UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 494

Does a Diversification Motive Influence Children's School Entry in the Ethiopian Highlands?

Annika Lindskog

March 2011

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Department of Economics
School of Business, Economics and Law at University of Gothenburg
Vasagatan 1, PO Box 640, SE 405 30 Göteborg, Sweden
+46 31 786 0000, +46 31 786 1326 (fax)
www.handels.gu.se info@handels.gu.se

EFMD EQUIS ACCREDITED

Does a diversification motive influence children's school entry in the Ethiopian highlands?

Annika Lindskog
annika.lindskog@economics.gu.se

Household-level diversification of human capital investments is investigated. A simple model is developed, followed by an empirical analysis using 2000-2007 data from the rural Amhara region of Ethiopia. Diversification would imply negative siblings' dependency and be more important in more risk-averse households. Hence it is investigated if older siblings' literacy has a more negative (smaller if positive) impact on younger siblings' school entry in more risk-averse households. Results suggest diversification across brothers, but are not statistically strong, and with forces creating positive sibling dependency dominating over diversification.

Keywords: Diversification, Education, Ethiopia, Uncertainty
JEL Codes: I21, D81, D13

1. Introduction

Investment in education is likely to be the most important investment decision made for most people, and also to have uncertain returns, i.e. as opposed to risky investments with a known variance they are truly unknown,. Still, the mplications of uncertainty for such investment have received little attention in economics. Returns to formal education are uncertain, but the alternative, more learning by doing, is not free of uncertainty either. This makes it difficult to come up with hypotheses about how uncertainty and risk aversion affect investments in education.

A common strategy for dealing with risky returns is diversification. In the rural Amhara region of Ethiopia, and in other rural areas of less-developed countries where there is extensive informal insurance, with parents relying on children for old-age support, diversification could mean some investment in formal education – perhaps directed towards employment in the "modern" sector – and some investment in traditional knowledge, acquired through learning by doing in the household and in the field.[1]

Diversification at the household level implies within-household education inequality, and thus, negative sibling-dependency in education (Lilleør, 2008b). If an older sibling has more education, diversification would mean that a younger one should get less, and spend more time acquiring traditional knowledge. Such diversification should relate to risk-aversion, and be stronger in more risk-averse households.

Are investment in education really affected by a household-level diversification motive? Unique data, that in addition to extensive information about children's schooling has information about risk-preferences of the household head, is used to investigate whether sibling-dependency in education was more negative in households with more risk-averse heads in rural Amhara during 2000-2006, as well as whether diversification took place across all siblings or was gender-specific.

[1] In the literature on child-schooling and labour, investment in traditional knowledge is often considered just child work, expanding current consumption possibilities, but without future rewards. There are a few exeptions though, including Bommier and Lambert (2000) and Lilleør (2008a). Rosenzweig and Wolpin (1985) and Grootaert and Kanbur (1995) also demonstrate the potential usefulness of household and farm-specific knowledge in rural areas of less-developed countries.

Annual school entry probabilities of boys and girls age 6 to 16 are estimated. School entry is analyzed since it, more than education decisions at later stages, is likely to be affected by parental preferences rather than child preferences and ability. Annual school entry probabilities are used so that the full sample of children can be used without problems of censoring. To control for time-constant unobserved parental preferences, a linear probability model with household fixed effects is used.

Total siblings' dependency in education turn out to be positive; hence other forces dominate diversification. The results still suggest diversification across brothers; older brothers' education does not have the same positive impact on boys' school entry in households with the most risk-averse heads. However, the diversification results are not statistically strong.

The next section provides a theoretical framework for the study, including the development of the model used. Section 3 then describes and explains the empirical approach, while Section 4 describes the data and variables while giving some background on the study area and on education in Ethiopia. Section 5 presents and discusses the results, and Section 6 summarizes and draws conclusions.

2. Theoretical framework

Economic theory about education is dominated by human capital theory, according to which people invest in education as long as the marginal benefit exceeds marginal costs. Marginal costs include direct costs as well as opportunity costs, while the marginal benefit mainly consists of increased future income (Becker, 1962; Ben-Porath, 1967). Though the literature on determinants of educational investment in less-developed countries is extensive, it has focused on the cost rather than the benefit side. Poverty and credit constraints, high opportunity costs of child time, and supply-side constraints such as lack of nearby schools or sufficient teachers, are generally considered the main factors keeping children out of school (Jacoby, 1994; Edmonds, 2006; Gitter and Barham, 2007; Orazam and King, 2008; Huisman and Smits, 2009).

2.1 Expected returns and uncertainty of returns to educational investment

According to theory, expected returns to education should matter for educational investment decisions, and empirical evidence from economics subfields suggests that they do. In part, returns to education are determined by the child's ability to transform time spent in school into marketable knowledge and skills. As data on test-scores has become more available, studies have been done on the impact of various school inputs on test-scores and on demand for education, and of test scores on school continuation and future labour-market outcomes (Card and Krueger, 1992; Glewwe, 2002; Glewwe and Kremer, 2006; Hanushek, 2008; Hanushek and Woessmann, 2008; Akresh et al., 2010). At least when the family is not too poor, demand for education has also been shown to respond to regional variations in returns to education (Anderson et al., 2003; Kochar, 2004; Kingdon and Leopold, 2008; Chamarbagwala, 2008),[2] as well as to subjective perceptions about returns to education (Attanasio and Kaufmann, 2009; Jensen, 2010).[3]

However, the effect of uncertainty of returns to education has received very little attention. There are many possible sources of uncertainty: about the quality of education, about the child's ability, health and survival throughout adulthood, about future market returns; and about the child's future filial transfers.

Theoretically, risky returns to education can decrease time in school (Lehvari and Weiss, 1974), but the relative riskiness of more versus less education is what matters (Kodde, 1986). Pouliot (2006) introduces uncertainty into Baland and Robinson's (2000) influential model on child labour, and demonstrates that, without perfect insurance markets, uncertainty could lead to inefficiently low levels of education, even with perfect credit markets and parents able to impose filial transfers from children. Estevan and Baland (2007) consider a particular source of uncertainty – young adult mortality – and again that, without perfect insurance markets, uncertainty could lead to inefficiently low levels of education when parents want filial transfers rather than planning on parental transfers (bequests) to their children.

[2] But Nerman and Owens (2010) find that returns do not determine demand in Tanzania.
[3] That schooling of children in poor households responds less to differences in expected returns than does that of children in richer households is suggestive of the importance of credit constraints and poverty.

There is little empirical evidence on the importance of uncertainty for educational investments, perhaps because uncertainty is impossible to observe and measure. In an Italian sample, Belzil and Leonardi (2007) find that educational investment is negatively related to risk-aversion, with a small but statistically significant effect. Attanasio and Kaufmann (2009) find that perceived subjective employment and wage risks affect the decision to continue into senior high school in urban Mexico, though again the effect is small.

As mentioned earlier, it is not obvious that investments in formal education are (perceived to be) more uncertain than the relevant alternative. In rural Ethiopia the alternative to formal education is learning by doing in the household and on the farm, which is investment in traditional knowledge. But rain-feed agriculture in the Ethiopian highlands is definitely not free of uncertainty. Furthermore, it is possible to view education as insurance, making the individual better able to manage in an uncertain future.[4]

Still, education is more likely to lead to migration and urban employment, with higher probability of unemployment, and less parental control of children. Empirical evidence on returns to education and on unemployment probabilities in Ethiopia is scarce and not always consistent, but to have more than a couple of years of education appears to yield high returns in cities, though not in rural areas – while unemployment rates are higher for the better educated (World Bank, 2005).

Lack of experience with education and with non-agricultural employment in rural Ethiopia may also create subjective uncertainty about the returns expected from education, which ought to be what matters for actual decisions. In the Dominican Republic teenagers living in neighborhoods with few well-educated were found to underestimate returns to education, while provision of correct information increased their schooling (Jensen, 2010). Since returns to both formal education and traditional knowledge are uncertain it is hard to hypothesise about the effect of uncertainty or of risk aversion on the level of education.

[4] The idea that education has a return primarily during times of change has been around since Schultz (1975) and Foster and Rosenzweig (1996).

2.2 Diversification of human capital investment

Independent of which investment is viewed as most uncertain – formal education or traditional knowledge – diversification is a possible strategy. To some extent, it is possible to diversify at the individual level by providing a child with both formal education and traditional knowledge. A reason for delayed school entry in rural areas of less-developed countries can be a desire that children should first gain some basic traditional knowledge (Bommier and Lambert, 2000).

However, scope for diversification is probably larger at the household level, where, for example, one sibling can get more formal education while the others spend more time acquiring traditional knowledge. Household-level diversification has been proposed to matter for rural-urban migration, an issue naturally connected investment in education (Levhari and Stark, 1982).

Lilleør (2008a; 2008b, 2008c) argues that household-level diversification of educational investment should be especially important where people rely on mutual support within the extended family for insurance and old-age support. Such diversification can result in negative sibling-dependency in education, i.e. more education of older siblings being negatively related to younger sibling's education. Using Tanzanian data, Lilleør's (2008b) finds negative sibling-dependency among sons – who are perhaps more likely to support parents when old - when a large share of older brothers is well-educated (the total effect is non-linear; when older brothers have less education the effect is positive).

A problem when attempting to investigate diversification by analysing sibling-dependency in education is that many other factors – e.g. credit constraints, unobserved parental preferences, or positive within-household education spillovers – could also affect siblings' dependency[5]. However, there should be more diversification in more risk-averse households, while there is no reason to expect risk aversion to matter for the other factors creating sibling-dependency. Thus, we can test for diversification as a motive influencing children's school entry by analysing how sibling-dependency differs in differently risk-averse households.

[5] In a companion paper (Lindskog, 2011) total sibling-dependency in education and what explains it is analyzed, using the same data from rural Amhara region of Ethiopia. Thus, a more detailed description of other mechanisms that can create sibling-dependency can be found there.

2.3 A simple model of diversification

The purpose of this model is to illustrate the motivation to diversify, and how this differs with risk aversion, not to offer a complete and realistic model of what determines children's school entry or total schooling. For now, let's therefore completely abstract from inter-temporal aspects of the human-capital investment decision, making essentially a one-period model, even though we will consider the impact of current human-capital investment on expected future income and consumption. Since savings/debts, as well as current-period costs and gains from formal education or from traditional knowledge, i.e. the fruits of labour, are left out of the model, credit constraints are not an issue.

The household consists of one parent, who is the decision-maker, and two children, one older and one younger. Parents get utility from consumption, and have a concave utility function, meaning that they are at least to some degree risk averse.

Children's time can be allocated to formal education (ed) or learning of traditional knowledge (tk), so that $tk^{old} = 1 - ed^{old}$ and $tk^{young} = 1 - ed^{young}$. Parent's future consumption depends on the random return to children's time invested in formal education (R^{ed}) and traditional knowledge (R^{tk}) where expectations on older and younger siblings' contributions are the same, conditional on their human capital; $E\left[R^{tk\_young}\right] = E\left[R^{tk\_old}\right] = \mu^{tk}$ and $E\left[R^{ed\_young}\right] = E\left[R^{ed\_old}\right] = \mu^{ed}$. We assume that human capital investment of the older child's time has already been made, leaving only the decision about the younger child's human capital investment.

The amount of formal education is chosen to maximize the expected utility of future consumption, $E[u(c)] = E\left[u\left(R^{tk}\left(2 - ed^{old} - ed^{young}\right) + R^{ed}\left(ed^{old} + ed^{young}\right)\right)\right]$, resulting in the following first order conditions:

$$\begin{aligned} E\left[u'(c)R^{ed}\right] &= E\left[u'(c)R^{tk}\right] \qquad && ed^{young} \in (0,1) \\ E\left[u'(c)R^{ed}\right] &> E\left[u'(c)R^{tk}\right] \qquad && ed^{young} = 1 \\ E\left[u'(c)R^{ed}\right] &< E\left[u'(c)R^{tk}\right] \qquad && ed^{young} = 0 \end{aligned} \tag{1}$$

The expected marginal utilities should thus be the same from investment in formal education and in traditional knowledge. At the extremes, when they differ "sufficiently", the younger child will specialize completely in either formal education or traditional knowledge.

Focusing on the interior solution, the first-order condition can be rewritten

$$E[u'(c)]\mu^{ed} + Cov(u'(c), R^{ed}) = E[u'(c)]\mu^{tk} + Cov(u'(c), R^{tk}) \qquad (2)$$

where the covariance terms are negative and dependent on the total amount invested in formal education, $ed^{old} + ed^{young}$. The diversification motive stems from this dependence of the covariance terms on the total level of investment in formal education. With more formal education, consumption depends more on the return to formal education and less on the return to traditional knowledge, making the right-hand side larger and the left-hand side smaller. Thus, if the expected return to formal education is not too different from that to traditional knowledge, parents will want to diversify and make consumption dependent on both types of human capital, rather than just one. When the older sibling has more formal education, the optimal level of formal education for the younger child consequently becomes smaller.

Now let's assume that $v(c)$ is the utility function of a more risk averse person than the person with utility function $u(c)$. This means that $v(c)$ is more concave than $u(c)$; formally, there is an increasing concave function $\psi(\cdot)$ such that $v(c) = \psi(u(c))$. This makes $v(c)$ more sensitive to dispersions of consumption around the expected value. [6] Thus, everything else equal, the absolute values of the covariance terms $Cov(v'(c), R^{ed})$ and $Cov(v'(c), R^{tk})$ are larger than the absolute values of $Cov(u'(c), R^{ed})$ and $Cov(u'(c), R^{tk})$. Moreover, if $u(c)$ is replaced by $v(c)$, in absolute terms, the larger of the two covariance terms, $Cov(u'(c), R^{ed})$ and $Cov(u'(c), R^{tk})$, will increase more than the smaller one. Everything else equal, the larger covariance term will, by equation (2), be the covariance term pertaining to returns to investments in the type of human capital with the highest expected return, and therefore where most investment should be made. Thus, an increase in risk-aversion will imply a shift from investments

[6] How sensitive $v(c)$ and $u(c)$ are to differences in expected returns depends on the specific functional forms. Sometimes $v(c)$ will be more sensitive at 'sufficiently' low levels of consumption. At 'sufficiently' high levels of consumption (sometimes always), $v(c)$ will be less sensitive.

in the type of human capital with the highest expected return, towards a more diversified combination.

3. Empirical approach

Sibling-dependency in education is investigated by estimating the effect of older siblings' literacy on the annual conditional school entry probability. The first sub-section describes and motivates the empirical model. The second sub-section brings up the possibility of sample-selection bias.

3.1 The annual school entry probability

While parents are almost invariably assumed in economics to be the sole household decision-makers, it is quite likely that children themselves influence education decisions more and more as they grow older and/or gain more school experience.[7] But, as noted earlier, school entry should be more influenced by parents' preferences, on risk for example, than by the child's preferences. School entry should also be less influenced by the child's relieved ability in school and 'taste' for school.

There are practical advantages with focusing on school entry, which ensures a fairly large sample with a fair degree of variation in the dependent variable. All children over a certain age have been eligible for school entry at some point (though not necessarily between 2000 and 2006), and information about ever having attended school, and school-entry age, is available for most children in the data.

The official school-entry age in Ethiopia is 7, but some enter at 6, and many enter later, particularly in rural areas. In the empirical analysis, a child is classified as eligible to enter school if between 6 and 16 years old and never attended before. If a children eligible to enter school does not, this means either that they will never enter or that they will enter later. Never attending school would mean complete specialization in traditional knowledge, while, as noted earlier, delayed school entry

[7] In urban Mexico, preferences of both the child and the mother have been found to matter for the decision to continue into senior secondary school, while only preferences of the child matter for the decision to start college (Attanasio and Kaufmann, 2009). It thus seems reasonable that preferences of the parents are most important for initial school entry.

may be the result of a desire that the child first acquire some basic traditional knowledge (Bommier and Lambert, 2000).

The annual conditional school entry probability is essentially a discrete-time “hazard” model. The advantage of modelling it instead of duration until school entry (or age at school entry) is that the data can be fully used while avoiding censoring problems. There is no need to restrict the sample to children old enough to know that most of those who haven’t entered will not do so later. This is especially useful for the results to remain relevant in a situation, such as Ethiopia, where schooling has increased massively in recent years.

The main explanatory variable used here is literacy of older siblings, which is clearly endogenous to parental characteristics affecting education decisions regarding all children in the family. Some of these characteristics, such as parental attitudes towards formal education versus investment in traditional knowledge, and towards child human capital (education and health) investment in general, are unobservable. Unobserved household effects must therefore be controlled for. Doing so with random effects would be problematic, since unobserved parental characteristics are likely to be correlated with older siblings’ literacy. For this reason, household fixed effects are used.

There could still be problems of time-varying shocks to the household, affecting education both of the child and of older siblings. Older siblings’ education is lagged, which would deal with time-varying household effects in a model without fixed effects. But with fixed effects, strict exogeneity is a must; that is, the explanatory variables must be uncorrelated with lags and leads of the error term (Arellano, 2003). One way to deal with time-varying shocks is to include measures of self-reported health and environmental shocks as done here. As opposed to variables measuring income or wealth over time, these should be exogenous to older siblings’ education.

To be able to include household fixed effects, a linear probability model is used, though the common procedure would be to estimate a conditional probability with a logit model. But using household dummies in a non-linear model, such as the logit, leads to biased estimates when not many observations per household are available (Lancester, 2000). A conditional fixed-effects logit model could be an alternative, but then observations' from households without variation in the dependent variable would not be used (Hsiao, 2003). The disregarded observations would be from households where no child enters school or where all children enter as soon as they become eligible, and thus disregarded households would be special with regards to preference for diversification. A disadvantage with the linear probability model is that we may end up with predicted probabilities below 0 or above 1.

The conditional school-entry probability of child *i* in household *h* and year *t* is

$$\Pr(y_{it} = 1) = \beta_0 + \sum_{k=1}^{3} \beta_k lit_{it}^{old} risk_h^k + \beta_4 sib_{it} + \beta_5 age_{it} + \beta_6 z_{ht} + \alpha_h + \tau_t + \varepsilon_{it} \qquad (3)$$

which depends on the education of older siblings measured by their literacy (lit_{it}^{old}), with the effect allowed to differ depending on three categories of risk preference of the household head $(risk_h^k)$, with $k = 1$ being the least risk averse and $k = 3$ the most. The school entry probability also depends on: the number of older siblings, (sib_{it}) capturing effects of birth-order rather than of the total number of siblings in a within-household model; child age (age_{it}); self-reported shocks to the household (z_{ht}); time-constant household characteristics (α_h) such as parental education, permanent income and unobserved parental attitudes towards education; year effects, (τ_t) capturing both the massive expansion in primary education in Ethiopia over the study period and possible effects of aging of the households in the panel; and an error term (ε_{it}).

The main hypothesis to be tested is $\beta_1 > \beta_2 > \beta_3$, i.e. that the effect of older siblings' literacy is more negative (smaller if positive, larger if negative) in more risk-averse households. If the diversification motive is strong enough to dominate factors suggesting positive sibling-dependency in education, we should also find $\beta_1, \beta_2, \beta_3 < 0$.

Estimations are done separately for boys and girls since it is quite possible that diversification is gender-specific, that it takes place across brothers or across sisters, rather than across siblings in general. For example, if older brothers have little formal education (and much traditional knowledge), parents might want the younger brother to have more formal education, and hence that he enters school early, but their preferences on the younger sister's education might not be affected.

3.2 Is there sample selection bias?

Since we are interested in the effects of older siblings' literacy, obviously only children with older siblings can be included in the estimations. So first-born children were excluded, when estimating gender specific effects both first-born sisters and first-born brothers. Smaller families can then be disproportionally excluded, which might be problematic as they might differ from larger families in important ways.

According to the child quantity-quality trade-off hypothesis, parents will beforehand make decisions about the number of children and how much to invest in their education and health (Becker and Lewis, 1973). More education-friendly parents may choose to have fewer children in order to invest more in each. Moreover, more risk-averse parents may choose to have more children (Cain, 1983).

To investigate possible sample-selection bias, excluded households are compared with all others in Table 1, with regard to educational indicators, risk aversion of the household head, and some other household characteristics.

Table 1: Comparison of included households and those excluded since they only had first-born children in the sample

	Excluded households		Included households		Equal means test	
	Mean	St. dev.	Mean	St. dev.	t-stats	p-value
Entry rate	0.31	0.26	0.28	0.25	0.94	0.335
Grade-progress rate	0.95	0.10	0.94	0.14	1.03	0.302
Literacy of household head	0.37	0.48	0.46	0.50	-1.69	0.091
Literacy of spouse	0.22	0.41	0.21	0.41	0.15	0.882
Household size	3.49	2.45	5.57	1.95	-7.78	0.00
Age of household head	48.03	20.76	46.50	14.08	0.59	0.552
Most riskaverse head	0.18	0.38	0.17	0.37	0.33	0.742
Middle riskaverse head	0.47	0.50	0.50	0.50	-0.65	0.515
Least riskaverse head	0.36	0.48	0.34	0.47	0.42	0.671

Excluded households tend to be much smaller, and are less likely to have literate heads (if the smaller size had anything to do with a quantity-quality trade-off, we would have expected smaller households to have more literate heads). Excluded households do not differ much from others with regard to children's school-entry and -progress rates,[8] or with regard to risk aversion.

4. The data and variables

The data used here comes from the Ethiopian Environmental Household Survey (EEHS), collected by the Ethiopian Development Research Institute (EDRI) in cooperation with the University of Gothenburg and, during the last round, with the World Bank. Four rounds of data have been collected, in 2000, 2002, 2005, and 2007. Interviews were conducted in April/May, towards the end of the Ethiopian school year, which starts in September and ends in June.

The sampled households were from 13 Kebeles in the South Wollo and East Gojjam zones of the Amhara region. The two zones were chosen to represent different agro-climatic zones in the Ethiopian highlands: There is less rainfall in South Wollo than in East Gojjam. Most households in the study areas make their living from rain-fed subsistence agriculture. Access to roads and to capital markets is quite limited.

Two of the Kebeles were added in the third round in order to evaluate a land certification program. The other eleven Kebeles were chosen randomly within the two zones. Within each 120 households were randomly selected. On average an interview took 1.6 days to complete. When a household was not located in a follow-up survey, it was replaced with another, randomly selected, household.

Most of the information on children's education was collected in the fourth round, when respondents were asked about the schooling history of all household members aged 6 to 24. It was attempted to collect data for household members no longer residing in the household, but less successfully, resulting in more missing and incomplete data for non-resident household members. Data from the fourth round, is her to create an annual panel on entry into first grade.

[8] But consistent with the empirical evidence from many less-developed countries, oldest siblings (excluded from the estimations) generally have rates than did their younger siblings (results not reported).

In the fourth-round sample there are 5,160 children aged 6 to 16 from 1,652 households. A requirement that the household be present in at least one previous round, and that there be relatively stable risk preferences of the household head over the third and fourth rounds, reduces the sample to 3,694 children from 1,171 households.[9] Excluding children with no older sibling further reduces the sample to 2,402 children from 875 households (using only children with both an older brother and an older sister, as is done in some estimations, reduces it to 1,766 children from 638 households). Of the remaining children, school entry data is available for 94.6%; 21.1% have never attended school and 73.5% have information about age at school entry. To be included in the estimation children have to be eligible for school entry at some point during 2000-2006 and information on explanatory variables has to be available. This leaves 1,094 children from 527 households in the final main sample.

A central explanatory variable is the household head's risk preference. In the third and fourth rounds the household head did risk aversion experiments, being asked to make pair-wise choices between plots that differed with respect to their yields in good times and bad, with a 50/50 probability of each. One plot had a higher expected yield, but a lower certain (bad times) yield. Based on a sequence of choices, each household was given a risk-preference rank of 1 to 5.

Risk preferences expressed at a specific time are likely to have both a time-constant part, i.e., underlying exogenous preference, and a context-dependent part, which might vary with income and wealth for example. Here focus is on the time-constant part. The mean from the two rounds was calculated and three dummies created for differently risk-averse heads.[10] To increase the reliability of the data households where risk preferences changed too much between the two rounds (about one third of the sample) were also excluded. This probably also eliminate households that experienced large income shocks between the rounds.[11]

[9] Including also households with less stable risk preferences don't qualitatively affect the results.

[10] An alternative is to compute a risk-aversion parameter and interact this with the literacy of older siblings. This doesn't qualitative change the results (results not reported). When computing the risk-aversion parameter, constant relative risk aversion was assumed. The mid-points of ranges were used for the three middle ranks and the least extreme end-points for the most and least risk-averse ranks. Means over the two rounds were then calculated.

[11] Using the same data, Damon et al., (2011) considered determinants of changed time preferences between rounds, and find environmental shocks to be the major determinant.

The data on the number of older siblings is from the last preceding round and includes both those living in the household and any who might have left. As noted, siblings' education is measured by their literacy rate. Though a rough indicator of investment in education, literacy has the advantage of few missing values.

To control for time-varying shocks to household income, dummies indicating the self-reported occurrence of health and environmental shocks are used. Health shocks are either the death or serious illness of a household member. Environmental shocks are mainly draughts and floods, but also other weather-related shocks as well as pests affecting plants or animals. Each dummy was set to one if the shock had occurred at least once between rounds.[12]

An age control is also included and possible gender effects are distinguished by running separate estimations for boys and girls. Summary statistics for all variables included in the model are reported in table A1 in the appendix.

5. Education in Ethiopia and among children in the data

There have been dramatic changes in primary education in Ethiopia recently, with massive increases in enrolment, albeit from a very low starting point. The changes started with the 1994 Education Reform, followed, so far, by three Education Sector Development Programs. The reform in 1994 abolished school fees, and since then decision-making has been decentralized and community involvement in schools has been encouraged. Moreover, many new schools have been built: The number of primary schools increased about 50% during 2000-2004, with the largest increase in rural areas. The budget share for education has also increased, from 13.8% in fiscal year 2000/01 to 19% in 2004/05. As a result, enrolment rates have steadily increased at all stages of education: The gross primary school enrolment rate rose from 34.0% in 1994/95 to 91.3% in 2005/2006, and net enrolment from 36.0% in 1999/2000 to

[12] If there was a shock during 2005-2006, the dummy was set to equal one for both these years. While there is information about the timing of the last shock in the data, there is no information about the timing of earlier shocks, so an annual shocks series could not be created. Estimations using an index of wealth in the preceding round instead of shocks where also run, which did not qualitative affect results. Since income and wealth could be endogenous to older siblings' education, the shocks variables were preferred despite their limitations.

77.5% in 2006/07.[13] Furthermore, the gender gap has been narrowed; the gender parity index increased from 0.6 in 1997/98 to 0.84 in 2005/2006. As is common with such large expansions in enrolment, the numbers of teachers and classrooms have not increased at pace with the number of pupils, raising concerns about reduced quality (Oumer, 2009; Ministry of Education, 2005; World Bank, 2005).

Ethiopia is a large and diverse country, and there are large regional variations in gross and net enrolment rates, as well as in gender disparities. In Amhara net enrolment in years 2004/2005 was 54.6% for boys and 53.1% for girls, both lower than the country averages, but Amhara is one of the few regions where net enrolment appears to be nearly as high for girls as for boys (Ministry of Education, 2005).

Using the school-entry data collected in the fourth round, Table 2 reports the shares of 8 and 11 year old children who had started school over time. It was common to start late, the share who has started by age 8 is around 20% in the mid-1990s and approach 60% after the mid-2000s, while a larger share of 11-year olds have started (around 30% in the mid-1990s and approaching 85% after the mid-2000s). Still in 2006, many children appear to never start school at all, or at least they had not yet done so by age 11.

Table 2: Share of 8 and 11 years old boys and girls that had started school over time

Year	Girls age 8	Boys age 8	Girls age 11	Boys age 11
1996	0.20	0.23	0.29	0.30
1997	0.22	0.13	0.40	0.35
1998	0.30	0.23	0.41	0.47
1999	0.32	0.28	0.54	0.51
2000	0.38	0.31	0.60	0.54
2001	0.45	0.41	0.65	0.51
2002	0.46	0.44	0.74	0.62
2003	0.43	0.45	0.69	0.63
2004	0.53	0.53	0.84	0.71
2005	0.56	0.58	0.80	0.75
2006	0.61	0.59	0.82	0.80
2007	0.56	0.35	0.84	0.86

Information on child age and if the child has yet started school is from the spring in the relevant year. Thus, a child that have not started could start in the autumn that year.

[13] The gross primary school enrolment rate is the ratio of number of pupils enrolled in primary school to the number of children in primary-school age. The net primary school enrolment rate is the ratio of the number of pupils in primary-school age enrolled in primary school to the total number of children in primary-school age.

6. Regression analysis

Estimations of annual school entry probabilities were run for girls and boys separately, treating all older siblings the same (columns 1 and 3, Table 3) and distinguishing sisters from brothers (columns 2 and 4, Table 3).

Table 3: Coefficents from linear estimations with household fixed effects of the effect of diversification on the annual school entry probability.

	Girls		Boys	
	(1)	(2)	(3)	(4)
Number of older siblings	-0.002		0.014	
	(0.012)		(0.011)	
Older siblings' literacy rate *low risk-aversion	0.123*		0.122	
	(0.069)		(0.079)	
Older siblings' literacy rate *middle risk-aversion	0.040		0.231***	
	(0.082)		(0.069)	
Older siblings' literacy rate *high risk-aversion	0.098		0.072	
	(0.147)		(0.096)	
Number of older sisters		-0.005		0.001
		(0.022)		(0.024)
Older sisters literacy rate *low risk-aversion		0.046		0.038
		(0.098)		(0.080)
Older sisters literacy rate *middle risk-aversion		0.156*		0.149*
		(0.090)		(0.085)
Older sisters literacy rate *high risk-aversion		0.042		0.140
		(0.063)		(0.134)
Number of older brothers		-0.001		0.073***
		(0.027)		(0.027)
Older brothers literacy rate*racy rate low risk-aversion		0.089		0.221**
		(0.063)		(0.086)
Older brothers literacy rate*middle risk-aversion		0.023		0.223**
		(0.080)		(0.098)
Older brothers literacy rate*high risk-aversion		0.004		-0.007
		(0.149)		(0.172)
Age	0.011**	0.009	0.026***	0.035***
	(0.005)	(0.007)	(0.005)	(0.008)
Year 2001	0.069**	0.077**	0.055**	0.083**
	(0.027)	(0.033)	(0.024)	(0.033)
Year 2002	0.137***	0.156***	0.095***	0.091**
	(0.033)	(0.039)	(0.025)	(0.036)
Year 2003	0.194***	0.237***	0.144***	0.143***
	(0.035)	(0.045)	(0.030)	(0.041)
Year 2004	0.220***	0.252***	0.145***	0.103*
	(0.044)	(0.057)	(0.038)	(0.056)
Year 2005	0.250***	0.275***	0.219***	0.201***
	(0.047)	(0.059)	(0.041)	(0.060)
Year 2006	0.276***	0.294***	0.155***	0.089
	(0.054)	(0.067)	(0.045)	(0.069)

Table 3 cont.

Health shock	0.027	0.027	-0.031	-0.049
	(0.041)	(0.052)	(0.039)	(0.050)
Environmental shock	-0.017	0.022	0.007	0.043
	(0.031)	(0.039)	(0.035)	(0.047)
Constant	-0.064	-0.105	-0.275***	-0.554***
	(0.073)	(0.105)	(0.071)	(0.139)
Observations	2194	1337	2553	1361
Children	693	413	794	420
Households	465	274	517	288

Standard errors, clustered at the household, in parenthesis.
*= p<0.10, **=p<0.05, ***=p<0.01

Older siblings' literacy seems to matter for school-entry probabilities, with a positive effect. In all four estimations there are statistically significant large positive effects for at least one risk-aversion category. Thus, diversification motive is not strong enough to dominate forces creating positive sibling-dependency. But in line with a diversification motive, the positive effects are found in low and middle, not high, risk-aversion households. For boys there is a statistically significant 22 percentage point increase in the school entry probability if older siblings are literate rather than illiterate, when the head has middle risk aversion. When the household head has low risk aversion the probability of school-entry increases with about 12 percentage points for girls and boys alike. For girls the effect is just significant at the ten percent level and for boys it is just not. Effects in high risk-aversion households, and effects on girls' school entry in middle risk-aversion households, are far from statistical significance. However, coefficients do not follow the hypothesized pattern $\beta_1 > \beta_2 > \beta_3$. For boys the effect is larger in mid- than low-risk aversion households. The null hypothesis of equal effects of older siblings' literacy in differently risk averse households cannot be rejected (Table 4).

Table 4: Is the effect of older siblings' literacy equal across levels of household risk-averseion (F-tests on equality of coefficients)?

	Boys		Girls	
Null hypothesis	F-statistic	p-value	F-statistic	p-value
Literacy of older siblings has an equal coefficient across levels of household risk aversion	1.05	0.351	0.31	0.731
Literacy of older sisters has an equal coefficient across levels of household risk aversion	0.53	0.588	0.64	0.530
Literacy of older brothers has an equal coefficient across levels of household risk aversion	0.73	0.481	0.28	0.754

Distinguishing older sisters from older brothers (columns 2 and 4, Table 3) gives results suggestive of diversification across brothers. In households with not too risk-averse heads, the school entry probability for boys is about 22 percentage point higher when older brothers are literate, while the effect is small and not statistically significant in households with the most risk-averse heads. Diversification across brothers, but not across sisters, is consistent with the findings of Lilleør (2008b). Again, the null hypotheses of equal effect of older sisters' and brothers' literacy in differently risk-averse households cannot be rejected (Table 4).

The number of older siblings had little effect on the annual school-entry probability (Table3). The exception is a positive effect for boys of having more older brothers; it might thus be beneficial to be later born among the boys. This says nothing about any adverse effect of being first-born, or being the first-born sister, often discussed in the literature, since first-born are not in the sample.[14]

Age increases school entry probability more for boys than for girls. This is consistent with boys being more likely to enter late. The omitted reference year is 2000; the coefficients of the year dummies reveal that, over time, girls' school-entry probability increased more than boys'. There are no statistically significant effects of health shocks or environmental shocks.

7. Summary and conclusions

Household-level diversification in human capital investment in a rural area in a less-developed country was investigated. Returns both to formal education and to investment in traditional knowledge are uncertain. A simple model shows that unless expected returns differ too much, parents therefore want to diversify, resulting in negative siblings' dependency in education. If older siblings have less formal education and more traditional knowledge, the young sibling will have more formal education. The desire for diversification should be stronger when the household head is more risk-averse, implying that sibling-dependency in education should be more strongly negative as well.

[14] Just comparing means, first-born children indeed appear to get less education than others.

The empirical analysis estimated the effect of older siblings' literacy on annual school entry probability, using data from rural Amhara region in Ethiopia. Rural Amhara is a place with extensive informal insurance and where parents are likely to depend on children in old age, hence a place where household-level diversification could be of importance. The effect of older siblings' literacy was allowed to differ across households with differently risk-averse heads. Estimations were done separately for girls and boys, and distinguishing effects of older sisters from effects of older brothers.

Total sibling-dependency in education was found to be positive, so any diversification was dominated by other forces. But in line with diversification across brothers, the effect of older brothers' literacy was smaller (more negative) in households with the most risk-averse heads. Having literate rather than illiterate older brothers was found to increase the annual school-entry probability with 22 percentage points in households where the head was not too risk-averse. In households with the most risk-averse heads there is no effect of older brothers' literacy, i.e. the effect of having literate older brothers was estimated to be around 22 percentage points lower. However, the coefficient of older brothers' literacy in the households with the most risk-averse heads, in addition to being close to zero, had a large standard error, so that the null hypothesis of an equal effect in these households could not be rejected. For girls no influence of diversification could be detected.

Summing up, diversification across brothers possibly influenced boys' school entry in rural Amhara in 2000-2006. Possible diversification across brothers, but not across sisters, has been found also in rural Tanzania (Lilleør, 2008b). However, the 'diversification across brothers' result in this paper was statistically weak, and any diversification was dominated by factors creating positive sibling-dependency in education.

Considering the very limited research on diversification of human capital investment, more research is certainly needed. The issue has important policy implications. If a desire for diversification is important to rural households, policies that makes it easier to combine school with farm work and household work – or that makes the curriculum more relevant for such traditional activities in rural areas – could result in more schooling for children from rural households with risk-averse decision-makers.

References

Akresh, R., E. Bagby, D. de Walque, H. Kazianga (2010), "Child Ability and Household Human Capital Investment Decisions in Burkina Faso", IZA Discussion Paper 5326.

Anderson, K.H., E.M. King, Y. Wang (2003), "Market Returns, Transfers and Demand for Schooling in Malaysia, 1976-89", *Journal of Development Studies* 39(3): 1-28.

Arellano, M. (2003), "Panel Data Econometrics", Oxford University Press, New York.

Attanasio, O., K. Kaufmann (2009), "Educational Choices, Subjective Expectations, and Credit Constraints", NBER Working Paper 15087.

Baland, J.M., J.A. Robinson (2000), "Is Child Labor Inefficient", *Journal of Political Economy* 108(4): 663-679.

Becker, G. (1962), "Investment in Human Capital: A Theoretical Analysis", *Journal of Political Economy* 70(5): 9-49.

Becker, G., H. Lewis (1973), "On the Interaction between the Quantity and Quality of Children", *Journal of Political Economy* 81(2): S279-S288.

Belzil, C., M. Leonardi (2007), "Can Risk Aversion Explain Schooling Attainments? Evidence from Italy", *Labor Economics* 14(6): 957-970.

Ben-Porath, Y. (1967), "The Production of Human Capital and the Life Cycle of Earnings", The Journal of Political Economy 75(4): 352-365.

Bommier, A., S. Lambert (2000), "Education Demand and Age at School Enrolment in Tanzania", *Journal of Human Resources* 35(1): 177-203.

Cain, M. (1983), "Fertility as an Adjustment to Risk", *Population and Development Review* 9(4): 688-302.

Card, D., A. Krueger (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools", *Journal of Political Economy* 100(1): 1-40.

Chamarbagwala, R. (2008), "Regional Returns to Education, Child Labour and Schooling in India", *The Journal of Development Studies* 44(2): 233-257.

Damon, M., Difalco, S., Köhlin, G. (2011), "Environmental Shocks and Rates of Time Preference: A Behavioral Dimension of Poverty Traps", mimeo.

Edmonds, E. (2006), "Child Labor and Schooling Responses to Anticipated Income in South Africa", *Journal of Development Economics* 81(2): 386-414.

Estevan, F., J.M Baland (2007), "Mortality risks, education and child labor", *Journal of Development Economics* 84(1): 118-137.

Foster, A. , M. Rosenzweig (1996), “Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution”, *American Economic Review* 86(4): 931-953.

Gitter, S.R., B.L. Barham (2007), “Credit, Natural Disasters, Coffee, and Educational Attainment in Rural Honduras”, *World Development* 35(3): 498-511.

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

Glewwe, P., M. Kremer (2006), “Schools, Teachers, and Education Outcomes in Developing Countries”, Chapter 16 in Hanuschek, E. and Welch, F. (eds.) *Handbook of the Economics of Education*, North-Holland, Amsterdam.

Grootaert, C., R. Kanbur (1995), “Child Labour: an Economic Perspective”, *International Labour Review* 134(2): 187-203.

Hanushek, E. (2008), “Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries”, *Journal of Human Capital* 2(1): 69–105.

Hanushek, E., L. Woessmann (2008), “The Role of Cognitive Skills in Economic Development”, Journal of Economic Literature 46(3): 607–668.

Huisman, J., J. Smits (2009), “Effects of Household- and District-Level Factors on Primary School Enrollment in 30 Developing Countries”, *World Development* 37(1): 179-193.

Hsiao, C. (2003), Analysis of Panel Data, Second Edition (Econometric Society Monographs No. 34), Cambridge University Press, Cambridge.

Jacoby, H. (1994), “Borrowing Constraints and Progress Through School: Evidence from Peru”, *Review of Economics and Statistics* 176(1): 151-160.

Jensen, R. (2010), “The (Perceived) Returns to Education and the Demand for Schooling”, *Quarterly Journal of Economics* 125(2): 515-548.

Kingdon, G., N. Leopold (2008) “Do Returns to Education Matter to Schooling Participation? Evidence from India”, *Education Economics*, 16(4), 329-350.

Kochar, A. (2004) “Urban Influences on Rural Schooling in India”, *Journal of Development Economics* 74(1): 113-136.

Kodde, D. (1986), “Uncertainty and the Demand for Education”, *The Review of Economics and Statistics* 68(3): 460-467.

Lancester, T. (2000), “The Incidental Parameters Problem Since 1948”, *Journal of Econometric Literature* 95: 391-413.

Levhari, D., O. Stark (1982), “On Migration and Risk in LDCs”, *Economic Development and Cultural Change* 31(1): 191-196.

Levhari, D., Y. Weiss (1974), "The Effect of Risk on the Investment in Human Capital", *The American Economic Review* 64(6): 950-963.

Lilleør, H.B. (2008a), "Can Future Uncertainty Keep Children out of School?", Working paper no 2008-06, Centre for Applied Microeconomtrics (CAM), University of Copenhagen.

Lilleør, H.B. (2008b), "Sibling Dependence, Uncertainty and Education. Findings from Tanzania", Working paper no 2008-05, Centre for Applied Microeconomtrics (CAM), University of Copenhagen.

Lilleør, H.B. (2008c), "Human Capital Diversification within the Household: Findings from Rural Tanzania", Working paper no 2008-04, Centre for Applied Microeconomtrics (CAM), University of Copenhagen.

Lindskog, A. (2011), "The Effect of Older Siblings' Literacy on School-Entry and Primary School Progress in the Ethiopian Highlands", Working Papers in Economics No. 495, Department of Economics, School of Economics, Business and Law, University of Gothenburg.

Ministry of Education (2005), Education Sector Development Program III (ESDP-III): Program Action Plan, Addis Ababa.

Nerman, M., T. Owens (2010), "Determinants of Demand for Education in Tanzania: Costs, Returns and Preferences", Working Papers in Economics No. 472, Department of Economics, School of Economics, Business and Law, University of Gothenburg.

Orazem, P.F, E.M. King (2007), "Schooling in Developing Countries: The Roles of Supply, Demand and Government Interventions", chapter 55 in T. Schultz and J. Strauss (eds.) Handbook of development economics, volume 4, North Holland.

Oumer, J. (2009), The Challenges of Free Primary Education in Ethiopia", International Institute for Educational Planning (IIEP), Paris.

Pouliot W (2006). "Introducing Uncertainty into Baland and Robinson's Model of Child Labor", *Journal of Development Economics,* 79(1): 264-272.

Rosenzweig, M., Wolpin, K. (1985), "Specific Experience, Household Structure and Intergenerational Transfers: Farm Family Land and Labor Arrangements in Developing Countries", *Quarterly Journal of Economics* 100(Supplement): 961-987.

Schultz, T. (1975), "The Value of the Ability to Deal with Disequilibria", *Journal of Economic Literature* 13(3): 827-856.

World Bank (2005), "Education in Ethiopia: Strengthening the Foundation for Sustainable Progress", Washinton DC.

Appendix

Table A1: Variables and summary statistics

Variable	Obs	Mean	Std. Dev.
Age in years	11723	11.063	3.022
Year=2000	11723	0.142	0.349
Year=2001	11723	0.145	0.352
Year=2002	11723	0.149	0.356
Year=2003	11723	0.148	0.355
Year=2004	11723	0.147	0.354
Year=2005	11723	0.142	0.349
Year=2006	11723	0.128	0.334
Health shock	11695	0.177	0.381
Environmental shock	11695	0.346	0.476
Number of older sisters	11723	1.713	1.298
Number of older brothers	11723	1.858	1.379
Older sisters' literacy rate	9620	0.480	0.462
Older brothers' literacy rate	11690	0.513	0.430
Household head belonged to the least risk-averse category	11723	0.341	0.474
Household head belonged to the middle risk-averse category	11723	0.490	0.500
Household head belonged to the most risk-averse category	11723	0.169	0.375