Household Income Shocks and Sibling Composition: evidence from rural Tanzania

Mathilde VALERO*

Aix-Marseille University, CNRS, EHESS, Centrale Marseille, AMSE.

Preliminary version: *October 2018*

Abstract

Exploiting exogenous variation in drought conditions across districts in Tanzania, this paper investigates the effects of climate change on children's education according to the number of siblings and their gender composition. The study highlights how liquidity-constrained households respond to income shocks when formal insurance schemes are not available. They select the children on which they invest according to their perceived educational returns. Whereas boys are not affected, the impact of drought on girls' likelihood of school dropout depends on their younger siblings' gender composition. Younger sisters offset the negative shock on girls, as they offer less competition for limited educational resources, due to higher uncertainty about their probability of educational success. Conversely, older sisters disfavor later-born daughters and reduce their likelihood of being in school during time of hardship. I find that drought decreases the educational attainment of girls with only older sisters by more than 1.3 years, compared to girls with only older brothers. Finally, I demonstrate that underpriviledged daughters help the family overcome climate shocks by increasing their work in agriculture and, for some of them, by leaving the household with their early-marriage.

JEL Classification: *D1, 01, I0.*

Keywords: *Household Income Shocks, Sibling composition, Education, Labor, Marriage.*

*mathilde.valero@univ-amu.fr

1 Introduction

This paper investigates the role of income shocks in altering parental investment decisions in children's human capital, depending on sibling composition.

According to the Unesco Institute for Statistics, Sub-Saharan Africa has the highest rates of education exclusion with over one-fifth of children between 6 and 11 years old being out of school in 2015 and one-third of youth between 12 and 14. Girls are still more likely than boys to suffer from inequality, as they have 1.5 times more chances of being out of primary school (UIS, 2017). Differences in labor market opportunities, cultural preferences, limited access to school and teachers eligibility requirements, or adult literacy have played an important role in explaining region and gender disparities in educational attainment (Orazem and King, 2007; Handa, 2002; Huisman and Smits, 2009). Recently, the international community aknowledged the vulnerability of Sub-Saharan Africa to climate shocks as a determinant preventing the realization of the Millennium Development Goals (Kreft et al., 2010), such as universal primary education or gender equality and women's empowerment. Climate variability is a growing development issue in Southern and Eastern Africa, experiencing more recurrent droughts and heavy rainfall during the last 30 to 60 years (Field, 2014). Being credit constrained for the majority of households living in these rural areas, families are forced to resort to other mechanisms to cope with income variations. Adjusting labor-supply, reducing food consumption, shrinking expenditures in education, in health or re-organizing the household through marriage are different options to insure subsistence needs against income shocks, which could be damaging to child educational attainment.

More importantly, the parents' reaction to adverse shocks mainly depends on within-family patterns. However, much less is known about whether climate shocks impact children differently in the family, keeping girls or the younger ones out of school. Indeed, natural disasters increase poverty (Carter et al., 2007) and poverty forces many families to choose which of their children to send to school. Parish and Willis (1993) demonstrate that children's educational outcomes depend on economic security and the nature of siblings, who must share family resources. Whereas in wealthy families siblings had no effect on educational opportunity, families with low incomes chose more carefully who in the home got educated. [1] Conversely, Lafortune and

[1] The results show that early born children in large families do poorly, and particularly poorly if they are female in Taiwan. Their model can adapt to economies where concerns about old age support are not severe, which is not the case in developing societies such as Tanzania.

Lee (2014) provide evidence that parents unable to finance their children's education may rely on some of their older children's labor income as a source of funding. Overall, these studies suggest that altruistic parents who attempt to finance optimal investment in their children's human capital are frustrated by credit constraints. In this end, this paper answers the following questions: Are climate shocks impacting children differently within the family? How does income instability affect a child's education according to the number and gender of siblings?

How families will react to income shocks and allocate resources is not trivial (Behrman, 1997). When confronted by scarce resources, parents can invest in more education for children with higher returns to schooling, that is children with the greatest academic potential and higher future employment opportunities (Becker and Tomes, 1976). These strategies are based on the expectation that all family members will reap the benefits from the success of any one family member. In contrast, families can reduce educational resources on offsprings with higher earnings potential as a concern of equity (Behrman et al., 1982).

Two possible mechanisms involving in a negative productivity shock can be highlighted through the literature; the substitution and income effect (Cogneau and Jedwab, 2012). The first one may affect human capital investment through wages by changing the relative price of time. For example, Shah and Steinberg (2017) identify that during a drought year, agricultural productivity decreases, which reduces opportunity costs of schooling. Since human capital production becomes less expensive, droughts lead to substitute agricultural work into higher human capital investment. In addition, parents might have time to devote to their children's education, due to fewer outside opportunities at the time of drought. In contrast, the income effect may negatively affect children's human capital as households do not have the capacity to pay for schooling. Indeed, decreasing school attendance appears to play a significant role in the self-insurance strategy of poor households from rainfall in rural India (Jacoby and Skoufias, 1997). Moreover, Jensen (2000) shows that children living in regions that experienced adverse weather shocks in Côte d'Ivoire had lower investments in education and suffered from malnutrition in the short-term. Rainfall shocks have also long-lasting consequences in Latin America, decreasing children's human capital accumulation (Caruso, 2017). What is less clear is whether specific children are associated with deeper effect of environmental shocks on their educational attainment. Recent papers attempted to dissociate impacts by child's gender, finding most of the time that girls suffer more from crops shocks in Tanzania due to cultural biases (Bandara et al., 2015), or from rainfall shocks that raise their early marriage (Corno and Voena, 2016).

Similarly, Björkman-Nyqvist (2013) pinpoints a significant negative effect on girls' education in Uganda, where values of child labor differ across sexes. Conversely, positive shocks, measured through early-life rainfall, improve girls' educational attainment, reflecting gender biases in household resource allocation (Maccini and Yang, 2009). However, the literature has neglected differential effects of shocks across children within the family, assuming that children with few or many siblings, sisters or brothers, are equally affected. One exception is the paper of Thomas et al. (2004), which investigates the consequences of the 1998 Indonesian financial crisis on children's education taking into account the number of younger and older household members. They find that the crisis decreases especially the enrollment of younger children, keeping the older children in school at the expense of the younger ones. The only similarity with my paper is the inclusion of sibship size in the analysis of income shocks.

To the best of my knowledge, there is no empirical evidence on the impact of climate variability on education taking into account siblings' characteristics, irrespective of their migration experience or whether siblings are still alive at the time of survey. This research contributes to the literature on climate variability, allowing heterogenous effect of shocks by the number of siblings as well as their gender composition. In this setting, I identify who may have suffered disproportionately large effects. Also, the paper provides additional evidence on the effect of climate change on children's education, using a more reliable and consistent measure of weather shocks. Indeed, the SPEI drought index is based not only on precipitation but also on evapotranspiration data as to reflect local drought conditions. Indeed, temperature variations, which are not included in related papers, have been shown to better explain spatial and temporal variation in agricultural income (Dell et al., 2012; Lobell et al., 2011). Besides, the SPEI indicator has multi-scalar characteristics which enables the identification of droughts with different variation and magnitude in the context of global warming (Harari and La Ferrara, 2013; Almer et al., 2017). Finally, the study is related to the literature on how gender gaps in perceived returns to investment cause unequal resource allocation across siblings.

This study provides a microeconomic panel model examining drought consequences on boys and girls' education, conditional on the number and sex ratio of older and younger siblings. Exogenous variation in rainfall across districts in a rural region of Tanzania is exploited as a reduced-form instrument for household income variability. I outline an empirical strategy based on interactions between drought and variables reflecting the effect of sibling size separately from sibling gender. I use constructed individual data over the period 1970-2010 from retrospective information on the school-leaving age and the level of education of the Kagera Health Development Survey (KHDS) database. Responses to economic shocks in the short run are firstly considered, focusing on a child's probability of school enrollment. Then, a long-term analysis investigates the potential long-lasting consequences of drought on human capital accumulation, which ultimately determines individual's economic status over the life cycle.

A qualitative analysis realized in 2018 across 4 districts of the Kagera region determines what would be the parents' educational investment strategies among their progeny under budget constraints. The field work highlights that parents' investment decisions are based on their expected returns to education, which may vary with uncertainty on the child's probability of school completion or educational success. Given the possibility that a girl gets pregnant, then be expelled from school and prevented from returning in Tanzania, daughters might present a larger default risk than sons and should receive less investment when scarce resources. Similarly, parents might invest on children with the higher perceived educational performances, depending on their beliefs and almost on child's grade at the time of shock. Drought may thus induce parents to focus on the most educable children through higher levels of attainment rather than spreading some education evenly among all of their children.

I find that drought does not have any impact on boys and affects girls' education differently depending on the number and sex composition of the young sibship. The education of girls with few younger siblings of opposite sex is disproportionately cut short during a drought. Additional young brothers offset the negative effect of income instability on girls, whereas additional young sisters even increase their probability of enrollment. Reversely, natural disasters experienced during childhood decrease the educational attainment of girls with only older sisters by 1.3 years of schooling.

Then, I provide potential channels looking at the effect of droughts on child labor and child marriage. Indeed, responses to shocks to protect family consumption may consist in taking selected girls out of school to save on costs but also to send them to work in agriculture, business or as substitutes for parents in doing household chores. Another alternative may be to marry daughters in order to reduce daily costs or to get money from the wedding, as bride price payments regulate marriage in Kagera. According to this norm, the groom's family has to give a transfer to the family of the bride at the time of marriage. On the one hand, human capital increases future bride price payments[2] and let them to smooth consumption throughout

[2] Ashraf et al. (2018) demonstrate that bride price payments are positively correlated with female education in

life. On the other hand, work and marriage ensures immediate benefits but interferes with the accumulation of children's human capital and may lead to long-run costs, with lower future earnings potential. [3] Allowing parents to invest differently in their offsprings would let them respond to income shocks while maximizing expected transfers.

The remainder of the paper is structured as follows. Section 2 presents the theoretical framework. Section 3 describes the data. Section 4 examines the empirical strategy and results are discussed in Section 5. Section 6 explores child labor and marriage as potential mechanisms while section 7 provides robustness checks. Finally, Section 8 concludes.

2 Theoretical framework

Why might liquidity-constrained households invest differentially in their offsprings, especially between males and females, between older and younger family members.

Economic theories on education predict that educational investments is made until the marginal costs is equal to the discounted expected future benefits (Becker, 1992). However, individuals in low-income countries may not be well informed about the returns, schooling decisions might actually be driven by their perceived returns to education (Jensen, 2010; Nguyen, 2008). I assume that parents are making decisions on children's education, decisions which may vary across individuals (Behrman, 1997). Indeed, there might be trade-offs among household members in responses to short-run shocks. Parents will then adopt their investment decisions on the basis of children's differential effect on the resources available to the family.

Gender or cultural preferences is one explanation of differential investment between girls and boys in many developing economies (Alesina et al., 2013). Whereas son preference is a well-established idea in Asia,[4] the idea is not clear in Africa, with most Sub-Saharan African countries displaying a preference for mixed-gender or no preference at all (Rossi and Rouanet, 2015). Nevertheless, in many African traditional societies, parents rely on their sons to look after them in old age since the girl will marry and leave the natal family. In such cases, the expected returns to invest in sons might be higher and it would be efficient to invest more in a

Indonesia and Zambia, suggesting that female education is valuable in the marriage market, in theory (Anderson, 2007).

[3] A large economic literature demonstrates that parents' incentives for children's human capital are mainly related to transfers for old-age support in the absence of savings and insurance.

[4] See related papers (Li and Wu, 2011; Kishore and Spears, 2014; Edlund, 1999; Rebeca and Javier, 2016)

son rather than a daughter (Strauss and Thomas, 1995).[5] However, there is no evidence that girls are not able to assist their parents in Tanzania nowadays. Indeed, a qualitative study has been conducted in 15 villages selected on their level of development,[6] across 4 districts in Kagera in the summer 2018. In each village, I interviewed a group of women, then men and finally a mixed-gender group of 6 individuals resulting in a sample of 270 individuals. Participants were parents involving in agriculture activities, with at least one child enrolled in primary or secondary school.

The field work provides some insights on positive changes in parental perceptions about daughter's contribution as an adult, although patrilocal exogamy, which prescribes that women have to live in their husband's family, remains. The rapid adoption of mobile phones in Sub-Saharan Africa facilitates communication among social networks in response to shocks (Aker and Mbiti, 2010). The reduction in transaction, transport, and time costs associated with the mobile money service facilitates remittances between members who are geographically separated (Munyegera and Matsumoto, 2016; Jack and Suri, 2014), and might thus increase daughters' possibility to help the parents in time of need (Becker et al., 2016). In addition, the emergence of women's rights to own property or access productive resources and family gender awareness training delivered in rural areas by NGOs contribute in reducing parental perceptions that daughters are less economically valuable than sons (Allendorf, 2007; Fernández, 2014).

In times of shocks, marginal costs increase and parents keep investment in children with higher returns. Indeed, Attanasio and Kaufmann (2009) showed that credit constraints break the link between expected returns, or risk perceptions, and schooling decisions. Differential investments in offsprings can arise even with equal concern of parents towards each of their children (Strauss and Thomas, 1995). Investment decisions involve uncertainty and are therefore shaped by the expectations of return and riskiness of the project (Delavande et al., 2011).[7] Expected returns are conditional on the ultimate educational outcome, which is uncertain when parents decide to invest in an additional year of schooling. Thus, variables that affect the odds that a child will actually complete the schooling level may influence his expected return to education (Altonji, 1991). Similarly, Cameron and Shah (2015) demonstrate in rural Indonesia that

[5] Whereas the assumption that differences in the treatment of boys and girls in Indonesia are mainly due to post-marital residential practice is advocated in most of the literature, Levine and Kevane (2003) find no empirical evidence on this pattern in Indonesia.

[6] Very remote, remote, urban and very urban are the selection criteria for the villages in each district.

[7] Most papers in the literature neglect the importance of risk as a determinant of educational choice, whereas risk perceptions are important predictors for high school attendance decisions (Attanasio and Kaufmann, 2009).

natural disasters change people's beliefs and reduce their risk-taking behaviors.[8] Consequently, the child facing lower human capital investment risks will have higher expected returns and will benefit from scarce resources at the expense of the others.

The qualitative study helps me to identify who are the children facing higher completion risks. (1) Younger children may present a larger default risk than older ones because parents are less able to identify their abilities and performances. Indeed, parents' investment decisions depend on their perception about a child's academic performance. High-ability children are more likely to be enrolled and receive more educational resources (Akresh et al., 2012). During focus group discussions, through the hypothetical scenario where parents have to select which children would benefit from scarce resources in times of needs, more than 80% of groups reported the child enrolled in secondary school rather than the one in primary school (Panel A in Table 1). As a matter of fact, the one in secondary school provides less uncertainty about his performance and probability of educational success. A secondary school child has already passed some exams, is closest to the final level of education[9] and will bring money sooner to help siblings who would still lagging behind. Households are thus associating higher educational levels with higher performances at the time of shock, although parental beliefs are often inaccurate.[10]

(2) Daughters may have lower perceived educational returns than sons, as they are less likely to complete schooling due to the risk of getting pregnant. Following the hypothetical question on which one of their children, among a girl and a boy, would benefit from scarce educational resources, the unprompted answer that girls can get pregnant and be expelled from school was raised among 30% of groups (Panel C, Table 1). Indeed, the Tanzanian law does not allow women who gave birth to study anymore, making parents to loose their investment. Tanzania's ban on pregnant girls attending government primary and secondary schools dates back to 1961 and was strengthened in June 2017 with the new presidency. Even though successive governments have made a push for girls' education, those that fall pregnant are routinely expelled from school, and prevented from returning.[11] The phenomenon would be strongest in poorer families,

[8]Experiencing a natural disaster causes people to perceive that they now face a greater risk of a future disaster. In their paper, however, Cameron and Shah (2015) do not find that changes in income that is associated to the natural disaster explain the more risk-averse behavior of households.

[9]With no financial constraints, parents would like their children to go to the highest level of education, that is university. However, the average level of education increased from standard 7, the end of primary school, to form 4, the middle of secondary school in recent years.

[10]Dizon-Ross (2018) demonstrates that parents' baseline beliefs about children's academic performance are incorrect, with frictions preventing the use of available information in low-income households. When comparing two of their children, one third of parents are mistaken about which child is higher-performing in Malawi.

[11]Through the policy, the government's goal is to stop "bad behaviors" and reduce teenage pregnancies. Effec-

where long distances to school forced female students to pay transport services with their body in the lack of family resources. Thus, being a girl alters the probability of school completion and then affects the ex ante rate of educational return. Girls present a larger default risk than sons and should thus receive less education (Parish and Willis, 1993) when credit constraints are important.

In sum, when confronted by scarce resources, parents may invest in a smaller number of "higher-quality" children, who show the greatest promise of educational success or school completion, so as to maximize the benefits of the investment for the whole family. Depending on uncertainty in their beliefs, daughters may have lower perceived educational performances than sons due to the risk of getting pregnant. Likewise, younger children may present a larger default risk than older ones because parents are less able to identify their abilities and performances. At similar educational backgrounds, the theoretical framework predicts that a drought might have larger negative effects on daughters than sons. Then, the choice between investing in older daughters compared to younger sons depends on the degree of uncertainty associated to a particular grade level and gender.

3 Data and descriptive statistics

In this study, I use information from a unique 19-year panel survey on the Kagera region, in Tanzania. The region is located in the northwestern corner of Tanzania near the lake Victoria and neighbors Uganda, Rwanda, Burundi. The survey interviewed 912 households, 6,353 individuals living in 51 villages for the first time in 1991, then three times at six-monthly intervals during 1991-1994. The KHDS 2004 and 2010 attempted to re-interview all respondents from previous waves, irrespective of whether the respondent had moved out of the original village, region, or country, or was residing in a new household. Compared to other panel surveys, the attrition rate is exceptionally low. Excluding the households in which all previous household members were deceased, 92% of households were recontacted (De Weerdt et al., 2012).

tively, nurses are checking girls at school with the administration of pregnancy tests and teachers are denouncing girls under government pressure. According to the 2015–2016 Tanzania Demographic and Health Survey, one in four women aged 15–19 are mothers.

3.1 Main variables of interest

3.1.1 Education data

My primary data source is the last wave of the Kagera Health Development Survey (KHDS) realized in 2010. The availability of retrospective information allows me to construct a panel of individuals from their age of entry into primary school, in 1970 for some of respondents, to 2010.

Particularly, I exploit information on school-leaving age for those who complete their education by 2010. Then, I define a dummy variable equals to 0 during school years and to 1 when dropping out of school, after which the individual is excluded from the sample. I thus measure annual school dropouts, conditional on being enrolled in school. The hypothesis is that negative rainfall may affect schooling decisions only when a child is supposed to be in school. The dependent variable includes data from the entry age into primary school, which can vary across individuals. The largest share of children, around 23% in my sample, attend primary school from about age 7 whereas the mean of the age of school entrance is 8.7 over the study period (Table 2, Panel A). On average, people acquire more than 6 years of schooling which correspond to the primary level.[12]

Figure 1 plots the distribution of school-leaving ages in the final sample, separately for males and females. A high dropout rate, standing at an average of 20%, occurs among students aged 16. A left-skewed distribution is more pronounced in the male than in the female sample, as females tend to leave school lightly before males and few of them remain in school beyond 18 years old. The difference in the distribution by gender, though, is small. Figure 2 represents the educational attainment of individuals who complete their schooling, by sex.[13] One striking fact is the high proportion of children, around 56% of boys and 60% of girls, who finish only 7 years of primary school. Few of them, around 10%, complete lower secondary education and only 4% of males continue their studies in upper secondary school. The transition to secondary school represents the point at which gender gap is the largest, with less than 1% of girls pursuing further study after standard 4.

One of the main limits for using annual school dropouts as the outcome is the exclusion

[12]Primary school in Tanzania covers grades 1-7 and secondary school from standard 1 to 6. Most children are unable to continue schooling beyond primary school, with the possiblity that a child can be legally employed at 15 years old (Alam, 2015).

[13]The figure is based on 1,149 individuals with information on school-leaving age. For 764 individuals with missing information, I assume that education is complete beyond their 18 years old.

of individuals who never went to school from the analysis. This concerns 11% of the 1,806 individuals with information on complete education by 2010 in my primary sample, including 39% of males and 61% of females. Among reasons for not attending any school, 34% reported the lack of financial resources, 24% did not have access to school and 15% had to work in agriculture or in the household. Also, the constructed measure does not take into account temporary interruptions in children's schooling during a year. Indeed, income variations may affect children's attendance in schools without stopping their education definitely.[14] However, repeated interruptions within the academic year might affect student's performances. In this way, I extend the analysis using the number of completed years of schooling in section 5.3, including individuals with no formal schooling who present specific characteristics.

3.1.2 Sibling's characteristics

To construct variables on sibling composition, I use a 2010 questionnaire that provides a list of eligible children of the 1991 household head, belonging to the same couple. Eligible children are those where the head was living with one of his biological child in 1991 and such child was under 13 years old in the baseline survey. The criteria reduce the sample to 456 families, with 2,558 children.[15] The questions concentrate on their current situation or if they are deceased, on situation before their death. Thereby, data allows me to know exactly family structures, including siblings who migrated or died. First, I measure a size effect, focusing on the number of younger and older siblings separately. I disaggregate the analysis by sibling subgroups to allow differences in the effect of having older or younger siblings during financial difficulties. The variables range from 0 to 4, with a value of 4 meaning that the individual has at least 4 younger or older siblings, depending on the selected variable. Secondly, I focus on indicators reflecting gender composition within sibling subgroups. I determine sex ratios as the number of younger (older) sisters divided by the total number of younger (older) siblings. Then, I generate a categorical variable equals to 0 for having only brothers, 1 for at least one sister and 2 for only sisters. Thus, I assess the effect from having a higher proportion of sisters among the young

[14] During focus group discussion, parents raised the possibility that children cannot go to school during few days as families do not have money to buy shoes, uniforms or meals that are required by school.

[15] Families with no child living in the household in 1991, or monoparental families and child headed households are not considered. This exclusion may lead to a downward bias in my estimations as such families may provide specific characteristics, such as HIV prevalence and adult mortality rates, which may increase their vulnerability to exogenous income shocks. In the robustness section, I will include them and reproduce the analysis on the first four waves of the KHDS database. Indeed, I will exploit a specific questionnaire on children residing elsewhere, available between 1991 and 1994, to complete information from the main survey and construct variables reflecting siblings' characteristics.

and old relatives. As reported in Panel A in Table 2, individuals have more than 5 siblings on average, with 2.8 sisters and 2.6 brothers. Panel B indicates that around 40% of observations have imbalanced sex ratios among younger and older siblings.

Exploiting data on school-leaving age and sibling structures, my final sample consists of 1,148 individuals, with 570 males and 578 females. The temporal dimension varies by individuals, spanning from the beginning to the end of their schooling, encompassing years between 1970 and 2010. Finally, the panel includes 11,074 observations.

3.2 Weather shocks and income variation

Drought, used as a reduced form instrument for household income variability at the district level, have been collected from the Global SPEI database. The SPEI index is based not only on precipitation but also on evapotranspiration data. Temperature have been shown to better explain spatial and temporal variation in agricultural income (Dell et al., 2012; Lobell et al., 2011). The drought index is expressed in units of standard deviations from the long-run average over the 12 months times scale, between 1950 and 2010, to estimate a probability density function from a log-logistic distribution. A negative value in a given year indicates unusual low water balance relative to what is normally experienced in a particular location. A particular advantage of the SPEI is that the values are split into extremely, severely or moderately dry, and near neutral conditions. I generate a dummy equals to 1 when the index is below -1, indicating a drought of any intensity.[16] The dummy is then linked to the 51 baseline villages of the Kagera survey through GPS coordinates for school years, that is from 1970 to 2010. As the SPEI index has a spatial resolution of 0.5-degrees, the location of villages selected in 6 districts[17] of the KHDS database results into 10 different grids with different coordinates. The main sample matches up to 10×40(years) unique grid cells across 51 villages in the Kagera region, each one being approximately 2,500 square kilometers in area. The panel nature of the data allows me to track where respondents lived as a child rather than their current location, which leads me to easily identify rainfall shocks experienced during school years. In my sample, 28% of observations are affected by a shock of any intensity (Table 2). In Figure 3, I plot the density of villages exposed to rainfall shocks between 1970 and 2010. The bars represent droughts occurring with any intensity. The figure shows that shocks are independent across

[16] In the robustness part, I restrict the analysis using only severe and extreme droughts, from values below -1.5 of the SPEI index.

[17] Biharamulo, Bukoba Rural, Bukoba Urban, Karagwe, Muleba and Ngara are the baseline districts.

time and location, and do not appear to be auto-correlated.

As pointed out by Corno et al. (2017), the relationship between rainfall shocks and agricultural output is well established in the literature. Let me now investigate how the measure of droughts may be considered as a good instrument for rural income volatility in Kagera. First, the sample under study is overwhelmingly rural with agriculture related as the main economic activity in more than 96% of communities at the baseline. Individuals are primarily engaged in producing bananas and coffee in the north and rain-fed annual crops (maize, sorghum, cotton) in the south (De Weerdt, 2010). According to the community questionnaires of 1991 and 2004, more than 90% of villages experienced at least one disaster respectively since 1985 and 2004, due firstly to droughts, then epidemic and flood. Moreover, De Weerdt and Hirvonen (2016) relate poor harvest due to adverse weather as the second largest negative shock experienced by panel respondents since 1994. Finally, among 270 individuals who participated in focus group discussions in 2018, each respondent confirmed the growing issue of climate change. In particular, droughts is referred in each discussion as the most significant negative phenomenon on harvest.

To validate the latter claim, I test how the measure of rainfall shocks affects crop yields and thereby rural income in Kagera. To do so, I combine droughts with food and consumption data in the six waves of the KHDS. Following Corno and Voena (2016), I examine the impact of droughts on the natural logarithm of consumption and food consumption per capita in the last 12 months.[18] Conducting the analysis at the household level, all estimates include survey and household fixed effects, so the impact of drought on household consumption is identified between-households and between-survey variation. Column 1 in Table 3 shows a consumption growth gap of 30 percentage points in households exposed to climate shocks. Similarly, experiencing a drought is associated with a decrease in food consumption per capita by 16% on average (column 2). Finally, I test the effect of droughts on the natural logarithm of individual expenditures (including expenses on clothes, jewelry, toys, books, restaurants, tobacco, newspapers, cosmetics etc), adding survey and individual fixed effects.[19] As reported in column 3, drought negatively and strongly affects expenditures. In particular, drought leads to a 0.22 unit decrease in individual expenditures on different items. Overall, these findings suggest that the

[18] Consumption per capita includes household expenditures on education, health and durables goods like home or vehicle repairs, and food consumption per capita. Food consumption per capita is the household food consumption during the rainy and dry seasons divided by the number of members living in the family.

[19] Information on individual expenditures are only available in KHDS 1991-1994 and KHDS 2010.

constructed measure of drought may be a relevant reduced-form instrument for negative income shocks in the region under analysis.

4 Identification Strategy

I first investigate the relationship between income shocks, sibling composition and educational outcomes using a simple approximation of a duration model as in Corno et al. (2017). More specifically, I examine the impact of income shocks on the hazard into school drop-out for girls and boys with different sibling composition, using the following reduced-form equation:

$$\begin{aligned}\text{Dropout}_{i,t} = \alpha_0 + \alpha_1\text{Drought}_{i,t} + \alpha_2\text{Siblings}_{i,t} + \alpha_3\text{Sex Composition}_{i,t} + \alpha_4\text{Drought}_{i,t} \\ \times \text{Siblings}_{i,t} + \alpha_5\text{Drought}_{i,t} \times \text{Sex Composition}_{i,t} + \zeta_i + \lambda_t + \eta_{i,t}\end{aligned} \tag{1}$$

where α_0, - , α_5, are vectors of parameters to estimate and ζ_i, λ_t, $\eta_{i,t}$ are error terms.

Since effects of income shocks likely vary by child's gender, regressions are estimated separately for girls and boys. *Dropout* i, t measures the probability of drop-out of school for individual i in time t. The dependent variable is a binary variable coded as 1 in the year the individual leaves school, and zero otherwise. Thus, the duration of interest is the time between the beginning and the end of child's schooling. Specifically, $Drought_{i,t}$ is a dummy indicator for a drought in a given year. To easily interpret the findings, variables reflecting siblings' characteristics are first measured per subgroup, among the young and the old sibship. Then, I include the old and young sibling variables in the same regression to estimate which subgroup has the largest effect. $Sibling_{i,t}$ represents the number of (younger and older) siblings, tallied as the number of biological brothers and sisters. In addition, Sex Composition$_{i,t}$ corresponds to sibling sex ratios, measured through a categorical variable. In equation 1, α_1 is the effect of drought for a child who has one younger (older) brother, the excluded category given simultaneous interactions with sibling gender and number. α_4 and α_5 capture how the shock differs for an individual with more younger (older) sisters than brothers, and at least 2 younger (older) siblings. In doing so, the focus is on the conditional expectation of having younger and older brothers or sisters when the family is affected by income shocks. Variations in estimations stem first from exposures to climatic events which vary across districts, then from the young and old sibling structures which differ across individuals in the family. Estimations are thus conducted between individuals, affected or not by drought, with a particular sibling composition.

A potential threat to my identification strategy comes from the fact that the number and sex composition of siblings is not exogenous, as it depends on the realization of parents' fertility decisions (Morduch, 2000). According to the old-age security hypothesis, fertility choices are influenced by the child role of investment-good or household asset. In developing countries, parents generally choose to invest in their future in the form of children (Banerjee and Duflo, 2011). Similar concerns emerge from the post-birth sex ratio if the latter varies due to endogenous mechanisms like migration or mortality. Sex ratios are constructed from all biological children, currently alive or deaceased, ever born to a mother, irrespective of their migration experience. Besides, the probability of school dropout may be influenced by individual characteristics, such as abilities or skills. In this line, I include individual fixed effects, ζ_i, which control for any omitted family-individual-specific unobservables that do not vary over time.[20] Following Jayachandran and Pande (2017), I then consider a different sample in robustness checks, where mothers have likely completed fertility, assuming that women above 49 years old may have fewer possibilities to give birth again.[21] Endogeneity concerns of family size and sex composition are thus reduced, assuming that parental preferences do not evolve across time. Besides, I include year fixed effects, λ_t, to control for shocks, such as new births, occurring in different time periods. Thus, differential impacts of drought on school dropout conditional on sibling structures are identified within individuals' school years over time, with a linear probability model. Despite the fact the outcome is not considered as a discrete variable distributed on two possible values, the linear probability model is consistent under the usual assumption that the error term and regressors are uncorrelated with each other. Conversely, the probit and logit models require arbitraty assumptions about the error term distribution and are less suited to the application of panel data methods. Finally, because I use data on multiple siblings within families, standard errors are clustered at the family level to correct for arbitrary within-family correlation.[22]

Another concern may arise if the location of the individual at the baseline survey, in 1991-

[20] Also, I test the results using family instead of individual fixed effects, which lead me to analyze between-sibling variations in drought and in subgroup size and gender composition.

[21] As most surveys, the KHDS fertility module is administered to women aged between 15-49 years old. Women above 49 are thus not considered as of child-bearing potential.

[22] Clustering standard errors at the village level does not affect the results.

1994, is different than the location during school years. Using specific questions on past migration in 2010, I am able to know if the individual grew up in the baseline village or nearby. Among 1,821 individuals who complete their schooling by 2010, 8% moved out of the original village in their teenage years. Similarly, child fostering is common in Africa, in which parents send their own biological children to live with another family, for child work and/or child schooling. Among individuals with information on their educational attainment, 6% of them migrated for their schooling so as to go in a boarding school or leave in another household, including 61% of males and 39% of females. Serra (2009) demonstrates that being fostered to a better-off household may improve the human capital of a child, even if this child works no less than what he would have done back home. Thereby, including fostered children may decrease the magnitude of my estimates, as these children may be less exposed to parents' income variations. When fostering a child, families may also be less exposed to adverse shocks, relying on this reallocation of resources as a form of insurance (Akresh, 2004). Nevertheless, I exclude school-age migrants from the analysis in robustness checks and find similar results.

Then, the number of younger and older siblings is correlated with age and birth order, that may have an independent effect on schooling dropout. Indeed, having a large number of older (younger) siblings reflects a high (low) birth order. To ensure the results are not driven by an age and birth order effect, I allow direct interactions between age and drought variables. The results are robust to controlling for age and birth order, providing evidence that droughts differently affect children due to a size and gender effect in sibling subgroups.

Besides, reduced form equations may be problematic if rainfall shocks affect education through channels other than income, as highlighted in the paper of Björkman-Nyqvist (2013). It is possible that rainfall shocks could have other indirect effects on education, notably through health. Children affected by any disease are less likely to attend school, which increases their probability of school dropout. Nevertheless, the possibility of contracting malaria may be equally distributed among girls or boys within the same family. Moreover, Shah and Steinberg (2017) empirically test the relationship between drought and malaria, finding no association between drought and a weaker or higher likelihood of malaria infection. Further evidence is provided in Sub-section 6.

Finally, as I am using retrospective information depending on reported age at school dropout, measurement error would be possible, leading to imprecision in estimates. In the same way, the constructed sibling structure depends on birth dates given during interviews. However, using self-reports age estimates is common in economic studies, and such measures are rather accurate in DHS studies for instance (Pullum and Staveteig, 2017).

5 Results

My main results examine the differential effect of income instability, measured with drought, on children's school dropout conditional on sibling composition. In a first stage, the analysis is conducted on girls and specific results for boys are presented in Sub-section 5.2. Lastly, sub-section 5.3 analyzes the heterogeneous long-term impacts of climate change on girls and boys' human capital.

5.1 Girls' educational attainment

Table 4 provides estimates on girls' probability of school dropout differences, according to sibling composition. All regressions include individual and year fixed effects, so that the results are robust to any household, individual or time-invariant characteristics which may influence educational attainment. The focus is on the role of being affected by a negative income shock when having younger siblings, as the effect of droughts does not vary with the old brotherhood. I estimate heterogeneous effect of income shocks within a girl's academic life across time. Experiencing at least one drought during school years concerns 483 among 584 girls in the sample. Similarly, coefficients on sibship-related variables identify only girls who have new brothers or sisters during their years of schooling. Due to limited cases, I only interpret coefficients on such sibling measures once interacted with drought.

In the first column, I examine how drought varies with the gender of younger siblings. I interact sibling sex composition with drought, controlling for their number. If those girls who experienced an agricultural shock and had more younger sisters than brothers are more likely to be enrolled in school in the same year, the interaction should be negative. It is significant at a 5% size effect. In contrast, the raise in the likelihood of school dropout following a drought is greater for girls who have only younger brothers (main effect in the first line). In column 2, drought does not have differential effects according to the young fraternity size. As the gender composition of children may influence fertility behavior,[23] I simultaneously interact sibling size

[23]Jayachandran and Pande (2017) demonstrate that son preference in India, especially for the eldest son, influences parents' fertility decisions. Indeed, when a daughter is born into a family with only girls, her parents are likely to continue having children in their quest for a son, exceeding their originally desired family size. A less obvious association between family size and sex composition is found in Tanzania. Regarding declared

and sex composition in column 3, representing my best specification presented in equation (1). The main effect measures now drought impact for girls with one younger brother. It is positive and highly significant. The interaction of drought and having only female young siblings is -0.08, significant at the 1 percent level. Similarly, the interaction of drought and having at most 3 young siblings becomes negative and significant, indicating that additional siblings may offset the negative impact of income instability.

In order to easily interpret the findings, Figure 4 represents the marginal effects of climate change on school dropout along the distribution of younger siblings by sex composition. That is whether the differences in drought levels are significant if the values of sibship variables are changed.[24] Climate change has a U-effect along the size of the young sibship, irrespective of the gender composition. After an initial increase, drought impact diminishes (i.e. gets closer to zero) as the number of young siblings increases up to a certain threshold. Then, the pattern of shock is different for girls with only sisters than those with brothers.[25] Indeed, for girls with only younger brothers, the difference in effects is first positive, then declines. The average marginal effect of drought is 7 percentage points for girls with one brother, significant at the 1% level and becomes insignificant with additional brothers. Having additional younger brothers may be positively correlated with one's higher educational grade. In times of shock, parents may thus be indifferent between educating the daughter with higher educational levels and the sons enrolled in lower levels. Conversely, girls with one younger sister are not affected and drought increases girls' chances of going to school with additional sisters. Everything else equal, I would expect a 8 percentage point decrease in the probability of girls' school dropout if I change the sibling composition from male to female, setting the size of the young brotherhood to the average (around 2 younger members).

To ensure that sibling size and gender effects do not proxy for birth order, I include the birth order of individuals (up to 5-and-higher birth order) in the controls. Child age is also correlated with the number of younger siblings as the child with a higher number of younger siblings will, by definition, be older. I therefore use age as covariates. Importantly, for each of these covariates, I include the interaction with the drought dummy. Column 4 shows that the addition of these

preferences, the focus group discussion analysis reports that most parents have no ideal gender composition, or prefer to have mixed-sex children, reflecting positive changes in attitudes arising within the last 10 years.

[24]The marginal results are based on regression in column 3 of Table 4, where both demographic variables are interacted with the shock measure.

[25]Similar patterns emerge from having only younger brothers and having at least one younger sister. The results are sensitive to extreme cases in gender composition of the young sibling group.

control variables does not reduce the magnitude and significance of Drought×Sex composition (Young) and Drought×Number of siblings (Young) coefficients.

Column (5) adds interactions of drought with the old sibling variables to take into account the whole sibling structure. The coefficient on drought in the first line is greater than with previous estimates. A drought raises the probability of leaving school by more than 10% for girls with one younger and one older brother. Once again, coefficients on interactions between drought and young sibship variables are significant, of similar magnitudes. Also, drought seems to have a negligeable effect on girls with 3 older siblings, but the coefficient is only significant at the 10% level and is insignificant with 4 older relatives.

Finally, columns 6 reproduces the estimation of equation (1), including family and cohort fixed effects. The latter specification investigates the effect of drought between siblings, who are exposed to identical family shocks and have not, at that time, a similar number of younger and older brothers or sisters. The estimation requiring at least 2 girls in the family, the sample size is reduced to 425 girls (out of 578). The effective sample size is even smaller as sex composition coefficients are identified only for girls with younger siblings. The interactions remain statistically significant, with similar magnitudes.

In summary, girls are not equally affected by family income loss according to the size and gender composition of the young sibship. Drought takes its greatest toll on the schooling of girls with few younger opposite-sex siblings. While additional younger brothers, which reflects higher educational grade, tend to offset the negative shock on girls' education, additional young sisters may even increase girls' probability of being enrolled in times of hardship. Parents sought to protect investments in the schooling of older girls in the absence of young males, at the expense of younger daughters, due to perceived educational retuns at time of shocks. Whereas younger sisters generally cause earlier school-leaving in developing economies (Vogl, 2013; Parish and Willis, 1993), they benefit older daughters during income fluctuations and incertitude.

5.2 Boys' probability of school dropout

Table 5 reports my second set of results: the impact of adverse rainfall shocks on the probability of school dropout for boys with different sibling structures. The first two columns investigate separately the effect of droughts according to gender and then the size of the young brotherhood. In column 1, the main effect of drought is non-significant, suggesting that boys may be not affected by income shocks. The interaction of drought and having a mixed-gender composition among young siblings is 0.03, significant at the 10% level. However, the coefficient when summing the main effect and the interaction term is not significant. Examining how drought impacts vary for boys with the gender and number of younger siblings, I find no significant effect in the next columns. Column 4 shows only that drought has a differential effect on boys according to birth order, suggesting that shocks will primarily affect older boys. The magnitude, though, is really small.

In sum, regardless of their educational grade level or performances, sons are broadly not affected by family income drop. When comparing boys and girls, the results show that perceived educational returns are greater for boys than for girls. More precisely, younger boys may have higher expected returns of schooling than older daughters who may be enrolled in a more advanced level at the time of shocks. The analysis so far suggests that parents have higher incertitude on girls due to fear of their possible pregnancy than on the younger children with uncertain educational capacity.

5.3 Long-run heterogeneous effects of droughts

In this part, I complement the analysis investigating the long-term differential effect of drought on children's human capital. Thus, the shock variable is converted to a cumulative measure indicating the number of droughts experienced during childhood over different age ranges, from 6 to 20 years old, to capture drought impacts on children enrolled in different educational grades. I extend the results to extreme and severe droughts to compare drought intensity. Rather than estimating the probability of school enrollment, the dependent variable reflects child's educational attainment with the number of complete school years of each respondent. The estimation includes individuals who never went to school who were excluded with the annual school dropouts outcome. The nature of data leads me to run a cross-sectional regression, exploiting past information up to 2010. I investigate the heterogeneous effect of cumulative droughts experienced during childhood on the human capital accumulation of a child i, living in family f, born in year y, such as the following:

$$
\begin{aligned}
\text{Education}_{i,f,y} = \iota_0 + \sum_a \iota_1 \text{Droughts at age a}_{i,f,y} + \iota_2 \text{Siblings}_{i,f,y} + \iota_3 \text{Sex Ratio}_{i,f,y} + \\
\sum_a \iota_4 \text{Droughts at age a}_{i,f,y} \times \text{Siblings}_{i,f,y} + \sum_a \iota_5 \text{Droughts at age a}_{i,f,y} \\
\times \text{Sex Ratio}_{i,f,y} + \kappa_f + \nu_y + \mu_{i,f,y}
\end{aligned} \tag{2}
$$

where ι_0, - , ι_5, are vectors of parameters to estimate and κ_f, ν_y, $\mu_{i,f,y}$ are error terms.

Age a denotes three schooling age groups: the 6-10, 11-15 and 16-20 years old group. Regressions are run separately for boys and girls in order to study gender specific effects. The specification includes family fixed effects (κ_f) and year of birth fixed effects (ν_y), to capture time-invariant family characteristics (i.e., preferences in fertility) and time-invariant cohort characteristics (e.g., education reforms and available services in some particular year) that may be related to child's attainment. With the use of family fixed effects, variations in rainfall stem from the fact that siblings are differently exposed to shocks within the household. Indeed, following Alam (2015), the hypothesis is that negative rainfall may affect schooling decisions only when a child is eligible to be in school at the time of shocks. If a child is inelegible for school, parental income variability simply cannot have an impact on their schooling. Also, I consider different measures of sibling composition per subgroup, including the number of younger (older) relatives and their sex ratio in the form of continuous variables instead of categorical variables, the latter assuming linearity across the values. Thus, Sex ratio$_{i,f,y}$ among the young (old) fraternity is ranked between 0 and 1, from a male to a female composition. Correspondingly, Siblings$_{i,f,y}$ describes the absolute number of younger (older) relatives. These variables reflect the resultant final sibling composition of 2010. To avoid an upward bias effect, I exclude 18 families among 472 households in the sample, where the number of children was extremely high, with more than 10 children per family.[26] Standard errors are clustered at the family level.

Table 6 presents the results for girls and boys separately and limits the analysis to severe droughts in the last 4 columns. The number of young relatives has a strong significant effect on how girls will be affected by a drought arising during their 16-20 years old (column 1). The shock will lead them to pursue higher education courses if there is a large number of young siblings in the family. An additional drought improves the educational attainment of girls aged 16-20 years old by one year at a 5% size test, to the disadvantage of younger cohorts (column

[26] Including these families in the estimation does not affect the significance of the results.

2). The negative coefficient on the interaction between drought arising during the first primary school years and old sibling sex ratio shows that young girls are most directly suffering from economic shocks in the presence of older sisters. The elicited effect is clearly identified in column 5, where the estimation is confined to extreme climatic events. Having more sisters than brothers enrolled in higher education increases the vulnerability of young girls to adverse shocks. Moving from an all-brother to all-sisters scenario in families exposed to droughts, will cause the shock to decrease completed years of schooling by more than 1.3 years[27], at a 5% level of significance, with an average number of older relatives. Parents are thus investing in the older rather than younger daughters as soon as they are affected by economic difficulties. Uncertainty in regards to a child's school completion is higher for the younger daughters who may combine no guarantee of future performance and risk of future early pregnancy.

Finally, the analysis on boys reveals that droughts taking place in their 6-10 years old will negatively affect those with a large number of older siblings. However, the results have to be interpreted with caution as the marginal effects of droughts are negative and significant only from 5 older siblings, when the sample size is shrinking drastically. Also, a downard effect on boys' educational attainment emerges with extreme droughts occurring at the end of primary school, between 11-15 years old in column 7. Holding the size of the young brotherhood to the mean, having an additional young sisters compared to an additional brother will decrease the human capital of boys around 0.98 years (0.40-1.39= -0.98). However, the coefficient is significant at the 10% level, thereby no much attention is devoted to the interpretation of this effect.

6 Potential Transmission Mechanisms

Girls are used as a riskcoping instrument when households have difficulties in sustaining consumption. Indeed, the results highlight that daughters with more young relatives will be favoured in same-sex families, at the expense of the younger ones during income shocks. Boys are not affected by drought, regardless of the size and gender of the brotherhood. Nevertheless, little evidence has been provided on the mechanisms involved. How do girls, especially the younger ones, help the household to absorb the shock?

The first answer following the analysis would be to save education costs. Indeed, education

[27] The coefficient associated to a male sex ratio and a size of the old sibship equals to the sample mean is 0.72. The marginal effect of severe drought is thus 0.72-2.04=-1.31

in Tanzania is not accessible to all children, since it has to be paid for. It is only since 2016 that education again becomes free in primary and secondary schools up to form 4. Excluding school fees, many expenditures still remain to be paid like uniforms, exercise books, lunch expenses, tutoring. Younger girls will be withdrawed from school to reduce household expenditures as they are the children with lower perceived educational returns at the time of shocks. On the one hand, girls are riskier investment due to fear of pregnancy and school expulsion associated; on the other hand, younger children are riskier investment because they have not yet demonstrated their abilities and are further away from the "final level" when the shock occurs. Saving on education costs may however not be sufficient to overcome shocks, in which case other alternatives increasing liquidity may have to be engaged.

Among strategies implemented to cope with shocks, reducing food consumption was one of the most frequent answer provided during qualitative field work. People reduces food quality as well as food consumption, with one meal a day instead of three. The consequence is that children are going to school on an empty stomach, tired and weakened, which may raise the probability of school dropout due to low performances.[28] Indeed, malnutrition is related to low cognitive performance later in life (Glewwe and King, 2001). At first sight, food may be equally divided among children within families and poor nutrition would produce similar effects on educational attainment across siblings. However, Behrman (1986) demonstrates in India that when food is scarcest, parents follow their investment strategy, favoring older children and exposing the more vulnerable children to greater malnutrition. Quisumbing (2003) finds similar patterns in Ethiopia where food aid programs redress imbalances among children in their nutritional status. Following these studies, I test the effect of environmental shocks on the number of days where the child was sick during the year, in logs, as the outcome. Sibling variables are allowed to interact with drought, testing whether drought effects are the same across subpopulations regarding nutritional status. Neither the main drought effect nor interactions are significant for boys and girls. Thus, it is unlikely this channel is driving the contemporaneous school dropout results.

Then, one means by which a household may cope with shocks is to spatially diversify its sources of income, locating its members in areas characterized by low covariances in income (Rosenzweig and Stark, 1989; Halliday, 2006; Beegle et al., 2011). Among men, labor-related

[28]Tanzanian students have to pass examinations at the end of primary school and then at each grades of secondary.

movements and migration out of the district more than doubled under severe drought in Ethiopia (Gray and Mueller, 2012). However, due to severe liquidity constraints and low necessary resources, poor households may be less able to increase their mobility (Cattaneoa and Peri, 2016; Gray and Mueller, 2012).

In this part, I quantitatively investigate through KHDS data two potential alternatives affecting the probability of school dropout for selected children: labor and marriage, which encompass the migration channel. Following the analysis above, I focus on girls who are likely to be affected by a negative income shock.

6.1 Child labor

Do younger girls help other relatives to share burdens through labor? Selected children may work on the labor market to find extra income or work as subtitutes for parents in doing household chores in times of need. Beegle et al. (2006) demonstrate that households increase their use of child labor in response to crop shocks in Kagera and in Tanzania (Bandara et al., 2015). A concern arises if drought equally affects families within the village and reduces child labor opportunities, which may invalidate selected daughters' employment as a way to increase family liquidity. Indeed, Shah and Steinberg (2017) demonstrate that bad rainfall years result in relative paucity of outside options and low wages, which diminish the use of child labor and increase school attendance. The qualitative work, though, pinpoints various income levels within villages due to different agricultural size plots, storage quantities or livestock, and small business outside agriculture. Differences in income and assets across families allow specific children to be hired by wealthier households to work principally in agriculture or as housemaid.

I use information on the number of weekly hours performed by an individual in agriculture and in household work from KHDS 1991-2004. Agricultural work includes any activities on plots, collective land and crops, caring for animals or transforming animal products. Household work gathers activities such as cleaning the house, preparing meals, collecting firewood and water, shopping for food, seeking medical care or caring for ill household members. Following Beegle et al. (2008), I restrict the analysis to individuals aged between 7 and 15 years old, a period in which child labor and education are simultaneous decisions.[29] Labor variables are merged with data from the 2010 sibling module, the latter is restricted to biological children of selected

[29] I also test the results on individuals aged between 7 and 18 years old, as higher education may encompass the 16-18 age group as well. The results are the same.

couples. To get the family structure for a wider sample of individuals, I exploit information on women's births, aged between 15 and 49 years old, available only in the first four waves of the KHDS between 1991 and 1994. Climate change is still a dummy for adverse drought conditions experienced during the year of interview which is interacted with sibling measures, per subgroup. Individual and survey fixed effects lead me to investigate the consequences of droughts depending on sibling size and sex composition within individuals across surveys. Focusing the analysis on children aged between 7 and 15 years old, the fifth wave of the KHDS database in 2004 does not provide additional information. Thus, estimations are performed between 1991 and 1994.

Results in the first three columns of Table 7 are restricted to female agricultural work as no significant effect are identified on female household chores and on boys' labor in general. Negative rainfall increases by 34 percent the number of worked hours in agriculture for girls with a single younger brother (coefficient on drought in column 1) and becomes non significant for girls with at least 2 younger brothers. Furthermore, the interaction on drought and having only younger sisters is negative at the 10% level. Drought does not have any impact on girls with younger sisters, irrespective of sibling size (sum of main effect and interaction terms). In column 2, I focus on the heterogenous effect of drought according to the old fraternity characteristics. The interaction of drought and having 2 older siblings is negative at the 1% level, but the marginal effect is non-significant. Finally, estimating the impact of climate change conditional on demographic measures of both the young and old sibling group, the main effect measures drought impact on the girl in the middle of a family with only brothers (column 3). The effect is positive and significant at the 1% level. Also, negative and significant terms emerge from drought interacted with having only younger sisters and 2 older siblings. First, the marginal effects suggest first that drought reduces the number of worked hours for girls with a high number of older siblings (at least 4), fixing other sibling variables to their mean. Secondly, drought reduces the number of worked hours in agriculture by 75% for girls with only younger sisters compared to those with only younger brothers, holding other siblings variables to their average.

In this way, the effect of climate change on girls' child labour likely varies by the size of the old brotherhood, but mostly by the young sibling sex composition. The findings conform with previous elicited effects: younger sisters help to offset the negative effect of drought on female labor in agriculture. However, in large families, drought will reduce also worked hours for girls with a lot of older siblings, suggesting that the youngest daughter may not be negatively impacted by drought through labor.

6.2 Child marriage

Finally, do girls are going to leave the household through marriage to alleviate family expenses? Indeed, parents may have incentives to marry their daughter to get bride price payments as a source of consumption smoothing. Corno et al. (2017) found that when aggregate income is temporarily low, marginal utility of consumption is higher, and households prefer to marry their daughters earlier in order to consume the marriage transfer. However, drought may reduce marriage payments due to lower income of the grooms' family in nearby areas. Gray and Mueller (2012) showed that moderate drought lower marriage-related mobility, reflecting a decreased ability to finance wedding expenses and new household formation. Moreover, the qualitative study has shown that the bride price is usually shared within the community, between all relatives. In most cases, bride price is not intended to a specific use and may not help parents to pay for education expenditures for instance. Nevertheless, daughters can leave the household through marriage to partially alleviate the liquidity shortages weighing on their families at the moment of shock. Besides, there is no cultural tradition governing the timing of marriage between daughters, allowing the younger girls to be married before the older ones, which would potentially explain the educational results.

Data on marriage were collected firstly on household members who were less than 17 years old at baseline in 1991-94 and have married at least once by 2010, and then merged with sibling data. The marriage age is then completed through a question on the length of marriage from the 2010 sibling module. As the length variable is presented in multiple range scales, I use the median of each interval to generate the age at first marriage for missing observations.[30] Similarly to the analysis on education, I convert the data into person-year panel format from their year of birth to 2010. I test the impact of drought in the current year on the timing of marriage, conditional on sibling structures. Thus, the outcome is a binary variable coded as 1 in the year the woman gets married, and zero otherwise. Considering that a 15 years-old girl may be at risk of early marriage,[31] the duration of interest is the time between her 15 and the

[30]The different intervals are: less than five years, between 5 and 9 years, between 10 and 14, between 15 and 20, more than 20.

[31]Although many African countries have established 18 as the minimum age of marriage for both boys and girls, weak enforcement has meant these laws have had little impact. Moreover, there are no rules about the timing of marriage between sisters, which suggests that the youngest can be forced to marry early before the

age when the girl marries, after which she exits the data. Since I am interested in parents' alternatives to higher education investment, I consider early marriage as a union where one member got married at 20 or younger. This concerns 53% of individuals among 645 women in the sample. Thus, women married after age 20 are right censored. The econometric strategy is similar to the one exposed in Sub-section 4, relying on interactions between drought and sibling composition measures among the young and old sibship.

Columns 4, 5 and 6 of Table 7 present the results based on the within estimator, using individual and year fixed effects to analyse variations in climate conditions within individuals. In all specifications, standard errors are clustered at the family level. The effect of drought on the likelihood of early marriage subtly differs along the size of the young brotherhood (column 4). Indeed, younger siblings will mitigate the negative effect of climate change on child marriage, as coefficients associated to its interaction with drought conditions are negative and significant almost at the 5% level. Estimating marginal effects, a drought increases by 3% the likelihood of early marriage for girls with a single younger brother or sister. I find no effects of rainfall shocks on the timing of marriage for girls with additional younger siblings. However, when I include interactions with measures reflecting the old sibling subgroup, it is the number of older relatives which seems to be the most important factor altering the impact of negative rainfall (column 5). The number of old siblings interacted with weather shocks displays positive and significant coefficients at the 5% level of significance, suggesting that additional older siblings magnify the negative effect of drought on girls' marriage. The overall significance test tells that drought increases the annual hazard into child marriage for girls with few younger siblings and a high number of older relatives by 3% (column 5, $p < 0.05$).

In large families, the marriage of the youngest daughters is thus a form of self-insurance following income losses. Whereas they are working in opposite direcion, both the young and old sibship size determine how drought will affect girls' probability of early marriage. However, the marriage results do not identify any gender effect coming from sibling composition pinpointed in the education analysis.

7 Robustness part

Finally, I have to perform additional estimations to test the robustness of my results.

oldest daughters.

It is reasonable to suppose that drought likely affects children who have school age siblings and has little impacts on children with very young household members or adults out of school. In that purpose, the regression in Table 8 focuses on siblings aged from 6 to 19 years old,[32] including the total number of school-age siblings and the number of sisters (excluding the male group) to separate the effect of size from sex composition. This additional specification allows me to include children who do not have any younger or older siblings. Those individuals were excluded from the main analysis due to the inclusion of sex ratio measures, which are constructed conditional on having younger or older siblings. The rise in the probability of school dropout when affected by a shock is greater for girls who do not have any school-age sibling (first line in column 1). Additional school-age sisters mitigates the negative effect of drought as the coefficient on the interaction with the number of sisters is negative and significant at the 5% level. When I disaggregate the analysis by sibling subgroups in column 2, 3 and 4, the results demonstrate that what matters is the presence of school-age younger sisters. The number of young relatives slighlty worsens the impact of drought but the effect is only significant at the 10% level. Whereas these specifications analyze in more details the heterogeneous effect of drought on individuals, they reduce the sample size which naturally decreases the level of precision in the coefficient estimates. Drought does not have any effect on girls between 6 and 9 years old (column 5). However, the absence of significance can be explained by a lack of statistical power as school dropout affects only 3% of girls in this specific age range. During the latest stages of primary school, between 10 and 14 years old, drought will decrease the likelihood of being out of school for girls with additional younger sisters. Similarly, it is the number of older sisters which may be a strong predictor of the effect of climate change. Indeed, during secondary school, between 15 and 19 years old, parents will shelter older sisters who may be enrolled in higher grades. Income losses will increase around 17% the probability of stopping school at the secondary level for a girl with an additional older sister enrolled in upper-secondary school.

Then, Table 9 does not allow sibling measures to vary over time, fixing the final family structure to the one in 2010. The results are similar. Besides, the last four columns reduce the analysis to families with complete fertility, that is families where the mother is more than 49 years old in 2010. Estimations corroborate previous findings; drought impact differs mainly by young sibling sex ratios and slightly by the young sibsize.

[32] The specific age range is based on primary school-entrance age and on school-leaving age from the 2010 survey, with more than 91% of individuals in the sample who complete their education by 19 years.

Finally, I test the robustness of results using a duration model with the maximum likelihood estimator. Indeed, some concerns may arise on the use of a linear probability model with a binary dependent variable which is considered as a continuous one. The objective of maximum likelihood estimation is to find the values of the estimated parameters that maximize the probability of observing the values of the dependent variable in the sample. Survival analysis is used to measure the time to an event of interest such as school drop-out, which occurs only once in this setting. I wish to analyze time until drop-out of school, measured in years. The individuals range in age from 7 years to the end of their schooling. I estimate a Cox proportional hazard model (Cox, 1972). Table 10 reports the results only on female education and disaggregates the analysis by drought intensities. The hazard ratio for a drought of any intensity in the last 4 columns is interpreted as the proportional change in hazard when drought is increased by one unit. I find in column 1 that drought results in a higher hazard and therefore a smaller survival time for girls with only young brothers (first line). Conversely, having only young sisters significantly decreases the risk of being out-of-school when a shock occurs (the interaction between drought and a young sex ratio equals to 2 has an hazard ratio inferior to 1, which indicates a negative coefficient). The findings are the same in column 3 when I interact drought with simultaneously young sibship size and gender. However, drought does not have a shallower effect on girls with more female-skewed sex ratio when old brotherhood variables are also included in the estimation. Nevertheless, the second part of the table clearly identifies a young-sibling gender effect which alter how severe drouths will impact daughters. Indeed, the interaction between severe drought and a sex ratio equals to 2 is significant in all specifications, at the 5% level, corroborating previous findings; severe droughts will have its largest effect on young girls in mixed-gender families.

8 Conclusion

This paper studies the existing relationship between economic conditions, sibling composition and children's outcomes when markets are incomplete, in a specific region of Tanzania. The study reveals that uninsured households withdraw specific children from school in response to unanticipated income shocks. The size but also sex ratio among the young brotherhood determine how children will be affected by drought. Being a girl, especially among the youngest, is associated with reduced likelihood of enrollment when the family is affected by a drop of income. Older sisters may also be negatively affected in household with opposite-sex young siblings. Thus, girls benefit from having more young sisters when a shock occurs, as they help share burdens and offer less competition. Globally, the finding indicates that parents adopt strategies worsening gender gap within household when the economic environment becomes uncertain. An increase in liquidity to buffer the shock may come mainly from employment and sligthly from early marriage amongst rural girls. The long-term analysis on temporary climate shocks points out permanent effects on young girls' human capital development and thus future earnings.

Considering that gender equality has been identified as one of the Millenium Development Goals, this paper shows that mechanisms protecting rural household from agricultural shocks through saving options, are likely to reduce gender gap both in educational attainment as well as on the marriage and labor market. These latter outcomes define the degree of risky environments that individuals will face when starting a family of their own. The paper would thus encourage governments to adapt to climate change, which seems to have not only negative effects on environment but also damaging consequences on gender-equity, child educational attainment and thereby economic development.

References

Aker, J. C. and I. M. Mbiti (2010). Mobile phones and economic development in africa. *The Journal of Economic Perspectives 24*(3), 207–232.

Akresh, R. (2004). Adjusting household structure: school enrollment impacts of child fostering in burkina faso. BREAD Working Paper No. 89.

Akresh, R., E. Bagby, D. D. Walque, and H. Kazianga (2012). Child ability and household human capital investment decisions in burkina faso. *Economic Development and Cultural Change 61*(1), 157–186.

Alam, S. A. (2015). Parental health shocks, child labor and educational outcomes: Evidence from tanzania. *Journal of Health Economics 44*, 161–175.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics 128*(2), 469–530.

Allendorf, K. (2007). Do women's land rights promote empowerment and child health in nepal? *World Development 35*(11), 1975–1988.

Almer, C., J. Laurent-Lucchetti, and M. Oechslin (2017). Water scarcity and rioting: Disaggregated evidence from sub-saharan africa. *Journal of Environmental Economics and Management 86*, 193–209.

Altonji, J. G. (1991). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics 11*(1, Part 1), 48–83.

Anderson, S. (2007). The economics of dowry and brideprice. *Journal of Economic Perspectives 21*(4), 151–174.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2018). Bride price and female education. *Journal of Political Economy*, Forthcoming.

Attanasio, O. and K. Kaufmann (2009). Educational choices, subjective expectations, and credit constraints. *No. w15087. National Bureau of Economic Research.*

Bandara, A., R. Dehejia, and S. Lavie-Rouse (2015). The impact of income and non-income shocks on child labor: Evidence from a panel survey of tanzania. *World Development 67*, 218–237.

Banerjee, A. V. and E. Duflo (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty.* PublicAffairs.

Becker, G. S. (1992). Human capital revisited. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*, pp. 15–28. The university of Chicago press.

Becker, G. S., K. M. Murphy, and J. L. Spenkuch (2016). The manipulation of children's preferences, old-age support, and investment in children's human capital. *Journal of Labor Economics 34*(S2), S3–S30.

Becker, G. S. and N. Tomes (1976). Child endowments and the quantity and quality of children. *Journal of political Economy 84*(4, Part 2), S143–S162.

Beegle, K., J. De Weerdt, and S. Dercon (2011). Migration and economic mobility in tanzania: Evidence from a tracking survey. *The Review of Economics and Statistics 93*(3), 1010–1033.

Beegle, K., R. H. Dehejia, and R. Gatti (2006). Child labor and agricultural shocks. *Journal of Development Economics 81*, 80–96.

Beegle, K., R. H. Dehejia, R. Gatti, and S. Krutikova (2008). The consequences of child labor: Evidence from longitudinal data in rural tanzania. Policy Research Working Paper no. 4677, World Bank, Washington, DC.

Behrman, J. R. (1986). Nutrition, health, birth order and seasonality. *Journal of Development Economics 28*, 43–62.

Behrman, J. R. (1997). Intrahousehold distribution and the family. *Handbook of population and family economics 1*, 125–187.

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

Björkman-Nyqvist, M. (2013). Income shocks and gender gaps in education: Evidence from uganda. *Journal of Development Economics 105*, 237–253.

Cameron, L. and M. Shah (2015). Risk-taking behavior in the wake of natural disasters. *The Journal of Human Resources 50*(2), 484–515.

Carter, M. R., P. D. Little, and T. Mogues (2007). Poverty traps and natural disasters in ethiopia and honduras. *World Development 35*(5), 835–856.

Caruso, G. D. (2017). The legacy of natural disasters: The intergenerational impact of 100 years of disasters in latin america. *Journal of Development Economics 127*, 209–233.

Cattaneoa, C. and G. Peri (2016). The migration response to increasing temperatures. *Journal of Development Economics 122*, 127–146.

Cogneau, D. and R. Jedwab (2012). Commodity price shocks and child outcomes: The 1990 cocoa crisis in cote d'ivoire. *Economic Development and Cultural Change 60*(3), 507–534.

Corno, L., N. Hildebrandt, and A. Voena (2017). Age of marriage, weather shocks, and the direction of marriage payments. *NBER Working Paper No. 23604*.

Corno, L. and A. Voena (2016). Selling daughters: age of marriage, income shocks and the bride price tradition. IFS Working Paper W1 /08.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological) 34*(2), 187–220.

De Weerdt, J. (2010). Moving out of poverty in tanzania: Evidence from kagera. *Journal of Development Studies 46*(2), 331–349.

De Weerdt, J., K. Beegle, H. Bie Lilleør, S. Dercon, K. Hirvonen, M. Kirchberger, and S. Krutikov (2012.). Kagera health and development survey, 2010: Basic information document. Working Paper no. 46, Rockwool Foundation.

De Weerdt, J. and K. Hirvonen (2016). Risk sharing and internal migration. *Economic Development and Cultural Change 65*(1), 63–86.

Delavande, A., X. Giné, and D. McKenzie (2011). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics 94*, 151–163.

Dell, M., B. F. Jones, and B. A. Olken (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics 4*(3), 66–95.

Dizon-Ross, R. (2018). Parents' beliefs about their children's academic ability: Implications for educational investments. *National Bureau of Economic Research* (No. w24610).

Edlund, L. (1999). Son preference, sex ratios, and marriage patterns. *Journal of Political Economy 107*(6), 1275–1304.

Fernández, R. (2014). Women's rights and development. *Journal of Economic Growth 19*(1), 37–80.

Field, C. B. (2014). *Climate change 2014–Impacts, adaptation and vulnerability: Regional aspects*. Cambridge University Press.

Glewwe, P. and E. M. King (2001). The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter? *The World Bank Economic Review 15*(1), 81–113.

Gray, C. and V. Mueller (2012). Drought and population mobility in rural ethiopia. *World Development 40*(1), 134–145.

Halliday, T. (2006). Migration, risk, and liquidity constraints in el salvador. *Economic Development and Cultural Change 54*(4), 893–925.

Handa, S. (2002). Journal of development economics. *Journal of Development Economics 69*, 103–128.

Harari, M. and E. La Ferrara (2013). Conflict, climate and cells: A disaggregated analysis. *Review of Economics and Statistics* (0).

Huisman, J. and J. Smits (2009). Effects of household- and district-level factors on primary school enrollment in 30 developing countries. *World Development 37*(1), 179–193.

Jack, W. and T. Suri (2014). Risk sharing and transactions costs: Evidence from kenya's mobile money revolution. *American Economic Review 104*(1), 183–223.

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

Jayachandran, S. and R. Pande (2017). Why are indian children so short? the role of birth order and son preference. *American Economic Review 107*(9), 2600–2629.

Jensen, R. (2000). Agricultural volatility and investments in children. *American Economic Review Papers and Proceedings 90*(2), 399–404.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics 125*(2), 515–548.

Kishore, A. and D. Spears (2014). Having a son promotes clean cooking fuel use in urban india: Women's status and son preference having a son promotes clean cooking fuel use in urban india: Women's status and son preference. *Economic Development and Cultural Change 62*(4), 673–699.

Kreft, S., S. Harmeling, C. Bals, W. Zacher, and K. van de Sand (2010). The millennium development goals and climate change: Taking stock and looking ahead. Technical report, Bonn: Germanwatch., www.germanwatch.org/klima/klimdg10e.htm.

Lafortune, J. and S. Lee (2014). All for one? family size and children's educational distribution under credit constraints. *American Economic Review Papers and Proceedings 104*(5), 365–369.

Levine, D. and M. Kevane (2003). Are investments in daughters lower when daughters move away? evidence from indonesia. *World Development 31*(6), 1065–1084.

Li, L. and X. Wu (2011). Gender of children, bargaining power, and intrahousehold resource allocation in china. *Journal of Human Resources 46*(2), 295–316.

Lobell, D. B., W. Schlenker, and J. Costa-Roberts (2011). Climate trends and global crop production since 1980. *Science*, 1204531.

Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *The American Economic Review 99*(3), 1006–1026.

Morduch, J. (2000). Sibling rivalry in africa. *The American Economic Review 90*(2), 405–409.

Munyegera, G. K. and T. Matsumoto (2016). Mobile money, remittances, and household welfare: Panel evidence from rural uganda. *World Development 79*, 127–137.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from madagascar. Technical report, MIT Working Paper.

Orazem, P. F. and E. M. King (2007). Schooling in developing countries: The roles of supply, demand and government policy. *Handbook of Development Economics 4*, 3475–3559.

Parish, W. L. and R. J. Willis (1993). Daughters, education, and family budgets taiwan experiences. *The Journal of Human Resources 28*(4), 863–898.

Pullum, T. W. and S. Staveteig (2017). An assessment of the quality and consistency of age and date reporting in dhs surveys,2000-2015. In *ICF*.

Quisumbing, A. R. (2003). Food aid and child nutrition in rural ethiopia. *World Development 31*(7), 1309–1324.

Rebeca, E. and H. Javier (2016). The missing link between parents' preferences and daughters' survival: The moderator effect of societal discrimination. *World Development 78*, 372–385.

Rosenzweig, M. R. and O. Stark (1989). Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of Political Economy 97*(4), 905–926.

Rossi, P. and L. Rouanet (2015). Gender preferences in africa: A comparative analysis of fertility choices. *World Development 72*, 326–345.

Serra, R. (2009). Child fostering in africa: When labor and schooling motives may coexist. *Journal of Development Economics 88*(1), 157–170.

Shah, M. and B. M. Steinberg (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy 125*(2), 527–561.

Strauss, J. and D. Thomas (1995). Human resources: Empirical modeling of household and family decisions. *Handbook of development economics 3*, 1883–2023.

Thomas, D., K. Beegleb, E. Frankenbergc, B. Sikokid, J. Strausse, and G. Teruelf (2004). Education in a crisis. *Journal of Development Economics 74*(1), 53–85.

UIS (2017). Global education monitoring report. *UNESCO Institute for Statistics*.

Vogl, T. S. (2013). Marriage institutions and sibling competition: Evidence from south asia. *The Quarterly Journal of Economics*, 1017–1072.

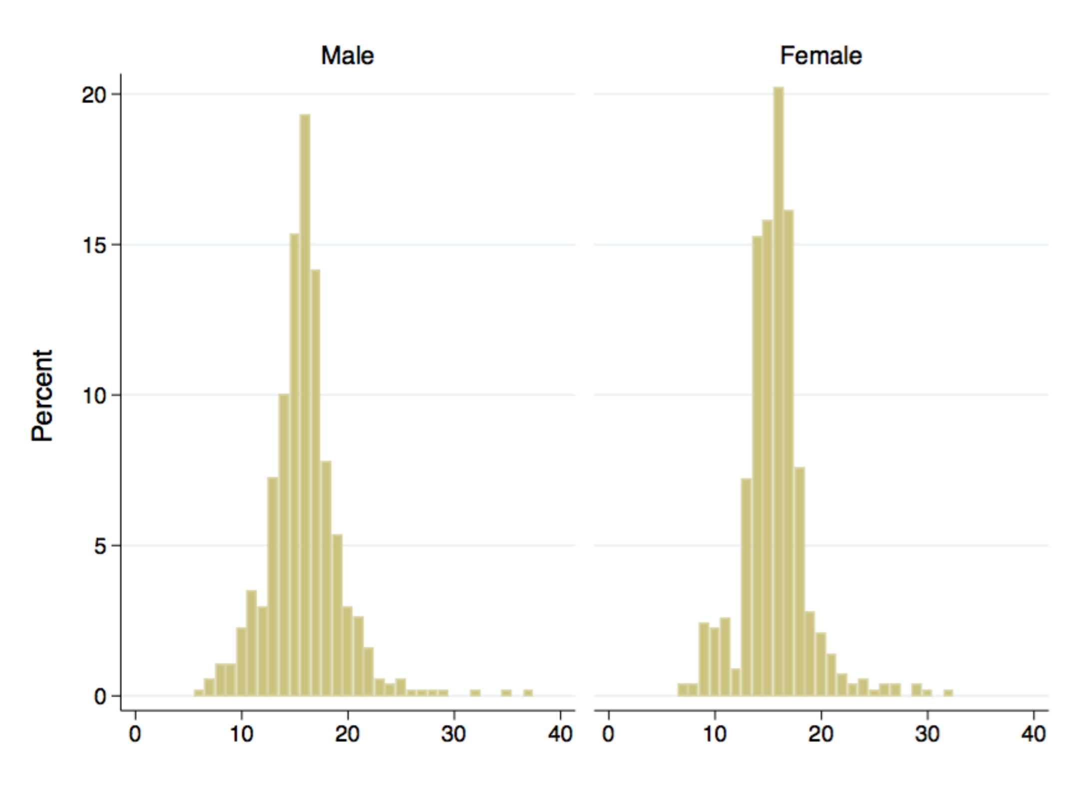

Figure 1: Distribution of school-leaving age, by sex.

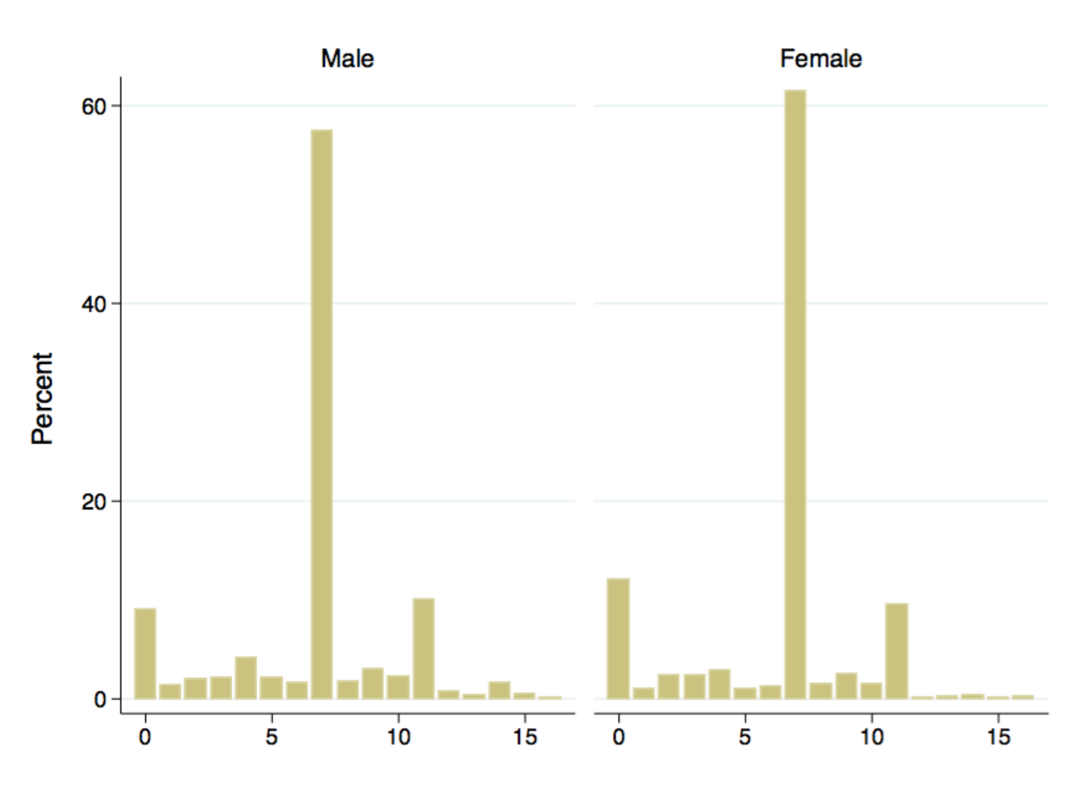

Figure 2: Distribution of educational attainment for individuals who have completed their schooling in 2010, by sex.

Table 1: Qualitative results

	Freq.	Percent
Panel A: Selection of children during shocks, by educational grade		
Primary child	1	2.33
Secondary child	36	83.72
No concensus	6	13.95
Panel B: Selection of children during shocks, by gender		
Girls	3	7.32
Both gender	38	92.68
Panel C: Unprompted answers on girls at risk of becoming pregnant		
Mentioned	15	33.33
No mentioned	30	66.67

The qualitative study is realized across 15 selected villages in 4 districts of Kagera in 2018, resulting in 270 respondents among 45 groups. In each village, a group of women, then women and finally a mixed-gender is selected. Respondents are parents involving in agriculture with at least one child enrolled (have been enrolled) in primary or secondary school.

Table 2: Summary statistics

Panel A: Summary statistics for continuous variables

Variable	Number of obs.	Mean	Std. Dev.	Minimum	Maximum
Number of siblings	11074	5.2	2.7	0	14
Number of sisters	10539	2.8	1.8	0	11
Number of brothers	10621	2.7	1.6	0	8
Age of school entrance *	11074	9.1	1.7	4	17
Educational attainment *	11019	7.3	2.2	0	16
Age of school dropout *	11074	16.6	3.3	7	37

Panel B: Summary statistics for dummy variables

Variable	Value	Number of obs.	Percentage
Gender composition among younger siblings	Only males	2,204	23.9
	Mixed gender	4,955	53.7
	Only females	2,069	22.4
Gender composition among older siblings *	Only males	1,652	18.4
	Mixed gender	5,381	59.9
	Only females	1,955	21.8
Number of younger siblings	0	1,846	16.7
	1	2,267	20.5
	2	2,335	21
	3	2,061	18.6
	More than 4	2,565	23.1
Number of older siblings *	0	2,086	18.8
	1	1,868	16.9
	2	1,812	16.4
	3	1,459	13.2
	More than 4	3,849	34.8
Severe droughts	1	1,524	13.8
Droughts	1	3,095	28

Notes: All variables are constructed between 1972 and 2010, from restrospective information of the 2010 modules on siblings and household members. The time span is determined by school years, which vary for each individual in the sample. Variables indicated with a (*) do not vary across time. Severe droughts variable includes any extreme or severe climate shocks whereas the droughts variable encompasses climate events of any intensity.

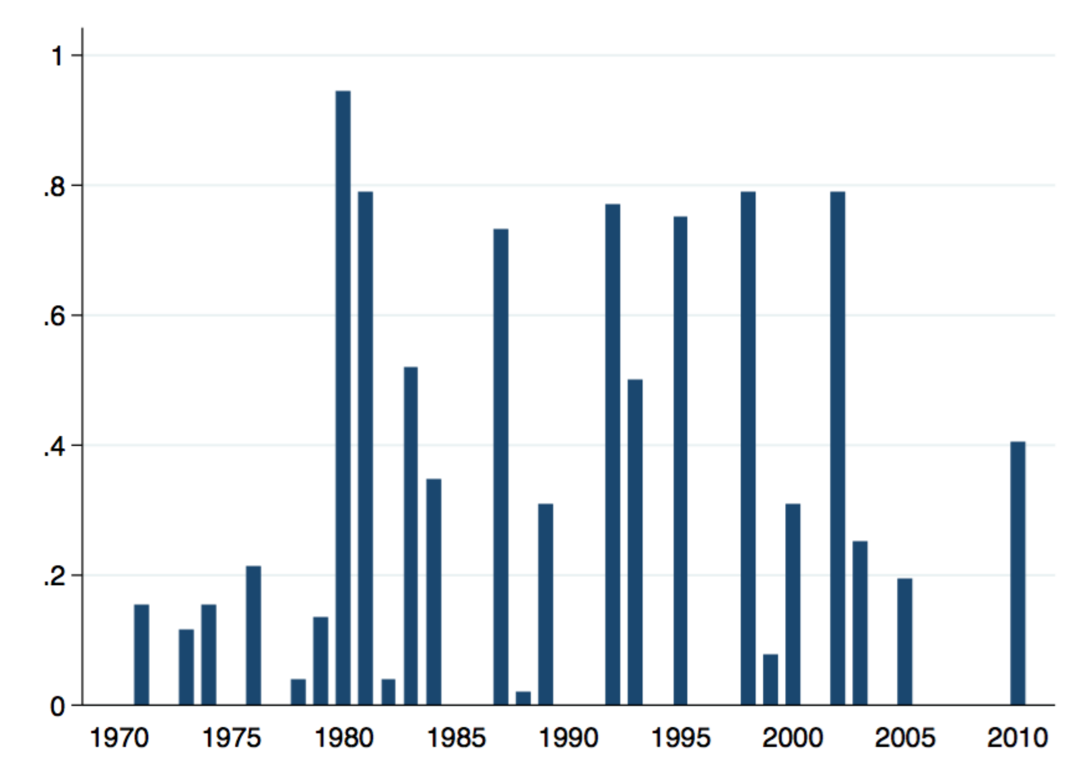

Figure 3: Evolution of droughts over time, by intensity.

Table 3: Impact of droughts on household consumption and individual expenditures

Dependent variable	Consumption p/c (in logs)	Food consumption p/c (in logs)	Individual expenditures (in logs)
Variables	(1)	(2)	(3)
Drought	-0.306***	-0.162***	-0.229***
	(0.062)	(0.043)	(0.065)
Household FE	Yes	Yes	No
Individual FE	No	No	Yes
Survey FE	Yes	Yes	Yes
Obs/#Individuals	9,346/ 3,752	9,268 /3,716	34,507/17,275
R-squared	0.319	0.501	0.282

Standard error in parentheses, clustered at the village level.

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

Results are estimated across the six waves of the KHDS database, between 1991 and 2010. All estimates include survey fixed effects. Household fixed effects mainly controls for family health status, which is proxy by the father's education in most cases. Consumption per capita includes household expenditures on education, health and durables goods like home or vehicle repairs, and food consumption per capita. Food consumption per capita is the household food consumption during rainy and dry seasons divided by the number of members in the family. Individual expenditures contains expenditures on clothes, jewelry, toys, books, restaurants, tobacco, newspapers, cosmetics etc.

Table 4: Effects of drought on girls' school dropout by number and sex of siblings

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Drought	0.046**	0.031	0.074***	0.061	0.109**	0.071***
	(0.021)	(0.021)	(0.024)	(0.041)	(0.044)	(0.025)
Drought×Sex composition (Young)						
Mixed gender	-0.027		0.001	-0.001	0.011	-0.007
	(0.023)		(0.028)	(0.027)	(0.029)	(0.030)
Only females	-0.086**		-0.088***	-0.090***	-0.083***	-0.081***
	(0.027)		(0.027)	(0.026)	(0.027)	(0.029)
Drought×Number of siblings (Young)						
2		-0.044	-0.062*	-0.066**	-0.057	-0.052*
		(0.026)	(0.029)	(0.029)	(0.032)	(0.030)
3		-0.034	-0.069**	-0.078**	-0.070**	-0.063*
		(0.025)	(0.030)	(0.030)	(0.032)	(0.032)
At least 4		-0.006	-0.046	-0.064*	-0.024	-0.029
		(0.028)	(0.036)	(0.035)	(0.041)	(0.039)
Drought×Sex Composition (Old)						
Mixed gender					0.002	
					(0.038)	
Only females					-0.022	
					(0.036)	
Drought×Number of siblings (Old)						
2					-0.034	
					(0.036)	
3					-0.089*	
					(0.042)	
At least 4					-0.067	
					(0.039)	
Drought×Age				0.004		
				(0.003)		
Drought×Birth Order				-0.010*		
				(0.006)		
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual & Year FE	Yes	Yes	Yes	Yes	Yes	No
Family & Cohort FE	No	No	No	No	No	Yes
Obs/#Individuals	4,627/484	4,627/484	4,627/484	4,627/484	3,686/386	4,627/484
R-squared	0.244	0.243	0.245	0.245	0.243	

Standard error in parentheses, clustered at the family level.

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

Drought is a dummy equals to 1 for a drought of any intensity in a given year. Controls include the number of younger siblings and their gender composition, which vary over time. The first five columns include individual and year fixed effects. Family and year of birth fixed effects are introduced in the last column.

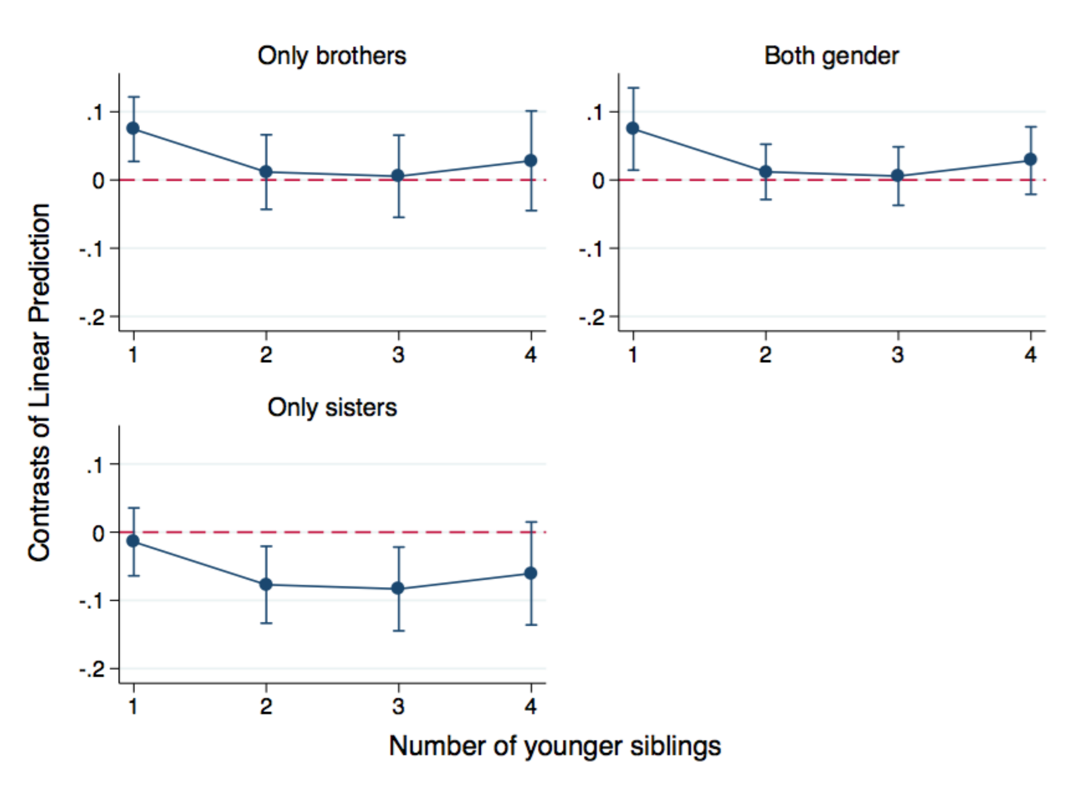

Figure 4: How the net effect of drought varies with the number of young siblings by their sex ratio?

Table 5: Effects of drought on boys' school dropout by number and sex of siblings

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Drought	-0.028	-0.027	-0.028	0.040	-0.036	-0.035
	(0.016)	(0.020)	(0.021)	(0.038)	(0.044)	(0.022)
Drought×Sex composition (Young)						
Mixed gender	0.035*		0.031	0.030	0.025	0.027
	(0.018)		(0.027)	(0.026)	(0.030)	(0.032)
Only females	0.001		0.001	-0.002	0.026	0.008
	(0.021)		(0.021)	(0.022)	(0.026)	(0.028)
Drought×Number of siblings (Young)						
2		0.009	-0.004	-0.008	-0.011	-0.007
		(0.022)	(0.024)	(0.024)	(0.027)	(0.029)
3		0.027	0.005	0.001	0.013	0.017
		(0.025)	(0.035)	(0.036)	(0.040)	(0.036)
At least 4		0.035	0.007	0.003	-0.010	0.017
		(0.027)	(0.035)	(0.035)	(0.041)	(0.041)
Drought×Sex ratio (Old)						
Mixed gender					0.001	
					(0.042)	
Only females					0.008	
					(0.027)	
Drought×Number of siblings (Old)						
2					-0.013	
					(0.049)	
3					0.019	
					(0.041)	
At least 4					0.001	
					(0.043)	
Drought×Age				-0.001		
				(0.002)		
Drought×Birth Order				-0.012**		
				(0.005)		
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual & Year FE	Yes	Yes	Yes	Yes	Yes	No
Family & Cohort FE	No	No	No	No	No	Yes
Obs/#Individuals	4,599/469	4,599/469	4,599/469	4,601/469	3,633/374	4,601/469
R-squared	0.226	0.225	0.226	0.226	0.234	

Standard error in parentheses, clustered at the family level.

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

Drought is a dummy equals to 1 for a drought of any intensity in a given year. Controls include the number of younger siblings and their gender composition, which vary over time. The first five columns include individual and year fixed effects. Family and year of birth fixed effects are introduced in the last column.

Table 6: Long-term effect of droughts, with different intensity, on a child's educational attainment by gender

Variables	Droughts of any intensity				Extreme or severe droughts			
	Girls		Boys		Girls		Boys	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drought 6-10	-0.397	0.361	-0.778	0.235	-0.470	0.442	-0.325	-0.471
	(0.480)	(0.595)	(0.555)	(0.435)	(0.525)	(0.819)	(0.628)	(0.700)
Drought 11-15	-0.427	-0.056	-0.088	-0.207	0.092	0.274	0.187	-0.469
	(0.480)	(0.511)	(0.513)	(0.461)	(0.547)	(0.748)	(0.622)	(0.774)
Drought 16-20	-0.813	0.928**	-0.634	-0.197	-0.195	0.402	0.308	-0.519
	(0.705)	(0.454)	(0.524)	(0.375)	(0.718)	(0.647)	(0.707)	(0.706)
Interactions with the young sibship								
Drought 6-10×Sex ratio	0.131		0.076		0.035		0.466	
	(0.475)		(0.515)		(0.696)		(0.630)	
Drought 11-15×Sex ratio	-0.168		-0.395		-0.049		-1.393*	
	(0.549)		(0.536)		(0.613)		(0.782)	
Drought 16-20×Sex ratio	-0.391		0.861		-0.260		0.531	
	(0.727)		(0.528)		(0.768)		(0.700)	
Drought 6-10×Number of siblings	0.105		0.143		-0.001		0.046	
	(0.093)		(0.090)		(0.134)		(0.123)	
Drought 11-15×Number of siblings	0.057		0.027		-0.077		0.079	
	(0.094)		(0.082)		(0.145)		(0.137)	
Drought 16-20×Number of siblings	0.217**		0.037		0.036		-0.182	
	(0.094)		(0.076)		(0.102)		(0.143)	
Interactions with the old sibship								
Drought 6-10×Sex ratio		-0.910*		0.750		-2.043**		1.050
		(0.517)		(0.502)		(0.798)		(0.681)
Drought 11-15×Sex ratio		0.395		-0.212		-0.459		0.064
		(0.540)		(0.380)		(0.801)		(0.660)
Drought 16-20×Sex ratio		-0.476		-0.339		-0.964		0.172
		(0.518)		(0.490)		(0.665)		(0.764)
Drought 6-10×Number of siblings		0.021		-0.214*		0.104		-0.107
		(0.139)		(0.116)		(0.125)		(0.135)
Drought 11-15×Number of siblings		-0.041		0.057		0.002		0.131
		(0.099)		(0.084)		(0.122)		(0.125)
Drought 16-20×Number of siblings		-0.210		0.034		0.019		0.168
		(0.131)		(0.103)		(0.142)		(0.146)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household & Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs/#Families	644/294	556/265	584/278	507/253	644/294	556/265	584/278	507/253
R-squared	0.170	0.174	0.182	0.250	0.153	0.181	0.188	0.248

Standard error in parentheses, clustered at the family level.

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

The regression is estimated from restrospective information in 2010, the family structure does not evolve over time. I restrict the analysis to girls who have complete their education and lived at the baseline village during school years. The drought variable represents the number of droughts experienced at different age ranges. In the last four columns, the variable is limited to severe and extreme climatic events. Controls include the number of younger siblings and their gender composition in columns 1, 3, 5, 7 whereas controls in the remaining columns include variables reflecting the old sibling group. These controls reflecting sibling groups do not vary over time. Estimates include household and birth of year fixed effects.

Table 7: Effect of droughts on female labor in agriculture and on the timing of marriage, conditional on the number and gender of siblings

Dependent variable	Hours worked in agriculture			Being married before 20		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Drought	0.344**	0.180	0.621***	0.024	-0.014	-0.002
	(0.151)	(0.198)	(0.239)	(0.015)	(0.012)	(0.021)
Drought×Sex Composition (Young)						
Mixed gender	-0.152		-0.124	0.011		0.009
	(0.192)		(0.208)	(0.014)		(0.016)
Only females	-0.356*		-0.319*	0.006		0.017
	(0.186)		(0.190)	(0.015)		(0.017)
Drought×Number of siblings (Young)						
2	-0.030		-0.001	-0.027*		-0.018
	(0.172)		(0.173)	(0.015)		(0.017)
3	-0.126		-0.148	-0.039**		-0.029
	(0.214)		(0.231)	(0.018)		(0.019)
At least 4	-0.151		-0.227	-0.044**		-0.028
	(0.228)		(0.256)	(0.018)		(0.021)
Drought×Sex Composition (Old)						
Mixed gender		0.233	0.032		0.001	-0.020
		(0.204)	(0.216)		(0.018)	(0.021)
Only females		0.115	-0.010		-0.006	-0.009
		(0.214)	(0.222)		(0.015)	(0.017)
Drought×Number of siblings (Old)						
2		-0.616***	-0.670***		0.029	0.041**
		(0.222)	(0.221)		(0.018)	(0.019)
3		-0.185	-0.197		0.037*	0.048**
		(0.236)	(0.233)		(0.020)	(0.023)
At least 4		-0.425*	-0.414*		0.020	0.043**
		(0.233)	(0.227)		(0.018)	(0.021)
Controls	Yes	No	Yes	Yes	No	Yes
Individual & Survey/Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs/#Individuals	1,576/535	1,542/527	1,319/449	7,678/786	7,368/762	6,059/626
R-squared	0.038	0.044	0.055	0.117	0.118	0.113

Standard error in parentheses, clustered at the family level.

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

Drought is a dummy equals to 1 for a drought of any intensity in a particular year. Columns 1 to 6 estimate the effect of drought on the number of worked hours in agriculture, in logs, for girls aged between 7-15 years old and are based on the first four waves of the KHDS database. Individuals and survey fixed effects are included. Columns 7, 8 and 9 estimate the effect of drought on the timing of early marriage, based on the constructed panel from information in 2010. The dependent variable is a binary variable equals to 1 the year of marriage if the girl is less than 20 years old. Time-invariant unobservables are controlled with individual and year fixed effects. Controls include the number of younger siblings and their gender composition, which vary over time and across surveys.

Table 8: Effect of droughts on girls' drop out of school according to school-age sibling compositions

Dependent variable	Girls probability of school dropout				6-9 group	10-14 group	15-19 group
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Drought	0.055*	0.015	0.019	0.022	0.018	0.032	0.112
	(0.030)	(0.013)	(0.019)	(0.020)	(0.015)	(0.029)	(0.090)
Drought×Number of sisters	-0.030**						
	(0.014)						
Drought×Number of siblings	0.007						
	(0.009)						
Drought×Number of sisters (Young)		-0.049**		-0.050***	-0.003	-0.045**	-0.018
		(0.019)		(0.019)	(0.010)	(0.019)	(0.050)
Drought×Number of siblings (Young)		0.023*		0.024*	-0.009	0.012	-0.030
		(0.013)		(0.013)	(0.010)	(0.013)	(0.040)
Drought×Number of sisters (Old)			0.014	0.014	-0.006	-0.005	0.179**
			(0.013)	(0.013)	(0.011)	(0.016)	(0.080)
Drought×Number of siblings (Old)			-0.010	-0.011	0.001	-0.005	-0.062
			(0.010)	(0.010)	(0.006)	(0.013)	(0.071)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs/#Individuals	5,513/578	5,404/578	5,404/578	5,404/578	1,728/578	2,651/560	1,025/396
R-squared	0.244	0.252	0.249	0.253	0.102	0.135	0.441

Standard error in parentheses, clustered at the family level.

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

Drought is a dummy equals to 1 for a drought of any intensity in a given year. Variables reflecting family composition are constructed from school-age siblings, aged between 6 to 19 years old. The last 3 columns stratified the analysis into 3 age groups. Estimations include individual and year fixed effects. Controls denote the number of school-age (younger and older) sisters and siblings, which vary over time.

Table 9: Robust effects of drought on girls' school dropout with fixed and complete family structures

					Families with completed fertility			
Variables	(1)	(2)	(3)	(4)	(1')	(2')	(3')	(4')
Drought	0.066***	0.046	0.037	0.103**	0.055*	0.059	0.030	0.117**
	(0.024)	(0.043)	(0.034)	(0.044)	(0.029)	(0.057)	(0.040)	(0.051)
Drought $\times$ Sex ratio (Young)								
Mixed gender	0.002	0.001		-0.001	-0.028	-0.028		-0.016
	(0.032)	(0.031)		(0.033)	(0.037)	(0.036)		(0.037)
Only females	-0.081***	-0.082***		-0.083***	-0.077**	-0.080**		-0.088**
	(0.029)	(0.029)		(0.030)	(0.036)	(0.035)		(0.034)
Drought $\times$ Number of siblings (Young)								
2	-0.045	-0.049		-0.045	-0.024	-0.030		-0.045
	(0.036)	(0.036)		(0.038)	(0.042)	(0.042)		(0.042)
3	-0.077**	-0.087**		-0.079**	-0.082*	-0.092**		-0.101**
	(0.035)	(0.035)		(0.037)	(0.042)	(0.043)		(0.044)
At least 4	-0.044	-0.058		-0.030	-0.024	-0.040		-0.043
	(0.036)	(0.036)		(0.041)	(0.047)	(0.046)		(0.049)
Drought $\times$ Sex ratio (Old)								
Mixed gender			0.024	0.001			0.038	0.002
			(0.034)	(0.038)			(0.038)	(0.043)
Only females			0.006	-0.019			0.015	-0.014
			(0.033)	(0.036)			(0.038)	(0.042)
Drought $\times$ Number of siblings (Old)								
2			-0.024	-0.024			-0.039	-0.034
			(0.034)	(0.035)			(0.042)	(0.042)
3			-0.080**	-0.076*			-0.092**	-0.076
			(0.039)	(0.041)			(0.046)	(0.046)
At least 4			-0.061*	-0.057			-0.078*	-0.065
			(0.036)	(0.039)			(0.042)	(0.044)
Drought $\times$ Age		0.004				0.003		
		(0.003)				(0.003)		
Drought $\times$ Birth Order		-0.009				-0.008		
		(0.006)				(0.008)		
Individual & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs/#Individuals	4,664/486	4,664/486	4,492/472	3,711/389	3,458/355	3,458/355	3,668/378	3,041/313
R-squared	0.243	0.244	0.241	0.241	0.235	0.236	0.233	0.236

Standard error in parentheses, clustered at the family level.

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

Sibling composition is constructed from information in 2010 and does not vary over years in this setting. The last four columns reduce the analysis to families with completed fertility, (ie) the mother is more than 49 years old in 2010. Drought is a dummy equals to 1 for a drought of any intensity in a given year. All estimates include individual and year fixed effects.

Table 10: Results of drought on female education survival analysis depending on sibling sex ratio and number

Variables	Droughts of any intensity				Extreme or severe droughts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drought	0.311*	0.151	0.401*	0.515	0.392*	-0.003	0.296	0.108
	(0.180)	(0.180)	(0.217)	(0.392)	(0.206)	(0.193)	(0.225)	(0.434)
Drought×Sex composition (Young)								
Mixed gender	-0.313		-0.171	-0.029	-0.272		-0.716*	-0.788*
	(0.207)		(0.321)	(0.377)	(0.245)		(0.376)	(0.406)
Only females	-0.610*		-0.617*	-0.451	-0.855**		-0.833**	-0.862**
	(0.314)		(0.317)	(0.342)	(0.385)		(0.387)	(0.421)
Drought×Number of siblings (Young)								
2		-0.163	-0.165	-0.173		-0.011	0.200	0.233
		(0.293)	(0.337)	(0.400)		(0.384)	(0.410)	(0.465)
3		-0.239	-0.305	-0.216		-0.216	0.077	0.159
		(0.265)	(0.336)	(0.372)		(0.333)	(0.396)	(0.447)
At least 4		-0.140	-0.219	-0.121		0.356	0.742*	0.960
		(0.224)	(0.393)	(0.533)		(0.236)	(0.415)	(0.588)
Drought×Sex composition (Old)								
Mixed gender				-0.185				0.074
				(0.325)				(0.374)
Only females				-0.192				-0.144
				(0.307)				(0.345)
Drought×Number of siblings (Old)								
2				-0.090				0.300
				(0.331)				(0.377)
3				-0.354				0.031
				(0.344)				(0.421)
At least 4				-0.237				-0.060
				(0.336)				(0.401)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,627	4,627	4,627	3,686	4,627	4,627	4,627	3,686

Standard error in parentheses, clustered at the family level.

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

Results are estimated with maximum likelihood. The table reports hazard ratios, that is exponentiated coefficients. Drought is a dummy equals to 1 for a drought of any intensity in a given year in the fours columns. Then, the variable is restricted to severe and extreme climatic events. Controls include the number of younger (older) siblings and their gender composition.