ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 679 November 2009

The Effects of Risk and Shocks on School Progression in Rural Indonesia

Christopher M Gilbert
Francesca Modena

"The Effects of Risk and Shocks on School Progression in Rural Indonesia"*

Christopher L. Gilbert and Francesca Modena+

Abstract

Many empirical and theoretical studies explore the effects of ex-ante risk and ex-post shocks on child education. While scholars share the opinion that shocks reduce investment in education, there is no general agreement over the effects of uncertainty on child schooling. This work uses the Indonesian Family Life Survey to explore the effects of ex-ante risk and ex-post shocks on school progression in rural Indonesia. We develop a model of household school transition decisions from elementary to junior education and from junior to senior school considering different sources of uncertainty related both to parental and adult income, and under the assumption that withdrawal from school is permanent. In this way, temporary interruptions in child schooling have long term impacts on the child human capital. We show that there is no simple answer to the question of how uncertainty affects schooling decisions. Econometric results suggest that uncertainty about parental income for the time the child may be potentially at school increases the probability of attending junior school while uncertainty about expected earnings from education has a negative and significant effect only for senior school attendance. Finally, positive (negative) income shocks increase (decrease) the probability of attending junior school.

Keywords: Education choice; shocks; uncertainty; dynamic optimization

JEL Classification: J24; C61; D81; O12

---

* We gratefully acknowledge Kaushik Basu, Federico Perali, Maria Luigia Segnana and Marco Francesconi. This paper is based on work carried out during a visit to the Department of Economics at Essex and to the European Centre for Analysis in the Social Sciences (ECASS) at the Institute for Social and Economic Research, University of Essex supported by the Access to Research Infrastructures action under the EU Improving Human Potential Programme.

+ Department of Economics and IDEC, University of Trento. E-mail: christopher.gilbert@unitn.it; francesca.modena_2@unitn.it.

1. INTRODUCTION AND LITERATURE REVIEW

In developing countries where incomes are volatile and on average low and financial and insurance markets are incomplete (Townsend, 1994; Morduch, 1990), shocks, whether ex ante risk or ex post, may have a large impact on child labour and child education. Many papers support the idea that, in poor countries, child time allocation may be used as a strategy to both insure ex ante against risk and to cope ex post with negative hardships in the face of incomplete insurance markets and credit constraints (Jacoby and Skoufias, 1997; Beegle *et al*., 2006; Baland and Robinson, 2000; Ranjan, 2001). Children may be used as a buffer against shocks, or they may be sent to work to earn money and support the family. Children may work as part-time family workers, helping their parents at home and in the fields. If necessary, children will drop out of school to save the costs of schooling and contribute additional labour.

Many empirical and theoretical studies analyze the extent to which child time allocation provides a means to cope ex post with shocks. Most of the literature finds evidence that transitory income shocks reduce child schooling and increase child labour (Sawada and Lokshin, 2001; Sawada, 2003; Jensen (2000); Thomas *et al*., 2004; Duryea *et al*., 2007, Gubert and Robilliard, 2008). Some works underline the role played by incomplete financial markets and credit rationing in determining child schooling decisions in the face of income shocks (Jacoby and Skoufias, 1997; Guarcello et al., 2003; Beegle, Dehejia, and Gatti, 2003 and 2006[1]).

Another strand of literature examines how ex ante risk affects education decisions. While many studies share the opinion that shocks reduce investment in education, there is no general agreement over the effects of uncertainty on child schooling. Kazianga (2005) shows that child time may be used as an ex ante measure to diversify risk in an uncertain environment. He finds that the less well developed are other formal or informal insurance mechanisms, the more is child labour used. Maitra *et al*. (2006) show that, in India, children serve as a buffer to deal with the uncertainty created by the labour market: they are neither sent to work nor sent to school, but instead kept idle. Fitzsimons (2007) finds evidence that, in Indonesia,

[1] Beegle, Dehejia, and Gatti (2003, 2006) find that in Tanzania households with durable assets can mitigate the negative effects of crop shocks on child labour.

aggregate village components of risk negatively affect children education[2]. Jalan and Ravallion (2001) find that risk increases child schooling in rural China. Guarcello, Rosati and Scaramozzino (2008) show that in Brazil increased uncertainty about labour market outcomes is associated with higher levels of schooling, consistent with a real options approach to education as an investment. Similarly, Hogan and Walker (2007) find that people spend more time in school when returns to education are uncertain.

The use of child time as a risk management strategy, both ex ante and ex post, may have long run consequences. Children who are withdrawn from school may not be able to restart school or recover the educational gap: in this way temporary schooling interruptions have lasting impacts. Guarcello *et al*. (2003) and Duryea *et al*. (2007) point out the potential long-term consequences of short term shocks for children's human capital. However, few papers propose a theoretical treatment of child schooling choices in the presence of risk taking into account the irreversibility or state dependence of school attendance. De Janvry et al. (2006) develop a model of child schooling decisions in the face of income shocks with a re-entry cost as an additional cost of schooling. This allows them to consider the dependence of child enrollment on previous schooling decisions[3].

We develop a model of parental decisions with regard to child schooling considering withdrawal from school as an absorbing state, that is children cannot re-enroll once they stop going to school. In this way, temporary interruptions in child schooling have long term impacts on the child human capital. This paper analyzes the role that both ex ante risk and ex post shocks play in determining child education, differentiating by school levels. We consider different sources of uncertainty related both to parental and adult income.

Given irreversibility of withdrawal from school, we find that in the face of parental income variability over the time the child may potentially be at school

[2] Whereas households are able to insure themselves against idiosyncratic risk without reducing child schooling.

[3] Using Mexican data, De Janvry et al. (2006) find evidence that shocks force children to drop out of school with long run effects on children's human capital given the state dependence of school attendance. They find that this state dependence is higher in secondary than in primary school, suggesting that school attendance is more flexible for children at the elementary than at the secondary level.

parents are more likely to send children to elementary school in the current period to give them the option to continue with higher schooling levels in the future (and hence earn higher earnings when they become adults). A positive/negative shock increases/decreases the probability of attending school for junior education.

As regard the effect of uncertainty of returns to education, conditional on having completed a given school level, the paper suggests that income risk reduces the incentive to enrol at senior school.

This paper analyzes the progression decisions from one education level to another (in particular, from elementary to junior secondary school, and from junior to senior school) using the Indonesian Family Life Survey. The focus on school transitions has been suggested by the fact that returns to education largely result from achieving a certain qualification, e.g. high school graduation, rather than from the actual time passed in schooling. Moreover, there are high drop out rates in Indonesia after graduation, and in particular after the completion of the primary level.

The paper is organized as follows. The theoretical model is presented in section II. Section III presents some evidence on child schooling in Indonesia and descriptive analyses of the data. Section IV discusses the empirical methodology and section V summarizes the results. Section VI concludes.

2. THE THEORETICAL MODEL

We set out a model which allows exploration of the effects of income uncertainty on school transition decisions. We focus on transition because the returns to education largely result from achieving a certain qualification, e.g. high school graduation, rather than from the actual time passed in schooling. The result is that, although in a small proportion of instances, schooling is terminated between transitions, most terminations occur at transition points (as shown in the next paragraph). We consider two sequential transitions – from elementary education (where attendance is compulsory) to junior education and from junior to senior education.

Irreversibility is an important feature of the education process – once a child has quit school, it is difficult to resume. We simplify by assuming complete irreversibility so that quitting is permanent. This has the implication that, in deciding whether to enrol a child in junior school, a family must take into the account that termination of the child's education at this initial transition point will preclude the possibility of his/her attending secondary school. The family therefore needs to anticipate the second transition decision at the time they make the decision with regard to the initial transition. Effectively, the family must solve a dynamic programming problem.

The other important feature of our model is that we see education as allowing children to realize their latent ability. We model the expected returns to senior education as depending on their ability (where ability is to be interpreted as earning ability). We suppose that families observe the earning ability of their children and take schooling decisions on the basis of this knowledge. For the economist, ability is unobserved and all children appear equal. By contrast, the child's earnings if s/he quits prior to senior school will not reflect ability. This feature of our model reinforces the need for anticipation at the initial transition since a high ability child may benefit rather little, in earnings terms, from attending junior school but needs to make this investment to be in a position to attend senior school.

Our model takes the family as the decision unit and supposes that all sources of income are aggregated within the family. Importantly, the child remains within the family when adult. In this framework, a principal benefit of education to the family is that of income diversification. If instead we were to take the opposite, individualistic, polar case of supposing that the child, once adult, becomes a separate unit with its own utility function, and that there are no inter-generational transfers, this benefit would disappear. We take the view that, in many developing countries the family perspective is more realistic than the individualistic alternative.

The analysis is based on a set of incomes and schooling costs set out in Table 1.

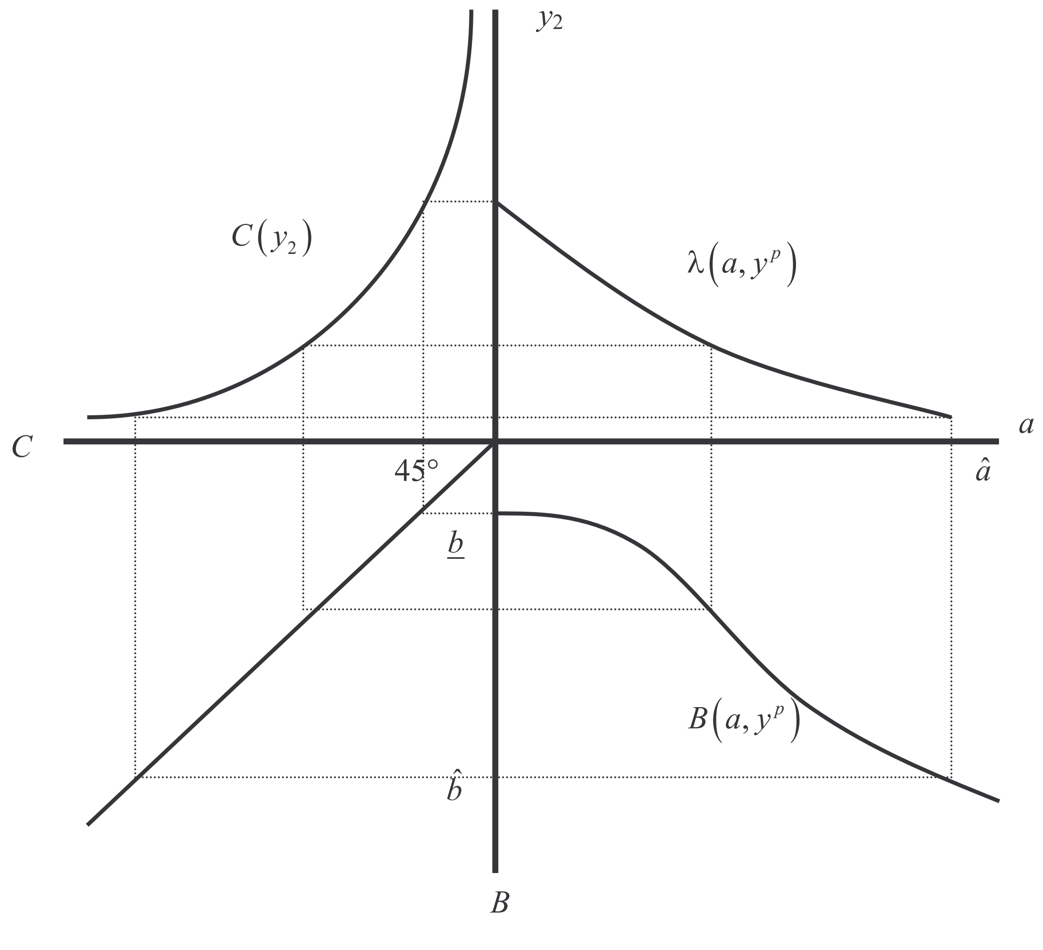
Table 1
Incomes and Schooling Costs

	Period	No post-elementary school education	Elementary + junior school	Elementary + junior + senior school
Parents' income	1		$y_1 = y^p + v_1$	
	2		$y_2 = y^p + v_2$	
	3		$y_3 = y^p + v_3$	
Child's income	1	w_1	0	0
	2	w_2	$x_2 = w_2 + \xi_2$	0
	3	w_3	$x_3 = w_3 + \xi_3$	$z_3 = w_3 + \xi_3 + \zeta_3 + a$
School cost	1	0	p	p
	2	0	0	q

We take permanent income y^p as constant and known by the family. We suppose that the transitory components of parental incomes, v_2 and v_3 are unknown until the start of the period, have zero expected value and are independent of schooling decisions. We suppose that additional schooling cannot reduce the child's future earnings, i.e. $\xi_2, \xi_3, \zeta_3 \geq 0$. For analytic simplicity, we suppose ξ_2 to be known at the start of period 1, but that the incremental adult incomes ξ_3 and ζ_3 are not revealed until period 3. We regard education to senior school level as allowing children to capture a latent earning ability a, which we regard as observed by the family but not the economist. For simplicity, we take school costs as known. For analytic tractability, we take w_1, w_2, w_3, ξ_2, p and q to be constant across the population.

Although we model the financial costs of schooling as independent of income level, the utility costs are declining in income – a poor family needs to make a greater sacrifice to maintain a child in school than does a comparable but richer family. This observation motivates definition of two functions which play a pivotal role in the analysis. The function $\mu\left(a, y^p\right)$ defines the threshold level of the parents' period 1 income y_1 such that, for a child of earning ability a and from a family with permanent income y^p, the child is enrolled in junior school for incomes $y_1 \geq \mu\left(a, y^p\right)$ and not enrolled for lower incomes. Similarly the function $\lambda\left(a, y^p\right)$ defines the threshold level of the parents' period 2 income y_2 such that, for a child of earning ability a and from a family with permanent income y^p and conditional on the child having attended junior school, the child is enrolled in senior school for incomes $y_2 \geq \lambda\left(a, y^p\right)$ and not enrolled for lower incomes.

Figure 1: The function λ(*a,y*ᵖ)



The functions $\mu\left(a, y^p\right)$ and $\lambda\left(a, y^p\right)$ may be seen as the outcome of balancing the costs of education, which depend on current parental income, with the benefits, which depend on the child's earning ability. This is illustrated in Figure 1 for the function $\lambda\left(a, y^p\right)$ for a particular value of y^p. The north-west quadrant of the figure shows the cost function $C\left(y_2\right)$. Costs are falling in income so $C'\left(y_2\right) < 0$ and, under assumptions which are made explicit in the appendix A, $C''\left(y_2\right) > 0$ so that the function is convex. The south-east quadrant of the figure shows the benefit function $B\left(a, y^{p}\right)$ which is rising in ability, $\frac{\partial B\left(a, y^{p}\right)}{\partial a} > 0$. The 45° line in the south-west quadrant gives the equality $B\left(a, y^{p}\right) = C\left(y_{2}\right)$ mapping out the function $y_{2} = \lambda\left(a, y^{p}\right)$ in the north-east quadrant. A similar diagram is available for the function $\mu\left(a, y^{p}\right)$.

Because period 2 income is not known at the time of the period 1 junior school enrolment decision, and because the underlying distribution is that of transitory income v_2, the analysis of that decision proceeds in terms of the function $\eta\left(a, y^{p}\right) = \lambda\left(a, y^{p}\right) - y^{p}$ which defines the senior school enrolment threshold and in terms of the family's period 2 transitory income v_2. The child is enrolled in senior school for incomes $v_{2} \geq \eta\left(a, y^{p}\right)$ and not enrolled for lower incomes.

Given these functions, the proportion of children enrolling at each school level is obtained as the complement of the appropriate income distribution integrated over ability and permanent income. Within this structure, the effects of income uncertainty are more complicated than has been previously appreciated. At the time of the initial transition (elementary to junior school) there are four possible sources of income uncertainty.

a) uncertainty about the parents' transitory income v_2 at the time the child could be attending senior school;

b) uncertainty about the child's incremental income ξ_{3} as adult, conditional on non-attendance of senior school;

c) uncertainty about the child's incremental income ζ_{3} as adult, conditional on attendance of senior school; and

d) uncertainty about the parents' transitory income v_3 once the child has become adult.

At the time of the second transition (junior to senior school) the first source of uncertainty has been resolved, but the final three remain. It follows that, in this model, there are seven potential effects of uncertainty on schooling.

In what follows, we analyze the impact of increases in uncertainty in each of the five income variables through the mean-preserving spread technique. The signs of these effects depend crucially on the distributional assumptions. We analyze the simplest possible case in which the distributions of all five income variables are independent. In practice, these variables are likely to be correlated to varying degrees. While acknowledgement of this possibility would increase realism, it would also complicate the analysis of the separate routes by which uncertainty affects school transition decisions, which is the objective of this study. The results of these analyses are tabulated in Table 2. Appendix A presents a detailed description of the theoretical model and derives the results.

Table 2
Predicted Effects of Uncertainty on School Enrolment

	Y_2	ξ_3	ζ_3	y_3
Progression to junior school	?	?	–	+
Progression to senior school	n.a.	+	–	+

The effects summarized in Table 2 are all behavioural in the sense that uncertainty with respect to future incomes affects the educational choices that families make. These effects should be distinguished from the aggregation effects of income variability. If period 1 income variability changes, a larger or smaller proportion of families may find that their period 1 incomes exceed the school enrolment threshold $\mu\left(a, y^{p}\right)$. Similarly, if period 2 income variability changes, a larger or smaller proportion of families may find that their period 2 incomes exceed the school enrolment threshold $\lambda\left(a, y^{p}\right)$. These are aggregation, not behavioural effects – a particular family's educational choice decisions are unaffected by the situations in which other families find themselves. It is not possible to sign these aggregation effects *a priori*.

To obtain estimable functions, we need to make linear approximations to the model. This approximation considers only first order effects, and eliminate those effects that rely on assumptions about the second and third derivatives of the utility function.

Making these assumptions, the equation for senior school shows that the child will be enrolled if (see Appendix A):

$$\delta a_h \geq \left(x_{h2}+q_h\right)\frac{1-\rho\dfrac{y_{h2}-\overline{y}_2}{\overline{y}_2}}{1-\rho\dfrac{y_h^p-\overline{y}^p}{\overline{y}^p}}-\delta\overline{\zeta}_{h3}+\tfrac{1}{2}\delta\rho y_h^p\sigma_{h\zeta}^2$$

$$\simeq\left(1-\rho\frac{v_{h2}}{\overline{y}_2}\right)\left(x_{h2}+q_h\right)-\delta\overline{\zeta}_{h3}+\tfrac{1}{2}\delta\rho y_h^p\sigma_{h\zeta}^2=\delta\hat{a}_h$$

Similarly, the child will attend junior school if (see Appendix A):

$$\gamma\delta a_h \geq \gamma\left(1-k\rho\sigma_{h2}\right)\left(q_h+w_{h2}+\xi_{h2}\right)-\gamma\delta\left(\overline{\zeta}_{h3}-\tfrac{1}{2}\rho y_h^p\sigma_{h\zeta}^2\right)$$
$$-\frac{1}{1-\Phi_h}\left[\gamma\left(\xi_{h2}+\delta\left(\overline{\xi}_{h3}-\tfrac{1}{2}\rho y_h^p\sigma_{h\xi}^2\right)\right)+\left(1-\rho\frac{v_{h1}}{\overline{y}_1}\right)\left(w_{h1}+p_h\right)\right]$$

According to these equations, period two and one income shocks (v_{h2} and v_{h1}) (interacted with the cost of schooling) increase the probability of attending senior and junior school respectively. As regard the effects of the different sources of risk on schooling decisions, they are the following:

Table 3
Predicted Effects of Uncertainty on School Enrolment – Linear approximation

	y_2	ξ_3	ζ_3
Progression to junior school	+ (interacted with the total cost of senior school)	+ (interacted with permanent income)	– (interacted with permanent income)
Progression to senior school	n.a.		– (interacted with permanent income)

3. DATA - EVIDENCE ON CHILD SCHOOLING IN INDONESIA

We exploit the role of risk on school progression in rural Indonesia using the Indonesian Family Life Survey (mainly the 1993 round)[4]. A large amount of information on household economic conditions (income sources, consumption, types of assets), and household and individual characteristics were collected in this survey, as well as education data for all household members.

As shown in Table 4, the Indonesian schooling system in 1993 consisted of four levels: elementary (compulsory), junior secondary school, senior secondary school and higher education. The public sector provided the greater part of the education services, and compulsory education was free. However, there was a strong private education sector as well.

Table 4
The Indonesian schooling system in 1993

	Year	*Age*
Elementary school (compulsory from 1984)	1-6	7-12[a]
Junior secondary school	7-9	13-15
Senior secondary school	10-12	16-18
Higher education	>12	>18

The table sets out the structure of the Indonesian schooling system.
Source: American Institutes for Research (2002).
[a] 6 years olds can also be admitted

Since the early 1980's Indonesian government has made a major effort to achieve the goal of universal elementary education, improving both the quantity and quality of education. These policies were largely successful in improving primary educational levels5, but the continuation rate from primary to junior secondary school remains low. This have been seen as causes of adult illiteracy and reliance on child labour (Manning, 2000)[6].

[4] Individual sampling weights are used for the descriptive analysis.
[5] IFLS data show that in 1993 the net enrolment ratio (enrolled children in the age group 7-12/ total # of children in the age group 7-12) for elementary school is 87%, while the gross enrolment ratio (enrolled children/total # of children in the age group 7-12) is 97%.
[6] A new policy was thus required to achieve universal education, and in 1994 compulsory basic education was increased to nine years: six years of primary education (for ages 7-12) and three years

To better describe this issue, Table 5 shows the distribution of out-of-school children, who have at some time been to school, according to the number of schooling years completed. We focus on the sample of children and young adults aged 7-17, since most children start school at age 7, and only a few individuals are at school after age 17. More than half of out-of-school children aged 7-17 years, and who have at some time been to school, completed 6 years of schooling, i.e. they left school after graduation from primary level.

School progression rates at different educational stage can be further analyzed estimating the conditional survival function. Here we follow the procedure of Sawada and Lokshin (2007). Let n_k denote the number of children who have completed education at least at stage k-1 (the set is not right-censored at education level k-1). Among these n_k students, define as h_k the number of children who have completed education at least at level k, so that $h_k = n_{k+1}$. Then h_k/n_k is the fraction of students who progressed to a higher stage of education, conditional on having completed stage k-1; this is an estimate of the conditional survival probability at education level k. Results suggest that in 1993 the probability of continuation after the completion of the elementary school in rural Indonesia was 65%[7], and the probability of continuation to senior secondary school was 73%.

These results show that the largest loss occur after the graduation from the elementary level, and justify the focus on the schooling decisions made at the transition points, and in particular on school progression at the point of the graduation from the primary level.

of junior secondary education (ages 13-15). As a result of the 1994 law, continuation rates from primary to junior secondary rose to nearly 72% in 1999-2000, and the net enrolment ratio at junior secondary school increased from 42% in 1992 to 62% in 2002, although with significant differences between rural and urban areas (lower in rural areas) and between poor and rich households.

[7] Manning (2000) and UNDP (2004) reported a continuation rate close to 62% in 1993/94.

Table 5
Distribution of out-of-school children according to schooling years completed and gender

School level	*Number of years completed*	*Male (%)*	*Female (%)*	*Male and female (%)*
Elementary – not completed	<6	28.2	26.9	27.6
Elementary – completed	6	58.9	62.1	60.5
Junior secondary – not completed	7-8	4.3	2.3	3.4
Junior secondary - completed	9	7.9	8.2	8.0
Senior secondary	> 9	0.7	0.5	0.6
TOT		*100*	*100*	*100*

The table reports the distribution of out-of-school children, who have at some time been to school, according to number of schooling years completed. Percentages are computed using individual sampling weights and on the sub-sample of children 7-17 years old that are not at school but have at some time been to school

4. EMPIRICAL STRATEGY

The theoretical analysis discussed in section 2 shows that income uncertainty is a key determinant of the schooling decisions. There are several different possible sources of uncertainty each with different effects on enrolment choices. For this reason, there cannot be a simple answer to the question of how uncertainty affects child education. We reflect the conclusions of the theoretical model by constructing an empirical model which includes the ex ante perceived variance both of parental incomes while the child remains at school and of the child's income as adult, and also takes into account income shocks (i.e. ex post outcomes).

An important feature of our model is that education allows children to realize their latent ability. Ability is unobserved by the investigator, and therefore household decisions become stochastic from the point of view of the economist.

4.1. Measures of income shocks and uncertainty

Income shocks

An income shock at time *t* is the difference between the actual household income y_{ht} at time *t* and the household's permanent income y_{ht}^{P} defined as expected household income, given all the information available at time *t*: $v_{ht} = y_{ht} - y_{ht}^{P}$,

Permanent income is inferred from an estimated equation (Cameron and Worswick, 2003)

$$y_{ht} = \alpha' X_{ht} + \varepsilon_{ht}, \quad (1)$$

where X_{ht} is a vector of household characteristics at time *t*. Results are reported in table B.1 (Appendix B).

Uncertainty of parental period two income

Uncertainty about parents' transitory income ν_{h2} is measured as variance of period 2 income conditional on all the information available at time 1. A two stage procedure is used. We first regress future household income on current household characteristics:

$$y_{h2} = \alpha' X_{h1} + \varepsilon_{h2} \quad (2)$$

where *t*=2 corresponds to 1997 and *t*=1 to 1993. The four year gap is appropriate in that junior secondary school lasts three years. At the second stage, we regress the logarithm of the squared predicted residuals ($\ln \hat{\varepsilon}_h^2$) on a constant and all first moments, second moments and cross products of the original regressors:

$$\ln \hat{\varepsilon}_h^2 = \delta_0 + \sum_{i=1}^{k} \sum_{j}^{k} \delta_{ij} X_{ih} + \upsilon_h \quad (3)$$

The exponential of fitted values from equation (3) are the estimates of parental period two income variability. These estimate supposes households know the distribution of shocks across the population but does not require them to have information in period 1 about variables in period 2.

The child's expected earnings as adult

Predicted child incomes follow from an income regression for all household members older than 10 years who report work information. We run separate regressions for males and females, including educational dummies to estimate returns to schooling, and allowing the return to education to vary across provinces. The income equation is

$$\ln y_i = \alpha_0 + \alpha_1 esper_i + \alpha_2 esper_i^2 + \sum_{j=1}^{J}\sum_{p=1}^{P} \beta_{jp} D_{ji} * P_{pi} + u_i \qquad (4)$$

where the dependent variable is log monthly income of individual *i*, *esper* and *esper*^{2} are experience and squared experience and P_{pi} are provincial dummies. The equation includes educational dummy variables D_{ji} (for elementary, junior secondary, senior secondary, and high school) which enter the specification multiplicatively. In the absence of information on actual experience, *esper* is replaced by "potential experience", measured as $age_i - schooling\ years_i - 7$, assuming people start school at the age of 7. Income is observed only for those who work: we use the Heckman (1978) procedure to control for sample selection. The coefficients on education and experience estimated from the income regression (4) are used to construct predicted earnings by gender, provinces and school level *j* ($\hat{x}_j$) (coefficients from equation (4) are estimated on the sample of all household members older than 10 years that report work information and then used to predict income for the sample of children).

$$\hat{x}_j = \hat{\alpha}_1 esper_i + \hat{\alpha}_2 esper_i^2 + \sum_{p=1}^{P} \hat{\beta}_{jp} D_j * P_p \ ,$$

Estimation results are given in appendix C.

Variation of the child's earnings as adult

Uncertainty about the child's incremental income as adult is estimated by regressing the square of predicted residuals from equation (4) on a constant and all first moments, second moments and cross products of the original regressors in (4). We estimate two equations: one over the sample of adults graduated from senior school and the second over the sample of adults that have completed the junior school. The fitted values from these regressions are used to construct the uncertainty measures appertaining to the child's incremental income as adult, conditional on having completed the senior and junior school respectively.

In order to implement the model, we need information on schooling costs. We infer the cost of different levels of schooling from regressing total household expenditure for education on the number of children attending each school level (instrumented using predicted numbers from Poisson models). Appendix D outlines the econometric execution.

4.2. Schooling decisions

Making a distributional assumption for the child's earning ability a_h, we may estimate the equation for the school attendance decision using standard discrete choice methods. The most straightforward assumption is normality, $a_h \sim N\left(\bar{a}, \omega^2\right)$. On this basis, we estimate conditional probability of attending school as a probit model. Children and young adults are divided in two cohorts: children who have graduated from elementary school and junior secondary graduates. The progression to junior secondary school is estimated on the first sub-sample, while the continuation to senior school refers to the second cohort. To better capture the continuation decisions made at the time of the interview, the samples for the junior and senior school are restricted to children aged less than 15 years, and less than 18 years respectively[8]. In those cases in which a household has children in both cohorts the same income variable will define v_{h1} for the child in the younger cohort and v_{h2} for the child in the older cohort.

From the probit equation for senior school we estimate the conditional probability that the child does not attend senior school, $\Phi_h = \Phi\left(\frac{\delta\left(\hat{a}_h - \bar{a}\right)}{\omega}\right)$, where $\Phi(.)$ is the standard normal distribution function. This is estimated for the sample of children graduated from the junior secondary school, but is applied to the data for children

[8] As previously noted, this paper focuses on the sample of children 7-17 years old.

graduated from the elementary school, and hence used in modelling the junior secondary enrolment decision.

5. RESULTS

5.1. Shocks, uncertainty and schooling

Tables 6 and 7 present the empirical results for senior and junior secondary school enrolment[9]. Probit models are estimated for the sub-sample of children and young adults that completed the previous school level[10].

We first comment on the results of the probit model for senior school decisions (Table 6). This was estimated both including and omitting the schooling cost variable and its interaction term. Uncertainty about the child's expected earnings as adult ($\sigma^2_{h\zeta}$), interacted with permanent income, have the expected negative sign, while period two income shock (v_{h2}), interacted with the cost of schooling, has no effect on senior school participation.

As regard the effects of risk and shocks on junior school, results reported in Table 7 partly confirm the theoretical predictions (given in Table 3). In line with the linear approximation, the parental income variance for the time the child may be potentially at school (σ_{h2}, interacted with the cost of senior school and multiplied by the probability of continuing to senior school) has a positive (and statistically significant) impact on the probability of enrollment. This result confirms the importance of considering the option value effect: in the face of parental period two income variability parents are more likely to send children to junior school in the current period to give them the option to continue with senior school in the future (and hence earn higher earnings when they become adults).

The variance of child's income as adult, conditional on attending secondary school ($\sigma^2_{h\zeta}$, interacted with permanent income, and multiplied by the probability of continuing to senior school), has the predicted negative sign, but it is statistically

[9] All these regressors predicted by the model are estimates, consequently coefficient standard errors may be biased (Davidson and MacKinnon, 2004).

[10] To control for possible sample selection problems we augment the probit models for senior and junior decisions with the corresponding selection terms calculated from a probit that estimates the predicted probability of completing the junior and elementary level respectively.

significant at the 10% level only in the model without the cost of senior school. The variance of child's income as adult, conditional on not attending secondary school ($\sigma^2_{h\xi}$, interacted with permanent income) is not statistically significant.

The period 1 income shock (v_{h1}, interacted with the cost of junior school) has the predicted positive effect and it is statistically significant at the 10% and 5% level for the probit model estimated with and without the cost of senior school respectively

5.2. The effects of cost of schooling, predicted earning and other variables

Other determinants of the school enrollment decisions predicted by the theoretical model are the cost of schooling and the expected returns of education in terms in earnings. In the empirical specification we discount both expected earning and costs using a discount factor of $\delta = 0.95$ *per annum* (Zimmerman and Carter, 2003)[11].

Table 6 shows that the estimated coefficients for the child's expected earnings as adult ($\overline{\zeta}_{h3}$) has the expected positive effect on the probability of attending senior school, and it is statistically significant. However, the coefficient for the total cost of senior school is estimated as positive (predicted sign negative). Hence we also estimated the probit model omitting this cost variable and its interaction term (see columns 4 and 5 of the table).

A number of other explanatory variables are included as controls in the probit equation for senior school enrolment. The probability of attending senior school decreases with the number of younger siblings aged 13-15 in the house[12], and if the spouse of the household head is illiterate[13]. As suggested by other papers, we find that the ownership of a farm decreases the probability that a child is in school, because this increases the child's work participation (Cockburn and Dostie, 2007). Other control variables are the number of good senior schools in the village (a

[11] Assuming 55-60 years as the retirement age (Leechor, 1996), and in view of the fact that the official age at which children should complete primary school is 12 years, the length of "rest of life" (period three) is approximately 37-42 years.

[12] Traditionally the first child drops out school or is not enrolled in school to provide education opportunities for the younger children. Our results suggest this may be the case. In future work we propose to introduce a birth order variable.

[13] The education of the spouse of the head is considered instead of that of mother's education because some children do not report to have a mother in the household. In most cases the spouse is the mother of the child.

positive and significant effect)[14] and the religion of the household (children living in a Muslim or Christian household are more likely to attend senior school compared to household with other religions).[15]

The results reported in Table 7 show that, in line with the model predictions, the total cost of junior school has a negative impact on the probability of enrollment. The cost of attending senior school has positive estimated coefficients on the probability of a child attending school (predicted sign negative), as was the case in the senior school progression estimates already reported. We also estimate the model without this cost variable – see Table 7 (columns 4). Expected earnings from junior ($\gamma\xi_{h2} + \gamma\delta\overline{\xi}_{h3}$) and senior school ($\overline{\zeta}_{h3}$) have the predicted sign, but only the latter variable is statistically significant at the 10% level in the model without the cost of senior school.

Among the other control variables, the number of siblings 13-15 years old and 16-18 years old have a positive and significant effect (the former at the 10% but not at the 5% level). We remark that at lower levels of schooling, school progression of a child is positively associated with the number of siblings of the same age or older (Sawada and Lokshin, 2007), while at higher levels of schooling (see the results of senior secondary school) the probability that the child attends school decreases with the number of younger siblings.

---

[14] An important issue in determining school attendance is the availability of school facilities. In Indonesia, lack of schooling infrastructure emerges as a problem at the secondary level: according to IFLS data in 1993 nearly 65% and 85% of villages did not have a junior or a senior secondary school respectively, and very few villages had more than one such school. Even when school facilities are available at a convenient distance, their quality may be inadequate (The PROBE Team, 1999). In 1993 nearly 40% of the junior and senior secondary schools are considered of good quality by the head of the village.

[15] "Christian" and "Muslim" have similar marginal effects.

Table 6
Probit model for senior secondary school

Variables	*Coeff*	*z*	*Coeff*	*z*
# of siblings age 13-15	-0.987	-4.08	-0.945	-4.13
Spouse of the head is illiterate*	-1.088	-2.83	-1.086	-2.89
Total cost of senior school/1000 $(x_h + q_h)$	0.488	1.81	-	-
$(x_h + q_h)$* v_{h2}/1000,000	-0.153	-0.41	-	-
$\overline{\zeta}_{h3}$ $(= z_h - x_h)$	0.022	2.30	0.017	1.80
$\sigma^2_{h\zeta}$ * permanent income (/1000)	-0.138	-1.99	-0.096	-1.40
household owns a farm*	-0.829	-2.77	-0.834	-2.88
household religion: islam*	0.663	1.54	0.787	2.05
household religion: cristian*	2.193	2.26	2.330	2.43
# good senior school in the village	1.497	2.21	1.550	2.38
Selection term	-1.342	-3.62	-1.358	-3.79
Intercept	1.519	2.21	2.269	3.81
Pseudo R-squared	0.346		0.318	
Log pseudo-likelihood	-61.0123		-63.617	
N	135			

The table reports the results of the probit model for senior school attendance. Dependent variable is a dummy that equals one if the child is attending school. The sample is children graduated from the junior school.

Table 7
Probit model for junior secondary school

Variables	*Coeff.*	*z*	*Coeff*	*z*
Total cost of junior secondary school $(w_{h1}+p_h)/100$	-0.146	-2.78	-0.103	-2.93
$(w_{h1}+p_h)*v_{h1}/1000{,}000$	0.075	1.69	0.097	2.24
Cost of senior school $(q_h+w_{h2}+\xi_{h2})*(1-\Phi_h)/1000$	0.362	1.10	-	-
$(q_h+w_{h2}+\xi_{h2})*\sigma_{h2}*(1-\Phi_h)/1000{,}000$	0.580	2.22	0.948	3.74
$\gamma\xi_{h2}+\gamma\delta\overline{\xi}_{h3}$	0.004	0.96	0.003	0.81
$\sigma^2_{h\xi}$ * permanent income/100,000	0.200	0.01	-0.686	-0.00
$\overline{\zeta}_{h3}*(1-\Phi_h)$	0.008	1.00	0.013	1.72
$\sigma^2_{h\zeta}*(1-\Phi_h)$ *permanent income/1000	-0.130	-1.12	-0.207	-1.71
# siblings age 13-15	0.496	2.55	0.470	2.44
# siblings age 16-18	0.436	3.22	0.411	3.19
intercept	0.343	1.10	0.159	0.56
Pseudo R-squared	0.11		0.11	
Log pseudo-likelihood	-190.83		-190.58	
N	312			

The table reports the results of the probit model for junior secondary decisions. Dependent variable is a dummy that equals one if the child is attending school. Sample: children graduated from the elementary school, aged 12-14 years. $(1-\Phi_h)$ is the probability of continuation to the senior secondary school.

6. CONCLUSIONS

This work uses the Indonesian Family Life Survey to explore the effects of ex-ante risk and ex-post shocks on child education. We develop a model that takes into account the irreversibility of withdrawal from school: drop out children cannot re-enroll, and hence they lose the opportunity to continue to further levels of schooling and achieve higher earnings when they become adults. In this way, temporary interruptions in child schooling have long term impacts on the child human capital.

Various papers have shown the use of child time allocation as a means to cope with income shocks ex post. Other studies explore how uncertainty affects child education. In this paper we combine these two streams of literature by examining the role of ex ante income risk and of ex post income shock occurrence on child education, and, in particular, on school progression from one school level to another. We consider the effect of different sources of uncertainty: uncertainty about the parents' transitory income at the time the child could be attending school and uncertainty about the child's incremental income as adult conditional on attending different school level. Shocks are the ex post counterpart of ex ante uncertainty, and relate to parental transitory income.

We have developed a model of household school transition decisions from elementary education to junior education and from junior to senior school under the assumptions that there are no savings and borrowing, that children continue to form part of the household once they become adults, and that withdrawal from school is permanent. This has the implication that a family needs to anticipate the second transition decision at the time they make the decision with regard to the initial transition. The family must solve a dynamic programming problem.

An important feature of our model is that education allows children to realize their latent ability. We model the expected returns to secondary education as depending on their ability, and this reinforces the need to anticipate at the initial transition. Ability is unobserved by the economist.

There is no simple answer to the question of how uncertainty affects schooling decisions. Our model predicts that income shocks increase the probability of attending school, uncertainty with respect to the earnings of the adult (ex-child) once graduated tends to discourage school attendance, while uncertainty with respect to parents' income over the time the child may remain at school should increase school participation. This second effect arises out of irreversibility: in the face of household income variability parents are more likely to send children to school to give them the option to continue with higher schooling levels in the future (and hence earn higher earnings when they become adults).

We estimate school progression by estimating probit models for both junior and senior secondary school on different cohorts: the sample for junior attendance is those children graduated from the elementary school, aged less than 15 years; the probit for senior school is estimated on the sample of junior secondary graduates with aged less than 18 years. Results suggest that uncertainty about parental income for the time the child may be potentially at school increases the probability of attending junior school, thus confirming the importance of the option value effect (the option being that of continuation to senior secondary school). Uncertainty about expected earnings from education has a negative and significant effect only for senior school attendance. As regards the ex-post counterpart of uncertainty, we find that positive/negative shocks increase/decrease the probability of attending junior school.

REFERENCES

American Institutes for Research (2002), *Case studies in secondary education reform*, Improving Educational Quality (IEQ) Project, Clementina Acedo (ed.).

Baland, J. and J. Robinson (2000), "Is child labor inefficient?, *Journal of Political Economy*, 108(4), 663-679.

Beegle, K, R. Dehejia, and R. Gatti (2003), "Do Households Resort to Child Labor to Cope with Income Shocks?", *World Bank Policy Research Working Paper* N. 3075, Washington, D.C.: The World Bank.

Beegle, K, R. Dehejia, and R. Gatti (2006), "Child Labour, Crop Shocks, and Credit Constraints", *Journal of Development Economics*, 81(1), 80-96.

Cameron, L.A. and C. Worswick (2003), "The Labour Market as a Smoothing Device: Labour Supply Responses to Crop Loss", *Review of Development Economics* 7(2), 327-341.

Cockburn, J. and B. Dostie (2007), "Child Work and Schooling: The Role of Household Asset Profiles and Poverty in Rural Ethiopia", *Journal of African Economies*, 16(4), 519-563.

Davidson, R. and J.B. MacKinnon (2004), *Econometric Theory and Methods*, Oxford University Press, New York.

de Janvry, A., F. Finan, and E. Sadoulet (2006), "Can Conditional Cash Transfers Serve as Safety Nets to Keep Children at School and Out of the Labor Market?", *Journal of Development Economics*, 79(2), 349-373.

Duryea, S., D. Lam, and D. Levison (2007), "Effects of Economic Shocks on Children's Employment and Schooling in Brazil", *Journal of Development Economics*, 84, 188-214.

Fitzsimons, E. (2007), "The Effects of Risk on Education in Indonesia", *Economic Development and Cultural Change*, 56(1), 1-25.

Guarcello, L., F. Mealli and F.C. Rosati (2003), "Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing and Insurance", *Social Protection Discussion Paper Series* N. 0322.

Guarcello, L., F. Rosati and P. Scaramozzino (2008), "Uncertainty, education, and the school-to-work transition: theory and evidence from Brazil", *UCW Working Paper* N. 40.

Gubert, F. and A.S. Robilliard (2008), "Risk and Schooling Decisions in Rural Madagascar: A Panel Data-Analysis", *Journal of African Economies*, 17(2), 207-238.

Heckman, J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", *Econometrica* 46(4), 931-60.

Hogan, V. and I. Walker (2007), "Education choice under uncertainty: Implications for public policy", *Labour Economics*, Elsevier, 14(6), 894-912.

Jacoby, H. and E. Skoufias (1997), "Risk, financial markets, and human capital in a developing country", *Review of Economic Studies*, 64(3), 311–335.

Jalan, J. and M. Ravallion (2001), "Behavioral Responses to Risk in Rural China", *Journal of Development Economics*, 66(1), 23-49.

Jensen, R. (2000), "Agricultural Volatility and Investments in Children", *American Economic Review*, 90(2), 399–405.

Kazianga, H. (2005), "Income risk and schooling decisions in Burkina Faso", mimeograph, Earth Institute, Columbia University.

Leechor, C. (1996), "Reforming Indonesia's Pension System", *World Bank Policy Research Working Paper* N. 1677.

Maitra, P., P. Bibhudutta, and S. Sfrangi (2006), "Idle Child: the Household's Buffer", Mimeo, Louisiana State University.

Manning, C. (2000), "Labour Market Adjustment to Indonesia's Economic Crisis: Context, Trends and Implications, *Bulletin of Indonesian Economic Studies*, 36(1), 105-136.

Menon, M. and F. Perali (2006), "The Shadow Wage of Child Labour: An Application to Nepal", *UCW Working Paper* N. 11.

Morduch, J (1990), "Risk, Production and Saving: Theory and Evidence from Indian Households", Harvard University. Mimeo.

Ranjan, P. (2001), "Credit constraints and the phenomenon of child labor", *Journal of Development Economics*, 64(1), 81-102.

Sawada, Y. (2003), "Income Risks, Gender, and Human Capital Investment in a Developing Country", *CIRJE F-Series* N. CIRJE-F-198, CIRJE, Faculty of Economics, University of Tokyo.

Sawada, Y. and M. Lokshin (2001), “Household Schooling Decisions in Rural Pakistan”, The World Bank, *Policy Research Working Paper Series* N. 2541.

Sawada, Y. and M. Lokshin (2007), “Obstacles to School Progression in Rural Pakistan: An Analysis of Gender and Sibling Rivalry Using Field Survey Data”, *CIRJE Discussion Paper* N.CIRJE-F-484.

The PROBE Team (1999), *Public Report on Basic Education in India*, Oxford University Press, USA.

Thomas, D, K. Beegle, E. Frankenberg, B. Sikoki, J. Strauss, and G. Teruel (2004), “Education during a Crisis”, *Journal of Development Economics*, 74(1), 53-86.

Townsend, R.M. (1994), “Risk and Insurance in Village India”, *Econometrica*, 62(3), 539-591.

UNDP (2004), *Indonesia: progress report on the Millenium development goals*, Jakarta: Coordinating Minister for People's Welfare, 2004.

Zimmerman, F. and M.R. Carter (2003), “Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints”, *Journal of Development Economics*, 71, 233-260.

APPENDIX A

We write variables whose values have yet to be revealed with a tilde.

We assume

(A1) Parental incomes are independent of schooling decisions.

(A2) Additional schooling cannot reduce the child's future earnings, i.e. $x_2 \geq w_2$ and $z_3 \geq x_3 \geq w_3$. These conditions imply $\xi_2, \xi_3, \zeta_3, a \geq 0$.

(A3) Family income in each period is given by $y_t = y^p + v_t \quad (t = 1,2,3)$ where y^p is a constant permanent component and $v_t \quad (t = 1,2,3)$ is a transitory component.

(A4) Independence: We assume ξ_3, ζ_3, v_1, v_2 and v_3 are independently distributed with distribution functions $H_\xi(\xi_3)$, $H_\zeta(\zeta_3)$, $G_1(v_1)$, $G_2(v_2)$ and $G_3(v_3)$ respectively. ξ_3 has support $\left[0, \hat{\xi}\right]$, ζ_3, has support $\left[0, \hat{\zeta}\right]$ and v_1, v_2 and v_3 have common support $[\underline{v}, \hat{v}]$ and satisfy $E(v_j) = 0 \quad (j = 1,2,3)$. Write $E(\xi_3) = \overline{\xi}$ and $E(\zeta_3) = \overline{\zeta}$.

(A5) Ability a is observed by the family but not the economist. It has finite support $[0, \hat{a}]$. Permanent income y^p is observed by the family but not directly be the economist. It has finite support $\left[\underline{y}, \hat{y}\right]$. Ability and permanent income are jointly distributed in the population with distribution function $F(a, y^p)$.

(A6) The utility function $u(.)$ has positive first and third and negative second derivatives, i.e. $u'(.) > 0$, $u''(.) < 0$ and $u'''(.) > 0$. The function satisfies $\lim_{m \to 0} u(m) = -\infty$ and $\lim_{m \to \infty} u'(m) = 0$.

The family's maximization problem is solved as a dynamic programme. The household first evaluates the expected decision to attend senior school and then moves back to the earlier decision with respect to the junior school enrolment decision conditional on the anticipated senior secondary school decision.

Consider the decision to allow a child to continue to senior school conditional on having completed junior school. The household will do this if

$$D\left(a, y^p, v_2\right) = B\left(a, y^p\right) - C\left(y_2\right) \geq 0 \qquad \text{(A.1)}$$

where[16]

$$\begin{aligned} B\left(a, y^p\right) &= \delta E\left[u\left(\tilde{y}_3 + \tilde{z}_3\right) - u\left(\tilde{y}_3 + \tilde{x}_3\right)\right] \\ &= \delta E\left[u\left(y^p + \tilde{v}_3 + w_3 + \tilde{\xi}_3 + \tilde{\zeta}_3 + a\right) - u\left(y^p + \tilde{v}_3 + w_3 + \tilde{\xi}_3\right)\right] \end{aligned}$$

and
$$C\left(y_2\right) = u\left(y_2 + x_2\right) - u\left(y_2 - q\right)$$

δ is the discount rate from period 3 to period 2 and the subscript 2 after the expectation operator indicates an expectation taken relative to period 2 information. $B\left(a, y^p\right)$ is the expected benefit from attending senior school and $C\left(y_2\right)$ is the expected cost. Note that y_2 is known at the time this decision is made and therefore, conditional on the value of current income, does not depend on the family's permanent income.

<u>Lemma 1:</u>

For every pair $\left(a, y_2\right)$ with $a \in \left[0, \hat{a}\right]$ and $y_2 \in \left[\underline{y} - \underline{v}, \hat{y} + \hat{v}\right]$ there exists a unique period 2 income $y_2 = \lambda\left(a, y^p\right)$ that $B\left(a, y^p\right) = C\left(\lambda\right)$. At income levels $y_2 > \lambda\left(a, y^p\right)$, the family will enroll the child in senior school. At income levels $y_2 < \lambda\left(a, y^p\right)$, the family will not enroll the child in senior school. The function $\lambda\left(a, y^p\right)$ has derivatives $\dfrac{\partial \lambda\left(a, y^p\right)}{\partial a} < 0$ and $\dfrac{\partial \lambda\left(a, y^p\right)}{\partial y^p} > 0$.

<u>Proof:</u>

Both $B\left(a, y^p\right)$ and $C\left(y_2\right)$ are continuous in their arguments. Differentiation of $C\left(y_2\right)$ gives

[16] .Since there is no temporal dependence, we do not need to distinguish expectations by dates,

$$\begin{aligned} C'(y_2) &= u'(y_2 + x_2) - u'(y_2 - q) < 0 \\ C''(y_2) &= u''(y_2 + x_2) - u''(y_2 - q) > 0 \end{aligned} \tag{A.2}$$

where the signs of the derivatives follow from assumption (A6). The same assumption implies $\lim_{y_2 \to q} C(y_2) = \infty$ and $\lim_{y_2 \to \infty} C(y_2) = 0$. Differentiation of $B(a, y^p)$, again in conjunction with assumption (A6), gives

$$\begin{aligned} &\frac{\partial B(a, y^p)}{\partial a} > 0 \quad \frac{\partial_2 B(a, y^p)}{\partial a^2} < 0 \\ &\frac{\partial B(a, y^p)}{\partial y^p} < 0 \quad \frac{\partial_2 B(a, y^p)}{\partial (y^p)^2} > 0 \end{aligned}$$

Then $B(a, y^p)$ has support $\left[\underline{b}, \hat{b}\right]$ where

$$\begin{aligned} \underline{b} &= B(0, \hat{y}) = \delta\left[E_2 u(\hat{y} + v_3 + \tilde{z}_3) - E_2 u(\hat{y} + v_3 + \tilde{x}_3)\right] \geq 0 \\ \hat{b} &= B(\hat{a}, \underline{y}) = \delta\left[E_2 u(\underline{y} + \tilde{v}_3 + \tilde{z}_3) - E_2 u(\underline{y} + \tilde{v}_3 + \tilde{x}_3)\right] > \underline{b} \end{aligned}.$$

It follows that for there exists an income $y_2 = \lambda(a, y^p)$ which satisfies $B(a, y^p) = C(\lambda)$. Differentiation of $\lambda(a, y^p)$ gives $\frac{\partial \lambda(a, y^p)}{\partial a} = \frac{1}{C'} \cdot \frac{\partial B(a, y^p)}{\partial a} < 0$, i.e. the higher the child's ability level the greater the advantage of senior school education and hence the lower the threshold income below which he will not be enrolled in senior school. Similarly, $\frac{\partial \lambda(a, y^p)}{\partial y^p} = \frac{1}{C'} \cdot \frac{\partial B(a, y^p)}{\partial y^p} > 0$. Although a family with high permanent income can more easily afford to keep a child at school, it has less need to do so. (This effect disappears within the individualistic framework, i.e. when we suppose that the child, once adult, becomes a separate unit with its own utility function, and that there are no inter-generational transfers).

Comments:

1. The function $\lambda(a, y^p)$ is graphed in Figure 1 for a fixed value of y^p. In period 2, when the family takes the progression decision with respect to senior school, the family's period 2 income y_2 is known. In period 1, when the family must anticipate that decision, only permanent income y^p and the period 2 transient income

distribution $G_2(v_2)$ are known. This makes it more natural to work in terms of the function $\eta(a,y^p)=\lambda(a,y^p)-y^p$. It is not possible to sign the derivative $\frac{\partial\eta(a,y^p)}{\partial y^p}=\frac{\partial\lambda(a,y^p)}{\partial y^p}-1$ *a priori* although it tends to be negative for reasonable values of the parameters. Note that we can express equation (1) as

$$D\left(a,y^p,\eta\left(a,y^p\right)\right)=0 \tag{A.3}$$

2. If we were to adopt an individualist approach, the benefit term in equation (A.1) would become $B\left(a,y^p\right)=\delta E\left[\upsilon(\tilde{z}_3)-\upsilon(\tilde{x}_3)\right]$ where υ(.) is the child's utility function as adult. The parents' transitory income v_3 ceases to enter the decision problem. As a consequence, the threshold λ in independent of the family's permanent income y^p,

3. The child will be enrolled in secondary school if period 2 transitory income satisfies $v_2\geq\eta\left(a,y^p\right)$. Let π^s be the proportion of children in the population of junior school graduates enrolling in senior school. It follows that

$$\pi^s=1-\int_{\underline{y}}^{\hat{y}}\int_0^{\hat{a}}\int_{\underline{v}}^{\eta(a,y^p)}dG_2\left(v_2\right)dF\left(a,y^p\right)=1-\int_{\underline{y}}^{\hat{y}}\int_0^{\hat{a}}G_2\left(\eta\left(a,y^p\right)\right)dF\left(a,y^p\right) \tag{A.4}$$

where $G_2(v_2)$ are the distribution function period 2 transitory income v_2 and $F\left(a,y^p\right)$ is the joint distribution function of ability a and permanent income y^p.

We now go back one period and consider the decision to attend junior school. If the child does not attend junior school, family utility is

$$u\left(y_1+w_1\right)+\gamma E_1u\left(\tilde{y}_2+w_2\right)+\gamma\delta E_1u\left(\tilde{y}_3+w_3\right) \tag{A.5}$$

Instead, if the child continues at school, family utility is

$$u\left(y_1-p\right)+E_1\max\left[\gamma u\left(\tilde{y}_2+x_2\right)+\gamma\delta u\left(\tilde{y}_3+\tilde{x}_3\right),\gamma u\left(\tilde{y}_2-q\right)+\gamma\delta u\left(\tilde{y}_3+\tilde{z}_3\right)\right] \tag{A.6}$$

The first set of term on the right hand side of equation (A.6) show the benefits of attending junior secondary school. The final term gives the value of the option of continuing to senior secondary school. Combining equations (A.6) and (A.7), the child will be enrolled in junior school if

$$K\left(y^{p}\right)+L\left(a,y^{p}\right)-J\left(y_{1}\right)\geq 0 \tag{A.7}$$

where $$J\left(y_{1}\right)=u\left(y_{1}+w\right)-u\left(y_{1}-p\right),$$

the period 1 utility cost of junior school enrolment,

$$\begin{aligned}K\left(y^{p}\right)&=\gamma E\left[\left(u\left(\tilde{y}_{2}+\tilde{x}_{2}\right)-u\left(\tilde{y}_{2}+w_{2}\right)\right)+\delta\left(u\left(\tilde{y}_{3}+\tilde{x}_{3}\right)-u\left(\tilde{y}_{3}+w_{3}\right)\right)\right]\\&=\gamma\left[\left(u\left(y^{p}+\tilde{v}_{2}+\xi_{2}+w_{2}\right)-u\left(y^{p}+\tilde{v}_{2}+w_{2}\right)\right)+\delta E\left(u\left(y^{p}+\tilde{v}_{3}+\tilde{\xi}_{3}+w_{3}\right)-u\left(y^{p}+\tilde{v}_{3}+w_{3}\right)\right)\right]\end{aligned} \tag{A.8}$$

is the discounted expected net benefit from junior school enrolment supposing that the child does not proceed to senior school, and

$$\begin{aligned}L\left(a,y^{p}\right)&=\gamma E\int_{\eta\left(a,y^{p}\right)}^{\hat{v}}\left(u\left(\tilde{y}_{2}-q\right)-u\left(\tilde{y}_{2}+\tilde{x}_{2}\right)\right)+\delta\left(u\left(\tilde{y}_{3}+\tilde{z}_{3}\right)-u\left(\tilde{y}_{3}+\tilde{x}_{3}\right)\right)dG_{2}\left(\tilde{y}_{2}-y^{p}\right)\\&=\gamma\int_{\eta\left(a,y^{p}\right)}^{\hat{v}}\left\{\left(u\left(y^{p}+\tilde{v}_{2}-q\right)-u\left(y^{p}+\tilde{v}_{2}+w_{2}+\xi_{2}\right)\right)\right.\\&\quad\left.+\delta E\left[u\left(y^{p}+\tilde{v}_{3}+\tilde{\xi}_{3}+\tilde{\zeta}_{3}+w_{3}+a\right)-u\left(y^{p}+\tilde{v}_{3}+w_{3}+\tilde{\xi}_{3}\right)\right]\right\}dG_{2}\left(\tilde{v}_{2}\right)\end{aligned} \tag{A.9}$$

is the discounted expected net benefit from junior school enrolment in the circumstance that the child proceeds to senior school. (For notational convenience, we take expectations over period 3 variables but integrate over the distribution $G_{2}\left(v_{2}\right)$ of v_2). Using equation (A.1), we may write condition (A.9) as

$$L\left(a,y^{p}\right)=\gamma E\int_{\eta\left(a,y^{p}\right)}^{\hat{v}}D\left(a,y^{p},\tilde{v}_{2}\right)dG_{2}\left(\tilde{v}_{2}\right) \tag{A.10}$$

Note that the period 2 threshold income $\eta\left(a,y^{p}\right)$ does not depend on any period 2 variables and hence can be evaluated by the family at the start of period 1.

<u>Lemma 2:</u>

For every pair $\left(a,y^{p}\right)$ with $a\in\left[0,\hat{a}\right]$ and $y^{p}\in\left[\underline{y},\widehat{y}\right]$ there exists a unique period 1 income $y_{1}=\mu\left(a,y^{p}\right)$ that $J\left(\mu\right)=K\left(y^{p}\right)+L\left(a,y^{p}\right)$. At income levels

$y_1 > \mu(a, y^p)$, the family will enroll the child in senior school. At income levels $y_1 < \mu(a, y^p)$, the family will not enroll the child in senior school. The function $\mu(a, y^p)$ has derivatives $\frac{\partial \mu(a, y^p)}{\partial a} < 0$ and $\frac{\partial \mu(a, y^p)}{\partial y^p} > 0$.

Proof:

This has the same structure as the proof of Lemma 1. Differentiating $J(y_1)$,

$$\frac{dJ(y_1)}{dy_1} = u'(y_1 + w) - u'(y_1 - p) < 0.$$

From assumption (A6), $\lim_{y_1 \to p} J(y_1) = \infty$ and $\lim_{y_1 \to \infty} J(y_1) = 0$. Differentiation of $L(a, y^p)$ with respect to *a*, using equation (A.10), gives

$$\frac{\partial L(a, y^p)}{\partial a} = \gamma \frac{\partial}{\partial a} E \int_{\eta(a, y^p)}^{\hat{v}} D(a, y^p, \tilde{v}_2) dG_2(\tilde{v}_2)$$

$$= \gamma\left[1 - G_2\left(\eta(a, y^p)\right)\right] E\left[\frac{\partial B(a, y^p)}{\partial a}\right] - \gamma E\left[D\left(a, y^p, \eta(a, y^p)\right)\right] g_2\left(\eta(a, y^p)\right) \frac{\partial \eta(a, y^p)}{\partial a} \tag{A.11}$$

where $g_2(v_2)$ is the density of v_2. The final term in equation (A.11), which arises from differentiation of the lower limit of integration in equation (A.10), is equal to zero from equation (A.3). The initial term is positive since $\frac{\partial B(a, y^p)}{\partial a} > 0$ from Lemma 1 implying $\frac{\partial L(a, y^p)}{\partial a} > 0$ as required.

Differentiation of $L(a, y^p)$ with respect to y^p, using equation (A.10), gives

$$\frac{\partial L\left(a, y^{p}\right)}{\partial y^{p}} = \gamma \frac{\partial}{\partial y^{p}} E \int_{\eta\left(a, y^{p}\right)}^{\hat{v}} D\left(a, y^{p}, \tilde{v}_{2}\right) dG_{2}\left(\tilde{v}_{2}\right)$$

$$= \gamma\left[1 - G_{2}\left(\eta\left(a, y^{p}\right)\right)\right] E\left[\frac{\partial B\left(a, y^{p}\right)}{\partial y^{p}}\right] - \gamma \int_{\eta\left(a, y^{p}\right)}^{\hat{v}} \frac{\partial C\left(y^{p} + \tilde{v}_{2}\right)}{\partial y^{p}} dG_{2}\left(\tilde{v}_{2}\right) \tag{A.12}$$

where a final term, equal to zero, as in equation (11), has been omitted. The initial term in equation (A.12) is negative since $\frac{\partial B\left(a, y^{p}\right)}{\partial y^{p}} < 0$ from Lemma 1. The second term is also negative since $C' < 0$ from equation (A.2). Hence $\frac{\partial L\left(a, y^{p}\right)}{\partial y^{p}} < 0$ as required.

From these two results, we can infer that $L\left(a, y^{p}\right)$ has support $\left[\underline{l}, \hat{l}\right]$ where $\underline{l} = L\left(0, \hat{y}\right)$ and $\hat{l} = L\left(\hat{a}, \underline{y}\right)$. It follows that for every ability-permanent income pair $\left(a, y^{p}\right)$ there is a unique period 1 income $\mu\left(a, y^{p}\right)$ which solves

$$J\left(\mu\right) = K\left(y^{p}\right) + L\left(a, y^{p}\right) \tag{A.13}$$

Further
$$\frac{d\mu\left(a, y^{p}\right)}{da} = \frac{1}{J'} \frac{\partial L\left(a, y^{p}\right)}{\partial a} < 0$$

implying that the higher the child's ability level *a*, given the family's permanent income y^p, the lower the threshold income below which he will not be enrolled in junior school. It is immediate from equation (7) that $\frac{\partial K\left(a, y^{p}\right)}{\partial y^{p}} < 0$. It follows that

$$\frac{d\mu\left(a, y^{p}\right)}{dy^{p}} = \frac{1}{J'}\left[\frac{\partial K\left(y^{p}\right)}{\partial y^{p}} + \frac{\partial L\left(a, y^{p}\right)}{\partial y^{p}}\right] > 0$$

implying that the higher is family's permanent income y^p, given the child's earning ability, the higher is the threshold income below which he will not be enrolled in junior school. As in the case of the senior school continuation decision, although a family with high permanent income can more easily afford ti keep a child at school, it has less need to do so.

Comment:

The child will be enrolled in junior school if period 1 transitory income satisfies $v_1 \geq \mu\left(a, y^p\right) - y^p$. Let π^j be the proportion of children in the population of elementary school graduates enrolling in junior school. It follows that

$$\pi^j = 1 - \int_{\underline{y}}^{\hat{y}} \int_0^{\hat{a}} \int_{\underline{v}}^{\mu\left(a, y^p\right) - y^p} dG_1\left(v_1\right) dF\left(a, y^p\right) = 1 - \int_{\underline{y}}^{\hat{y}} \int_0^{\hat{a}} G_1\left(\mu\left(a, y^p\right) - y^p\right) dF\left(a, y^p\right) \tag{A.14}$$

where $G_1\left(v_1\right)$ are the distribution function period 1 transitory income.

The senior school enrolment decision

We are now in a position to look at the impact of income uncertainty. We start with the senior school enrolment decision and evaluate the impact of uncertainty with respect to the period 3 earnings distribution functions $H_\xi\left(\xi_3\right)$, $H_\zeta\left(\zeta_3\right)$ and $G_3\left(y_3\right)$ .We first prove a lemma.

Lemma 3:

Consider a set of mean-preserving spreads applied in turn to the distributions of ξ_3, ζ_3 and v_3. The derivatives of the mean-preserving spreads on the expected benefit $B\left(a, y^p\right)$ defined in equation (1) with respect to the distributions $H_\xi\left(\xi_3\right)$ and $G_3\left(v_3\right)$ are positive and that with respect to $H_\zeta\left(\zeta_3\right)$ is negative.

Proof:

1. First consider a mean-preserving spread $\zeta_3 \to \zeta_3 + \theta_\zeta\left(\zeta_3 - \overline{\zeta}\right)$ where $\overline{\zeta} = E\left(\zeta_3\right)$. Differentiating $B\left(a, y^p, \theta_\zeta\right)$ with respect to the spread parameter θ_ζ, we obtain

$$\frac{\partial B\left(a, y^p, \theta_\zeta\right)}{\partial \theta_\zeta} = \delta E_2\left[u'\left(y^p + \tilde{v}_3 + w_3 + \tilde{\xi}_3 + \left(1+\theta_\zeta\right)\tilde{\zeta}_3 + a - \theta_\zeta\overline{\zeta}\right)\left(\tilde{\zeta}_3 - \overline{\zeta}\right)\right]$$

$$= \delta Cov_2\left(u'\left(y^p + \tilde{v}_3 + w_3 + \tilde{\xi}_3 + \left(1+\theta_\zeta\right)\tilde{\zeta}_3 + a - \theta_\zeta\overline{\zeta}\right), \tilde{\zeta}_3 - \overline{\zeta}\right) < 0$$

(A.15)

2. Now consider first a mean-preserving spread $\xi_3 \to \xi_3 + \theta_\xi\left(\xi_3 - \overline{\xi}\right)$ where $\overline{\xi}_3 = E\left(\xi\right)$.

Differentiating:

$$\frac{\partial B\left(a, y^p, \theta_\xi\right)}{\partial \theta_\xi} = \delta E_2\left[u'\left(y^p + \tilde{v}_3 + w_3 + \left(1+\theta_\xi\right)\tilde{\xi}_3 + \tilde{\zeta}_3 + a - \theta_\xi\overline{\xi}\right)\right.$$
$$\left.\left(-u'\left(y^p + \tilde{v}_3 + w_3 + \left(1+\theta_\xi\right)\tilde{\xi}_3 - \theta_\xi\overline{\xi}\right)\right)\left(\tilde{\xi}_3 - \overline{\xi}\right)\right] \quad \text{(A.16)}$$

There is a complication here that the child's earning ability *a* may be too large to permit linear approximation. We circumvent this problem by bounding the derivative by taking first order approximations around the two income arguments. By virtue of the third derivative component of assumption (A6),

$$Cov_2\left(u''\left(y^p + \tilde{v}_3 + w_3 + \left(1+\theta_\xi\right)\tilde{\xi}_3 + a + \tilde{\zeta}_3 - \theta_\xi\overline{\xi}\right), \tilde{\xi}_3\right) \geq \frac{1}{\delta\left(a+\overline{\zeta}\right)}\frac{\partial B\left(a, y^p, \theta_\xi\right)}{\partial \theta_\xi}$$
$$\geq Cov_2\left(u''\left(y^p + \tilde{v}_3 + w_3 + \left(1+\theta_\xi\right)\tilde{\xi}_3 - \theta_\xi\overline{\xi}\right), \tilde{\xi}_3\right) > 0$$

(A.17)

Hence $\dfrac{\partial B\left(a, y^p, \theta_\xi\right)}{\partial \theta_\xi} > 0$.

3. Consider a mean-preserving spread $v_3 \to \left(1+\theta_3\right)v_3$. The proof is identical to that of the mean-preserving spread in ξ_3.

$$Cov_2\left(u''\left(y^p + \left(1+\theta_3\right)\tilde{v}_3 + w_3 + \tilde{\xi}_3 + a + \tilde{\zeta}_3\right), \tilde{v}_3\right) \geq \frac{1}{\delta\left(a+\overline{\zeta}\right)}\frac{\partial B\left(a, y^p, \theta_{y_3}\right)}{\partial \theta_{y_3}}$$
$$\geq Cov_2\left(u''\left(y^p + \left(1+\theta_3\right)\tilde{v}_3 + w_3 + \tilde{\xi}_3 y^p +\right), \tilde{v}_3\right) > 0$$

(A.18)

again using assumption (A6). Hence $\frac{\partial B\left(a, y^{p}, \theta_{3}\right)}{\partial \theta_{3}} > 0$ (This effect disappears if we adopt the individualistic approach the child's earnings as adult).

Theorem 1:

Increasing mean-preserving spreads in the distribution $H_3\left(\xi_3\right)$ of the child's incremental period 3 income ξ_3 (i.e. income as adult), conditional on non-enrolment in senior school, and in the distribution $G_3\left(y_3\right)$ of the parents' period 3 income y_3, both increase enrolment in senior school.

Proof:

From Lemma 3, in both cases $\frac{\partial B\left(a, y^{p}, \theta_{r}\right)}{\partial \theta_{r}} > 0 \quad \left(r = \xi, 3\right)$. It follows that

$$\frac{\partial \lambda\left(a, y^{p}, \theta_{r}\right)}{\partial \theta_{r}} = \frac{1}{C'\left(y_1\right)} \frac{\partial B\left(a, y^{p}, \theta_{r}\right)}{\partial \theta_{r}} < 0 \quad \left(r = \xi, 3\right)$$

Increases in income uncertainty associated with ξ_3 and y_3 therefore both decrease the income threshold at which a child will be enrolled in senior school. Since $\frac{\partial \eta\left(a, y^{p}, \theta_{r}\right)}{\partial \theta_{r}} = \frac{\partial \lambda\left(a, y^{p}, \theta_{r}\right)}{\partial \theta_{r}} \quad \left(r = \xi, 3\right)$

$$\frac{d\pi^{s}}{d\theta_{r}} = -\int_{\underline{y}}^{\hat{y}} \int_{0}^{\bar{a}} g_2\left(\eta\left(a, y^{p}, \theta_{r}\right)\right) \frac{\partial \eta\left(a, y^{p}, \theta_{r}\right)}{\partial \theta_{r}} dF\left(a, y^{p}\right) > 0 \quad \left(r = \xi, 3\right) \tag{A.19}$$

where $g_2\left(v_2\right) = \frac{dG_2\left(v_2\right)}{dv_2}$ is the density of v_2. Senior school enrolment rises as required. However, variability of the parental income y_3 is irrelevant to the senior school progression decision if the individualistic approach to the child's earnings as adult is appropriate.

Theorem 2:

An increasing mean-preserving spread in the in the distribution $H_\zeta(\zeta_3)$ of the incremental component of the child's period 3 income ζ_3 (i.e. income as adult), conditional on enrolment in senior school, decreases enrolment in senior school.

Proof:

Consider a mean-preserving spread $\zeta_3 \rightarrow \zeta_3 + \theta_\zeta(\zeta_3 - \zeta^*)$. The enrolment threshold $\eta(a, y^p, \theta_\zeta)$ is now defined by

$$C(\eta(a, y^p, \theta_\zeta)) = B(a, y^p, \theta_\zeta)$$
$$= \delta E_2\left[u\left(y^p + \tilde{v}_3 + w_3 + \tilde{\xi}_3 + (1+\theta_\zeta)\tilde{\zeta}_3 + a - \theta_\zeta\bar{\zeta}\right) - u\left(y^p + \tilde{v}_3 + w_3 + \tilde{\xi}_3\right)\right]$$

From Lemma 3, $\dfrac{\partial B(a, y^p, \theta_\zeta)}{\partial \theta_\zeta} < 0$. It follows that

$\dfrac{\partial \eta(a, y^p, \theta_\zeta)}{\partial \theta_\zeta} = \dfrac{\partial \lambda(a, y^p, \theta_\zeta)}{\partial \theta_\zeta} = \dfrac{1}{C'(y_1)} \dfrac{\partial B(a, y^p, \theta_\zeta)}{\partial \theta_\zeta} > 0$ implying that an increase in period 3 income dispersion increases the enrolment threshold for all levels of ability a. Hence

$$\frac{d\pi^s}{d\theta_\zeta} = -\int_{\underline{y}}^{\hat{y}}\int_0^{\bar{a}} g_2\left(\eta(a, y^p, \theta_\zeta)\right)\frac{\partial \eta(a, y^p, \theta_\zeta)}{\partial \theta_\zeta} dF(a, y^p) < 0 \qquad \text{(A.20)}$$

Senior school enrolment falls as required.

Comment:

The result set out in Theorem 2 is the standard result that income uncertainty reduces the incentive to invest. In this case, the relevant income ζ_3 arises directly out of the investment (school enrolment) decision. The results demonstrated in Theorem 1, which go in the opposite direction, have a different origin. In these cases, increased income uncertainty affects the family's income in the same way irrespective of the decision to enrol the child in senior school. Uncertainty with respect to the child's incremental income ξ_3, conditional on quitting after junior school, reduces the attractiveness of quitting. Uncertainty with respect to parental income y_3 reduces the family's expected utility. The child's income as adult compensates for this decline in utility following an adverse income shock. This income diversification effect works to increase senior school enrolment.

The junior school enrolment decision

We now return to period 1 and consider start the junior school enrolment decision commencing with the impact of uncertainty with respect to the period 3 earnings distribution functions $H_\xi(\xi_3)$, $H_\zeta(\zeta_3)$ and $G_3(y_3)$. We first prove a pair of lemmas.

Lemma 4:

Consider a set of mean-preserving spreads applied in turn to the distributions of ξ_3, ζ_3, v_2 and v_3. The derivatives of the mean-preserving spreads on the expected benefit $K(y^p)$ defined in equation (A.8) with respect to the distribution $H_\zeta(\zeta_3)$ is zero, that with respect to the distribution $H_\xi(\xi_3)$ is negative and those with respect to $G_2(v_2)$ and $G_3(v_3)$ are positive.

Proof:

1. First consider a mean-preserving spread $\zeta_3 \to \zeta_3 + \theta_\zeta(\zeta_3 - \bar{\zeta})$. We see from (A.8) that $K(y^p)$ is independent of ζ_3. It follows that $\dfrac{\partial K(y^p)}{\partial \theta_\zeta} = 0$.

2. Consider first a mean-preserving spread $\xi_3 \to \xi_3 + \theta_\xi(\xi_3 - \bar{\xi})$. From equation (A.8),

$$\frac{\partial K(y^p, \theta_\xi)}{\partial \theta_\xi} = \gamma\delta E\left[\left(u'\left(y^p + \tilde{v}_3 + (1+\theta_\xi)\tilde{\xi}_3 + w_3 - \theta_\xi\bar{\xi}\right)\right)\left(\tilde{\xi}_3 - \bar{\xi}\right)\right] < 0$$

3. Now consider a mean-preserving spread $v_3 \to (1+\theta_3)v_3$. Again, from equation (A.9) and using assumption (A6),[17]

$$\frac{\partial K(y^p, \theta_3)}{\partial \theta_3} = \gamma\delta E\left[\left(u'\left(y^p + (1+\theta_3)\tilde{v}_3 + \tilde{\xi}_3 + w_3 - \theta_3\bar{\xi}\right) - u'\left(y^p + (1+\theta_3)\tilde{v}_3 + w_3\right)\right)\tilde{v}_3\right] > 0$$

[17] This derivative is zero in the individualistic case.

4. Finally, consider a mean-preserving spread $v_2 \to (1+\theta_2)v_2$. From equation (A.8) and using assumption (A6),

$$\frac{\partial K(y^p,\theta_2)}{\partial \theta_2} = \gamma\Big[\Big(u'\big(y^p + (1+\theta_2)\tilde{v}_2 + \xi_2 + w_2\big) - u'\big(y^p + (1+\theta_2)\tilde{v}_2 + w_2\big)\Big)\tilde{v}_2\Big] > 0$$

<u>Lemma 5:</u>

Consider a set of mean-preserving spreads applied in turn to the distributions of ξ_3, ζ_3, v_2 and v_3. The derivatives of the mean-preserving spreads on the expected benefit $L(a,y^p)$ defined in equation (A.10) with respect to the distributions $H_\xi(\xi_3)$ and $G_3(v_3)$ are positive, that with respect to $H_\zeta(\zeta_3)$ is negative and that with respect to $G_2(v_2)$ cannot be signed *a priori*.

<u>Proof:</u>

1. First consider a mean-preserving spread $\zeta_3 \to \zeta_3 + \theta_\zeta(\zeta_3 - \bar{\zeta})$. Using equation (10), differentiation of $L(a,y^p,\theta_\zeta)$ with respect to the spread parameter θ_ζ, we obtain

$$\frac{\partial L(a,y^p,\theta_\zeta)}{\partial \theta_\zeta} = \gamma E\int_{\eta(a,y^p,\theta_\zeta)}^{\hat{v}} \frac{\partial D(a,y^p,\tilde{v}_2,\theta_\zeta)}{\partial \theta_\zeta} dG_2(\tilde{v}_2) - \gamma E\big[D(a,y^p,\eta,\theta_\zeta)\big] g_2(\eta)\frac{\partial \eta(a,y^p,\theta_\zeta)}{\partial \theta_\zeta} \tag{A.21}$$

The final term in equation (A.21) is zero from equation (A.3). The initial term simplifies to give

$$\frac{\partial L(a,y^p,\theta_\zeta)}{\partial \theta_\zeta} = \gamma\Big[1 - G_2\big(\eta(a,y^p,\theta_\zeta)\big)\Big] E\left[\frac{\partial B(a,y^p,\theta_\zeta)}{\partial \theta_\zeta}\right] \tag{A.22}$$

From Lemma 3, $\frac{\partial B(a,y^p,\theta_\zeta)}{\partial \theta_\zeta} < 0$, implying $\frac{\partial L(a,y^p,\theta_\zeta)}{\partial \theta_\zeta} < 0$.

2. Consider first a mean-preserving spread $\xi_3 \to \xi_3 + \theta_\xi(\xi_3 - \bar{\xi})$. Analogously with equations (A.21) and (A.22)

$$\frac{\partial L\left(a,y^{p},\theta_{\xi}\right)}{\partial\theta_{\xi}}=\gamma\left[1-G_{2}\left(\eta\left(a,y^{p},\theta_{\xi}\right)\right)\right]E\left[\frac{\partial B\left(a,y^{p},\theta_{\xi}\right)}{\partial\theta_{\xi}}\right] \tag{A.23}$$

From Lemma 3, $\frac{\partial B\left(a,y^{p},\theta_{\xi}\right)}{\partial\theta_{\xi}}>0$. Hence $\frac{\partial L\left(a,y^{p},\theta_{\xi}\right)}{\partial\theta_{\xi}}>0$.

3. Now consider a mean-preserving spread $v_3 \rightarrow (1+\theta_3)v_3$. The proof is identical to that of the mean-preserving spread in ξ_3 and establishes $\frac{\partial L\left(a,y^{p},\theta_{3}\right)}{\partial\theta_{3}}>0$.[18]

4. Finally consider a mean-preserving spread $v_2 \rightarrow (1+\theta_2)v_2$. Here we follow a different strategy in acknowledgement of the fact that $\eta\left(a,y^{p}\right)$ is unaffected by changes in the distribution $G_2(v_2)$. Noting that the condition $(1+\theta_2)v_2 \geq \eta$ implies , $v_2 \geq \frac{\eta}{1+\theta_2}$, we may rewrite equation (11) as

$$L\left(a,y^{p},\theta_{2}\right)=\gamma E\int_{\eta/(1+\theta_2)}^{\hat{v}} D\left(a,y^{p},(1+\theta_2)\tilde{v}_2\right)dG_2(\tilde{v}_2)$$

Differentiation with respect to θ_2 gives

$$\frac{\partial L\left(a,y^{p},\theta_{2}\right)}{\partial\theta_{2}}=-\gamma E\left[D\left(a,y^{p},\eta\right)\right]-\gamma\int_{\eta/(1+\theta_2)}^{\hat{v}} C'\left(y^{p}+(1+\theta_2)\tilde{v}_2\right)\tilde{v}_2 dG_2(\tilde{v}_2) \tag{A.24}$$

The initial term in expression (A.24) is zero, using equation (A.3). Define $\bar{v}(\theta_2)=\frac{1}{1-G_2\left(\eta/(1+\theta_2)\right)}\int_{\eta/(1+\theta_2)}^{\hat{v}}\tilde{v}_2 dG_2(\tilde{v}_2)$, the conditional mean of v_2. Then

$$\frac{\partial L\left(a,y^{p},\theta_{2}\right)}{\partial\theta_{2}}\simeq-\gamma(1+\theta_2)\left[\left(1-G_2\left(\frac{\eta}{1+\theta_2}\right)\right)C'\left(y^{p}\right)\bar{v}_2+C''\left(y^{p}\right)\int_{\eta/(1+\theta_2)}^{\hat{v}}\left(\tilde{v}_2-\bar{v}_2\right)^2 dG_2(\tilde{v}_2)\right] \tag{A.25}$$

[18] This derivative is zero in the individualistic case.

The first term in square brackets in expression (A.25) is positive and the second negative. It follows that $\frac{\partial L\left(a, y^{p}, \theta_{2}\right)}{\partial \theta_{2}}$ cannot be signed *a priori*.

Theorem 3:

An increase in the dispersion of the distribution $H_{\zeta}\left(\zeta_{3}\right)$ of the child's period 3 incremental income as adult ζ_3, conditional on attending junior and senior school, decreases junior school enrolment.

Proof:

Consider a mean-preserving spread with respect to ζ_3. Using equation (A.7) and the result from Lemma 4 that $K\left(y^{p}\right)$ is independent of θ_{ζ}, the threshold $\mu\left(a, y^{p}, \theta_{\zeta}\right)$ solves

$$J\left(\mu\left(a, y^{p}, \theta_{\zeta}\right)\right)=K\left(y^{p}\right)+L\left(a, y^{p}, \theta_{\zeta}\right)$$

Hence $$\frac{\partial \mu\left(a, y^{p}, \theta_{\zeta}\right)}{\partial \theta_{\zeta}}=\frac{1}{J'} \frac{\partial L\left(a, y^{p}, \theta_{\zeta}\right)}{\partial \theta_{\zeta}}$$

From Lemma 5, $\frac{\partial L\left(a, y^{p}, \theta_{\zeta}\right)}{\partial \theta_{\zeta}}<0$. It follows that $\frac{\partial \mu\left(a, y^{p}, \theta_{\zeta}\right)}{\partial \theta_{\zeta}}>0$. An increase in the dispersion of the child's period 3 incremental income as adult ζ_3, conditional on attending both junior and senior school, therefore decreases junior school enrolment.

Comment:

This is the standard result that increased income variability reduces the incentive to invest in education. It extends Theorem 2 which established the same result for senior school enrolment.

Theorem 4:

An increase in the dispersion of the distribution $H_{\xi}\left(\xi_{3}\right)$ of the child's period 3 incremental income ξ_3, conditional on the child having attended junior but not senior school, can either increase or decrease junior school enrolment.

Proof:

Consider the mean-preserving spread with respect to ξ_3. Using equation (A.7), the threshold $\mu\left(a, y^p, \theta_\xi\right)$ solves $J\left(\mu\left(a, y^p, \theta_\xi\right)\right) = K\left(y^p, \theta_\xi\right) + L\left(a, y^p, \theta_\xi\right)$

Hence $$\frac{\partial\mu\left(a, y^p, \theta_\xi\right)}{\partial\theta_\xi} = \frac{1}{J'}\left[\frac{\partial K\left(y^p, \theta_\xi\right)}{\partial\theta_\xi} + \frac{\partial L\left(a, y^p, \theta_\xi\right)}{\partial\theta_\xi}\right]$$

From Lemmas 4 and 5 respectively, $\frac{\partial K\left(y^p, \theta_\xi\right)}{\partial\theta_\xi} > 0$ and $\frac{\partial L\left(a, y^p, \theta_\xi\right)}{\partial\theta_\xi} < 0$. The overall effect is therefore ambiguous – an increase in the dispersion of the child's period 3 incremental income ξ_3 can either increase or decrease junior school enrolment.

Comment:

Two opposing effects are at work. Increased variability of the child's incremental income ξ_3, conditional on attending junior but not senior school, reduces the utility benefit of junior school enrolment. This is reflected in the derivative $\frac{\partial K\left(y^p, \theta_\xi\right)}{\partial\theta_\xi}$.

The second effect comes through the option to continue to secondary school. Increased variability of ξ_3 increases the utility value of the income stream arising from secondary education which increases the value of the continuation option and hence also the incentive to enroll in junior school. This is reflected in the derivative $\frac{\partial L\left(a, y^p, \theta_\xi\right)}{\partial\theta_\xi}$. The overall effect of increased variability in ξ_3 is therefore to reduce the attractiveness of junior school *per se* but to increase the attractiveness of eventual enrolment in senior school, for which graduation from junior school is a precondition.

Theorem 5:

An increase in the dispersion of the distribution $G_3\left(v_3\right)$ of parental period 3 transitory income v_3 increases junior school enrolment.

Proof:

Consider a mean-preserving spread with respect to v_3. Using equation (A.7), the threshold $\mu\left(a, y^p, \theta_3\right)$ solves $J\left(\mu\left(a, y^p, \theta_3\right)\right) = K\left(y^p, \theta_3\right) + L\left(a, y^p, \theta_3\right)$

Hence $$\frac{\partial \mu\left(a, y^p, \theta_3\right)}{\partial \theta_3} = \frac{1}{J'}\left[\frac{\partial K\left(y^p, \theta_3\right)}{\partial \theta_3} + \frac{\partial L\left(a, y^p, \theta_3\right)}{\partial \theta_3}\right]$$

From Lemmas 4 and 5 respectively, $\frac{\partial K\left(y^p, \theta_3\right)}{\partial \theta_3} > 0$ and $\frac{\partial L\left(a, y^p, \theta_3\right)}{\partial \theta_3} > 0$. It follows that $\frac{\partial \mu\left(a, y^p, \theta_3\right)}{\partial \theta_3} < 0$. An increase in the dispersion of the parents' period 3 transitory income v_3 increases junior school enrolment. (This effect is absent if we adopt the individualistic approach to the child's earnings as adult).

Theorem 6:

An increase in the dispersion of the distribution $G_2\left(v_2\right)$ of parental period 2 transitory income v_2 may either increase or decrease junior school enrolment.

Proof:

Consider a mean-preserving spread with respect to v_2. Using equation (A.7), the threshold $\mu\left(a, y^p, \theta_2\right)$ solves $J\left(\mu\left(a, y^p, \theta_2\right)\right) = K\left(y^p, \theta_2\right) + L\left(a, y^p, \theta_2\right)$

Hence $$\frac{\partial \mu\left(a, y^p, \theta_2\right)}{\partial \theta_2} = \frac{1}{J'}\left[\frac{\partial K\left(y^p, \theta_2\right)}{\partial \theta_2} + \frac{\partial L\left(a, y^p, \theta_2\right)}{\partial \theta_2}\right]$$

From Lemmas 4 and 5 respectively, $\frac{\partial K\left(y^p, \theta_2\right)}{\partial \theta_2} > 0$ while the sign of $\frac{\partial L\left(a, y^p, \theta_2\right)}{\partial \theta_2}$ is uncertain. It follows that the sign of $\frac{\partial \mu\left(a, y^p, \theta_2\right)}{\partial \theta_2}$ is also uncertain. An increase in the dispersion of the parents' period 2 transitory income v_2 may either increase or decrease junior school enrolment.

Comment:

There are three effects here:

1. For a child who is not expected to proceed to senior school, the more uncertain are his parents' period 2 earnings ν_2 , the more valuable will be the child's earnings. This effect is reflected in the derivative $\dfrac{\partial K\left(y^p, \theta_2\right)}{\partial \theta_2}$ and is the same result as obtained in Theorem 5 with respect to the variability of parental transitory income in period 3.

2. The second effect arises out of the option to continue to senior school. On average families that send children to senior school will have positive period 2 transitory incomes. This effect is reflected in the first term in square brackets in equation (A.25). In the event of a high realization for period 2 transitory income, the utility cost of senior school education will be relatively low and families will choose to exercise this option and thereby benefit from the cost saving. Conversely, in the event of a low realization, the utility cost of senior school education will be relatively high but families will choose not to exercise and will therefore not suffer these higher costs. Families capture the benefit from of low education costs without suffering the pain of high costs. Controlling for permanent income and the child's earning ability, utility costs of senior school education will be lower for those families who do enroll their children than for those who do not. This potential benefit increases the attractiveness of junior school enrolment.[19]

3. The third effect arises from convexity of the secondary school cost function $C\left(y_2\right)$. This convexity implies that, for those families which do enroll their children, the expected cost of senior school education rises with the variability of parental period 2 transitory income thereby making senior school enrolment less attractive. This effect is reflected in the second term in square brackets in equation (A.25). It is not possible to state a priori whether this effect will be sufficiently large to offset the two preceding effects.

<u>Theorem 7:</u>

[19] Because this is a pure option effect, it does not depend on the curvature of the utility function and would be present even if utility were linear in income. In that circumstance, this would be the only route by which income uncertainty impact on the education decision.

An increase in the variability of the period 1 transitory income distributions $G_1(v_1)$ has an ambiguous effect on junior school enrolment. An increase in the variability of period 2 transitory income distribution $G_2(v_2)$ has an ambiguous effect on senior school enrolment.

Proof:

Consider a mean-preserving spread with respect to v_1. From equation (A.14)

$$\pi^j(\theta_1) = 1 - \int_{\underline{y}}^{\hat{y}} \int_0^{\hat{a}} \int_0^{\mu(a,y^p)} dG_1^r(r_1)\, dF(a, y^p)$$

where $r_1 = (1+\theta_1)v_1$. Rewriting the integral in terms of v_1, we obtain

$$\begin{aligned} \pi^j(\theta_1) &= 1 - \int_{\underline{y}}^{\hat{y}} \int_0^{\hat{a}} \int_0^{\mu(a,y^p)/(1+\theta_1)} dG_1(v_1)\, dF(a, y^p) \\ &= 1 - \int_{\underline{y}}^{\hat{y}} \int_0^{\hat{a}} \left[1 - G_1\left(\frac{\mu(a, y^p)}{1+\theta_1}\right)\right] dF(a, y^p) \end{aligned}$$

Differentiating with respect to the spread parameter θ_1,

$$\frac{d\pi^j(\theta_1)}{d\theta_1} = \frac{1}{(1+\theta_1)^2} \int_{\underline{y}}^{\hat{y}} \int_0^{\hat{a}} \mu(a, y^p)\, g_1\left(\frac{\mu(a, y^p)}{1+\theta_1}\right) dF(a, y^p) \qquad \text{(A.28)}$$

Since the threshold income $\mu(a, y^p)$ may exceed or fall short of permanent income y^p, it is not possible to sign expression (A.28) on an *a priori* basis. The result for the impact of a change in the variability of the period 2 transitory income distribution $G_2(v_2)$ on senior school enrolment π^s follows from question (2) on an identical argument.

Linear approximation

To obtain estimable functions, we need to make linear approximations to the model. First consider the conditional decision to attend senior school for a child belonging to household *h* with vector of characteristics χ_h. The condition for this is given by equation (1):

$$B\left(a_h, y_h^p\right) - C\left(y_{h2}\right) \geq 0$$

The household knows the cost $C\left(y_{h2}\right)$ with certainty in period 2. Approximating around the population mean income, $\overline{y}_2$, we obtain

$$\begin{aligned} C\left(y_{h2}\right) &= u\left(y_{h2} + x_{h2}\right) - u\left(y_{h2} - q_h\right) \\ &\simeq \left(x_{h2} + q_h\right) u'\left(y_{h2}\right) \simeq \left(x_{h2} + q_h\right)\left[1 - \rho \frac{y_{h2} - \overline{y}_2}{\overline{y}_2}\right] u'\left(\overline{y}_2\right) \end{aligned} \tag{A.29}$$

where $\rho = -\dfrac{\overline{y}_2 u''\left(\overline{y}_2\right)}{u'\left(\overline{y}_2\right)}$, the coefficient of relative risk aversion, which we suppose to be constant across both time and households. On the benefit side, also from equation (A.1)

$$\begin{aligned} B\left(a_h, y_h^p\right) &= \delta E\left[u\left(\tilde{y}_{h3} + \tilde{z}_{h3}\right) - u\left(\tilde{y}_{h3} + \tilde{x}_{h3}\right)\right] \\ &= \delta E\left[u\left(y_h^p + \tilde{v}_{h3} + w_{h3} + \tilde{\xi}_{h3} + \tilde{\zeta}_{h3} + a_h\right) - u\left(y_h^p + \tilde{v}_{h3} + w_{h3} + \tilde{\xi}_{h3}\right)\right] \end{aligned}$$

Proceeding in the same manner

$$\begin{aligned} B\left(a_h, y_h^p\right) &= \delta\left[\left(\overline{\zeta}_{h3} + a_h\right) - \tfrac{1}{2}\rho\sigma_{h\zeta}^2\right] u'\left(y_h^p\right) \\ &\simeq \delta\left[1 - \rho \frac{y_h^p - \overline{y}^p}{\overline{y}^p}\right]\left[\left(\overline{\zeta}_{h3} + a_h\right) - \tfrac{1}{2}\rho y_h^p \sigma_{h\zeta}^2\right] u'\left(\overline{y}^p\right) \end{aligned} \tag{A.30}$$

Here $\sigma_{h\zeta}^2 = \dfrac{Var\left(\zeta_{h3} \mid \chi_h\right)}{\left(\overline{y}_h^p\right)^2}$, the squared conditional coefficient of variation of ζ_{h3}.

Note that equation (A.30) lacks a term in $\sigma_{h\xi}^2$ which would depend on the third derivative of the utility function. Now make the assumption that $\overline{y}_2 = \overline{y}^p$, i.e. the period 2 common income shocks are zero. Using the approximations (A.29) and (A.30), condition (A.1) implies that the child will attend senior school if

$$\begin{aligned} \delta a_h &\geq \left(x_{h2} + q_h\right) \frac{1 - \rho \dfrac{y_{h2} - \overline{y}_2}{\overline{y}_2}}{1 - \rho \dfrac{y_h^p - \overline{y}^p}{\overline{y}^p}} - \delta\overline{\zeta}_{h3} + \tfrac{1}{2}\delta\rho y_h^p \sigma_{h\zeta}^2 \\ &\simeq \left(1 - \rho \frac{v_{h2}}{\overline{y}_2}\right)\left(x_{h2} + q_h\right) - \delta\overline{\zeta}_{h3} + \tfrac{1}{2}\delta\rho y_h^p \sigma_{h\zeta}^2 = \delta\hat{a}_h \end{aligned} \tag{A.31}$$

Equation (A.31) tells us that the conditional probability of attending senior school depends

- negatively on the total cost of schooling $x_{h2} + q_h$;
- positively on the interaction of this cost with the period 2 income shock v_{h2};
- positively on the child's expected earnings $\overline{\zeta}_{h3}$ as adult; and
- negatively on the squared conditional coefficient of variation of the child's income as adult $\sigma^2_{h\zeta}$ (interacted with permanent income).

If we make a distributional assumption for the child's earning ability a_h, we may estimate equation (A.31) using standard discrete choice methods. The most straightforward assumption is normality, $a_h \sim N\left(\overline{a}, \omega^2\right)$. On this basis, write the estimated conditional probability that the child does not attend senior school as $\Phi_h = \Phi\left(\frac{\delta\left(\hat{a}_h - \overline{a}\right)}{\omega}\right)$ where $\Phi(.)$ is the standard normal distribution function.

We turn now to the decision to attend junior school given by equation (A.7)

$$K\left(y_h^p\right) + L\left(a_h, y_h^p\right) - J\left(y_{h1}\right) \geq 0$$

The expansion of $J\left(y_{h1}\right)$ follows the same lines as that of $C\left(y_{h2}\right)$ in equation (A.29).

$$\begin{aligned} J\left(y_{h1}\right) &= u\left(y_{h1} + w_h\right) - u\left(y_{h1} - p_h\right) \\ &\simeq \left(w_{h1} + p_h\right) u'\left(y_{h1}\right) \simeq \left(w_{h1} + p_h\right)\left[1 - \rho\frac{y_{h1} - \overline{y}_1}{\overline{y}_1}\right] u'\left(\overline{y}_1\right) \end{aligned} \tag{A.32}$$

Similarly, the expansion of $K\left(y_h^p\right)$ follows that of $B\left(a_h, y_h^p\right)$ in equation (A.30):

$$\begin{aligned} K\left(y_h^p\right) &= \gamma E\left[\left(u\left(\tilde{y}_2 + \tilde{x}_2\right) - u\left(\tilde{y}_2 + w_2\right)\right) + \delta\left(u\left(\tilde{y}_3 + \tilde{x}_3\right) - u\left(\tilde{y}_3 + w_3\right)\right)\right] \\ &= \gamma\left[\xi_{h2} + \delta\left(\overline{\xi}_{h3} - \tfrac{1}{2}\rho y_h^p \sigma^2_{h\xi}\right)\right] u'\left(y_h^p\right) \\ &= \gamma\left(1 - \rho\frac{y_h^p - \overline{y}^p}{\overline{y}^p}\right)\left[\xi_{h2} + \delta\left(\overline{\xi}_{h3} - \tfrac{1}{2}\rho y_h^p \sigma^2_{h\xi}\right)\right] u'\left(\overline{y}^p\right) \end{aligned}$$

(A.33)

The expansion of the final term, $L\left(a_h, y_h^p\right)$, is more complicated. Define $\hat{y}_{h2} = E\left[y_{h2} \mid y_{h2} \geq \eta\left(a_h, y_h^p\right)\right]$. From equation (A.9)

$$L\left(a_h, y_h^p\right) = \gamma\left(1-\Phi_h\right) E \int_{\eta\left(a_h, y_h^p\right)}^{\hat{v}} \left\{\left(u\left(\tilde{y}_{h2} - q_h\right) - u\left(\tilde{y}_{h2} + \tilde{x}_{h2}\right)\right)\right.$$
$$\left. + \delta\left(u\left(\tilde{y}_{h3} + \tilde{z}_{h3}\right) - u\left(\tilde{y}_{h3} + \tilde{x}_{h3}\right)\right)\right\} dG_2\left(\tilde{y}_{h2} - y_h^p \mid y_{h2} \geq \eta\left(a_h, y_h^p\right)\right)$$

Expanding the initial term around $\hat{y}_{h2}$ and the second term around y_h^p

$$L\left(a_h, y_h^p\right) = \gamma\left(1-\Phi_h\right)\left[-\left(q_h + w_{h2} + \xi_{h2}\right)u'\left(\hat{y}_{h2}\right) + \delta\left(a_h + \overline{\zeta}_{h3} - \tfrac{1}{2}\rho y_h^p \sigma_{h\zeta}^2\right)u'\left(y_h^p\right)\right]$$

Now $u'\left(\hat{y}_{h2}\right) \simeq \left[1 - \rho\dfrac{\hat{y}_{h2} - y_h^p}{y_h^p}\right]u'\left(y_h^p\right) \simeq \left(1 - k\rho\sigma_{h2}\right)u'\left(y_h^p\right)$ for some $k > 0$ which will depend on the distribution of the period 2 income shocks v_{h2}. Hence

$$L\left(a_h, y_h^p\right) \simeq \gamma\left(1-\Phi_h\right)\left[-\left(1-k\rho\sigma_{h2}\right)\left(q_h + w_{h2} + \xi_{h2}\right) + \delta\left(a_h + \overline{\zeta}_{h3} - \tfrac{1}{2}\rho y_h^p \sigma_{h\zeta}^2\right)\right]u'\left(y_h^p\right)$$
$$= \gamma\left(1-\Phi_h\right)\left(1 - \rho\frac{y_h^p - \overline{y}^p}{\overline{y}^p}\right)\left[-\left(1-k\rho\sigma_{h2}\right)\left(q_h + w_{h2} + \xi_{h2}\right) + \delta\left(a_h + \overline{\zeta}_{h3} - \tfrac{1}{2}\rho y_h^p \sigma_{h\zeta}^2\right)\right]u'\left(\overline{y}^p\right) \tag{A.34}$$

Equation (A.34) lacks a term in σ_{h3}^2 which would depend on the third derivative of the utility function. Combining equations (A.7), (A.32), (A.33) and (A.34), the child will attend junior school if

$$\begin{aligned} \gamma\delta a_h \geq\ & \gamma\left(1-k\rho\sigma_{h2}\right)\left(q_h + w_{h2} + \xi_{h2}\right) - \gamma\delta\left(\overline{\zeta}_{h3} - \tfrac{1}{2}\rho y_h^p \sigma_{h\zeta}^2\right) \\ & - \frac{1}{1-\Phi_h}\left[\gamma\left(\xi_{h2} + \delta\left(\overline{\xi}_{h3} - \tfrac{1}{2}\rho y_h^p \sigma_{h\xi}^2\right)\right) + \left(1 - \rho\frac{v_{h1}}{\overline{y}_1}\right)\left(w_{h1} + p_h\right)\right] \end{aligned} \tag{A.35}$$

Equation (A.35) tells us that the conditional probability of attending junior school depends

- negatively on the total cost of attending junior school $w_{h1} + p_h$, divided by the probability $1 - \Phi_h$ of continuing to senior school;
- positively on the interaction of this cost with the period 1 income shock v_{h1};
- negatively on the total cost of attending senior school $q_h + w_{h2} + \xi_{h2}$;
- positively on the interaction of this cost with the coefficient of variation of the parents' period 2 income;

- positively on the total discounted expected income from attending junior school $\gamma\xi_{h2} + \gamma\delta\overline{\xi}_{h3}$ divided the probability $1-\Phi_h$ of continuing to senior school;
- negatively on the squared conditional coefficient of variation of the child's income as adult $\sigma^2_{h\xi}$, conditional on not attending senior school, interacted with permanent income and divided the probability $1-\Phi_h$ of continuing to senior school;
- positively on the child's expected earnings $\overline{\zeta}_{h3}$ as adult conditional on attending senior school;
- negatively on the squared conditional coefficient of variation of the child's income as adult $\sigma^2_{h\zeta}$, conditional on attending secondary school, interacted with permanent income.

APPENDIX B – MEASURES OF RISK AND SHOCKS

As discussed in section 4, income shock at time t is defined as the difference between the actual household income y_{ht} at time t and the household's permanent income y_{ht}^{P} defined as expected household income, given all the information available at time t. Table B.1 reports the results of equation (1) that estimates permanent income.

Table B.1
1993 income equation estimates

Variables	*Coeff.*	*t*
1992 farm assets	0.01	1.32
1992 business non-farm assets	0.02	2.30
1992 liquid non business assets	0.03	1.23
1992 illiquid non business assets	0.02	4.30
Hh owns a farm	-21.48	-0.22
Hh owns a business	517.98	7.19
Head does not work	598.34	1.95
Head employee	907.40	3.25
Head self employed	245.84	0.89
Head complete primary educ	189.01	3.52
Head junior secondary educ	946.26	6.03
Head senior secondary educ	1711.07	8.02
Head high educ	2884.97	6.30
Nr. of income earner	111.68	3.22
Household size	36.60	0.46
Household size^(2)	1.57	0.22
1993 crop loss	-43.70	-0.41
Intercept	-255.60	-0.69
Number of obs= 2296		
R-squared= 0.30		

The table records the results from equation (1) and estimates the predicted household permanent income. Dependent variable is 1993 parental income. Provincial dummies are included as additional regressors. Both income and assets are measured in thousands of rupiah. Standard errors are robust.

Uncertainty about parents' transitory income ν_{h2} is measured as variance of period 2 income conditional on all the information available at time 1. As discussed in section 4, a two stage procedure is used. The first step regress future household income on current household characteristics (equation 2). Table B.2 reports the results of equation (2), where the dependent variable is 1997 household income and independent variables are 1993 household characteristics.

Table B.2
1997 income equation estimates

Variables	*Coeff.*	*t*
1992 farm assets	0.01	1.32
1992 business non-farm assets	0.05	1.30
1992 liquid non business assets	0.05	3.23
1992 illiquid non business assets	0.02	4.32
Hh owns a farm	-78.11	-0.83
Hh owns a business	354.97	4.66
Head does not work	-151.66	-0.58
Head employee	-217.07	-0.92
Head self employed	-112.22	-0.49
Head complete primary educ	225.69	3.44
Head junior secondary educ	620.03	3.39
Head senior secondary educ	561.15	4.18
Head high educ	1502.17	3.01
Nr. of income earner	43.46	1.30
Household size	4.06	0.06
Household size^(2)	-0.51	-0.09
1993 crop loss	-116.05	-1.14
Intercept	633.66	2.00
Number of obs= 1870		
R-squared= 0.13		

The table records the results from equation (2) and estimates the first stage of the procedure to calculate the variance of period 2 household income. Dependent variable is 1997 household income. Provincial dummies are included as additional regressors. Both income and assets are measured in thousands of rupiah. Standard errors are robust.

APPENDIX C - INCOME LEVELS BY EDUCATION

Table C.1 reports the results of the selection model (Heckman procedure) that estimates equation (4)[20]. Dummies for different levels of completed schooling are included, with no schooling acting as the omitted category (i.e. there is no constant term in the regression equation). Off 5982 females and 5534 males, we observe the log monthly income for 1437 and 2636 working individuals (females and males respectively). The large proportions of zeros, especially for the self-employed sector, suggests that monthly income may contain an important seasonal component. To control for this, the selection equation includes the dummies for the month of interview. Results are in line with standard income selection equation estimates. The impact of experience is greater for females than for males, but females have also a higher negative coefficient on experience squared. Returns to education are higher for men than for women, and they are highest in West Java (where Jakarta is located) and (for men) in the province of South Kalimantan.

The coefficients on education and experience estimated from the income regression are used to construct predicted earnings by gender, provinces and school level[21]

[20] For females, we cannot reject the hypothesis that coefficients on dummies for w_1, w_2, w_3, x_2, x_3 are equal (F(4, 1426) = 1.41, Prob>F = 0.23). We therefore generate a single dummy for females graduated from elementary and junior secondary school, for any experience level. This suggests that it is only senior secondary school attendance that substantially increases female income.

[21] We do not differentiate between income of children working on the family farm and income from wage employment. In Indonesia, as in many developing countries, most children do not work for a wage, but as unpaid family workers in the family farm or business. Moreover, in many developing countries, labour markets are imperfect. It is therefore difficult to infer children's contributions to household income by using observed market wages and self-reported farm or business profits (Menon and Perali, 2006). A shadow wage approach may be more suitable in evaluating the income contribution of child labour, and we will consider this approach in future extensions of this work.

Table C.1
Individual income equation: Heckman selection model

		Female	Male
Dependent variable: log income (robust se in parenthesis)			
Male: dummy elementary school*provincial dummies Female: dummy elementary or junior secondary school*provincial dummies	North Sumatra	1.913*** (0.437)	4.160*** (0.264)
	West Sumatra	1.731*** (0.401)	4.480*** (0.243)
	South Sumatra	1.846*** (0.522)	3.665*** (0.257)
	Lampung	1.501** (0.469)	3.716*** (0.258)
	West Java	1.960*** (0.392)	4.013*** (0.210)
	Central Java	1.525*** (0.373)	3.803*** (0.242)
	Di Yogyakarta	1.449*** (0.433)	3.862*** (0.294)
	East Java	1.774*** (0.375)	4.192*** (0.230)
	Bali	1.858*** (0.426)	4.193*** (0.255)
	West Nusa Tenggara	1.658*** (0.436)	3.832*** (0.281)
	South Kalimantan	1.785*** (0.446)	4.320*** (0.282)
	South Sulawesi	1.490*** (0.389)	4.194*** (0.330)
Male: dummy junior secondary school*provincial dummies	North Sumatra	-	4.233*** (0.316)
	West Sumatra	-	4.742*** (0.344)
	South Sumatra	-	4.025*** (0.292)
	Lampung	-	4.019*** (0.346)
	West Java	-	4.512*** (0.287)
	Central Java	-	4.612*** (0.292)
	Di Yogyakarta	-	4.075*** (0.351)
	East Java	-	4.668*** (0.246)
	Bali	-	4.569*** (0.494)
	West Nusa Tenggara	-	4.599*** (0.312)
	South Kalimantan	-	4.836*** (0.322)
	South Sulawesi	-	4.596*** (0.492)
Dummy senior secondary school*provincial dummies	North Sumatra	2.930*** (0.417)	4.822*** (0.262)
	West Sumatra	3.211*** (0.322)	4.669*** (0.302)
	South Sumatra	2.454*** (0.408)	4.801*** (0.350)
	Lampung	dropped	4.167*** (0.276)
	West Java	3.912*** (0.408)	4.818*** (0.219)
	Central Java	2.895*** (0.363)	4.672*** (0.245)
	Di Yogyakarta	3.402*** (0.374)	4.571*** (0.334)
	East Java	3.142*** (0.303)	4.684*** (0.217)
	Bali	2.604*** (0.467)	4.749*** (0.233)
	West Nusa Tenggara	3.063*** (0.428)	4.699*** (0.224)
	South Kalimantan	2.991*** (0.383)	5.046*** (0.213)
	South Sulawesi	3.656*** (0.305)	4.888*** (0.219)
Experience		0.096*** (0.013)	0.028* (0.011)
Experience square		-0.001*** (0.000)	-0.0004* (0.000)

	Female	Male
Dependent variable: log income (robust se in parenthesis)		
Selection (robust se in parenthesis)		
Non business assets/1000	-0.002 (0.000)	0.001 (0.000)
married	-0.308*** (0.048)	0.660*** (0.067)
unschooled	-1.607*** (0.271)	-0.977*** (0.177)
incomplete primary	-1.541*** (0.267)	-0.803*** (0.169)
highest grade completed: elem	-1.367*** (0.266)	-0.647*** (0.168)
highest grade completed: jrsec	-1.254*** (0.271)	-0.559** (0.173)
highest grade completed: srsec	-0.421 (0.273)	0.012 (0.175)
exper	0.096*** (0.005)	0.098*** (0.005)
exper2	-0.001*** (0.000)	-0.002*** (0.000)
# household members age 0-5	-0.122*** (0.025)	-0.026 (0.025)
# household members age 6-9	-0.023 (0.030)	-0.083** (0.031)
# household members age 10-14	-0.064* (0.025)	-0.063* (0.025)
# household members age 15-17	-0.093* (0.036)	0.002 (0.034)
Month of interview: September	0.235** (0.086)	0.663*** (0.084)
Month of interview: October	0.237** (0.076)	0.614*** (0.074)
Month of interview: November	0.438*** (0.073)	0.586*** (0.074)
Month of interview: December	0.192** (0.069)	0.227*** (0.068)
Intercept	-0.194 (0.271)	-0.994*** (0.176)
Mills (robust se in parenthesis)		
lambda	0.539** (0.192)	-0.868*** (0.127)
N	*5982*	*5534*
Censored	*4545*	*2898*
Uncensored	*1437*	*2636*

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table reports the results of equation (4). Dependent variable is log monthly income. Sample: all individuals aged 10 years or old with information on labour. Dummies for "illiterate" and "high education" are included in the income regression, with no schooling acting as the omitted category. In the selection equation omitted categories are "high education" for educational dummies and "January" for the month of interview.

APPENDIX D - COST OF SCHOOLING

Total educational expenditure can be written as $E_h = p_{h1}N_{h1} + p_{h2}N_{h2} + p_{h3}N_{i3}$, where p_{hj} is the cost of the school level *j* for the household h, and N_{hj} is the number of children in the household attending education level *j*. Expenditure for the different levels of schooling is not directly observable but may be inferred from regressing total expenditure for education on the number of children attending each school level. The regression also controls for household income and provinces. We use the following model which is additive in a set of three multiplicative components, one for each level of schooling:

$$E_h = y_h^{\eta} e^{\gamma' P_h} \sum_{j=1}^{3} \beta_j N_{hj} \tag{D.1}$$

where E_h is the total educational expenditure of family *h*, y_h is household income, *P* are province dummies, N_{hj} is the number of children attending school level *j*, and β_j is the coefficient associated with N_{hj}. The cost of sending an additional child to school level *j* is given by $y_h^{\hat{\eta}} e^{\hat{\gamma}' P_h} \hat{\beta}_j$. Assuming constant economies of scale, that is assuming linearity of educational expenditure on the number of children attending school[22], the marginal effect is also the average effect, and is our measure of the cost of schooling.

As already noted, equation (D.1) includes household income, thus allowing the cost of schooling to vary across households. According to this formulation, households estimate the cost of the various levels of schooling as their expected expenditures given their incomes. Alternatively, one might view income-related expenditures as discretionary, implying that these expenditure components should be omitted from the cost variables. If this were the case, the estimated cost would become $\bar{y}^{\hat{\eta}} e^{\hat{\gamma}' P_h} \hat{\beta}_j$, where $\bar{y}$ is the sample average income. We construct both

[22] The square number of children attending school is not statistically significant in the educational expenditure equation, thus rejecting the hypothesis of a quadratic relationship between expenditure and the number of children.

estimates of the cost of schooling: the first conditional on household income, the second conditional on average income.

Household income and the number of children attending school may both be endogenous. Potential instruments for household income are the education of the head and the value of non-business assets owned by the household the year before the interview. The numbers of children attending each school level can be instrumented with the predicted count variables estimated using Poisson regression models on the sub-sample of households with at least one child in the reference age group (children aged 7-12 for the elementary school, 13-15 for junior secondary school, and 16-18 for senior secondary school). Because of the high percentage of zeroes in the number of children enrolled in junior and senior secondary school, the count variables are estimated with the zero-inflated Poisson (ZIP) model. Table D.1 reports the results of equation D.1, and Table D.2 summarizes the estimated costs of schooling, by level of school and gender. Both set of estimates, conditional on household and on average income, are presented. As expected, schooling costs are higher for higher education levels (there are school fees at junior and senior levels and many children are obliged to go to school outside the home village). Estimates show a higher cost for girls than for boys, especially for secondary school.

Table D.1
Educational expenditure regression

Number of children attending school (β_j)	**Coeff.**	**t**
# male attending elementary school	130.58	5.60
# female attending elementary school	135.65	5.92
# male attending junior secondary school	206.71	4.47
# female attending junior secondary school	385.00	4.88
# male attending senior secondary school	435.80	5.69
# female attending senior secondary school	794.78	8.62
Income interaction effect (η)	**Coeff.**	**t**
Household income	0.35	4.20

N =1974
R-squared = 0.11

The table reports the IV estimated coefficients from equation (D.1). Dependent variable is the total household annual educational expenditure. Provincial dummies are included in the regression. Household income is instrumented with head's educational dummies and the value of household non-business assets. The number of children in the household attending school are instrumented with the predicted numbers obtained from Poisson and Zero inflated Poisson models. The sample is households with at least one child attending school.

Table D.5
Descriptive statistics of predicted cost of schooling

Variable	*Mean*	*Std. Dev.*	*Min*	*Max*
Conditional on household income				
Male-elementary	105.13	47.90	30.34	355.51
Female-elementary	109.21	49.76	31.52	369.31
Male-junior secondary	166.42	75.83	48.04	562.79
Female- junior secondary	309.94	141.23	89.47	1048.17
Male-senior secondary	350.84	159.87	101.27	1186.48
Female-senior secondary	639.84	291.55	184.69	2163.81
Conditional of average income				
Male-elementary	117.50	29.18	71.30	171.82
Female-elementary	122.06	30.31	74.07	178.48
Male-junior secondary	186.01	46.19	112.87	271.99
Female- junior secondary	346.43	86.04	210.22	506.57
Male-senior secondary	392.14	97.39	237.96	573.41
Female-senior secondary	715.16	177.61	433.97	1045.74
# households=2288				

The table reports the predicted costs of schooling, by level of school and gender. The first box presents household specific estimates that are conditional on household income ($y_h^{\hat{\eta}} e^{\hat{\gamma}' P^h} \hat{\beta}_j$), the second shows the predicted values calculated using the average income ($\overline{y}^{\hat{\eta}} e^{\hat{\gamma}' P^h} \hat{\beta}_j$). These latter estimated varies by province, gender and school level.