowing the coefficients on household consumption (i.e. household income) to vary across activities, renders sibling composition relevant for choices for the following reason. The imposition of common coefficients on household consumption implies that investment in education is independent of current family income and household structure. This can be seen by comparing the probabilities of school enrolment in the two households, (3.5) and (3.11), which are the same in both households and are invariant to changes in the value of Y if $\alpha_S = \alpha_W = \beta_{SS} = \beta_{SW} = \beta_{WS} = \beta_{WW}$.[15] However, in the more plausible case in which liquidity constraints exist, current household income will affect education choices. I incorporate this by allowing the parameters on consumption in each choice to be different. Changing the value of Y in this case not only changes (3.5) and (3.11), but also affects them differently in different household types.

In order to compare the wage elasticities in the one-child and two-child households, I compare expressions (3.6) and (3.12) under the assumptions that

$$\beta_{SS} = \alpha_S, \quad \beta_{WW} = \alpha_W, \quad \beta_{WS} = \beta_{SW}, \quad \alpha_1 = \beta_1 = \beta_2$$
$$w_1 = w_2 = w, \quad p_1 = p_2 = p$$

The first two equalities imply that the utility from consumption is the same in each household, conditional on all children in the household either being

[15]Indeed they only change in response to changes in the parameters or school costs, but are always equal across household types. This is the standard pure investment model of schooling in which there are no liquidity constraints and income does not matter for education, with parents investing in each child's schooling until the expected marginal returns are equal to the marginal costs.

in work or in school and controlling for family size. The assumption that $\beta_{WS} = \beta_{SW}$ means that in a two-child household in which only one child works, the parent attains the same utility from consumption regardless of which child is contributing to income. Assuming that $\alpha_1 = \beta_1 = \beta_2$ is equivalent to saying that a unit of human capital is valued equally by the parent regardless of which child it accumulates to, and that human capital is valued equally across both households. The assumption that the wages and direct costs of education are the same across all children is relevant in the case in which the children within the household are of similar ages. I also assume that

$$\alpha_S > \alpha_W$$
$$\beta_{SS} > \beta_{SW} > \beta_{WW}$$

This is akin to assuming that if the parent were somehow to face the same consumption value for each possible activity or set of activities, then the utility from sending the child to school is higher than the utility from sending the child to work, due to the accumulation of human capital in school and not in work. In the two-child household, the option of having one child in school and one in work is also preferred to having both in work, and less preferred to having both in school, again due to the value the parent places on human capital accumulation.[16] Under these assumptions

$$\left|\frac{\partial P_S^0}{\partial w}\right| < \left|\frac{\partial P_S^1}{\partial w}\right|$$

As the changes in the choice probabilities must sum to zero when an exogenous variable changes, it follows that

$$\frac{\partial P_W^0}{\partial w} < \frac{\partial P_W^1}{\partial w}$$

This simple static model points towards the wage having a more adverse effect on education, and a correspondingly more positive effect on work participation, in large households compared to its effect in small households. The model has been developed in order to focus specifically on a comparison of the wage elasticities of children from different household sizes, which

[16]See for example Train (1986).

is the empirical focal point of this chapter. Whilst we have seen that it is theoretically plausible to expect the wage elasticity to be increasing in sibling size, there are however, other sibling-related factors that are not accounted for and that are likely to be important channels through which sibling size and structure are relevant to the wage elasticity. In particular, the activities of other siblings are likely to be important. For example, conditional on sibling size and structure, if many of the siblings are working and contributing to household income, it is intuitively plausible to expect the wage elasticity of a particular child to be different to the case in which many of the siblings are in school.[17] The extension of the model to incorporate a more complete analysis of the interactions between siblings and child wages, will be an important extension of this work.

3.3 Estimation Methodology

In the two-child case outlined above, the *set* of activities of children in the household is the dependent variable. In a conditional logit model, this would mean that there are four outcomes to consider and therefore four dependent variables in the estimation. Were I to extend it to a three-child household, there would be eight possible combinations of activities, and so on, with the number of dependent variables proliferating rapidly. In a multinomial logit model on the other hand, in which the choice variable is the activity of one child only, there are only as many dependent variables as there are choices for the child, and in this sense the complexity of the estimation is reduced.[18] Therefore, I consider the choice variable to be the activity of one child only.

[17]Note also that we have not distinguished between income and substitution effects of the wage. Indeed, structural modelling of child labour supply is an important future research agenda.

[18]A further complication in the conditional logit model is that for the two-child household, for example, it would necessitate having access to the values of the wages and the costs of schooling for both children under all four sets of activities: SS, SW, WS and WW. Whilst it is plausible to assume that the spot wage in the market is the same for a child, regardless of the activity of the other child, this is not necessarily the case for children working on the family enterprise. The child wage in this case is the marginal product of the child, which may vary for a child depending on whether the other child is in work or in school, and strong assumptions would need to be made in order to impute it.

I classify the number of siblings in the household into distinct groups on the basis of the age and gender of the siblings, and allow for the effect of the wage to vary by sibling composition, through interacting the wage with the number of siblings in each group. Under the above theoretical model, I expect the effect of the wage on the activity of the child, to be increasing in the number of siblings. The essential difference is that the independent variables are no longer choice-specific, whilst the parameters are. The model is therefore estimated using a multinomial logit.[19]

In the empirical estimation, the choice set of the parent in household h ($h = 1, \cdots, H$) for child i consists of three mutually exclusive and exhaustive categories: full- and part-time school (S), full-time work (W), and some other non-school activity, such as housework or looking for work (O).[20] I denote the activity of child i in household h as y_{ih}. I condition the analysis on a vector of observed child, household and locality characteristics, X_{ihv}, that I expect to be relevant factors in the decision. In addition, I allow for a common unobserved household-level component to choices across children within the same household. One can think of this as representing underlying parental preferences, or unobserved differences in the costs of and returns to education across households, that may affect the maximisation of household welfare and therefore observed choices.[21] I approximate the latent utility to household h from activity k for child i by

$$U^*_{ihk} = \gamma_{1_k} w_{iv} + \sum_g \gamma_{2g_k} n_{gh} + \sum_g \gamma_{3g_k} (w_{iv} * n_{gh}) + \gamma_{4_k} X_{ihv} + \epsilon_{ihk} \quad (3.13)$$

where the child wage w_{iv}, depends on the village in which he lives - I return to this below - n_{gh} represents the number of siblings in the household in group g and $w_{iv} * n_{gh}$ denotes interactions between the child wage and the number of group g siblings in the household. To allow for a common household-level unobserved component, I assume that the random component of utility is

[19]See Amemiya (1981) for a thorough review of the multinomial logit model.

[20]Three activities can readily be incorporated into the theoretical model, but the expanded choice set would yield no new insights of particular relevance here.

[21]Ignoring such unobserved heterogeneity in nonlinear models may lead to biased parameter estimates. This is along the lines of Deb and Rosati (2002).

characterised by a one-factor structure

$$\epsilon_{ihk} = \phi_k \lambda_h + \zeta_{ihk}$$

where λ_h is an unobserved latent factor representing household preferences, for example, and ζ_{ihk} is a random component to choices. I assume that

- **Assumption 1** The random variable λ_h is independent of $\zeta_{ihk}\ \forall\ i, h, k$, and all λ_h and ζ_{ihk} are independent across households and individuals.
- **Assumption 2** The term ζ_{ihk} is an extreme value random variable and is independent of all other $\zeta_{i',h'',k'''}$ except for $i = i', h = h'', k = k'''$.

The probability that child i in household h is observed in activity k is therefore

$$Pr(y_{ih} = k | w_{iv}, n_{gh}, X_{ihv}, \lambda_h) =$$

$$\frac{exp[\gamma_{1_k} w_{iv} + \sum_g \gamma_{2g_k} n_{gh} + \sum_g \gamma_{3g_k}(w_{iv} * n_{gh}) + \gamma_{4_k} X_{ihv} + \phi_k \lambda_h]}{\sum_{k'} exp[\gamma_{1_{k'}} w_{iv} + \sum_g \gamma_{2g_{k'}} n_{gh} + \sum_g \gamma_{3g_{k'}}(w_{iv} * n_{gh}) + \gamma_{4_{k'}} X_{ihv} + \phi_{k'} \lambda_h]} \tag{3.14}$$

for all k and k'.

The joint probability of children's activities in household h is the product of (3.14) over all children in the household

$$l_h = \prod_{i=1}^{n_h} \left[\prod_k \left(Pr(y_{ih} = k | w_{iv}, n_{gh}, X_{ihv}, \lambda_h) \right)^{1.(y_{ih}=k)} \right] \tag{3.15}$$

where n_h denotes the number of children in the household. As λ_h is unobserved, it is integrated out of (3.15). To avoid any potential bias arising from the incorrect specification of a functional form for $f(\lambda_h)$, I model it non-parametrically by assuming that it has discrete support and that households can be described by a finite number, M, of latent types, with proportions π_m of each type. One may think of the discrete distribution as an approximation to some underlying continuous density (although it is also intuitive to think of the underlying population as being divided into M discrete groups).[22] The contribution of household h to the log likelihood function is therefore

$$L_h = log(\sum_{m=1}^{M} \pi_m \prod_{i=1}^{n_h} (\prod_{k} [Pr(y_{ih} = k|w_{iv}, n_{gh}, X_{ihv}, \lambda_h = r_m)]^{1.(y_{ih}=k)})) \tag{3.16}$$

and the sample log likelihood is the sum of (3.16) across all households in the sample, $L = \sum_{h=1}^{H} L_h$.

The above exposition takes the values of each of the conditioning variables as known. However, the child wage is clearly observed for working children only, in which case it is endogenous for this sub-sample.[23] I obtain an estimate of the available wage for each child in the sample from the wages of working children, by regressing the observed child wages on the age and gender of the entire sample of children, along with the average adult agricultural wage in the village and state dummy variables. I therefore estimate the following equation by OLS

$$w_{iv} = \theta_0 + \theta_1 D_s + \theta_2 Z_{iv} + \theta_3 A_v + \nu_{iv} \tag{3.17}$$

where w_{iv} is the wage of child i in village v, D_s is a vector of state dummies, Z_{iv} is a vector of child characteristics (age, age squared and gender), A_v is the average agricultural adult wage in the village, which is the within-village average of the wages of household heads who report working in agriculture, and ν_{iv} is a random error term affecting wages, with $E(Z_{iv}, \nu_{iv}) = E(A_v, \nu_{iv}) = 0$ for all i, v. The predicted $\hat{w}_{iv}$ from (3.17) is the estimate of the potential child wage. The key identifying assumption is that the average agricultural adult wage in the village affects the activity of the child only through its effect on the child wage. This is confirmed by very low partial $\bar{r}^2$'s from regressions of the main activity of the child on the instrument, after con-

[22]See Heckman and Singer (1985).

[23]In any case, interest centres on the effect of the wage at the extensive rather than at the intensive margin. See Bhalotra (2000) who uses a selection procedure to correct for the endogeneity of the wage and to examine the responsiveness of work hours of working children to the wage.

trolling for all other exogenous variables, including average village wealth.[24] The results from this regression are displayed in table 3.1, for both 8 to 12 year olds and 13 to 17 year olds. Whilst the predictive power of the instrument is confirmed - the child wage is closely related to the adult agricultural wage, being on average between 60 to 72 per cent of the adult agricultural wage - this is also suggestive that the child and the adult labour markets are closely linked and for reasons discussed above, leads to concerns about the exclusion restriction assumption. Future work will address this issue further.

3.4 Data

I estimate the model using cross-sectional socio-economic census data that was collected across marginalised rural areas throughout Mexico between 1996 and 1999. This survey - the Survey of Household Socio-Economic Characteristics (Encuesta de Caracteristicas Socioeconomicas de los Hogares, ENCASEH) - was conducted with a view to aiding in the targeting of marginalised localities and households that would be eligible for the implementation of a welfare programme, Progresa, that was introduced in selected villages across seven states in rural Mexico in 1998. The data contains a cross-section of information on a large variety of individual and household characteristics and is supplemented by locality data that can be matched to households. To make the sample sizes manageable, I restrict the analysis to the seven states in which Progresa was subsequently (selectively) imple-

[24] The theoretical validity of this instrument is, however, questionable on the basis that factors that affect the adult labour market are also likely to directly affect the child labour market, apart from their indirect impact through their effect on the child wage. To the extent that the adult wage proxies for such factors, it may therefore be expected to have a direct impact on the activity of the child. The importance of such factors is likely to depend on the degree of substitutability in production in adult and child labour markets (see for example Basu and Van (1998)). Failure to adequately control for common labour market factors in the child activity regressions may render the exclusion restriction invalid, with the coefficient on the predicted child wage comprising both the effects of the wage and of other labour market influences that are common to both children and adults. To the extent that such labour market factors are difficult to observe and to control for, this is of some concern here.

mented. This data covers just over 1.5 million 8 to 17 year olds across just over 1 million households in approximately 13,000 rural villages.[25]

I consider the activities of 8-12, 13-14 and 15-17 year olds, separately by gender. These age ranges broadly correspond to primary, post-primary and secondary school ages.[26] Access to primary education in Mexico is universal, as seen in the high net primary enrolment ratio of 103% in 2000-2001.[27] The net secondary enrolment ratio in 2000-2001 was 60%. The economic activity of children may be any one of full-time school, full-time work (either formal market work or work on the family enterprise), part-time school (which includes part-time work), or some other activity. Table 3.2 shows the main activity of the child, by age group and gender. Until age 12, full-time school enrolment is approximately 88% for both males and females, full-time work participation is extremely low at between 0.8% and 1.6%, and participation in other activities is slightly higher for females than for males, at 8.8%. Full-time school enrolment drops considerably after age 12, for both males and females, with an even sharper drop after age 14.[28] At the same time, the proportions in full-time work increase, more noticeably for males. The proportions in the other category increase slightly for males, and much more so for females, suggesting that this activity for females largely comprises housework.[29] Males are more likely to be in part-time school throughout, although the proportions are generally low, at between 2.5% and 6.8%. A closer look at the activities by age and gender in table 3.3 shows that by age 17, 52% of females report other as their main activity, compared to 22% re-

[25] Whilst ideally I would like to consider the activities of 5-7 year olds as well, information on work participation is only collected for children aged 8 plus.

[26] After age 17, the proportion enrolled in school is approximately 2% for females and 4% for males across all of Mexico.

[27] This is the proportion of the population of the official age for primary education, according to national regulations, that is actually enrolled in primary school. See the Human Development Report 2003: www.hdr.undp.org.

[28] This was the primary motivation behind the Progresa intervention in Mexico - to provide financial incentives at the point in the education cycle at which individuals were at a high risk of dropping out, in the transition from primary to secondary school.

[29] This is consistent with many studies on gender time allocation in LDCs. See for example, Pitt and Rosenzweig (1990) and references therein.

porting full-time work and 23% reporting full-time school. For males on the other hand, 57% are in full-time work by age 17 compared to 17% in other. The precise activities of males in the other category are unclear. They may consist of looking for work and leisure/idleness. However, the proportions in this activity are quite substantial by age 17, and given that the sample consists of individuals living in the poorest marginalised communities in rural Mexico, the extent of leisure is likely to be very low. I therefore consider the other category for males to comprise a substantial number of children who are looking for work.[30]

In the empirical analysis, as the proportions in part-time school are relatively low and also relatively stable throughout, I combine the full- and part-time school categories (FS and PS) and examine the determinants of full-/part-time school (S), full-time work (W) and other (O). For 8 to 12 year olds, due to fact that the proportions in full-time work are extremely low, I consider only two possible choices: school (full- or part-time) and other (which includes full-time work).

A key focus of this analysis is the relationship between the number of siblings and the economic activities of children. As can be seen in table 3.4, the percentage of the sample residing in large households is substantial. Approximately 27% of 8 to 12 year olds and 30% of 13 to 17 year olds have five or more siblings. This is useful for the analysis insofar as it allows me to examine differences in economic activities at low and high numbers of siblings, without having to rely on out-of-sample estimates and a corresponding compromise in precision.

A list and description of the variables used in the analysis, disaggregated

[30]In terms of providing informal evidence that the other category consists largely of males looking for work, it should be noted that males in both the other and work categories are more likely to have less well-educated parents compared to those in school (the mean parental years of education are 2.3, 2.1 and 3.2 respectively). Further, males reporting other as their main activity, are ten percentage points less likely than those in work to live in households that own land, thus suggesting that access to work may be more difficult for them and job search may subsequently be higher.

according to whether they pertain to the individual, the household or the locality, is provided in table 3.5.[31] I control for characteristics of the household head, including age, gender, years of education and work status. Other household level variables include the presence of parents in the household, ownership of the house and the number of adults present. Locality-level characteristics include the proportion working in agriculture, the proportion owning land, literacy and fertility rates, the proportion of females working, the average village income and its cross-sectional coefficient of variation.

In tables 3.6 and 3.7, I compare the characteristics of individuals across the three activities, for females and males respectively. To briefly summarise them, the average age of the full-time male worker is 15.4 years, compared to 11.4 for males in full-time school. For the male sample, the age of the household head is lower in households with full-time child workers and years of education of the household head are also lower. The proportion of household heads in agriculture is higher in households with a full-time child worker, as is land ownership. The locality characteristics are broadly similar across activities. The patterns for females are similar, although female full-time workers are not noticeably more likely than non-workers to live in agricultural/land-owning households, suggesting that it is males who are more likely to help out on the family farm.

3.5 Results

I estimate (3.16) separately by gender within each age group.[32] After some experimentation, I find that the model with three points of support fits

[31] Unfortunately, the quality of data on school characteristics at the village level is poor and cannot be used.

[32] Note that n_h in (3.16) will refer to the total number of children in the household, within the sub-sample of interest. Likelihood ratio tests that make pairwise comparisons by gender, between the estimates within each age group, indicate highly significant structural differences in activities between the genders and strongly reject the pooling of males and females. Note that in the estimation, the coefficients in the other category are normalised to 0 for identification. The coefficient on the factor loading in school, ϕ_S, is normalised to 1 in order to identify ϕ_W and the points of support, r_m.

the data best.[33] Due to the non-linearity of the multinomial logit model, the marginal effect of a variable on the choice probabilities depends on the full pattern of coefficient estimates, and neither sign nor magnitude can be directly deduced from the estimates. I use the coefficient estimates to derive the marginal probabilities for each explanatory variable. The marginal effect of a regressor X_m on the probability of choice j is

$$\frac{\partial P_j}{\partial X_m} = P_j[\beta_{jm} - \sum_{k=1}^{K} P_k \beta_{km}]$$

where P_j is the probability of choice j, β_{jm} is the multinomial logit coefficient for variable m in state j and K is the total number of choices.[34]

Tables 3.8, 3.10, 3.12, 3.14 and 3.16 display the marginal effects of the main parameters of interest - the sibling groups, the child wage, and interactions between them - for all six subgroups.[35] The interpretation of the interactions is complicated by the fact that they are both continuous variables. To simplify interpretation and to gain some intuition as to their magnitudes, for each sub-sample I simulate the proportions in each of the activities at two values of the child wage - the 25th and 75th percentiles - and four values of the number of siblings per group - 0, 1, 2 and 3. These

[33]I compare models with one, two and three points of support on the basis of the Bayesian Information Criterion. Three points of support yields the best fit for all subsamples. The values of the log likelihood functions change very little when I increase the number of points of support to four.

[34]Standard errors of the marginal effects are constructed by calculating the marginal effects for each of 500 parameter vectors drawn from a multivariate normal distribution with mean given by the point estimates of the parameters and covariance matrix given by the sandwich estimate. The standard deviations of the sample of marginal effects are the estimates of the standard errors.

[35]The reported standard errors have been adjusted for the fact that the child wage is predicted. The sign and significance of the remaining variables, which are included as controls, are broadly in line with previous findings in the literature. To summarise them, household-level variables that are generally favourable for school enrolment include the presence of more adults in the household, having a more educated household head, the presence of a mother or both parents in the household, land ownership, full ownership of the house and an agricultural head who owns land. Negative household-level factors include the head working in agriculture as a labourer and having a father but no mother present in the household.

values are chosen to represent the width of the distributions of the wage and sibling composition. As mentioned already, the large number of observations is advantageous in the sense that I observe a non-negligible number of children with the maximum number of siblings (three) in a particular group and therefore the estimates that I obtain do not rely on out-of-sample predictions. This is however not the case for the 13 to 14 year old sibling age range, for whom only a very small proportion of each of the sub-samples has three such siblings. In interpreting the estimates for these it should therefore be borne in mind that they may be less precise due to out-of-sample predictions. I set the remainder of the variables not involved in the particular interaction to their mean values. The predicted proportions in each of the activities are displayed in the tables immediately after the marginal effects, for each sub-sample.

In order to elucidate the main patterns for the reader, I confine comparisons in the forthcoming sections, to the two extremes of the sibling distribution, rather than making comparisons across all four values (which may be seen in the tables referred to above). First, in order to highlight the relationship between the number of siblings and participation in each of the activities, I display the difference in participation in each of the activities across 0 and 3 siblings. Comparisons are made separately at low and high child wages, which provides insight into whether the relationship between the number of siblings and economic activities varies according to the value of the child wage. I therefore evaluate $Pr(y_{ih} = j|n_g = 3, w_p) - Pr(y_{ih} = j|n_g = 0, w_p)$, where y_{ih} denotes the activity of individual i in household h, n_g denotes the number of siblings in group g and $p = l, h$, where w_h is the 75^{th} percentile child wage and w_l is the 25^{th} percentile wage. These figures are reported in the upper panel of each of the tables that follow. Second, in order to directly focus on whether the wage elasticities vary with the number of siblings, I compare the change in the proportions in each of the activities with respect to the wage, across different sibling sizes. I therefore calculate $Pr(y_{ih} = j|n_g = r, w_h) - Pr(y_{ih} = j|n_g = r, w_l)$, for $r = 0, 3$. These figures are presented in the lower panel of each table.[36]

[36]Note that in all of the below, unless otherwise stated, all differences are significant at

8-12 Year Olds: Females and Males

I begin by discussing the results for 8 to 12 year olds. Marginal effects on school participation are displayed for both females and males in table 3.8, whilst table 3.9 shows changes in school enrolment in response to sibling composition and number, at different wage levels. In the table below, I summarise the key findings in the way discussed above.

8-12 Year Olds

		Females		**Males**	
$Pr(y_{ih} = S \mid n_g = 3, w_p) - Pr(y_{ih} = S \mid n_g = 0, w_p)$					
		$p = l$	$p = h$	$p = l$	$p = h$
Siblings					
$0 - 4$		-3	-5	-3	-5
$5 - 12$	F	0	1	-1	-1
	M	0	0	-2	-2
$13 - 14$	F	2	2	2	2
	M	-1	1	-1	0
$15+$	F	0	0	1	0
	M	-3	-4	-2	-4
$Pr(y_{ih} = S \mid n_g, w_h) - Pr(y_{ih} = S \mid n_g, w_l)$					
		$n_g = 0$	$n_g = 3$	$n_g = 0$	$n_g = 3$
$0 - 4$		0	-2	1	-1
$5 - 12$	F	-1	0	0	0
	M	-1	-1	0	0
$13 - 14$	F	0	0	1	1
	M	-1	1	0	1
$15+$	F	-1	-1	1	0
	M	0	-1	1	-1

S=full-/part-time school, n_g=number of group g siblings, F=female, M=male, w=wage, l=low, h=high.

There are some noteworthy points in this table (nonetheless the results

conventional levels.

are less remarkable than for the older age groups, which largely reflects the near-universality of primary education in Mexico). In the upper panel, we see that there is significantly lower enrolment in school amongst both male and female children who have a relatively high number of very young siblings, at both low and high levels of the wage. For males only, having a greater number of 5 to 12 year old sisters or brothers is also associated with lower school enrolment. Having more sisters aged 13 to 14 is associated with more school participation for both males and females, suggesting that having such older siblings frees up resources for education - however for females, this is insignificant at conventional levels. Sisters aged 15 plus are neutral to education, whilst brothers aged 15 plus are associated with significantly lower participation in education. This is somewhat surprising, and as we will see and discuss below, is borne out across all sub-samples.

As can be seen in the lower panel of the table, the wage elasticities are generally low - changes in school participation in response to changing the wage from the 75th to the 25th percentile are small - and there is no systematic relationship for these subgroups, between the wage elasticities and sibling composition.

13-14 Year Olds: Females and Males

I now turn to the findings for male and female 13 to 14 year olds, which are summarised in the table below.[37] This analysis is more informative due to the fact that for the 13-plus age ranges, participation in education is substantially lower than for primary-school children and this allows me to examine a more disaggregate choice set which includes school, work and

[37] As discussed in section 3.4, it is reasonable to expect the other activity to represent housework for females and looking for work for males. From hereon, I therefore refer to the terms interchangeably.

other.

13-14 Year Olds

		Females						Males					
		$Pr(y_{ih} = j \mid n_g = 3, w_p) - Pr(y_{ih} = j \mid n_g = 0, w_p)$											
		$p = l$			$p = h$			$p = l$			$p = h$		
	$j =$	S	W	O	S	W	O	S	W	O	S	W	O
Siblings													
$0-4$		-11	2	9	-13	3	10	-9	4	5	-11	5	6
$5-12$	F	-3	2	1	-5	1	4	-3	2	1	-5	2	3
	M	-3	1	2	-6	1	5	-3	2	1	-5	3	2
$13-14$	F	6	1	-7	-3	4	-1	-4	0	4	-7	2	5
	M	4	-1	-3	0	0	0	-2	3	-1	-4	5	-1
15+	F	2	2	-4	-1	1	0	1	-2	1	-1	-3	4
	M	-8	0	8	-9	0	9	-4	3	1	-5	4	1
		$Pr(y_{ih} = j \mid n_g, w_h) - Pr(y_{ih} = j \mid n_g, w_l)$											
		$n_g = 0$			$n_g = 3$			$n_g = 0$			$n_g = 3$		
		S	W	O	S	W	O	S	W	O	S	W	O
$0-4$		1	0	-1	-1	1	0	-3	0	3	-5	1	4
$5-12$	F	1	1	-2	-1	0	1	-3	1	2	-5	1	4
	M	1	1	-2	-2	1	1	-2	0	2	-4	1	3
$13-14$	F	1	0	-1	-8	3	5	-2	0	2	-5	2	3
	M	1	0	-1	-3	1	2	-3	0	3	-5	2	3
15+	F	1	1	-2	-2	0	2	-3	1	2	-5	0	5
	M	1	0	-1	0	0	0	-2	0	2	-3	1	2

S=full-/part-time school, W=full-time work, O=other, n_g=number of group g siblings, F=female, M=male, w=wage, l=low, h=high.

As can be seen from the above, for 13 to 14 year olds, having more siblings is generally, with only a few exceptions for females, associated with significantly lower participation in education. There are, however, noticeable differences across sibling types and the lower enrolment in education is exacerbated at high values of the child wage. This underlines the importance of considering interactions between the wage and sibling composition. The negative association between participation in education and the number of siblings is most pronounced for 0-4 year old siblings. This is intuitively reasonable, insofar as they represent a pure drain on household resources and are unable to contribute to household income, whilst siblings above this age are less disadvantageous in the sense that they are more likely to be able to perform some types of (direct and indirect) income-generating tasks. Females are more likely to engage in housework at the expense of education, the greater the number of very young siblings in the household, suggesting that they are likely to look after young children. Males on the other hand, are equally likely to be drawn from education into work or some other activity (look for work), suggesting that they are a more important means of directly contributing to household resources than females.

Having relatively more 5 to 12 year old siblings, either male or female, is also associated with less participation in education. This conforms to expectations, as such siblings are likely to be enrolled in primary school, therefore placing more pressure on household resources. Again it appears that females are more likely to stay at home, especially at higher values of the wage, with males equally likely to be either working or looking for work.

The importance of disaggregating siblings by gender and of allowing for their effect to vary across different wages, is borne out for both 13 to 14 and 15-plus year old siblings. Interestingly, for females who face low wages, having more sisters of the same age is associated with higher participation in education and less housework. This conforms to expectations in the sense that there are more females in the household that may be used to perform household chores, thus freeing up resources and enabling some others to participate in education. However, the benefit for education disappears at high values of the child wage, and indeed in this case, participation in education is lower the greater the number of such sisters, whilst full-time work participation is noticeably higher. This suggests that the wage incentive for work is strong enough to outweigh any potential benefits of having a large number of such siblings. For males, at low and even more so at high child wages, having more sisters of a close age is associated with lower participation in education and increased job search. Indeed the higher the wage, the higher is participation in full-time work also. The wage interactions are also important for 13 to 14 year old brothers. Such siblings are beneficial for the education of females at low wages only and have no effect at high wages. For males on the other hand, whilst they are generally detrimental to participation in education, these associations are not significantly different from zero at conventional levels.

Having a large number of older siblings is in general not beneficial for the education of 13 to 14 year olds - apart from for females who face low wages, for whom having a higher number of older sisters is associated with significantly more enrolment in education, but also more participation in full-time work - however the magnitudes are low. At high wages, there are fewer discernible differences in the activities of females, across different numbers of older sisters. For males facing high wages, having older sisters is associated with less participation in full-time work and a higher propensity to look for work. Having a large number of older brothers is markedly disadvantageous for female education, with the reallocation transferring entirely to housework. For males, having more older brothers is also associated with less participation in education and increased full-time work, although the magnitudes are smaller than for females. The findings for older brothers are somewhat surprising, particularly so for older males, whom one might expect to contribute substantially to household resources and thus to enable others to receive more education. The adverse effect of older brothers for education is observed for both females and males, so the favouring of boys for education due to higher perceived returns is an unlikely explanation. However, males tend to participate more in full-time work, the higher the number of older brothers they have, suggestive of work complementarities between males within the household, or some unobserved parental preferences for work. Future work will investigate these issues further.

Whilst the importance of the interactions between the wage and sibling composition are inherent in the above analysis, the lower panel is more directly informative as to the own-wage elasticities of the various activities. We see that for females, the response of enrolment in education to an increase in the wage is actually positive or zero when the number of siblings is low, and negative at a very high number of siblings. Further, the wage elasticity appears to be increasing in the number of same-age (and particularly same-gender) siblings only, and otherwise is relatively invariant to sibling size. For males, the wage elasticity of education is negative throughout and increasing in sibling size, as can be seen from the relatively larger changes in participation in activities in response to changes in the wage, at higher numbers of siblings. Further, the higher the wage, the higher the job search of males, whilst the effect of the wage on participation in full-time work is small and insignificant. These findings of wage elasticities that are increasing in sibling size are consistent with the model in section 3.2 and highlight the potentially important interactions between schooling costs and household composition.

15-17 Year Olds: Females and Males

I now turn to the results for 15-17 year olds, the key points of which are summarised below.

15-17 Year Olds

		Females						**Males**					
		$Pr(y_{ih} = j \vert n_g = 3, w_p) - Pr(y_{ih} = j \vert n_g = 0, w_p)$											
		$p = l$			$p = h$			$p = l$			$p = h$		
	$j =$	S	W	O	S	W	O	S	W	O	S	W	O
Siblings													
$0-4$		-9	4	5	-11	5	6	-8	6	2	-10	8	2
$5-12$	F	-3	4	-1	-7	5	2	-2	3	-1	-5	5	0
	M	-3	3	0	-6	3	3	-1	2	-1	-5	5	0
$13-14$	F	-2	8	-6	-6	9	-3	-4	4	0	-8	5	3
	M	0	1	-1	-5	2	3	-1	3	-2	-4	4	0
15+	F	3	6	-9	-2	7	-5	-2	-3	5	-5	-1	6
	M	-5	0	5	-7	1	6	-3	6	-3	-6	7	-1
		$Pr(y_{ih} = j \vert n_g, w_h) - Pr(y_{ih} = j \vert n_g, w_l)$											
		$n_g = 0$			$n_g = 3$			$n_g = 0$			$n_g = 3$		
		S	W	O	S	W	O	S	W	O	S	W	O
$0-4$		2	-1	-1	0	0	0	0	0	0	-2	2	0
$5-12$	F	2	-2	0	-2	-1	3	0	0	0	-3	2	1
	M	1	-1	0	-2	-1	3	0	0	0	-4	3	1
$13-14$	F	2	-1	-1	-2	0	2	0	1	-1	-4	2	2
	M	2	-2	0	-3	-1	4	-1	1	0	-4	2	2
15+	F	1	-1	0	-4	0	4	0	0	0	-3	2	1
	M	2	-2	0	0	-1	1	0	1	-1	-3	2	1

S=full-/part-time school, W=full-time work, O=other, n_g=number of group g siblings, F=female, M=male, w=wage, l=low, h=high.

Similarly to the findings for 13 to 14 year olds, having more siblings, regardless of their age or gender, is associated with less education for 15 to 17 year olds. The magnitudes, however, again vary across different sibling groups and the associations are generally also stronger at higher wages. The main difference between the results for the 13 to 14 and 15 to 17 year old children, is that for the latter group, the observed lower participation in education feeds through mainly to higher participation in full-time work, particularly for males.

In line with the previous results, very young siblings are most disadvantageous for participation in education. It is interesting, however, that females are now equally likely to be in full-time work or housework, if they have a high number of such siblings. Presumably this is because it is easier and more acceptable to send older females to full-time work whilst younger females are more likely to be engaged in household tasks - as shown in the previous results. Males on the other hand, are more likely to be in full-time work at the expense of education, particularly so at high wages.

For females, having a large number of 5 to 12 year old siblings (regardless of the gender of such siblings) is mainly associated with more full-time work, although when the child wage is high, they also engage in more housework, and enrolment in education is noticeably lower. For males, there is again more participation in full-time work, the higher the number of such siblings, and again more so at high wages, with subsequent adverse implications for enrolment in education.

For females, having many 13 to 14 year old sisters is associated with lower participation in both education and housework and higher participation in full-time work, at both low and high wages. This is consistent with younger siblings fulfilling household chores and older siblings being displaced into formal work instead. For males, having more 13 to 14 year old sisters is also mostly associated with increased work participation and at high wages it also feeds through to increased job search. For females facing high wages, having more brothers aged 13 to 14 is associated with increased work and housework, whilst at low wages there are no discernible differences in activities across differing numbers of such siblings. Males are more likely to be observed in full-time work if they have a large number of 13 to 14 year old brothers, regardless of the wage they face, although at high wages, education is most adversely affected. Again, all of these results point to important resource competition within the household, but with such competition being an important function of the opportunity costs of school for children.

In terms of siblings of the same age or older, for females there are important interactions with the wage. It is somewhat beneficial for education to have more older sisters, if the child wage is low. However, the substantially lower housework is also associated with higher participation in full-time work. At high wages however, the observed limited advantage for education disappears and the overall effect is higher work participation. For males on the other hand, having more 15 plus year old sisters tends to be associated with a higher propensity to be looking for work. Similar to the previous findings for 13 to 14 year olds, same-age and older brothers are again detrimental for education. However, even though females engage more in housework as a result, these differences are insignificant at conventional levels. Males, on the other hand, participate significantly more in full-time work the higher the number of such brothers they have, irrespective of the wage. Again this is somewhat puzzling, but suggestive of complementarities between the work activities of males in the household.

Again, the lower panel of the table gives an indication as to the effect of the wage across different sibling compositions. For females with very few siblings, enrolment in education is slightly higher at high values of the wage, and becomes negative as the number of siblings increases. For males, the wage elasticities of all activities are mostly zero when the number of siblings, regardless of composition, is very low. At a large number of siblings on the other hand, work increases and education decreases in response to increases in the available wage, with less discernible impacts on looking for work.

To conclude, whilst previous literature has documented the (generally negative) association between the number of siblings and the educational attainment of children, the foregoing analysis has pointed to resource competition among siblings within the household importantly varying with the opportunity costs of education. In particular, I find that the competition across siblings for resources is strongest, the higher the number of relatively young and/or relatively old siblings, and is generally more intense at higher opportunity costs of school. The strong negative association between the number of very young siblings and education is unsurprising, given that such children represent a pure drain on resources and do not contribute to household income. However, the detrimental nature of older brothers for education is surprising, as one may expect such siblings to contribute to (or to at least free up) household resources and to thus facilitate the education of younger children. There is some evidence that there may be work complementarities between males within the household, as males are more likely to be observed in full-time work if they have more older brothers. However, this could also indicate unobserved parental preferences for work. Future work will investigate these interactions further.

3.6 Conclusion and Policy Appraisal

In this chapter, I have examined the relationship amongst the wage, sibling composition and the economic activities of children in rural Mexico. A number of key findings have emerged. Firstly, the association between participation in education and the number of siblings is generally negative, particularly for very young siblings. Secondly, for males, school participation of 13 to 14 year olds is decreasing in the child wage, and the association is stronger the more siblings the child has, regardless of the ages and genders of siblings, with most of this feeding through to increased job search. For 15 to 17 year old males with a low number of siblings, participation in each of the activities is invariant to changes in the wage. However, if the number of siblings is relatively high, school participation significantly decreases in response to increases in the wage, with most of this decrease feeding through to increased full-time work participation. For 13 to 17 year old females with relatively few siblings on the other hand, school participation is higher, the higher the wage. If the child has a relatively large number of siblings however, higher wages lead to a decrease in school participation, and there is increased participation in housework rather than in full-time work.

Whilst the associations between sibling composition and the activities of children are informative, an interesting extension of this work will be to consider how the effect of the wage varies according to the activities of other children in the household. This research is a step towards a more complete analysis of intra-household decision-making that will take account of the fact that parents make decisions about several children simultaneously, with liquidity constraints forcing competition across siblings for scarce resources. One way of gaining new insights into the allocation of resources across children, is to use the variation induced by various educational subsidy programmes in Mexico and Colombia, as a means of assessing whether subsidies introduce distortions in the allocation of schooling across household members. The particular feature of the programmes that may be useful to exploit is the following: subsidies are payable to the household up to some maximum threshold. If the household is 'technically' entitled to more than this amount, on the basis of the sum of individual entitlements within the household, one can examine how households respond to the threshold in terms of the allocation of schooling across children. I consider this to be an important area of future work.

I wrap up the preceding chapters with a brief appraisal of child labour and education policies in LDCs. A key underlying motivation for the foregoing analyses of education and work choices, has been to contribute to policy design, for which the identification of the key factors in education and work choices is central. The analyses have pinpointed the importance of insurance market interventions, along with significant variation in the effects of schooling costs on the activity of the child, by family composition. Designing the best policy response to child labour and under-investment in education, is a complex task. There is no single approach that will eradicate the pervasive phenomenon. It is widely accepted that a complete ban on child labour is a blunt instrument if imperfectly enforced, in the sense of not providing alternatives (such as subsidised education) to compensate for any short-term adversities incurred by the household due to the loss in income from children working. It may even have the unintended detrimental effect of pushing households into deeper poverty and/or of forcing children into hazardous or illegal forms of work. The problem is that a complete ban is poorly targeted and fails to address the underlying causes of child labour. More effective policy approaches recognise that the elimination of child labour will be a gradual process and not only provide income-compensated alternatives to work, but also establish legal standards and protection for working children. The 'optimal' policy is likely to comprise many elements, as substantiated by the historical decrease in child labour in developed countries, which took place under a variety of concurrent factors such as legislative bans, changes in compulsory education laws and increasing economic prosperity.

A variety of policies aimed at reducing child labour have been widely implemented in recent years. Some directly target education, others focus on actual work, and most are a combination of the two. To begin with, there has been a relatively recent move towards directly subsidising education. As one of the highest costs of attending school in less developed economies is the opportunity cost in the form of foregone wages, means-tested households are compensated for the loss in foregone earnings, by receiving subsidies conditional on school attendance. Such programmes are widely observed in many Latin American countries. The Progresa programme has been found to be effective in increasing enrolment in marginalised communities in rural Mexico, especially amongst post-primary children.[38] Such programmes of course rely on the availability of schools (notwithstanding any issues to do with school quality - indeed the provision of broad access to education, rather than improvements in school quality, is the focus of most policies). A huge deterrent to education is the lack of schools in the village, or having to travel far to school.[39] In addition, more school flexibility (for example, decreased hours

[38] It has been expanded to urban areas and renamed Opportunidades. See Schultz (2001) for evidence as to the effectiveness of Progresa. Other programmes that are currently being implemented include Familias en Accion in Colombia, Food For Education in Bangladesh and Bolsa Escola in Brazil; along with programmes in Honduras, Nicaragua and Argentina.

[39] See Grootaert and Kanbur (1995) and Lavy (1996). There is evidence that living in close proximity to a secondary school increases the probability that young children will attend school, consistent with the notion of primary school serving as a gateway to secondary school (Lavy (1996)).

during harvest time) so that children will not fall behind and subsequently drop out of school, is also worth considering in the design of such programmes. However, the evaluation of these programmes to date says little about what has happened to child labour. Whilst participation in education has increased, the effects of cheaper schooling on child labour are theoretically ambiguous and depend on the substitutability/complementarity between child labour and schooling. The increase in participation in schooling may come out of leisure time. Ravallion and Wodon (2000) find a significant negative effect of a conditional school stipend in rural Bangladesh on child labour that is, however, smaller than the strong positive effect on the probability of attending school.

A second type of policy of relevance to investment in education is intervention in credit markets. Standard human capital theory predicts that an individual will invest in schooling up to the point at which the marginal benefit of an additional year of schooling (increased future earnings) equals the marginal cost (in terms of foregone earnings). If individuals do not have full access to credit markets, they will under-invest in education. Credit markets generally do not function well in less developed economies, especially in rural areas.[40] There is thus a role for intervention in credit markets in order to make loans available for families to send their children to school. However, the provision of repayable loans is likely to have very different effects on education to the previously discussed policy of making money freely available (conditional on attending school). This is because there can be no supposition that even if borrowing constraints were relaxed and loans for education were made more widely available, parents would borrow to send their children to school. Apart from intergenerational commitment problems in terms of the inability of parents to enforce repayments (transfers) from children for past investment in education[41], households may simply not want to take on extra debt for education if they are myopic and/or focused on day to day survival.

[40] See Besley (1995a, 1995b).

[41] See Baland and Robinson (2000).

As discussed in chapter two, well-functioning insurance markets may have positive implications for investment in education. This is because insurance may cushion households against falls in consumption due to unexpected income falls. Uncertain environments and the absence of insurance, may lead households to use the labour of their children to build up buffer stocks against unforeseen events. The effects are exacerbated in imperfect capital market settings.[42] A whole set of policies aimed at developing capital markets and at improving risk-coping mechanisms and providing safety nets has been recently stressed by the World Bank.[43] However, there is a paucity of evidence on the effects of insurance market failures on child labour, partly due to the difficulty of measuring risk and of observing formal and informal insurance mechanisms within villages. In the light of the evidence presented in chapter two however, this is an area in which increased understanding might very well aid policy.

Fourthly, it is likely that trade policies will have important consequences for investment in education and child labour. Many Latin American countries have recently undergone dramatic processes of trade liberalisation (for example, Colombia in 1991-1992). This has led to LDCs increasing specialisation in the production of labour intensive goods, with a subsequent increase in the demand for low-skilled workers. The returns to college education have increased, whilst the returns to secondary education have decreased.[44] This change in incentives has important implications for investment in education. In order to realise the high returns to college education, indigent families must commit to a very high level of investment in education. However, liquidity constraints may generate under-investment in education. In addition, the riskiness of the environment means that there is generally substantial uncertainty on the part of parents over their ability to keep children in school throughout all of primary and secondary levels. Thus

[42]See Jacoby and Skoufias (1997) and Guarcello et al (2002), who show that credit rationing and shocks significantly influence child work and school attendance; see also Beegle et al (2003) for evidence of the ex-post use of child labour in response to shocks, which is exacerbated in thin credit markets.

[43]See World Bank (2001) and Holzmann and Jorgensen (2002).

[44]See Attanasio et al (2003).

the theoretical effects of altering the returns to education are ambiguous. The increase in the relative price of the exported good due to trade liberalisation, also has ambiguous predictions for the incidence of child labour.[45] Thus trade sanctions or import tariffs against countries that use child labour do not necessarily reduce the incidence of child labour.[46]

In conclusion, the recognition that policy design will be significantly advanced by the continued analysis of micro-level data, is borne out by the increased emphasis on the collection of such data for micro-level research, across a large number of LDCs. This greatly facilitates more comprehensive research into the complex array of factors that underlie child time allocation decisions.

[45]Edmonds and Pavcnik (2002) provide evidence that reductions in child labour may in fact be increasing in the price of exported goods.

[46]See the model developed in Ranjan (2001).

3.7 Tables

Table 3.1: Estimates from Equation (3.17)

	Log weekly child wage	
	8 to 12 year olds	**13 to 17 year olds**
Age	-0.2417*	0.3688**
	(0.1047)	(0.0259)
Age squared	0.0152**	-0.0101**
	(0.0050)	(0.0008)
Male	0.1552**	0.1395**
	(0.0188)	(0.0078)
Proportion of agricultural workers in village	0.2606**	0.1303**
	(0.0715)	(0.0138)
Proportion female workers aged 18+ in village	-1.3075**	-0.6260**
	(0.1130)	(0.0647)
Average village income	0.0302	0.0082
	(0.0244)	(0.0069)
Log agricultural wage of household heads in village	0.6054**	0.7176**
	(0.0406)	(0.0137)
Constant	1.6319**	-2.2660**
	(0.5949)	(0.2256)
N	9,364	123,254
R^2	0.18	0.30

Dependent variable is the log of the weekly child wage. Also control for state dummy variables.
* denotes significance at 5% level or less.

Table 3.2: Activity by Gender and Age Group

	N^i	N^{hh}	Full-time school	Full-time work	Work and School	Other
Males						
8-12	413,134	327,341	0.8776	0.0164	0.0367	0.0693
			(0.3278)	(0.1270)	(0.1881)	(0.2539)
13-14	156,433	149,729	0.6460	0.1375	0.0677	0.1489
			(0.4801)	(0.3467)	(0.2514)	(0.3594)
15-17	204,939	185,021	0.3252	0.4354	0.0600	0.1793
			(0.4651)	(0.4965)	(0.2357)	(0.3866)
Females						
8-12	399,680	317,646	0.8840	0.0080	0.0198	0.0882
			(0.3202)	(0.0890)	(0.1391)	(0.2837)
13-14	152,211	145,718	0.6076	0.0605	0.0249	0.3070
			(0.4905)	(0.2385)	(0.1539)	(0.4657)
15-17	181,409	165,342	0.3252	0.1722	0.0227	0.4799
			(0.4529)	(0.3673)	(0.1408)	(0.4991)

Table 3.3: Economic Activity by Age

	Females				Males			
	FS	**W**	**PS**	**O**	**FS**	**W**	**PS**	**O**
Age								
8	0.91	0.005	0.015	0.07	0.91	0.005	0.02	0.065
9	0.92	0.005	0.015	0.06	0.91	0.005	0.025	0.06
10	0.905	0.005	0.02	0.07	0.89	0.01	0.04	0.06
11	0.89	0.005	0.02	0.085	0.87	0.02	0.05	0.07
12	0.80	0.02	0.02	0.16	0.81	0.04	0.06	0.09
13	0.67	0.04	0.02	0.27	0.71	0.09	0.06	0.14
14	0.55	0.08	0.03	0.35	0.59	0.18	0.07	0.16
15	0.41	0.13	0.02	0.43	0.44	0.31	0.07	0.18
16	0.31	0.18	0.02	0.49	0.31	0.45	0.06	0.19
17	0.23	0.22	0.02	0.52	0.22	0.57	0.05	0.17

FS=full-time school, W=full-time work. PS=part-time school, O=other.

Table 3.4: Distribution of Siblings

	Number of siblings						
	0	**1**	**2**	**3**	**4**	**5**	**6+**
Proportions							
8-12	0.0467	0.1294	0.1981	0.1991	0.1621	0.1166	0.1480
13-14	0.0560	0.1111	0.1720	0.1917	0.1718	0.1297	0.1676
15-17	0.0674	0.1191	0.1687	0.1806	0.1599	0.1262	0.1782

Table 3.5: List and Description of Variables in the Analysis

Individual	
age	Child age
age2	Child age squared
gender	=1 if male
predcw	Potential log child wage
Siblings	
sibs0-4	Age 0 to 4 : Number of
fsibs5-12, msibs5-12	Age 5 to 12: Number of, by gender
fsibs13-14, msibs13-14	Age 13 to 14: Number of, by gender
fsibs15plus, msibs15plus	Age 15 plus: Number of, by gender
Household Head	
agehead	Age of household head
malehead	=1 if male
yrseduchead	Years of education
workhead	Works full-time
workincomehead	Income from work
agrichead	Works in agriculture
ownland	Owns land
agricpost	Works in agriculture and owns land
Other Household	
moth	Mother present
fath	Father present
mothfath	Both parents present
ownfully	Own house fully
numadults	Number of adults in household

Table 3.5 contd.

Locality	
pragricwork	Proportion adults working in agriculture
prfemwork	Proportion of females aged 18+ working
fertility	Ratio of 0-4 year olds to 15-44 year old females
prownland	Proportion owning land
prindhead	Proportion of indigenous heads
prlit15-34	Literacy rate amongst 15 to 34 year olds
meanv	Mean village wage of heads
vcv	Coefficient of variation of income of heads

Table 3.6: Mean Characteristics by Activity: Females 8 to 17

	S	W	O
Individual			
age	11.2 (2.44)	15.18 (1.75)	14.09 (2.38)
predcw	4.52 (0.36)	4.84 (0.35)	4.82 (0.36)
Siblings			
sibs0-4	0.57 (0.78)	0.49 (0.76)	0.49 (0.76)
sibs5-12	1.49 (1.12)	1.64 (1.33)	1.56 (1.28)
sibs13-14	0.39 (0.54)	0.45 (0.55)	0.39 (0.54)
sibs15plus	1.00 (1.24)	1.28 (1.21)	1.24 (1.25)
Household Head			
agehead	44.5 (11.9)	47.9 (11.8)	47.3 (11.5)
malehead	0.90 (0.29)	0.83 (0.36)	0.89 (0.30)
yrseduchead	3.33 (3.05)	2.16 (2.29)	2.30 (2.38)
workhead	0.91 (0.27)	0.88 (0.31)	0.90 (0.29)
agrichead	0.63 (0.48)	0.64 (0.47)	0.67 (0.46)
ownland	0.52 (0.61)	0.58 (0.68)	0.54 (0.62)
agricpost	0.40 (0.49)	0.43 (0.49)	0.41 (0.49)
Other Household			
moth	0.92 (0.26)	0.92 (0.25)	0.91 (0.27)
fath	0.84 (0.35)	0.80 (0.39)	0.84 (0.35)
mothfath	0.83 (0.37)	0.78 (0.41)	0.82 (0.38)
ownfully	0.90 (0.28)	0.92 (0.26)	0.91 (0.27)
numadults	2.68 (1.19)	2.93 (1.32)	2.91 (1.30)

Table 3.6 contd.

	Females 8-17		
	S	**W**	**O**
Locality			
pragricwork	0.70 (0.24)	0.71 (0.24)	0.73 (0.22)
prfemwork	0.18 (0.09)	0.23 (0.10)	0.17 (0.08)
fertility	0.61 (0.18)	0.65 (0.18)	0.62 (0.18)
prownland	0.47 (0.23)	0.49 (0.23)	0.47 (0.22)
prindhead	0.25 (0.38)	0.31 (0.41)	0.23 (0.37)
prlit15-34	0.89 (0.09)	0.86 (0.12)	0.86 (0.11)
meanv	6.28 (0.83)	6.16 (0.84)	6.27 (0.84)
vcv	0.36 (0.06)	0.37 (0.06)	0.36 (0.06)

S=full-/part-time school, W=full-time work, O=other.

Table 3.7: Mean Characteristics by Activity: Males 8 to 17

	S	W	O
Individual			
age	11.4 (2.5)	15.4 (1.7)	13.4 (2.7)
predcw	4.69 (0.36)	5.03 (0.32)	4.93 (0.38)
Siblings			
sibs0-4	0.56 (0.77)	0.44 (0.73)	0.51 (0.77)
sibs5-12	1.47 (1.13)	1.55 (1.33)	1.54 (1.25)
sibs13-14	0.38 (0.54)	0.44 (0.55)	0.39 (0.54)
sibs15plus	1.00 (1.23)	1.24 (1.20)	1.20 (1.26)
Household Head			
agehead	44.6 (11.8)	48.5 (11.5)	47.3 (11.6)
malehead	0.90 (0.29)	0.87 (0.33)	0.89 (0.31)
yrseduchead	3.29 (3.02)	2.11 (2.25)	2.32 (2.49)
workhead	0.92 (0.27)	0.88 (0.31)	0.89 (0.31)
agrichead	0.63 (0.48)	0.69 (0.46)	0.64 (0.47)
ownland	0.53 (0.61)	0.59 (0.61)	0.49 (0.63)
agricpost	0.40 (0.49)	0.46 (0.49)	0.36 (0.48)
Other Household			
moth	0.92 (0.26)	0.91 (0.27)	0.91 (0.27)
fath	0.85 (0.35)	0.83 (0.37)	0.84 (0.36)
mothfath	0.83 (0.36)	0.80 (0.39)	0.81 (0.38)
ownfully	0.90 (0.28)	0.92 (0.26)	0.90 (0.29)
numadults	2.69 (1.19)	2.94 (1.30)	2.84 (1.26)

Table 3.7 contd.

	Males 8-17		
	S	**W**	**O**
Locality			
pragricwork	0.71 (0.24)	0.75 (0.22)	0.71 (0.23)
prfemwork	0.18 (0.09)	0.19 (0.09)	0.17 (0.09)
fertility	0.61 (0.18)	0.64 (0.18)	0.61 (0.18)
prownland	0.48 (0.23)	0.49 (0.22)	0.45 (0.22)
prindhead	0.26 (0.38)	0.25 (0.38)	0.21 (0.36)
prlit15-34	0.89 (0.09)	0.86 (0.11)	0.86 (0.12)
meanv	6.26 (0.84)	6.17 (0.82)	6.35 (0.84)
vcv	0.36 (0.06)	0.37 (0.06)	0.36 (0.05)

S=full-/part-time school, W=full-time work, O=other.

Table 3.8: Marginal Effects, Female and Male 8-12 Year Olds

	Females	**Males**
	School	**School**
Sibs0-4	0.0212 (0.0072)*	0.0368 (0.0060)*
Fsibs5-12	0.0268 (0.0063)*	0.0301 (0.0079)*
Fsibs13-14	-0.0114 (0.0151)	0.0029 (0.0155)
Fsibs15plus	0.0150 (0.0078)*	0.0284 (0.0087)*
Msibs5-12	0.0238 (0.0073)*	0.0139 (0.0084)
Msibs13-14	-0.0286 (0.0153)*	-0.0010 (0.0146)
Msibs15plus	0.0022 (0.0064)	0.0211 (0.0083)*
Wage	-0.0048 (0.0068)	0.0097 (0.0073)
Wage*sibs0-4	-0.0078 (0.0016)*	-0.0108 (0.0013)*
Wage*Fsibs5-12	-0.0065 (0.0014)*	-0.0076 (0.0017)*
Wage*Fsibs13-14	0.0038 (0.0034)	0.0006 (0.0034)
Wage*Fsibs15plus	-0.0035 (0.0018)*	-0.0060 (0.0019)*
Wage*Msibs5-12	-0.0058 (0.0017)*	-0.0044 (0.0018)*
Wage*Msibs13-14	0.0070 (0.0035)*	-0.0001 (0.0032)
Wage*Msibs15plus	-0.0028 (0.0014)*	-0.0066 (0.0018)*
N	399,680	413,134
log L	-129,954	-140,508

Also control for variables listed in table 3.5. * denotes significance at 5% level or less.

Table 3.9: Participation in Activities: Female and Male 8-12 Year Olds

FEMALES		All								
# siblings 0-4			**0**	**1**	**2**	**3**				
Wage	25^{th}	S	0.89	0.88	0.87	0.86				
	75^{th}	S	0.89	0.88	0.86	0.84				
			Females				Males			
# siblings 5-12			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89
	75^{th}	S	0.88	0.88	0.89	0.89	0.88	0.88	0.88	0.88
# siblings 13-14			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.88	0.89	0.89	0.90	0.89	0.88	0.88	0.88
	75^{th}	S	0.88	0.89	0.89	0.90	0.88	0.89	0.89	0.89
# siblings 15plus			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.89	0.89	0.89	0.89	0.89	0.89	0.87	0.86
	75^{th}	S	0.88	0.88	0.88	0.88	0.89	0.88	0.87	0.85
MALES		All								
# siblings 0-4			**0**	**1**	**2**	**3**				
Wage	25^{th}	S	0.88	0.87	0.86	0.85				
	75^{th}	S	0.89	0.87	0.86	0.84				
			Females				Males			
# siblings 5-12			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.88	0.88	0.87	0.87	0.88	0.87	0.87	0.86
	75^{th}	S	0.88	0.88	0.87	0.87	0.88	0.88	0.87	0.86
# siblings 13-14			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.87	0.88	0.89	0.89	0.88	0.87	0.87	0.87
	75^{th}	S	0.88	0.88	0.89	0.90	0.88	0.88	0.88	0.88
# siblings 15plus			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.87	0.88	0.88	0.88	0.88	0.87	0.87	0.86
	75^{th}	S	0.88	0.88	0.88	0.88	0.89	0.87	0.86	0.85

S=full-/part-time school.

Table 3.10: Marginal Effects, Female 13-14 Year Olds

	School	**Full-time Work**	**Other**
Sibs0-4	0.0860 (0.0112)*	-0.0175 (0.0065)*	-0.0685 (0.0104)*
Fsibs5-12	0.1488 (0.0134)*	-0.0075 (0.0038)*	-0.1413 (0.0140)*
Fsibs13-14	0.5546 (0.0707)*	-0.1049 (0.0356)*	-0.4496 (0.0658)*
Fsibs15plus	0.1819 (0.0176)*	-0.0146 (0.0062)*	-0.1673 (0.0177)*
Msibs5-12	0.2110 (0.0189)*	-0.0382 (0.0128)*	-0.1727 (0.0185)*
Msibs13-14	0.2628 (0.0368)*	-0.0503 (0.0179)*	-0.2124 (0.0340)*
Msibs15plus	0.0497 (0.0109)*	0.0015 (0.0042)	-0.0513 (0.0091)*
Wage	0.0446 (0.0096)*	0.0026 (0.0030)	-0.0473 (0.0078)*
Wage*Fsibs0-4	-0.0315 (0.0031)*	0.0065 (0.0023)*	0.0249 (0.0028)*
Wage*Fsibs5-12	-0.0358 (0.0033)*	0.0035 (0.0014)*	0.0323 (0.0033)*
Wage*Fsibs13-14	-0.1138 (0.0164)*	0.0238 (0.0082)*	0.0900 (0.0151)*
Wage*Fsibs15plus	-0.0380 (0.0038)*	0.0046 (0.0018)*	0.0334 (0.0037)*
Wage*Msibs5-12	-0.0491 (0.0048)*	0.0091 (0.0032)*	0.0399 (0.0045)*
Wage*Msibs13-14	-0.0530 (0.0082)*	0.0102 (0.0037)*	0.0428 (0.0075)*
Wage*Msibs15plus	-0.0193 (0.0027)*	0.0006 (0.0009)	0.0187 (0.0026)*
N		152,211	
log L		-108,491	

Also control for variables listed in table 3.5. Standard errors in parentheses. * denotes significance at 5% level or less.

Table 3.11: Participation in Activities: Female 13-14 Year Olds

			All							
# siblings 0-4			**0**	**1**	**2**	**3**				
Wage	25^{th}	S	0.65	0.61	0.57	0.54				
		W	0.04	0.05	0.06	0.06				
		O	0.31	0.34	0.37	0.40				
	75^{th}	S	0.66	0.61	0.57	0.53				
		W	0.04	0.06	0.06	0.07				
		O	0.30	0.33	0.37	0.40				
			Females				Males			
# siblings 5-12			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.64	0.63	0.62	0.61	0.64	0.63	0.62	0.61
		W	0.04	0.05	0.05	0.06	0.04	0.05	0.05	0.05
		O	0.32	0.32	0.33	0.33	0.32	0.32	0.33	0.34
	75^{th}	S	0.65	0.63	0.61	0.60	0.65	0.63	0.61	0.59
		W	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.06
		O	0.30	0.32	0.33	0.34	0.30	0.32	0.34	0.35
# siblings 13-14			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.63	0.65	0.67	0.69	0.63	0.64	0.65	0.67
		W	0.05	0.05	0.06	0.06	0.05	0.04	0.04	0.04
		O	0.32	0.30	0.27	0.25	0.32	0.32	0.31	0.29
	75^{th}	S	0.64	0.63	0.62	0.61	0.64	0.64	0.64	0.64
		W	0.05	0.06	0.07	0.09	0.05	0.05	0.05	0.05
		O	0.31	0.31	0.31	0.30	0.31	0.31	0.31	0.31
# siblings 15plus			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.63	0.63	0.64	0.65	0.65	0.62	0.60	0.57
		W	0.04	0.05	0.05	0.06	0.05	0.05	0.05	0.05
		O	0.33	0.32	0.31	0.29	0.30	0.33	0.35	0.38
	75^{th}	S	0.64	0.64	0.63	0.63	0.66	0.63	0.60	0.57
		W	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05
		O	0.31	0.31	0.31	0.31	0.29	0.32	0.35	0.38

S=full-/part-time school, W=full-time work, O=other/housework.

Table 3.12: Marginal Effects, Male 13-14 Year Olds

	School	Full-time Work	Other
Sibs0-4	0.0879 (0.0217)*	-0.0585 (0.0176)*	-0.0293 (0.0114)*
Fsibs5-12	0.0735 (0.0167)*	-0.0412 (0.0131)*	-0.0323 (0.0096)*
Fsibs13-14	0.1413 (0.0621)*	-0.0955 (0.0375)*	-0.0457 (0.0340)
Fsibs15plus	0.1224 (0.0231)*	-0.0756 (0.0204)*	-0.0468 (0.0118)*
Msibs5-12	0.0891 (0.0212)*	-0.0395 (0.0142)*	-0.0496 (0.0134)*
Msibs13-14	0.0890 (0.0647)	-0.0561 (0.0439)	-0.0324 (0.0341)
Msibs15plus	0.0326 (0.0162)*	-0.0514 (0.0163)*	0.0188 (0.0085)*
Wage	-0.1009 (0.0206)*	0.0279 (0.0091)*	0.0730 (0.0186)*
Wage*Sibs0-4	-0.0277 (0.0056)*	0.0173 (0.0048)*	0.0103 (0.0031)*
Wage*Fsibs5-12	-0.0191 (0.0044)*	0.0110 (0.0038)*	0.0081 (0.0024)*
Wage*Fsibs13-14	-0.0332 (0.0126)*	0.0209 (0.0075)*	0.0123 (0.0070)
Wage*Fsibs15plus	-0.0238 (0.0048)*	0.0127 (0.0041)*	0.0110 (0.0029)*
Wage*Msibs5-12	-0.0221 (0.0050)*	0.0108 (0.0036)*	0.0112 (0.0032)*
Wage*Msibs13-14	-0.0217 (0.0139)	0.0151 (0.0094)	0.0066 (0.0075)
Wage*Msibs15plus	-0.0116 (0.0039)*	0.0143 (0.0043)*	-0.0026 (0.0017)
N		156,433	
log L		-108,407	

Also control for variables listed in table 3.5. Standard errors in parentheses. * denotes significance at 5% level or less.

Table 3.13: Participation in Activities: Male 13-14 Year Olds

			All							
# siblings 0-4			**0**	**1**	**2**	**3**				
Wage	25^{th}	S	0.76	0.73	0.70	0.67				
		W	0.11	0.12	0.14	0.15				
		O	0.13	0.15	0.16	0.18				
	75^{th}	S	0.73	0.69	0.65	0.62				
		W	0.11	0.13	0.15	0.16				
		O	0.16	0.18	0.20	0.22				
			Females				Males			
# siblings 5-12			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.75	0.74	0.73	0.72	0.75	0.74	0.73	0.72
		W	0.11	0.12	0.12	0.13	0.11	0.12	0.13	0.13
		O	0.14	0.14	0.15	0.15	0.14	0.14	0.14	0.15
	75^{th}	S	0.72	0.71	0.69	0.67	0.73	0.71	0.69	0.68
		W	0.12	0.12	0.13	0.14	0.11	0.12	0.13	0.14
		O	0.16	0.17	0.18	0.19	0.16	0.17	0.18	0.18
# siblings 13-14			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.74	0.73	0.72	0.70	0.74	0.74	0.73	0.72
		W	0.12	0.12	0.12	0.12	0.12	0.12	0.13	0.15
		O	0.14	0.15	0.16	0.18	0.14	0.14	0.14	0.13
	75^{th}	S	0.72	0.69	0.67	0.65	0.71	0.70	0.69	0.67
		W	0.12	0.13	0.13	0.14	0.12	0.13	0.15	0.17
		O	0.16	0.18	0.20	0.21	0.17	0.17	0.16	0.16
# siblings 15plus			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.74	0.75	0.75	0.75	0.75	0.74	0.73	0.71
		W	0.12	0.11	0.10	0.10	0.11	0.12	0.13	0.14
		O	0.14	0.14	0.15	0.15	0.14	0.14	0.14	0.15
	75^{th}	S	0.71	0.71	0.71	0.70	0.73	0.71	0.69	0.68
		W	0.13	0.12	0.11	0.10	0.11	0.12	0.14	0.15
		O	0.16	0.17	0.18	0.20	0.16	0.17	0.17	0.17

S=full-/part-time school, W=full-time work, O=other/housework.

Table 3.14: Marginal Effects, Female 15-17 Year Olds

	School	**Full-time Work**	**Other**
Sibs0-4	0.0946 (0.0579)	-0.0586 (0.0223)*	-0.0360 (0.0233)
Fsibs5-12	0.2188 (0.0509)*	-0.0563 (0.0291)	-0.1625 (0.0482)*
Fsibs13-14	0.2525 (0.0938)*	-0.1012 (0.0273)*	-0.1512 (0.0391)*
Fsibs15plus	0.3249 (0.0663)*	-0.0533 (0.0504)	-0.2715 (0.0751)*
Msibs5-12	0.2660 (0.0562)*	-0.0920 (0.0320)*	-0.1740 (0.0500)*
Msibs13-14	0.2694 (0.0848)*	-0.0162 (0.0319)	-0.2531 (0.0476)*
Msibs15plus	0.0101 (0.0268)	-0.0153 (0.0412)	0.0051 (0.0653)
Wage	0.0661 (0.0204)*	-0.0593 (0.0116)*	-0.0068 (0.0182)
Wage*sibs0-4	-0.0311 (0.0108)*	0.0177 (0.0070)*	0.0133 (0.0095)*
Wage*Fsibs5-12	-0.0493 (0.0081)*	0.0165 (0.0065)*	0.0327 (0.0076)*
Wage*Fsibs13-14	-0.0561 (0.0176)*	0.0291 (0.0110)*	0.0269 (0.0155)
Wage*Fsibs15plus	-0.0651 (0.0114)*	0.0170 (0.0073)*	0.0481 (0.0100)*
Wage*Msibs5-12	-0.0585 (0.0092)*	0.0224 (0.0074)*	0.0360 (0.0090)*
Wage*Msibs13-14	-0.0568 (0.0166)*	0.0057 (0.0085)	0.0511 (0.0138)*
Wage*Msibs15plus	-0.0089 (0.0055)	0.0049 (0.0043)	0.0040 (0.0050)
N		181,409	
log L		-157,201	

Also control for variables listed in table 3.5. Standard errors in parentheses. * denotes significance at 5% level or less.

Table 3.15: Participation in Activities: Female 15-17 Year Olds

			All							
# siblings 0-4			**0**	**1**	**2**	**3**				
Wage	25^{th}	S	0.33	0.30	0.27	0.24				
		W	0.16	0.17	0.19	0.20				
		O	0.51	0.53	0.54	0.56				
	75^{th}	S	0.35	0.31	0.27	0.24				
		W	0.15	0.16	0.18	0.20				
		O	0.50	0.53	0.55	0.56				
			Females				Males			
# siblings 5-12			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.33	0.32	0.31	0.30	0.33	0.32	0.31	0.30
		W	0.16	0.17	0.18	0.20	0.16	0.17	0.18	0.19
		O	0.51	0.51	0.51	0.50	0.51	0.51	0.51	0.51
	75^{th}	S	0.35	0.32	0.30	0.28	0.34	0.32	0.30	0.28
		W	0.14	0.16	0.18	0.19	0.15	0.16	0.17	0.18
		O	0.51	0.52	0.52	0.53	0.51	0.52	0.53	0.54
# siblings 13-14			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.32	0.32	0.31	0.30	0.32	0.32	0.32	0.32
		W	0.16	0.18	0.21	0.24	0.17	0.17	0.18	0.18
		O	0.52	0.50	0.48	0.46	0.51	0.51	0.50	0.50
	75^{th}	S	0.34	0.32	0.30	0.28	0.34	0.32	0.30	0.29
		W	0.15	0.17	0.21	0.24	0.15	0.16	0.17	0.17
		O	0.51	0.51	0.49	0.48	0.51	0.52	0.53	0.54
# siblings 15plus			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.32	0.33	0.34	0.35	0.33	0.31	0.29	0.28
		W	0.15	0.17	0.19	0.21	0.17	0.17	0.17	0.17
		O	0.53	0.50	0.47	0.44	0.50	0.52	0.54	0.55
	75^{th}	S	0.33	0.33	0.32	0.31	0.35	0.33	0.30	0.28
		W	0.14	0.16	0.19	0.21	0.15	0.15	0.16	0.16
		O	0.53	0.51	0.49	0.48	0.50	0.52	0.54	0.56

S=full-/part-time school, W=full-time work, O=other/housework.

Table 3.16: Marginal Effects, Male 15-17 Year Olds

	School	**Full-time Work**	**Other**
Sibs0-4	0.1014 (0.0301)*	-0.0967 (0.0305)*	-0.0046 (0.0123)
Fsibs5-12	0.2036 (0.0260)*	-0.1725 (0.0227)*	-0.0311 (0.0106)*
Fsibs13-14	0.2175 (0.0381)*	-0.0899 (0.0379)*	-0.1276 (0.0202)*
Fsibs15plus	0.1658 (0.0324)*	-0.1577 (0.0273)*	-0.0080 (0.0120)
Msibs5-12	0.2616 (0.0237)*	-0.2017 (0.0226)*	-0.0599 (0.0127)*
Msibs13-14	0.2659 (0.0340)*	-0.1459 (0.0347)*	-0.1199 (0.0189)*
Msibs15plus	0.1583 (0.0213)*	-0.0919 (0.0248)*	-0.0663 (0.0114)*
Wage	-0.0064 (0.0098)	0.0257 (0.0108)*	-0.0193 (0.0036)*
Wage*sibs0-4	-0.0303 (0.0057)*	0.0290 (0.0059)*	0.0012 (0.0026)*
Wage*Fsibs5-12	-0.0448 (0.0042)*	0.0395 (0.0042)*	0.0052 (0.0023)*
Wage*Fsibs13-14	-0.0488 (0.0072)*	0.0237 (0.0069)*	0.0250 (0.0042)*
Wage*Fsibs15plus	-0.0339 (0.0057)*	0.0286 (0.0053)*	0.0053 (0.0024)*
Wage*Msibs5-12	-0.0553 (0.0041)*	0.0444 (0.0042)*	0.0109 (0.0027)*
Wage*Msibs13-14	-0.0556 (0.0065)*	0.0330 (0.0068)*	0.0225 (0.0040)*
Wage*Msibs15plus	-0.0375 (0.0040)*	0.0264 (0.0049)*	0.0111 (0.0026)*
N		204,939	
log L		-181,280	

Also control for variables listed in table 3.5. Standard errors in parentheses. * denotes significance at 5% level or less.

Table 3.17: Participation in Activities: Male 15-17 Year Olds

			All							
# siblings 0-4			**0**	**1**	**2**	**3**				
Wage	25^{th}	S	0.38	0.35	0.33	0.30				
		W	0.43	0.45	0.47	0.49				
		O	0.19	0.20	0.20	0.21				
	75^{th}	S	0.38	0.35	0.31	0.28				
		W	0.43	0.46	0.49	0.51				
		O	0.19	0.19	0.20	0.21				
			Females				Males			
# siblings 5-12			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.38	0.37	0.36	0.36	0.38	0.38	0.37	0.37
		W	0.43	0.44	0.45	0.46	0.43	0.43	0.44	0.45
		O	0.19	0.19	0.19	0.18	0.19	0.19	0.19	0.18
	75^{th}	S	0.38	0.36	0.34	0.33	0.38	0.36	0.34	0.33
		W	0.43	0.45	0.47	0.48	0.43	0.45	0.47	0.48
		O	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
# siblings 13-14			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.38	0.37	0.35	0.34	0.38	0.37	0.37	0.37
		W	0.43	0.44	0.46	0.47	0.43	0.44	0.45	0.46
		O	0.19	0.19	0.19	0.19	0.19	0.19	0.18	0.17
	75^{th}	S	0.38	0.35	0.33	0.30	0.37	0.36	0.34	0.33
		W	0.44	0.45	0.47	0.49	0.44	0.45	0.47	0.48
		O	0.18	0.20	0.20	0.21	0.19	0.19	0.19	0.19
# siblings 15plus			**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Wage	25^{th}	S	0.38	0.37	0.37	0.36	0.38	0.37	0.36	0.35
		W	0.44	0.43	0.42	0.41	0.42	0.44	0.46	0.48
		O	0.18	0.20	0.21	0.23	0.20	0.19	0.18	0.17
	75^{th}	S	0.38	0.36	0.35	0.33	0.38	0.36	0.34	0.32
		W	0.44	0.44	0.43	0.43	0.43	0.45	0.47	0.50
		O	0.18	0.20	0.22	0.24	0.19	0.19	0.19	0.18

S=full-/part-time school, W=full-time work, O=other/housework.

Chapter 4

The Impact of a Conditional Education Subsidy on the Activities of Young Adults in the UK

4.1 Introduction

The sharp decrease in school enrolment at the end of primary education observed in Mexico in the previous chapter, is unique to LDCs. In most developed economies, schooling is mandatory until age 16. Upon completion of compulsory education however, there is a noticeable decline in participation in education, that is disproportionate across income groups. In 2002, there was an 82% enrolment rate in England and Wales amongst 16 year olds with a parent in a managerial/professional occupation, compared to 59% amongst individuals with an unskilled manual parent.[1] By age 18, the corresponding participation rates were 55% and 26%. Further, individuals from the top three social classes were almost three times as likely to enter higher education compared to those from the bottom three, with participation rates of 49% and 18% respectively. A variety of factors contribute

[1]The overall average participation rate was 71%: a 75% enrolment rate for females and a 66% enrolment rate for males.

to this observed differential investment in education across socio-economic groups. In the literature, they are broadly categorised into 'short-term' and 'long-term' factors. Short-term factors encompass those that are important for education choices relatively late on in the education cycle, such as income at the time of making college decisions, whilst long-term factors are those that are relevant from the earliest stages and throughout the entire educational development of individuals and that contribute to the development of cognitive and non-cognitive skills and motivation. Most policies aimed at increasing participation in education focus on improving short-term factors, insofar as they are relatively easier to target.[2] The Education Maintenance Allowance (EMA) in the UK is one such example, being a conditional education subsidy, equivalent to a decrease in the cost of schooling for those who choose to remain, that was targeted at low-income individuals upon completion of post-compulsory education. In this chapter we examine its impact on post-compulsory choices between education, work and other activities. Whilst the theoretical effect of the EMA on enrolment is positive (assuming education is a normal good), it represents just one possible intervention amongst a host of others and indeed we cannot say anything about its relative merits vis-á-vis other policies. However, it is useful to consider the potential magnitude of the effectiveness of the subsidy, or another way of thinking about it, the relative importance of short-term factors for education choices. This may be at least partially ascertained from the empirical evidence as to the differential importance of long- and short-term factors in education choices, a brief overview of which we now turn to.

The short-term explanation for differential investment in education across income groups points to low-income families having fewer resources to invest in education, which is relevant in the presence of binding borrowing constraints. Proponents of this explanation point to (at least) two pieces of indirect evidence of credit constraints, which are by no means uncon-

[2]Early intervention in the form of policies to foster cognitive and non-cognitive development (ability) is costly, not least of all because of the difficulty in pinpointing the precise mechanisms that underlie such development and that must be targeted. See Carneiro and Heckman (2003).

troversial.[3] The first is a relatively large response to conditional education subsidies for low-income individuals. However, as conditional subsidies decrease the relative costs facing individuals, one would expect the privately optimal level of schooling to increase even in the absence of liquidity constraints. In addition, the precise magnitude of a 'large' response is unclear. The second piece of evidence is a high income elasticity of education, which is however only valid if one controls for *all* long run factors that are both correlated with income and that affect education choices.[4]

However, social class differences in the education system are engendered from the very earliest years.[5] Education decisions from age 16 onwards are closely driven by prior choices and investments into human capital. Throughout the education cycle, parents make choices regarding school quality and tuition, along with home investments that enhance the transmission of education values to children and that affect the effort and time that children devote to study. Such factors shape the cognitive and non-cognitive development of individuals and therefore affect the stock and quality of human capital. Liquidity constraints that are operative throughout the education cycle may affect the affordability of inputs such as, for example, the choice of school and extra tuition. A recent array of papers, mostly US-based, examines the importance of long-term factors for later education choices vis-á-vis short-term credit constraints. Cameron and Heckman (1998, 2001) find that most of the effect of family income on schooling is due to long-run family background effects that produce ability rather than short-term liquidity constraints. The importance of family income and other family background factors for education choices from age 16 onwards, greatly diminishes after controlling for variables that are correlated with wealth and that affect education choices, such as ability. Keane and Wolpin (2001) esti-

[3]It is difficult to find direct evidence as to the importance of liquidity constraints for education choices. An ideal experiment would involve randomly and unconditionally allocating money to some individuals, from different backgrounds, at the time of making their post-compulsory education decisions. Another useful experiment would involve giving money to families at different points in the life-cycle.

[4]See Cameron and Heckman (1998).

[5]See the White Paper (2003).

mate a structural model of school, work and savings decisions of young males in the US and find that a substantial relaxation of borrowing constraints has almost no effect on college attendance decisions of youth from low-income families and stress the importance of targeting government policies instead on the factors that generate unequal outcomes much earlier in the life cycle. Kane (1994) finds that most of the increase in the high school graduation rate of blacks in the US in recent decades is due to an increase in the average education of their parents. A number of recent papers use instrumental variables to exploit exogenous variation in income.[6] Acemoglu and Pischke (2000) use changes in income inequality in the US over time to instrument parental income and find a positive and significant effect of family income on college enrolment decisions. This finding is supported by Blanden et al (2002), who instrument income using changes in the UK tax system and find a significantly positive effect of instrumented income on the post-16 decision to stay on in education, even after controlling for unobserved heterogeneity linked to early childhood and family background.

Therefore, evidence as to the relative importance of long-term and short-term factors for education is mixed, translating into analogous ambiguity about the appropriate policy response. The effectiveness of tuition grants is also debated. Keane and Wolpin (2001) find evidence of large effects of parental transfers on college attendance, whilst Kane (1994) finds no significant effect of tuition grants on college enrolment rates of blacks, thus casting doubt on the effectiveness of means-tested financial aid on enrolment. The general consensus is that there is a role for both long- and short-term interventions, but that they must be properly targeted. Carneiro and Heckman (2003) point out that whilst there is scope for intervention to alleviate short-term liquidity constraints, this will only go a small way towards narrowing

[6]This is because estimates of the income elasticity of education may be downward biased due to measurement error and transitory movements in income measured at a point in time, thus attenuating the effect of income on education. The attenuation bias will be exacerbated if other variables correlated with permanent income, such as parental education, are included as controls, thus possibly erroneously understating the estimated effect of income on education. See Haveman and Wolfe (1995).

the enrolment gap across income groups. At the same time, early interventions are more costly and more difficult to target, but are likely to have larger long-term impacts.

In this chapter we estimate a dynamic discrete-choice model of participation in education, work and other activities in the UK and examine the effect of a conditional education subsidy, the Education Maintenance Allowance (EMA), on these choices. We follow individuals for three years after compulsory education, observing their activity at a number of different points in time. We allow for a persistent unobserved component to individual decision-making and a key contribution of the analysis is to account for attrition from the sample through time that is possibly non-random in an unobserved dimension. We also control for a large range of observed long-term family background factors, the importance of which was discussed above, so as not to confound the effectiveness of the subsidy with such characteristics. Simulations from the model indicate that the EMA is effective in increasing participation in post-compulsory education and that there is scope for increasing the generosity of the subsidy as a means of further increasing participation. The positive effects are specifically on participation in education without a job, with negative, but smaller, effects on enrolment in education with a job. We also find that accounting for attrition does not make much difference to the structural parameter estimates.

The chapter proceeds as follows. Section 4.2 discusses the EMA in some more detail and the outcome variables of interest. In section 4.3 we outline the model used in the estimation. Section 4.4 describes the data used in the analysis. Sections 4.5 and 4.6 discuss the results and simulations respectively, whilst section 4.7 concludes.

4.2 The Education Maintenance Allowance

The EMA programme was introduced in ten Local Education Authorities (LEAs) in England in September 1999, providing a means-tested benefit to 16-18 year olds remaining in education after the end of compulsory schooling (year 11).[7] It was not randomly allocated to LEAs, but was implemented fully in certain areas. The subsidy, which consists of weekly payments, a retention bonus each term and an achievement bonus at the end of the term, may be claimed for up to two years (three years for individuals with special needs). The ongoing pilot scheme pertains to two cohorts - those who completed compulsory education in 1999 or 2000. Individuals are eligible for the full amount of the EMA if their joint parental taxable income is below £13,000. This is tapered away linearly such that individuals with joint parental income of £30,000 are eligible for the minimum amount of £5 per week and individuals whose parents earn more than £30,000 are not eligible for any award. The pilot scheme consists of four variants, differing in relation to the amounts paid to the student and to whom it is paid (the student or the parent). Details are shown in table 4.1.[8] We also have data on a number of carefully selected control areas which did not receive any subsidy. In this chapter, we use data on all pilot and control individuals who are fully or partially eligible for the EMA and who finished compulsory education at the end of May 1999 (cohort 1), thus conditioning the empirical analysis on low-income individuals.[9]

We observe and model the main activities of individuals within the pilot and control LEAs at seven discrete points in time, $t = 1, \cdots, 7$, for three years immediately after the completion of compulsory education. Each period corresponds to an academic term: we observe an individual's main activity in three terms in each of waves 1 and 2 and in the first term of wave 3. In waves 1 and 2 the individual's choice set includes full-time education

[7]LEAs, of which there are 154 in England, provide certain planning and support functions which are essential to guarantee adequate school provision within a particular area; they also support school improvement especially by helping schools which are underachieving, and with regard to special educational needs and school transport.

[8]We do not directly examine the effectiveness of the achievement and retention bonuses in this study. For an analysis of this see Dearden et al (2003).

[9]Previous evaluation of this programme, using propensity score matching to compare changes in enrolment rates in post-compulsory education amongst eligible pilots and controls after the introduction of the programme, has identified a disproportionate increase in post-compulsory education for eligible males, with no statistically significant effect on university participation. Attrition is however not modelled. See Dearden et al (2003).

(separately classified into with and without a job), full-time work, part-time work and unemployment/other. In wave 3 it includes university, other full-time education (both of which are classified into with and without a job), full-time work, part-time work and unemployment/other.[10]

The overall attrition rate from wave 1 to wave 3 for cohort 1 eligibles is 42%, approximately half of which has occurred by wave 2. Table 4.2 provides a summary of the collection and organisation of the data. It should be noted from this that whilst activities in t=2,3 relate to wave 1, they are collected retrospectively in wave 2 and such data are affected by wave 2 attrition; similarly for t=5,6 which relate to activities in wave 2, yet are collected in wave 3.

4.3 Estimation Methodology

The specification of the econometric model that we estimate is tailored to the structure of the data that we have described above. The sequential discrete choice model outlined below is related to that of Cameron and Heckman (1998, 2001), who highlight the importance of controlling for dynamic selection so as not to obtain biased parameter estimates. However, we account for attrition by modelling the joint probability of both the choice made by the individual and of attrition, importantly allowing for attrition to depend on the same unobserved factors that may (partially) generate the individual choices.[11] This is important because attrition that is informative about the choices of individuals, in a way that is possibly unobserved by us, will lead to biased and inconsistent parameter estimates if ignored. To begin with,

[10]The part-time work category includes individuals in part-time education. This was deemed preferable to including them amongst individuals in full-time education with a job, as the group in part-time education is not eligible for any EMA, so this would confound two groups. In any case, the numbers of eligible individuals in part-time education is extremely low. Note also that the other full-time education category in wave 3 mostly comprises individuals who are repeating their A/AS levels.

[11]If one does not model attrition, whilst the inclusion of the incomplete information on attritors in the estimation improves efficiency, it does not reduce any bias. See Hsiao (1986).

we outline the basic model that does not account for attrition. In section 4.3.2, the model is extended to account for possible non-random attrition from the panel.

4.3.1 Individual Choices

The dependent variable is the main activity of individual i at time t, denoted y_{it}. We assume that the individual can choose among K states in each time period, and can move freely across states at each point in time. Therefore each of the variables $(y_{i1}, \cdots, y_{iT})$ may take values in $\{1, \cdots, K\}$. The observed choice is the optimal outcome of an underlying utility maximisation by the individual. Let the latent variable V^*_{ijt} represent the utility to individual i from being in state j at time t. The variation in decisions across individuals is due to observable characteristics, including the previous choice of the individual, as well as unobserved person-specific effects that persist across time, and a random unobserved factor. Assume that V^*_{ijt} is linear and additive in these arguments

$$V^*_{ijt} = \sum_{k=1}^{K} \delta_{kjt}.1\{y_{it-1} = k\} + \beta_{jt} X_{it} + \epsilon_{ijt} \tag{4.1}$$

where y_{it-1} represents the state chosen by the individual in the previous time period[12], X_{it} includes individual, family and area characteristics and ϵ_{ijt} is an unobserved factor affecting choices.[13] The coefficients δ_{kjt} and β_{jt} represent the effect of the individual's lagged activity and other observed characteristics respectively, on the utility derived from the current activity. Explicitly allowing for lagged state dependence in the model, guards against overstating the degree of unobserved heterogeneity. The individual will be observed to choose state j if it provides the maximum utility across all possible choices

$$y_{it} = j.1\{V^*_{ijt} > V^*_{imt}\} \quad \forall \quad m \neq j$$

[12]Implicit in (4.1) is the assumption that the process follows a first-order Markov process i.e. conditional on one's choice in the previous period, choices in preceding periods are not informative about the current choice of the individual.

[13]In the empirical analysis, X_{it} are the values of the observed individual and area-level characteristics at the date of the first interview and are thus time-invariant. We therefore drop the t subscript from hereon.

The unobserved error term is assumed to follow a one-factor structure[14]

$$\epsilon_{ijt} = \alpha_{jt}\eta_i + u_{ijt} \tag{4.2}$$

This allows for one component of the unobserved trait affecting choices, η_i, to be a persistent characteristic of the individual (ability, for example) and to have a different effect across choices and in different time periods, as captured by the factor loading α_{jt}. u_{ijt} captures unobserved, random influences on individual choices. The econometric interpretation of this latent factor specification considers η_i to be an unobserved covariate that affects the outcome and that is common to all states and time periods, and the α's to be the regression coefficients, that vary across states and time periods.

Inserting (4.2) into (4.1), we obtain

$$V^*_{ijt} = \sum_{k=1}^{K} \delta_{kjt}.1\{y_{it-1} = k\} + \beta_{jt}X_i + \alpha_{jt}\eta_i + u_{ijt} \tag{4.3}$$

From (4.3), there are three possible sources of persistence (apart from observable characteristics X_i) that affect V^*_{ijt} and therefore the observed choice of the individual in each time period: true state dependence as captured by the lagged dependent variable y_{it-1}, persistent person-specific (scalar) unobserved heterogeneity η_i and possible serial correlation in the error term u_{ijt}. Clearly, they all have very different policy implications, making it important to distinguish between them.[15]

We make the following assumptions on the error term

- **Assumption 1** The random variable η_i is independent of u_{ijt} $\forall$ i, j, t, and all η_i and u_{ijt} are independent across individuals.
- **Assumption 2** The term u_{ijt} is an extreme value random variable and is independent of all other $u_{i',j'',t'''}$ except for $i = i', j = j'', t = t'''$.

[14]See Heckman (1981) for a discussion of one-factor models.

[15]See Heckman (1991), especially on the necessity of having access to multi-spell data to distinguish between them.

The contribution to the likelihood function of individual i at time t may now be written as

$$Pr(y_{it} = j | y_{it-1}, X_i, \eta_i) = \frac{exp(\sum_{k \geq 1} \delta_{kjt} 1.(y_{it-1} = k) + \beta_{jt} X_i + \alpha_{jt} \eta_i)}{\sum_{j' \geq 1} exp(\sum_{k \geq 1} \delta_{kj't} 1.(y_{it-1} = k) + \beta_{j't} X_i + \alpha_{j't} \eta_i)} \quad (4.4)$$

for all j and j'.[16]

It is clear from (4.4) that the dependence in individual choices across time, conditional on observables and on the previous choice of the individual, is due to the unobserved η_i. Therefore, piecewise estimation of each period's sub-likelihood functions could lead to inconsistent parameter estimates due to the wrong likelihood being maximised. The problem with estimating (4.4) in a particular time period is immediately clear due to the correlation between y_{it-1} and η_i, as y_{it-1} is itself a function of η_i. In addition, through time, uncorrected heterogeneity in the form of η_i will contaminate parameter estimates through a process of dynamic selection, which is likely to lead to a different correlation structure between observed characteristics and η_i as time goes on. In this non-linear setting, we would ideally like to condition on η_i in order to eliminate person-specific unobservables. However, because η_i is unobserved, and due to the non-linearity of the probabilities, which precludes first-differencing or mean deviations as a way of purging the equation of the individual-specific component, we integrate it out over all possible choices across periods 1 to T, to obtain a sequence of choice probabilities for a particular individual that is conditional on observables only

$$\int_{\eta_i} \prod_{j=1}^{K} [Pr(y_{iT} = j | y_{iT-1}, X_i, \eta_i)]^{1.(y_{iT}=j)} \prod_{j=1}^{K} [Pr(y_{iT-1} = j | y_{iT-2}, X_i, \eta_i)]^{1.(y_{iT-1}=j)}$$
$$\cdots \prod_{j=1}^{K} [Pr(y_{i1} = j | X_i, \eta_i)]^{1.(y_{i1}=j)} dF(\eta_i) \quad (4.5)$$

where we have assumed that at the initial stage of the process, $X_i \perp \eta_i$ (**Assumption 3**), which allows us to integrate with respect to the marginal

[16] See McFadden (1974). As the coefficients in this model are only estimated up to a scale factor, the coefficients for the reference choice, unemployment/other, are set to zero.

and not the conditional distribution of η_i.[17] Note also that as we observe the choices of individuals from when they complete compulsory education, there is no reason for y_{i1} to be generated by the same process as $y_{i2}, \cdots, y_{iT}$. We assume that $y_{i1} \perp y_{i0}$ - the individual's choice in the first observed period is independent of his activity in the period before the sampling begins and therefore the last probability in (4.5) is not conditioned on y_{i0}.[18]

We impose no distributional assumptions on the persistent unobserved factor η_i and instead model it using the non-parametric maximum likelihood estimator (NPMLE).[19] This assumes that whilst endowment heterogeneity is unobserved, we know that there are M types of individual. Therefore $F(\eta_i)$ is modelled as a univariate discrete mixing distribution with a finite number of support points M and unrestricted locations for these support points for η_i: the unknown vectors $(r_1, \cdots, r_M)$ to which the M unknown probabilities $(\pi_1, \cdots, \pi_M)$ are attached. The likelihood contribution for each individual is a finite mixture of the type-specific likelihoods, with the likelihood of each individual's observed choice sequence weighted by their probability of being type m.[20] The parameters α_{jt} on the unobserved hetero-

[17] The fixed effects approach, in which no assumptions are made on the relationship of the individual-specific effect with the explanatory variables, is the other commonly used method of estimating dynamic non-linear models. For a comparison of fixed and random effects estimators see Arellano and Honore (2002).

[18] The initial conditions problem is only relevant when one starts observing individuals when the process is already in progress, in which case the first observation is a function of the dependent variable in periods before the sample starts, and/or if there is serial correlation in the unobserved transitory component (which we assume is not the case - see [Assumption 2] - serial correlation only enters through permanent unobserved heterogeneity). For a treatment of models with initial conditions problems see Heckman (1981) and Wooldridge (2002).

[19] This strategy is due to Laird (1978). Heckman and Singer (1984) discuss the sensitivity of the structural parameter estimates in such models to the functional form of the unobserved heterogeneity.

[20] Monte Carlo evidence suggests that the nonparametric estimation method is quite successful in terms of structural parameter estimates, but estimated mixing distributions may be very inaccurate - see Heckman and Singer (1985). Furthermore, the likelihood can be quite ill-conditioned, leading to the existence of a large number of local maxima. In the final estimation, we generate several starting values to ensure that the optimisation

geneity term in (4.4), are loading parameters that are estimated jointly with $\{r_1, \cdots, r_M, \pi_1, \cdots, \pi_{M-1}\}$.[21] Therefore the sequence of choice probabilities for an individual in (4.5) may be written as

$$\prod_{m=1}^{M} \left\{ \left[\prod_{t=2}^{T} \left(\prod_{j=1}^{K} [Pr(y_{it} = j | y_{it-1}, X_i, \eta_i = r_m)]^{1.(y_{it}=j)} \right) \right] \right.$$
$$\left. \prod_{j=1}^{K} [Pr(y_{i1} = j | X_i, \eta_i = r_m)]^{1.(y_{i1}=j)} Pr(\eta_i = r_m) \right\} \tag{4.6}$$

4.3.2 Attrition

If the relationship between an individual's decision to attrit and the main activity chosen is purely through observables, we can control for this by including the relevant observables in the equation determining the main activity. However, the possibility of sample selection bias arises when one examines a sub-sample of individuals (non-attritors), and when the unobserved factors determining inclusion in the sub-sample (the probability of not attriting) are correlated with both the observables and the unobservables influencing the primary outcome of interest. If this is the case, the model outlined in section 4.3.1 will yield inconsistent estimates of the structural parameters of the model, being based on a sub-sample of individuals without taking account of the possibility that attrition may be informative on the unobserved determinants of individual choices (the η_i in the above exposition). This means that **Assumption 3** ($X_i \perp \eta_i$) is likely to be violated for the sub-sample of non-attritors, and if this is ignored, we will obtain biased and inconsistent estimates of the effects of observables on individual choices.

In order to deal with this, we augment the model presented in section 4.3.1 to model the (unobserved) dependence between individual choices and attrition. This has not been addressed in previous evaluations of this programme. Whilst the primary reason for this is to correct for the effects of

algorithm converges to the same point.

[21] As we leave the mean of η_i unrestricted, the structural parameter vector does not include an intercept. We normalise the factor loading in full-time education with a job in time period one to one, in order to identify the remainder of the factor loadings.

attrition on estimates of the structural parameters of the model, it will also enable us to assess the importance of unobservables for attrition.[22] The dependence between an individual's choice and the decision to attrit or not, is modelled by allowing for those unobserved persistent factors affecting choices, η_i, to also affect the probability of attrition. We require exclusion restrictions to identify the model, i.e. variables that affect the attrition decision but that have no independent effect on the choice of the individual, except through their effect on attrition. In other words, we require variables Z that are independent of the error term in the choice equation, i.e. $Z \perp \eta_i, u_{it} | (X_i, y_{it-1})$.

As discussed already, whilst individual choices are modelled in seven time periods, attrition can only take place in wave 2 or wave 3 and therefore is only modelled at most twice. Further, as we have seen in table 4.2, for individuals who attrit in wave 2, we only observe their activity in one time period, t=1. This is because data on t=2 and t=3 activities in wave 1, are collected retrospectively in wave 2. For individuals who attrit in wave 3, we observe their activities in t=1, $\cdots$,4. Again, the retrospective collection of the data means that information on wave 2 activities in t=5, t=6 is collected in wave 3. For individuals who never attrit on the other hand, we observe their activity in t=1, $\cdots$,7. This is relevant for the notation below.

Formally, let S_{iw} be a binary variable taking the value one if individual i does *not* attrit in wave w, for $w = 2, 3$ (by definition, there is no attrition in wave 1). This is driven by the individual's latent propensity not to attrit, S^*_{iw}, which we assume to be a linear and additive function of the individual's previous activity, observable characteristics of the individual and their environment, as well as unobserved individual traits, which are independent

[22]For readings on non-random attrition from panel surveys see Hausman and Wise (1979), Heckman (1979), Fitzgerald et al (1998) and Ziliak and Kniesner (1998). All of these studies emphasise the fact that the primary goal of accounting for attrition is to correct for the effects of attrition on estimates of the structural parameters of the model, rather than any inherent interest in attrition *per se*.

by assumption

$$S_{iw}^{*}=\varphi_{kw}\sum_{k\geq 1}1.(y_{iw-1}=k)+\gamma_{w}X_{i}+\omega_{w}Z_{i}+v_{iw} \tag{4.7}$$

where y_{iw-1} is the activity of the individual in the previous wave, X_i are time-invariant observed individual, family and area characteristics that affect attrition, and are distinguished from Z_i, which are the time-invariant instruments that are used for identification - we return to a discussion of these in section 4.5. As in section 4.3.1, we assume that the error term may be decomposed into a one-factor structure

$$v_{iw}=\delta_{w}\eta_{i}+\xi_{iw} \tag{4.8}$$

As can be seen from comparing (4.2) and (4.8), the person-specific permanent errors affecting choices and attrition are assumed to be drawn from the *same* mixing distribution, with the loading factors differing across both.

The observed survey participation status of the individual in wave w, S_{iw}, is a function of his propensity not to attrit. In particular,

$$\begin{aligned} S_{iw} &= 1 \quad \text{if} \quad S_{iw}^{*}>0, \quad \forall \quad i, \quad \text{for} \quad w=2,3 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

As in section 4.3.1, we make the following assumptions on the components of the error term

- **Assumption 4** The random variable η_i is independent of ξ_{iw} $\forall$ i, w. All η_i and ξ_{iw} are independent across individuals.
- **Assumption 5** The term ξ_{iw} is an extreme value random variable and is independent of all other $\xi_{i'w''}$ except for $i=i', w=w''$.

Under these assumptions, the conditional probability that the individual does not attrit from the sample in wave w may be written using a logit functional form, in which the dependent variable is 1 for each period someone is in the sample, and 0 when (and if) an individual leaves the sample

$$Pr(S_{iw}=1|y_{iw-1},X_{i},Z_{i},\eta_{i})=\frac{exp(\sum_{k\geq 1}\varphi_{kw}.1(y_{iw-1}=k)+\gamma_{w}X_{i}+\omega_{w}Z_{i}+\delta_{w}\eta_{i})}{1+exp(\sum_{k\geq 1}\varphi_{kw}.1(y_{iw-1}=k)+\gamma_{w}X_{i}+\omega_{w}Z_{i}+\delta_{w}\eta_{i})} \tag{4.9}$$

where as before, η_i is integrated out due to the fact that it is not observed.[23] The δ_w indicate the importance of unobserved heterogeneity for attrition. If they are not statistically different from zero, and under the random effects and one-factor assumptions, there is no selection on unobservables.[24] However, a finding that unobserved traits affect both attrition and choices, underlines the importance of basing inference about the distribution of $y_{it}|X_i, y_{it-1}$ on the joint distribution of $(y_{it}, S_{iw})|X_i, y_{it-1}, Z_i$.

We denote the individual log likelihood contributions by l_i^A, where A is the individual's attrition status which may be any of na (do not attrit), $aw3$ (attrit in wave 3) or $aw2$ (attrit in wave 2). So for example, the likelihood contributions for individuals who choose the same activity, k, in each period in which they remain in the sample, for each of the three groups are (which are detailed in Appendix A)

$$\begin{aligned}
l_i^{na} =& Pr(S_{i3} = 1, S_{i2} = 1, y_{i7} = \cdots = y_{i1} = k, X_i, \eta_i = r_m) \\
l_i^{aw3} =& Pr(S_{i3} = 0, S_{i2} = 1, y_{i4} = \cdots = y_{i1} = k, X_i, \eta_i = r_m) \\
l_i^{aw2} =& Pr(S_{i2} = 0, y_{i1} = k, X_i, \eta_i = r_m)
\end{aligned}$$

The log likelihood function across the whole sample of individuals is the sum of the individual log likelihood contributions of the three groups of individuals

$$L = \sum_{i=1}^{N^{na}} l_i^{na} + \sum_{i=1}^{N^{aw3}} l_i^{aw3} + \sum_{i=1}^{N^{aw2}} l_i^{aw2}$$

4.4 Data and Descriptives

The individual and family background data used in the analysis is mostly based on a face-to-face interview with both the parents and the young adults

[23] Note that for modelling S_{i3}, the young person's activity in the previous wave (w=2) is observed at three different points, y_{i2}, y_{i3}, y_{i4}. We condition on the most recent activity, y_{i4}. For modelling S_{i2}, the young person's activity in the previous wave (w=1) is observed at one point, y_{i1}.

[24] Note that a more general specification would allow the error terms in both equations to come from separate mixture distributions, with dependence between them. See Van der berg and Lindeboom (1998).

at the beginning of the school year when the subsidy first became available. At this point, a wide range of pre-programme information including family composition, parental background and detailed family income data was collected. In subsequent waves, the young person was re-interviewed by telephone and provided information on their main activity, their updated academic achievement, and detailed part-time and full-time work information, where applicable. We also observe a host of characteristics on the areas in which the individuals reside. This is important because although control areas were chosen to be as similar as possible to pilot areas in a wide range of observable dimensions, Dearden et al (2003) provide evidence that pilot areas are relatively more deprived. A list of the variables used in the analysis is presented in table 4.3.

The measure of 'ability' that we control for is the individual's English and maths GCSE test scores (such tests are administered at the end of compulsory schooling). These better capture acquired rather than innate ability, i.e. ability that is itself a function of prior education. The precise amount of the EMA that one is eligible for, is determined exactly by income: the EMA is tapered away linearly as parental income increases. Therefore, although the model is able to predict the impact of the EMA programme without functional form assumptions as to how income affects participation, the impact of changing the EMA depends what we assume about how participation varies with income in the first place. We control for income and its square, and therefore the effect of changing the EMA is identified on the basis of nonlinearities in income.

4.4.1 Main Activity

As discussed in section 4.3, the discrete decision period is each term of the academic year. Tables 4.4 to 4.7 display the proportions in various activities for eligible pilots and controls, both overall and separately by gender, across terms in waves 1 to 3. Noteworthy features are the fact that the proportion in full-time education with a job is higher in control than in pilot areas, throughout each time period in waves 1 and 2. The opposite pattern holds for full-time education without a job, in which the participation of pilots is consistently higher than that of controls. There are more eligible controls than eligible pilots in full-time work and the proportions in part-time work, unemployed or doing something else, are extremely similar across pilots and controls throughout.

Whilst there are notable gender differences in participation rates in education and work, the pilot-control comparison shows similar patterns for both males and females. In particular, participation in education without a job is consistently higher in pilot than in control areas, for both males and females. However, enrolment in this activity is higher for males than for females, in both pilot and control areas. Enrolment in education with a job on the other hand, is generally greater in control than in pilot areas for both genders, and is higher for females than for males. Full-time work participation is higher in control areas for both genders, and is higher for males than for females. The differences across pilots and controls in part-time work and unemployment/other are very small, for both genders.

Differences in wave 3 activities across pilots and controls are much less noticeable. Nonetheless, the proportions in university with a job in pilot areas are slightly higher than in control areas, and participation in full-time work is higher in control areas.

Transitions across states within waves

The advantage of termly data is that it enables us to capture individuals dropping out or interrupting education temporarily, and returning later. Tables 4.8 and 4.9 present transitions across the five states for all eligible pilots and controls, for waves 1 and 2 respectively and give an indication as to the extent to which individuals move across states in each wave. It also shows the proportion of individuals attriting, by their previous activity. Individuals move across states less as the year progresses, i.e. there is more movement across states between September and February, than between February and May. The majority of the returning to education after a temporary absence occurs between September and February of each of waves 1 and 2 (mostly the former wave), with very few individuals returning to education from a non-education activity between February and May. For example of those who did not choose education in September of wave 1 (2,039 individuals), 11.4% of them have returned to education by February of that wave, and 37.4% have attrited (thus their activity is not observed); the corresponding figures in wave 2 are 4.4% and 28.2% respectively.

It is also instructive to examine the young person's activity in September of wave 3, by their lagged activity (May of wave 2). Table 4.10 shows that amongst those in university (with or without a job), just over 2.7% are entering from a non-education activity. The corresponding figure for other full-time education is unsurprisingly higher at 8.7%, as it includes individuals repeating exams. It is worth noting that the relatively high proportion of individuals in part-time work who were in full-time education with a job in the previous period (44.8%), may be explained by the fact that individuals may still have been making their choices by the time they were interviewed, and the data collected in wave 4 may be more informative in this regard.

4.4.2 Exploratory Analysis of Attrition

As a joint model of the outcome and attrition probabilities is computationally complex, a simple exploratory analysis of attrition from the panel is useful, in order to determine the necessity of modelling the attrition process itself. Table 4.11 provides a descriptive overview of the types of individual who attrit in waves 2 or 3, along with those who do not attrit by wave 3, by a range of observable background characteristics. We see that on average, those who drop out of the sample are of lower socio-economic background. However, the table is only informative on the extent to which observable characteristics are associated with attrition. If it is the case that individuals from lower socio-economic backgrounds differ in a number of unobservable dimensions to those from higher socio-economic backgrounds, then this means that attrition cannot be fully accounted for by observable traits only. Further, even though we account for the lagged activity of the individual in the modelling of attrition, which is an important factor in the attrition process (attrition in waves 2 and 3 tends to be higher for individuals who did not stay on in post-compulsory education - perhaps due to individuals moving out of home when they start to work - see tables 4.8 and 4.9), it is likely that the interviewer in wave 1 has an important impact on future attrition, over and above the effects of socio-economic variables and the individual's previous activity. We turn to a discussion of the importance of interviewers and unobservables in the attrition process, in the next section.

4.5 Results

4.5.1 Attrition

As discussed in section 4.3.2, modelling the joint processes of attrition and activities necessitates access to a variable that affects attrition but that can be excluded from the main equation explaining individual choices. The first wave of interviews was conducted face-to-face, with subsequent interviews carried out over the telephone. Data has been collected on the age and experience of interviewers who conducted the interviews in wave 1. We code the age of the interviewer into eight mutually exclusive categories as shown below.

Instruments for Attrition

Variable	Age Group	% of Interviewers	% indivs stay on in wave 2	% indivs stay on in wave 3 given that stayed on in wave 2
age1	≤ 30	2.8	0.67	0.85
age2	$31-40$	9.5	0.73	0.80
age3	$41-45$	11.5	0.75	0.80
age4	$46-50$	16.4	0.75	0.79
age5	$51-55$	18.9	0.75	0.78
age6	$56-60$	22.0	0.75	0.78
age7	$61-65$	12.1	0.70	0.77
age8	66+	6.9	0.71	0.75

Whether one may expect older interviewers to better convey the importance of survey response, due to their relatively higher level of experience, or whether younger interviewers may relate better to the interviewees, it seems intuitively reasonable to expect this variable to have an effect on attrition, without having any independent effect on the individual's choice. The table above points towards positive associations between very young interviewers (below age 30) and relatively older interviewers (above age 61), and attrition in wave 2. In wave 3, attrition seems to be higher, the older the wave 1 interviewer. Table 4.12 shows the parameter estimates for attrition in waves 2 and 3, estimated from the model with correlated unobserved heterogeneity across attrition and individual choices. Interestingly, in the causal analysis, the instruments are more relevant for wave 3 than for wave 2 attrition. In wave 2, the only significant effect is for the youngest category of interviewers ($\leq$ 30), for which there is a positive and significant effect on the probability of attriting in wave 2. However, the low proportion of interviewers in this age range is of some concern. On the other hand, conditional on not attriting in wave 2, older interviewers (aged 46 plus) have a significantly positive effect on the probability of attriting in wave 3.[25] The evidence therefore points to young people being more responsive, in terms of re-interview reliability, to younger interviewers than to older ones. This suggests that lower age gaps between interviewers and young people, rather than the experience of the interviewer, may be more important to young people and may help to reduce attrition.

Observable individual and family characteristics that have positive and significant effects on the probability of attriting in wave 2 include being male, having more older siblings, and having a father in full-time work. Negative predictors of attrition in this wave include having both parents present in the household, having a mother in part-time work, having a mother with O-level qualifications or higher, having a high maths GCSE score, living in a rural area and having stayed on in post-compulsory education in wave 1. On the other hand, living in a rural area is positively associated with wave

[25]Note that after controlling for age of the interviewer, experience has very little effect on attrition.

3 attrition (perhaps suggestive of young people moving away from home to further their education) and individuals who are partially eligible for the EMA and who live in a pilot areas are also more likely to attrit. Negative effects again come from the presence of both parents in the household, in addition to having a father in part-time work, having a high English GCSE score and having been in any activity other than unemployment/other in the previous period.

In terms of the importance of unobserved heterogeneity for attrition, the parameters on the unobserved heterogeneity terms are significant in both the wave 2 and wave 3 attrition equations. We examine the correlations in unobserved preferences for attrition and activities chosen, by comparing the loading factors α_t and δ_w on the unobserved heterogeneity.[26] This shows that unobserved preferences for education with a job or full-time work, are positively correlated with an unobserved tendency to stay on in the sample in wave 2. There is on the other hand, a negative correlation between unobserved preferences for education without a job or part-time work and one's propensity to stay on in the sample in wave 2. For individuals who stay on after wave 2, unobserved preferences for non-university education without a job are positively correlated with an unobserved tendency to stay on in the sample in wave 3, whilst individuals with unobserved preferences for university or other non-university education with a job, or full-time/part-time work, are more likely to attrit in wave 3.

4.5.2 Goodness of Fit

To assess the importance of unobserved heterogeneity and attrition in the data, we compare two models, both of which control for unobserved heterogeneity. In the first, we allow for attrition to depend on the same unobserved factors affecting choices, i.e. $\delta_w \neq 0$ in (4.9). In the second we assume that $\delta_w = 0$. Under the null hypothesis of unobserved heterogeneity having no effect on attrition, parameters from both models are consistent.

[26]The parameter estimates from the model for each time period are shown in tables 4.23 to table 4.26 in Appendix B.

The first criterion that we use to judge which model fits the data better, is the Bayesian information criterion (BIC). This weighs the trade-off between increased information and decreased reliability (due to the number of parameters increasing) and takes the sample size into account. The model is chosen with the largest sample BIC. Comparing rows (1) and (2) in table 4.13 shows that the model allowing for correlated unobserved heterogeneity in attrition is favoured using this criterion.

A comparison of the parameter estimates from the two models is useful in order to see whether controlling for correlated unobserved heterogeneity makes much difference to the parameter estimates. The structural parameter estimates in the models with and without correlated unobserved heterogeneity across the main and attrition equations differ by very little.[27] This is an important finding, given that these covariate effects are the parameters of interest. Thus even though the unobserved heterogeneity terms are significantly dependent, the parameter estimates are not very sensitive to this.

We next compare the model that accounts for an unobserved persistent component in individual decision-making to one that does not.[28] Comparing rows (2) and (3) in table 4.13 shows that the former is preferred over a model that treats all decisions across time for a given individual as independent of each other, conditional on observables. A Wald test of the joint hypothesis that all of the parameters relating to unobserved heterogeneity (including the attrition unobserved heterogeneity parameters) are zero gives a value of 176.09 which is distributed as χ^2 with 26 degrees of freedom. The null hypothesis of no unobserved heterogeneity is clearly rejected (the critical value is $\chi^2_{26} = 38.88$). The two points of support indicate the presence of two types that differ in some unobserved dimension.[29] We return to a more

[27]These are not reported to save space and are available upon request.

[28]As attrition makes very little difference to the structural estimates, it is valid to compare models in which we either control or do not control for attrition. Here the two models that we compare both account for attrition that is uncorrelated with unobserved heterogeneity.

[29]Estimates of the parameters of the model would be inconsistent only if the number of types was mis-estimated. Increasing the number of points of support to three led to both

detailed description of these individual types in section 4.5.3.

As a further assessment of how well the final model that accounts for unobserved heterogeneity and correlated attrition fits the data, its estimates are used to predict the proportions in each activity across pilots and controls in each time period. These predicted proportions are then compared to the actual proportions. Table 4.14 shows the actual and predicted differences across pilots and controls in the proportions in the various activities. The predicted proportions mimic the actual ones very closely, providing additional evidence that the model fits the data well.

4.5.3 Descriptive Analysis of Individual Types

We have so far been discussing the notion of unobserved heterogeneity in somewhat impalpable terms. Here we explore unobserved heterogeneity somewhat further. Within the random effects framework, each point of support and associated probability describes a latent class or type of individual, which may represent some common preferences or ability, for example. Table 4.15 shows the percentage point difference in participation in each activity between type 1's and 2's. It indicates that the split of the sample is on the basis of preferences for work, rather than on the basis of preferences for education or non-education. Type 1 individuals, constituting approximately 85% of the sample, have stronger unobserved preferences for education without a job, whilst type 2 individuals are more likely to be observed in either education with a job or full-time work.

We can investigate the properties of unobserved heterogeneity further. Whilst it is not possible to estimate each individual's endowment type, we can use Bayes rule to estimate posterior type probabilities for each individual in the sample conditional on the estimated parameters, the observed choices

the parameters estimates and the value of log likelihood remaining virtually unchanged compared to two points of support. Note also that the parameters in terms 2 and 3 in each of waves 1 and 2 are constrained to be the same in order to make the model more parsimonious. Maximisation is highly time-consuming due to the non-linearity of the function and the large number of parameters to be estimated.

and initial schooling. This allows us to correlate the conditional probabilities with observable factors. This is purely descriptive, but is nonetheless a useful specification check as to how intuitively reasonable the 'types' are.

We estimate the probability of observing a particular sequence of activities for each person, S_i^j, where j may represent any of the sequences listed in table 4.16. Due to the number of possible sequences proliferating rapidly, we aggregate education with and without jobs into 'education', E, and full-time work, part-time work and unemployment/other into 'non-education'.[30] So for example, if j represents the sequence of education in every time period, the probability of observing sequence j for individual i, conditional on observables and on being type r_m, $m = 1, 2$ is

$$Pr(S_i^{eeeee}|X_i, \eta_i = r_1) = Pr(y_{i1} = e, y_{i2} = e, y_{i3} = e, y_{i4} = e, y_{i5} = e|X_i, \eta_i = r_1)$$
$$Pr(S_i^{eeeee}|X_i, \eta_i = r_2) = Pr(y_{i1} = e, y_{i2} = e, y_{i3} = e, y_{i4} = e, y_{i5} = e|X_i, \eta_i = r_2)$$

For each individual, we find his/her most probable sequence of activities. Denoting this by $Pr(S_i^*|X_i, \eta_i = r_m)$, this is the maximum probability across all sequences

$$Pr(S_i^*|X_i, \eta_i = r_1) = max_j\left[Pr(S_i^j|X_i, \eta_i = r_1)\right] \quad \text{across all} \quad j$$
$$Pr(S_i^*|X_i, \eta_i = r_2) = max_j\left[Pr(S_i^j|X_i, \eta_i = r_2)\right] \quad \text{across all} \quad j$$

The posterior probability of being in latent class r_m, for $m = 1, 2$, given sequence S_i^* for individual i is then

$$Pr(\eta_i = r_1|X_i, S_i^*) = \frac{\pi_1.Pr(S_i^*|X_i, \eta_i = r_1)}{\pi_1.Pr(S_i^*|X, \eta_i = r_1) + \pi_2.Pr(S_i^*|X_i, \eta_i = r_2)}$$

$$Pr(\eta_i = r_2|X_i, S_i^*) = \frac{\pi_2.Pr(S_i^*|X_i, \eta_i = r_2)}{\pi_1.Pr(S_i^*|X, \eta_i = r_1) + \pi_2.Pr(S_i^*|X_i, \eta_i = r_2)} \quad (4.10)$$

where $\pi_1 = Pr(\eta_i = r_1) = 0.85, \pi_2 = Pr(\eta_i = r_2) = 1 - \pi_1 = 0.15$ are the estimated proportions of types 1 and 2 respectively.

[30]Note that here and in the simulations below, we aggregate data from time periods 2 and 3 in each of waves 1 and 2.

Table 4.16 shows the predicted proportions in each of the listed sequences of activity in pilot and control areas, separately by type. The proportion of pilots that is always in education is around 4 percentage points higher than the proportion of controls. In terms of differences across types, type 2's are more likely than type 1's to remain in education throughout the period, for both pilots and controls. This is interesting, as we have seen already that type 2's are also more likely to be in full-time education with a job, with type 1's more likely to be in full-time education without a job. A relatively higher proportion of type 1's compared to 2's on the other hand stay in education for the first two waves and then drop out of education in wave 3. Type 2's are also slightly more likely to be observed in non-education across all periods.

To summarise, type 2's are more likely to be observed in education throughout the whole period. Further, they are also more likely to be observed in full-time education with a job, compared to type 1's. This may be due to unobserved preferences for education with a job or higher ability for example (that would render them more likely to stay in education; being of high ability is also consistent with their having more time to work part-time whilst in school). In table 4.17 we show the results from regressing the probability that an individual is type 1, as estimated from equation (4.10), on a set of observable characteristics. Type 1 individuals are the 'lower ability' group, in the sense that there is a negative correlation between the probability of being type 1 and GCSE test scores. Males are more likely to be type 1, as are individuals living in rural areas. It is also interesting that having a mother with O-levels or higher is positively associated with high-ability. Interestingly, there is no significant correlation between living in a pilot area and being type 1. This is reassuring, suggesting as it does that pilot and control areas are similar in terms of unobserved characteristics.[31]

Finally, it is worth noting that the existing empirical evidence on the

[31] Therefore comparisons across pilots and controls, conditional on observables, are likely to provide an unbiased estimate of the effect of the programme. In the programme evaluation terminology, the Conditional Independence Assumption is likely to hold.

links between part-time work and educational attainment is mixed, with some studies finding that part-time work in school adversely affects academic attainment and others finding no significant associations.[32] However, the evidence here suggests that selection bias may affect any observed associations, with high-ability individuals more likely to select into part-time work whilst in school. The links between family background, ability and part-time work in school and the subsequent effects on academic attainment, is an interesting topic for future research.

4.6 Simulations

4.6.1 Post-Compulsory Second-Level Participation

As the parameter estimates from the estimated model have no direct interpretation, we focus on results from the simulations that are carried out using the parameter estimates, which are easily interpretable and can be used to answer policy questions of interest. We simulate the impact of a number of hypothetical policy decisions, in order to compare the current policy package with other alternatives. We set the actual EMA entitlement to zero, decrease it by 50%, increase it by 50% and double it, and estimate the predicted proportions in the activities in each time period. We report unconditional probabilities throughout and all probabilities are calculated at the mean value of all other individual, family and area characteristics.[33] Estimates pertain to the sample of eligible individuals only and are based on the estimated parameters from the dynamic model controlling for unobserved heterogeneity and correlated non-random attrition.

Tables 4.18 to 4.20 show the results from these simulations. Each figure in the table represents the proportion in the listed state in the simulated scenario minus the proportion under the actual scenario. Thus a negative number on an activity in the lower-EMA simulations means that the EMA

[32]See Singh (1998) and Dustman et al (1996a) respectively.

[33]Thus for each period, we estimate the probability that the individual is in a particular state, j, over all previous activities, k: $Pr(y_{it} = j|X) = \sum_k Pr(y_{it} = j|y_{it-1} = k, X).Pr(y_{it-1} = k|X)$.

has a positive effect on participation in those activities, whilst a negative number in the higher-EMA simulations represents activities in which increasing the EMA from its current amount has a negative effect on participation.

In brief, the results show that the EMA has a positive effect on participation in full-time education without a job and a negative effect on full-time education with a job, with an overall positive net effect on education. The effects on full-time work are generally negative, particularly in the first wave, and the impacts on unemployment/other are also negative. In terms of the magnitudes, the largest effects are on participation in the education categories. We discuss these more fully below.

Firstly, setting the EMA to zero for all pilots reduces the participation rate in education without a job by between 5 and 10 percentage points for all terms in waves 1 and 2. At the same time, participation in education with a job significantly increases by between 2.5 and 4 percentage points, and the (significant) increases in full-time work are between 1.5 and 3 percentage points. In the first wave, there is a small negative effect on unemployment/other, and in term 1 of wave 2, the sign of this effect changes, but remains small at 1.1 percentage points. The overall effects of reducing the EMA by 50% are similar, but of lower magnitudes. The effect on education without a job is to reduce participation by between 3 and 5 percentage points, compared to a positive impact on education with a job of between 1.3 and 2.4 percentage points. The overall effect on participation in education is negative.

Increasing the EMA by 50% on the other hand, increases participation in education without a job - by a maximum of around 6 percentage points for both the latter part of wave 1 and September of wave 2, and by 2 to 3 percentage points in the other terms of waves 1 and 2. Again, most of this increase is drawn from education with a job and full-time work. The net effect on education is positive. Doubling the EMA leads to similar inferences: larger significant increases in the proportions enrolled in education without a job, being highest in February/May of wave 1 at 15.5 percentage points, and

September of wave 2 at 14.1 percentage points. The corresponding decreases in participation in education with a job are 4.2 and 4.3 percentage points respectively, and in full-time work, 6.4 and 7.0 percentage points. There are also decreases in unemployment/other, most notably 3.6 percentage points in February/May of wave 1 and 3.5 percentage points in September of wave 2.

Therefore the general finding is that the EMA leads to an increase in participation in full-time education without a job, which is greater than the induced decrease in participation in education with a job. It is not just that the EMA induces individuals to shift from education with a job to education without a job, but it also draws individuals into education from full-time work and other states.

4.6.2 University Participation

As a descriptive tool, comparing enrolment in higher education across pilot and control areas, is informative as to the association between the EMA and higher education. If the EMA has an effect on university participation, and if we are confident that there are no unobservables between pilot and control areas differentially affecting stay-on rates, we would expect enrolment rates to be higher in pilot than in control areas. However, table 4.20, which shows the difference in predicted participation in activities in wave 3 across pilots and controls, provides some evidence that the EMA is not associated with higher university participation (or with different participation in any other activity in fact).

The EMA was designed as a means of encouraging individuals to stay on in post-compulsory education and we have seen that it has indeed been effective in achieving these objectives. Here we use the estimated model to examine the effect on university enrolment, of augmenting the pool of individuals in term 1 of post-compulsory second-level schooling. This is an interesting policy question in and of itself, in essence addressing the issue of the response of university enrolment to an $x\%$ increase in the post-compulsory education enrolment immediately following year 11. The parameter of interest is therefore $\frac{\partial(uni)}{\partial(educ_1)}$, where $educ_1$ refers to participation in education in September 1999. We examine this by increasing the observed participation in full-time education in term 1 of the first wave and simulating the effect on the predicted proportions in university in the final time period. We assume that in order to be in university, the individual must attend education in all time periods. So for example, amongst the group of individuals who are *not* in education in term 1 of wave 1[34], we shift $x\%$ of them into education in this period and predict the 'new' overall number in education in period 2, conditional on having been in education in period 1. We then predict the number in education in period 3 conditional on having been in education in period 2 (under the new scenario), and so on up to the last time period.

We find that participation in university is not very responsive to increases in the pool of term 1, wave 1 school participants. In the table below, we see that shifting 10% of individuals who chose not to remain in education in term 1 of wave 1, into education in this term (i.e. increasing participation by 92 individuals), leads to a 0.007 percentage point increase in observed university participation. Therefore, amongst the 92 extra individuals that we move into education in the first term of wave 1, the model predicts that only 7 of them go on to university in wave 3. This low pattern also holds for higher increases in term 1 participation: if we move 460 individuals out of work/other and into education in September of 1999, only 34 of them continue on to higher education in wave 3. We can conclude from this that the EMA induces only a very modest increase in university participation, compared to its effect on post-compulsory second-level participation, for cohort 1 eligible individuals.

4.6.3 Family Background

In analysing the effect of the EMA, we have controlled throughout for an extensive array of family background and other observable characteristics.

[34]Note that this analysis is carried out on the sample of non-attritors only. Of these, 917 were not in education in term 1 of wave 1.

Simulated University Participation

% increase in term 1 participation amongst non-participants	Number in university	Percentage point increase from baseline
0	949	-
10	956	0.007
20	962	0.014
30	970	0.022
50	983	0.036
80	1,002	0.056

Controlling for as many such factors as possible is important, so as not to spuriously attribute any observed disproportionate increases in education in pilot relative to control areas, to the EMA. In this section, we use a principal components analysis to summarise a variety of family background variables and to examine the influence of such characteristics on participation in various activities. This turns out to be an informative way of constructing an index to summarise a large set of background variables. We simulate the effect of changing the EMA on individuals from two different background types, which we refer to as 'upper' and 'lower', to describe relatively well off and relatively less well off households respectively.

We perform a factor analysis on all of the variables used in the estimation, in order to identify the principal components of the data.[35] The first factor is marked by high loadings on the locality characteristics, whilst the second factor is marked by high loadings on the household characteristics. Insofar as we are interested in the associations between family background characteristics and participation in education and work activities, we use the second factor to classify individuals into two groups, on the basis of whether they are above or below the median factor score (which is the value of individual cases for the factors). The table below shows that individuals

[35]We extract 14 out of 29 factors, which fully capture 90% of the variation in the variables.

in the upper range (i.e. above the median) are from relatively more affluent backgrounds compared to those in the lower range.[36]

Principal Components

	Upper	**Lower**
Background Characteristics		
Number older siblings	0.98	1.03
Number younger siblings	0.87	1.10
Maths GCSE score	4.17	3.15
English GCSE score	4.57	3.65
	Percentage	
Both parents present	0.94	0.15
Mother works	0.71	0.48
Father works	0.84	0.03
Mother has O-level or higher	0.48	0.39
Father has O-level or higher	0.57	0.05
Nonwhite	0.10	0.09
Owns house	0.83	0.43
1 or 2 parents in work when born	0.92	0.69
Low income	0.09	0.53

Table 4.21 shows the predicted differences in participation in each of the five activities, across the above two groups, for waves 1 and 2. In the first column, we compare participation under the current EMA package. In the second and third columns, we simulate changes in participation for the two groups under two new scenarios, in which we set the EMA to zero and double the generosity of the subsidy. This is useful in order to ascertain which household types may be more or less sensitive to changes in the EMA.

[36]It should be noted that Cameron and Heckman (1998) examine the importance of family and individual variables for education outcomes for blacks and whites, by decomposing the schooling gaps across the two groups into the contribution made by *each* explanatory variable, thus allowing one to assess the importance of each variable for education. In this analysis, we instead construct an index to summarise family background characteristics and assess the importance of this index for education and work choices.

Under the actual programme, we see that the two most notable differences in participation across the two groups, are in education with a job and unemployment. Individuals from relatively well off backgrounds participate more in education with a job, with participation approximately 6 percentage points higher on average for this group in each time period. At the same time, such individuals are less likely to be unemployed, but the differences between the groups are lower at between 2 and 5 percentage points. These findings are interesting, as previous literature has suggested that individuals from relatively more affluent backgrounds engage more in part-time work whilst in school, which has been partly attributed to their having parents with better connections to the workplace which may be beneficial in terms of finding employment.[37] It is also notable that individuals from less well off backgrounds tend to participate more in education without a job compared to those from the other group.

We next examine the effects on participation in the two groups when we set the EMA entitlement to zero. Column (2) in the table shows the results from this simulation. The two groups respond in fairly similar ways. Both decrease participation in education without a job in every time period, and participation in education with work increases. However, similar to the simulations in section 4.6.1, the overall effect on education seems to be negative for both groups. Participation in each of the other activities changes only very slightly.

When the EMA generosity is doubled, the net effect on education is positive for both groups.[38] Individuals from upper backgrounds increase participation in education without a job by between 5 and 13 percentage points, which exceeds the decrease in education with a job of around 4 per-

[37] See for example Dustman et al (1996b).

[38] Note that a doubling of the EMA entitlement represents a much larger increase for individuals in the 'lower' category, as they are more likely to be eligible for the full EMA entitlement, whereas individuals in the 'upper' category are more likely to be on the taper. Therefore, we are in some sense intermingling the effects of differential increases in the EMA by group, with family background effects.

centage points. Unemployment decreases from the latter part of wave 1 onwards. For lower-background individuals, doubling their EMA entitlement is even more effective for education, with participation in education without a job increasing substantially by between 7-16 percentage points. Again, from terms 2/3 of wave 1, most of these individuals are drawn from full-time work and unemployment. There a modest decrease in education with a job of between 1 and 4.5 percentage points.

Finally in wave 3, we use the model to predict participation in each of the seven activities, including higher education. We see that individuals from relatively less well off backgrounds are less likely to participate in higher education, are more likely to be in other full-time education (particularly without a job) and are also more likely to be unemployed.

To summarise, these findings reveal differential patterns of activity between individuals from different socio-economic backgrounds. In particular, individuals from relatively less well off backgrounds are less likely to engage in work whilst in school, but enrolment in post-compulsory second-level education is fairly similar across the two groups. Further, university enrolment is lower amongst individuals from relatively less well off backgrounds. However, the findings do not reveal any substantial differences in responsiveness to the subsidy across the two groups of individual.

4.7 Conclusion

Previous evaluation of the EMA has shown that it has a positive and significant effect on education enrolment for the first two waves, with no discernible impact in the third wave. This analysis, using a dynamic discrete choice model controlling for unobserved heterogeneity and allowing for attrition to be non-random in an unobserved dimension, also bears out these findings. The EMA increases participation in education without a job and increasing the generosity of the EMA would augment such participation even further. This suggests that the EMA may be leading individuals who choose to remain in education, to participate less in part-time work and in this sense, the subsidy may represent a substitute for part-time wages. The positive effect is partly at the expense of participation in education with a job. However, the overall net effect on participation in post-compulsory second-level education is positive. A complete understanding of the costs and benefits for individuals of switching from education with a job into education without a job, requires further analysis. Indeed the effects of work whilst in school on educational attainment are still widely debated. Future work will investigate these interactions. After its withdrawal in wave 3, it does not have any discernible impact on participation in higher education. Another important finding of this work is that controlling for attrition has no fundamental effect on the structural parameter estimates. This suggests that the effects of attrition may be accounted for by controlling for observable characteristics alone.

In terms of future work, more waves of data will enable us to examine longer term effects of the subsidy, including participation in, and attachment to, higher education. We will also examine the effects of having a part-time job whilst in full-time post-compulsory education on academic performance, along with issues to do with gender differences in various aspects of the programme.

4.8 Tables

Table 4.1: The Four EMA Variants

Variant	**Maximum weekly EMA award £**	**Weekly payment paid to**	**Termly Retention bonus £**	**Achievement bonus £**
1	30	Young person	50	50
2	40	Young person	50	50
3	30	Parent	50	50
4	30	Young person	80	140

Table 4.2: Data Summary

Year	Activity in month	Wave data collected	Sample size	Outcome
1999	September	1	$N^{na} + N^{aw3} + N^{aw2}$	y_1
2000	February	2	$N^{na} + N^{aw3}$	y_2
2000	May	2	$N^{na} + N^{aw3}$	y_3
2000	September	2	$N^{na} + N^{aw3}$	y_4
2001	February	3	N^{na}	y_5
2001	May	3	N^{na}	y_6
2001	September	3	N^{na}	y_7

N^{na} =number of individuals who do not attrit, N^{aw3} =number of individuals who attrit in wave 3, N^{aw2} =number of individuals who attrit in wave 2.

Table 4.3: List of Variables

Variable	
Male	male
Number of older siblings	siboldno
Number of younger siblings	sibyngno
GCSE maths score	mathsscore
GCSE English score	engscore
Dummy for missing GCSE maths score	zmathsscore
Dummy for missing GCSE English score	zengscore
Nonwhite	nonwhite
In education previous period	educl
In work previous period	workl
Person had serious illness between 0 and 1 years old	ill01
1 or 2 parents in work when born	parwork
Mother and father figure present	mothfath
Father figure present	fathfig
Mother in full or part time work	mothwk
Father in full or part time work	fathwk
Mother has O levels or higher	mthq
Father has O levels or higher	fthq
Dummy for father variables missing	zfathvar
Family income	fims
Family income squared	fim2s
Family receives means-tested benefit	lowinc
Owner occupier	ownocc
Lives in rural area	rural
EMA entitlement (=0 for controls)	emaest
Health, Deprivation and Disability	health
Education, Skills and Training	educatio
Housing	housing
Geographical Access to services	access
Percentage not going to university	pcnotuni
Percentage not staying on post-16	pcnotsta

Table 4.4: Participation in Activities, Cohort 1 Waves 1 and 2

	Eligibles				
	EJ	**ENJ**	**FTW**	**PTW**	**UEO**
1999/2000					
September					
P=4,543	0.24 (0.43)	0.50 (0.50)	0.12 (0.33)	0.03 (0.18)	0.11 (0.31)
C=2,744	0.30 (0.46)	0.40 (0.49)	0.14 (0.35)	0.05 (0.21)	0.12 (0.32)
February					
P=3,338	0.32 (0.47)	0.45 (0.50)	0.14 (0.35)	0.03 (0.17)	0.06 (0.23)
C=2,049	0.39 (0.49)	0.33 (0.47)	0.18 (0.39)	0.03 (0.18)	0.06 (0.25)
May					
P=3,338	0.34 (0.47)	0.39 (0.49)	0.17 (0.37)	0.03 (0.17)	0.07 (0.25)
C=2,049	0.40 (0.49)	0.28 (0.45)	0.22 (0.41)	0.04 (0.20)	0.07 (0.25)
2000/2001					
September					
P=3,338	0.34 (0.47)	0.33 (0.47)	0.20 (0.40)	0.04 (0.19)	0.09 (0.29)
C=2,049	0.40 (0.49)	0.21 (0.41)	0.26 (0.44)	0.04 (0.19)	0.09 (0.29)
February					
P=2,582	0.37 (0.48)	0.31 (0.46)	0.21 (0.41)	0.04 (0.19)	0.07 (0.26)
C=1,628	0.41 (0.49)	0.21 (0.40)	0.27 (0.44)	0.04 (0.20)	0.07 (0.26)
May					
P=2,582	0.37 (0.48)	0.28 (0.45)	0.24 (0.43)	0.04 (0.20)	0.07 (0.26)
C=1,628	0.39 (0.49)	0.20 (0.40)	0.28 (0.45)	0.05 (0.22)	0.07 (0.26)

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other; P=pilots, C=controls.

Table 4.5: Male Participation in Activities, Cohort 1 Waves 1 and 2

	Eligible Males				
	EJ	**ENJ**	**FTW**	**PTW**	**UEO**
1999/2000					
September					
P=2,270	0.19 (0.39)	0.50 (0.50)	0.15 (0.19)	0.04 (0.36)	0.12 (0.33)
C=1,352	0.24 (0.42)	0.40 (0.48)	0.19 (0.39)	0.05 (0.22)	0.13 (0.33)
February					
P=1,629	0.27 (0.44)	0.48 (0.50)	0.17 (0.38)	0.03 (0.16)	0.06 (0.23)
C=988	0.33 (0.47)	0.34 (0.48)	0.23 (0.42)	0.03 (0.17)	0.07 (0.25)
May					
P=1,629	0.28 (0.45)	0.42 (0.49)	0.21 (0.41)	0.02 (0.15)	0.06 (0.24)
C=988	0.32 (0.47)	0.31 (0.46)	0.27 (0.44)	0.03 (0.18)	0.07 (0.26)
2000/2001					
September					
P=1,629	0.29 (0.45)	0.35 (0.48)	0.24 (0.43)	0.03 (0.18)	0.09 (0.28)
C=988	0.32 (0.47)	0.24 (0.43)	0.30 (0.46)	0.03 (0.18)	0.10 (0.30)
February					
P=1,245	0.33 (0.47)	0.32 (0.47)	0.25 (0.43)	0.03 (0.16)	0.07 (0.26)
C=781	0.33 (0.47)	0.24 (0.43)	0.33 (0.47)	0.03 (0.18)	0.07 (0.26)
May					
P=1,245	0.33 (0.47)	0.30 (0.46)	0.27 (0.44)	0.03 (0.17)	0.07 (0.25)
C=781	0.32 (0.47)	0.23 (0.42)	0.34 (0.47)	0.04 (0.19)	0.08 (0.27)

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other; P=pilots, C=controls.

Table 4.6: Female Participation in Activities, Cohort 1 Waves 1 and 2

	Eligible Females				
	EJ	**ENJ**	**FTW**	**PTW**	**UEO**
1999/2000					
September					
P=2,273	0.29 (0.45)	0.49 (0.50)	0.10 (0.30)	0.03 (0.16)	0.10 (0.30)
C=1,392	0.36 (0.48)	0.40 (0.49)	0.09 (0.29)	0.04 (0.20)	0.11 (0.31)
February					
P=1,709	0.37 (0.48)	0.42 (0.49)	0.12 (0.32)	0.03 (0.18)	0.06 (0.23)
C=1,061	0.45 (0.50)	0.31 (0.47)	0.14 (0.35)	0.04 (0.19)	0.06 (0.24)
May					
P=1,709	0.40 (0.49)	0.37 (0.48)	0.13 (0.33)	0.04 (0.19)	0.07 (0.26)
C=1,061	0.47 (0.50)	0.25 (0.44)	0.17 (0.37)	0.05 (0.22)	0.06 (0.24)
2000/2001					
September					
P=1,709	0.39 (0.49)	0.32 (0.47)	0.16 (0.37)	0.04 (0.20)	0.09 (0.29)
C=1,061	0.48 (0.50)	0.19 (0.390)	0.21 (0.41)	0.04 (0.20)	0.08 (0.27)
February					
P=1,337	0.40 (0.49)	0.29 (0.46)	0.18 (0.39)	0.05 (0.21)	0.08 (0.27)
C=847	0.48 (0.50)	0.17 (0.38)	0.22 (0.41)	0.05 (0.22)	0.07 (0.26)
May					
P=1,337	0.40 (0.49)	0.26 (0.44)	0.21 (0.41)	0.05 (0.23)	0.08 (0.27)
C=847	0.46 (0.50)	0.18 (0.38)	0.23 (0.42)	0.06 (0.24)	0.07 (0.25)

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other; P=pilots, C=controls.

Table 4.7: Participation in Activities, Cohort 1 Wave 3

	Eligibles		
	All	**Males**	**Females**
2001/2002			
September			
UJ			
P	0.10 (0.30)	0.08 (0.27)	0.11 (0.32)
C	0.08 (0.26)	0.06 (0.24)	0.09 (0.29)
UNJ			
P	0.14 (0.35)	0.14 (0.35)	0.15 (0.35)
C	0.13 (0.33)	0.11 (0.31)	0.15 (0.35)
OEJ			
P	0.11 (0.32)	0.11 (0.31)	0.12 (0.32)
C	0.13 (0.33)	0.12 (0.33)	0.13 (0.34)
OENJ			
P	0.09 (0.29)	0.11 (0.32)	0.07 (0.26)
C	0.07 (0.25)	0.09 (0.29)	0.05 (0.21)
FTW			
P	0.35 (0.48)	0.37 (0.48)	0.32 (0.47)
C	0.39 (0.49)	0.42 (0.49)	0.36 (0.48)
PTW			
P	0.09 (0.29)	0.07 (0.25)	0.11 (0.31)
C	0.09 (0.29)	0.07 (0.25)	0.12 (0.32)
UEO			
P	0.12 (0.32)	0.12 (0.33)	0.12 (0.32)
C	0.12 (0.33)	0.14 (0.35)	0.11 (0.31)
P	2,582	1,245	1,337
C	1,628	781	847

UJ=university with a job, UNJ=university without a job, OEJ=other full-time education with a job, OENJ=other full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other; P=pilots, C=controls.

Table 4.8: Transitions, Cohort 1 Wave 1, Eligible Pilots and Controls

		Number in listed state, wave 1				
Term1 →	**Term2**	**→**		**Term 3**		
EJ 1,907 →	ej 1,189	→ej 1106	enj 32	ftw 27	ptw 19	ueo 5
	enj 232	→ej 44	enj 176	ftw 4	ptw 2	ueo 6
	ftw 68	→ej 0	enj 0	ftw 63	ptw 1	ueo 4
	ptw 25	→ej 2	enj 0	ftw 3	ptw 18	ueo 2
	ueo 15	→ej 2	enj 0	ftw 4	ptw 2	ueo 7
	att 378	(19.8 %)				
ENJ 3,341 →	ej 579	→ej 509	enj 39	ftw 16	ptw 10	ueo 5
	enj 1,813	→ej 158	enj 1590	ftw 27	ptw 5	ueo 33
	ftw 85	→ej 1	enj 0	ftw 82	ptw 0	ueo 2
	ptw 37	→ej 0	enj 1	ftw 4	ptw 28	ueo 4
	ueo 67	→ej 0	enj 1	ftw 7	ptw 2	ueo 57
	att 760	(22.7 %)				
FTW 948 →	ej 19	→ej 15	enj 0	ftw 2	ptw 0	ueo 2
	enj 33	→ej 3	enj 26	ftw 2	ptw 0	ueo 2
	ftw 486	→ej 2	enj 3	ftw 465	ptw 4	ueo 12
	ptw 26	→ej 0	enj 0	ftw 8	ptw 16	ueo 2
	ueo 39	→ej 0	enj 0	ftw 11	ptw 3	ueo 25
	att 345	(36.4 %)				
PTW 278 →	ej 45	→ej 39	enj 0	ftw 5	ptw 1	ueo 0
	enj 20	→ej 3	enj 13	ftw 0	ptw 3	ueo 1
	ftw 77	→ej 0	enj 1	ftw 73	ptw 2	ueo 1
	ptw 36	→ej 1	enj 2	ftw 4	ptw 29	ueo 0
	ueo 22	→ej 0	enj 1	ftw 6	ptw 1	ueo 14
	att 78	(28.1 %)				
UEO 813 →	ej 38	→ej 34	enj 2	ftw 1	ptw 1	ueo 0
	enj 77	→ej 6	enj 64	ftw 4	ptw 0	ueo 3
	ftw 134	→ej 1	enj 0	ftw 124	ptw 0	ueo 9
	ptw 47	→ej 0	enj 0	ftw 8	ptw 33	ueo 6
	ueo 178	→ej 0	enj 2	ftw 24	ptw 8	ueo 144
	att 339	(41.7 %)				
N_1 = 7,287	N_2=5,387			N_3=5,387		

Table 4.9: Transitions, Cohort 1 Wave 2, Eligible Pilots and Controls

		Number in listed state, wave 2					
Lag	**Term1 →**	**Term2**	**→**		**Term 3**		
ej 1,566		ej 1,350	→ej 1,225	enj 61	ftw 35	ptw 24	ueo 5
enj 321		enj 174	→ej 46	enj 119	ftw 7	ptw 0	ueo 2
ftw 40	EJ 1,957 →	ftw 68	→ej 2	enj 0	ftw 66	ptw 0	ueo 0
ptw 20		ptw 30	→ej 3	enj 0	ftw 7	ptw 20	ueo 0
ueo 10		ueo 12	→ej 1	enj 0	ftw 3	ptw 2	ueo 6
		att 323	(16.5 %)				
ej 148		ej 220	→ej 179	enj 24	ftw 2	ptw 10	ueo 5
enj 1,318		enj 918	→ej 85	enj 801	ftw 11	ptw 6	ueo 15
ftw 28	ENJ 1,546 →	ftw 32	→ej 1	enj 3	ftw 25	ptw 2	ueo 1
ptw 8		ptw 9	→ej 0	enj 2	ftw 1	ptw 5	ueo 1
ueo 44		ueo 44	→ej 0	enj 2	ftw 5	ptw 3	ueo 34
		att 323	(20.9 %)				
ej 146		ej 11	→ej 7	enj 0	ftw 4	ptw 0	ueo 0
enj 103		ej 10	→ej 1	enj 6	ftw 3	ptw 0	ueo 0
ftw 828	FTW 1,200 →	ftw 754	→ej 1	enj 2	ftw 724	ptw 2	ueo 25
ptw 49		ptw 31	→ej 2	enj 0	ftw 4	ptw 22	ueo 3
ueo 74		ueo 74	→ej 0	enj 1	ftw 22	ptw 4	ueo 47
		att 320	(26.7 %)				
ej 43		ej 8	→ej 7	enj 1	ftw 0	ptw 0	ueo 0
enj 32		enj 10	→ej 0	enj 9	ftw 0	ptw 0	ueo 1
ftw 15	PTW 201 →	ftw 44	→ej 0	enj 0	ftw 38	ptw 2	ueo 4
ptw 87		ptw 63	→ej 2	enj 1	ftw 5	ptw 52	ueo 3
ueo 24		ueo 27	→ej 0	enj 0	ftw 6	ptw 2	ueo 19
		att 49	(24.4 %)				
ej 43		ej 24	→ej 22	enj 2	ftw 0	ptw 0	ueo 0
enj 117		enj 20	→ej 3	enj 13	ftw 1	ptw 2	ueo 1
ftw 94	UEO 483 →	ftw 95	→ej 1	enj 0	ftw 81	ptw 3	ueo 10
ptw 20		ptw 28	→ej 1	enj 2	ftw 7	ptw 15	ueo 3
ueo 209		ueo 153	→ej 0	enj 1	ftw 15	ptw 14	ueo 123
		att 163	(33.7 %)				
N_4 = 5,387	N_5=4,209			N_6=4,209			

Table 4.10: Transitions, Cohort 1 Wave 3, Eligible Pilots and Controls

Number in listed state, wave 3			
Lag	**Term1**	**Lag**	**Term1**
ej 295		ej 356	
enj 62		enj 168	
ftw 8	UJ 372	ftw 895	FTW 1,527
ptw 6		ptw 49	
ueo 1		ueo 59	
ej 271		ej 171	
enj 295		enj 66	
ftw 6	UNJ 577	ftw 21	PTW 382
ptw 2		ptw 89	
ueo 3		ueo 35	
ej 387		ej 61	
enj 73		enj 121	
ftw 26	OEJ 497	ftw 110	UEO 505
ptw 8		ptw 30	
ueo 3		ueo 183	
ej 48			
enj 265			
ftw 7	OENJ 350		
ptw 6			
ueo 24			
	N_7 = 4,210		

UJ=university with a job, UNJ=university without a job, OEJ=other full-time education with a job, OENJ=other full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other; P=pilots, C=controls.

Table 4.11: Cross-tabulations of Characteristics and Attrition, Cohort 1

		Eligibles	
Attrition Status	**Do not attrit**	**Attrit wave 2**	**Attrit wave 3**
Number of observations	4,210	1,900	1,178
Variable			
Male	0.4812	0.5289	0.5025
Number of older siblings	0.9423	1.1126	1.0110
Number of younger siblings	0.9373	1.0516	1.0951
GCSE maths score	4.0708	2.9932	3.5119
GCSE English score	4.5064	3.4479	3.9771
Non-white	0.0796	0.1111	0.1104
Young person ill b/w 0 and 1	0.2207	0.2353	0.2326
1 or 2 parents in work when born	0.8527	0.7400	0.8014
Mother and father figure present	0.6261	0.4342	0.5441
Father figure present	0.7722	0.6184	0.6952
Mother in full or part time work	0.6606	0.4995	0.5620
Father in full or part time work	0.5083	0.3416	0.4244
Mother has O levels or higher	0.4879	0.3559	0 4261
Father has O levels or higher	0.3734	0.2168	0.3073
Father variables missing	0.3686	0.5316	0.4312
Income is $\geq$ £13,000 per annum	0.4658	0.3168	0.3829
Family receives means-tested benefit	0.2480	0.4011	0.3472
Owner occupier	0.7157	0.4895	0.6188
Lives in rural area	0.2993	0.2153	0.2717
Lives in pilot area	0.6133	0.6342	0.6418
Lives in pilot area and eligible for taper	0.2789	0.1958	0.2436
Health, Deprivation and Disability	0.8314	0.9823	0.9124
Education, Skills and Training	0.5959	0.8353	0.6657
Housing Index	0.3276	0.5408	0.3952
Geographical Access to services	-0.3182	-0.4622	-0.3462
Percentage not going to university	88.47	89.61	88.83
Percentage not staying on post-16	66.29	67.99	66.83

Table 4.12: Attrition Parameters, Waves 2 and 3

	Not attrit wave 2	Not attrit wave 3
male	-0.13 (0.06)*	0.06 (0.12)
siboldno	-0.06 (0.02)*	0.07 (0.05)
sibyngno	0.00 (0.03)	-0.04 (0.05)
mothfath	0.43 (0.11)*	0.66 (0.22)*
mothftwk	0.08 (0.08)	0.09 (0.17)
mothptwk	0.22 (0.07)*	0.23 (0.15)
fathftwk	-0.21 (0.10)	0.31 (0.19)
fathptwk	-0.16 (0.19)	1.55 (0.61)*
mthq	0.12 (0.06)*	0.07 (0.13)
fthq	0.13 (0.08)	0.17 (0.16)
fathfig	0.10 (0.09)	0.82 (0.19)*
zfathvar	-0.11 (0.10)	1.27 (0.20)*
lowinc	-0.09 (0.07)	0.28 (0.16)
rural	0.20 (0.07)*	-0.36 (0.13)*
pilot	-0.08 (0.07)	0.13 (0.15)
taper	0.12 (0.10)	0.41 (0.21)*
ema tap	0.07 (0.12)	-0.58 (0.25)*
mathsscore	0.08 (0.02)*	-0.03 (0.05)
engscore	0.05 (0.03)	0.20 (0.06)*
eductl/4	0.30 (0.09)*	0.90 (0.20)*
worktl/4	-0.04 (0.10)	0.46 (0.21)*
Interviewer Age		
age1	-0.40 (0.19)*	-0.17 (0.28)
age2	0.07 (0.13)	-0.24 (0.14)
age3	0.18 (0.12)	-0.22 (0.14)
age4	0.11 (0.12)	-0.61 (0.10)*
age5	0.06 (0.11)	-0.95 (0.19)*
age6	0.15 (0.11)	-1.01 (0.23)*
age7	-0.11 (0.12)	-0.49 (0.11)*
δ_w	-0.03 (0.01)*	0.50 (0.15)*

We also control for ethnicity and missing maths and English GCSE dummies. Standard errors in parentheses. Points of support $\eta_1 = 0.41(0.17)$, $\eta_2 = 0.09(0.04)$. * denotes significance at the 5% level or less.

Table 4.13: Model Selection

(1) Correlated Attrition, Unobserved Heterogeneity	
log L	-39779.8
Number of parameters, k	762
Sample size, n	26,484
BIC	-82930.0
(2) Uncorrelated Attrition, Unobserved Heterogeneity	
log L	-41434.3
Number of parameters, k	760
Sample size, n	26,484
BIC	-86230.0
(3) Uncorrelated Attrition, No Unobserved Heterogeneity	
log L	-42989.5
Number of parameters, k	736
Sample size, n	26,484
BIC	-89234.3

BIC= $2 * logL(\hat{\theta}_p; X) - k * log(n)$, where $logL$ is the negative of the maximised log likelihood, $\hat{\theta}_p$ are the parameters estimated in the p^{th} model, k is the number of unspecified parameters in the p^{th} model, and n is the number of observations.

Table 4.14: Comparison of Actual and Predicted Proportions

	Actual	**Predicted**
	Percentage of Pilots-Percentage of Controls	Percentage of Pilots-Percentage of Controls
Sept 99		
EJ	-0.0636	-0.0486
ENJ	0.1000	0.0892
FTW	-0.0157	-0.0127
PTW	-0.0142	-0.0160
UEO	-0.0063	-0.0118
Feb/May 00		
EJ	-0.0624	-0.0586
ENJ	0.1175	0.1181
FTW	-0.0438	-0.0438
PTW	-0.0082	-0.0102
UEO	-0.0029	-0.0054
Sept 00		
EJ	-0.0635	-0.0725
ENJ	0.1173	0.1459
FTW	-0.0524	-0.0590
PTW	-0.0004	-0.0062
UEO	-0.0010	-0.0080
Feb/May 01		
EJ	-0.0319	-0.0370
ENJ	0.0890	0.0962
FTW	-0.0509	-0.0512
PTW	-0.0066	-0.0096
UEO	0.0004	0.0017
Sept 01		
UJ	-0.0425	0.0097
UNJ	-0.0032	0.0123
OEJ	-0.0037	-0.0024
OENJ	0.0208	0.0230
FTW	0.0171	-0.0377
PTW	-0.0118	-0.0033
UEO	0.0233	-0.0017

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other, UJ=university with a job, UNJ=university without a job, OEJ=other full-time education with a job, OENJ=other full-time education without a job.

Table 4.15: Difference in Participation in Activities across Types

	Type 1 - Type 2				
	Sept99	**Feb/May00**	**Sept00**	**Feb/May01**	**Sept01**
EJ	-0.2481	-0.2685	-0.4377	-0.2587	-0.1811
ENJ	0.4564	0.2887	0.4935	0.2644	0.0896
FTW	-0.1142	-0.0528	-0.1255	-0.0426	0.2759
PTW	0.0009	0.0223	0.0444	0.0233	-0.0211
UEO	-0.0950	0.0103	0.0252	0.0135	-0.0306
UJ	-	-	-	-	-0.1048
UNJ	-	-	-	-	-0.0277
Number Type 1	6,194	4,579	4,579	3,578	3,578
Number Type 2	1,093	808	808	631	631

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other. EJ and ENJ in Sept 01 include repeat A/AS levels as well as third-level colleges. Number of type 1's and 2's are estimated from the model. Percentage point differences in unconditional probabilities are reported:

$$\sum_j \left[Pr(y_{it} = k|y_{it-1} = j, X, \eta_i = r_1).Pr(y_{it-1} = j|X, \eta_i = r_1)\right] -$$

$$\sum_j \left[Pr(y_{it} = k|y_{it-1} = j, X, \eta_i = r_2).Pr(y_{it-1} = j|X, \eta_i = r_2)\right]$$

Table 4.16: Predicted Sequences of Choices

	Pilots	Controls		Pilots	Controls	
	Type 1			Type 2		
			EMA Effect			EMA Effect
EEEEE	0.2278	0.1875	0.0403	0.2956	0.2775	0.0181
EEEEN	0.2714	0.2381	0.0333	0.1003	0.0973	0.0030
EEENN	0.0784	0.0679	0.0105	0.0416	0.0412	0.0004
EENNN	0.0489	0.0582	-0.0093	0.0401	0.0406	-0.0005
ENNNN	0.0936	0.1278	-0.0342	0.0706	0.0884	-0.0178
NNNNN	0.0276	0.0482	-0.0206	0.0841	0.1103	-0.0262
NEEEE	0.0173	0.0163	0.0010	0.0720	0.0670	0.0050
NNEEE	0.0016	0.0018	-0.0002	0.0072	0.0075	-0.0003
NNNEE	0.0009	0.0013	-0.0004	0.0067	0.0061	0.0006
NNNNE	0.0017	0.0024	-0.0007	0.0108	0.0129	-0.0021
NNEEN	0.0031	0.0035	-0.0004	0.0048	0.0045	0.0003
NENNN	0.0102	0.0158	-0.0056	0.0317	0.0323	-0.0006
NEENN	0.0121	0.0140	-0.0019	0.0223	0.0247	-0.0024
OTHER	0.1792	0.1897	-0.0105	0.1774	0.1572	0.0202

'E' includes education with and without a job, 'N' includes full-time and part-time work and unemployment/other. The 'OTHER' path includes all paths not otherwise listed. Sample of non-attritors only.

Table 4.17: Relationship of Type to Family Background Characteristics

	Type 1	
Male	0.0060	(0.0012)*
Number older siblings	0.0010	(0.0005)*
Mother in full or part time work	0.0036	(0.0014)*
Father in full or part time work	0.0030	(0.0019)
Mother has O levels or higher	-0.0045	(0.0012)*
Father has O levels or higher	0.0005	(0.0015)
Father figure present	0.0018	(0.0021)
Lives in rural area	0.0048	(0.0013)*
Had serious illness between 0 and 1 years old	-0.0028	(0.0014)*
Family income	-0.0036	(0.0019)
Family income squared	0.0006	(0.0005)
Nonwhite	0.0111	(0.0023)*
GCSE maths score	-0.0043	(0.0005)*
GCSE English score	-0.0041	(0.0006)*
Pilot	0.0014	(0.0011)
Proportion of type 1	85%	
N	4,209	
R^2	0.2690	

The table reports the coefficients on the listed variables in a regression in which the dependent variable is probability of being type 1. Standard errors in parentheses. * denotes significance at the 5% level or less.

Table 4.18: Change the EMA Generosity, Wave 1

	Δ in proportions: simulated less actual				
	EJ	**ENJ**	**FTW**	**PTW**	**UEO**
Simulation					
No EMA					
SEPT 99	0.0259	-0.0673	0.0150	0.0098	0.0168
	(0.0055)*	(0.0093)*	(0.0036)*	(0.0021)*	(0.0020)*
FEB/MAY 00	0.0395	-0.0704	0.0214	0.0044	0.0050
	(0.0120)*	(0.0145)*	(0.0056)*	(0.0039)	(0.0023)*
Decrease EMA by 50%					
SEPT 99	0.0129	-0.0306	0.0071	-0.0039	0.0061
	(0.0025)*	(0.0043)*	(0.0017)*	(0.0009)*	(0.0009)*
FEB/MAY 00	0.0209	-0.0391	0.0143	0.0016	0.0020
	(0.0054)*	(0.0072)*	(0.0072)*	(0.0045)	(0.0014)
Increase EMA by 50%					
SEPT 99	-0.0121	0.0282	-0.0066	0.0045	-0.0055
	(0.0021)	(0.0037)*	(0.0014)*	(0.0006)*	(0.0007)*
FEB/MAY 00	-0.0205	0.0615	-0.0271	-0.0039	-0.0099
	(0.0063)*	(0.0122)*	(0.0090)*	(0.0059)	(0.0029)*
Double EMA					
SEPT 99	-0.0235	0.0541	-0.0126	-0.0074	-0.0104
	(0.0039)*	(0.0067)*	(0.0025)*	(0.0011)*	(0.0013)*
FEB/MAY 00	-0.0424	0.1553	-0.0646	-0.0120	-0.0362
	(0.0121)*	(0.0264)*	(0.0160)*	(0.0109)	(0.0056)*

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other. Standard errors in parentheses are estimated using 500 random draws from the normal distribution of estimated parameters. * denotes significance at the 5% level or less.

Table 4.19: Change the EMA Generosity, Wave 2

	Δ in proportions: simulated less actual				
	EJ	**ENJ**	**FTW**	**PTW**	**UEO**
Simulation					
No EMA					
SEPT 00	0.0602	-0.0943	0.0291	-0.0056	-0.0107
	(0.0074)*	(0.0152)*	(0.0079)*	(0.0060)	(0.0060)*
FEB/MAY 01	0.0427	-0.0500	0.0087	-0.0009	-0.0004
	(0.0144)*	(0.0156)*	(0.0072)	(0.0057)	(0.0024)
Decrease EMA by 50%					
SEPT 00	0.0277	-0.0486	0.0194	-0.0044	0.0059
	(0.0034)*	(0.0086)*	(0.0048)*	(0.0039)	(0.0034)
FEB/MAY 01	0.0240	-0.0277	0.0058	-0.0010	-0.0011
	(0.0075)*	(0.0084)*	(0.0052)	(0.0042)	(0.0017)
Increase EMA by 50%					
SEPT 00	-0.0239	0.0630	-0.0315	0.0063	-0.0139
	(0.0023)	(0.0127)*	(0.0058)*	(0.0069)	(0.0051)*
FEB/MAY 01	-0.0240	0.0311	-0.0107	0.0018	0.0018
	(0.0072)*	(0.0104)*	(0.0092)	(0.0078)	(0.0036)
Double EMA					
SEPT 00	-0.0439	0.1415	-0.0703	0.0082	-0.0355
	(0.0035)*	(0.0266)*	(0.0094)*	(0.0167)	(0.0087)*
FEB/MAY 01	-0.0476	0.0745	-0.0288	0.0035	-0.0016
	(0.0136)*	(0.0289)*	(0.0218)	(0.0201)	(0.0104)

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other. Standard errors in parentheses are estimated using 500 random draws from the normal distribution of estimated parameters. * denotes significance at the 5% level or less.

Table 4.20: Difference in Activity across Pilots and Controls, Wave 3

	Δ in proportions: pilot less control						
	UJ	**UNJ**	**OEJ**	**OENJ**	**FTW**	**PTW**	**UEO**
SEPT 01	0.0077	0.0088	-0.0109	0.0142	-0.0136	-0.0046	-0.0016
	(0.0050)	(0.0094)	(0.0084)	(0.0112)	(0.0159)	(0.0146)	(0.0058)

UJ=university with a job, UNJ=university without a job, OEJ=other full-time education with a job, OENJ=other full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other. Standard errors in parentheses are estimated using 500 random draws from the normal distribution of estimated parameters. * denotes significance at the 5% level or less.

Table 4.21: Activities under Various EMA Generosities, by Background, Waves 1 and 2

EMA	Actual (1)		No Ema (2)		Double (3)	
Group	**Upper**	**Lower**	**Upper**	**Lower**	**Upper**	**Lower**
Activity						
Sept 99						
EJ	0.2646	0.2089	0.2924	0.2329	0.2397	0.1876
ENJ	0.5350	0.5211	0.4884	0.4530	0.5847	0.5896
FTW	0.1238	0.1286	0.1364	0.1456	0.1300	0.1369
PTW	0.0277	0.0377	0.0349	0.0500	0.0311	0.0435
UEO	0.0489	0.1037	0.0577	0.1283	0.0531	0.1156
Feb/May 00						
EJ	0.3315	0.2713	0.3705	0.3114	0.2900	0.2311
ENJ	0.3534	0.3798	0.2890	0.3021	0.4867	0.5439
FTW	0.2096	0.1945	0.2287	0.2186	0.1527	0.1313
PTW	0.0385	0.0529	0.0415	0.0589	0.0279	0.0363
UEO	0.0687	0.1014	0.0700	0.1087	0.0424	0.0572
Sept00						
EJ	0.2826	0.2349	0.3424	0.2956	0.2346	0.1893
ENJ	0.3613	0.3917	0.2737	0.2894	0.4865	0.5437
FTW	0.2120	0.1903	0.2380	0.2230	0.1527	0.1223
PTW	0.0519	0.0579	0.0505	0.0564	0.0630	0.0693
UEO	0.0883	0.1204	0.0953	0.1353	0.0629	0.0752
Feb/May 01						
EJ	0.3837	0.3202	0.4242	0.3659	0.3409	0.2752
ENJ	0.1974	0.2387	0.1512	0.1833	0.2668	0.3195
FTW	0.2890	0.2772	0.2964	0.2877	0.2619	0.2432
PTW	0.0495	0.0623	0.0486	0.0616	0.0519	0.0648
UEO	0.0803	0.1016	0.0795	0.1013	0.0783	0.0971

EJ=full-time education with a job, ENJ=full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other.

Table 4.22: Activities in Wave 3, by Background

Group	Actual EMA Upper	Actual EMA Lower
Activity Sept 01		
UJ	0.0596	0.0385
UNJ	0.1231	0.0752
OEJ	0.0728	0.0867
OENJ	0.0859	0.1417
FTW	0.4664	0.4332
PTW	0.1066	0.1090
UEO	0.0852	0.1154

UJ=university with a job, UNJ=university without a job, OEJ=other full-time education with a job, OENJ=other full-time education without a job, FTW=full-time work, PTW=part-time work, UEO=unemployment/other.

4.9 Appendix A

Let $S_{iw} = 1$ if individual i does not attrit in wave w, for $w = 2, 3$. Further $S_{i3} = 1 \implies S_{i2} = 1$. The individual contributions to the log likelihood function that jointly models attrition and the main choice of the individual are derived under the assumptions that

$$
\begin{aligned}
y_{it} \perp Z_i | X_i, y_{it-1}, S_{i3} = 1 \quad & \forall \; i, \quad \text{for} \quad t = 5, 6, 7 \\
y_{it} \perp Z_i | X_i, y_{it-1}, S_{i2} = 1 \quad & \forall \; i, \quad \text{for} \quad t = 2, 3, 4 \\
y_{i1} \perp Z_i | X_i \quad & \forall \; i \\
S_{i3} \perp y_{i3}, y_{i2}, y_{i1} | X_i, Z_i, y_{i4} \quad & \forall \; i \\
y_{it} \perp y_{it-k} | y_{it-1} \quad & \forall \; i, \quad \forall \; t > 2, \; \forall \; 2 \leq k < t
\end{aligned} \tag{4.11}
$$

We show the actual log likelihood contributions for three groups of individual: those who do not attrit, those who attrit in wave 3 and those who attrit in wave 2, under the structure of our data (see table 4.2). For the exposition, we assume that the individual chooses the same activity in all time periods. For simplicity we omit the i subscripts.

1. Individuals who have not attrited by the end of wave 3

$$
\begin{aligned}
l^{na} = Pr(S_{i3} = 1, & S_{i2} = 1, y_{i1} = \cdots = y_{i7} = k | X_i, Z_i) = \\
log \sum_{m=1}^{M} \pi_m \Big(& Pr(y_{i7} = k | y_{i6} = k, S_{i3} = 1, X_i, \eta_i = r_m). \\
& Pr(y_{i6} = k | y_{i5} = k, S_{i3} = 1, X_i, \eta_i = r_m). \\
& Pr(y_{i5} = k | y_{i4} = k, S_{i3} = 1, X_i, \eta_i = r_m). \\
& Pr(S_{i3} = 1 | y_{i4} = k, S_{i2} = 1, X_i, Z_i, \eta_i = r_m). \\
& Pr(y_{i4} = k | y_{i3} = k, S_{i2} = 1, X_i, \eta_i = r_m). \\
& Pr(y_{i3} = k | y_{i2} = k, S_{i2} = 1, X_i, \eta_i = r_m). \\
& Pr(y_{i2} = k | y_{i1} = k, S_{i2} = 1, X_i, \eta_i = r_m). \\
& Pr(S_{i2} = 1 | y_{i1} = k, X_i, Z_i, \eta_i = r_m). \\
& Pr(y_{i1} = k | X_i, \eta_i = r_m) \Big)
\end{aligned} \tag{4.12}
$$

2. Individuals who attrit in wave 3 conditional on not attriting in wave 2

$$
\begin{aligned}
l^{aw3} = Pr(S_{i3} = 0, S_{i2} = 1, y_{i1} = \cdots = y_{i4} = k | X_i, Z_i) = \\
log \sum_{m=1}^{M} \pi_m \Big(& Pr(S_{i3} = 0 | y_{i4} = k, S_{i2} = 1, X_i, Z_i, \eta_i = r_m). \\
& Pr(y_{i4} = k | y_{i3} = k, S_{i2} = 1, X_i, \eta_i = r_m). \\
& Pr(y_{i3} = k | y_{i2} = k, S_{i2} = 1, X_i, \eta_i = r_m). \\
& Pr(y_{i2} = k | y_{i1} = k, S_{i2} = 1, X_i, \eta_i = r_m). \\
& Pr(S_{i2} = 1 | y_{i1} = k, X_i, Z_i, \eta_i = r_m). \\
& Pr(y_{i1} = k | X_i, \eta_i = r_m) \Big)
\end{aligned}
\tag{4.13}
$$

3. Individuals who attrit in wave 2

$$
\begin{aligned}
l^{aw2} = Pr(S_{i2} = 0, y_{i1} = k | X_i, Z_i) = \\
log \sum_{m=1}^{M} \pi_m \Big(Pr(S_{i2} = 0 | y_{i1} = k, X_i, Z_i, \eta_i = r_m). Pr(y_{i1} = k | X_i, \eta_i = r_m) \Big)
\end{aligned}
\tag{4.14}
$$

The functional forms for the choice probabilities are given by equation (4.4) and for the attrition probabilities by equation (4.9) in the text.

4.10 Appendix B

Table 4.23: Parameter Estimates September 1999

	EJ	ENJ	FTW	PTW
Variables				
male	-0.55 (0.09)*	-0.09 (0.08)	0.32 (0.09)*	0.07 (0.13)
siboldno	-0.02 (0.04)	0.00 (0.04)	-0.02 (0.04)	0.01 (0.06)
sibyngno	0.04 (0.04)	-0.03 (0.04)	0.13 (0.04)*	-0.02 (0.06)
mothfath	-0.18 (0.17)	0.06 (0.16)	0.00 (0.17)	-0.52 (0.23)*
mothwk	0.61 (0.11)*	0.16 (0.10)	0.34 (0.11)*	0.61 (0.16)*
fathwk	-0.03 (0.15)	-0.24 (0.14)	0.15 (0.15)	0.60 (0.22)*
mthq	0.35 (0.10)*	0.29 (0.10)*	0.02 (0.10)	0.22 (0.15)
fthq	0.24 (0.12)*	0.35 (0.12)*	0.23 (0.13)	-0.04 (0.18)
fathfig	0.05 (0.14)	-0.27 (0.13)*	0.17 (0.14)	0.29 (0.18)
zfathvar	-0.34 (0.15)*	-0.44 (0.14)*	0.08 (0.15)	-0.05 (0.21)
lowinc	-0.14 (0.12)	0.07 (0.11)	-0.33 (0.12)*	0.19 (0.16)
rural	0.82 (0.12)*	0.37 (0.12)*	0.11 (0.13)	0.89 (0.17)*
ill01	0.36 (0.10)*	0.26 (0.10)*	0.05 (0.11)	-0.01 (0.15)
parwork	0.21 (0.11)*	0.06 (0.10)	0.01 (0.11)	-0.02 (0.15)
ownocc	0.41 (0.11)*	0.42 (0.10)*	0.04 (0.11)	-0.02 (0.16)
fims	0.26 (0.11)*	0.04 (0.10)	0.32 (0.11)*	0.35 (0.16)*
fim2s	-0.08 (0.04)*	-0.02 (0.03)	-0.06 (0.04)	-0.18 (0.08)*
emaest	0.23 (0.17)	1.52 (0.19)*	0.13 (0.18)	-0.02 (0.21)
nonwhite	-0.40 (0.18)*	0.71 (0.15)*	-0.82 (0.19)*	-0.51 (0.28)
mathsscore	0.19 (0.04)*	0.16 (0.04)*	-0.04 (0.04)	0.05 (0.06)
engscore	0.44 (0.04)*	0.36 (0.04)*	0.00 (0.04)	0.03 (0.06)
zmathsscore	-0.63 (0.26)*	-0.47 (0.22)*	-0.55 (0.21)*	0.20 (0.31)
zengscore	0.55 (0.27)*	0.42 (0.23)	-0.48 (0.22)*	-0.80 (0.33)*
health	-0.26 (0.08)*	0.18 (0.08)*	-0.17 (0.08)*	-0.18 (0.10)
educatio	0.21 (0.06)*	0.14 (0.06)*	0.13 (0.07)	0.34 (0.08)*
housing	-0.01 (0.06)	-0.08 (0.06)	0.13 (0.07)	-0.10 (0.10)
access	-0.15 (0.07)*	0.11 (0.07)	0.15 (0.08)	-0.18 (0.11)
pcnotuni	-3.76 (0.27)*	-3.40 (0.29)*	-1.17 (0.18)*	-4.72 (0.42)*
pcnotsta	1.38 (0.24)*	0.59 (0.20)*	1.19 (0.22)*	2.18 (0.37)*
α_{j1}	1.0 (0)	6.94 (2.79)*	-0.033 (0.05)	3.20 (1.15)*

Standard errors in parentheses. Points of support η_1 = 0.41(0.17), η_2 = 0.09(0.04). * denotes significance at the 5% level or less.

Table 4.24: Parameter Estimates February/May 2000

	EJ	ENJ	FTW	PTW
Variables				
male	-0.04 (0.10)	0.42 (0.10)*	0.39 (0.10)*	-0.23 (0.13)
siboldno	-0.11 (0.05)*	-0.09 (0.04)*	0.03 (0.04)	-0.05 (0.06)
sibyngno	-0.03 (0.05)	-0.09 (0.05)	0.01 (0.05)	-0.15 (0.06)*
mothfath	-0.12 (0.18)	0.19 (0.17)	-0.37 (0.17)*	-0.36 (0.23)
mothwk	0.76 (0.11)*	0.24 (0.11)*	0.67 (0.12)*	0.37 (0.15)*
fathwk	0.08 (0.15)	-0.09 (0.15)	0.40 (0.15)*	0.51 (0.19)*
mthq	-0.06 (0.11)	-0.01 (0.11)	-0.25 (0.11)*	0.02 (0.14)
fthq	-0.01 (0.13)	-0.05 (0.13)	-0.09 (0.14)	-0.31 (0.18)
fathfig	0.17 (0.16)	-0.05 (0.16)	0.33 (0.16)*	-0.14 (0.19)
zfathvar	0.04 (0.15)	-0.17 (0.14)	0.16 (0.16)	-0.36 (0.20)
lowinc	-0.10 (0.13)	0.06 (0.13)	0.08 (0.13)	0.31 (0.17)
rural	0.33 (0.14)*	-0.01 (0.14)	-0.02 (0.16)	0.37 (0.19)*
ill01	-0.07 (0.11)	-0.07 (0.11)	0.02 (0.11)	-0.21 (0.15)
parwork	0.31 (0.12)*	0.07 (0.11)	0.22 (0.12)	0.03 (0.15)
ownocc	0.59 (0.12)*	0.37 (0.12)*	0.32 (0.12)*	0.03 (0.15)
fims	0.26 (0.13)*	0.24 (0.13)	0.50 (0.18)*	0.77 (0.23)*
fim2s	-0.07 (0.03)*	-0.09 (0.04)*	-0.13 (0.07)	-0.38 (0.10)*
emaest	0.00 (0.28)	1.49 (0.28)*	-0.65 (0.26)*	-0.23 (0.32)
nonwhite	-0.18 (0.21)	0.77 (0.20)*	-0.23 (0.22)	0.44 (0.27)
mathsscore	0.23 (0.04)*	0.20 (0.04)*	0.10 (0.05)*	-0.04 (0.06)
engscore	0.24 (0.05)*	0.16 (0.05)*	0.00 (0.05)	0.10 (0.07)
zmathsscore	0.33 (0.26)	0.59 (0.24)*	-0.16 (0.23)	-0.02 (0.31)
zengscore	-0.29 (0.28)	-0.06 (0.25)	-0.26 (0.25)	-0.17 (0.33)
educ lag	4.65 (0.18)*	4.48 (0.14)*	1.02 (0.13)*	1.40 (0.19)*
work lag	1.40 (0.20)*	0.91 (0.18)*	3.08 (0.13)*	2.47 (0.18)*
health	-0.31 (0.10)*	0.07 (0.10)	-0.24 (0.10)*	-0.29 (0.13)*
educatio	0.04 (0.08)	0.05 (0.08)	-0.24 (0.08)*	-0.07 (0.10)
housing	0.29 (0.08)*	0.21 (0.08)*	0.49 (0.09)*	0.27 (0.11)*
access	-0.04 (0.09)	0.16 (0.09)	-0.01 (0.09)	-0.07 (0.12)
pcnotuni	-5.19 (0.41)*	-4.81 (0.37)*	-2.72 (0.36)*	-4.22 (0.51)*
pcnotsta	1.23 (0.29)*	0.12 (0.25)	1.52 (0.36)*	2.51 (0.46)*
α_{j2}	-3.43 (1.43)*	2.26 (0.94)*	-2.12 (0.83)*	0.92 (0.38)*

Standard errors in parentheses. Points of support $\eta_1 = 0.41(0.17)$, $\eta_2 = 0.09(0.04)$. * denotes significance at the 5% level or less.

Table 4.25: Parameter Estimates September 2000

	EJ	ENJ	FTW	PTW
Variables				
male	-0.43 (0.13)*	0.20 (0.13)	0.06 (0.12)	-0.25 (0.17)
siboldno	-0.11 (0.06)	-0.05 (0.05)	-0.08 (0.05)	-0.06 (0.07)
sibyngno	-0.05 (0.06)	-0.14 (0.06)*	0.01 (0.05)	-0.06 (0.08)
mothfath	-0.11 (0.26)	0.39 (0.25)	0.23 (0.24)	0.04 (0.32)
mothwk	0.38 (0.15)*	-0.41 (0.16)*	0.35 (0.15)*	0.14 (0.20)
fathwk	0.16 (0.20)	0.06 (0.20)	0.05 (0.20)	0.38 (0.27)
mthq	0.41 (0.14)*	0.43 (0.14)*	0.27 (0.14)*	0.43 (0.18)*
fthq	0.02 (0.18)	-0.01 (0.18)	0.05 (0.18)	-0.07 (0.23)
fathfig	-0.09 (0.20)	-0.26 (0.21)	0.24 (0.19)	-0.05 (0.26)
zfathvar	-0.09 (0.23)	-0.01 (0.23)	0.30 (0.21)	0.11 (0.28)
lowinc	-0.36 (0.17)*	0.20 (0.17)	-0.34 (0.16)*	-0.11 (0.22)
rural	0.37 (0.17)*	-0.04 (0.18)	-0.04 (0.16)	0.74 (0.21)*
ill01	0.01 (0.15)	-0.10 (0.15)	0.04 (0.14)	-0.02 (0.19)
parwork	0.17 (0.17)	-0.12 (0.17)	0.16 (0.15)	-0.15 (0.21)
ownocc	0.52 (0.16)*	0.11 (0.16)	0.30 (0.15)*	0.05 (0.20)
fims	-0.29 (0.16)	-0.12 (0.15)	-0.04 (0.16)	-0.17 (0.18)
fim2s	0.10 (0.06)	0.01 (0.06)	0.05 (0.07)	0.09 (0.06)
emaest	-1.31 (0.34)*	1.73 (0.36)*	-0.81 (0.29)*	0.66 (0.35)
nonwhite	-0.48 (0.25)	0.78 (0.24)*	-0.57 (0.25)*	0.08 (0.33)
mathsscore	0.22 (0.06)*	0.18 (0.06)*	0.02 (0.05)	-0.21 (0.07)*
engscore	0.25 (0.06)*	0.27 (0.06)*	-0.10 (0.06)	0.11 (0.08)
zmathsscore	0.32 (0.39)	0.81 (0.35)*	-0.26 (0.30)	-0.48 (0.40)
zengscore	-0.41 (0.41)	0.33 (0.36)	-0.61 (0.32)	0.13 (0.42)
educ lag	4.52 (0.27)*	3.11 (0.19)*	1.25 (0.16)*	1.14 (0.20)*
work lag	1.79 (0.29)*	0.07 (0.24)	2.93 (0.17)*	1.94 (0.22)*
health	-0.16 (0.12)	0.42 (0.12)	-0.06 (0.11)	0.16 (0.14)
educatio	0.12 (0.09)	0.08 (0.09)	-0.07 (0.09)	-0.03 (0.11)
housing	0.14 (0.10)	0.00 (0.10)	0.17 (0.10)	-0.12 (0.13)
access	-0.15 (0.10)	0.06 (0.11)	0.04 (0.10)	-0.26 (0.13)*
pcnotuni	-2.70 (0.47)*	-5.31 (0.53)*	0.39 (0.28)	-2.70 (0.46)*
pcnotsta	-0.72 (0.40)	-1.41 (0.40)*	-0.76 (0.28)*	-0.39 (0.36)
α_{j3}	-6.65 (2.76)*	6.58 (2.72)*	-4.38 (1.78)*	1.72 (0.68)*

Standard errors in parentheses. Points of support $\eta_1 = 0.41(0.17)$, $\eta_2 = 0.09(0.04)$. * denotes significance at the 5% level or less.

Table 4.26: Parameter Estimates February/May 2001

	EJ	ENJ	FTW	PTW
Variables				
male	0.03 (0.13)	0.40 (0.13)*	0.33 (0.12)*	-0.42 (0.14)*
siboldno	-0.17 (0.06)*	-0.16 (0.06)*	-0.08 (0.05)	-0.07 (0.06)
sibyngno	0.10 (0.06)	0.00 (0.06)	0.17 (0.05)*	0.02 (0.07)
mothfath	0.04 (0.24)	0.53 (0.25)*	-0.20 (0.21)	-0.34 (0.25)
mothwk	0.64 (0.15)*	-0.07 (0.15)	0.53 (0.14)*	0.17 (0.17)
fathwk	0.17 (0.19)	-0.17 (0.19)	0.05 (0.17)	0.26 (0.22)
mthq	-0.03 (0.14)	0.06 (0.14)	-0.25 (0.13)*	-0.02 (0.16)
fthq	0.32 (0.17)	0.28 (0.17)	0.08 (0.15)	0.19 (0.19)
fathfig	-0.61 (0.19)*	-0.88 (0.20)*	-0.42 (0.17)*	-0.86 (0.22)*
zfathvar	-0.06 (0.21)	-0.15 (0.21)	-0.25 (0.20)	-0.54 (0.23)*
lowinc	0.26 (0.17)	0.17 (0.17)	-0.29 (0.16)	0.10 (0.19)
rural	0.87 (0.19)*	0.71 (0.19)*	0.46 (0.18)*	0.71 (0.21)*
ill01	0.18 (0.15)	0.02 (0.15)	-0.02 (0.13)	-0.14 (0.16)
parwork	0.17 (0.17)	0.10 (0.17)	0.40 (0.14)*	-0.02 (0.17)
ownocc	0.61 (0.16)*	0.26 (0.16)	0.10 (0.14)	0.11 (0.17)
fims	0.24 (0.15)	0.15 (0.16)	0.01 (0.17)	0.63 (0.24)*
fim2s	-0.07 (0.04)	-0.07 (0.04)	-0.06 (0.05)	-0.24 (0.10)*
emaest	-0.49 (0.32)	0.91 (0.35)*	-0.34 (0.32)	-0.03 (0.42)
nonwhite	-0.21 (0.27)	0.85 (0.27)*	-0.05 (0.25)	-0.03 (0.33)
mathsscore	0.17 (0.06)*	0.20 (0.06)*	0.05 (0.05)	-0.16 (0.06)*
engscore	0.40 (0.06)*	0.38 (0.07)*	0.16 (0.06)*	0.31 (0.07)*
zmathsscore	0.32 (0.39)	0.87 (0.37)*	-0.32 (0.30)	-0.22 (0.37)
zengscore	0.26 (0.41)	1.30 (0.39)*	0.02 (0.32)	0.54 (0.42)
educ lag	5.74 (0.22)*	6.39 (0.27)*	1.73 (0.17)*	2.22 (0.22)*
work lag	0.79 (0.26)*	1.76 (0.31)*	3.75 (0.14)*	2.93 (0.20)*
health	-0.34 (0.13)*	-0.05 (0.14)	-0.32 (0.12)*	-0.37 (0.15)*
educatio	0.15 (0.10)	0.13 (0.10)	0.00 (0.09)	0.21 (0.11)*
housing	0.14 (0.11)	0.15 (0.11)	0.19 (0.10)	-0.06 (0.13)
access	-0.22 (0.11)*	-0.02 (0.12)	-0.02 (0.11)	-0.19 (0.13)
pcnotuni	-9.18 (0.71)*	-10.80 (0.75)*	-3.39 (0.57)*	-4.86 (0.69)*
pcnotsta	4.91 (0.75)*	4.00 (0.76)*	2.49 (0.60)*	2.77 (0.66)*
α_{j4}	-2.63 (1.20)*	2.97 (1.16)*	-1.63 (0.71)*	0.78 (0.28)*

Standard errors in parentheses. Points of support η_1 = 0.41(0.17), η_2 = 0.09(0.04). * denotes significance at the 5% level or less.

Table 4.27: Parameter Estimates September 2001

	UJ	**UNJ**	**OEJ**	**OENJ**
male	-0.37 (0.16)*	-0.23 (0.14)	-0.12 (0.14)	0.40 (0.16)*
siboldno	-0.28 (0.08)*	-0.24 (0.07)*	-0.14 (0.07)*	-0.07 (0.07)
sibyngno	0.01 (0.08)	0.01 (0.07)	0.12 (0.07)	0.09 (0.07)
mothfath	-0.11 (0.34)	0.20 (0.31)	-0.21 (0.29)	-0.01 (0.31)
mothwk	0.28 (0.19)	0.21 (0.18)	0.44 (0.17)*	-0.02 (0.19)
fathwk	0.24 (0.25)	-0.13 (0.22)	0.13 (0.23)	-0.23 (0.24)
mthq	-0.09 (0.17)	-0.16 (0.16)	-0.18 (0.15)	-0.04 (0.17)
fthq	-0.30 (0.19)	0.02 (0.18)	-0.18 (0.18)	0.22 (0.20)
fathfig	-0.21 (0.28)	-0.14 (0.26)	0.12 (0.24)	-0.79 (0.27)*
zfathvar	-0.18 (0.31)	0.18 (0.27)	0.27 (0.27)	-0.44 (0.29)
lowinc	0.14 (0.23)	0.00 (0.21)	0.15 (0.20)	0.32 (0.21)
rural	-0.97 (0.21)*	-0.28 (0.18)	-0.33 (0.18)	0.27 (0.21)
ill01	0.24 (0.18)	-0.12 (0.17)	0.19 (0.16)	-0.07 (0.19)
parwork	0.43 (0.23)	0.66 (0.22)*	0.45 (0.20)*	0.63 (0.21)*
ownocc	0.75 (0.21)*	0.67 (0.19)*	0.58 (0.18)*	0.56 (0.19)*
fims	-0.38 (0.22)	-0.45 (0.20)*	0.03 (0.24)	-0.71 (0.22)*
fim2s	0.21 (0.08)*	0.19 (0.08)*	0.01 (0.11)	0.26 (0.08)*
emaest	0.74 (0.41)	0.50 (0.41)	-0.35 (0.37)	0.46 (0.42)
nonwhite	0.14 (0.31)	0.59 (0.27)*	-0.35 (0.29)	0.50 (0.28)
mathsscore	0.30 (0.07)*	0.49 (0.07)*	-0.12 (0.06)	-0.17 (0.07)*
engscore	0.49 (0.08)*	0.49 (0.08)*	-0.04 (0.07)	-0.14 (0.08)
zmathsscore	1.16 (0.68)	0.51 (0.70)	-0.69 (0.48)	0.45 (0.40)
zengscore	0.03 (0.78)	1.79 (0.66)*	-1.57 (0.50)*	-1.64 (0.45)*
educ lag	3.14 (0.39)*	3.14 (0.36)*	4.91 (0.45)*	2.68 (0.24)*
work lag	0.86 (0.42)*	0.16 (0.41)	2.40 (0.47)*	-0.30 (0.32)
health	0.03 (0.15)	0.32 (0.13)*	-0.08 (0.14)	0.13 (0.16)
educatio	0.21 (0.12)	0.10 (0.11)	0.23 (0.10)	0.23 (0.12)
housing	0.09 (0.13)	0.24 (0.12)*	0.05 (0.12)	0.08 (0.13)
access	0.07 (0.13)	0.12 (0.12)	0.04 (0.12)	-0.11 (0.13)
pcnotuni	-6.68 (0.67)*	-10.26 (0.75)*	-4.03 (0.63)*	-5.04 (0.64)*
pcnotsta	0.40 (0.41)	2.07 (0.55)*	0.82 (0.36)*	2.11 (0.57)*
α_{j5}	-5.07 (2.13)*	-1.59 (0.69)*	-4.53 (1.84)*	4.66 (2.04)*

Standard errors in parentheses. Points of support $\eta_1 = 0.41(0.17)$, $\eta_2 = 0.09(0.04)$. * denotes significance at the 5% level or less.

Table 4.26 contd.

	FTW	PTW
male	-0.04 (0.12)	-0.59 (0.14)*
siboldno	-0.12 (0.05)*	-0.14 (0.06)*
sibyngno	0.06 (0.06)	0.07 (0.07)
mothfath	-0.15 (0.23)	0.32 (0.30)
mothwk	0.50 (0.14)*	0.37 (0.17)*
fathwk	0.26 (0.18)	-0.05 (0.22)
mthq	-0.22 (0.13)	-0.08 (0.15)
fthq	-0.20 (0.15)	0.18 (0.19)
fathfig	0.11 (0.19)	-0.23 (0.24)
zfathvar	0.22 (0.21)	0.37 (0.26)
lowinc	0.00 (0.16)	0.13 (0.20)
rural	-0.22 (0.15)	0.11 (0.18)
ill01	0.04 (0.13)	-0.06 (0.17)
parwork	0.37 (0.15)*	0.48 (0.20)*
ownocc	0.52 (0.14)*	0.47 (0.18)*
fims	-0.10 (0.21)	-0.51 (0.21)*
fim2s	0.04 (0.09)	0.18 (0.09)*
emaest	0.11 (0.31)	0.16 (0.41)
nonwhite	-0.68 (0.24)*	-0.62 (0.31)*
mathsscore	0.00 (0.05)	-0.04 (0.06)
engscore	0.11 (0.06)	0.15 (0.07)*
zmathsscore	0.03 (0.31)	0.15 (0.43)
zengscore	-0.53 (0.33)	-0.49 (0.46)
educ lag	1.78 (0.19)*	1.48 (0.21)*
work lag	2.73 (0.18)*	1.22 (0.21)*
health	-0.27 (0.11)*	-0.04 (0.13)
educatio	-0.03 (0.09)	0.02 (0.11)
housing	0.19 (0.10)	0.09 (0.13)
access	0.04 (0.10)	-0.22 (0.12)
pcnotuni	-1.62 (0.41)*	-1.81 (0.46)*
pcnotsta	0.95 (0.39)*	-0.60 (0.35)
α_{j5}	-3.80 (1.58)*	-1.49 (0.61)*

Standard errors in parentheses. Points of support $\eta_1 = 0.41(0.17)$, $\eta_2 = 0.09(0.04)$. * denotes significance at the 5% level or less.

Chapter 5

Conclusion

In this thesis, I have examined a variety of aspects of education and work choices in both developing and developed economies. Each of the analyses has focused on different factors affecting such decisions, whilst controlling for as many other factors as possible that have in some sense been peripheral to the particular analysis. Despite considering choices across two distinct types of economic environment, the set of conditioning variables in each analysis widely overlapped. However, whilst some variables - such as gender, age and parental education - were universally important for choices, notable differences emerged across economies, reflecting fundamental differences between low- and high-income settings.

The importance of risk and uncertainty for human capital accumulation was assessed using measures of idiosyncratic and aggregate village-level risk based on a time series of wage data in Indonesia. Results pointed to an adverse effect of aggregate village-level risk on the accumulated years of schooling of children, with important implications for the perpetuation of persistent poverty in low-income countries due to the lack of adequate insurance markets.

The analysis of the relationship between sibling structure, the child wage and interactions thereof, and education and work choices, was examined in rural Mexico, a country in which large household sizes are observed and drop-out rates from school upon completion of primary education are high. This pointed to the importance of considering interactions between the number of siblings and the child wage, in analysing their effect on child activities. The results revealed larger changes in participation in activities in response to changes in the wage, the higher the number of siblings.

In the UK on the other hand, such factors as village risk and child labour were less relevant for observed choices. Instead, the analysis focused on the effectiveness of a conditional education subsidy on post-compulsory second-level education and work choices and subsequent university enrolment. The programme was found to have been effective in increasing post-compulsory second-level enrolment in education, specifically in education without a job, but had no subsequent effect on university attendance.

An important theme throughout the analyses of less developed economies, has been the importance of child labour and the subsequent effect on human capital accumulation. Our understanding of the underlying causes of child labour has been greatly enhanced in recent years, particularly due the increased availability of large scale micro data sets, which has greatly advanced research in this area. However, it is still very much an area of research in which there are many gaps in our knowledge, and as such is an area with enormous potential for development. Below, I outline what I consider to be important directions for future research.

To begin with, I refer again to a central theme of this thesis, which concerns the interactions between risk and education choices. The propensity of households to withdraw children from school in response to negative shocks to income, is informative on the existence of liquidity constraints in such economies. However, in chapter two, I found evidence of incomplete insurance and credit markets having important implications for education, even in the absence of actual shocks. Importantly, persistent (perceived and actual) risk and uncertainty at the village-level feeds through to education choices of households, suggesting that deep-rooted uncertainty and risk, which are difficult to observe (and thus to target), affect the accumulation of human capital of children through time. This reveals underlying imperfections in insurance and credit markets, in the absence of which households would not have to rely on children as a form of precautionary saving. This has potentially important consequences for the long-term welfare of children and the persistence of poverty in LDCs. However, this is only one such analysis of this kind, and I believe that future policy design may be greatly aided by further probing of the robustness of this finding across different countries. The analysis in chapter two was based on constructed measures of risk and uncertainty, which I believe to be both credible and intuitively appealing. However, there has been a relatively recent shift in tailoring surveys in LDCs towards more direct measurements of risk at the household level.[1] This will be useful insofar as it will minimise the assumptions that must be made in the measurement of risk and is an extremely tantalising area of future research, with significant implications for the welfare of children in LDCs.

In chapter three, I examined associations between education and work choices, and the sibling composition of the child, in rural Mexico. I also allowed for the effects of the child wage to vary by sibling composition, which provided evidence that competition amongst siblings for scarce resources within the household, may vary with the opportunity costs of schooling. An immediate extension of chapter three is to consider not only the number, but also the activities of other children in the household, and to examine how the child's sensitivity to costs is affected. The analysis highlighted an important and under-developed area of research, that concerning the intra-household decision-making process. It is likely that parents make choices for children simultaneously, and there are many reasons why the structure of siblings is relevant to such choices, particularly in the presence of binding liquidity constraints. Ability differences among siblings are likely to be important (but difficult to observe) and may lead to parents choosing education levels to either compensate or reinforce such differences across children.[2] Further,

[1]For example, in the Familias en Accion survey in Colombia, household heads are asked about their expectations of future income, which may offer some important insights into their perception of risk.

[2]See the innovative paper by Behrman et al (1994), which is one of the few that aims

the relationship between schooling investments and children's endowments depends on such factors as how parents value the children's distribution of wealth and earnings, the complementarity of endowments and schooling returns in the labour market and the total financial resources of parents. There is no conclusive evidence in existing empirical studies as to how the family allocates resources across children and this is an interesting and challenging area for future research. As discussed in chapter three, one way that it may be possible to gain some insight into this issue, is to exploit the design of certain existing policies in order to examine how parents choose to allocate resources across children. In particular, the Progresa subsidy in Mexico is payable to households with eligible children, up to some maximum amount. What this implies is that even if more children in the household are technically entitled to receive a subsidy, the upper threshold precludes them from obtaining it. This feature of the policy may be useful in order to examine whether there are differential implications for the education of some children in the household compared to others and, if so, to identify which children are adversely affected.

Much of the existing work on child labour, including the analyses in this thesis, focuses almost exclusively on choices at the extensive margin. There is very little research on the determinants of the actual choice of work hours. This is an important shortfall, as the amount of labour supplied is not only indicative of welfare consequences in terms of foregone leisure, but the effects of child labour on human capital accumulation are likely to vary greatly across the wide variation in work hours and work activities. Again, the improvements in data collection in LDCs in recent years will greatly aid in the development of research in this area.

The dynamic effects of child work on the long-term welfare of children has, to my knowledge, not yet been explored in the empirical literature. This is a challenging area of research. It would ideally require a long-term analysis in which individuals were followed from childhood to adulthood, and

to identify the response of human capital investment in siblings to endowment differences among siblings. See also the seminal work by Becker and Tomes (1976).

outcomes such as health, wages and other indicators of welfare were observed at various points. However, this is clearly a demanding data requirement, and even in developed economies such long-term longitudinal data sets are relatively few. However, various longitudinal data sets have been collected in LDCs, and indeed the IFLS data, which was analysed in chapter two, has follow-up surveys in 1997 and 2000. In future work I will use these waves of data to gain potentially important insights into the welfare of individuals up to 7 years later, and to examine human capital and other indicators of welfare for individuals who did and did not engage in work as children.[3]

There has been an increased focus of policymakers in recent years on the conditional subsidising of education.[4] We saw in chapter four that this is an effective policy in the UK, in terms of encouraging participation in education, for low-income individuals. Whilst there is also compelling evidence on the effectiveness of such policies for education in LDCs[5], very little is known about the implications for child labour. Indeed, increased school attendance in response to a decrease in the price of schooling, may simply reduce child leisure rather than child work time.[6] Future work will examine these issues in Colombia, a country in which such a policy is currently being implemented. In addition, our understanding of the important issues will also be greatly enhanced by formal structural modelling and rigorous econometric estimation of such models. This will be an important advancement up on existing reduced form research.

Finally, in terms of education and work choices in the UK, the focus of chapter four was on decisions from post-compulsory education onwards. However, education choices are closely driven by prior choices and experiences. In particular, the educational development of children from very early ages is an extremely important predictor of their educational attainment at

[3]In addition, this data will be useful in order to build upon chapter two, to further analyse measurements of risk and uncertainty using more waves of income data.

[4]There is evidence that unconditional transfers have only a very small effect on school enrolment and child labour. See Behrman and Knowles (1999).

[5]See Schultz (2001) for evidence on the post-primary Progresa programme in Mexico.

[6]See Ravallion and Wodon (2000).

later stages of the education cycle.[7] This points towards the importance of considering the entire education cycle of individuals, from when they first commence their formal education, in order to identify the key factors and influences in such choices throughout the cycle. By examining education choices from age 16 onwards, the problem is that we ignore a very relevant part of this cycle, which we only capture to some degree through controlling for the individual's acquired measure of ability, which is a function of past choices. An important future area of work is therefore to use UK data to consider the influence of socio-economic factors throughout the entire educational paths of individuals.

[7]See Cameron and Heckman (1998, 2001).

Bibliography

[1] ACEMOGLU, D. AND S. PISCHKE (2000), "Changes in the Wage Structure, Family Income and Children's Education", mimeo: MIT and LSE.

[2] AHN, T., J. KNODEL, D. LAM AND J. FREIDMAN (1998), "Family Size and Children's Education in Vietnam", *Demography,* 35(1): 57-70.

[3] ALTONJI, J, F.HAYASHI AND L.KOTLIKOFF (1997), "Parental Altruism and Inter Vivos Transfers: Theory and Evidence", *Journal of Political Economy,* 105(6): 1121-1166.

[4] AMEMIYA, T. (1981), "Qualitative Response Models, A Survey", *Journal of Economic Literature,* 19(4): 483-586.

[5] ARELLANO, M. AND B. HONORE (2002), "Panel Data Models: Some Recent Developments", *Handbook of Econometrics,* Vol.5 Chapter 53, North-Holland, 3229-3296.

[6] ASHWORTH, K., J. HARDMAN, Y. HARTFREE, S. MAGUIRE, S. MIDDLETON, D. SMITH, L. DEARDEN, C. EMMERSON, C. FRAYNE, C. MEGHIR (2003), "Education Maintenance Allowance: The First Two Years - A Quantitative Evaluation", Department for Education and Employment.

[7] ASRA, A. (1993), "Working Children in Bandung, Indonesia 1993", Jakarta, Central Bureau of Statistics and the International Labour Office.

[8] ASRA, A. (1996), "Working Children in Medan and Deli Serdang", Jakarta, Central Bureau of Statistics and the International Labour Office.

[9] ATTANASIO, O., P. GOLDBERG AND N.PAVCNIK (2003), "Trade Reforms and Wage Inequality in Colombia", *Journal of Development Economics*, forthcoming.

[10] BALAND, J. AND J. ROBINSON (2000), "Is Child Labor Inefficient?", *Journal of Political Economy,* 108(4): 663-679.

[11] BASU, K. AND P.H. VAN (1998), "The Economics of Child Labor", *American Economic Review,* 88(3): 412-427.

[12] BASU, K. (1999), "Child Labor: Cause, Consequence and Cure with Remarks on International Labor Standards", *Journal of Economic Literature,* 37(3): 1083-1119.

[13] BASU, K. (2001), "Gender and Say: A Model of Household Behavior with Endogenously-Determined Balance of Power", mimeo: Cornell University.

[14] BECKER, G. (1991), A Treatise on the Family, Enl. Ed. Cambridge, Mass.: Harvard University Press.

[15] BECKER, G. AND N. TOMES (1976), "Child Endowments and the Quantity and Quality of Children", *Journal of Political Economy,* 84(4): S143-162.

[16] BEEGLE, K., R. DEHEJIA AND R. GATTI (2003), "Child Labor, Income Shocks and Access to Credit", Development Research Group, The World Bank.

[17] BEHRMAN, J. AND P. TAUBMAN (1986), "Birth Order, Schooling and Earnings", *Journal of Labor Economics*, 4: S121-S145.

[18] BEHRMAN, J., M. ROSENZWEIG AND P. TAUBMAN (1994), "Endowments and the Allocation of Schooling in the Family and in the Marriage Market", *Journal of Political Economy,* 102 (6): 1131-1174.

[19] BEHRMAN, J. AND C. KNOWLES (1999), "Household Income and Child Schooling in Vietnam", *World Bank Economic Review,* 13(2): 211-256.

[20] BESLEY, T. (1995a), "Nonmarket Institutions for Credit and Risk Sharing in Low-Income Countries", *Journal of Economic Perspectives,* 9(3): 115-127.

[21] BESLEY, T. (1995b), "Savings, Credit and Insurance", Chapter 36 in Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*, Volume 3A.

[22] BHALOTRA, S. (2000), "Is Child Work Necessary?", mimeo: LSE.

[23] BLANDEN, J., P. GREGG AND S. MACHIN (2002), "Education and Family Income ", mimeo: CEE.

[24] BLUNDELL, R., L. DEARDEN, C. MEGHIR AND B.SIANESI (1999), "The Returns from Education and Training to the Individual, the Firm and the Economy: A Review of the Theoretical and Empirical Literature", Technical Report to the DfEE Skills Task Force, London: Institute for Fiscal Studies.

[25] CAIN, M. (1982), "Perspectives on Family and Fertility in Developing Countries", *Population Studies*, 36(2): 159-175.

[26] CAMERON, L. AND D. COBB-CLARK (2001), "Old-Age Support in Developing Countries: Labour Supply, Intergenerational Transfers and Living Arrangements", IZA Discussion Paper No.289.

[27] CAMERON, S. AND J. HECKMAN (1998), "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males", *Journal of Political Economy* 106(2): 262-333.

[28] CAMERON, S. AND J. HECKMAN (2001), "The Dynamics of Educational Attainment for Black, Hispanic and White Males", *Journal of Political Economy* 109(3): 455-499.

[29] CARNEIRO, P. AND J. HECKMAN. (2003), "The Evidence on Credit Constraints in Post Secondary Schooling", *Economic Journal* 108(4): 663-679.

[30] DEARDEN, L., C. EMMERSON, C. FRAYNE AND C. MEGHIR (2003), "The Impact of Financial Incentives on Education Choice", mimeo: IFS.

[31] DEATON, A. (1992), "Understanding Consumption", Oxford University Press.

[32] DEB, P. AND F. ROSATI (2002), "Determinants of Child Labor and School Attendance: The Role of Household Unobservables", mimeo: Hunter College Department of Economics.

[33] DURAISAMY, M. (2000), "Child Schooling and Child Work in India", Econometric Society World Congress Contributed Papers No. 0837.

[34] DUSTMAN, C., N. RAJAH AND A. VAN SOEST (1996a), "Part-time Work, School Success and School Leaving", mimeo: UCL.

[35] DUSTMAN, C., N. RAJAH AND S. SMITH (1996b), "Earning and Learning: Education Policy and the Growth of Part-time Work by Full-time Pupils", *Fiscal Studies,* 17(1): 79-103.

[36] EATON, J. AND H. ROSEN (1980), "Taxation, Human Capital and Uncertainty", *American Economic Review,* 70(4): 705-715.

[37] EDMONDS, E. (2002), "Understanding Sibling Differences in Child Labor", mimeo: Dartmouth College.

[38] EDMONDS, E. AND N. PAVCNIK (2002), "The Effect of Trade Liberalization on Child Labor", mimeo: Dartmouth College.

[39] EJRNAES, M. AND K. PÖRTNER (2002), "Birth Order and the Intrahousehold Allocation of Time and Education", mimeo: University of Copenhagen.

[40] EMERSON, P. AND A. SOUZA (2002), "Birth Order, Child Labor and School Attendance in Brazil", mimeo: Vanderbilt University.

[41] FITZGERALD, J., P. GOTTSCHALK AND R. MOFFIT (1998), "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics", *Journal of Human Resources* 33(2): 251-299.

[42] FREIJE, S. AND L. LOPEZ-CALVA (2001), "Child Labor, School Attendance, and Poverty in Mexico and Venezuela", mimeo: El Colegio de Mexico.

[43] GALASSO, E. (1999), "Intra-household Allocation and Child Labor in Indonesia", mimeo: Boston College.

[44] GARG, A. AND J. MORDUCH (1998), "Sibling Rivalry and the Gender Gap: Evidence from Child Health Outcomes in Ghana", *Journal of Population Economics*, 11(4): 471-493.

[45] GOMES, M. (1984), "Family Size and Educational Attainment in Kenya", *Population and Development Review*, 10: 647-60.

[46] GROOTAERT, C. (1999), "Child labor in Cote d'Ivoire", in C. Grootaert and H. Patrinos, eds., *The Policy Analysis of Child Labor: A Comparative Study*: World Bank, Washington.

[47] GROOTAERT, C. AND R. KANBUR (1995), "Child Labour: An Economic Perspective", *International Labour Review*, 135(2): 187-203.

[48] GUARCELLO, L., F. MEALLI AND F. ROSATI (2002), "Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing and Insurance", mimeo: Understanding Children's Work Agency.

[49] HAUSMAN AND WISE (1979), "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment", *Econometrica* 47(2): 455-474.

[50] HAVEMAN, R. AND B. WOLFE (1995), "The Determinants of Children's Attainments: A Review of Methods and Findings", *Journal of Economic Literature*, 33(4): 1829-1878.

[51] HAZARIKA, G. AND A. BEDI (2003), "Schooling Costs and Child Labour in Rural Pakistan", *Journal of Development Studies*, 39(5): 29-64.

[52] HECKMAN, J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1): 153-162.

[53] HECKMAN, J. (1981), "Statistical Models for Discrete Panel Data", in C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Panel Data with Econometric Applications*, Cambridge, Mass: MIT Press.

[54] HECKMAN, J. (1991), "Identifying the Hand of Past: Distinguishing State Dependence from Heterogeneity", *American Economic Review* Papers and Proceedings 81(2): 75-79.

[55] HECKMAN, J. AND B. SINGER (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Durtion Data", *Econometrica* 52(2): 271-320.

[56] HECKMAN, J. AND B. SINGER (1985), "Social Science Duration Analysis", in J. Heckman and B. Singer, eds., *Longitudinal Analysis of labor Market Data*, Cambridge: Cambridge University Press.

[57] HOLZMAN, R. AND S. JORGENSEN (2002), "Social Risk Management: A New Conceptual Framework for Social Protection and Beyond", *The World Bank.*

[58] HSIAO, C. (1986), "Analysis of Panel Data", Cambridge: Cambridge University Press.

[59] JACOBY, H.G. AND E. SKOUFIAS (1997), "Risk, Financial Markets and Human Capital in a Developing Country", *Review of Economic Studies*, 64(3): 311-335.

[60] JENSEN AND NIELSEN (1997), "Child Labour or School Attendance? Evidence from Zambia", *Journal of Population Economics*, 10(4): 407-424.

[61] KAESTNER, R. (1997), "Are Brothers Really Better? Sibling Sex Composition and Educational Achievement Revisited", *Journal of Human Resources*, 32(2): 250-284.

[62] KANE, T. (1994), "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education", *Journal of Political Economy* 102(5): 878-911.

[63] KEANE, AND K. WOLPIN (2001), "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment", *International Economic Review* 42(4): 1051-1103.

[64] KOCHAR, A. (1995), "Explaining Household Vulnerability to Idiosyncratic Income Shocks", *American Economic Review*, 85(2), Papers and Proceedings: 159-164.

[65] KODDE, D. (1986), "Uncertainty and the Demand for Education", *Review of Economics and Statistics*, 68(3): 460-467.

[66] LAIRD, N. (1978), "Nonparametric Maximum Likelihood Estimation of a Mizing Distribution", *Journal of the American Statistical Association*, 73: 805-811.

[67] LAVY, V. (1996), "School Supply Constraints and Children's Educational Outcomes in Rural Ghana", *Journal of Development Economics*, 51(2): 291-313.

[68] LELAND, H. (1968), "Saving and Uncertainty: the Precautionary Demand for Saving", *Quarterly Journal of Economics*, 82(3): 465-473.

[69] LEVHARI, D. AND Y. WEISS (1974), "The Effect of Risk on the Investment in Human Capital", *American Economic Review*, 64(6): 950-963.

[70] MANNING, C. (2000), "The Economic Crisis and Child Labour in Indonesia", mimeo: ILO.

[71] MCFADDEN, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior" in P. Zarembka, ed., Frontiers in Econometrics, New York, Academic Press, 105-142.

[72] MORDUCH, J. (1995), "Income Smoothing and Consumption Smoothing", *Journal of Economic Perspectives* 9(3): 103-114.

[73] PARISH, W. AND R. WILLIS (1993), "Daughters, Education and Family Budgets: Taiwan Experiences", *Journal of Human Resources*, 28(4): 863-898.

[74] PATRINOS, H.A. AND G. PSACHAROPOULOS (1997), "Family Size, Schooling and Labour in Peru: An Empirical Analysis", *Journal of Population Economics*, 10(4): 387-405.

[75] PAXSON, C. (1992), "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand", *American Economic Review*, 82(1): 15-33.

[76] PITT, M. AND M. ROSENZWEIG (1990), "Estimating the Intrahousehold Incidence of Illness: Child Health and Gender Inequality in the Allocation of Time", *International Economic Review*, 31(4): 969-989.

[77] PÖRTNER, C. (2001), "Children as Insurance", *Journal of Population Economics*, 14(1): 119-136.

[78] PSACHAROPOULOS, G. (1985), "Returns to Education: A Further International Update and Implications", *Journal of Human Resources*, 20(2): 583-604.

[79] PSACHAROPOULOS, G. (1994), "Returns to Investment in Education: A Global Update", *World Development*, 22(9): 1325-1343.

[80] RANJAN, P. (1999), "An Economic Analysis of Child Labor", *Economics Letters*, 64(1): 99-105.

[81] RANJAN, P. (2001), "Credit Constraints and the Phenomenon of Child Labor", *Journal of Development Economics*, 64(1): 81-102.

[82] RAUT, L. AND L. TRAN (1998), "Motives for Investment in Human Capital of Children: Evidence from Indonesian Family Life Survey Data", mimeo: Economics Working Paper Archive at WUSTL.

[83] RAVALLION, M. AND Q. WODON (2000), "Does Child Labor Displace Schooling? Evidence on Behavioral Responses to an Enrolment Subsidy", *Economic Journal*, 110: C158-C176.

[84] RAY, R. (2000), "Child Labor, Child Schooling, and their Interaction with Adult Labor: Empirical Evidence for Peru and Pakistan", *World Bank Economic Review*, 14(2): 347-67.

[85] ROSENZWEIG, M. (1982), "Educational Subsidy, Agricultural Development and Fertility Change", *Quarterly Journal of Economics*, 97(1): 67-88.

[86] ROSENZWEIG, M. (1988), "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries", *Economic Journal*, 98: 1148-1170.

[87] ROSENZWEIG, M. AND H. BINSWANGER (1993), "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments", *Economic Journal* 103: 56-78.

[88] ROSENZWEIG, M. AND R. EVENSON (1977), "Fertility, Schooling and the Economic Contribution of Children in Rural India: An Econometric Analysis", *Econometrica* 45(5): 1,065-1,079.

[89] ROSENZWEIG, M. AND O. STARK (1989), "Consumption Smoothing, Migration and Marriage: Evidence from Rural India", *Journal of Political Economy*, 97(4): 905-926.

[90] ROSENZWEIG, M. AND K. WOLPIN (1980), "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment", *Econometrica*, 48(1): 227-240.

[91] SANDMO, A. (1970), "The Effect of Uncertainty on Saving Decisions", *Review of Economic Studies*, 37(3): 353-360.

[92] SAWADA, Y. AND M. LOKSHIN (2001), "Household Schooling Decisions in Rural Pakistan", mimeo, World Bank, Country Economics Department.

[93] SCHULTZ, T.P. (1993), "Returns to Women's Education", in E.M. King and M. Hill, eds., *Women's Education in Developing Countries: Barriers, Benefits and Policies*, Baltimore: John Hopkins University Press.

[94] SCHULTZ, T. (2001), "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program", mimeo: Yale University Economic Growth Centre.

[95] SIANESI, B. AND J. VAN REENEN (2003), "The Returns to Education: Macroeconomics", *Journal of Economic Surveys* 17(2): 157-200.

[96] SINGH, K. (1998), "Part-time Employment in High School", *Journal of Educational Research*, 91(3): 131-139.

[97] STRAUSS, J. AND D. THOMAS (1995), "Human Resources: Empirical Modelling of Household and Family Decisions", Chapter 34 in Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*, Volume 3A.

[98] TOWNSEND, R. (1994), "Risk and Insurance in Village India", *Econometrica*, 62(3): 539-591.

[99] TRAIN, K. (1986), Qualitative Choice Analysis: Theory Econometrics, and an Application to Automobile Demand, Cambridge: The MIT Press.

[100] UDRY, C. (1994), "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria", *Review of Economic Studies* 61(3): 495-526.

[101] VAN DER BERG, G. AND M. LINDEBOOM (1998), "Attrition in Panel Survey Data and the Estimation of Multi-State Labor Market Models", *Journal of Human Resources* 33(2): 458-478.

[102] WEALE, M. (1992), "Externalities from Education", in G. Johnes and E. Cohn, eds., *Recent Developments in the Economics of Education*, Basingstoke: Macmillan.

[103] WHITE PAPER (2003), "The Future of Higher Education", Department for Education and Skills.

[104] WOOLDRIDGE, J. (2002), "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity", *CeMMAP Working Paper 18*, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

[105] World Bank (2001), "Social Protection Sector Strategy - From Safety Nets to Springboard", *The World Bank, Washington D.C.*.

[106] Ziliak, J. and T. Kniesner (1998), "The Importance of Sample Attrition in Life Cycle Labor Supply Estimation", *Journal of Human Resources* 33(2): 507-530.