Income Risk and Household Schooling Decisions in Burkina Faso

Harounan Kazianga

Oklahoma State University[1]

Forthcoming, World Development

harounan.kaziange@okstate.edu

[1] Contact: Harounan Kazianga, Oklahoma State University, Department of Economics, 324 Business Building, Stillwater, OK 74708. Email: harounan.kazianga@okstate.edu

Abstract

I study the effects of income uncertainty on household schooling decisions. Households with more volatile incomes have a greater incentive to build a buffer stock to insure against unforeseen adverse shocks, and non-enrollment can be part of such a strategy. I use data from rural Burkina Faso, where school attainment is low and income shocks are frequent, to show that income uncertainty reduces several educational outcomes, including enrollment, education expenditures and years of education completed. The findings suggest that income uncertainty has larger welfare costs in terms of human capital than is implied by studies that only focus on realized income shocks.

Key words: education, income risk, Africa, Burkina Faso

JEL classification: D99; I21; O1

1. INTRODUCTION

In this paper, I examine a feature of household income in less developed areas that has received little attention in connection to investments in education: income uncertainty. If removing children from school is an option when households are exposed to negative income shocks (e.g., Beegle et al., 2006; Sawada, 2003) and negative income shocks are frequent[1], then prudent households may optimally choose not to enroll their children before the shocks materialize. *A priori*, this would be of lesser concern if returns to education were linear, i.e., if regardless of the grade at which child drops out, her education were to generate some positive returns[2]. However, there is some evidence to suggest increasing returns to education in low-income settings. Returns to education in the formal sector are typically small or non-existent at low levels of educational attainment (Bennell, 2002; Kazianga, 2004; Shady, 2003; Schultz, 2004, 2003)[3]. In addition, in the absence of technological innovation, returns to education in the agricultural sector are small, especially in sub-Saharan Africa (e.g., Appleton and Balihuta, 1996; Canagarajah et al., 1998; Joliffe, 1998).

Low levels of human capital, including education, health and nutrition, have direct consequences on welfare. Inequality in human capital outcomes, in addition to being of interest *per se*, also has direct and indirect impacts on income inequality. Education is crucial in augmenting individual earnings and improving the prospects for economic growth in general. Therefore, a better understanding of the constraints that poor households face when making decisions regarding education is critical in effectively addressing poverty. The examination of which constraints are the most important and which policies can best promote education has generated a vast literature in economic research (see Schultz, 1988 for a review).

Education is an irreversible investment with delayed, and possibly increasing, returns. From an economic perspective, holding expected income constant, risk-averse households that face uninsurable risk would allocate more resources to liquid assets than to irreversible investments. In particular, the precautionary motive for holding liquid assets may prevent households from undertaking productive investments (even when they can self-finance them), especially those that are irreversible (Fafchamps and Pender, 1997).

Hence, understanding how income uncertainty impacts decisions about schooling can shed light on the barriers to schooling that poor households face in low-income countries. I test the extent to which households facing higher income risk are more likely to reduce their investments in the human capital of their children to build saving stocks to offset future income shocks. More specifically, I test whether and to what extent income uncertainty acts as a barrier to educational attainment in rural areas, given school supply, household wealth and child characteristics.

The empirical work uses data from rural Burkina Faso, an environment in which income risk is pervasive and education levels are among the lowest in the world (UNESCO, 2005). Burkina offers an interesting setting for testing the effects of income variance on education for two reasons. First, levels of schooling in Burkina have been notably low. Total years of schooling average about 0.6 years for men aged 50 to 54 and 2.6 years among the youngest cohort (e.g., Schultz, 2003). Women in the same cohorts receive about half of the male schooling level, which suggests a persistent gender gap. For children aged 7 to 15 years, the average enrollment rate was about 36 percent in 2003, with wide disparities between boys and girls and between rural and urban areas (e.g., UNESCO, 2005). In the sample villages studied in this paper, the proportion of children between 7 and 15 who have ever attended school increased from 29.1 percent in 1995 to 34.4 percent in 2004, which indicates that the increase in education levels was modest. In light of the considerable evidence linking economic growth to education, it may be argued that such low levels of education are likely to have adverse effects on both individual welfare and long-term economic growth.

Second, households in Burkina face frequent crop failures, primarily due to drought spells. In the 1990s, the country confronted three major crop failures, in 1990/1991, 1995/1996 and 1997/1998, or roughly a major crop failure every three years (Zoungrana et al., 1999). Given that more that 80% of the population lives in rural areas, and virtually all of the rural population depends on rain-fed subsistence agriculture for their livelihoods, frequent crop failures translate into high income volatility. The extent to which such income volatility (in addition to exposure to negative income shocks) impacts household education choices has received little attention in economic research on low-income settings in general. In the specific case of rural Burkina Faso, Kazianga and Udry (2006) have shown that uncertainty about future income is an important determinant of current decisions on consumption and livestock holdings. In particular, they have established that conditional on current income shocks, households with higher income variance chose to save less, in the forms of livestock sales and grain storage drawn-down (e.g., Park, 2006). This paper extends these results to examine how income uncertainty affects households' education choices. Understanding how income uncertainty affects education choices can provide additional insights into the costs of incomplete financial markets in rural economies and how the lack of insurance in risky environments can contribute to the perpetuation of poverty. The main contribution of this paper is to resolve the puzzle of the co-existence of high (or increasing) returns on education in low-income countries (e.g., Kazianga, 2004; Psachalopoulos and Patrinos, 2004; Schultz, 2005) and low levels of human capital by using precautionary saving motives rather than binding credit constraints.

The paper is related to two strands of literature. The first strand tests how imperfect financial markets impact human capital acquisition (Duryea, 1998; Duryea and Arends-Kuenning, 2003; Jacoby and Skoufias, 1997; Jensen, 2000). This branch of research shows that exposure to income shocks is detrimental to education when households cannot rely on formal or informal mechanisms to smooth out negative income shocks. In particular, in the face of negative income shocks, households divert child time away from education and toward labor to generate immediate income (Beegle et al., 2006). This paper differs substantially from this line of work, however. Instead of examining how parents alter (*ex-post*) child time reallocation when faced with negative income shocks, the paper is more concerned with the cumulative effects of living in a risky environment. If, in anticipation of negative income shocks, households refrain from enrolling their children, then income uncertainty (not exposure to negative income shocks) becomes the main cause of observed low enrollment rates. In fact, enrollment rates may remain low even if the shocks do not materialize. This line of reasoning would imply that using child time to cope with negative income shocks *ex-post* could lead to a succession of enrollments and de-enrollments and/or low attendance. Ultimately, most individuals would have at least some level of education. In contrast, income uncertainty implies that a large fraction of individuals would never enroll. Hence the welfare costs of income risk and incomplete financial markets might be higher when households' *ex-ante* behavior is taken into account[4].

Second, the paper is related to a vast literature that examines how income uncertainty influences household saving and consumption behavior (e.g., Browning and Lusardi, 1996; Carroll, 1997; Carroll and Kimball, 2001; Kimball, 1991), as well as production behavior (e.g., Dercon and Christiaensen, 2010). A fundamental result in the precautionary savings literature is that the presence of uninsured risk leads prudent agents to save more than they would if there were no uncertainty (e.g., Aiyagari, 1994; Leland, 1968). The existing literature on precautionary savings focuses on the effects of income uncertainty on current consumption or asset portfolio allocation, with little attention to human capital acquisition. This paper departs from this strand of work by examining the effects of income uncertainty on education in an environment where income risk is pervasive and education levels are very low.

The closest related work is the study by Fitzsimons (2007) that tests the effects of income uncertainty on education in the context of Indonesia. I use, however, a different identification strategy than that used by Fitzsimons (2007). Furthermore, the settings are different. Enrollment rate in the study areas covered by Fitzsimons (2007) is about 80 percent; hence it is difficult to disentangle the effects of exposure to shocks which may have lead to temporary or permanent interruption from the effects of income uncertainty which influences the decision to enroll a child. In addition, while Fitzsimons (2007) finds a large impact of aggregate risk and a relatively small impact of idiosyncratic shocks, in the context of rural Burkina where households fail to insure against idiosyncratic income shocks (Kazianga and Udry, 2006), one would anticipate a stronger impact of idiosyncratic risk.

Controlling for current economic shocks, household wealth levels and child characteristics, I find that income uncertainty reduces a number of education outcomes including current enrollment status, education expenditures per child, the number of years of education completed and the probability of ever having been enrolled. The results indicate that, in addition to current income shocks and wealth levels (which have been found to determine education choices), income uncertainty has a separate effect on households' education choices. It is then plausible that the welfare and long-term costs of incomplete financial markets and income risk are higher than previously implied by studies that were exclusively focused on the impacts of the use of child time to cope with negative income shocks *ex-post*. In particular, *ex-post* adjustments to negative shocks imply a smaller (but positive) accumulated total years of education on average. However, income uncertainty can induce a situation in which a large fraction of the population never enrolls at all, especially when returns on education are non-linear.

The rest of the paper is organized as follows. The second section provides a brief review of the literature on parental income shocks and children's education. The third section describes the surveys and data used. The fourth section presents the empirical approach. The fifth section discusses the empirical results. The sixth provides some robustness checks and the seventh section concludes.

2.INCOME SHOCKS AND SCHOOLING DECISIONS: A BRIEF REVIEW

There is a sizable literature that examines the effects of income shocks on households (e.g., Alderman, 1996; Alderman and Paxson, 1994; Deaton, 1992; Morduch, 1999; Rosenzweig and Wolpin, 1993; Townsend, 1994). A subset of this larger strand of work establishes a link between transitory shocks to parental income and children's academic achievement. In particular, recent empirical work shows the role that shocks play in decisions regarding schooling. In one of the earliest studies, Jacoby (1994) examines the relationship between borrowing constraints and progression through school among Peruvian children. He concludes that a lack of access to credit is detrimental to the acquisition of human capital because children in households with borrowing constraints begin withdrawing from school earlier than those with access to credit.

Jacoby and Skoufias (1997) provide further evidence on the relationships between human education and the incompleteness of financial markets. Using data on school attendance patterns from six Indian villages, the authors find that households use fluctuations in school attendance as a form of self-insurance. Sawada (2003) and Sawada and Lokshin (2009) show that children's propensity to enroll in and drop out of school in rural Pakistan responds to transitory shocks. He finds that transitory income has a larger effect than permanent income, implying that transient income variation is a greater barrier to education than chronic poverty *per se*. Duryea (1998) examines the role of transitory shocks to household income on children's advancement through school in Brazil. Her estimates suggest that children whose fathers experience spells of unemployment (her proxy for income shocks) are less likely to advance to the next grade. These findings corroborate results uncovered by Jacoby (1994) in Peruvian villages. While these papers focus on non-diversifiable income fluctuations due to credit market imperfections, I look at income uncertainties arising from insurance market imperfections.

Conceptually, households' education choices in response to negative income shocks could operate in two ways. On the one hand, as in Jacoby and Skoufias (1997), when households are confronted with negative income shocks, parents may opt to have children engage in immediate income-generating activities, presumably at the cost of time allocated to education. If time reallocation operates at the margin, it may lead to lower attendance rates without children dropping out of school. On the other hand, exposure to a negative shock may induce schooling (permanent or temporary) interruption, i.e., parents decide to take their children out of school (Sawada, 2003; Sawada and Lokshin, 2009). In either case, in the long run, lower attendance rates and early dropout would translate into fewer years of completed education, but years of education would be non-zero for most individuals.

More precisely, using child time as part of an *ex-post* risk coping strategy implies that fewer years of education are completed than would have been under complete financial markets or in a risk-free world. Furthermore, only a small fraction of the population would never enroll because parents have the option of enrolling their children and then taking them out when faced with negative income shocks. In contrast, income uncertainty, especially in conjunction with an increasing rate of return on education, would induce a situation where forward looking households might choose to never enroll their children, i.e., at any point in time a sizeable fraction of the population (school age and above) never enrolls.

3.DATA AND DESCRIPTIVE STATISTICS

The data come from two surveys conducted in rural Burkina Faso in 1995 and 2004/2005. The survey covers six villages in three different regions with different agricultural and non-agricultural potential: the Namentenga province located in a Sudanian type region, the Soum province located in a Sahelian region and the Kossi province located in a Northern-Guinean type region. The main activity in the Sahelian region is herding. Agriculture and rearing small animals dominate in the Sudanian region. Overall, the populations in the three locations consist of subsistence farmers. Opportunities for cash crops are limited, except in the Northern-Guinean region, where cotton farming is important. For the purpose of this paper, it is worth noting that there is a school in each of these villages, so distance to school should be a minor concern[5].

In each village, 50 households were randomly selected to participate in a general household survey in 1995. A follow up survey, which tracked the original households, was conducted by the author between November 2004 and March 2005. Individuals who had left these households but still resided in the same villages at the time of the follow-up survey were also included. In total, 369 households were surveyed in the second round. This new sample consists of 125 newly formed households (from marriages and divisions of the 300 households) and 244 households that were part of the original sample. In addition to general information on household income, wealth and consumption, the follow-up survey collected detailed information on household size dynamics, education, fertility and immunization. This paper exploits the detailed information on the land holdings and educational histories of individual household members.

---Insert Table 1 here---

Table 1 summarizes key educational outcomes for school age children (i.e., children aged 7 to 15). The top two panels (panels A and B) show educational outcomes in 1995 and 2004. The education variable contained in the 1995 survey indicates whether an individual has ever enrolled. Although this variable may appear limited *a priori*, it still conveys information in an environment where approximately one in three children has ever attended school. The figures indicate that the likelihood of having ever enrolled increased for both boys and girls (from 29 percent in 1995 to 34.4 percent in 2004), although a sizeable gender gap still exits (40.1 percent of boys have ever been enrolled as opposed to 28.9 percent of girls). Enrollment rates improved in all villages, except in the Sudanian Niéga village, where in comparison to 1995, fewer children had ever been enrolled in school in 2004.

Data on current enrollment status are only available for the 2004 round and are summarized in panel C. The average current enrollment rate is about 26.3 percent, and this figure is consistent with figures from national surveys, which report an enrollment rate of 22 percent for rural areas (according to the 2003 release of the Burkina Demographic and Health Survey data). Overall, villages located in the Northern Guinean region (villages 5 and 6), tend to have the highest enrollment rates. A potential explanation is that cotton (which is a cash crop) provides farmers in these villages with a more reliable source of income. In addition, given current farming technologies; the return to education is potentially higher on cash crop farms (cotton) than on subsistence farms[6]. A puzzling result is the relatively higher enrollment rates in the Sahelian villages (villages 3 and 4). Although not well-documented in this version of the paper, prolonged interventions from NGOs could explain this pattern. Another caveat is that being close to a local town does not necessarily imply higher enrollment rates. Villages 1 (Niéga), 3 (Béléhédé) and 5 (Kéréna) are closer to the local town than the other villages in the same region. With the notable exceptions of Kéréna and Dissankuy[7], the enrollment rate is lower in villages that are closer to the local town.

In panel D, I summarize per student education-related expenses. Although primary education is officially free, parents are still required to pay various fees, including parent association fees, books and notebooks. The table shows the unconditional means, and the means conditional on being enrolled, at the time of the survey. Households spend the equivalent of approximately \$3 a year on boys' education and about \$2 on girls' education, although there are large differences across villages. Conditional on being enrolled, these figures increase to \$8 for boys and \$3 for girls. Although these figures appear small, they should be put in the context of these poor villages. First, education related expenditures are large relative to household income. The conditional mean education related expenditures (for boys and girls combined) corresponds to almost 20 percent of crop income per adult equivalent[8]. Second, households are required to make payment in cash and in a timely manner. This could be additional obstacles to cash constrained households with seasonal income opportunities.

4.EMPIRICAL MODEL AND IDENTIFICATION

Following Beegle et al. (2006) and Sawada (2003), among others, I use an empirical model of the following form, where I assume that the standard deviation of income shocks is a good proxy for income risk.

$$s_{ihv} = \alpha_1 std_{hv} + \alpha_2 x_{ihv} + \alpha_3 x_{hv} + \alpha_4 x_v + \varepsilon_{ihv} \tag{1}$$

Where s_{ihv} is the educational outcome for child i in household h in village v, std_{hv} is the estimated standard deviation of the income shocks for household h in village v, x_{ihv} summarizes child characteristics, x_{hv} summarizes household characteristics, x_v summarizes village characteristics, and ε is an error term. The $\alpha' s$ are parameters to be estimated. The theory predicts that α_1 should be negative (i.e., higher income variance reflects more uncertainty). The relatively long time span between the two rounds of the survey suggests that attrition might be an issue. Furthermore, estimating regression 1 requires a measure of the standard deviation of income shocks. I discuss these two issues below in subsections 4.1 and 4.2.

(a) Attrition

Although the 1995 sample was drawn randomly from village census data, the 2004 sample may not be random because households may leave selectively. The main concern is that land holdings (that I use in the identification strategy) and education (the outcome of interest) are potentially correlated with the decision to leave the villages and, hence, the sample. This would in turn bias the estimation results. For these reasons, this subsection provides a discussion of sample attrition as it pertains to the data.

As previously discussed, among the 300 households included in the 1995 survey, 248 of them remained in 2004. The attrition rate is about 17.33 percent over the 10 year interval, which corresponds to an annual attrition rate of 1.88 percent[9]. This level of attrition falls within the range of attrition observed for panel surveys with comparable interval lengths.

---Insert Table 2 here---

Table 2 presents the summary statistics by attrition status, using the 1995 data. "Leavers" refer to households who were dropped from the survey in 2004 and "stayers" refer to households that remained in the survey. The last row of the table reports the absolute t-value of the mean difference. This preliminary exploration implies that only female headship and household composition, particularly the presence of adult and school-aged girls in 1995, are important for attrition. Significant differences between stayers and leavers in the observables suggest that they could also differ in unobservables. If this is the case, consistent estimations require that attrition be addressed appropriately (see Alderman et al., 2001 for a comparison of attrition rates in developing countries).

To address attrition, I adopt the inverse probability weighting (IPW) method proposed by, for example, Fitzgerald et al. (1998)**.** IPW is based on the key assumption that sample attrition is ignorable with respect to the dependent variable, conditional on the observables in the attrition equation Wooldridge (2002). The IPW procedure consists of two stages. In the first stage, data from the first round are used to estimate the probability of remaining in the survey in the second round. The inverses of the predicted probabilities are used to weight the second-round data, essentially giving more weight to households who are more likely to leave, conditional on observables. I have not pursued a selection on unobservables approach. This stems from the lack of credible exclusion restrictions that would define variables that predict the probability of dropping from the sample but are not associated with children's educational outcomes. All of the regression results reported below are based on the weighted sample.

---Insert Table 3 here---

Table 3 presents Probit estimations of the conditional probabilities of being in the survey in the second round. In addition to household characteristics, I also control for the ability of the enumerators to track the households. Enumerators were selected and assigned to villages based on experience and ethnic background (i.e., each enumerator was required to be able to communicate in the language spoken in the village), but religious beliefs were not a criterion. As the survey required that both the enumerators and the supervisors reside in the villages for a prolonged time[10], religion might have served as one of the networks that enumerators could rely on to track hard-to-find households. Hence, households whose head's religion matches the enumerator's or supervisor's religion would have been more likely to be resurveyed in the second round[11].

(b) Measures of income risk

Where agriculture is essentially rain-fed, rainfall deviations and heterogeneity in households' land holdings (in terms of soil types and topo-sequence) can be used to recover a measure of income shocks. To the extent that production on different types of land responds differently to similar rainfall levels, and land allocation is made at the beginning of the season when the level of rainfall is unknown, the cross-product of soil types and rainfall realization provides a measure of the income shock that is both exogenous and unanticipated (e.g., Alderman and Paxson, 1994; Fafchamps et al., 1998; Paxson, 1992). Furthermore, absent an active land market, a household's stock of land (which may be different from land farmed in any given year) reflects its ability to cope with rainfall risk.

Following this vein of the literature, I use data from 1995 and 2004 to estimate the following regression:

$$y_{ivt} = z_{ivt}\beta_1 + F_{vt}X_{ivt}\beta_2 + \gamma_{vt} + \gamma_i + \varepsilon_{ivt} \qquad (2)$$

where y_{itv} is crop income (total output value, net of all purchased inputs and hired labor), z_{ivt} is a set of household demographic variables, X_{ivt} represents the area of plots of specific soil types cultivated by the farmer, F_{vt} is the deviation of current rainy season rainfall from the long-term mean, γ_{vt} is a village-year fixed effect, γ_i is a household fixed effect and ε_{ivt} is an error term. Households are indexed by i, villages by v and time by t.

---Insert Table 4 here---

The estimation results of Equation (2) are reported in Table 4, using rainfall data provided by Burkina Faso Office of Meteorology which records rainfall data across the country. The first column does not control for aggregate shocks. The second column includes village-year dummies to control for aggregate shocks. The third column allows for the village-specific effects of rainfall deviations. With data for only two years, this last specification assumes that rainfall deviations capture all village fixed effects. However, the income response to rainfall variations interacted with land is stable between columns 2 and 3, suggesting that rainfall deviations are the most important factors explaining yearly variations across villages. Therefore, I treat column 3 as my preferred specification and use these estimates to predict income shocks for the remaining years and derive the variance of income shocks. In the last two rows, F-tests of the joint significance of the excluded instruments are reported. The instruments are jointly significant in all of the regressions. The null hypothesis that these interactions are jointly non-significant is rejected at the one percent level across all specifications (the F statistics range from 8.76 to 12.86). The F-statistic for my preferred specification is 12.86, which is larger than the threshold recommended by Stock and Yogo (2005).

Using estimates from Equation 2, idiosyncratic shocks are given by $F_{vt}X_{(iv)t}\beta_2$. If households have rational expectations concerning the distribution of income shocks due to their expected rainfall (Kazianga and Udry, 2006), then income variance is given as[12]:

$$\widehat{var}(y^T_{ivt+1}) = \frac{1}{33}\sum_{t=1971}^{2003}\left(F_{vt}X_{iv}\beta_2 + \beta_v F_{vt} - (\bar{F}_v X_{iv}\beta_2 + \beta_v \bar{F}_v)\right)^2. \quad (3)$$

The starting year (1971) is the earliest period when rainfall data are available for all villages. The measures of income variance are entirely characterized by landholdings and rainfall deviations and do not require extra information at the household level. Hence, historical land holdings can be used to derive the history of income shocks for each household. Because I control for aggregate shocks (village dummies interacted with year dummies) my measure of income risk relies entirely on idiosyncratic risk. I use the standard deviation of the unpredictable income shocks as a measure of income risk (e.g., Guiso, Jappelli and Terilizze, 1996; Heaton and Lucas, 2000)[13]. The income risk varies from 7,810 to 968,580, with a sample mean of 94,780. In the regressions reported below, income risk has been divided by 10,000.

A concern with my measure of risk is that risk-averse households could change their mixes of land holdings to reduce their exposures to risk. In this case, my estimate would not get at the "raw" exposure to risk. To mitigate this concern, I use total landholdings instead of cultivated land. Assuming that there is no active land market, the household cannot change its mix of land holdings. Thus, for each household, total land holdings provide a proxy for household income volatility over time. Information on land areas and acquisition dates was used to reconstruct the historical land holdings for each household between 1995 and 2004.

---Insert Table 5 here---

Table 5 summarizes average land holdings by household, including number of plots, average area in hectares and means of land acquisition. The figures confirm that the land market is very thin. It is apparent that land is essentially acquired through one's family or through the village as inheritances or gifts. Other means of land acquisition (including borrowing or purchasing) account for a small fraction of the land stock. Information about land areas and acquisition dates was then used to reconstruct the historical land holdings for each household between 1995 and 2004.

5. RESULTS AND DISCUSSION

(a) Observed enrollment in 1995 and 1996-2004 income variance[14]

Because equation (3) captures past income shocks, it may be difficult to distinguish a situation where households choose ex-ante not to enroll their children in school because of anticipated future income shocks from the more conventional explanation where negative income shocks lead to children being withdrawn from school, if children who withdraw are less likely to return even once income has improved[15]. To overcome this issue, I examine the 1995 observed enrollment response to future income risk, by using available information for the period from 1996 to 2004 to evaluate equation (3) [16]. The income risk, evaluated using data from 1996 to 2004, varies between 542 and 1,210,113, with a sample average of 99,896. The intuition is that variations in rainfall observed between 1996 and 2004 would affect enrollment outcomes observed in 1995 only to the extent that households' beliefs about future rainfall variations influence enrollment decisions. It is still possible that past shocks predict future ones. Such a pattern of shocks would, however, be consistent with the argument that households account for future income shocks when making enrollment decisions. A remaining concern is that cumulated previous shocks may have deterred enrollment so that previous shocks (instead of future ones) are keeping children out of school. While I cannot convincingly separate the effects of previous shocks from future ones, I provide some suggestive evidence in section 6 where I discuss some robustness checks.

---Insert Table 6 here---

In Table 6, I report mean marginal effects, estimated using Logit specifications. Because the regressions include predicted variables, I use bootstrapped standard errors[17]. In column 1, I show the combined results for boys and girls. Income variance has a negative effect on enrollment status, and the point estimate is significant at the 10 percent level. I show the results of separate regressions for boys and girls in columns 2 and 3. Income variance has a negative impact on boys' enrollment but has no discernible effect on girls' enrollment. When income risk increases by CFA[18] 10,000, enrollment for boys and girls decreases by 0.033 percentage points (column 1). The effect is, however, driven by boys' enrollment, which would decrease by 0.043 percentage points if income risk were to increase by CFA 10,000.

In columns 4-6, I control for current crop income and proxies for household wealth, including livestock, the value of farm equipment and the value of household durable goods. These variables are potentially endogenous and may be correlated with estimated income risk. The estimates of income risk remain, however, essentially when unchanged, suggesting that the income risk variable is not picking up effects on enrollment that are attributable to current income and/or to current household wealth[19].

(b). Current educational outcomes and income risk

I now use the 2004 cross-sectional survey, which has more detailed information on education. In 2004, in addition to current enrollment, the survey collected information on years of education completed and school-related expenditures for each child who was enrolled at the time of the survey. In addition to income variance, the explanatory variables in all of these regressions include child characteristics (gender, number of siblings of school age, whether a child is a paternal or a maternal orphan), parents' characteristics (whether the father and mother are literate), current household income, household wealth (expressed as the value of durable goods and farm equipment, land area measured in hectares per adult and livestock holdings) and household structure (number of adult males and females, and elderly males and females), as well as village and religion dummies.

---Insert Table 7 here---

The marginal effects from Logit estimations of current enrollment status are displayed in Table 7[20]. Columns 1 and 4 contain estimation results for boys and girls taken together. Columns 2 and 3, and 5 and 6 contain separate estimations for boys and girls, respectively. In the last three columns, I include contemporary income shocks, measured as crop income shocks and livestock losses (from theft and deaths), to control for any contemporary shock effects that might be confounded with the variance effects.

In column 1, the estimated marginal effects imply that a CFA 10,000 increase in income risk is associated with a.007 reduction in boys' and girls' enrollments (significant at the five percent level). The reduction is larger for boys' enrollment (0.013, column 2), and the point estimate is significant at the one percent level. The effect on girls' enrollment (column 3), although negative, is not statistically significant.

In columns 4-6, I also control for current income shocks, approximated by predicted income shocks and livestock losses. The statistical significance is somehow weaker, but the effect of income risk on enrollment still remains quite large: a CFA 10,000 increase in income risk is associated with a 0.012 decrease in enrollment for boys and girls taken together (column 4), and the point estimate is significant at the five percent level. Enrollment decreases by 0.13 for both boys (column 5) and girls (column 6), but only the estimated effect for boys is significant at the 10 percent level.

----Insert Table 8 here---

I now turn to exploring the effect of income risk at the intensive margin. In Table 8, I show Tobit estimates of the effect of income risk on education-related expenditures per student, measured in CFA 1000. I report the average partial effects (e.g. Wooldridge, 2010). The estimation results for boys and girls combined are shown in column 1, for boys only in column 2 and for girls only in column 3. The point estimates indicate that income risk has a negative and significant effect on education-related expenditures. If income risk increases by CFA 10,000, expenditures per student fall by CFA 37.2 for all children (significant at the five percent level), by CFA 47.6 for boys (significant at the 10 percent level) and by CFA 36.1 for girls (not statistically significant). The estimated income effects are larger when current shocks are controlled for (columns 4-6). The point estimate for boys (column 5) is, however, no longer significant at the 10 percent level.

The central hypothesis of the precautionary saving model is that households that are more exposed to income risk accumulate more assets to build a buffer that makes it possible to absorb the income losses associated with negative shocks (e.g., Caballero, 1990). In particular, the canonical model of precautionary saving is that optimal consumption is a function of physical, financial and human assets, as well as income variance. Assuming that children's education enters the utility function separately, the regressions shown in Table 8 provide a direct test of precautionary savings. Controlling for household human, financial and physical assets, I find that higher income volatility is associated with lower current consumption in the form of education-related expenditures. Taken together with the findings on enrollment and years of education completed, it appears that both the precautionary motive and irreversible nature of investments in education combine to keep education levels exceptionally low. Fafchamps and Pender (1997) reach similar conclusions from analyzing investments in irrigation in India.

Current enrollment status and education expenses reflect current household education choices and do not necessarily account for past decisions that could provide useful information about the effects of income uncertainty. To account for previous decisions, I consider the number of years of education completed.

---Insert Table 9 here----

Tobit estimation results (average partial effects) of years of education completed are reported in Table 9. The estimates of the effect of income risk imply that children from households with riskier incomes receive fewer years of education. In columns 1-3, the point estimates for boys and girls combined, for boys only and for girls only, respectively, are all significant at the one percent level. The reported marginal effects indicate that a CFA 10,000 increase in income risk is associated with 0.036 fewer years of education completed for the pooled sample of boys and girls (column 1). For boys and girls only, the reductions are 0.039 and 0.036 years, respectively. After controlling for current income shocks in columns 4-6, the point estimates are larger in magnitude, but the statistical significance weakens for boys. The decrease in years of education completed now corresponds to 0.058 years for boys and girls combined (column 4), 0.04 years for boys (column 5) and 0.087 years for girls (column 6).

In all of the specifications discussed above (Tables 6-9), the coefficient of income risk is identified even after controlling for current household wealth and current income shocks. The effects of current income shocks (approximated by predicted crop income shocks and livestock losses) are consistent with the findings of previous studies, i.e., that negative income shocks are detrimental to child education (e.g., Beegle et al., 2006; Jacoby and Skoufias, 1997). Hence, it is likely that income risk is not only picking up some effects attributable to household wealth or current shocks. Instead, the results support the hypothesis that in addition to realized income shocks, income volatility has separate effects on households' decisions on education. In particular, households with more volatile income invest less in their children's education.

While the girls' human capital has been found to be more elastic to household resources than boys' education (e.g. Alderman and Gertler, 1997), my results are nuanced. At the extensive margin (ever being enrolled and currently enrolled), I found boys' education is more elastic to income risk than girls education. In contrast, at the intensive margin (education related expenditures and number of years of education completed), girls' education is more elastic to income risk than boys' education. One could speculate that boys' labor is more productive in agriculture, and therefore boys' labor is more useful in smoothing income fluctuations than girls' labor. For instance, boys tend to engage in farming and herding, where the returns are higher, whereas girls engage more in household chores, so that they are not as subject to decreased human capital investments due to income uncertainty. My findings are broadly consistent with Edmonds (2006) who found in South Africa that schooling and child labor were more sensitive to the timing of old age pension income for boys than girls.

6. ROBUSTNESS CHECKS

I now provide two forms of robustness checks. First, I verify the validity of the measured income risk given by Equation 3. I test how estimated income risk is correlated with self-reported income shocks, rainfall shocks and land characteristics. Second, I test whether my empirical results are robust to an alternative measure of income risk proposed by Dardanoni (1991).

(a) Validity of measured income risk

I check the robustness of the estimated unpredictable income shocks by examining their correlation with self-reported shocks. The data contain directly solicited information on income shocks between 1995 and 2005. In each household, two adults (the head of household and another adult) were independently asked to rate the years between 1995 and 2004 as good, average or bad. In Table 10, I investigate the extent to which these self-reported income shocks correlate with estimated income risk.

---Insert Table 10 here---

The dependent variable in Table 10 is the estimated standard deviation of income shocks (from Equation 3). To facilitate reading the table, income risk is divided by 10,000. In column 1, the self-reported negative and positive shocks have similar effects on estimated income variance. The estimates also suggest that the estimated income risk is not only picking up negative income shocks, but actually measures income volatility. In column 2, I also control for rainfall measured in the rainy season (May-August) and household landholdings. The estimates confirm that greater rainfall is associated with lower income risk. Moreover, the land topography and texture of the soil substantially influence income risk. In particular, land in the upper and in the lower ends of the topography is associated with lower income risk. Conversely, lateritic land, sandy land and gravel are associated with higher income risk. Overall, it appears that estimated income risk conveys sufficient information on household income volatility.

(b) Alternative measure of income risk

In this section, I re-estimate the regressions shown in tables 6-9 based on an alternative definition of income uncertainty. Following Dardanoni (1991), I divide the sample into many homogeneous groups and use the income variance within each group as an index of earnings uncertainty for each household within that group. To obtain the homogenous groups, I use land size and a Simpson diversity index to approximate the diversity of each household's landholdings. I then construct terciles of total land size and land diversity. This results in 24 groups of households, where households within the same groups have similar landholdings in hectares and similar mix of lands based on the land characteristics (i.e. mid-slope land, plain, gravel, sand, lateritic soils and other soils). The underlying assumption is that in each village, households with similar land size and land diversity face similar income uncertainty.

---Insert Table 11 here---

Consistent with the analysis above, I approximate income risk using the standard deviation of income in each homogenous group. The sample average income risk is 25,000. Regression results using this alternative definition of income risk are summarized in Table 11, where panels A, B, C and D correspond to Tables 6 (the first six columns), 7, 8 and 9, respectively.

The estimated size effects shown in Table 11 are larger than those reported in tables 6-9. However, the results remain qualitatively unchanged. In particular, across the four educational outcomes examined, income risk consistently has a negative effect, and the effect is larger for boys than for girls. Moreover, the effect of income risk remains large and statistically significant after controlling for current income and household wealth.

(c) Effect of income risk on enrollment of younger children

As discussed previously, children who are out of school at the time of the survey may have dropped out or may have ever enrolled because of past shocks. If past shocks are positively correlated with income risk, then my estimates on current enrollment and ever enrolled may be picking up exposure to past shocks.

One could, however, argue that the previous shocks have a minimal effect on enrollment of younger children who were supposed to start school at the time of the survey. For these children, after controlling for current shocks and income risk would matter only if it predicts future income shocks. Although previous income shocks could have pushed household into poverty, current household wealth (landholdings, livestock, farm equipment and household durable goods) would control for household poverty level.

---Insert Table 12 here---

In Table 12, I report Logit estimates (mean marginal effects) of current children who were six and seven years old at the time of the survey. This age range corresponds to the official school entry age and therefore these children would be less likely to have been kept out school because of previous income shocks. The effect of income risk is qualitatively similar to the results shown in Table 7. Income risk has a negative impact on boys' enrollment, and the effect persists even after controlling for current income shocks. This supports the interpretation that income risk has a separate effect on investment in education which is different from the cumulated effects of previous negative income shocks.

7. CONCLUSION

The objective of this paper was to evaluate the extent to which households' income uncertainty influences investments in their children's education. Controlling for current income shocks, household wealth, child and parent characteristics, I find that income standard deviation (my proxy for income uncertainty) has a significant and negative effect on a range of educational outcomes that reflect both current education choices and accumulated education.

The finding that income uncertainty is detrimental to education has both analytical and policy implications. From an analytical perspective, the finding implies that focusing only on households' *ex-post* responses to negative shocks may not account for the full costs of income risk. First, income uncertainty is sufficient to maintain a low enrollment rate, even if negative shocks do not frequently materialize. Second, forward-looking households may allocate child time *ex-ante* (e.g., by only enrolling a few of their children and having the others work full time) to minimize the impact of negative income shocks on school attendance and the probability of dropping out. Non-linearities in the returns to education may exacerbate such behavior. It is then possible that empirical tests may find little (or no) response of education decisions to negative income shocks, while income uncertainty still has a significant negative impact.

The data used come from six villages in three provinces of Burkina Faso. Although these villages and provinces are very diverse, they are not representative of Burkina Faso. Therefore, the conclusions might have a limited external validity. Nevertheless, the results remain very suggestive that high income volatility, in addition to realized income shocks, may explain why a large proportion of school age children are out of school in low income countries.

References

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics* (3), 659–684.

Alderman, H. (1996). Saving and economic shocks in rural Pakistan. *Journal of Development Economics.* 51(2) 343-365

Alderman, H., J. R. Behrman, H. P. Kohler, J. A. Maluccio, and S. C. Watkins (2001). Attrition in longitudinal household survey data: Some tests for three developing-country samples. *Demographic Research* (4), 79–123.

Alderman, H. and P. Gertler. 1997. Family Resources and Gender Differences in Human Capital Investments: The Demand for children's Medical Care in Pakistan. In Haddad, Hoddinott, and Alderman (eds.) *Intrahousehold Resource Allocation in Developing Countries: Models, Methods and Policies*." Johns Hopkins University Press, Baltimore.

Alderman, H. and C. H. Paxson (1994). Do the poor insure? a synthesis of the literature on risk and consumption in developing countries. In E. L. Bacha (Ed.), *Economic in a Changing World*, Volume Development, Trade and the Environment, Chapter 3, pp. 48–78. London: Macillan.

Appleton, S. and A. Balihuta (1996). Education and agricultural productivity: Evidence from Uganda. *Journal of International Development* (3), 415–466.

Beegle, K., R. Dehejia, and R. Gatti (2006). Child labor and agricultural shocks. *Journal of Development Economics* (1), 80–96.

Bennell, P. (2002). Hitting the target: Doubling primary school enrollments in Sub-Saharan Africa by 2015. *World Development* (7), 1179–1194.

Browning, M. and A. Lusardi (1996). Household saving: Micro theories and micro facts. *Journal of Economic Literature* (4), 1797–1855.

Caballero, R. (1990). Consumption Puzzles and Precautionary Savings. *Journal of Monetary Economics*. 25(1) 113-136.

Canagarajah, S., D. Mazumdar, and X. Ye (1998). *The Structure and Determinants of Inequality and Poverty Reduction in Ghana*. Washington DC: The World Bank.

Carroll, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *Quarterly Journal of Economics* , 1–55.

Carroll, C. D. and M. S. Kimball (2001). Liquidity constraints and precautionary saving. Working Paper 2001, Department of Economics, The Johns Hopkins University, Baltimore, MD.

Case, A. and C. Ardington (2006). The impact of parental death on school outcomes: Longitudinal evidence from South Africa. *Demography* (3), 401–420.

Chetty, R. and A. Looney (2005). Consumption smoothing and the welfare consequences of social insurance in developing economies. *National Bureau of Economic Research Working Paper 11709*.

Dardanoni, V., (1991). Precautionary saving under income uncertainty: A cross-sectional analysis. *Applied Economics*. 23(2), 153–160.

Deaton, A. (1992). *Understanding Consumption*. New York: Oxford University Press.

Dercon, S. (2005). *Insurance Against Poverty*. Oxford University Press.

Dercon, S. and L. Christiaensen (2010). Consumption risk, technology adoption and poverty traps: evidence from Ethiopia. *Journal of Development Economics*.

Duryea, S. (1998). Children's advancement through school in Brazil: The role of transitory shocks to household income. Working Paper 376, Inter-American Development Bank, Office of the Chief Economist.

Duryea, S. and M. Arends-Kuenning (2003). School attendance, child labor and local labor market fluctuations in urban Brazil. *World Development* (7), 1165–1178.

Edmonds, E. V. 2006. Child labor and schooling responses to anticipated income in South Africa. *Journal of Development Economics*. 81(2) 386-414

Fafchamps, M. and J. Pender (1997). Precautionary Saving, Credit Constraints, and Irreversible Investment: Theory and Evidence from Semiarid India. *Journal of Business and Economic Statistics.* 15(2), 180-94

Fafchamps, M., C. Udry, and K. Czukas (1998). Drought and saving in West Africa: Are livestock a buffer stock? *Journal of Development Economics* (2), 273–305.

Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998). An analysis of sample attrition in panel data. *Journal of Human Resources* (3), 251–299.

Fitzsimons, E. (2007). The effects of risk on education in Indonesia. *Economic Development and Cultural Change* , 1–25.

Guiso, Luigi, Tullio Jappelli, and Daniele Terilizzesse (1996). Income risk, borrowing constraints and portfolio choice, *American Economic Review* 86, 158-172.

Heaton, John, and Deborah Lucas (2000), Portfolio choice and asset prices: The importance of entrepreneurial risk, *Journal of Finance* 55, 1163-1198.

Jacoby, H. (1994). Borrowing constraints and progress through school: Evidence from Peru. *The Review of Economics and Statistics* (1), 151–160.

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

Jensen, R. (2000). Agricultural volatility and investments in children. *American Economic Review Papers and Proceedings* (2), 399–404.

Joliffe, D. (1998). Skills, schooling and household income in Ghana. *World Bank Economic Review* (1), 81–104.

Kazianga, H. (2004). Schooling returns for wage earners in Burkina Faso. *Economic Growth Center, Yale Univeristy Discussion Paper 892*.

Kazianga, H. and C. Udry (2006). Consumption smoothing? livestock, drought and insurance in rural Burkina Faso. *Journal of Development Economics* , 413–446.

Kimball, M. (1991). Precautionary savings in the small and the large. *Econometrica* (1), 53–73.

Leland, E. H. (1968) Saving and Uncertainty: The Precautionary Demand for Saving. *The Quarterly Journal of Economics*, 82(3) pp. 465-473

Morduch, J. (1999. Between the state and the market: Can informal insurance patch the safety net? *World Bank Research Observer* (2), 187–207.

Park, A. F. (2006). Risk and household grain management in developing countries. *Economic Journal*. 116(514): 1088-1115

Paxson, C. H. (1992). Using weather variability to estimate the response of saving to transitory income in Thailand. *American Economic Review* (1), 15–33.

Rosenzweig, M. R. and K. I. Wolpin (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in India. *Journal of Political Economy* (2), 223–44.

Sawada, Y. (2003). Income risks, gender, and human capital investment in a developing country. Mimeo.

Sawada, Y. and M. Lokshin (2009). Obstacles to school progression in rural Pakistan: An analysis of gender and sibling rivalry using field survey data. *Journal of Development Economics.* 88(2) 335–347

Schultz, T. P. (1988). Education investments and returns. In H. Chenery and T. N. Srinivassan (Eds.), *Handbook in Development Economics*. Amsterdam: North Holland: Elsevier.

Schultz, T. Paul (2004). Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education. *Journal of African Economies*, 13, pp. 95-148.

Schultz, T. P. (2003). Human resource research agenda and survey evidence in Africa: Monitoring efficiency and improving equity. Center Discussion Paper 875, Economic Growth Center, Yale University, New Haven, Connecticut, USA.

Shady, N. R. (2003). Convexity and sheepskin effects in the human capital earnings function: Recent evidence for Filipino men. *Oxford Bulletin of Economics and Statistics* (2), 171–196.

Stock, J. H. and M. Yogo. 2005. Asymptotic Distributions of Instrumental Variables Statistics with Many Instruments in Identification and Inference for Econometric Models: *Essays in Honor of Thomas Rothenberg*, ed. D.W.K. Andrews and J.H. Stock. Cambridge: Cambridge University Press.

Townsend, R. M. (1994, May). Risk and insurance in village India. *Econometrica* (3), 539–91.

UNESCO (2005). Global education digest 2005: Comparing education statistics across the world. Technical report, UNESCO Institute for Statistics, Montreal: Quebec, Canada.

Wooldridge, J. M. (2010) *Econometric Analysis of Cross Section and Panel Data*. 2nd Ed. The MIT Press

Wooldridge, J. (2002). Inverse probability weighted M-estimators for sample selection, attrition, and stratification. *Portuguese Economic Journal* (2), 117–139.

Zoungrana, A., B. Sawadogo, S. Zerbo, I. Tagnan, N. Terpend, S. Sanon, and D. Michiels (1999). Plan national d'organisation et de coordination des secours d'urgence et de réhabilitation: Volet sécurité alimentaire. Rapport technique, Comité National de Secours d'Urgence et de Réhabilitation, Ouagadougou, Burkina Faso.

[1]See Dercon (2005) for a recent review of the literature on income risk in developing countries.

[2]Note that this may still be sub-optimal because the marginal returns are not necessarily equalized to the marginal costs of investments.

[3]For instance, Bennell (2002) reports that completion of secondary school (or 6-8 years of education) is the minimum entry requirement for formal sector jobs in most sub-Saharan African countries. For the specific case of Burkina Faso, Kazianga (2004) estimates the rate of return on education to be about 9 percent for primary education for men and women, 16.4 percent for women and 14.3 percent for men at the secondary level, and 18.1 percent for women and 23.4 percent for men at the university level. After controlling for endogenous choice between the private and the public sectors, the rate of return on primary education is almost zero for women.

[4]See Chetty and Looney (2005) for a recent related discussion.

[5]Given the dwelling pattern, especially in the Sudanian region, distance to school is likely to vary substantially across households.

[6]This is because cotton farming necessitates the use of modern inputs (fertilizers and pesticides). Presumably, farmers with formal education could learn how to use these inputs more quickly.

[7]Where cash crop −cotton− opportunities exist.

[8] Table A-1 in the appendix shows the descriptive statistics for all variable used in the estimations, including crop income per adult equivalent.

[9]Annual attrition rate is calculated as $1-(1-q)^{1/T}$, where q is the overall attrition rate and T is the number of years covered by the panel (Alderman et al., 2001).

[10] There were four survey teams. Each survey team consisted of one supervisor and four enumerators, with two enumerators per village. Each team spent about two months at the survey

sites.

[11]I include these variables to have a richer specification of the probability of attrition. These variables cannot be treated as instruments, as they do not predict the probability of moving out of the survey area.

[12]The earliest year that rainfall data are available for all six villages is 1971.

[13]Guiso, Jappelli and Terilizze (1996) and Heaton and Lucas (2000) use the standard deviation of labor income growth to measure risk.

[14] All regressions discussed in the text control for attrition. Estimations which do not correct for attrition are shown in Tables A2-A5. Overall, ignoring attrition biases the point estimates slightly upward. The standard errors are also relatively smaller resulting in more coefficients which are statistically significant. The main story remains, however, essentially unchanged whether attrition is controlled for or not. .

[15]To help picture this concern, consider two children *i* and *j* who are six years old in 1995. Over the next three years, child *i*'s family has incomes of 500, 0, 1000. Child *j*'s family has incomes of 500, 500, and 500. The means are the same, but the variance differs. Now we observe these children in 2004 and find that child *i* is less likely to be currently enrolled in school (in addition to having fewer years of schooling and a family that spends less on schooling). My interpretation is that child *i*'s family knew that its income was more risky, so it chose to invest little in child education, perhaps preferring to put the resources away for savings in case of a shock. However, it could also be that when incomes are very low, kids are taken out of school, and once withdrawn from school, are less likely to return.

[16]The income variance is measured as:

$$\widehat{var}(y^{F}_{hvt+1}) = \frac{1}{9}\sum_{t=1996}^{2004}\left(F_{vt} X_{hvt}\hat{\beta}_{k} + \hat{\beta}_{v} F_{vt}\right.$$

---

[17] Standard errors are bootstrapped for all regressions which include predicted variables.

[18] CFA (Communaute Financiere Africaine) is the local the currency, and at the time of the survey, the exchange rate was about one US dollar for CFA 500.

[19] The main variable of interest (income risk) may still be endogenous because it is constructed from rainfall which is itself a choice variable, due to purposive migration to more favorable weather (including rain) conditions. Note, however, that all estimations include village fixed effects. Hence, village level time-invariant weather conditions (which could attract migrants) are controlled for.

[20]The estimations do not account for late entry (i.e., some school age children who are not enrolled may enroll in the future). Likewise, I do not address right censoring (that years of education completed is at least equal to current years of education for those who are still attending school) when estimating Tobit regressions of years of education completed. Instead, I include age dummies in all regressions.

Table 1: Summary of education outcomes in 1995 and 2004 (children 7-15 years old)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Village						6 villages
		Niega	Kognere	Belehede	Petega	Kerena	Dissankuy	
Panel A: ever enrolled, 1995								
	Boys	29.27	54.00	30.43	46.38	76.92	39.24	37.24
	Girls	16.18	6.94	22.97	25.00	42.22	19.72	20.63
	Boys&Girls	23.33	10.32	26.57	37.21	58.33	30.00	29.09
	n	139	117	133	109	78	139	715
Panel B: ever enrolled, 2004								
	Boys	20.31	24.21	49.38	44.23	78.72	40.00	40.08
	Girls	10.67	14.46	22.83	40.00	64.29	35.71	28.66
	Boys&Girls	15.11	19.66	35.26	41.96	71.91	38.04	34.44
		128	165	160	104	82	255	894
Panel C: currently enrolled, 2004								
	Boys	10.94	20.00	33.33	34.62	61.70	32.67	30.47
	Girls	4.00	13.25	17.39	36.67	42.86	27.78	21.97
	Boys&Girls	7.19	16.85	24.86	35.71	52.81	30.43	26.27
	n	128	165	160	104	82	255	894
Panel D: education expenditures (CFA 1,000)								
Boys	unconditional mean	1.318	1.918	1.154	0.263	1.226	2.433	1.629
	conditional on being enrolled	6.132	9.375	2.320	0.760	1.987	5.691	4.254
Girls	unconditional mean	0.247	0.639	0.256	0.210	0.749	2.533	0.959
	conditional on being enrolled	2.833	4.495	1.141	0.573	1.625	8.412	3.928
Boys&Girls	unconditional mean	0.578	1.213	0.643	0.221	0.730	2.375	1.158
	conditional on being enrolled	5.143	7.586	1.881	0.657	1.848	6.824	4.119
n		128	165	160	104	82	255	894

Table 2: Summary statistics by attrition status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Attrition	Non food cons.	Female head	Head age	Crop income	Cattle	Goat/Sheep	Asset value	adults	pre-school	boys	girls	hhsize
Stayers	10.882	0.045	48.757	27817.640	0.462	1.571	13608.070	4.243	2.271	1.283	1.279	9.077
Leavers	10.706	0.135	46.231	20502.350	0.532	1.902	8426.704	3.135	1.750	0.923	0.788	6.596
Total	10.851	0.060	48.318	26545.420	0.474	1.629	12706.960	4.050	2.181	1.221	1.194	8.645
t-test (absolute t value)	0.061	2.500	1.125	1.458	0.458	0.853	0.854	3.875	2.005	1.755	2.792	3.4638

Table 3: Determinants of attrition

	(1)	(2)
	leaver=1	leaver=1
Head Female	0.669	0.72
	[0.352]*	[0.456]
Head literate	0.138	0.056
	[0.104]	[0.141]
Head age	-0.046	-0.065
	[0.037]	[0.040]
Head age squared	0	0.001
	[0.000]	[0.000]*
Crop Income	-0.009	-0.011
	[.004]**	[.004]**
Cattle	0.087	0.088
	[0.050]*	[0.070]
Goat and Sheep	-0.003	0.017
	[0.043]	[0.032]
Asset	-0.001	-0.001
	[.006]	[0.007]
Pre-school	0.005	0.010
	[0.039]	[0.050]
Boys 7-15	-0.044	-0.019
	[0.094]	[0.070]
Girls 7-15	-0.207	-0.221
	[0.099]**	[0.102]**
Adults	-0.099	-0.112
	[0.064]	[0.062]*
Head Christian		1.059
		[0.484]**
Enumerator Christ.		0.405
		[0.559]
Head christ. & Enum christ.		-0.727
		[0.964]
Head trad. & Enum christ.		-1.843
		[0.702]***
Controler chris.		0.383
		[0.183]**
Head christ. & Enum christ.		-0.774
		[0.545]
Head trad. & Enum christ.		0.304
		[0.197]
Constant	0.416	1.325
	[0.931]	[1.079]
Observations	299	299
chi2-test instrum.		27.39

Table 4: Determinants of income

VARIABLES	(1) Specfication1	(2) Specification3	(3) Specification3
Rainfall interacted with type of land			
Mid-slope	0.0923***	0.103***	0.103***
	[0.0219]	[0.0221]	[0.0221]
Plain	0.121***	0.149***	0.149***
	[0.0231]	[0.0236]	[0.0236]
Gravel	-0.104***	-0.137***	-0.137***
	[0.0296]	[0.0301]	[0.0301]
Sand	-0.0735***	-0.106***	-0.106***
	[0.0249]	[0.0259]	[0.0259]
Laterite	-0.193***	-0.224***	-0.224***
	[0.0292]	[0.0299]	[0.0299]
Other	-0.128***	-0.148***	-0.148***
	[0.0211]	[0.0215]	[0.0215]
Rainfall deviation	1.399**	0.564***	0.788***
	[0.597]	[0.105]	[0.145]
hh composition	yes	yes	yes
Age dummies	yes	yes	yes
hh wealth	yes	yes	yes
Village dummy	yes	yes	yes
Year dummy	yes	No	yes
Village*year dummy	no	No	yes
Village*devrain	no	no	no
Constant	-254.5*	3.740	-35.11
	[135.6]	[104.0]	[99.39]
Observations	660	660	660
R-squared	0.830	0.839	0.839
F-test instrum.	8.755	10.22	12.86

Robust standard errors in brackets (standard errors are clustered at the household level)
*** p<0.01, ** p<0.05, * p<0.1

Table 5: land size per household and means of land acquisition

Villages		Hh land	Received from: Family	Received from: Village	Other sources
Niega	# plots	9.91	9.26	0.40	0.25
	Area (ha)	4.41	4.08	0.21	0.13
Kognere	# plots	7.58	6.27	1.23	0.08
	Area (ha)	3.26	2.75	0.49	0.01
Belehede	# plots	3.81	3.46	0.35	0.00
	Area (ha)	7.01	6.21	0.80	0.00
Petega	# plots	2.57	2.11	0.41	0.05
	Area (ha)	4.18	2.96	1.08	0.14
Kerena	# plots	4.58	3.15	0.55	0.89
	Area (ha)	4.91	3.46	0.45	1.00
Dissankuy	# plots	7.32	3.09	3.34	0.88
	Area (ha)	9.20	3.41	4.25	1.53
Total	# plots	6.37	4.80	1.19	0.38
	Area (ha)	5.61	3.77	1.33	0.51

Other sources includes borrowing, renting and purchasing

Table 6: Logit estimation of ever having enrolled based on 1995 data (mean marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income risk	-0.0327**	-0.0426**	-0.0227	-0.0269*	-0.0442**	-0.0144
	[0.01503]	[0.02037]	[0.03045]	[0.01428]	[0.02039]	[0.02841]
Current Income				0.0000	-0.0003	0.0005
				[0.00064]	[0.00100]	[0.00101]
Male	0.1472***			0.1519***		
	[0.03531]			[0.03458]		
Hh. Composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
Household wealth	no	no	no	yes	yes	yes
Vill. fixed effects	yes	yes	yes	yes	yes	yes
Observations	715	369	346	715	369	346

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.. Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female seniors (65 and older), the dependency ratio (children and elderly divided by adults), child gender, father's and mother's literacy and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Table 7: Logit estimation of current enrollment using 2004 data (mean marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income variance	-0.0069**	-0.0131***	-0.0024	-0.0119**	-0.0129*	-0.0129
	[0.003]	[0.005]	[0.005]	[0.005]	[0.008]	[0.008]
Current income shock				0.0008	0.0002	0.0013
				[0.001]	[0.001]	[0.001]
Livestock losses				0.0203***	0.0235**	0.0202**
				[0.006]	[0.011]	[0.009]
hh composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
hh. wealth	yes	yes	yes	yes	yes	yes
vill. Fixed effects	yes	yes	yes	yes	yes	yes
Observations	894	454	440	894	454	440

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.
Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elderly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Table 8: Tobit estimation of education expenditures using 2004 data (mean marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income risk	-0.0372**	-0.0476*	-0.0361	-0.0692***	-0.0622	-0.1032**
	[0.016]	[0.029]	[0.022]	[0.025]	[0.041]	[0.044]
Current income shock	0.6367***	0.3577	0.7165***	0.5867***	0.2697	0.6602***
	[0.174]	[0.319]	[0.245]	[0.176]	[0.323]	[0.251]
Livestock losses				0.0063*	0.0049	0.0095*
				[0.004]	[0.007]	[0.005]
hh composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
hh. wealth	yes	yes	yes	yes	yes	yes
vill. Fixed effects	yes	yes	yes	yes	yes	yes
Observations	894	454	440	894	454	440

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.
Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elederly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Table 9: Tobit estimation of years of education completed using 2004 data (mean marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income risk	-0.0356***	-0.0392***	-0.0356***	-0.0584***	-0.0401*	-0.0866***
	[0.009]	[0.015]	[0.013]	[0.014]	[0.022]	[0.024]
Current income shock				0.0043**	0.0016	0.0072**
				[0.002]	[0.004]	[0.003]
Livestock losses				0.0894***	0.1240***	0.0765**
				[0.023]	[0.041]	[0.030]
hh composition	Yes	yes	yes	yes	yes	yes
Age dummies	Yes	yes	yes	yes	yes	yes
hh wealth	Yes	yes	yes	yes	yes	yes
vill. fixed effects	Yes	yes	yes	yes	yes	yes
Observations	894	454	440	894	454	440

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.
Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elderly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Table 10: Correlation between estimated income risk, and self-reported income shocks, rainfall and household landholdings

	(1)	(2)
negative shock	1.7994***	1.0493**
	[0.627]	[0.485]
positive shock	1.9164***	0.7871**
	[0.512]	[0.396]
Rainfall		-0.0163***
		[0.002]
upper slope		-1.0604***
		[0.085]
low land		-1.2971***
		[0.100]
mid slope		0.5481
		[0.346]
plain		0.0291
		[0.095]
Laterite		1.4734***
		[0.117]
Sandy land		1.5569***
		[0.105]
Gravel		1.8969***
		[0.107]
Other land		1.3521***
		[0.082]
Constant	3.8167***	8.9123***
	[0.414]	[0.874]
Observations	660	660
R-squared	0.163	0.491

Robust standard errors in brackets (standard errors are clustered at the household level)

*** p<0.01, ** p<0.05, * p<0.1

Table 11: Estimation results based on an alternative measure of income risk

	(1)	(2)	(3)	(4)	(5)	(6)
	boys & Girls	boys	girls	boys & girls	boys	girls
Panel A: ever enrolled, 1995						
Income risk	-0.0096***	-0.0132***	-0.0045	-0.0077**	-0.0109**	-0.0026
	[0.00324]	[0.00475]	[0.00460]	[0.00317]	[0.00481]	[0.00427]
Current income shock				0.0006	-0.0019	0.0038
				[0.00294]	[0.00459]	[0.00357]
Panel B: current enrollment, 2004						
Income risk	-0.0019	-0.0064**	0.0041*	-0.0019	-0.0062**	0.0038*
	[0.002]	[0.003]	[0.002]	[0.002]	[0.003]	[0.002]
Current income shock				0.0209***	0.0260***	0.0195**
				[0.006]	[0.008]	[0.008]
Panel C: Current education-related expenditures, 2004						
Income risk	-0.0143	-0.0435**	0.0022	-0.0147	-0.0430**	0.0017
	[0.016]	[0.021]	[0.004]	[0.016]	[0.021]	[0.004]
Current income shock				0.0665**	0.1023	0.0513**
				[0.034]	[0.068]	[0.021]
Panel D: years of education completed, 2004						
Income risk	-0.0135*	-0.0248**	0.0053	-0.0134*	-0.0239**	0.0043
	[0.007]	[0.011]	[0.010]	[0.007]	[0.010]	[0.010]
Current income shock				0.1090***	0.1393***	0.0954***
				[0.025]	[0.039]	[0.031]

Robust standard errors in brackets. Standard errors are clustered at the village level.

*** p<0.01, ** p<0.05, * p<0.1
Regressions also control for child age dummies, household composition and wealth as described above.

Table 12: Income risk and current enrollment of children aged 7 and 8 years (Logit marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income risk	-0.0150***	-0.0217**	-0.0058	-0.0200	-0.0337**	-0.0021
	[0.005]	[0.010]	[0.004]	[0.013]	[0.015]	[0.014]
Current income shock				0.0019	0.0050**	0.0034**
				[0.001]	[0.002]	[0.001]
Livestock losses				-0.0312*	-0.0239	0.0182
				[0.019]	[0.029]	[0.022]
hh composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
hh. wealth	yes	yes	yes	yes	yes	yes
vill. Fixed effects	yes	yes	yes	yes	yes	yes
Observations	187	102	85	187	102	85

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.
Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elderly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Appendices:

Table A-1 summary statistics of all variables used in the estimations.

	(1)	(2)	(3)	(4)
Variables	Mean	St. deviation	Minimum	Maximum
Dependent variables				
Unconditional education expenses (CFA 1'000 per child)	1.107	4.273	0.000	5.175
Currently enrolled (0/1)	0.263	0.440	0.000	1.000
Ever been enrolled (0/1)	0.344	0.475	0.000	1.000
Years of education (years)	1.135	1.891	0.000	8.000
Independent variables				
Income risk (10'000)[a]	9.478	16.970	0.781	96.858
Income risk (10’000)[b]	12.691	20.779	0.054	121.011
Income risk (10’000)[c]	26.612	18.994	8.337	81.986
Livestock losses (CFA 1'000 per AE)	1.556	4.108	0.000	66.606
Current Income (CFA 1'000 per AE)	21.572	33.597	0.000	194.017
Land holdings (hectares per AE)	0.703	0.927	0.000	6.578
Goat/sheep (number hh)	11.565	16.838	0.000	110.000
Cattle (number hh)	5.048	12.254	0.000	94.000
Durable goods (CFA 1'000 per AE)	8.959	14.336	0.000	92.734
Farm equipment (CFA 1'000 per AE)	1.909	5.908	0.000	132.365
Age (years)	10.960	2.637	7.000	15.000
Head child	0.688	0.464	0.000	1.000
number of other boys	0.640	1.389	0.000	9.000
Number of other girls	0.552	1.103	0.000	8.000
Child is Male	0.507	0.500	0.000	1.000
Mother is literate	0.037	0.189	0.000	1.000
Maternal orphan	0.057	0.232	0.000	1.000
Father is literate	0.101	0.302	0.000	1.000
Paternal orphan	0.073	0.261	0.000	1.000
Number of adult males	2.322	1.881	0.000	12.000
Number of adult females	3.265	2.468	0.000	14.000
Number of senior men	0.165	0.383	0.000	2.000
Number of senior women	0.235	0.472	0.000	2.000
Dependence ratio	0.572	0.136	0.100	1.000

a: income risk is measured as the standard deviation of income shocks, expressed in 10,000, and using rainfall information from 1971 to 2003.

b: income risk is measured as the standard deviation of income shocks, expressed in 10,000, and using rainfall information from 1996 to 2004.

c: income risk is measured as the standard deviation of income, expressed in 10,000. Income variance is calculated based on the method proposed by Dardanoni (1991), and described in page 16.

Table A2: Logit estimation of ever having enrolled based on 1995 data (mean marginal effects) without correcting for attrition

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income risk	-0.0347**	-0.0471**	-0.0224	-0.0275*	-0.0466**	-0.0127
	[0.01559]	[0.02030]	[0.03130]	[0.01491]	[0.02091]	[0.03021]
Current Income				-0.0001	-0.0006	0.0003
				[0.00064]	[0.00109]	[0.00078]
Male	0.1510***			0.1541***		
	[0.03289]			[0.03254]		
Hh. Composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
Household wealth	no	no	no	yes	yes	yes
Village fixed effects	yes	yes	yes	yes	yes	yes
Observations	715	369	346	715	369	346

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.
Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elderly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Table A-3: Logit estimation of current enrollment using 2004 data (mean marginal effects) without correcting for attrition

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income variance	-0.0062**	-0.0122**	-0.0012	-0.0115**	-0.0128*	-0.0125
	[0.003]	[0.005]	[0.005]	[0.005]	[0.008]	[0.008]
Current income shock				0.0008	0.0003	0.0014
				[0.001]	[0.001]	[0.001]
Livestock losses				0.0193***	0.0222*	0.0197*
				[0.007]	[0.012]	[0.011]
hh composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
hh. wealth	yes	yes	yes	yes	yes	yes
vill. Fixed effects	yes	yes	yes	yes	yes	yes
Observations	894	454	440	894	454	440

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.
Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elderly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Table A-4: Tobit estimation of education expenditures using 2004 data (mean marginal effects) without correcting for attrition

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income variance	-0.0419***	-0.0557**	-0.0372	-0.0801***	-0.0782*	-0.1067**
	[0.016]	[0.028]	[0.023]	[0.024]	[0.044]	[0.042]
Current income shock	0.6379***	0.2990	0.7392***	0.5789***	0.1908	0.6813**
	[0.179]	[0.316]	[0.285]	[0.185]	[0.313]	[0.296]
Livestock losses				0.0075**	0.0066	0.0100**
				[0.003]	[0.007]	[0.005]
hh composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
hh. wealth	yes	yes	yes	yes	yes	yes
vill. Fixed effects	yes	yes	yes	yes	yes	yes
Observations	894	454	440	894	454	440

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.
Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elderly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.

Table A-5: Tobit estimation of years of education completed using 2004 data (mean marginal effects) without correcting for attrition

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys & Girls	Boys	Girls	Boys & Girls	Boys	Girls
Income variance	-0.0361***	-0.0414***	-0.0338***	-0.0623***	-0.0475**	-0.0883***
	[0.009]	[0.015]	[0.012]	[0.014]	[0.023]	[0.023]
Current income shock				0.0049**	0.0027	0.0076***
				[0.002]	[0.004]	[0.003]
Livestock losses				0.0811***	0.1090**	0.0701**
				[0.025]	[0.044]	[0.034]
hh composition	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
hh. wealth	yes	yes	yes	yes	yes	yes
vill. Fixed effects	yes	yes	yes	yes	yes	yes
Observations	894	454	440	894	454	440

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Standard errors are bootstrapped (500 replications), accounting for clustering at the village level.

Household composition includes the number of male and female children (5-16 years old), male and female adults (16-65 years old), male and female elderly (65 and older), the dependency ratio (children and seniors divided by adults), child gender, father's and mother's literacy, and whether the child is an orphan. Household wealth variables include landholdings, livestock (number of cattle, goats and sheep), value of household durable goods and value of farm equipment.