Aix-Marseille Université - École Doctorale de Sciences Économiques et de Gestion d'Aix-Marseille N°372

Faculté d'Économie et de Gestion

Aix-Marseille School of Economics

Numéro attribué par la bibliothèque

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |

**Thèse pour le Doctorat en Sciences Économiques**

*Présentée et soutenue publiquement par*

**Mathilde VALERO**

*le 18 septembre 2020*

en vue de l'obtention du grade de docteur d'Aix-Marseille Université

# Education and Gender in Developing Countries

*Jury :*

| | |
|---|---|
| **Siwan ANDERSON** | Vancouver School of Economics, *Rapporteur* |
| **Mathieu COUTTENIER** | ENS Lyon, *Rapporteur* |
| **Habiba DJEBBARI** | Aix-Marseille University, *Examinateur* |
| **Alain VENDITTI** | Aix-Marseille University *Directeur de thèse* |

*L'université d'Aix-Marseille n'entend ni approuver, ni désapprouver les opinions particulières du candidat :*

*Ces opinions doivent être considérées comme propres à leur auteur.*

*A ma famille*

# Remerciements

Ce travail n'aurait pas été possible sans ceux qui ont cru en moi et qui m'ont accompagné, d'une façon ou d'une autre, dans cette aventure.

Je tiens avant tout à remercier mes directeurs de thèse, Alain Venditti et Marc Sangnier. Malgré la configuration peu ordinaire, vous m'avez tous les deux accordé votre soutien et votre confiance. Marc, je tiens à vous remercier en particulier pour votre encadrement, votre patience et votre disponibilité. Vous, ainsi que Nicolas Berman et Roberta Ziparo, avez réellement contribué à mener à bien cette recherche. J'ai pu à la fois apprendre de vos compétences professionnelles, mais aussi de vos grandes qualités humaines lors de nos réunions de groupe à l'AMSE.

Je souhaite également remercier Siwan Anderson, Mathieu Couttenier et Habiba Djebbari. Non seulement pour avoir accepté de faire parti de mon jury de thèse mais aussi pour m'avoir soutenue au cours de ma recherche. Siwan en tant que membre de mon comité de deuxième année, Habiba pour ces nombreuses conversations passionnantes dans votre bureau. Enfin, Mathieu, un grand merci pour m'avoir si bien accueillie lors de cette dernière année de thèse avec toute l'équipe de l'ENS Lyon. Aussi, je vous remercie pour les suggestions et les commentaires dont vous m'avez fait part dans vos rapports de pré-soutenance, qui ont contribué à améliorer la version finale de ma thèse.

J'aimerais ensuite remercier ma co-auteur Pauline Morault pour les échanges productifs et amicaux que nous avons pu avoir, notamment au cours de mon séjour à l'Université Catholique de Louvain. Merci pour ton soutien et ta réactivité lors de ces changements de dernières minutes au téléphone.

Mes remerciements s'adressent à l'ensemble des membres de l'AMSE, notamment les chercheurs qui ont commenté mes travaux à de multiples occasions. Parmi eux, je remercie

Sebastien Bervoets pour ses conseils et son soutien. Merci aux organisateurs des séminaires externes, grâce à qui j'ai pu bénéficier d'intéressantes réflexions intellectuelles avec Richard Akresh, Lucia Corno, Esther Duflo, Garance Génicot. Aussi, merci à l'Economic Development Initiatives (EDI), à Johanna Choumert et Joachim De Weerdt de m'avoir aidé dans la mise en place de mon projet de terrain au Kagera. Enquêter auprès de la population rurale de Tanzania, avec l'aide de mon ami Privatus Karugendo, restera une expérience inoubliable.

Je souhaite à présent remercier l'équipe du château Lafarge, qui a rendu cette expérience si chaleureuse. En particulier Patricia, Elisabeth, Jordan, Nandita, Patrice, Bruno, Tanguy, Marc, Eva, Anne-Sophie et bien d'autres. Merci à Bernadette pour ta gentillesse et ta disponibilité.

Enfin, cette thèse n'aurait pu aboutir sans le soutien inconditionnel de mes amis et de mes proches. Tout d'abord, les Doctorants. Un merci tout particulier à Laurène pour ton oreille attentive et ta présence, Rémi pour ton aide et ton amitié. Fatima, pour ta compassion et ta gentillesse. Une thèse sans les moments de joie de la team du cours sextius n'aurait pu être envisageable. Merci à Cyril, Kelly, Brice, Guillaume, Morgan, Laura d'avoir été présents.

Voici maintenant le moment de remercier mes amies de longue date qui m'ont suivie et soutenue dans ce parcours. Merci pour votre présence et pour votre humour qui m'a permis de relativiser dans tous ces moments de doutes. Laurie, je te remercie pour tes appels du dimanche soir, tes conseils et tes encouragements.

Enfin, un grand merci à ma famille, notamment Damien, Julie pour vos relectures. Une ligne particulière à ma Maman, qui a partagé et vécu à travers moi toutes les émotions de ces longues années. Merci à mes grands-parents pour la tendresse dont ils m'ont entouré. Papi, ton exemple de courage et d'acharnement ont été une source d'inspiration au cours de cette thèse. Ma gratitude se dirige enfin vers Jérémy, avec qui je partage cette expérience depuis le début. Merci de m'avoir suivi à Aix-en-Provence puis à Lyon pour de si beaux moments. Je te remercie pour ton soutien sans faille, et ta patience à toute épreuve lors de tous ces moments de travail et d'acharnements.

# Contents

# List of Tables

# List of Figures

# General Introduction

*"Basic education is not just an arrangement for training to develop skills, it is also a recognition of the nature of the world, with its diversity and richness, and an appreciation of the importance of freedom and reasoning as well as friendship. The need for that understanding - that vision - has never been stronger."* **Amartya Sen, Commonwealth education conference, 2003**.

In the past decade, millions of children around the world have gained access to educational opportunities. In developing countries, enrolment in primary education has reached 91% and basic literacy skills have improved tremendously. However, five years after the adoption of Sustainable Development Goal 4 in 2015, and the promise to provide universal primary and secondary education by 2030, there has been no progress in reducing the global number of out-of-school children.[1]

In 2018, more than 59 million primary age children remained out of school (UIS, 2019). As represented in Figure 1, more than half live in sub-Saharan Africa and South-West Asia. Access to higher education is even more dramatic. For instance in Tanzania, secondary school enrollment rate was lower than 31 percent in 2016 (UIS, 2017). In this line, Figure 2 shows that lower secondary out-of-school rate is nearly twice as large as the rate in primary education. Two potential reasons explain the recent increase in the number of out-of-school children and adolescents. According to UNESCO (2015), a number of countries have difficulties addressing the rising demand for education from a school-age population that grows

[1] The Sustainable Development Goal 4 aims to "ensure that all girls and boys complete free, equitable and quality primary and secondary education, leading to relevant and effective learning outcomes" by 2030.

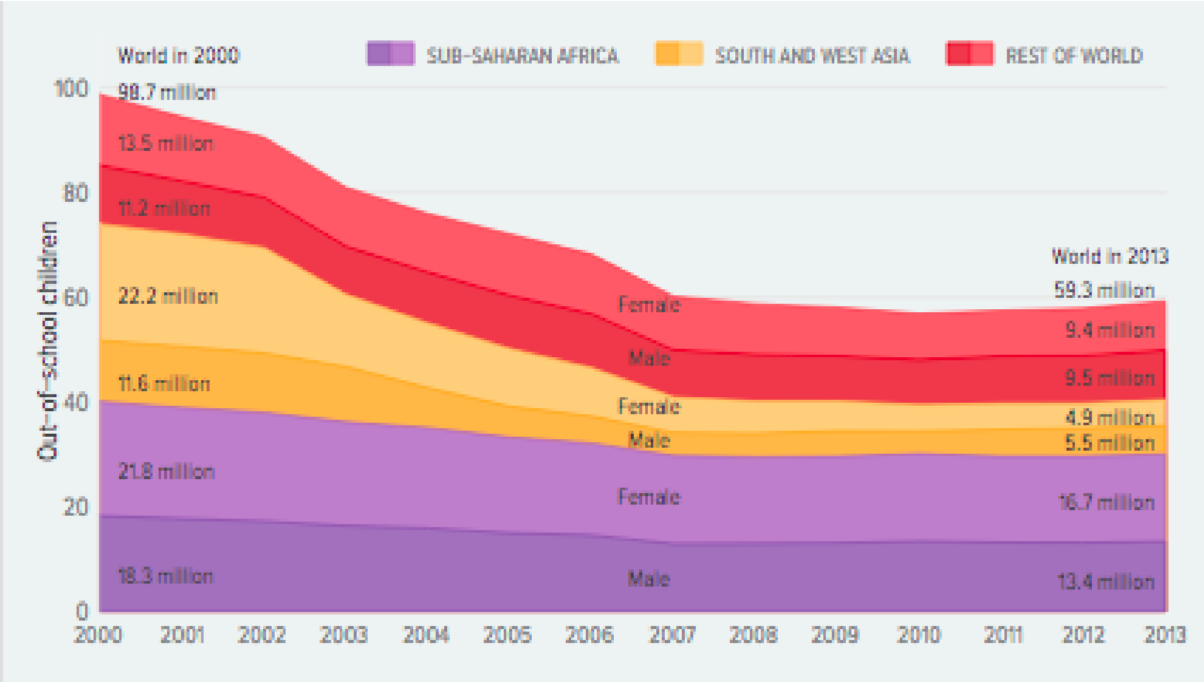



**Figure 1** – Number of out-of-school children of primary school age, 2000-2013.
Source: UNESCO Institute for Statistics database.

continuously. At the same time, the tremendous progress seen during the last century was mainly due to large-scale measures to improve access to education, such as the abolition of tuition fees and the construction of new schools.

These issues are even more relevant today when additional challenges, such as climate change, jeopardize the progress made so far in particular regions. Recently, the international community aknowledged the vulnerability of developing countries to climate shocks as a determinant preventing the realization of the Millennium Development Goals (Kreft, Harmeling, Bals, Zacher and van de Sand, 2010), such as universal primary education or gender equality and women's empowerment. In Southern and Eastern Africa, more recurrent droughts and heavy rainfall affect the region 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, which could be damaging to the child's educational attainment.

In this context, targeted interventions are needed to reach the most marginalised children and youngsters who are out of school today. These children are often from the most socially marginalized communities, children living in extreme poverty or children excluded because

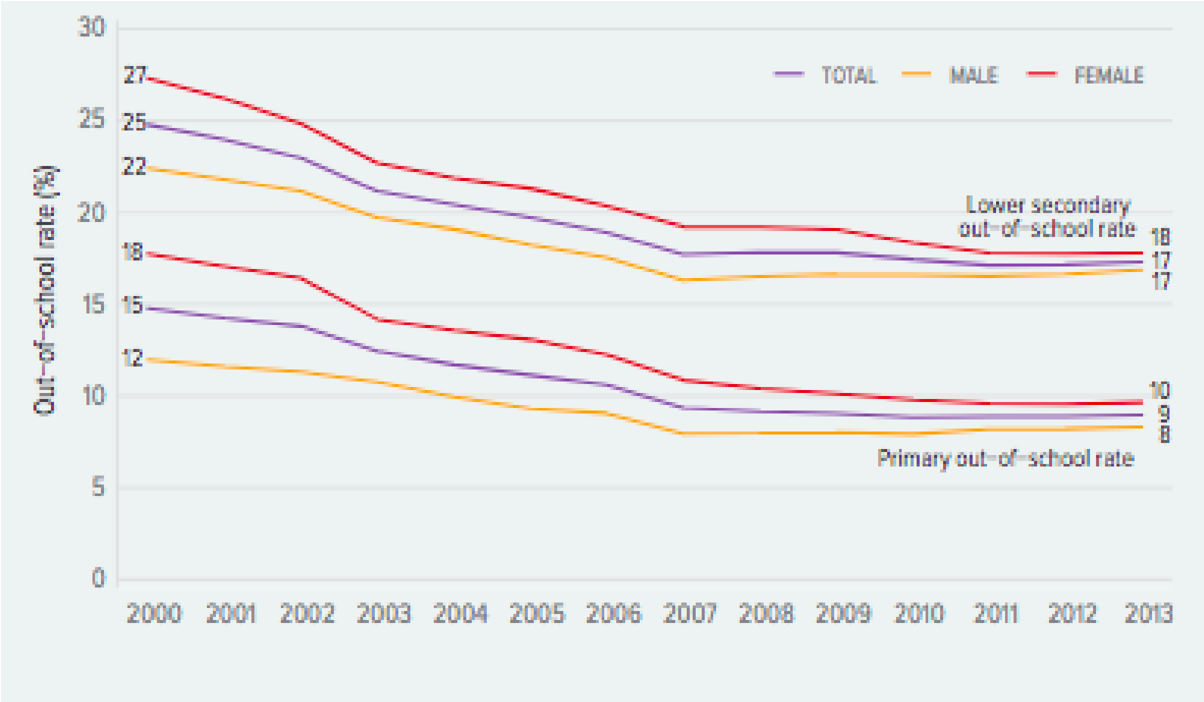



**Figure 2** – Global out-of-school rate, children of primary and lower secondary age, 2000-2013. Source: UNESCO Institute for Statistics database.

of gender barriers. Indeed, girls continue to face the greatest barriers. Their disadvantage starts early and becomes greater during their teenage years mainly due to child marriage and adolescent pregnancy. Approximately 39% of girls in sub-Saharan Africa are married before the age of 18 with men that are on average 4.8 years older. This large age gap between spouses is associated with difficulty for women to exercise their decision-making power within the household, domestic violence, early widowood: resulting in greater inequality. In addition, gender inequality in education has potential negative impacts on a wide range of other development outcomes, not only for the girls themselves, but also for their families, communities and societies.

To explore these concerns, this thesis investigates how families' educational choices are strategically made in a risky environment. The research deals with the gender gap in education and analyzes the choices within the family which contribute to its emergence. A microeconomic approach allows me to identify the evolution of family educational choices. The goal is to provide additional empirical evidence on the effect of family characteristics on human capital investments, to highlight new findings.

From the point of view of economic development and according to Behrman (2010), education is "the acquisition of knowledge and skills through experiences from conception onwards over the life cycle that increase productivity". The first stage of the life-cycle is preschool education and physical development, from conception through about age 5 or 6, determined by nutrition, health or infectious diseases provided in the home, neighborhood, and institutions. Then, education during the school-age years (from age 6 or 7) is determined by formal schooling but also by out-of-school experiences ranging from homework to labor market work, all conditional on stage 1 outcomes. The three chapters of this thesis focus on factors that drive the demand for education of school-aged children in this second-stage of the educational process. Finally, education during postschool is determined in part by formal programs such as training programs, experiences in labor markets, household production, and other activities.

Over the last two decades, researchers have done extensive work on the state of education in developing countries. Household and community resources have been identified as one of the main determinants of the demand for education.[2] Decisions on whether to send children to school are based on comparing the costs and benefits of doing so. In most cases, parents are not basically reluctant to send their children to school provided affordable, effective and safe schooling opportunities. However, there are several reasons why household decisions may not be socially optimal in developing countries. The difficulty of getting to school and the cost of schooling prevent the most vulnerable families to send their children to school. Excluding school fees, many expenditures still remain to be paid like uniforms, exercise books, lunch expenses, tutoring. Parents are often forced to pay for additional tutoring to help their children to pass examinations in primary and secondary school, due to a poor quality education. As pointed out by Epstein and Yuthas (2012), "in particular places, opportunity costs may be large since children have to produce income working on the family

[2] See Glewwe and Muralidharan (2016) for a review of recent studies and particular education policies.

farm or selling in the marketplace." Besides, families may decide to underinvest in children's education if they expect that they will not be able to obtain the returns for this education.

The latter argument is documented in most empirical studies on gender inequalities in education. Indeed, in a patrilocal society where daughters tend to leave the natal family upon marriage, the expected private returns perceived by the parents might be lower for girls than for boys since parents rely on their sons to look after them in old-age (Strauss and Thomas, 1995, Becker, Murphy and Spenkuch, 2016). Also, differences in schooling costs between girls and boys are explaining gender disparities in poor countries. As girls are considered as more productive than sons in home production (Björkman-Nyqvist, 2013, Edmonds, 2006, Dammert, 2010), have lower labor-market opportunities (Butcher and Case, 1994, Strauss and Thomas, 1995) and marriage represents an alternative educational choice for girls in societies engaged in bride price payments (Corno and Voena, 2016, Corno, Hildebrandt and Voena, 2017), girls face a higher opportunity cost in terms of foregone earnings.

Many questions related to education inputs and incentives remain hanging in the literature to date. Some examples include the following: How educational choices are made within the family in a risky environment? Does investment strategies differ by child's gender? What accounts for differences in educational outcomes for boys and girls? In particular, the literature abstracts from the possibly important roles of intra-household allocations and the role of heterogeneous endowments. Most of the analytical modeling and most of the empirical work has been within a one-period framework with consensus parental preferences. The nonconsensus models of household behavior emphasize that different household members, usually husband and wife, have different preferences and different command over resources (Browning, Chiappori and Weiss, 2014). Similarly, different roles and endowments of household members, including children, sets a rational ground for intra-household inequalities. The consideration of the heterogeneity of situations children face should be taken into account in the design and evaluation of policy programs. As a result, this thesis addresses the need to better understand this channel, to identify who are the more disadvantaged members in order to create adequate opportunities for them. However, understanding the nature of and the impact of intrahousehold allocations on children's educationl through empirical data is a relevant challenge.

The three chapters of this analysis use empirical methods to discuss those aspects of education and gender inequality in developing countries. Chapter 1 serves as an introductory chapter and sets the stage for the analysis of family educational choices. The study introduces the broad determinants of children's human capital with a focus on family backgrounds such as mothers' endowments. Specifically, the chapter investigates the relationship between women's economic rights and children's education in several developing countries. In addition to shedding some light on the mechanisms of the transmission of family resources, this exploration highlights the existence of heterogeneous behaviours and preferences within families. Accordingly, the chapter which follow provides an empirical application in a rural context and think about the role of heterogeneous endowments across family members. Chapter 2 revisits the link between income shocks and educational achievement by considering the effect of sibling composition in a rural region of Tanzania. The paper devotes a particular attention to whether children with relatively more sisters than brothers are better ensured during income shocks and whether the effects are different between girls and boys. Chapter 3 gives additional insights on the nature of interhousehold relations among relatives, including parents and children. A large body of the literature on education assumes that parents' incentives for children's human capital are mainly related to transfers for old-age support in the absence of savings and insurance. The old-age security hypothesis often explains any differential treatment in the allocation of educational resources across gender. To test the latter assumption, Chapter 3 explores how a shock on education costs affects boys' and girls' education and the resulting old-age support transfers from offspring to elderly parents in Indonesia.

The following addresses more precisely the contents of each chapter of the thesis.

## Chapter One

Chapter 1 sets the stage for the analysis by offering a reflection on education and gender. The chapter presents a critical review of the concept under study, its operationalization and the exploration of its main determinants. The study is based on a joint work with Christophe Muller (Aix-Marseille School of Economics).

We start from the fact that allocating resources to mothers is often beneficial to economic and social development. Following most studies on the subject including Doepke, Tertilt and Voena (2012), we begin by considering that: (a) the resources controlled by the household head have a major impact on the realized share of expenses that he/she prefers; (b) women are more concerned about children's human capital. However, there is no global evidence on these issues across countries. The following question then naturally follows: Is there a general positive impact of women's economic opportunity rights on children's schooling in developing countries? Part of a macroeconomic perspective, the study examines the relationship between women's economic rights and children's human capital in developing countries.

We propose a refined measure of women's empowerment. In particular, we focus on rights reflecting women's agency, or how much autonomy women enjoy to act on their own in the economic and family environment, especially the access to a bank account or to a job without the husband's permission, the possibility to go to court, to sign legal contracts on their own and to be legally designated as the household head. We assess the variation in women's legal rights both across countries and over time, by using the "50 Years of Women's Legal Rights Database" (The World Bank, 2012). We also identify main drivers of education outcomes in developing countries through an extensive literature review, and convert microeconomic measures of schooling costs and returns into aggregate variables.

Finally, this chapter attempts to correct for the endogeneity of lagged women's opportunity rights to goes beyond the reported correlations. However, the analysis is limited to the choice of available instruments in a macroeconomic setting. Nevertheless, we develop an intrumental variable strategy, based on internal and external instruments.

We first find that endowing women with equal opportunities than men in justice, on the labour and financial market, and within the family, increases children's average years of schooling in the latest stages of their human capital accumulation, between 15-19 years old. We then allow for a heterogeneous impact of women's economic rights, distinguishing between the educational attainment of boys and girls. Greater opportunity rights allocated to married women improve both girls and boys' human capital, with similar magnitude. The finding suggests that higher returns for girls' schooling on the labor-market is not the only mechanism at stake.

## Chapter Two

The context of the chapter is as follows: income shocks induce changes in households' resources and time allocations, which are crucial determinants of youths' schooling. The literature shows that households respond to income shocks by varying the amount of schooling and resources provided to girls while boys are to a large extent sheltered (Björkman-Nyqvist, 2013). Implicit in many studies is the notion that there might be trade-offs among household members either in long-run investment strategies or in responses to short-run opportunities or shocks. However, when capital and labor markets are imperfect, choice sets narrow, and parents must choose how to ration available funds and time between their children. The literature, though, abstract away heterogenous effects of income shocks within the family, making these impacts hard to predict.

In this end, the second chapter investigates the role of income shocks on children's education, depending on the number of biological sibling and their gender composition in a rural region of Tanzania. Several approaches have been undertaken in identifying exogenous variations in income. In this chapter, I construct a microeconomic panel model exploiting exogenous variation in rainfall across districts as a reduced-form instrument for household income variability. Specifically, measures of exposure to climate shocks at a localized level estimate the relative effect of exposure to (unusual) agricultural conditions on the probability of secondary school dropout. I then allow for an heterogeneous impact of income shocks, distinguishing between children with relatively more (younger) brothers or sisters. In addition, special attention is devoted to the fact that the number and sex composition of siblings is not completely exogenous, as it depends on the realization of parental fertility decisions.

I first find that children with sisters do particularly worst during economic hardship. In contrast, abnormal climate conditions during the growing seasons decrease the likelihood of school dropout for children with only brothers. The results are mainly driven by the gender composition of the younger siblings and the magnitude of the heterogeneous effects are larger for girls than for boys. Boys increase their total work, including household work, and constitute a reliable asset for families affected by income shortfalls. The short-term effects have long-lasting consequences for selected children, since climate shocks result in a substantial loss of human capital on average.

Globally, the findings indicate that parents adopt strategies decreasing human capital and worsening gender gap within household when the economic environment becomes uncertain. Heterogeneity impacts should thus be taken into account in the design and evaluation of policy programs.

## Chapter Three

This chapter is a joint work with a former PhD candidate, Pauline Morault (University of Cergy-Pontoise). This chapter engages a broader idea on the motives for children's education: parents generally choose to invest in their future in the form of children as children potentially become a source of income for old age, with limited markets for credit and insurance (Becker, Murphy and Spenkuch, 2016). However, factors affecting the old-age transfer such as adult-children's time allocated to the labor market, the quality of the spouse and the control over family resources, are endogenously determined with their education. Indeed, parents might anticipate the consequences of their educational decisions on monetary support and assistance that they expect to receive during old-age. From a methodological perspective, studies on the intergenerational parent-child exchange should account for the endogeneity of children's human capital.

In this paper, we exploit the existence of a quasi-natural experiment in Indonesia to study these issues empirically and to overcome the endogeneity issue inherent to the research question. We rely on the fact that changes in children's educational attainment were greatly influenced by an exogenous shock on education costs: the massive construction of primary schools during the 1970s. Exogenous variations in exposure to the reform across cohorts and space allow us to identify the effect of a decrease in education costs on children's educational attainment. In a second-stage, we use exposure to the reform as an instrumental variable and investigate the causal impact of children's human capital on transfer to elderly parents. More importantly, we allow heterogeneous effects across married sons and married daughters.

Finally, our goal is to better understand how schooling affects the assistance to elderly, by child's gender. To do so, we construct a model which focuses on two main channels. The first one is the impact of education on expected wages, the labor-market returns to education. The second is the impact of education on the quality of the spouse and on the bargaining power within the newly formed household, that we name the marriage-market returns to education. Empirical results are consistent with this model in which both labor- and marriage-markets returns play a role. We find that a decrease in education costs raises children's educational attainment for both males and females. However, contrary to theoretical findings, we do not identify stronger effects associated to a particular gender.

# References

Becker, Gary S, Kevin M Murphy and Jörg 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.

Behrman, Jere R. 2010. Investment in education—Inputs and Incentives. In *Handbook of Development Economics*. Vol. 5 Elsevier pp. 4883–4975.

Björkman-Nyqvist, Martina. 2013. "Income shocks and gender gaps in education: Evidence from Uganda." *Journal of Development Economics* 105:237–253.

Browning, Martin, Pierre-André Chiappori and Yoram Weiss. 2014. *Economics of the Family*. Cambridge University Press.

Butcher, Kristin F. and Anne Case. 1994. "The Effect of Sibling Sex Composition on Women's Education and Earnings." *The Quarterly Journal of Economics* 109(3):531–563.

Corno, Lucia and Alessandra Voena. 2016. "Selling daughters: age of marriage, income shocks and the bride price tradition." IFS Working Paper W1 /08.

Corno, Lucia, Nicole Hildebrandt and Alessandra Voena. 2017. "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments." *NBER Working Paper No. 23604* .

Dammert, Ana C. 2010. "Siblings, child labor, and schooling in Nicaragua and Guatemala." *Journal of Population Economics* 23:199–224.

Doepke, Matthias, Michèle Tertilt and Alessandra Voena. 2012. "The Economics and Politics of Women's Rights." *Annu. Rev. Econ.* 4(1):339–372.

Edmonds, Eric V. 2006. “Child labor and schooling responses to anticipated income in South Africa.” *Journal of Development Economics* 81:386–414.

Epstein, Marc J and Kristi Jane Yuthas. 2012. “Redefining education in the developing world.” *Stanford Social Innovation Review* .

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

Glewwe, Paul and Karthik Muralidharan. 2016. Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*. Vol. 5 Elsevier pp. 653–743.

Kreft, Sönke, Sven Harmeling, Christoph Bals, Winfried Zacher and Klemens van de Sand. 2010. The Millennium Development Goals and climate change: Taking stock and looking ahead. Technical report Bonn: Germanwatch.

Strauss, John and Duncan Thomas. 1995. “Human resources: Empirical modeling of household and family decisions.” *Handbook of development economics* 3:1883–2023.

The World Bank. 2012. Historic Data: 50 years of women’s legal rights. Technical report The World Bank, Washington, DC.

UIS. 2017. “Global Education Monitoring Report.” *UNESCO Institute for Statistics* .

UIS. 2019. UIS Releases New Education Data and SDG 4 Indicators for 2018 School Year. Technical report.

UNESCO. 2015. “A Growing Number of Children and Adolescents Are Out of School as Aid Fails to Meet the Mark.” *Policy paper 21, fact sheet 33* .

# 1 Women's Economic Opportunity Rights and Children's Education in Developing Countries

This paper examines whether the emergence of women's legal rights in developing countries has affected the human capital of the next generation. While previous authors have investigated this question at a micro level for specific reforms and specific countries, there exists almost no global evidence about these issues across countries. We fill this gap by investigating married women's opportunity rights as a determinant of children's education in 75 developing countries from 1970 to 2010. We find evidence that reforms that have improved the legal capacity of married women have contributed to raising educational attainment of the 15-19 years old cohorts.

***Keywords***: *Social Institutions, Gender Equity, Child Education, Developing Countries.*

## 1.1 Introduction

The promotion of gender equality and women's empowerment is one of the Millennium Development Goals that have been universally pursued since 2010. While significant progress towards reaching these goals has already been achieved, the extent of women's rights remains largely insufficient in many developing countries. Women's access to economic resources are still limited. Particularly, female individual economic and social rights are often found to be inferior to those of their male counterparts. According to the 2016 Women, Business and the Law report, 155 of the 173 economies covered have at least one law impeding women's economic opportunities. This concerns almost all countries in the Middle East and North Africa and in Sub-Saharan Africa. In 2010, husbands can still legally oppose to their wives working on the labour market in 18 of these economies. In some countries like Cameroon, Ivory Coast or Congo-Brazzaville, the husband retains exclusive control on matrimonial assets, and local and national financial institutions require the husband's consent to allow his wife to access financial services. Gender inequality in legal rights affects women's ability to save, borrow, pay for expenses or insure themselves against risk. In these conditions, women's decision making is severely limited by these legal restrictions, which badly hamper their economic activities.

In this paper, we investigate from a global cross-country perspective the role of women's economic legal rights in promoting children's education in developing countries. We expand both the country coverage and the time dimension considered in the literature, by using the "50 Years of Women's Legal Rights Database" (The World Bank, 2012). We focus on married women's economic rights measured by the following questionnaire items: access to a bank account or to a job without the husband's permission, the possibility to go to court, to sign legal contracts in their own and to be legally designated as the household head. Taken together, these variables measure women's agency, as they capture how much autonomy women enjoy that would allow them to act on their own in their economic and family environment.

Our empirical strategy is based on estimating dynamic panel data models. While estimating panel regressions partly help us to attenuate endogeneity issues through country fixed effects, we also control for other drivers of education outcomes, by converting microeconomic measures of schooling costs and returns into aggregate variables. The macroeconomic setting has the advantage of ruling out many issues resulting from some unobserved individual heterogeneity and some individual measurement errors as they get smoothed out through the aggregation process. However, macro-level education data questions the causal interpretation of findings. To reinforce the interpretation of the elicited effects, we use an instrumental strategy valid under certain conditions that we fully exposed in Section 4.2. We complement the analysis by the typical System-GMM internal instruments, on the one hand; and with a

Difference-in-Differences analysis, on the other hand. We find that social institutions that improve women's opportunies in their private and social life are associated with more children education. The magnitude of the effect is equivalent over the 40 years studied period to the effect of an GDP p.c. annual growth of more than 30 percent. This result supports policies aimed at alleviating gender disparities in legal rights for development purposes.

Strengthening economic rights for women seems to be a natural policy avenue to improve their economic opportunities, and as a result children schooling if it depends on women's incomes. In particular, access to a banking account or to a job without the husband's permission may help women to increase their economic participation, as it should not only enhance motivations and incentives of these women, but it should also much extend their opportunities. Opportunity rights govern women's access to the labour and financial markets. They also alter their general acquisition paths to productive and economic resources and their capacity to spend them freely, notably on their children. There are other reasons, which are discussed more in details below, why advances in women's legal rights may be beneficial to children education.

Note that we focus exclusively on rights directly improving women's economic opportunities, while we neglect other rights such as the right to acquire or inherit a property. Beyond data issues, this selection is supported in many developing countries by the large gap often observed between the official statement of female property rights and their actual implementation, for example due to social pressure from religion and traditions. As a matter of fact, in many countries, multiple normative regulation systems operate in parallel, and customary or religious laws often prevail on the legal constitution (Htun and Weldon, 2012).[1] Gender discrimination is particularly relevant for married women as habits and social norms affect wives differently from unmarried women. Indeed, the institution of marriage is typically associated with social, economic and physical restrictions for women. In particular, married women face more legal discrimination than unmarried one, as the marriage contract legally

[1] However, the neglected rights may also affect investment incentives, as property rights may allow for collateral and could contribute to financial autonomy. Therefore, we also partly examine their effects in Sub-section 5.2. We find that endowing women with property and inheritance rights equal to those of men has no significant effect. Moreover, the positive effect associated with women's opportunity rights still remains when we include property rights in the estimation.

transfers some of their decision-making capacity to their husbands. For instance, unmarried women do not have to request a man's permission for conducting their economic activities in Indonesia or in Yemen, while wives need this authorization (The World Bank, 2012).

Gender equality matters not only for equity reasons, but also because women specifically contribute to economic development by raising the economic prospects of the next generation (The World Bank, 2011). A growing literature demonstrates the particular benefits of women's access to economic resources, not only for themselves, but also for their families and especially for their children. Indeed, female resources are often found to be associated to greater investment in children human capital. In a cross-country analysis, Branisa, Klasen and Ziegler (2013) demonstrate how gender inequality in social institutions is related to lower female secondary education.[2] An exogenous rise in female income increases children's educational attainment in China (Qian, 2008), while a rise in female assets leads to better child's schooling outcomes in Bangladesh, Indonesia and South Africa (Quisumbing and Maluccio, 2003). Also, women's perception of the division of households assets upon divorce affects children's education in Ethiopia (Kumar and Quisumbing, 2012). Similarly, Geddes, Lueck and Tennyson (2012) in US and Deininger, Goyal and Nagarajan (2013) in India point out the association between women's economic rights and girls' school enrollment.

A few presumed mechanisms can be highlighted through the literature. In Ethiopia, expanding wives' access to marital property and removing restrictions to working outside the home are found to increase the women's share in occupations with higher educational requirements and paid work (Hallward-Driemeier and Gajigo, 2015). Female land rights have been found, in diverse specific countries, to be a positive determinant of household income (Deere, Durán, Mardon and Masterson, 2005) and to reduce fertility (Fernandez, 2014), child mortality (Eswaran, 2002, Field, 2003) and child labor (Field, 2007). Besides, women's share of business assets, savings and farmland have been found to affect household budget shares (Doss, 2005). More importantly for children education, women's asset ownership

[2]Social institutions related to gender inequality is measured with the Social Institutions and Gender Index (SIGI), aggregating deprivation of women in; family code, civil liberties, physical integrity, son preferences and ownership of rights.

has been linked to decrease in household consumption of some male-favored goods[3] and increases in spending on more durable goods, such as children's health or education (Thomas, 1990). Married woman may be more able to exercise their preferences on household spending when they can participate in household decision making. Their bargaining power could be increased with their asset endowment (Wiig, 2013, Field, 2003, Doss, 2013), such as inheritance or land rights (Allendorf, 2007). All this evidence invites us to examine closely the policies likely to enhance mothers' information, female bargaining power within households and women's returns to effort notably through broader access to markets in order to better understand how they may be connected to more efficient human capital accumulation. However, at odds with this literature, Edmonds (2006) shows in South Africa that child schooling can also be enhanced when resources are given to male members instead.[4]

The heterogeneity of the results from these microeconomic studies raises the question of possible special characteristics of local contexts that may partly drive these results. In contrast, a global perspective covering most developing economies will assist us in testing our main hypothesis of interest. That is: Is there a general positive impact of women's economic opportunity rights on children's schooling in developing countries? The aim of this research is to provide a global assessment of this issue for a broad set of developing countries, and thus to fill a gap in this literature. An exception to the specific-country orientation of the literature is the study of Hallward-Driemeier, Hasan and Rusu (2013*b*) which analyses the correlation of married and unmarried legal rights with various socio-economic outcomes in a sample of high and low-income countries. They found that the removal of gender gaps in rights is correlated with greater participation of women in the labor force, lower fertility and lower infant mortality and finally higher female educational enrollment, although without attempting to delineate the causal dimensions of these correlations.

[3]See Anderson and Baland (2002), Bobonis (2009), Brown (2009), Wang (2014), Menon, Rodgers and Nguyen (2014)

[4]Edmonds demonstrated through a social pension program in South Africa that child schooling, mainly boys' attendance to school, increased more in family with an eligible male compared to those with an eligible female. Moreover, in family with joint male and female eligibility, the magnitude of this rise was similar to male eligibility alone. The main reason advanced was the gender differences in access to credit, men being more liquidity constrainted due to higher mortality risk and riskier behavior.

Our second contribution, which goes further than the literature, is to address the issues of identification and endogeneity of the female rights in the determination of children's schooling. Indeed, the findings on consequences and origins of women's rights suggest that causality between economic development and women's rights may run in both directions (Doepke, Tertilt and Voena, 2012). According to Duflo (2012), "women empowerment and economic development are closely related: in one direction, development alone can play a major role in driving down inequality between men and women; in the other direction, empowering women may benefit development. The temporal structure of our model reduces the likelihood of reverse causality without eliminating it completely.[5] For example, gender equality and the human capital of the future generations may be influenced by observed and unobserved common factors, such as international influences[6] or technological change.[7] In another example, Doepke and Tertilt (2009) claim that the high intensity in human capital of modern technologies may incite men to grant additional rights to women in general because men care about their own daughters' education and future income. Thus, along economic development, a growing interest in education by men may foster both women's legal rights and future educational attainment. Therefore, endogeneity of these rights may arise from simultaneity and unobserved common factors, leading to biased estimates. We try to deal with this endogeneity issue and derive the conditions under which the reported correlations could be interpreted as causal evidence.

The remainder of the paper is structured as follows. Section 2 discusses the theoretical relationship between women's empowerment and children education. Section 3 describes the data. Section 4 examines our empirical strategy. Section 5 addresses issues of causality and highlights the main results. Section 6 presents the strategy and the results associated with the difference-in-differences estimations. Finally, Section 7 concludes.

[5]In the estimations, the indicator reflecting women's legal rights is lagged five years.

[6]The Convention on the Elimination of all forms of Discrimination Against Women (CEDAW) is one of the international forces that promote the implementation of women's legal rights as well as girls' access to education in a country.

[7]Technical progress may increase women's returns to education, which may alter cultural perceptions about the economic role of women in society, thereby leading to the implementation of gender equal rights.

## 1.2 Theoretical framework

It seems fair to say that a consensus about a common theoretical framework has not yet been achieved in the literature regarding the determination of schooling, especially at aggregate level. However, Haveman and Wolfe (1995) claim that "the attainment of children depends on three primary factors: (1) the choices made by the society that determine the opportunities available to both children and their parents; (2) the choices made by the parents regarding the quantity and quality of family resources devoted to children; (3) the choices that children make given the investments in and opportunities available to them." Accordingly, as a simplifying analytic framework, we consider that society acts first by reforming specific married women's rights; that is: the economic and institutional environment that households will face when making their own decisions about their children's education. Later on, we will also consider the possibility that female rights are not predetermined.

Let us consider a few potential explanatory channels of enhancement in children's schooling, starting from women's rights. Allowing women to work outside home, to sign contracts or to open a bank account, raises their economic opportunities. Moreover, better women's prospects on the labor market should increase female employment, which in turn raises the marginal value of female time and family income. On the one hand, with additional earnings, parents can spend more resources on children's education if the latter is a normal good. In parallel, a rise in household income due to broader economic opportunities for female members may limit the need for child labor and thereby facilitate an increase in children schooling time.[8] On the other hand, if that husbands and wives make joint labor supply decisions, greater women's employment opportunities may partly substitute to men's employment (Mammen and Paxson, 2000). In some cases, this substitution effect may reduce family income and affect negatively children's education. As well, greater women's opportunities on the labor market should raise the expected benefits from mother's schooling (Jensen, 2010). There are other reasons why mother's human capital can be positively or negatively related to child attainment; positively, as educated mothers value more the edu-

[8] See Edmonds (2007), Ravallion and Wodon (2000), Baland and Robinson (2000).

cation of their children (Behrman, Foster, Rosenweig and Vashishtha, 1999), and negatively, as they spend less time at home educating them (Behrman and Rosenweig, 2002). More mother's education is also associated with better children's health. Healthier children will be more productive, will live longer and thereby the expected return to their education is higher. Moreover, healthier children learn better at school (Alderman, Behrman, Lavy and Menon, 2001).

Besides, women gaining a greater command of resources from new laws improving their job prospects may have greater bargaining power within the household. What matter are a woman's abilities to leave her household as an outside option, which increase her influence on household's decisions. Then, if mothers value children education more than fathers do, an increase in women's bargaining power may result in greater household resources devoted to children education.[9]

A last, but related, channel through which married women's rights reforms may influence children's education is fertility. Given a stronger woman's bargaining power, for example owing to a higher educational level with an increase in her time marginal value for the household, or from more legal rights in our case, a woman may be more prone to reduce her fertility (Currie and Moretti, 2003). In turn, children in smaller families may accumulate more human capital than those in larger families where they have to compete with other siblings for limited family resources. This is consistent with a quantity-quality trade-off, such as in Becker and Lewis (1973).

Clearly, all these mechanisms yield different implications for understanding how women's economic and social rights affect the resources devoted to children within the family, and decisions about children's education. Notably, positive and negative effects of female rights are possible through some of the listed channels, which yields an ambiguous theoretical conclusion. What is now needed is some empirical evidence to establish which of the positive and negative channels are dominant and whether they are globally significant. However, in an empirical context, the above theoretical framework will imply a possible endogeneity of

[9]Related references demonstrating the higher preference of mothers, as compared to fathers, for children's education can be found in the introduction.

**Table 1.1** – Summary statistics

| Variables | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Education (15-19 age group) | 657 | 5.74 | 2.44 | 0.23 | 12.07 |
| Opportunity rights index | 611 | 4.02 | 1.43 | 0.00 | 5.00 |
| GDP per capita* | 605 | 6.96 | 1.01 | 4.49 | 9.22 |
| Democratic index | 629 | -0.37 | 6.46 | -10.00 | 10.00 |
| Child Mortality | 648 | 106.99 | 74.52 | 8.00 | 395.23 |
| Fertility rate | 657 | 4.80 | 1.77 | 1.11 | 8.87 |
| Adult education (above 25 years old) | 657 | 4.31 | 2.60 | 0.00 | 11.74 |
| Women's education | 657 | 3.74 | 2.77 | 0.00 | 11.81 |
| Urban population | 657 | 39.91 | 20.31 | 3.34 | 90.79 |
| Education spending (% of GDP)* | 412 | 4.13 | 2.73 | 0.81 | 44.33 |
| Agriculture value-added (% of GDP) | 570 | 23.75 | 14.16 | 2.47 | 68.54 |
| Share of female in labor market* | 365 | 3.90 | 0.47 | 2.27 | 4.50 |
| Opportunity rights index in host countries | 721 | 0.03 | 0.05 | 0.01 | 0.53 |

* In logs.

the rights through a variety of unobserved confounding factors. For example, a tendency to modernity of a given country in the medium term may affect both gender equality in legal rights and human capital accumulation. Dealing with endogeneity biases in the estimations will be at the core of our empirical strategy. Before discussing this part, we present the data used in the empirical analysis.

## 1.3 Data

The data we use are extracted from aggregate databases at country level. There are four kinds of variables: (1) Educational outcomes, our dependent variables of interest; (2) Women's rights, our variable of interest; (3) Control variables, and, finally, (4) Instrumental variables, which we discuss further in sub-section 4.2. The data covers 75 countries and 40 years. Table 1.1 reports a few descriptive statistics, and the list of the covered countries is presented in the Appendix in Table A.1.

### 1.3.1 Education

A major pratical issue is to avail of educational outcome measures that are consistent across countries. Data on enrolment rates are widely available, but these data, which measure the current flows of schooling, do not adequately measure the aggregate stock of human capital available (Barro and Lee, 1993). Moreover, enrolment rates refer to the registered number of students at the beginning of each school year, which may be substantially higher than the actual number of children that attend school during the year. Educational attainment, based on the accumulation of the education flows, is at best a proxy for the component of the human capital stock obtained through schooling. However, higher attainment rates should reflect the general higher educational level in the considered country.

In this study, we measure children's attainment with the average years of schooling of the cohort of individuals aged between 15 and 19 years old. This is the youngest cohort available in the Barro and Lee (2013)'s database. The focus on this age group is justified by its time location just beyond the compulsory top school age in most developing countries. According to the Unesco Institute for Statistics, the duration of compulsory education is on average 7 years in low income countries in 2000, 8 years in 2010, while it remains close to 9 years in middle income countries across the period. See Table A.2 for more details. Also, the outcome allows effects both on primary and secondary education. Indeed, given that five-year averages are used for all variables, the cohort is aged 10-15 years in time t-1 (see details in section 4). If we assume that resources available to females affect mainly primary education in a contemporaneous way, a change in the women's opportunity rights index in time t-1 would result in significant effects for the cohort aged 15-19 years old in time t. On the contrary, if we think that a change in women's endowments impacts education with some lags, in this case, our empirical strategy reflects changes in inputs and incentives among secondary-school aged children (aged 15-19 years old).

In Barro and Lee (2013), the average years of schooling of the 15-19 age group are constructed from census/survey figures compiled by UNESCO, Eurostats and others, which report the distribution of educational attainment of each cohort. Then, assuming that changes in enrollment lead to corresponding changes in attainment with time lags, these authors extrapolate the missing attainment data by using the available estimates for the same age group in the previous period adjusted by the changes in enrollment for the corresponding age groups. The average number of years of schooling of the 15-19 age group may be expressed as:

$$S_{i,t} = \sum_{j} h_{j,i,t}.DUR_{j,i,t} \tag{1.1}$$

where $h_{j,i,t}$ indicates the frequency of educational attainment $j$ for the 15-19 age group $j$ in country $i$ and year $t$, and $DUR_{i,t}$ denotes the corresponding mean number of schooling years to attain attainment $j$ in country $i$. Information for each country on changes in the duration of primary school and of the two levels of secondary education over time comes from UNESCO data (Statistical Yearbook). For higher education, the missing observations are completed by assuming a four years duration for all countries and all years. Thus, aggregating the duration and the distribution of educational attainments for the 15-19 age group yields the average years of schooling in the 15-19 cohort, in each country. In the sample of the 75 developing countries, the average years of schooling of the 15-19 age group is around 6 years, with a minimum of 2 months and a maximum of 12 years. By region, Europe-Asia and Latin America stand above the overall mean with respectively 7.7 and 6.8 years of education, while Sub-Saharan Africa is far behind with 4.4 years.

### 1.3.2 Women's rights

The legal rights variables are taken from the "50 years of women's legal rights" database (The World Bank, 2012). These data provide, for each country and each covered year, information on gender gaps in terms of matrimonial rules and on women's ability to perform activities independently from their husband. Based on their published legislation, countries receive the value of zero if the considered right is not equal for men and for women, and a value of one if women and men enjoy the same right equally. However, these indicators based on the law as it is included in the country's constitution may imperfectly describe the actual context. For example, traditional and informal family usages, and lack of awareness of the law by individuals, may prevent women from opening a bank account, although the legal rights are formally enforced in the country.

In this study, we use the responses to the following questions to construct a synthetic index of female rights: (1) *Can a married woman legally get a job or pursue a trade or profession in the same way as a married man?* The question assesses whether married women need permission from her husband to access a job, can be forced to leave their job through court if working contrary to her husband's wishes or to the interest of the family, which may be considered as a form of disobedience; (2) *Can a married woman legally open a bank account in the same way as a married man?* Specific provisions may limit the capacity of a married woman to open a bank account in her own name, such as the need of permission from her husband or additional documentation; (3) *Can a married woman legally be the "head of household" in the same way as a married man?* The answer is negative if there is an explicit provision stating that only husbands can be so designated since they represent the family or as they are the default family member who receives the family document that is necessary for access to services; (4) *Can a woman legally sign a contract in the same way as a married man?* The question investigates if a woman has a limited legal capacity to sign a binding contract or needs the signature, consent or permission of her husband to legally bind herself. This question does not concern restrictions on signing contracts specifically related to marital property; (5) *Can a woman legally initiate proceedings in the same way as a married man?* The latter question refers to the possibility to engage in justice without any consent of a male family member.

Free opportunities to apply for jobs and to open a bank account allow women to earn and own money amounts that may be used or invested according to their own wishes. The right to enter into contracts in their own name enables women to start a new business for instance, and to protect their earnings and their property. Access to justice facilitates their defense against gender-based violence and discrimination in all spheres of their family and economic life. Finally, the right to be legally considered as the family head alleviates power imbalances with the husband and differences in decision making. This right may for instance

**Table 1.2** – Distribution of the women's opportunity rights index

| | Overall | | Between | | Within |
|---|---|---|---|---|---|
| Values | Freq. | Percent. | Freq. | Percent. | Percent. |
| 0 | 33 | 5.40 | 7 | 9.72 | 58.73 |
| 1 | 16 | 2.62 | 3 | 4.17 | 62.04 |
| 2 | 36 | 5.89 | 9 | 12.50 | 53.09 |
| 3 | 83 | 13.58 | 15 | 20.83 | 69.66 |
| 4 | 90 | 14.73 | 20 | 27.78 | 52.76 |
| 5 | 353 | 57.77 | 50 | 69.44 | 80.50 |

prevent guardianship system in which an adult female living without a husband must submit to be legally recognized.

Using dummy variables for each of these specific rights yields too little dispersion for useful econometric analysis. As a response to this difficulty we aggregate all these specific right indicators by summing them into an index that ranges between 0 and 5, and reflects general opportunity rights. This composite index captures more variations in reforms than any specific right index. Table 1.2 shows the distribution of the constructed index. More than 57 percent of the observations reach the maximum, while only few observations (5.4 percent) are equal to 0, the minimum value. The between dimension reveals that 7 countries ever had Opportunity index equals 0 and 50 ever had Opportunity index equals 5, for a total of 104 countries with variations in women's rights across the period. Moreover, the within dimension presented in the last three columns tells us that conditional on a country ever giving no rights to women, 59% of its observations have Opportunity rights equal 0. Similarly, 80% of observations of a particular country which ever provided full rights to women remain equal to the higher level of gender equality in rights across the period. This fact is comforted by Figure 1.1 that shows that most rights have evolved in parallel to a common trend, which supports our strategy to use a composite index. The right to initiate legal proceedings in justice or to sign contracts in its own name was almost always established at an earlier time than the rights giving women access to financial ressources or to the labour market. The possibility to be recognized as the head of family irrespective of the gender is lagging far

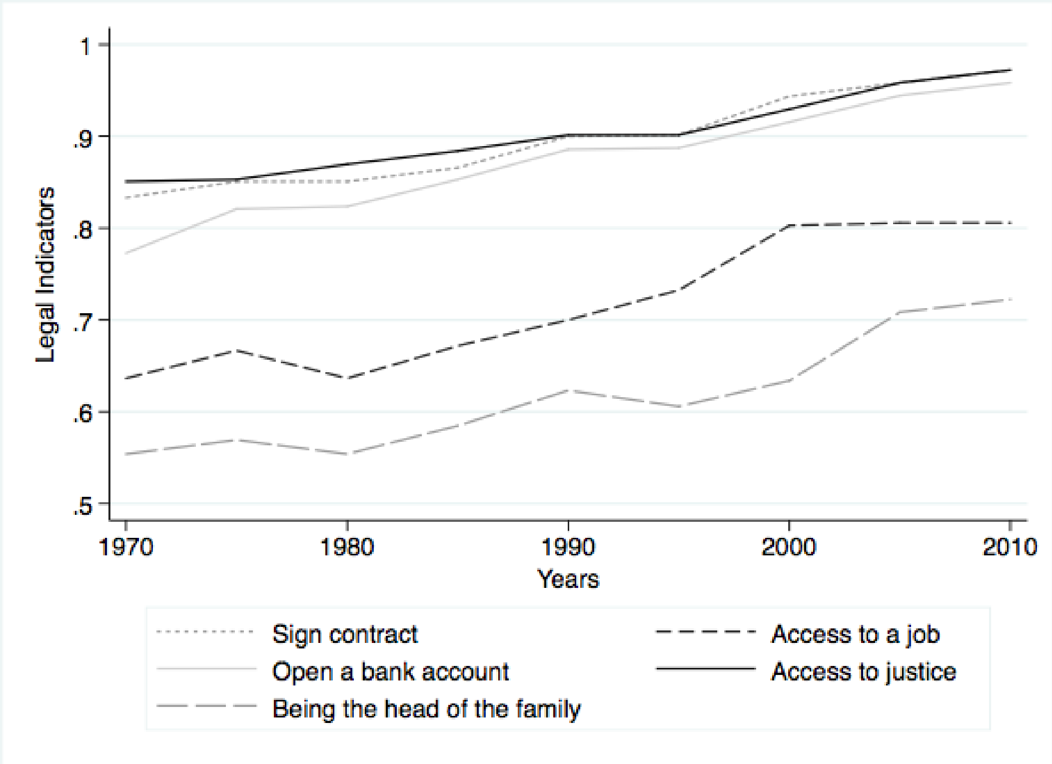



**Figure 1.1** – Evolution of women's individual rights between 1970 and 2010

behind the other rights. Significant improvements in legal rights have been achieved since 1970. Latin America, East Asia and Pacific or Sub-Saharan Africa have reduced the number of legal constraints to women by more than 50 percent (Hallward-Driemeier, Hasan and Rusu, 2013*a*). Despite some general progress, substantial gender disparities in legal rights persist across countries. In 2010, 14 countries still legally prevent women from working outside home without their husband's consent. They are majorily located in Sub-Saharan Africa (Figure 1.2), where delays in children education attainment are also recorded.[10] An obvious positive similar trend over years between the average years of education of the 15-19 age group and the level of women's opportunity rights for the whole sample of countries is displayed in Figure 1.3.

[10]In Bolivia, Cameroon, Congo Democratic Republic, Ivory Coast, Gabon, Iran, Jordan, Mali, Mauritania, Niger, Sudan, Syria, Togo and Yemen.

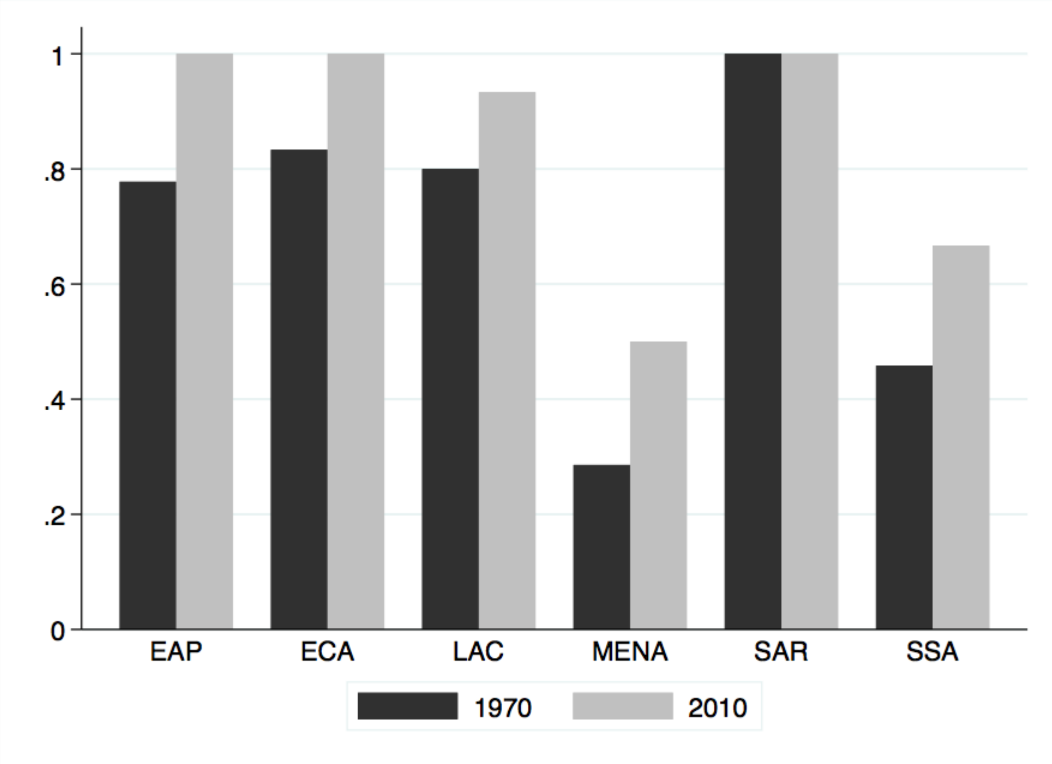



**Figure 1.2** – Evolution of the right to work without the husband's permission from 1970 to 2010, by region

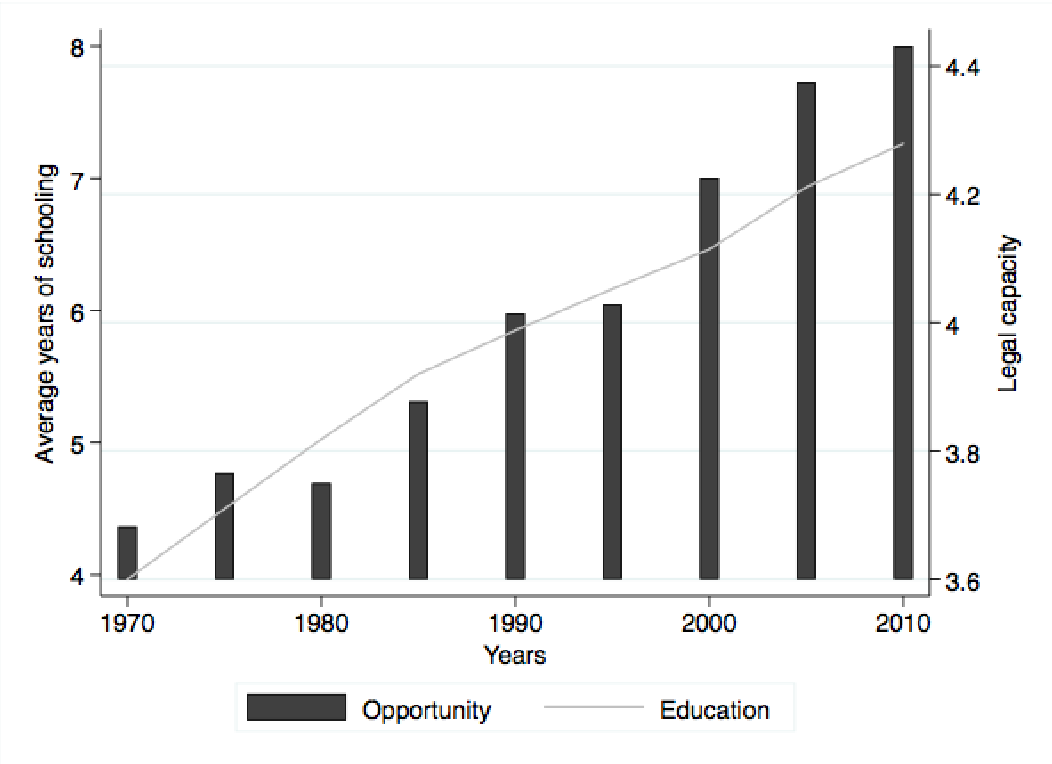



**Figure 1.3** – Correlation between women's opportunity rights and children's education across time

### 1.3.3 Controls

According to Behrman (2010), educational outcomes in a microeconomic setting can be seen as determined by a combination of household demand decisions and educational supply provider decisions. However, when modelling the education process at a country level instead of a micro perspective, one must also account for institutional and economic conditions that alter the costs and benefits of individual investment in human capital.

In that respect, in our empirical equation determining the education level of the 15-19 years old cohort in each country, we first control for the logarithm of GDP per capita, in constant prices (US$ PPP, in 2005), which measures the general level of living standards that is likely to be correlated to several of these costs and benefits. In particular, it is generally believed that children's average years of schooling are much determined by the expected returns in human capital investment in the country. In most contexts, many jobs available on the labor market in developed countries require advanced skills, which directly creates incentives for private investments in education. The GDP per capita, which reflects the level of economic development, can partly be seen as a proxy for the human capital needs in a country's labor market.[11] Moreover, costly children's education is constrained by the resources available to the families. Thus, GDP per capita help us to account for the general level of these resources at a macroeconomic level. Higher per capita income should be associated with higher aggregate demand for education, not only owing to higher returns to education, on the relaxation of liquidity constraints, but also to the fact that education is a normal good.[12]

Incorporating child mortality, defined as the number of deaths per 1,000 live births below five years, controls for general health status. This is because healthier children may be better to develop their capacity and be more able to study. They will work longer as they will live longer, which further increases the expected return of investment in their education.[13]

[11] In the robustness Sub-Section 5.2, we account for the country specialization in the industry/services sector, by controlling for the agriculture value added, which is correlated with the availability of high-skill occupations.

[12] See Bils and Klenow (2000).

[13] See Lorentzen, McMillan and Wacziarg (2008), Cervellati and Sunde (2005, 2013).

We include the mean educational level of adult people, aged 25 and above, to proxy parents' education at country level, since first, parents' schooling is correlated with family income and particularly their own individual wages. In addition, parent's education may reflect parents' and especially mother's access to information, and preferences for higher child quality. Finally, there are directly externalities of parents' education on the efficiency of children's learning process. Thus, returning to the aggregate level, higher aggregate parents' education may raise the demand for education in the aggregate children's education equation. In the same line, we replace the average years of schooling of adult people with only adult female's schooling in Sub-Section 5.2. This may account for the within-household female bargaining position, which may indicate how resources are allocated to children's education. Unfortunately, aggregate adult education and aggregate female adult education are too correlated to be included together. The effect of the female labor force participation rate is also tested in Sub-Section 5.2. by including the share of women in the labor market as a control variable, since it may be a potential channel of women's rights on child schooling.

A polity index, constructed from the Polity IV database (Marshall, Gurr and Jaggers, 2017) is included to control for any potential effect coming from the supply side. Indeed, very unequal societies, which tend to be politically and socially unstable, tend to invest less in the educational system than more democratic countries (Perotti, 1996). One reason may be that autocratic regimes have interest in limiting the human capital of their citizens so as to maintain current political status quo and to protect the regime against newly arising claims and social unrest. Also, a democratic country may be less exposed to civil wars, which may have a direct negative impact on the education infrastructure and the delivery of schooling services over time. The indicator is computed by substracting the institutionalized autocracy score from the democracy score. The latter variables are constructed adding, with different weights, codings of three elements: the competitiveness of political participation, the openness and competitiveness of executive recruitment, and the constraints on the chief executive. In the sample of countries and years, the polity index ranges from -10 to +10, where a higher score means that country $i$ at date $t$ is more democratic. In the robustness Sub-Section 5.2, we include the urbanization rate as it also associated with the level of education infrastructure and services. This rate is likely to be correlated with the availability of schools, teachers, school supplies which proxy some components of the supply side of the education system. Changes in the quality and quantity of educational supply services are also related to government expenditure on education, which is also tested in Sub-Section 5.2. The sign of the effect of public spending on educational outcomes has been found to be ambiguous and to depend on institutional quality.[14]

Finally, the inclusion of the fertility rate in an additional specification accounts for the possibility of Beckers' quantity-quality tradeoff for children. In particular, new births in a budget-constrained family directly reduces the resources devoted to each child, if we assume that parents treat equally their children. With more limited resources, children may reach a lower level of education. Moreover, the presence of numerous children reduces the mother's time allocated to each child at home, for educating them for instance, which should have a negative effect on their level of attainment. Further, parents' marginal rates of substitution of children quantity and quality change with higher income, inciting them to favour higher human capital for their children and value less a large offspring. In that case, parents with few children should have a higher valuation of their children education. Finally, from a global perspective, including the fertility rate may also control for the average number of pupils in classes, which has a negative impact on their productivity and ability to learn, and thereby reduces their educational attainment.[15]

All these variables, which correspond to attemps to capture the potential transmission mechanisms described in Section 2, are included in our regressions to control for the above alternative explanations of education levels. However, some mechanisms are difficult to capture at a country level. For example, it is difficult to measure correctly women's bargaining power within their own household, since bargaining power is fundamentally unobservable (Doss, 2013). We now turn to our econometric strategy to confront the diverse explanations with the data.

[14] See Gregorio and Lee (2002), Rajkumar and Swaroop (2008), Baldacci, Clements, Gupta and Cui (2008), Bennell (2002).

[15] Following Aghion, Persson and Rouzet (2012), since education takes the share of people in the relevant age-group into account, here the 15-19 age group, we do not control for population growth in these specifications.

## 1.4 Econometric strategy

### 1.4.1 Specification of the education equation

Because the dependent variable is constructed from Barro and Lee (2013)'s data that is made of five-year averages for many education variables, and in order to reduce measurement error, five-year averages are used for all variables. In each regression, we lagged the opportunity laws index by five years. Indeed, most reforms affecting female rights are little likely to produce immediately powerful effects. It takes time for individuals to incorporate the information about these reforms and effectively change their behavior. Lagging this index has also the advantage of alleviating some potential simultaneity issues between the education and rights variables. As a result of this discussion, we estimate the following equation:

$$Schooling_{i,t} = \alpha_0 + \alpha_1 Opportunity_{i,t-1} + \alpha_2 X_{i,t} + \nu_i + \chi_t + \epsilon_{i,t} \tag{1.2}$$

where $\alpha_0$, $\alpha_1$, $\alpha_2$ are vectors of parameters to estimate and $\nu_i$, $\chi_t$, and $\epsilon_{i,t}$ are error terms. The dependent variable $Schooling_{i,t}$ refers to the average years of schooling of the 15-19 years old individuals in country $i$ and period $t$. The $Opportunity_{i,t-1}$ term describes the lagged level of women's opportunity rights. Our main coefficient of interest is $\alpha_1$ that captures the effect of the changes in these legal rights in country $i$ and period $t$. To save on degrees of freedom $X_{i,t}$ includes only the control variables which have been found to have significant effects and do not suffer from colinearity and missing data issues. These variables are: the logarithm of GDP per capita, the index of democracy, the average years of schooling of people aged over 25 and the child mortality rate. Urbanization, the female labor force participation rate and the logarithm of government expenditures in education are tested only in the robustness Sub-Section 5.2, since they have may missing values and including them reduces the sample. The agriculture value-added, which is almost colinear with the logarithm of GDP per capita, will also be included separately. The specification entails country ($\nu_i$) and year ($\chi_t$) fixed effects. $\epsilon_{i,t}$ is a centered idiosyncratic error term for which appropriate semi-parametric restrictions are imposed to guarantee the consistency of the used estimators.

In particular, this error is first assumed to satisfy strict exogeneity restrictions, or at least predeterminedness. In that case, we can use the within estimator in order to deal with the unobserved heterogeneity.

One challenge is that several regressors in the education equation are likely to be endogenous, as they may be determined simultaneously, may be influenced by education levels or may be affected by the same unobserved factors than education. In our setting, this would be the case for example if the accumulation of education across generations contributes to explain reductions in mortality. To assess the potential magnitude of these endogeneity issues, we will present the results without any control variables, except for the log GDP per capita, and then present separately the results with the inclusion of these controls for comparison. Besides, special attention will be devoted to the fertility rate. Indeed, a tradeoff between “quality" and “quantity" of children (Becker and Lewis, 1973) implies that children schooling and fertility decisions are jointly determined. Therefore, the fertility rate could not be included in the education equation, without some apprehension for reverse causality and simultaneity that would bias the estimation. Thus, we decide to withdraw this variable from the baseline specification and rather test its inclusion separately. Moreover, endowing women with rights to access the labor and financial markets may impact women's decisions regarding their own fertility. To investigate this potential channel, we will introduce the fertility rate as the dependent variable in another equation, while controlling for the same covariates used in the education equation.

Finally, we estimate Equation (3), which includes the lagged dependent variable as a regressor. This specification allows us to better capture the dynamics involved in human capital accumulation. Moreover, it helps us to eliminate some of the possible biases that may result from some unobserved simultaneous determinants of educational attainment. The following equation requires GMM estimation:

$$Schooling_{i,t} = \beta_0 + \beta_1 Opportunity_{i,t-1} + \beta_2 X_{i,t} + \beta_3 Schooling_{i,t-1} + \sigma_i + \lambda_t + \upsilon_{i,t} \tag{1.3}$$

where $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$ are vectors of parameters to estimate and $\sigma_i$, $\lambda_t$ and $\upsilon_{i,t}$ are error terms.

### 1.4.2 Dealing with the endogeneity of opportunity rights

As discussed before, there might be concerns regarding the direction of causality as gender equality may be both a cause and a consequence of economic development (Duflo, 2012). More generally, there are three possible sources of endogeneity of the opportunity right index: (1) measurement error, (2) simultaneity and (3) unobserved common factors.

Regarding measurement error, we use aggregate indicators of inequality in legal rights, measured at country level, which reduces the concern for measurement errors associated with individual records. Simultaneity arises when the explanatory variable is correlated with the regression error term because of joint and contemporary mechanisms. In our case, one may worry about legal rights and children's human capital being determined simultaneously through new legislation. For example, legal rights are partly affected by international agreements like the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW). This convention constrains state parties to change or abolish laws, regulations, customs and practices which discriminate against women. However, simultaneously, by ratifying it, state parties commit to advancing women's integration in education. Hence, the dynamic structure of the model may not suffice to remove all simultaneity as both women's legal rights and the human capital of future generations may be impacted in a complex way by common international influences.

As a generalisation of the previous argument, women's economic rights may be correlated with some omitted variables that affect their children's human capital accumulation, including dynamically, for example a tendency of the society toward modernity in the medium term. Modernity may include technological progress. Indeed, if technology is such that human capital is irrelevant, men may prefer patriarchy because they would have more power (Doepke and Tertilt, 2009). In contrast, in contexts with high returns to education, men may care more about the education of their offspring, including their daughters, and support women's rights. As a consequence, skill-biased technology change from a macroeconomic perspective may drive both future human capital and the expansion of women's rights. In the same way, information and communication technology (ITC) improves student's access to skills, which should have a direct impact on future educational attainment, on the one hand. On the other hand, ICT facilitates access to discussions and networks of peers, which may help women to challenge patriarchy. Finally, obvious endogeneity issues may arise from the presence of the lagged dependent variable in the model, when this is the case. Fixed effects allow us to deal with fixed omitted characteristics at the country level, while they may not be enough to purge the endogeneity issues.

As a result of this discussion, we try to adress these endogeneity issues and estimate two IV-type estimators (2SLS and System-GMM). The instrumentation strategy is based on a combination of external and internal instruments. For the 2SLS-FE estimator, the index of women's opportunity rights is instrumented with the ten-years lagged level of the rights index and with an external instrument, based on the expected dissemination of women's opportunity rights from each country's emigrants.

A number of attempts was devoted to find a good external instrument.[16] Some attention has been devoted to the Convention on the Elimination of all Forms of Discrimination Against Women (CEDAW), as a source of instruments in the macroeconomic literature on women's empowerment (Ferrant and Kolev, 2016). Indeed, the ratification of CEDAW encouraged states to review family laws (Simmons, 2000). By ratifying the CEDAW, state parties commit themselves to take appropriate decisions ending discrimination against women and girls in all spheres of their private, social and economic life, including measures aiming at eliminating gender gaps in education. Note that the latter claim may somewhat threaten the possibility to use the ratification of the CEDAW as a valid instrument in education equations. In addition, religious affiliation, civil liberties (Dollar and Gatti, 1999) or freedom

[16] Neumayer and De Soysa (2011) or Juhn, Ujhelyi and Villegas-Sanchez (2013) addressed the question of whether trade and investment linkages can diffuse the empowerment of women via spatial dependence. They study the effect of general openness to trade and foreign direct investment (FDI), understood as the extent of a country's integration into the global economy. These variables would have been promising instruments for women's economic and social rights in developing countries, except that international trade and general human capital accumulation may reinforce each other (Kim and Kim, 2000, Borensztein, de Gregorio and Lee, 1998). Therefore, we decided not to proceed with them.

of the press were tested, but these indicators may be correlated with omitted variables. For instance, an unobservable trend to modernity over time may bring shifts in norms and society values and promotes gender egalitarian attitudes as well (Inglehart and Norris, 2003). Finally, the country's legal origin (Acemoglu and Johnson, 2005) or the rule of law (Dollar and Gatti, 1999) were impossible to use as IV along with the presence of country and time fixed effects.

Then, we turned to the average level of women's rights in migrant-receiving countries as a promising candidate for an instrument. This supposes that the following criteria are met: (i) The decision to migrate is not related to the situation of women's rights in the origin country; (ii) Migrants do not select the country of destination according to its gender equality in social institution; (iii) Migrants are affected by exogenous shocks (social norms and beliefs in the receiving country) who are going to transmit in their origin country.

Regarding condition (i), women's capacity and incentives to migrate are limited in gender inequal societies due to low levels of female education and resources. Indeed, the more (less) educated are more (less) likely to emigrate since international migration is characterized by positive selection (Chiquiar and Hanson, 2005, Grogger and Hanson, 2011). We can reasonably consider that international migration are limited to males in gender inequal countries. Cerrutti and Massey (2001) found that wives generally appear to conform to the profile of tied movers following husbands rather than acting as independent migrants. In addition, women's decisions are constrained by patriarchal norms, whereby men ultimately determine whether and when wives should join them. Accordingly, we can plausibly assume that men, who are positively selected into migration, might be less likely to migrate due to a low level of gender equality in social institutions in the origin country. This argument provides plausible reasons to exclude potential bias arising from the decision to migrate. Condition (ii) naturally arises. Men, who are positively selected into international migration and select destinations independently of the wifes' wishes, are less likely to select destinations according to their level of gender equality in social institutions. Moreover, one can expect that migrations to countries with different gender equality rules in social institutions may influence institutional changes in the sending countries (condition (iii)). This is claimed by the litterature on transboundary transfers of norms (Spilimbergo, 2009, Beine, Docquier and Schiff, 2013), which supposes that political values are absorbed by immigrants, then transmitted to the origin countries where they contribute to reshaping attitudes and creating new norms about women. Lodigiani and Salomone (2015) find that international migration is associated with an increase in female political empowerment in the origin country. In the same way, Ferrant and Tuccio (2015) provide evidence on the link between South-South migration and discrimination against women in social institutions.

Our specific instrument measures the level of women's opportunity rights across the top 5 destinations of each country emigrants, weighted by the share of population from country $i$ who effectively migrate to each specific destination. We construct the instrument "Right Abroad" from the bilateral migration database of Özden, Parsons, Schiff and Walmsley (2011) and the "50 years of women's legal rights" database (The World Bank, 2012) as follows:

$$RightAbroad_{i,t} = \frac{1}{N_{i,t}} \sum_{i=1}^{N_{i,t}} \frac{Migr_{i,j,t}}{Pop_{i,t}} D_{j,t} \tag{1.4}$$

where $N_{i,t}$ denotes the number of target countries, up to five. $Migr_{i,j,t}$ refers to the number of people born in country $i$ who migrate to country $j$ at time $t$, while $Pop_{i,t}$ controls for the number of people living in country $i$ at time $t$. $D_{j,t}$ is the opportunity right index in country $j$ at time $t$, which is the sum of the indices for every individual law. Our external instrument is a proxy of the expected exposure to gender-equal rights from countries $j = 1$ to $N_{i,t}$, in the home country $i$ at time $t$.[17] We lagged this instrument by two periods (each of 5 years) as migrants need time to integrate and to transmit their new values to their home country.

Regarding the exclusion restriction, we assume that the level of women's rights in country $j$, where migrants reside, does not have a direct significant effect on the education system of country $i$, except through its influence on institutional changes in the sending countries. Indeed, we do not consider merely the percentage of migrants in all destinations, which may affect education directly through the transfer of education or fertility norms among others

[17] In order to ensure that the instrument does not capture only a population effect, we add the logarithm of population living in country $i$ at time $t$ as an additional covariate in Sub-Section 5.2.

(Bertoli and Marchetta, 2015), but we encompass instead the expected level of dissemination of the host countries' gender legal rights, conditionally to be an emigrant of country $i$. In Sub-Section 5.2, we include the average years of schooling of the 15-19 age cohort in the host country, weighted the probability that an individual from an origin country migrate to this destination, as a control variable in the second-stage equation. This specification may control for the expected transfer of education norms by emigrants from the host countries, which may be correlated with our external instrument that is the expected transfer of gender norms through migration. All instruments are included in the following first-stage regressions for the opportunity rights variable:

$$Opportunity_{i,t-1} = \gamma_0 + \gamma_1 Opportunity_{i,t-2} + \gamma_2 \sum_{j=1}^{5} RightAbroad_{j,t-2} + \gamma_3 Z_{i,t} + \delta_i + \rho_t + \zeta_{i,t} \tag{1.5}$$

where $\gamma_0$,_, $\gamma_3$ are vectors of parameters to estimate and $\delta_i$, $\rho_t$ and $\zeta_{i,t}$ are error terms. In respect of the internal instrument $Opportunity_{i,t-2}$, how plausible is the exclusion restriction that ten-year lagged legal rights affect education in time $t$ in no other way than through its effects on opportunity rights in time $t-1$? The intuition here is that a long lag of ten years is likely to dilute most reform effects on schooling, particularly if other society changes are also happening. Indeed, decisions on investment in child education are continually updated over time, adjusting to changing economic environment. In addition, the results of econometric tests indicate that the conditions for instrumenting the opportunity rights index by the internal and external instruments are respected. Indeed, the F-statistics of the first-stage regression, between 112 and 95 depending on the inclusion of covariates, demonstrate the power of the considered instruments (Staiger and Stock, 1997) and is above the critical values simulated by Stock and Yogo (2012) for the 5 percent level. The P-values of the Hansen test, between 0.11 and 0.15, indicate that the null hypothesis of over-identifying restrictions are not rejected, which support the used instrumental variables.

As mentionned before, an alternative specification of our schooling equation includes a lagged dependent variable term that may capture some of the dynamics involved in human capital accumulation in aggregate societies. This lagged effect may also help us to control for some endogeneity channels of the women's rights variables and control for other covariates that may not be strictly exogenous. However, it generates a new potential endogeneity issue associated with the lagged dependent variable. As a response to this issue, we use the System-GMM estimator of Blundell and Bond (1998), which simultaneously solve equations in levels and differences. The model is first estimated in first differences with instruments constructed from lagged levels as in Arellano and Bond (1991). However, the model in levels is also instrumented with the variables in lagged differences as they may remain good predictors for the endogenous variables, even when the series are very persistent. We use the Windmeijer (2005) correction procedure to correct the bias in the two-step standard errors. The status of covariates as exogenous, predetermined or endogenous variables has implications on the list of instruments used in the GMM analysis. Thus, we classify the logarithm of GDP per capita, the lagged dependent variable, the average years of schooling of people aged more than 25 and the democracy index as predetermined variables, which allow us to include their lag one as instruments.[18] Endogenous variables differ from predetermined variables only in that the former allow for correlation between the explanatory variable and the error term in time t, whereas the latter do not. Besides, since education and health evolve quite simultaneously, we categorize child mortality and women's opportunity rights as endogenous variables. As the number of periods is fairly large (9 periods representing 40 years), an unrestricted set of lags would introduce a huge number of instruments, with possibly some large small sample bias. This is why we collapse the instrument set so as to condense the information into a lower number of instruments. Collapsed instruments for predetermined variables in levels and in differences are constructed from lags one to eight, and from lags two to eight for the endogenous variables. We investigate the validity of the moments conditions used for the System-GMM estimates by testing the null hypothesis that the error term is not second-order serially correlated, an hypothesis confirmed with P-values around 0.7. The P-values of the corresponding Hansen's test of overidentifying restrictions,

[18]Indeed, unforcastable errors today are likely to much affect future changes in the economic growth and the human capital of both young and adult people. Generally, the more people are educated, the more they will support values related to liberty and equality.

between 0.15 and 0.22, support the use of the internal instruments based on the selected lags. Finally, we test the sensitivity of our specification by reducing the number of instruments to three or four lags, and test an alternative specification including the expectancy of women's opportunity rights of the migrants' destinations as an external instrument. This does not alter the results.

## 1.5 Results

### 1.5.1 Baseline results

Table 1.3 reports the estimations results for equations (2.2) and (2.3). The first three columns display the results for the within estimator. The 2SLS-FE results, based on the internal and external instruments, are presented in columns (4) to (6), while the last three columns show the results for the dynamic specification estimated using the System-GMM estimator. For each estimation method, the results are first shown without any controls, except the logarithm of GDP per capita (in columns (1), (4) and (7)), then the baseline controls are added (in columns (2), (5) and (8)). Finally, the fertility rate is included as another control variable in columns (3), (6) and (9). The results are not sensitive to the specification setup.

The estimated coefficient of the index of women's opportunity rights is positive and significant at the 1 percent level. Nevertheless, the coefficient is on average greater with the dynamic GMM estimator than with the 2SLS-FE, is and always higher than in FE regressions. The inclusion of the lagged dependent variable is supported by the results for System-GMM, where the coefficient on women's rights is magnified. The gender-equality reforms produce a larger impact on children's human capital at the beginning of their implementation, while their effects extend over time due to the inertia described by the lagged dependent variable coefficient. In all regressions, country and time fixed effects are included, so that the results account for any country-specific time-invariant characteristics and any global period-specific factors which may influence institutional gender equity and education. These characteristics

**Table 1.3** – Effects of women's opportunity rights on children's schooling

| Dependent variable | Average years of schooling of the 15-19 cohort | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variables | FE | | | 2SLS-FE | | | System GMM | | |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Opportunity, in t-1 | 0.147*** | 0.272*** | 0.251*** | 0.242*** | 0.334*** | 0.304*** | 0.192** | 0.361*** | 0.405*** |
| | (0.050) | (0.045) | (0.045) | (0.085) | (0.078) | (0.0809) | (0.092) | (0.107) | (0.104) |
| Income, pc | 1.070*** | 0.715*** | 0.706*** | 1.051*** | 0.696*** | 0.705*** | 0.358*** | 0.669* | 0.729** |
| | (0.147) | (0.133) | (0.132) | (0.181) | (0.164) | (0.163) | (0.110) | (0.383) | (0.358) |
| Child mortality | | -0.012*** | -0.011*** | | -0.012*** | -0.011*** | | -0.006*** | -0.006*** |
| | | (0.001) | (0.001) | | (0.001) | (0.002) | | (0.001) | (0.001) |
| Adult education | | 0.767*** | 0.715*** | | 0.774*** | 0.731*** | | -0.101 | -0.040 |
| | | (0.067) | (0.070) | | (0.078) | (0.080) | | (0.119) | (0.149) |
| Democratic index | | -0.019* | -0.018* | | -0.017 | -0.015 | | -0.003 | 0.001 |
| | | (0.009) | (0.009) | | (0.010) | (0.010) | | (0.015) | (0.019) |
| Fertility rate | | | -0.176** | | | -0.191** | | | 0.059 |
| | | | (0.073) | | | (0.085) | | | (0.118) |
| Education, in t-1 | | | | | | | 0.838*** | 0.653*** | 0.609*** |
| | | | | | | | (0.053) | (0.112) | (0.107) |
| F-Statistics | | | | 112 | 101 | 95 | | | |
| Hansen J-test (p-values) | | | | 0.13 | 0.11 | 0.15 | 0.15 | 0.22 | 0.21 |
| AR2 (p-values) | | | | | | | 0.72 | 0.70 | 0.69 |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes | No | No | No |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 513 | 507 | 507 | 447 | 442 | 442 | 513 | 507 | 507 |
| Number of countries | 72 | 72 | 72 | 69 | 69 | 69 | 72 | 72 | 72 |
| R-squared | 0.61 | 0.72 | 0.72 | 0.54 | 0.66 | 0.66 | | | |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard error in parenthesis. Opportunity is the women's opportunity right index that we constructed, ranging from 1 to 5. Income, pc represents the GDP per capita in logs. Child mortality is the under-five child mortality rate, per 1000 live births. Adult education represents the average years of schooling among individuals aged 25 years old and more. Democratic index is an indicator ranging from -10 to 10 which reflects the degree of democracy in the country. Education is the average years of education of the 15-19 age cohort. From columns (4) to (9), we take into account the potential endogeneity of the women's right index. IV estimations use the historical level of the legal rights index in the country and in the main destinations of the emigrants. Both variables are lagged by two period, which corresponds to 15 years. AR2 means second order autocorrelations.

include religion, colonial history and many other unobservable fixed features of the studied societies and periods.

In conclusion, all estimates indicate that women's legal opportunities are associated with higher children's educational attainment. For every unit increase in the Opportunity rights index, one can expect between a 0.14 and a 0.4 unit increase in the average years of schooling of the 15-19 age cohort, holding all other variables constant. Thus, additional reforms improving the equality of opportunity between men and women in the areas of finance, justice, labor and the family, increase the average years of schooling of the 15-19 up to 4 months. Over the fourty years period, the impact on children's education is similar to the equivalent effect of a rise of 43 percent of the logarithm of GDP per capita. Indeed, the coefficient of 0.7 on per capita income estimated with the 2SLS-FE including the control variables (column (6)), shows that an increase of 43 percent of the logarithm of GDP per capita would lead to a positive coefficient of 0.3, a magnitude associated with an additional female legal right.

Regarding the 2SLS-FE results, column (1) in Table 1.4 reports the results of equation 2.4 (the first-stage of the estimation in column (6) of Table1.3) and the related tests, using the internal and external instruments. The first-stage estimates indicate that both the ten-year lagged legal rights and the expected influence of the average level of women's rights in the migrant-receiving countries have significant coefficients at 1 percent level. Surprisingly, the latter variable reduces the probability of female reforms in the country of origin, when estimated with (column 1) and without (column 3) the internal instrument. Nevertheless, the second-stage estimates in panel B column (1) reveal the significance of the coefficient of the opportunity right index as well as the significance of the coefficient of all the control variables, except for the democracy index which has now no significant effect.

With respect to the System-GMM estimator, arguably the most appropriate estimator in the presence of a lagged dependent variable, columns (7) to (9) of Table 1.3 report the estimation results for the specifications described in Section 4.2. A one unit increase in the lagged dependent variable is associated with a 0.6 or 0.8 unit increase in children's average years of schooling. Incorporating the lagged dependent variable in the model amplifies the magnitude of coefficients on women's opportunity rights, suggesting a larger effect of women's rights

**Table 1.4 –** Endogeneity of women's opportunity rights: first-stage and second-stage estimates using 2SLS-FE

Panel A: First-stage estimations - 2SLS-FE

| Dependent variable | Women's rights | | | |
|---|---|---|---|---|
| Variables | (1) | (2) | (3) | (4) |
| Opportunity, in t-2 | 0.631*** | | | 0.566*** |
| | (0.047) | | | (0.056) |
| Opportunity, in t-3 | | 0.234*** | | |
| | | (0.079) | | |
| Opportunity in host countries, | -3.416** | -7.415*** | -6.008*** | -2.337 |
| in t-2 | (1.432) | (2.248) | (1.495 ) | (1.757) |
| Country FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| F-Statistics | 101 | 13.73 | 16.14 | 53 |
| Hansen J-test (p-values) | 0.11 | 0.26 | | 0.09 |

Panel B: Second-stage estimations - 2SLS-FE

| Dependent Variable | Average years of schooling of the 15-19 | | | |
|---|---|---|---|---|
| Variables | (1) | (2) | (3) | (4) |
| Opportunity, in t-1 | 0.334*** | 0.372** | 0.633** | 0.564*** |
| | (0.078) | (0.177) | (0.250) | (0.110) |
| Income, pc | 0.696*** | 0.671*** | 0.784*** | 0.878*** |
| | (0.164) | (0.225) | (0.152) | (0.184) |
| Child mortality | -0.012*** | -0.011*** | -0.014*** | -0.011*** |
| | (0.001) | (0.002) | (0.002) | (0.002) |
| Adult education | 0.774*** | 0.774*** | 0.853*** | 0.875*** |
| | (0.078) | (0.103) | (0.100) | (0.091) |
| Democratic index | -0.017 | -0.016 | -0.014 | -0.008 |
| | (0.010) | (0.012) | (0.010) | (0.011) |
| Education in host countries | | | | 0.610 |
| | | | | (1.234) |
| Country & Year FE | Yes | Yes | Yes | Yes |
| Observations | 442 | 379 | 500 | 373 |
| R-squared | 0.66 | 0.59 | 0.66 | 0.62 |
| Number of countries | 69 | 68 | 69 | 68 |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard error in parenthesis. Opportunity is the women's opportunity right index that we constructed, ranging from 1 to 5. Income, pc represents the GDP per capita in logs. Child mortality is the under-five child mortality rate, per 1000 live births. Adult education represents the average years of schooling among individuals aged 25 years old and more. Democratic index is an indicator ranging from -10 to 10 which reflects the degree of democracy in the country. Education in host countries is the average years of education of the 15-19 age cohort in the main destinations of emigrants. Opportunity in hos countries is the legal rights index in the main destinations of emigrants.

on children's schooling in the short term. Then, including the external instrument as a strictly exogenous variable in the estimation does not affect the significance nor the magnitude of the coefficient of interest. As well, the use of fewer lags for the instruments, from lags 1 to 3 for the pretermined variables, and lags 2 to 4 for the endogenous variables, does not affect the significance of the results (not shown and available upon request).

It is worth noting that the coefficient of the control variables have all the expected sign. The logarithm of GDP per capita presents a strong positive coefficient, between 0.35 and 1, significant at the one percent level in most specification. Resources invested in education, which improve the education services and relax liquidity constraints, largely increase children's educational attainment. Higher preferences for education, reflected by the average level of adult's schooling, which is significant in most regressions except in the GMM estimations, enhance also children's human capital. A one unit increase in parent's preferences in education leads to a 0.7 increase in the average years of schooling of the 15-19 age group. Interestingly, greater women's opportunities may affect parent's and especially mother's preferences for children's education as the opportunity right to acquire a job raises their returns to schooling. However, the inclusion of adult schooling in the estimation does not eliminate the significant effect of the opportunity rights index or reduce the magnitude of its coefficient. We may think that the elicited positive relationship between female opportunity rights and children's educational attainment is not fully explained by higher parent's preferences in education. In contrast, child mortality which reflects general health status and decrease children's returns in education, negatively affects the average years of schooling among the 15-19 years old. Once again, the inclusion of this variable does not affect the coefficient of female legal rights, suggesting that the main channel through which married women's rights reforms increase children's education is not children's health. Regarding the index of democracy, its coefficient is generally not significant.

Besides, the fertility rate, when included, presents a significant and negative coefficient which seems to validate Beckers' theory about the parental choice of quality versus quantity of children. Fewer children in the family augments the household capacity to educate well every one. However, the fertility rate does not have any impact on children's human capital when estimated with System-GMM. An explanation may be advanced: fertility rate may have small effects on parents' choices in the short term. Changes in women's bargaining power due to higher economic opportunities may impact fertility, which may in turn affect children's human capital. Indeed, the inclusion of fertility rate in the estimation slightly undermines the coefficient of the opportunity rights index. Nevertheless, the significance of women's rights remains, which suggests that a reduction in fertility, due to higher female bargaining power, does not explain all the variation in education resulting from a favourable change in law.

Finally, we put the fertility rate on the left-hand side of the equation as a dependent variable, investigating potential effects of women's opportunity rights on fertility decisions. We select the same covariates used in the education analysis, that is the logarithm of GDP per capita, child mortality rate, the adult average years of schooling and the democracy index. Results are presented in Table 2.16, where column (1) shows the regression estimated with the within estimator. The 2SLS-FE estimation is exposed in column (2), where we attempt to correct endogeneity of the female rights index with the internal and external instruments presented in Sub-Section 4.2. Finally, column (3) presents the results with System-GMM, using the lagged fertility rate as an explanatory variable. As previously mentionned, withdrawing laws that restrict women's economic independence reduces fertility, as coefficients of the opportunity rights index are negative and significant at 1 percent level in most specifications. An additional gender-equality law reduces the average fertility rate of 11 percentage points in the FE estimation, and 20 percentage points in the System-GMM estimation. Therefore, these results suggest that changes in reforms may reduce the number of desired children in the family, notably through higher female bargaining power. Once again, the magnitude of the coefficient is higher with the GMM estimator, using the same structure of lags as before. The logarithm of GDP per capita is not significant except in the System-GMM, where its coefficient becomes positive, suggesting that higher resources increase the fertility rate at a macroeconomic level in the short-run while its effect disappear when we extend the period of analysis. Besides, child mortality is always positive and significant at 1 percent level. The weaker the probability is to have children in good health, the higher the number of desired

**Table 1.5 –** Effects of women's opportunity rights on fertility

| Dependent variable | Fertility rate | | |
|---|---|---|---|
| | FE | 2SLS-FE | System GMM |
| Variables | (1) | (2) | (3) |
| Opportunity, in t-1 | -0.119*** | -0.153*** | -0.200** |
| | (0.029) | (0.049) | (0.083) |
| Income, pc | -0.050 | 0.048 | 0.437** |
| | (0.087) | (0.102) | (0.214) |
| Child mortality | 0.004*** | 0.004*** | 0.010*** |
| | (0.001) | (0.001) | (0.001) |
| Adult education | -0.292*** | -0.223*** | -0.077 |
| | (0.044) | (0.048) | (0.102) |
| Democratic index | 0.004 | 0.007 | 0.001 |
| | (0.006) | (0.006) | (0.008) |
| Fertility rate, in t-1 | | | 0.540*** |
| | | | (0.108) |
| Country FE | Yes | Yes | No |
| Year FE | Yes | Yes | Yes |
| F-Statistics | | 101 | |
| Hansen J-test (p-value) | | 0.16 | 0.10 |
| AR2 (p-values) | | | 0.21 |
| Observations | 507 | 442 | 507 |
| Number of countries | 72 | 69 | 72 |
| R-squared | 0.812 | 0.791 | |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parenthesis. Opportunity is the women's opportunity right index that we constructed, ranging from 1 to 5. Income, pc represents the GDP per capita in logs. Child mortality is the under-five child mortality rate, per 1000 live births. Adult education represents the average years of schooling among individuals aged 25 years old and more. Democratic index is an indicator ranging from -10 to 10 which reflects the degree of democracy in the country. In columns (2) and (3), we take into account the potential endogeneity of the women's opportunity right index. IV estimations use the historical level of the legal rights index in the country and in the main destinations of the emigrants. Both variables are lagged by two period, which corresponds to 15 years. AR2 means second order autocorrelations.

children is. The average number of years of schooling of the adult population is negative and significant in the first two columns, while higher preferences for children's education does not affect parent's fertility decisions in the short-run, in column (3). Finally, a one unit increase in the past level of fertility rate corresponds to a 0.5 unit increase in the current fertility rates, demonstrating that parent's fertility decisions evolve quite slowly over time.

### 1.5.2 Extensions and Robustness

A sensitivity analysis is conducted regarding the instrumentation strategy with the 2SLS-FE estimator. Although the results of the Arellano-Bond autocorrelation tests were supportive, concerns may still arise from the lags used with internal instrument, as women's reforms implemented ten years ago may still have small additional effects on the current state of education. As a response, we reinforce the exclusion restriction by lagging the index of opportunity rights one more additional period before to use it as internal instrument, resulting with a gap of fifteen years between the dependent variable and the instrument. The first stage results, which are presented in column (2) in panel A of Table 1.4, indicate that the opportunity rights index lagged with two periods, fifteen years hence, is a strong predictor of women's gender equality in legal rights five years ago. The lagged expected exposure to women's economic rights in the migration-receiving countries is associated to a lower probability of women's reforms in origin countries. The F-statistics of the first-stage regression associated with this specification is 13, above the critical values of Stock and Yogo (2012), which removes the concern for weak instruments. In addition, the P-value of the Hansen Test (0.26) in the second-stage estimation indicates that the exclusion restriction is respected. Then, we perform the estimation using only the external instrument, withdrawing any potential concern resulting from the use of internal instruments. In that case, the F-statistics of 16 in column (3) demonstrates that the influence of women's reforms adopted in other countries through migrants remains a strong instrument of the right index. Regarding the second-stage estimates, panel B indicates that the results are robust to the temporal structure of lags of the internal instrument (column (2)), and to the use of the external instrument exclusively (column (3)). However, the index of women's opportunity rights increases with the inclusion of additional lags and with the use of the external instrument exclusively (respectively around 0.3 and 0.6 percentage points).

The external instrument measures the expected influence of the average level of women's rights in migrant-receiving countries for each country of origin. Concerns may arise regarding the exclusion restriction: if we consider that the measure is strongly correlated with the education norms in host countries, the variable would affect children's human capital not only through the adoption of women's reforms, but also through the transfer of education norms. For that purpose, we construct the variable “EducAbroad" from the bilateral migration database of Özden et al. (2011) and the Barro and Lee (2013)'s database. Then, the variable measures the expected exposure to education norms from countries $j = 1$ to $N_{i,t}$, in the home country $i$ at time $t$, conditionally to be an emigrant:

$$EducAbroad_{i,t} = \frac{1}{N_{i,t}} \sum_{i=1}^{N_{i,t}} \frac{Migr_{i,j,t}}{Pop_{i,t}} Schooling_{j,t} \tag{1.6}$$

The only difference with the external instrument “RightAbroad" is that $Schooling_{j,t}$ refers to the average years of schooling of the 15-19 age cohort in country $j$ at time $t$. As the correlation between the “RightAbroad" instrument and the “EducAbroad" variable is around 0.3 significant at the 5 percent level, we include the expected transfer of education norms from destinations in the home country as a control variable in the second-stage estimation. Results are presented in column (4) in Table 1.4. The inclusion of education norms, which reduces its potential link with female opportunity rights in the migrant-receiving countries, does not affect the magnitude nor the significance of the coefficient of the women's rights index. This estimation then suggests that the external validity condition is respected with the use of the expeted transfers of gender norms from migrant-receiving countries. In addition, the inclusion of the logarithm of population living in country $i$ at time $t$ does not change the results, suggesting that the constructed external instrument does not capture only a population effect (results not shown but available).

**Table 1.6 –** Robust effects of women's legal rights: using individual rights separately

| Dependent variable | Average years of schooling of the 15-19 cohort | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variables | 2SLS-FE | | | | | System GMM | | | | |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Access to a job | 1.390*** | | | | | 0.446 | | | | |
| | (0.361) | | | | | (0.477) | | | | |
| Open a bank account | | 1.688*** | | | | | 1.406*** | | | |
| | | (0.392) | | | | | (0.340) | | | |
| Being the head of the family | | | 1.695*** | | | | | 0.909 | | |
| | | | (0.450) | | | | | (0.642) | | |
| Access to justice | | | | 1.975*** | | | | | 1.347*** | |
| | | | | (0.462) | | | | | (0.377) | |
| Sign legal contracts | | | | | 1.615*** | | | | | 1.426*** |
| | | | | | (0.372) | | | | | (0.361) |
| Education, in t-1 | | | | | | 0.678*** | 0.638*** | 0.717*** | 0.616*** | 0.583*** |
| | | | | | | (0.125) | (0.092) | (0.112) | (0.093) | (0.096) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| F-Statistics | 46 | 71 | 48 | 60 | 77 | | | | | |
| Hansen J-test (p-values) | 0.04 | 0.24 | 0.04 | 0.25 | 0.24 | 0.14 | 0.20 | 0.13 | 0.20 | 0.25 |
| AR2 (p-values) | | | | | | 0.72 | 0.53 | 0.62 | 0.58 | 0.46 |
| Country FE | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 442 | 443 | 442 | 443 | 442 | 513 | 515 | 510 | 521 | 513 |
| Number of countries | 69 | 69 | 69 | 69 | 69 | 72 | 72 | 72 | 72 | 72 |
| R-squared | 0.634 | 0.648 | 0.625 | 0.642 | 0.658 | | | | | |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parenthesis. Controls include the logarithm of GDP per capita, the average years of schooling of the population aged more than 25, the democracy index and the child mortality rate. Education is the average years of education of the 15-19 age cohort. IV estimations use the historical level of the legal rights index in the country and in the main destinations of the emigrants. Both variables are lagged by two period, which corresponds to 15 years. AR2 means second order autocorrelations.

We perform additional estimations to test the robustness of our results to the diverse components of the opportunity index. First, the aggregate women's opportunity index is split into its five components. Then, each individual right is included separately into different regressions. Unfortunately, the high correlation between these variables precludes their simultaneous inclusion in the regression. What we perform is therefore a kind of "horse race" exercise. Table 1.6 shows the results obtained when using the 2SLS-fixed effects and the System-GMM estimators.[19] Reforms allowing women to acquire a bank account, sign contract or initiate legal proceedings in court without their husband's permission have all positive and significant effects on children's education, when they are considered separately. The size of effects is greater than with the global index. Indeed, each additional individual right increases the average years of schooling among the 15-19 from 4 months to more than a year and a half. The coefficients are almost significant at one percent level, except for the right to acquire a job and to be legally considered as the head of household. The coefficients of the latter rights are strongly significant with the 2SLS-FE estimator, whereas their coefficients become insignificant when the dynamic of human capital accumulation is examined.

In another attempt, we disaggregate the analysis by gender. We replace the dependent variable with the average years of schooling for girls aged between 15-19, then for boys. Indeed, granting married women additional rights on the labor market may increase the expected benefits for girls' schooling more. We see in Table 1.7 that there is no significant difference according to gender. Women's opportunity rights increase the average years of schooling for both girls and boys, as all the coefficients of the variable of the opportunity index are significant and positive. The magnitude of the coefficient is slightly greater for girls than for boys, especially with the 2SLS-FE estimator.[20] Having one additional right

[19]In these specifications, we use the same internal and external instruments mentionned previously with the composite index.

[20]Note that the 2SLS-FE estimation is based on the fifteen-year lagged level of women's legal rights as internal instrument. The use of lagged periods reduces the sample size to 370 observations. The external instrument is still the expected exposure to women's rights from the migrant-receiving countries.

**Table 1.7 –** Effects of women's opportunity rights on children's schooling, by gender

| Dependent variable | Girls' Education | | | Boys' Education | | |
|---|---|---|---|---|---|---|
| Variables | FE | 2SLS-FE | System GMM | FE | 2SLS-FE | System GMM |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) |
| Opportunity, in t-1 | 0.259*** | 0.391** | 0.317*** | 0.288*** | 0.365* | 0.315** |
| | (0.047) | (0.189) | (0.097) | (0.052) | (0.197) | (0.126) |
| Income, pc | 0.702*** | 0.825*** | 0.571 | 0.700*** | 0.635** | 0.795 |
| | (0.139) | (0.240) | (0.355) | (0.153) | (0.251) | (0.494) |
| Child mortality | -0.014*** | -0.014*** | -0.008*** | -0.010*** | -0.009*** | -0.003 |
| | (0.001) | (0.002) | (0.002) | (0.001) | (0.002) | (0.001) |
| Adult education | 0.749*** | 0.716*** | -0.083 | 0.796*** | 0.825*** | -0.138 |
| | (0.071) | (0.110) | (0.114) | (0.078) | (0.115) | (0.126) |
| Democratic index | -0.020** | -0.019 | 0.001 | -0.017 | -0.014 | -0.007 |
| | (0.010) | (0.013) | (0.017) | (0.011) | (0.013) | (0.016) |
| Girls' education, in t-1 | | | 0.664*** | | | |
| | | | (0.127) | | | |
| Boys' education, in t-1 | | | | | | 0.687*** |
| | | | | | | (0.107) |
| Country FE | Yes | Yes | No | Yes | Yes | No |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| F-Statistics | | 13.73 | | | 13.73 | |
| Hansen J-test (p-values) | | 0.08 | 0.11 | | 0.67 | 0.28 |
| AR2 (p-values) | | | 0.88 | | | 0.77 |
| Observations | 507 | 379 | 507 | 507 | 379 | 507 |
| Number of countries | 72 | 68 | 72 | 72 | 68 | 72 |
| R-squared | 0.755 | 0.629 | | 0.597 | 0.485 | |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard error in parenthesis. Opportunity is the women's opportunity right index that we constructed, ranging from 1 to 5. Income, pc represents the GDP per capita in logs. Child mortality is the under-five child mortality rate, per 1000 live births. Adult education represents the average years of schooling among individuals aged 25 years old and more. Democratic index is an indicator ranging from -10 to 10 which reflects the degree of democracy in the country. Girls' education is the average years of education among girls aged 15-19 years old. Boys' education is analogously defined among boys only. IV estimations use the historical level of the legal rights index in the country and in the main destinations of the emigrants. Both variables are lagged by two period, which corresponds to 15 years. AR2 means second order autocorrelations.

increasing women's legal capacity raises educational attainment of girls and boys between 2.5 and 4 months on average.

In Table 1.8, we include additional control variables: agriculture value added,[21] the logarithm of government spending on education, the urbanization rate; and we replace adult schooling with women's average years of schooling, the latter accounting potentially better for the specific channel of mother's schooling. Given that female employment may be a key transmission variable, special attention is devoted to this variable that is also included. A shortcoming of including all these controls is that the sample size drops because of missing values. Despite all these changes, the estimates of the coefficient of the opportunity index remains the same.

We also exploit a specific information from the World Bank database on the possibility that customary or religious law may be superior to the constitution. Through this information, we restrict the analysis only to countries where religious norms do not prevail on the law. This reduces our sample to 330 observations and does not affect nor the significance nor the magnitude of coefficients of the women's opportunity index (results not shown, but available).

Therefore, the robustness part consolidates the elicited effect of women's opportunity rights on children's educational attainment in developing countries. Our instrumentation strategy, which accounts for potential endogeneity of the female rights, is not sensitive to the use of particular instruments, as the variation of lags of the internal instrument or the use of the external instrument exclusively does not affect the results. Moreover, the inclusion of additional covariates, which does not affect the significance nor the magnitude of the coefficient of the women's rights index, emphasizes and validates our identification strategy.

In the next section, we investigate further the foundation of the effect of the women's opportunity rights on children's schooling.

[21] In that case, we have to drop the logarithm of GDP per capita when testing the agriculture variable due to multicolinearity.

**Table 1.8 –** Robust effects of women's opportunity rights on children's education: additional controls

| Dependent variable | Average years of schooling of the 15-19 | | |
|---|---|---|---|
| | FE | 2SLS-FE | System GMM |
| Variables | (1) | (2) | (3) |
| Opportunity, in t-1 | 0.210*** | 0.249* | 0.321** |
| | (0.073) | (0.138) | (0.155) |
| Child mortality | -0.010*** | -0.010*** | -0.010*** |
| | (0.002) | (0.002) | (0.004) |
| Women's adult education | 0.594*** | 0.604*** | 0.066 |
| | (0.122) | (0.126) | (0.134) |
| Democratic index | -0.034* | -0.033* | -0.009 |
| | (0.018) | (0.017) | (0.027) |
| Urbanization | 0.015 | 0.012 | -0.015 |
| | (0.022) | (0.022) | (0.014) |
| Education spendings (in logs) | 0.037* | 0.037* | 0.026* |
| | (0.0201) | (0.019) | (0.013) |
| Agriculture value added | -0.003 | -0.008 | -0.018 |
| | (0.015) | (0.016) | (0.026) |
| Female labor (in logs) | -0.283 | -0.205 | -0.783 |
| | (0.539) | (0.537) | (1.159) |
| Education, in t-1 | | | 0.572*** |
| | | | (0.152) |
| Country FE | Yes | Yes | No |
| Year FE | Yes | Yes | Yes |
| F-Statistics | | 32 | |
| Hansen J-test (p-values) | | 0.54 | 0.09 |
| AR2 (p-values) | | | 0.87 |
| Observations | 262 | 255 | 262 |
| Number of countries | 70 | 64 | 70 |
| R-squared | 0.53 | 0.53 | |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard error in parenthesis. Opportunity is the women's opportunity right index that we constructed, ranging from 1 to 5. Income, pc represents the GDP per capita in logs. Child mortality is the under-five child mortality rate, per 1000 live births. Women's adult education represents the average years of schooling among women aged 25 years old and more. Democratic index is an indicator ranging from -10 to 10 which reflects the degree of democracy in the country. Urbanization is the share of the urban population. Education spendings represents the government expenditures on education, in logs. Female labor is the female labor force participation rate, in logs. In columns (2) and (3), we take into account the potential endogeneity of the women's opportunity right index. IV estimations use the historical level of the legal rights index in the country and in the main destinations of the emigrants. Both variables are lagged by two period, which corresponds to 15 years. AR2 means second order autocorrelations.

## 1.6 Difference-in-differences

The adoption of one law in a given country can be viewed as a treatment for this country at the time this law is passed. The treatment is the implementation of an additional law giving women equal opportunities than men in their private and social life. Precisely, our treatment definition, the women's opportunity index moves from $\{0, 1, 2, 3, 4\}$ to 5. This specification allows for the largest variation in the treatment variable.

The treatment effect is estimated by comparing the average change over time in the outcome variable, the average years of schooling of the 15-19 age group, for the treatment group of countries compared to the average change over time for a control group. The treatment group is made of the countries that eliminated gender disparities in all the considered opportunity rights during the period under study. The control group includes the countries for which gender disparities in legal rights remain present; that is: countries for which the index of women's opportunities is less than 5 throughout the period. Finally, we exclude all countries which did not suffer from any gender inequality in opportunity rights from the initial period in 1970. This categorization results in 20 countries in the treatment group and 7 countries in the control group.[22] The estimation method is an extension of the typical difference-in-differences method. The statistical conditions for the extension of this method to treatments occurring at different periods are exposed in Muller (2018). They mostly account to assuming some homogeneity of the effect of the treatment over time for each country.

Regarding the specification, we first define two dummy variables, respectively for the treatment group, and for the post-treatment period. In Equation 1.7, the effect of the policy is measured by the coefficient $\beta_1$ of the interaction of these two variables. As expected, the mean characteristics of the countries of the treatment group and of the countries of the control group differ. These differences are jointly controled by a few observable characteristics that

[22]The control group includes Cameroon, Ivory Coast, Gabon, Jordan, Mali, Niger, Syrian Arab Republic. The treatment group encompasses Benin, Bolivia, Brazil, Central African Republic, Dominican Republic, Guatemala, Indonesia, Iran, Lesotho, Morocco, Namibia, Paraguay, Peru, Philippines, Rwanda, Senegal, South Africa, Turkey, Yemen, Zimbabwe.

can be included in Equation 1.7 (in $X_{i,t}$). Finally, country and year fixed effects are added. The estimation is conducted with the Ordinary Least-Squares method.

$$Schooling_{i,t} = \beta_1 Treat_{i,t} + \beta_2 X_{i,t} + \varphi_i + \nu_t + \zeta_{i,t} \tag{1.7}$$

As previously mentionned, the dependent variable $Schooling_{i,t}$ refers to the average years of schooling of the 15-19 age group in country $i$ and year $t$. The effect of the treatment is measured through the coefficient of the variable $Treat_{i,t}$, which represents the interaction of the dummy of the treatment group and the dummy for the post-implementation period. Moreover, educational attainment is determined by the sum of time-invariant country effects $\varphi_i$, time dummies $\nu_t$ that are common across countries, and some observable controls $X_{i,t}$. $X_{i,t}$ includes variables used as controls in the previous models: logarithm of GDP per capita, child mortality, adult schooling and a policy index.

We first check if the hypothesis of common trend without treatment for the two groups is plausible. This is a key assumption the used difference-in-differences strategy that the outcome in the treatment and control group would follow the same time trend in the absence of the treatment. Figure 1.4 plots, separately for the treatment group and for the control group, the changes in the average years of children's education for periods before the implementation of an additional law giving married women complete equality in legal rights. The graphics shows that the common trend assumption is respected. Indeed, the average years of education of children in the control and treatment group follow the same pre-trend during the period under study, althought the initial point is higher for treated countries than untreated ones.

Regarding the results, Table 1.9 indicates that the effect of the treatment is positive and significantly different from zero, in the two tied specifications. The magnitude of the coefficient of the variable $Treat_{i,t}$ is 0.41 in column 2, with the inclusion of covariates, time and country dummies, and much higher at 1.78 without controls, which seems to show the importance of these controls. The magnitude close to the one obtained with previous estimations, which is comforting. Consequently, in this situation, adopting a law to end gender inequality

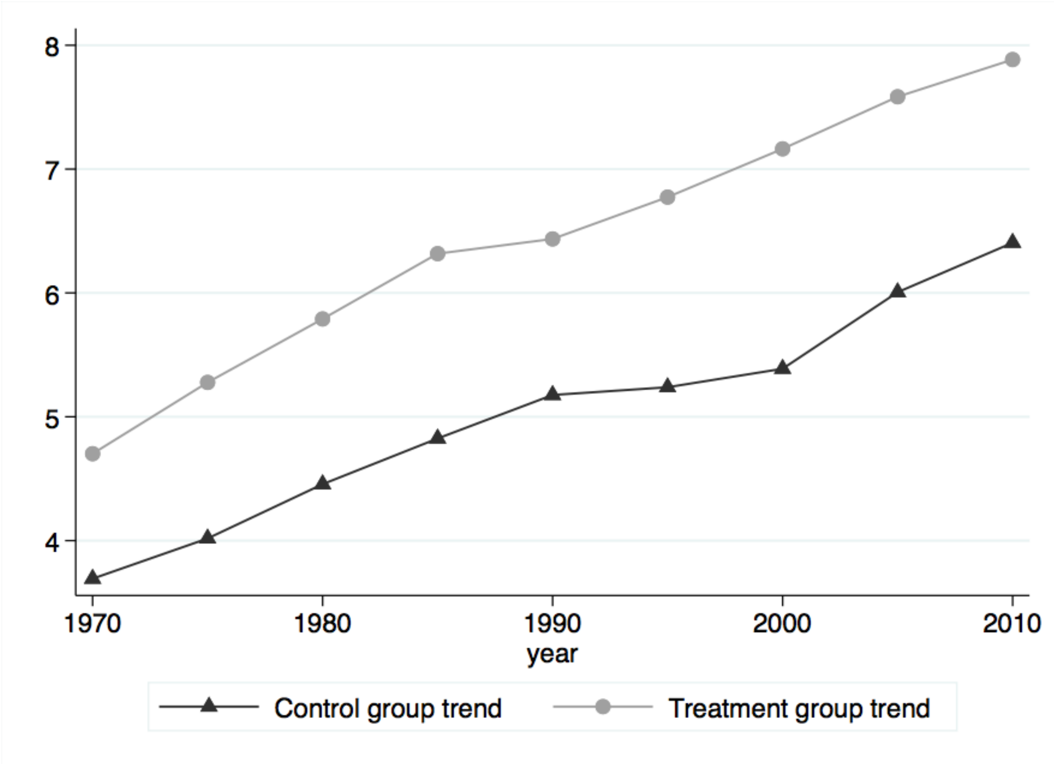



**Figure 1.4** – Common trend assumption

in economic opportunities increases the average years of schooling of the 15-19 age group by around 4 months. This is a confirmation of the results already achieved with standard panel data methods.

## 1.7 Conclusion

Concerns are sometimes raised that gender inequalities in legal rights may negatively impact children human capital. In this paper, we estimate the effect of laws promoting gender equality in economic opportunities on children schooling for the 15-19 age group. We pursue an aggregate approach of these issues at country level over the period 1970-2010, for 75 developing countries, with dynamic panel estimations. The possible endogeneity of social institutions in the education model is attenuated by using an external instrument based on the gender-equity context of the host country for migrants, and internal instruments. Under certain conditions which make our external instrument valid, we provide some evidence on how government reforms that improve equality between men and women in their private and

**Table 1.9 –** Estimation results using the difference-in-differences strategy

| Dependent variable | Average years of schooling of the 15-19 | |
|---|---|---|
| Variables | (1) | (2) |
| Equality in rights | 1.778*** | 0.417*** |
| | (0.219) | (0.134) |
| Income, pc | | 0.301** |
| | | (0.148) |
| Child mortality | | -0.005*** |
| | | (0.001) |
| Adult education | | 0.712*** |
| | | (0.066) |
| Democratic index | | -0.028** |
| | | (0.011) |
| Country & Year FE | No | Yes |
| Observations | 322 | 312 |
| Number of countries | 39 | 39 |
| R-squared | | 0.801 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parenthesis. Equality in rights is the treatment effect: the implementation of an additional law giving women equal opportunities than men in their economic life. Income, pc represents the GDP per capita in logs. Child mortality is the under-five child mortality rate, per 1000 live births. Adult education represents the average years of schooling among individuals aged 25 years old and more. Democratic index is an indicator ranging from -10 to 10 which reflects the degree of democracy in the country.

social life contribute to improve children educational attainment. Endowing women with equal opportunities than men in justice, on the labour and financial market, and within the family, is found to increase children's average years of schooling in the latest stages of their human capital accumulation. The analysis disaggregated by sex tells us that greater opportunity rights allocated to married women improve both girls and boys' human capital, with a similar magnitude, suggesting that higher returns for girls' schooling on the labor market is not the only mechanism at stake.

Our results are in line with Qian (2008), in which the causality is more reliable. Our paper comforts the external validity since we generalize the results to other countries and situations than China. The implication for policy makers is straightforward: factors that increase the economic value of women are also likely to increase the education investment in all children.

The main limitation of the study is the use of a composite index, which does not provide any information on what is the most important right affecting children's education. Although many countries have made efforts to review and reform legislation, translating gender laws from theory into practice at community level remains a challenge. Thus, our analysis is constrained by the nature of the data, which do not measure the degree of gender law enforcement. A future research will be to replicate the analysis with macroeconomic panel data measuring law enforcement through surveys, not available so far. Moreover, the elicited effect of an additional law giving women more economic opportunities on children's educational attainment may be alleviated under asymmetry of information, a relevant assumption in developing countries. Indeed, women in rural region may not be endowed with same information, ignoring particular changes in the legislation that may improve their empowerment and economic status. Thus, supporting women's equal opportunity rights through changes in legislation should be associated with policies promoting women's information, especially in rural areas. Similarly, women's literacy program will help them to benefit from these gender rights, as the right to open a bank account or to work will be inefficient in changing women's behaviour if women do not have the abilities to access these services.

# References


Acemoglu, Daron and Simon Johnson. 2005. "Unbundling Institutions." *Journal of Political Economy* 113(5):949–995.

Aghion, Philippe, Torsten Persson and Dorothee Rouzet. 2012. "Education and Military Rivalry." *National Bureau of Economic Research Working Paper 18046* .

Alderman, Harold, Jere R. Behrman, Victor Lavy and Rekha Menon. 2001. "Child Health and School Enrollment: A Longitudinal Analysis." *The Journal of Human Resources* 36(1).

Allendorf, Keera. 2007. "Do Women's Land Rights Promote Empowerment and Child Health in Nepal?" *World Development* 35(11):1975–1988.

Anderson, Siwan and Jean-Marie Baland. 2002. "The Economics of Roscas and Intrahousehold Resource Allocation." *The Quarterly Journal of Economics* 117(3):963–995.

Arellano, Manuel and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *The Review of Economic Studies* 58(2):277–297.

Baland, Jean-Marie and James A. Robinson. 2000. "Is Child Labor Inefficient?" *Journal of Political Economy* 108(4):663–679.

Baldacci, Emanuele, Benedict Clements, Sanjeev Gupta and Qiang Cui. 2008. "Social Spending, Human Capital, and Growth in Developing Countries." *World Development* 36(8):1317–1341.

Barro, Robert J. and Jong-Wha Lee. 1993. "International comparisons of educational attainment." *Journal of Monetary Economics* 32:363–394.

Barro, Robert J. and Jong-Wha Lee. 2013. "A new data set of educational attainment in the world, 1950–2010." *Journal of Development Economics* 104:184–198.

Becker, Gary and H Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy* 81(2):279–288.

Behrman, Jere R. 2010. Investment in education—Inputs and Incentives. In *Handbook of Development Economics*. Vol. 5 Elsevier pp. 4883–4975.

Behrman, Jere R., Andrew D. Foster, Mark R. Rosenweig and Prem Vashishtha. 1999. "Women's Schooling, Home Teaching, and Economic Growth." *Journal of Political Economy* 107(4):682–714.

Behrman, Jere R. and Mark R. Rosenweig. 2002. "Does increasing women's schooling raise the schooling of the next generation?" *The American Economic Review* 92(1):323–334.

Beine, Michel, Frederic Docquier and Maurice Schiff. 2013. "International Migration, Transfers of Norms and Home Country Fertility." *Canadian Journal of Economics* 46(4):1406–1430.

Bennell, Paul. 2002. "Hitting the Target: Doubling Primary School Enrollments in Sub-Saharan Africa by 2015." *World Development* 30(7):1179–1194.

Bertoli, Simone and Francesca Marchetta. 2015. "Bringing It All Back Home – Return Migration and Fertility Choices." *World Development* 65:27–40.

Bils, Mark and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *The American Economic Review* 90(5):1160–1183.

Blundell, Richard and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics* 87:115–143.

Bobonis, Gustavo J. 2009. "Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment." *Journal of Political Economy* 117(3):453–503.

Borensztein, Eduardo, José de Gregorio and Jong-Wha Lee. 1998. "How does foreign direct investment affect economic growth?" *Journal of International Economics* 45(1):115–135.

Branisa, Boris, Stephan Klasen and Maria Ziegler. 2013. "Gender Inequality in Social Institutions and Gendered Development Outcomes." *World Development* 45:252–268.

Brown, Philip H. 2009. "Dowry and Intrahousehold Bargaining: Evidence from China." *The Journal of Human Resources* 44(1):25–46.

Cerrutti, Marcela and Douglas S Massey. 2001. "On the auspices of female migration from Mexico to the United States." *Demography* 38(2):187–200.

Cervellati, Matteo and Uwe Sunde. 2005. "Human Capital Formation, Life Expectancy, and the Process of Development." *The American Economic Review* 95(5):1653–1672.

Cervellati, Matteo and Uwe Sunde. 2013. "Life Expectancy, Schooling, and Lifetime Labor Supply: Theory and Evidence Revisited." *Econometrica* 81(5):2055–2086.

Chiquiar, Daniel and Gordon H Hanson. 2005. "International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States." *Journal of political Economy* 113(2):239–281.

Currie, Janet and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *The Quarterly Journal of Economics* 118(4):1495–1532.

Deere, Carmen Diana, Rosa Luz Durán, Merrilee Mardon and Tom Masterson. 2005. "Female land rights and rural household incomes in Brazil, Paraguay and Peru." Agriculture and Rural Development Internal Report. Washington, D.C.: The World Bank.

Deininger, Klaus, Aparajita Goyal and Hari Nagarajan. 2013. "Women's inheritance rights and intergenerational transmission of resources in India." *Journal of Human Resources* 48(1):114–141.

Doepke, Matthias and Michèle Tertilt. 2009. "Women's Liberation: What's in It for Men?" *The Quarterly Journal of Economics* 124(4):1541–1591.

Doepke, Matthias, Michèle Tertilt and Alessandra Voena. 2012. "The Economics and Politics of Women's Rights." *Annu. Rev. Econ.* 4(1):339–372.

Dollar, David and Roberta Gatti. 1999. *Gender inequality, income, and growth : are good times good for women ?* Vol. 1 Development Research Group, The World Bank Washington, DC.

Doss, Cheryl. 2005. "The Effects of Intrahousehold Property Ownership on Expenditure Patterns in Ghana." *Journal of African Economies* 15(1):149–180.

Doss, Cheryl. 2013. "Intrahousehold Bargaining and Resource Allocation in Developing Countries." *World Bank Research Observer* 28(1):52–78.

Duflo, Esther. 2012. "Women Empowerment and Economic Development." *Journal of Economic Literature* 50(4):1051–1079.

Edmonds, Eric V. 2006. "Child labor and schooling responses to anticipated income in South Africa." *Journal of Development Economics* 81:386–414.

Edmonds, Eric V. 2007. "Child Labor." *Handbook of Development Economics* 4:3607–3709.

Eswaran, Mukesh. 2002. "The empowerment of women, fertility, and child mortality: Towards a theoretical analysis." *Journal of Population Economics* 15:433–454.

Fernandez, Raquel. 2014. "Women's rights and development." *Journal of Economic Growth* 19:31–80.

Ferrant, Gaëlle and Alexandre Kolev. 2016. "Does gender discrimination in social institutions matter for long-term growth?: Cross-country evidence." Documents de travail du Centre de développement de l'OCDE, no. 330, Éditions OCDE, Paris, https://doi.org/10.1787/5jm2hz8dgls6-en.

Ferrant, Gaëlle and Michele Tuccio. 2015. "South–South Migration and Discrimination Against Women in Social Institutions: A Two-way Relationship." *World Development* 72:240–254.

Field, Erica. 2003. "Fertility Responses to Land Titling: The Roles of Ownership Security and the Distribution of Household Assets." *V Cambridge MA: Harvard University* .

Field, Erica. 2007. "Entitled to Work: Urban Property Rights and Labor Supply in Peru." *The Quarterly Journal of Economics* 122(4):1561–1602.

Geddes, Rick, Dean Lueck and Sharon Tennyson. 2012. "Human Capital Accumulation and the Expansion of Women's Economic Rights." *Journal of Law and Economics* 55(4):839–867.

Gregorio, Jose De and Jong-Wha Lee. 2002. "Education and Income Inequality: New Evidence from Cross-Country Data." *Review of Income and Wealth* 48(3).

Grogger, Jeffrey and Gordon H Hanson. 2011. "Income maximization and the selection and sorting of international migrants." *Journal of Development Economics* 95(1):42–57.

Hallward-Driemeier, Mary and Ousman Gajigo. 2015. "Strengthening Economic Rights and Women's Occupational Choice: The Impact of Reforming Ethiopia's Family Law." *World Development* 70:260–273.

Hallward-Driemeier, Mary, Tazeen Hasan and Anca Bogdana Rusu. 2013*a*. "Women's legal rights over 50 years: progress, stagnation or regression?".

Hallward-Driemeier, Mary, Tazeen Hasan and Anca Bogdana Rusu. 2013*b*. "Women's Legal Rights over 50 Years What Is the Impact of Reform?" *World Bank Policy Research Working Paper Series* (6617).

Haveman, Robert and Barbara Wolfe. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature* 33(4):1829–1878.

Htun, Mala and Laurel Weldon. 2012. *Sex Equality in Family Law: Historical Legacies, Feminist Activism, and Religious Power in 70 Countries*. Washington, DC: World Bank.

Inglehart, Ronald and Pippa Norris. 2003. *Rising Tide. Gender Equality and Cultural Change around the World*. Cambridge University Press.

Jensen, Robert T. 2010. "Economic Opportunities and Gender Differences in Human Capital: Experimental Evidence from India." (No. w16021). National Bureau of Economic Research.

Juhn, Chinhui, Gergely Ujhelyi and Carolina Villegas-Sanchez. 2013. "Trade Liberalization and Gender Inequality." *The American Economic Review* 103(3):269–273.

Kim, Se-Jik and Yong Jin Kim. 2000. "Growth gains from trade and education." *Journal of International Economics* 50(2):519–545.

Kumar, Neha and Agnes R. Quisumbing. 2012. "Beyond "Death Do Us Part": The Long-Term Implications of Divorce Perceptions on Women's Well-Being and Child Schooling in Rural Ethiopia." *World Development* 40(2):2478–2489.

Lodigiani, Elisabetta and Sara Salomone. 2015. "Migration-Induced Transfers of Norms: The Case of Female Political Empowerment." *University Ca' Foscari of Venice, Dept. of Economics Research Paper Series No. 19/WP/2015.* .

Lorentzen, Peter, John McMillan and Romain Wacziarg. 2008. "Death and Development." *Journal of Economic Growth* 13:81–124.

Mammen, Kristin and Christina Paxson. 2000. "Women's Work and Economic Development." *The Journal of Economic Perspectives* 14(4):141–164.

Marshall, Monty G., Ted Robert Gurr and Keith Jaggers. 2017. Polity IV Dataset version 2016. Technical report College Park, MD: Center for International Development and Conflict Management, University of Maryland.

Menon, Nidhiya, Yana Van Der Meulen Rodgers and Huong Nguyen. 2014. "Women's Land Rights and Children's Human Capital in Vietnam." *World Development* 54:18–31.

Neumayer, Eric and Indra De Soysa. 2011. "Globalization and the Empowerment of Women: An Analysis of Spatial Dependence via Trade and Foreign Direct Investment." *World Development* 39(7):1065–1075.

Özden, Çağlar, Christopher R. Parsons, Maurice Schiff and Terrie L. Walmsley. 2011. "Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960-2000." *World Bank Economic Review* 25(1):12–56.

Perotti, Roberto. 1996. "Growth, Income Distribution, and Democracy: What the Data Say." *Journal of Economic Growth* 1:149–187.

Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: the effect of sex-specific earnings on sex imbalance." *The Quarterly Journal of Economics* 123(3):1251–1285.

Quisumbing, Agnes R. and John A. Maluccio. 2003. "Resources at Marriage and Intrahousehold Allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa." *Oxford Bulletin of Economics and Statistics* 65(3):283–327.

Rajkumar, Andrew Sunil and Vinaya Swaroop. 2008. "Public spending and outcomes: Does governance matter?" *Journal of Development Economics* 86:96–111.

Ravallion, Martin and Quentin Wodon. 2000. "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy." *The Economic Journal* 110(462):158–175.

Simmons, Beth A. 2000. "International Law and State Behavior: Commitment and Compliance in International Monetary Affairs." *The American Political Science Review* 94(4):819–835.

Spilimbergo, Antonio. 2009. "Democracy and Foreign Education." *The American Economic Review* 99(1):528–543.

Staiger, Douglas and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3):557–586.

Stock, James H and Motohiro Yogo. 2012. *Testing for weak instruments in linear IV regression*. Working paper, Harvard University.

The World Bank. 2011. World Development Report 2012: Gender Equality and Development. Technical report The World Bank, Washington, DC.

The World Bank. 2012. Historic Data: 50 years of women's legal rights. Technical report The World Bank, Washington, DC.

Thomas, Duncan. 1990. "Intra-Household Resource Allocation: An Inferential Approach." *The Journal of Human Resources* 25(4):635–664.

Wang, Shing-Yi. 2014. "Property rights and intra-household bargaining." *Journal of Development Economics* 107:191–201.

Wiig, Henrik. 2013. "Joint Titling in Rural Peru: Impact on Women's Participation in Household Decision-Making." *World Development* 52:104–119.

Windmeijer, Frank. 2005. "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators." *Journal of Econometrics* 126(1):25–51.

# 2 Heterogeneous Effect of Income Shocks on Educational Choices: Evidence from Rural Tanzania

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. I find that unfavourable climate shocks for agricultural crops increase the probability of school dropout for children with sisters and decrease the probability for children with relatively more brothers. Disaggregating the analysis by sibling subgroups, climate shocks differently affect children according to the sex composition of the young sibship. Children suffer an additional penalty during climate shocks the larger the share of girls among the younger siblings. Sibling characteristics alter the impact of income shocks on girls' education in particular. The results are driven mainly by the child labor mechanism. Indeed, underpriviledged children help the family overcome income shortfalls by increasing their labor in housework and agriculture.

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

## 2.1 Introduction

Recurrent droughts and heavy rainfall afflict Southern and Eastern Africa since the last 60 years (Field, 2014). As a result, a large body of the literature investigates whether climate shocks, or more generally fluctuations in parental income, are related to education disparities in these regions. Because of credit constraints, rural households have different options to insure subsistence needs against shocks: adjusting labor-supply, reducing food consumption, shrinking expenditures in education, in health or re-organizing the household through marriage. These alternatives could be damaging to child educational attainment.

However, most of the existing studies assume that income shocks effects are uniformly distributed and result in symetric outcomes across children. Rather than accepting a static link between income shocks and educational outcomes, a key challenge is to understand which conditions aggravate or mediate the impact of shocks such as drought. Parish and Willis (1993) demonstrate that children's education depend first on economic security but also on the nature of siblings who must share family resources. Duryea, Lam, and Levison (2007) demonstrate that households feeling the pinch of economic hardship reallocate their resources to subsist to the negative economic shock. These empirical evidences already allude to the fact that there is bound to be heterogeneity depending on the family composition. Indeed, sibling composition might exerce strong influence among families impacted by negative shocks who might choose more carefully who in the home got educated. The role of household composition when analyzing the response to income variation might be stronger for girls in particular, who are often affected by her siblings whereas men's educational choices have not (Butcher and Case, 1994). How would such a complication affect the analysis of the effect of income shocks on children's education?

I stress the role of heterogeneity across children and asses whether the family composition has a mitigating effect on the school enrollment response to an income shock. In particular, this paper explores the effect of income shocks on children's human capital depending on the composition of siblings in a rural region of Tanzania. In many low-income countries, a large proportion of the population depends on agriculture for its primary means of support, whether through subsistence agriculture, the production of cash crops, or as hired farm labor. Over 80% of Tanzanian households depend on the agricultural sector, which produces about 95% of the country's food requirements (FAO, 2019).[1] Tanzania, like most

[1]The agricultural sector contributes about 30% of GDP (at current prices) in 2016. Food crops production dominates the sector, totalling 15.5% of GDP with livestock accounting for 7.7%. The sector is dominated by smallholders that do farming at subsistence level, organized in some 8,000 villages with an average holding of less than 2 ha per family.

of Sub-Saharan countries, is particularly sensitive to climate extremes due to its reliance on rain-fed subsistence agriculture. Approximately 26% of Tanzanians live below the official World Bank poverty line and spend a large fraction of their budget on food.

In this study, I focus on the Kagera region, exploiting the Kagera Health Development Survey (KHDS).[2] KHDS data has serveral advantages to study how parents reallocate resources differently within the family following climate shocks. Firstly, Tanzania is one of the most country with low access to higher education. Secondary school enrollment rate is as low as 31 percent in 2016 (UIS, 2017). The percentage for girls is even lower, due mainly to early marriage and adolescent pregnancy.[3] Secondly, the KHDS survey is particularly rich in terms of educational data, which I exploit in my empirical work. The availability of information on age at school entry and school drop-out makes the data particularly suited for this resarch compared to existing surveys in Sub-Saharan Africa. Lastly, because the KHDS respondents are rural, involving mostly subsistence farmers, rainfall variation across space and time should generate corresponding variation in agricultural output and thus household income.

I construct a microeconomic panel model exploiting exogenous variation in rainfall across districts as a reduced-form instrument for household income variability. Retrospective information from the KHDS allows me to re-construct a panel of individuals between each respondent's age of entry into school to 2010. My identification strategy relies on random variations in the distribution of dry and wet climate shocks across space and time. Specifically, I use data from the Standard Precipitation Evapotranspiration Index (SPEI) to construct measures of exposure to dry and wet climate shocks at a localized level and estimate the relative effect of exposure to (unusual) agricultural conditions on the probability of school dropout among children with different sibling composition. The estimate is based on difference-in-difference

[2] Kagera contributes to 3.95% of the Tanzania's GDP current market prices in 2016, ranked 13 over the 25 regional economies. Kagera closely represents the large part of Tanzania, except Dar es Salaam and Mwanza regions which contribute significantly to the total GDP at current prices. From the year 2007 to 2016, the average contribution by the Dar es Salaam and Mwanza regions was about 26.7%, of which, 17% was contributed by Dar es Salaam. The share of the two regions is attributed to the concentration of economic activities such as manufacturing, mining, real estate, accommodation and food services.

[3] An estimated 5.1 million children aged 7 to 17 do not attend school. Adolescent pregnancy led to almost 3,700 girls dropping out of primary and secondary education in 2016.

by interacting climate shocks with variables that reflect the size and sex composition of siblings. I pay particular attention to whether children with relatively more sisters than brothers are better ensured during adverse climate shocks. I include individual and year fixed-effects to control for unobservables that do not vary over time. I take into account gender differences in educational returns and allow for the interdependence of decisions regarding siblings arising from the liquidity constraint. I first investigate the short-run responses to economic shocks, focusing on a child's probability of school enrollment as the outcome. Then, a long-term analysis explores the potential long-lasting consequences of income shocks on human capital accumulation, which ultimately determines individual's economic status over the life cycle.

In 2018, I visited a sub-sample of the KHDS villages in order to collect qualitative data on coping strategies following income shocks. Incorporating the qualitative research into large quantitative household surveys yields a deeper understanding of parental beliefs and adaptative strategies in rural areas. Particularly, I conducted the qualitative survey to test the hypothesis underlying my quantitative findings. On average, negative productivity shocks do not change the probability to drop-out of school. The result masks large heterogeneity. Indeed, children's schooling is sensitive to the sex composition of siblings during a shock. The education of both men and women responds to the percentage of siblings that are females during the loss of income. The age of siblings is also important to predict the effect of climate shocks since opportunity costs of education might vary by age group (Fafchamps and Wahba, 2006) and returns to education might be non-linear (Kuepie, Nordman, and Roubaud, 2009). Accordingly, I disaggregate the analysis by sibling subgroups. I find that climate shock raises the odds of school drop out for children with relatively more younger sisters while it decreases the likelihood for children with relatively more younger brothers. Estimates suggest that changing the sibling composition from younger sisters to younger brothers decreases the probability of school dropout by 8 percentage points for children affected by the shock, setting the size of the young sibship to its average. The magnitude of the effect becomes larger when I restrict the analysis on girls only, suggesting that sibling sex composition is a larger predictor of the impact of income shocks on girls' secondary school dropout. The findings are in line with theory predicting that boys constitute an economic asset in rural families during economic crises due to higher potential labor income and higher productivity. Finally, the long-term analysis indicates that drop-out of school results in a substantial loss of human capital on average. In the long run, natural disasters experienced during the first primary shool years reduce the educational attainment of both girls and boys by more than 1 year of schooling. Meanwhile, an additional climate shock during secondary school increases the human capital of girls with relatively more brothers, whereas it does not produce heterogeneous effect for boys.

They are different candidate channels explaining the relationship between income shocks, daughters' education and sibling composition. Responses to shocks to protect family consumption may consist in taking selected children 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. Accordingly, I test the effect of climate shocks on child labor and provide evidence that support for this channel. Boys increase their work in housework, which mainly consists in collecting firewood following shocks. Although I do not observe strong heterogeneous effect on boys' labor, girls with a female-skewed sibling sex ratio increase their labor in agriculture during economic hardship. 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.[4] On the one hand, human capital increases future bride price payments and lets them to smooth consumption throughout life. Ashraf, Bau, Nunn, and Voena (2018) demonstrate that bride price payments are positively correlated with female education in Indonesia and Zambia, suggesting that female education is valuable in the marriage market, in theory (Anderson, 2007). On the other hand, work and marriage ensure immediate benefits but interfere with the accumulation of children's human capital and may lead to

[4]Using the KHDS survey, Corno and Voena (2016) demonstrate that rainfall shocks increase girls' early marriage due to the bride price norm. According to the norm, the groom's family has to give a transfer to the family of the bride at the time of marriage. However, the study does not investigate the heterogeneous effect of climate variability on girls' educational outcomes depending on sibling composition. In addition, my study differs from their paper using a consistent measure of climate shocks in Kagera, based on rainfall and evapotranspiration data.

long-run costs, with lower future earnings potential.[5] I test for the channel of marriage and find that daughters with relatively more younger brothers are less likely to be married in times of income shocks. Though, the effect is observed only for girls with at least 20 years old (i.e., above secondary school age). Similarly, reducing food consumption during bad economic times might result in nutritional deprivation that would affect educational performances. However, I do not find evidence that the results are driven by a deterioriation in children's health.

This research builts on the broader literature that deals with the question of how income shocks affect education choices. Two possible mechanisms involving in a negative productivity shock can be highlighted; 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 shock. In contrast, the income effect may negatively affect children's human capital as households do not have the capacity to pay for schooling. Most papers identify the latter effect. Income fluctuations decrease children's school attendance in India (Jacoby and Skoufias, 1997), in Brazil (Duryea et al., 2007), or in Mexico (De Janvry, Finan, Sadoulet, and Vakis, 2006), educational investments and malnutritional status in Côte d'Ivoire (Jensen, 2000) and human capital accumulation in Latin-America (Caruso, 2017). Although volatile income in an environment of incomplete insurance or capital markets clearly leads to lost opportunities for educational investments, what is less clear is whether some children suffer disproportionately large effects. Indeed, a major strand in the income shocks literature assumes that the

[5] A large economic literature demonstrates in fact that parents' incentives for children's human capital are mainly related to transfers for old-age support in the absence of savings and insurance. Allowing parents to invest differently in their offsprings would let them respond to income shocks while maximizing expected transfers.

children are identical and neglects the heterogeneity assumption. However, income shocks might contraint parents' abilities to pursue altruistic goals for all of their children and impede to invest in equally.

Recent studies though, dissociate impacts by child's gender and find most of the time that girls suffer more from negative income shocks. The outcome reflects gender bias in household resource allocation (Maccini and Yang, 2009) through differences in child labor values across sexes (Björkman-Nyqvist, 2013), cultural preferences (Bandara, Dehejia, and Lavie-Rouse, 2015) or alternatives to girls' schooling. In line with the latter idea, Corno and Voena (2016) find that adverse shocks during teenage years increase the probability of early marriage and early fertility among women in the Kagera region. Beyond their gender, children's education might no longer be independent of family characteristics during economic hardship. According to Butcher and Case (1994), children's education may be affected by the size and composition of his or her sibship if the family faces borrowing constraints.[6] As a result, I empirically test the argument as it is crucial to learn more about the family determinants of resilience to the vagaries of shocks, such as siblings' size and gender composition. Beyond average effects, this study will delve into distributional impacts of income shocks to examine in particular, whether children are more vulnerable to climate shocks due to the presence of brothers relative to sisters, and hence, the consequences of climate shocks for resulting inequality. To the best of my knowledge, no study to date has presented a comprehensive evaluation of sibling composition in mitigating the negative impact of income shocks. Nevertheless, Thomas, Beegleb, Frankenbergc, Sikokid, Strausse, and Teruelf (2004) investigate the consequences of the 1998 Indonesian financial crisis on children's education according to the number of younger and older members in the family.

[6] A number of studies demonstrate that the sex composition of the sibship is of great importance. Vogl (2013) found that relative to younger brothers, younger sisters increase a girl's marriage risk and thereby earlier school-leaving in South Asia. In contrast, Morduch (2000) and Parish and Willis (1993) showed that having sisters is an important predictor of schooling outcomes, especially older sisters who help to care for younger children, take wage employment to pay school fees and postpone child labor for the younger ones. Similarly, Ejrnæs and Pörtner (2004) demonstrate that birth order effects are more pronounced in families where the first-born child is a boy, which goes against the expected pattern driven by parents investing more in sons than in daughters, while for Dammert (2010), effects are independent of the sex of the elderly. Finally, Lafortune and Lee (2014) examine the possibility that a child's schooling can increase with the number of siblings instead of being diminished.

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. In comparison with this research, I elicit parental adaptative strategies following shocks using data on biological siblings instead of current household characteristics, which addresses gender-selective child-fostering, migration, and marriage patterns.

The paper proceeds as follows. Section 2 presents the theoretical framework which relies on theories and qualitative findings. Section 3 describes the data. Section 4 establishes the empirical strategy with a focus on challenges to identification. Section 5 presents the results and Section 6 explores potential mechanisms. Section 7 goes through various robustness checks and Section 8 concludes.

## 2.2 Conceptual framework

This section first exposes economic theories in line with the empirical question. I support the theoretical assumptions using a unique qualitative survey that I realized in 2018 in the second subsection.

### 2.2.1 Theories

The theoretical framework is consistent with a model of one parent, the decision-maker, who optimally chooses the level of education of heterogenous children, maximizing an objective utility function subject to budget constraints.[7] The objective function depends on current consumption and children's development status, due to parental altruism or because children can provide support during old age (Attanasio, 2015). In a first period, parents have to make choices about the level of consumption and investment in children's education. With perfect information and access to credit market, parental educational investment in each child is made until the marginal costs is equal to the discounted expected future benefits (Becker,

[7] Contrary to the collective model of Chiappori (1988, 1992), preferences are assumed to be the same between spouses. Also, the number of children is taken as given. See Ejrnæs and Pörtner (2004) for an analysis of intrahousehold allocation with endogenous fertility.

1992).[8] Parents cannot borrow or save to finance their human capital investments in this model. Children have identical inherent abilities and talent at birth. However, schooling costs and returns differ by child's gender. I assume that girls face a higher opportunity cost in terms of foregone earnings.[9] The focus is on how parents make decisions to re-allocate resources between offsprings following exogenous income shocks. An income shock increases parents' marginal utility of consumption and marginal costs of schooling. Following the loss of income, sibling composition might enter the production function of the child's human capital through its effect on the budget constraint. I distinguish two competing explanations predicting the effect of the shock on a child's education by the family demographic composition.

According to the predictions based on the pure investment model à la Becker and Tomes (1979) or Becker (1992), as the cost of investing in human capital is negatively related to children's endowments, parents would invest more human capital in better-endowed children in periods of income shortfalls.[10] If boys receive a higher return for each level of schooling, I would expect both women's and men's schooling to be influenced by the percentage female

[8]Some papers assume that parents are endowed with imperfect information about educational returns and that schooling decisions are rather driven by parental beliefs about investment returns (Akresh, Bagby, Walque, and Kazianga, 2012, Boneva and Rauh, 2018), which may be inacurrate (Jensen, 2010, Attanasio, 2015). When comparing two of their children, Dizon-Ross (2018) found that one third of parents are mistaken about which child is higher-performing in Malawi.

[9]In a patrilocal society where daughters tend to leave the natal family upon marriage, the expected private returns perceived by the parents might be lower for girls than for boys since parents rely on their sons to look after them in old-age (Strauss and Thomas, 1995, Becker, Murphy, and Spenkuch, 2016). However, Levine and Kevane (2003) find no evidence that the overall pattern of rough equality in the treatment of boys and girls in Indonesia masks differences according to post-marital residential practice. Differences in returns between girls and boys based on labor-market opportunities is another explanation listed in the literature. Even in the presence of differences in labor market opportunities in favor of boys, it does not follow that women would receive less education than men though. If there is positive assortative mating on education, then schooling girls may improve their marriage market prospect and future income (Butcher and Case, 1994, Strauss and Thomas, 1995). Also, Ashraf et al. (2018) consider that bride price payments increase with female education all the same, providing an incentive for additional female schooling. Nonetheless, marriage represents an alternative educational choice in societies engaged in bride price payments (Corno and Voena, 2016, Corno, Hildebrandt, and Voena, 2017).

[10]A wealth effect might induce parents to compensate less well endowed children whereas efficiency induces them to reinforce better-endowed children. Efficiency often dominates for investments in human capital whereas the wealth effect dominates for investments in nonhuman capital as the cost of investing in nonhuman capital is much more independent of children's endowments (Becker and Tomes, 1976, 1979, Yi, Heckman, Zhang, and Conti, 2015). Allowing a second period where parents can compensate less-endowed children reconciliates the investment model (Becker and Tomes, 1976, 1979) with the separable earnings-transfers model where parental aversion to offsprings' earnings inequality plays a role (Behrman, Pollak, and Taubman, 1982).

among siblings in the family during shocks. In line with this theory, Morduch (2000) found in Tanzania that moving from an all-brothers to all-sisters scenario raises completed years of school by 0.44 years in Tanzania. This sets in motion rivalry for scarce resources in which parents favor sons, and children do better with sisters. If climate shocks only exacerbate the binding constraints, I should find similar results with larger magnitudes.

However, there might be counter-balancing advantages to having brothers rather than sisters during agricultural shocks, even when biases favor males. This supposes that girls and boys have different comparative advantage in labor activities. During an income shock, households are forced to look for alternative income generating activities and food and therefore, the demand for children's participation in home production (Björkman-Nyqvist, 2013) or labor raises.[11] The shock might increase labor for the child who has a comparative advantage in the activity generating larger benefits. If I assume that boys are more productive in labor activities and have a larger work capacity, their opportunity costs of time or shadow wages would be greater. As a result, boys might constitute a relevant economic asset in rural families during economic crises. In this case, children would benefit from having relatively more brothers in the family. The greater the share of the child's earnings in the shadow prices, the smaller the contribution of each child to the family income. The argument does not necessarily implicates that boys' education will be deeper affected by the shock in a society where child work and schooling are not totally incompatible (Serra, 2009). Indeed, competition between work and school is less rigid than typically argued since the higher probability of work does not displace school enrollment (Kruger, 2007). All in all, the influence of sibling gender composition during an income shock is ambiguous.

In addition, regardless of his gender, the presence of a sibling might help to share the burden associated to a negative income shock, since a new member can enter the workforce and bring extra resources into the household (Parish and Willis, 1993, Lafortune and Lee, 2014). Nevertheless, siblings in large families might reduce the amount of resources available to each children due to more competition for scarce resources (Black, Devereux, and Salvanes,

[11] The argument holds only if the income effect dominates the substitution effect (Shah and Steinberg, 2017).

2005, Ponczek and Souza, 2012). I derive the following testable predictions:

a) If boys' expected returns to schooling are still greater than their potential labor income, children with relatively more brothers than sisters might do worse during economic hardship.

b) Climate shocks might have a non-linear effect along sibship size.

The age of siblings might also be important. On the one hand, productivity and wages rise with child age (Fafchamps and Wahba, 2006). Consequently, I expect older children to work more than young ones during family income shortfall. Older brothers (i.e., with complete education) might bring relatively more resources into the household and help the family to mitigate income risks compared to older sisters since males are endowed with higher earning abilities and labor market opportunities.[12] Similarly, older girls (unmarried and living in the household) may have to spend time on housework if there are young children present. The argument is in line with Lafortune and Lee (2014) where parents being not able to pay for children's schooling rely on some of their older children's labor income as a source of funding.

On the other hand, existing evidence indicates that returns to education are non-linear (Kuepie et al., 2009), in Tanzania in particular (Söderbom, Teal, Wambugu, and Kahyarara, 2006, World Bank, 2004a). Wedgwood (2007) highlighted a threshold effect with benefits (higher incomes) being more pronounced for those with complete secondary education. On this basis, as households have already invested in the schooling of older children who are close to the completion of a particular grade, it might be more prudent to continue to protect education-related expenditures for older children and keep them in school during a shock. If the returns to investment in education of the younger child have greater variance than some alternatives (i.e., the investment in older siblings' education), the opportunity cost of investing in the younger child increases and the incentives for such investment fall (Behrman, 2010).[13] Propositions:

[12] The argument constrats with Rosenzweig and Schultz (1989) where the marriage of daughters to locationally distant, dispersed yet kinship-related households is a manifestation of implicit interhousehold contractual arrangements facilitating consumption smoothing in India.

[13] This is consistent with the paper of Akresh et al. (2012), where sibling ability rivalry exerts a strong impact on the parents' decision of which child to send to school. While a child with a higher-ability

c) The size of the young/old sibship might play a role in determining the differential effect of drought within the family.
d) If we assume that expected returns of boys exceed those of girls for any educational grade, a child with relatively more younger sisters is more likely to stay in school in times of shock. The argument holds only if the productivity in child labor (agriculture or household labor) does not raise exponentially with age.

Finally, due to differences in educational returns between boys and girls, child's gender might also be an important predictor of the impact of shocks on educational attainment. According to the literature, I would expect that girls' education to be largely affected by climate shocks compared to boys (Bandara et al., 2015, Björkman-Nyqvist, 2013, Maccini and Yang, 2009, Skoufias and Parker, 2006). Similarly, heterogeneous effects of climate shocks according to siblings' characteristics might be more pronounced among girls since women's educational choices are often affected by her siblings, and that men's choices have not (Butcher and Case, 1994).

In the following sub-section, I test the assumptions made in this theoretical part using a qualitative survey.

### 2.2.2 Qualitative survey-based

The main aim of the qualitative study that I realized in June-July 2018 is to test the empirical hypothesis: (1) income shocks induce trade-offs between offsprings regarding their investment in education; (2) educational returns differ by the child's gender (i.e., given (1), the impact of income shocks depends on the gender composition of siblings); (3) educational returns are non-linear along the age distribution (i.e., given (1), the composition of younger/older siblings matters during economic hardship).

sibling is 15% less likely to be currently enrolled, having two higher-ability siblings lowers a child's probability of enrollment by 30%.

The qualitative study is based on focus group discussions conducted across 4 districs in Kagera. On the basis of budget and time constraints, I first selected districts around the city of Bukoba, the Kagera's regional capital. They are Bukoba Rural, Muleba, Missenyi and Bukoba Urban. They represent 47% of the Kagera's population according to the 2012 national census. In each district, I classify all villages by their level of development: very remote, remote, urban and very urban.[14] I then randomly selected one village in each category.[15] See Table B.2 in the appendix for an exhaustive list. Finally, I interviewed a group of 6 women, then a group of 6 men and a mixed-gender group, randomly selected among parents involving in agriculture activities with at least one child enrolled (or has been enrolled) in primary or secondary school. Qualitative information was collected through an interactive discussion by a collection of the 6 participants and two facilitators as one group in one place. Overall, the sample consists of 270 individuals interviewed in 15 villages of Kagera.

Before testing the hypothesis, I identified drought as the most common cause of severe crop losses. Among the resulting coping mechanisms, interrupting children's education, advising them to fail at the exam to stop their education or reducing educational expenditures are listed in more than 40% of the groups (Table 2.1, Panel A). Looking for an additional work, mostly for women and children, is the most frequent statement coupled with reducing the quantity and/or the quality of food consumption. Both of these mechanisms alter children's school attendance and performances.

I first test the assumption that educational returns vary across girls and boys. Should that occur, it would provide insights on the importance of the child's gender and the sibling sex composition in an analysis of income shocks. The qualitative findings show that patrilocal exogamy, where parents anticipate larger transfers from their educated sons in times of need, does not explain gender differences in human capital investments (Table 2.1, Panel B). "In the past, we thought investing money on girls was not profitable since they left the household and contributed to another family. Mentalities have begun to change when people from the

[14] I thank you the members of each district council for their hospitality and collaboration. I am grateful to Education Development Initiatives (EDI) for their help on the survey design.

[15] Notice that only three villages were selected in Bukoba Urban since no "very remote" village was identified.

village discovered that families who sent their daughters to school received more help and transfers than families with educated sons." Indeed, the rapid adoption of technology in Sub-Saharan Africa facilitates communication among social networks in response to shocks (Aker and Mbiti, 2010).[16] "Even if girls are living far away from the parents, everybody has a mobile phone nowadays." Moreover, "the emergence of women's rights to own property, access productive resources and family gender awareness training delivered in rural areas by NGOs contribute in reducing perceptions that daughters are less economically valuable than sons" (common statements).[17] Gender preferences, though, might still remain in certain families as "the girl cannot help her parents without asking her husband's consent as she remains under his control" (a claim made in a group of men).

In addition, Panel E gives information on the child providing the largest assistance in farming or in the household. Although, overall, "all children help and are treated equally in both type of labor", 32% of the groups indicated that "girls are actually providing the largest assistance", suggesting larger opportunity costs on girls' education in terms of foregone earnings. In addition, while 77% of the groups did not mention any gender division of child labor, girls work more in the household and boys tend to have a comparative advantage in agricultural production (livestock and firewood).

The following exercise gives more insights on the existence of both differential treatments by child's gender and competition between siblings during income shocks. I constructed the hypothetical scenario where parents have to choose to invest in one child out of two, who only differ by sex. The unprompted answer that a girl can get pregnant and be expelled from school is reported in 30% of groups (Panel C). "If you invest in a girl, at the end of the day, she can misbehave and becomes pregnant. Then, she drops out secondary education, unable to complete form four and you lose your investment. Investing in a girl is thus more risky than investing in a boy."[18] The phenomenon would be even stronger in

[16]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' support during old-age.

[17]This is consistent with Allendorf (2007) and Fernández (2014).

[18]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

poorer families. "Long distances to school force female students to pay transport services with their body in the lack of family resources. Besides, droughts and floods lead to a lot of pregnancy as girls tend to look for diverse options to increase liquidity so as to help their parents."

In the main, the qualitative survey corroborates the hypothesis: marginal costs differ by the child's gender because alternatives to schooling are larger for girls than boys and girls' academic achievement is more uncertain; siblings are interdependent for limited family resources during adverse events.

Secondly, I test the hypothesis that educational returns are non linear, suggesting that the composition of younger/older siblings is a better predictor of the effect of income shocks. In particular, I test whether parents allocate different resources on younger and older children at the time of shocks. To do so, I constructed an hypothetical scenario where parents had to select one child among two children who would benefit from scarce resources in times of needs. In this scenario, the two children differed only by the number of school years already completed. I find that parents perceive late investments (children enrolled in secondary school) to be significantly more productive than early investments (children enrolled in primary school) with the secondary child selected in more than 80% of groups (Panel D). This is consistent with the results of Boneva and Rauh (2018). As a matter of fact, "families have already spend a lot of money on the child in secondary school who is near to finish form 4, associated with more income and more opportunities to go out the village." Besides, "the child in secondary will bring money sooner to help siblings who will still lagging behind". This is consistent with high-ability children being more likely to be enrolled and receiving more educational resources. As a result of lower uncertainty about the abilities of children with more education who already passed some exams, families associate years of schooling with greater potential and future performances at the time of shock. Finally, in terms of children's alternatives to

from school and prevented from returning. Through the policy, the government's goal is to stop "bad behaviors" and reduce teenage pregnancies. According to the 2015 - 2016 Tanzania Demographic and Health Survey, one in four women aged 15 - 19 are mothers. In practice, nurses are checking girls at school with the administration of pregnancy tests and teachers are denouncing girls under government pressure.

**Table 2.1 –** Qualitative results (statistics)

| | Freq. | Percent |
|---|---|---|
| Panel A: Risk-coping mechanisms | | |
| Doing extra labor (including children) | 35 | 77.8 |
| Reducing food consumption | 36 | 80.0 |
| Selling livestock, timber, storage, property | 16 | 35.5 |
| Asking for money or loan | 19 | 42.2 |
| Interrupting / reducing educational expenditures | 19 | 42.2 |
| Using storage | 2 | 4.4 |
| Planting drought-resistant crops | 7 | 15.5 |
| Reducing health expenditures | 11 | 24.4 |
| Marrying daughters | 4 | 8.9 |
| Extra money from prostitution of women | 2 | 4.4 |
| Moving to another village | 2 | 4.4 |
| Panel B: Selection of children's human capital during shocks, by gender | | |
| Girls | 3 | 7.3 |
| Both gender | 38 | 92.6 |
| Panel C: Unprompted answers on girls at risk of becoming pregnant | | |
| Mentioned | 15 | 33.3 |
| No mentioned | 30 | 66.6 |
| Panel D: Selection of children's human capital during shocks, by educational grade | | |
| Primary child | 1 | 2.3 |
| Secondary child | 36 | 83.7 |
| No concensus | 6 | 13.9 |
| Panel E: Child labor (assistance to parents) | | |
| All children | 19 | 43.2 |
| Girls work more | 14 | 31.8 |
| Older children work more | 11 | 25.0 |
| *Division of labor by child's gender* | | |
| Girls help only in household work | 7 | 16.0 |
| Boys help only in livestock & firewood | 10 | 22.7 |
| No division of labor | 34 | 77.3 |

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 (has been enrolled) in primary or secondary school.

schooling by age or birth order, only 25% of groups reported that eldest children provide the largest assistance in the family, working more than the younger ones. The small proportion suggests no great evidence on substitution of younger for older siblings in the market and domestic work. In sum, focus groups discussions suggest that educational returns are lower in primary than in secondary school and siblings are in competition for limited resources in times of shock.

The quantitative analysis is now discussed in details within this economic environment.

## 2.3 Data and descriptive statistics

This study focuses on the Kagera region. The region is located in the northwestern corner of Tanzania near the lake Victoria and neighbors Uganda, Rwanda, Burundi. For a total area of about 28,000 square kilometers, Kagera represents the fifteenth-largest region of the country. Kagera experiences a bi-modal rainfall pattern (March-May are the long rains or *Masika*, and October-November rains are the short rains or *Vuli*). Agriculture is the major economic activity, for most, subsistence farming, with limited use of wage labor. Rural household owns only 1.2 ha on average, below the national average of 2.0 hectares. The main cropping system is the traditional method, which combines cash crops (bananas and coffee) with annual food crops such as maize, beans, cassava, sweet potatoes and sorghum in certain areas. Most of households' cash income comes from annual crop farming activities followed by permanent crop farming (see Table B.2 in the appendix). Fishing and livestock farming are culturally important, but despite their potential they contribute little to the region economy and remain underdeveloped. The industrial base of Kagera region is still very small: the most important industries are those involved in the processing of cash crops such as coffee and cotton (United Republic of Tanzania, 2012).

In this study, I use information from the Kagera Health Development Survey (KHDS), a unique 19-year panel survey conducted by the World Bank and the University of Dar es Salaam. Originally, the main purpose of the KHDS was to measure the impact of adult mortality and morbidity on the welfare of individuals and households, given that HIV infection rates among adults were as high as 24% in the regional capital of Bukoba in the 1980s (Beegle, Dehejia, and Gatti, 2006). The survey interviewed 915 households living in 51 villages for the first time in 1991, then three times at six-monthly intervals during 1991-1994.[19] The KHDS 2004 and 2010 attempted to re-interview all 6,353 original household members, irrespective of whether the respondent had moved out of the original village, region, or country, or was residing in a new household. By 2010, at least one person were interviewed in 92% of the initial households. Precisely, 68% (4,336) of the original respondents were interviewed, 20% had died and 12% were untraced (De Weerdt, Beegle, Bie Lilleør, Dercon, Hirvonen, Kirchberger, and Krutikov, 2012).

A concern inherent to empirical studies relates to whether households exit (or enter) the sample in a way that is correlated with climate shocks or children's education. The attrition rate in the KHDS, though, is exceptionally low compared to other panel surveys. In addition, I conduct robust estimates using information on educational attainment available for all original children, irrespective of their location or their mortality status in 2010.[20] A second concern is that the survey design included an over-sample of households with a sick adult member or households that had experienced a recent adult death in 1991. It is possible that those households respond differently to shocks and over-estimate my results. However, the paper focuses on heterogeneous income effects according to sibling composition. The sibling module, which provides a roster of the children ever born, restricts the analysis on families in which both parents were alive at the baseline (see sub-section 3.1.2 for details), probably less affected by adult mortality.[21]

### 2.3.1 Main variables of interest

[19]The sample selection was drawn in two stages. In the first stage, based on the 1988 Tanzanian census, the census clusters were randomly selected after stratifying them according to mortality rates and agro-climatic zones. Households were then divided into high and low adult mortality rates, based on illness and death of households in the 12 months before the enumeration process. Finally, households were randomly sampled from the groups (World Bank, 2004b).

[20]If the individual was not found or he/she had passed away, information on the highest grade completed was provided from the most knowledgeable person of the individual (parents or siblings).

[21]As a result, the analysis of the remaining households would understate the heterogeneous effects of income shocks according to sibling composition since monoparental families (or families with high adult mortality) might overreact to the income shock.

#### 2.3.1.1 Education data

The survey contains detailed education data on household members aged 7 and above, such as, if they were ever enrolled in school, their school attendance in the last week, the highest grade they completed in each of all six survey rounds, and the time it took to reach school from home. For the main analysis, my primary data source is the last wave of the Kagera Health Development Survey (KHDS) realized in 2010. Particularly, the survey asks to all previous household members and their children “At what age did you start your schooling? How old were you when you completed your highest grade?” The availability of retrospective information allows me to construct a panel of individuals from their age of entry into primary school to 2010.

The Tanzanian educational system operates on the 7-4-2-3 system: 7 years of primary school, followed by 4 years of secondary school (Ordinary Level) leading to 2 years of Advanced Level. After the 13th year of secondary school students may take the Advanced Certificate exam and attend college for 3 to 4 years. In my sample, children attend school from 8.7 to 15.5 years old on average and complete the primary level (Table 3.1). The difference in educational attainment by gender is small (6.8 years of schooling for boys, 6.6 for girls). Though, 10% of girls do not have completed any education against 5% of boys in the main sample. One striking fact is the high proportion of children (70%) who complete only 7 years of primary school. Most children are unable to continue schooling beyond primary school, with the fact that secondary school students are only taught in English and the possiblity to be legally employed at 15 years old (Alam, 2015). Around 9% of children complete lower secondary education and less than 3% pursue study after standard 4.

In this study, I measure secondary school dropout and focus on individuals who leave school before they have completed standard 4 (the lower secondary level corresponding to 11 years of education), regardless of age.[22] Accordingly, I define a dummy variable equals

[22] I consider all individuals aged between 6-22 years old in the main analysis and exclude 3% of children who ended their education beyond 22 years old. In the robustness part, I tighten the definition and consider a secondary school dropout as any leave before the final year of standard 4 among secondary school age children (i.e., less than 18 years old at the drop-out). Primary drop-out school is hardly testable given the low prevalence rate of drop-out before standard 7. Using a dummy variable for a primary drop-out, the variable takes the value of 1 for only 2.23% of my observations.

to 0 during school years and to 1 when dropping out of school, after which the individual is excluded from the sample. Given the year of birth and the year of entry into primary school which vary across individuals, my data ranges from 1975 to 2010. 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 first limit with annual school dropouts as the outcome is the exclusion of individuals who never went to school. This concerns 8% of the 1,897 individuals who complete their education by 2010. 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. Secondly, the constructed measure does not take into account temporary interruptions in children's schooling. Indeed, income variations may affect school attendance without stopping education definitely. In this way, I reproduce the analysis on the six waves of the KHDS using a dummy variable equals to one if respondents attended school during the 12 months preceding the survey in section 7. Thirdly, the construction of the main dependent variable relies on children's birth dates, school starting and ending age. Recent papers, such as in Larsen, Headey, and Masters (2019), alert on the need to understand any biases introduced by errors in accurate birth dates data, especially for analysis of month-to-month changes in early childhood. In my setting, main variables are constructed on a year-to-year basis to analyze the effect of unafourable climate conditions on permanent school dropouts in a given year. Year of birth suffering less from self-reporting bias. In addition, a particular concern is wether nonresponse rates for school starting and ending age is correlated with the individual's level of education. Indeed, the actual age of children (or age during school years) may be unknown because of low numeracy and literacy, lack of birth registration, and limited celebration of birthdays or regular use of conventional calendars. I account for the latter and investigate the long-term effect of drought on academic performance using years of schooling as the dependent variable in section 5.3. Nevertheless, I acknowledge a source of measurement error as a limitation of the analysis and allow results to suffer from attenuation bias.

**Table 2.2 –** Descriptive Statistics

| Variables | Obs | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| **Panel A: constructed data from retrospective information in KHDS-3** | | | | | |
| Shock | 8,377 | 0.55 | 0.42 | 0.00 | 1.77 |
| Shock (0/1) | 8,377 | 0.16 | 0.36 | 0 | 1 |
| Father is alive | 8,377 | 0.93 | 0.25 | 0 | 1 |
| Mother is alive | 8,377 | 0.96 | 0.17 | 0 | 1 |
| Sex ratio* | 8,377 | 0.49 | 0.24 | 0 | 1 |
| Number of siblings* | 8,377 | 5.75 | 2.39 | 1 | 10 |
| Number of younger siblings* | 8,377 | 2.75 | 2.11 | 0 | 7 |
| Number of older siblings* | 8,377 | 2.82 | 2.26 | 0 | 7 |
| Probability of school dropout | 8,377 | 0.12 | 0.32 | 0 | 1 |
| Age | 8,377 | 12.46 | 3.01 | 6 | 22 |
| Age of school entrance* | 8,377 | 8.78 | 1.64 | 4 | 17 |
| Age of school dropout* | 8,377 | 16.13 | 2.39 | 6 | 22 |
| Education* | 8,183 | 7.30 | 1.82 | 0 | 11 |
| **Panel B: KHDS surveys** | | | | | |
| School attendance | 2,251 | 0.68 | 0.47 | 0 | 1 |
| Hours in school last week | 799 | 22.80 | 15.89 | 0 | 50 |
| Hours in agriculture last week | 2,084 | 8.53 | 10.35 | 0 | 62 |
| Time to school (minutes) | 1,003 | 22.13 | 14.13 | 0 | 60 |
| Distance to school (km) | 1,410 | 35.50 | 120.35 | 0 | 900 |
| Probability to be sick | 2,446 | 0.38 | 0.49 | 0 | 1 |
| Farming | 6,063 | 0.99 | 0.12 | 0 | 1 |
| Business outside farming | 6,159 | 0.42 | 0.49 | 0 | 1 |
| Livestock | 6,061 | 0.85 | 0.36 | 0 | 1 |
| Number of shambas | 5,514 | 2.52 | 1.39 | 0 | 5 |
| Materials of main dwelling | 6,254 | 0.73 | 0.44 | 0 | 1 |
| Food expenditures pc | 6,209 | 212245 | 154179 | 7079 | 2125136 |
| Non-food expenditures pc | 6,263 | 107275 | 309634 | 3232 | 1.13e+07 |
| Consumption pc | 6,209 | 320161 | 410135 | 20709 | 1.34e+07 |
| Household size | 6,263 | 7.36 | 3.19 | 1 | 27 |

Panel A describes the variables used in the main analysis exploiting the panel constructed from retrospective information in 2010. The time span is determined by school years, which vary between 1956 and 2010 for each individual in the sample. Variables indicated with a (*) do not vary across time. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Shock (0/1) is a dummy variable equals to one for any extreme climatic events (Spei in absolute values $\geq$1). Education designates the educational attainment of children who complete their education in 2010. Panel B describe variables using different waves of the KHDS. Farming is a dummy equals to one if the household is engaged in farming activities. Similarly, business outside farming equals to one if the household holds a buisiness outside agriculture. Livestock equals to one if the household raises livestock. Material of main dwelling equals to one if walls of the main dwelling are constructed with mud or bamboos and zero for stronger materials (ciment, bricks etc). pc: per capita.

#### 2.3.1.2 Sibling's characteristics

The analysis in this paper is based on comparing the effect of climatic shocks on children with different number of brothers and sisters. For that reason, the sample is limited to individuals with at least one sibling. Only 1% of children in the sample have no siblings. To construct variables on sibling composition, I use the 2010 sibling 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.[23] 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 which addresses gender-selective child-fostering, migration, and marriage patterns.

First, I focus on indicators reflecting gender composition within siblings. I determine sex ratios as the number of brothers divided by the total number of siblings. Thus, I assess the effect from having a higher proportion of brothers among the relatives. To allow differences in the effect of having older and younger siblings during income shocks, I disaggregate the analysis by sibling subgroups. Analogously, sex ratio among younger (older) siblings is the number of younger (older) brothers divided by the total number of younger (older) siblings. Finally, as gender ratios might not be independent of family size, I construct variables which measure the number of siblings (from 1 to at least 10), the number of younger siblings (from 1 to at least 7) and older siblings (from 1 to at least 7). As reported in Table 3.1, children have on average 6 siblings of balanced gender. Figure 2.1 represents the distribution of the number of sisters and brothers, younger and older, relevant for the interpretation of the effects in section 5.

Exploiting data on school-leaving age and siblings' characteristics, my final sample consists of 1,105 individuals (549 males, 556 females) from 376 families. The temporal dimen-

[23] Children in monoparental families or orphans are not considered. This exclusion may lead to a downward bias in my estimates as their families may provide specific characteristics, such as HIV prevalence and adult mortality rates, which may increase their vulnerability to exogenous income shocks.

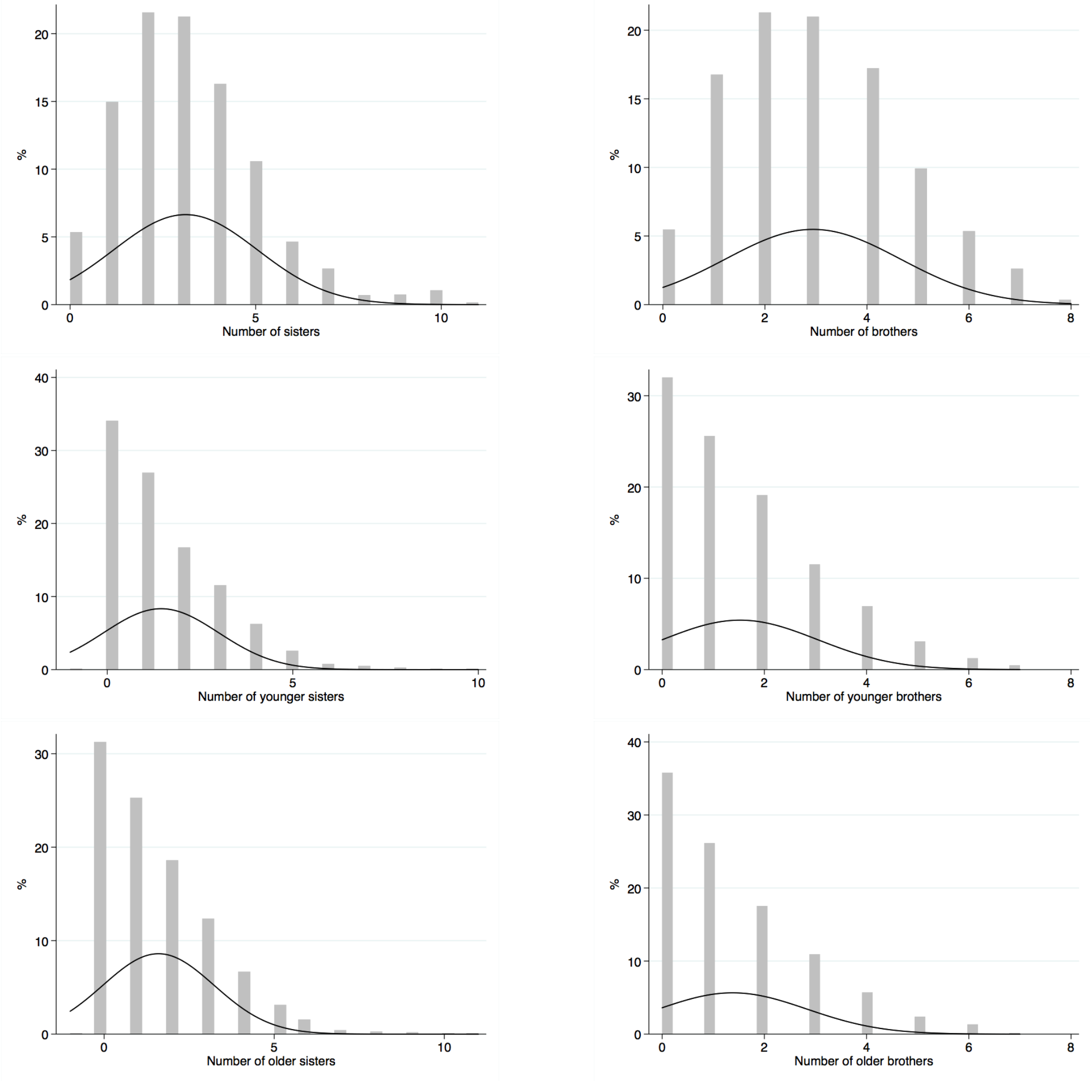



**Figure 2.1** – Family composition: distribution of the number of sisters and brothers, younger and older

sion varies across individuals, spanning from the beginning to the end of their schooling. Finally, the main sample includes 8,377 observations.

### 2.3.2 Climate shocks and income variation

This paper shares with a smaller set of other papers a strong concern with measuring climate variability, a reduced form instrument for household income variability at the village level. I use a more reliable and consistent measure of shocks with the Standardized Precipitation-Evapotranspiration Index (Spei) developed by Vicente-Serrano, Beguería, and López-Moreno (2010). The advantage of drought indices like the Spei is that they are not only based on precipitation but also consider potential evapotranspiration. The growing cycle of a plant does not depend only on the rainfall quantity but also on the evaporation of water from the soil and the transpiration of crops. Temperature have been shown to better explain spatial and temporal variation in agricultural income (Dell, Jones, and Olken, 2012, Lobell, Schlenker, and Costa-Roberts, 2011). The Spei uses the monthly difference between precipitation and potential evapotranspiration to estimate a probability density function from a log-logistic distribution. The function is transformed to a standard Gaussian variate with zero mean and standard deviation of one that expresses the water balance in units of standard deviations from the long-run average (calculated over the period 1901-2015). A value of zero means that the water balance is exactly at its long-run average; a value of plus one (minus one) indicates unusual high (low) water balance relative to what is normally experienced in a particular location. The fundamental strength of the index is that it is able to detect unusual climate conditions at different time scales to identify climate shocks of different variation and magnitude (Harari and La Ferrara, 2013, Almer, Laurent-Lucchetti, and Oechslin, 2017).[24] The time scale over which water deficits accumulate becomes important and separates the different types of droughts into meteorological (1 month timescale), agricultural (3–6 month timescale) and hydrological droughts (12 month timescale) (Vicente-Serrano et al., 2010, Homdee, Pongput, and Kanae, 2016).[25]

As the purpose of the current study is to estimate the effect of negative income shocks, I focus on agricultural shocks encompassing both drought and flood, and use the Spei in absolute value at the six-month time scale.[26] Because the main hypothesized channel linking climate shocks to educational investments operates through shocks to agricultural incomes, I need to isolate the component of climate variability that is relevant for agriculture. I follow Harari and Ferrara (2018) and measure climate conditions during the growing season of

[24] The Palmer Drought Severity Index (PDSI) is also widely used in the literature on climate shocks, based on a simplified water balance equation. However, the PDSI lacks the multiscalar character essential for differentiating among different drought types.

[25] As an example, a six-month time scale means that data from the current month and of the past five months will be used for computing the Spei value for a given month.

[26] Different accumulation periods (the one-month and the three-month time scales) and alternative measures such as dummies equals to one during adverse droughts or adverse wet conditions are tested in section 7.

the local crops, which is when crops are most sensitive to unfavorable conditions. These variables are created based on the rainfall patterns and the specific planting and harvesting calendar for the main crops in Kagera. Annual crop production takes place during the two rainfall patterns. The Vuli planting season is around October - November, with harvesting in late January - February; the Masika, or main planting seasons, starts in late February - March, with harvesting in July - August (World Bank, 2015). Accordingly, the climate indicator is computed by averaging monthly Spei over the growing season months in Kagera (March, April, May and October, November, December). The variable involves both spatial variation and within-year variation in the timing of shocks.

Before proceeding, one important caveat needs to be discuss. The climate variable does not exploit any spatial variation in crop cover, depending on the amount of the commodity produced in each municipality or based on the smallholder farmers' cropping patterns (Arndt, Farmer, Strzepek, and Thurlow, 2012, Imbert, Seror, Zhang, and Zylberberg, 2016). Indeed, most crops are still exclusively produced for self-production by subsistence farmers in the Kagera region. The two largest cash crops in the Kagera region, coffee and banana, are intercropping and produced in most of the Kagera districts. This prevents the use of additional variation between coffee and non-coffee areas for instance, such as in Dube and Vargas (2013). Also, it is fairly common to find cash crops grown in association with annual food crops such as maize, beans, cassava bananas or sweet potatoes.[27] Similarly, I do not allow heterogenous effect of climate shocks on income and consumption based on the smallholder farmer's ability to reallocate farm resources between activities. In Kagera, 71% of households were involved in growing crops only and 27% were involved in both crop production and livestock keeping in 2007 with similar farming size on average (United Republic of Tanzania, 2012). With this caveat in mind, I show that the assumption of homogenous shocks does not prevent the existence of sufficient geographical and time variations to estimate my effects.

The panel nature of the data allows me to track where the parents lived when the respondent was a child rather than the respondent's current location, which helps to identify climate

[27] See Table B.3 in the appendix.

shocks experienced during school years. The climate variable, *Shock*, is then linked to the 51 baseline villages of the Kagera survey through GPS coordinates during the school years of each individual. As the Spei index is gridded by longitude and latitude lines with a degree of precision of 0.5 (approximately 50 km), the location of households selected in the KHDS database results into 12 different grids with different coordinates. It is often admit that yearly averages from limited weather station data, over small period of time, fail to capture large spatial heterogeneity in climate shocks. However, using quite long time periods, my main sample matches up to 12×35(years) unique grid cells across the region. My source of climate variation is thus at the grid cell/year level. In Appendix, Figure B.1 provides maps of the raw Spei data averaged over 10 periods of time in the Kagera region. Lightest shaded areas are the darkest local areas of Kagera and darkest shaded areas the wettest. Figure 2.2 represents the distribution of the Spei index for the 1975-2010 period and its evolution over the decades. The vast majority of communities (19%) have experienced at least one severe dry year (maximum Spei value below -1.5) and 13% have experienced at least one severe flood year (maximum Spei value above 1.5). Finally, Figure 2.3 shows the distribution of the Shock variable (i.e., Spei index in absolute values) used in the main analysis. Shocks are independent across time and location, and do not appear to be auto-correlated. Nevertheless, I address potential spatial autocorrelation through spatial econometric techniques (see Section 4 for details). The average of *Shock* is 0.55 (Table 3.1) and its standard deviation is 0.42. 16% of observations are affected by a climate shock of significant intensity (i.e.*Shock*$\geq$*1)*.

The relationship between climate shocks and agricultural output is well established in the literature. Let me now investigate how my measure of climate shock 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. 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 to droughts, epidemic and flood (in order of importance). 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

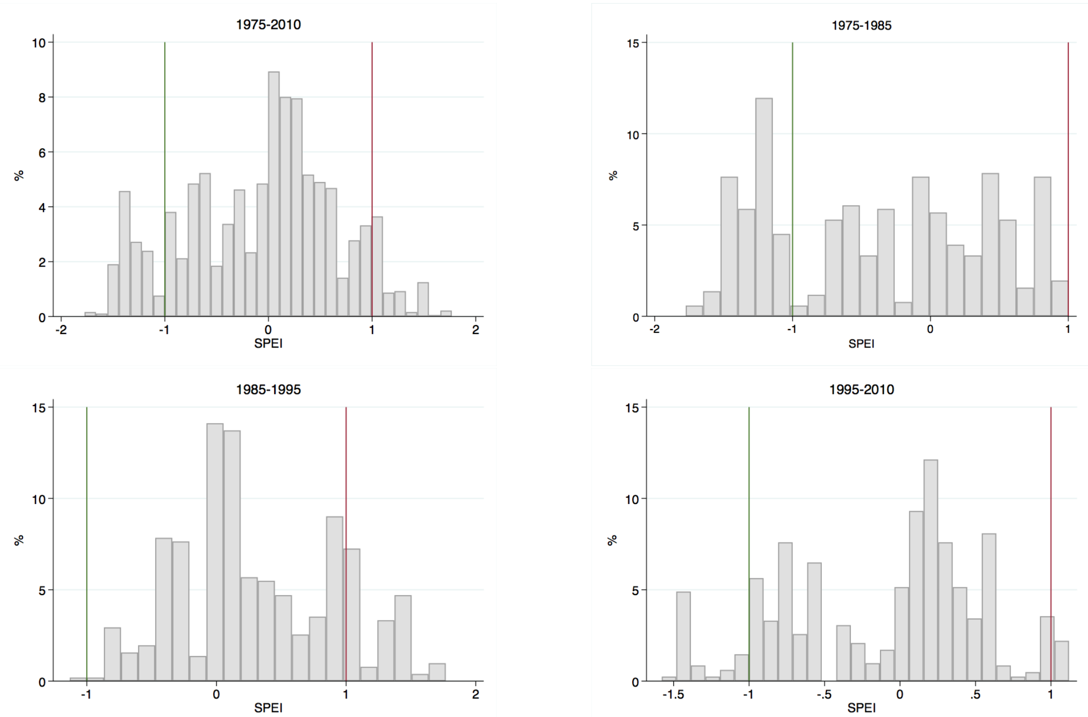



**Figure 2.2** – Spei density for the KHDS communities for the 1975-2010 period.
Notes: A value of 0 refers to what is normally experienced in the particular location, a value of +1 to a significant wet conditions, and a value of -1 to a significant drought.

individuals who participated in focus group discussions in 2018, each respondent confirmed the growing issue of climate change. In particular, droughts and heavy rainfalls are referred in each discussion as the most significant negative phenomenons on harvest.

To validate the latter claim, I test how my measure of shocks affects household consumption, a proxy for household income. Following Corno and Voena (2016), I examine the impact of *Shock* on the natural logarithm of food consumption per capita, non-food consumption per capita and total consumption per capita during the last 12 months of the survey.[28] The recall periods of purchased and home-produced food items were reduced from 12 to six months for waves 2-4 of KHDS 1991-94. The comparable consumption expenditure aggregate therefore only includes wave 1, 5 and 6. All components of the consumption aggregate were deflated using information from the KHDS price questionnaires in 2010. Estimates include survey and household fixed effects, so the impact of drought on household

[28]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. Non-food consumption per capita includes household expenditures on education, health, remittances, durables goods like home or vehicle repairs. Consumption per capita pools both categories.

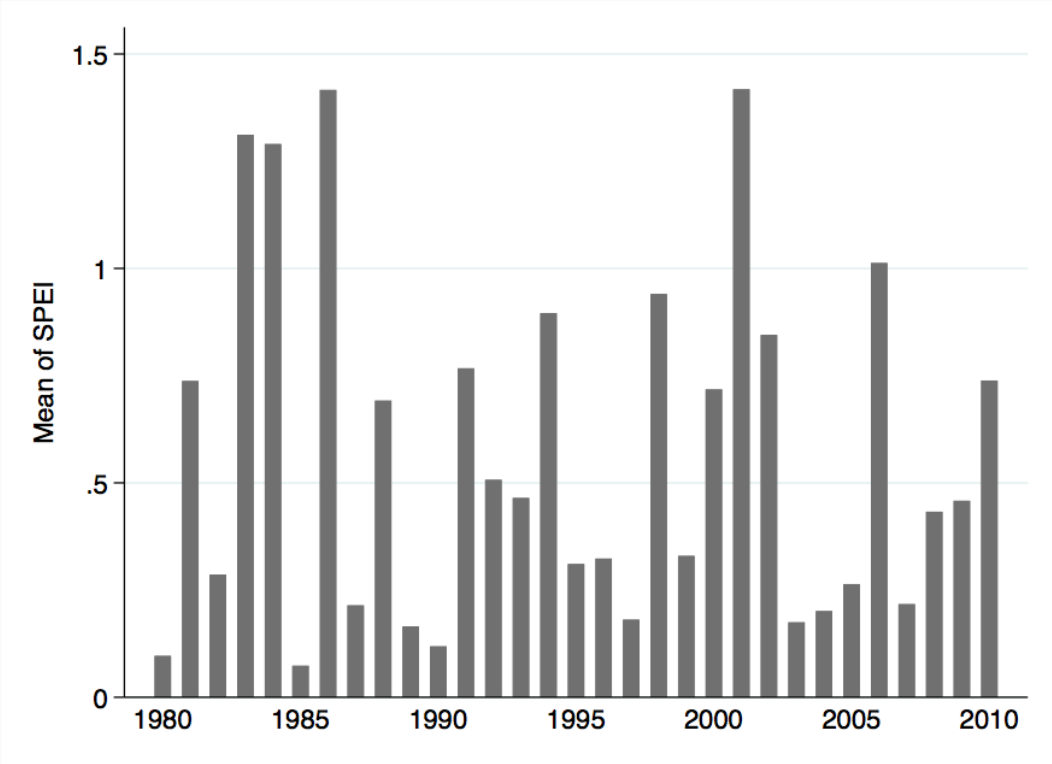



**Figure 2.3** – Distribution of the Spei index in absolute values, by year

consumption is identified between-households and between-survey variation. Climate shock used in baseline estimates is computed during the growing seasons, between March and May and between October and December of a given year. However, consumption is measured over the 12 months before the survey took place, between March and December. Lagged shocks, in period t–1 might be more likely to influence consumption recorded in year t relative to shocks measured in year t. As expected, when I introduce $Shock$ contemporaneously, the indicator becomes insignificant in Table 2.3, except in column 4. Indeed, Shock in time t positively affects non food expenditures per capita. Whereas I do not find a clear explanation for this finding, the contemporaneous shock does not affect global consumption though (column 4). $Shock$ in time t–1 is negatively correlated with consumption per capita, driven mainly by a reduction in non-food expenditures. The coefficient on non-food consumption per capita is statistically significant at 1% level (column 5). Higher values of $Shock$ in time t–1, signaling unfavorable climate conditions, lead to a non-food consumption growth gap of 29%. The magnitude of the effect decreases to 15% when I consider both food and non-food consumption per capita (column 8), given that high values of the indicator does not produce large effect on food-consumption (the coefficient on $Shock$ in t–1 is not significant, column 2).

Finally, I test the results using the first lag of the Spei index, in absolute values, calculated among months out of the growing seasons in Kagera. The identity of households in the Kagera region has implications for the relevant timing of climate shocks. Wage laborers may be most economically vulnerable to early season climate fluctuations, when cropping decisions are made. In contrast, smallholder farmers might be vulnerable to shocks occurring throughout the growing agricultural cycle. As land-owning smallholders represent the vast majority of households in the Kagera region, the yields of the crops cultivated by these farmers might be of primary interest. As expected, climate outside the growing season has no effect on consumption (columns 3, 6 and 9). This suggests that the mechanism operates through low agricultural yields, with cultivation out of the growing season simply constituting a far smaller share of annual agricultural income. If other channels were involved, I would expect to find an effect of climate throughout the year. Overall, these findings suggest that the constructed measure of climate shocks may be a relevant reduced-form instrument for negative income shocks in the region under analysis.

To give additional intuitions, I represent the probability of school dropout among children who experienced unusual weather conditions in the given year ($Shock \geq 1$) compared to those with normal conditions ($Shock < 1$) according to the number of siblings in Figure 2.4 and their sex composition in Figure 2.5. Climate shocks tend to not produce effect on offsprings' school enrollment in relative small families (between 1 and 3 siblings). In larger families (7 and more siblings), the mean of school dropout is actually weaker for children who experienced a negative income shock during the same year. The effect of shock might also diverge by siblings' sex composition, with a mean of school dropout lower for children with more brothers than sisters. The following section goes beyond these stylized facts and presents the identification strategy.

**Table 2.3 –** The effect of drought on household consumption

| Dependent variable | Food p/c | | | Non-food p/c | | | Consumption p/c | | |
|---|---|---|---|---|---|---|---|---|---|
| Specifications | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Shock, in t | -0.022 | | | 0.128** | | | 0.012 | | |
| | (0.050) | | | (0.052) | | | (0.046) | | |
| Shock, in t-1 | | -0.104 | | | -0.292*** | | | -0.156** | |
| | | (0.076) | | | (0.067) | | | (0.064) | |
| Shock out season, in t-1 | | | -0.013 | | | -0.040 | | | -0.036 |
| | | | (0.096) | | | (0.070) | | | (0.085) |
| Observations | 1,844 | 1,869 | 1,844 | 1,844 | 1,869 | 1,844 | 1,844 | 1,869 | 1,844 |
| Number of households | 907 | 915 | 907 | 907 | 915 | 907 | 907 | 915 | 907 |
| R-squared | 0.000 | 0.003 | 0.000 | 0.002 | 0.013 | 0.006 | 0.000 | 0.000 | 0.000 |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors corrected for spatial autocorrelation in parentheses. The dependent variables are measured in logs. Food consumption per capita is the household food consumption during the last 12 months divided by the number of members. Non-food consumption per capita represents exenditures in education, health, durables goods, remittances during the last 12 months, divided by household size. Consumption per capita includes both food and non-food expenditures of the family during the last 12 months. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Shock out season represents the average of the spei index calculated during the months out of the growing seasons, in absolute value. Results are estimated using the survey in 1991, 2004 and 2010. All estimates include survey and household fixed effects.

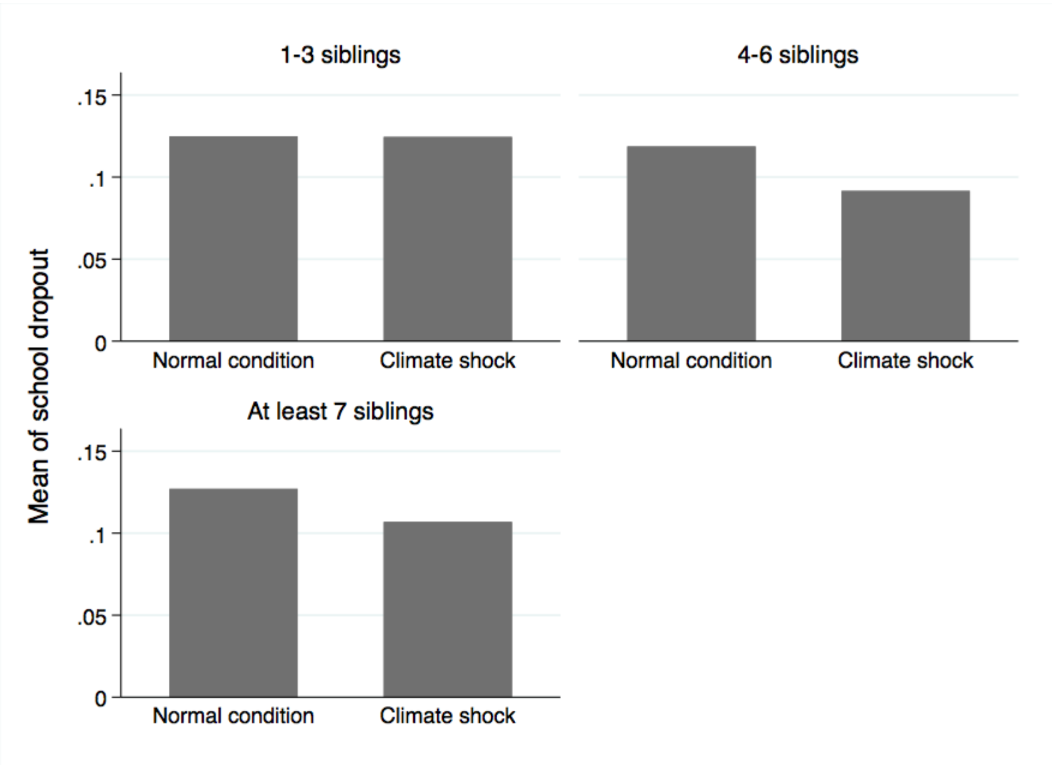



**Figure 2.4 –** Relationship between school dropout and weather conditions, by number of siblings

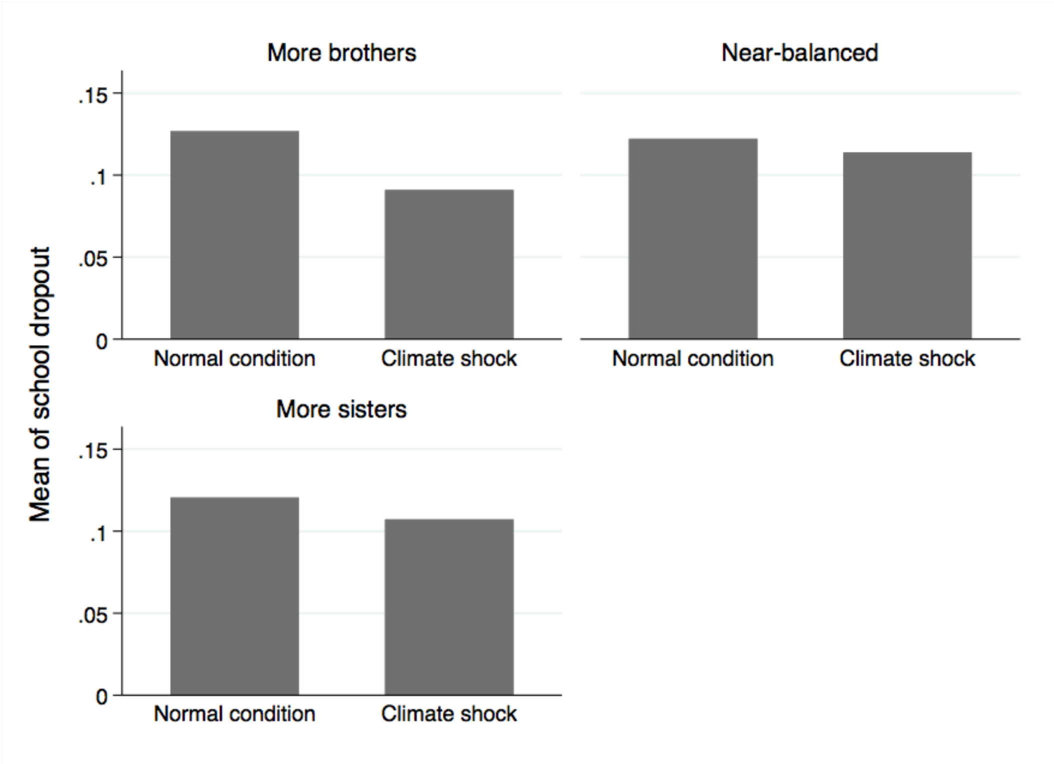



**Figure 2.5** – Relationship between school dropout and weather conditions, by siblings' gender composition

## 2.4 Identification Strategy

My objective is to study the way in which income shocks affect the likelihood of secondary school dropout. In a first step, I examine the impact of income shocks on the hazard into school drop-out using a simple approximation of a duration model as in Corno et al. (2017). I estimate a specification of the form:

$$\text{Dropout}_{i,t} = \alpha_0 + \alpha_1 \text{Shock}_{j(i),t} + \alpha_2 \text{X}_{i,t} + \delta_i + \iota_t + \epsilon_{i,t} \tag{2.1}$$

where $\text{Dropout}_{i,t}$ measures the probability of drop-out of school for individual $i$ in time $t$ and $\epsilon_{i,t}$ is the error term. The dependent variable is a binary variable coded as 1 in the year the child leaves school, and zero otherwise. Thus, the duration of interest is the time between the beginning and the end of the child's schooling. $\text{Shock}_{j(i),t}$ denotes the grid-specific measure of climate. The variable is the Spei drought index calculated during the growing seasons in absolute value. $\text{X}_{i,t}$ represents additional controls such as the father's and the mother's live status. Main specifications include contemporaneous climate shocks, that is, those occurring within the same calendar year in which school dropout is observed. Drought can also have delayed direct effects on education through the timing of economic outcomes. Indeed, negative income shocks may be temporarily offset through drawing down savings or accessing credit, with individuals affecting the education of their offsprings only after these means of smoothing consumption have been exhausted. However, in a context with no access to financial services and lack of assets, households are more likely to react to contemporaneous climate shocks. In the robustness part, I examine the impacts of shocks occurring in the previous calendar year and do not find significant effects.

All estimations control for time $\iota_t$ and individual $\delta_i$ fixed effects. The latter capture time-invariant child and family characteristics that may affect the average likelihood of school drop-out (i.e., abilities, skills, mother's education, place of residence for non-migrant children). Times fixed effects account for any common variations in the school dropout rates in the region over time such as government reforms in school education. Overall, fixed effects help to compare year-to-year variations in the exposure to local conditions that do not follow the secular time trend and are distinct from individual shocks. 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. Nevertheless, the results are robust to alternative estimations such as maximum likelihood (see section 7). Finally, serial correlation within villages and spatial correlation between grids have to be corrected when I estimate the standard errors. Following Conley (1999) and Hsiang (2010), weights in the variance-covariance matrix for standard errors are uniform up to 300 km. In addition, nonparametric estimates of village-specific serial correlation are estimated using linear weights that decay to zero after a lag length of 2. In the robustness part, I test the sensitivity of results to the distance and the lag length cut-off.

In a second step, I investigate the relationship between income shocks, sibling composition and educational outcomes. I explore heterogeneous effects and include interactions terms between $\text{Shock}_{j(i),t}$ and sibling-specific characteristics. I estimate the following equation:

$$\begin{aligned}\text{Dropout}_{i,t} = \alpha_3 + \alpha_4 \text{Shock}_{j(i),t} + \alpha_5 \text{Shock}_{j(i),t} \times \text{Sex ratio}_i + \alpha_6 \text{Shock}_{j(i),t} \times \text{Siblings}_i \\ + \alpha_7 \mathbf{X}_{i,t} + \zeta_i + \lambda_t + \eta_{i,t}\end{aligned} \tag{2.2}$$

where $\eta_{i,t}$ is the error term. The individual-specific sex ratio is the number of brothers divided by total number of siblings. Parameter $\alpha_5$ reflects the heterogenous effect of climate shocks among children with siblings of different sex composition. In doing so, the focus is on the conditional expectation of having relatively more brothers than sisters when the individual is affected by negative shocks. As mentionned in section 2, the sign of $\alpha_5$ is theoretically ambiguous. Assume that an increase of $\text{Shock}_{j(i),t}$ represents an exogenous decrease in family i's income. According to the educational returns theory, having relatively more brothers should increase the likelihood of school dropout if parents reallocate resources to better endowed children following the loss of income. $\alpha_5$ should be positive in this case. On the contrary, if boys' potential income in labor activities is relatively higher, brothers might provide an alternative strategy to smooth consumption (i.e., a reliable asset) during income shortfalls. $\alpha_5$ should be negative in this case.

Sibship gender composition might not be independent of family size and might influence directly fertility behavior. If there is a distribution of tastes over sons and daughters, those parents biased in favor of sons may systematically stop having children after a son but continue to have children if they have a daughter.[29] Calculating a simple correlation

[29] As pointed out by Butcher and Case (1994) family composition may be correlated with parental bias in favor of one sex. Jayachandran and Pande (2017) demonstrate that son preference in India, especially for the eldest son, influences parents' fertility decisions. 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. Whereas son preference is a well-established idea in Asia (Li and Wu, 2011, Kishore and Spears, 2014, Edlund, 1999, Rebeca and Javier, 2016), there is little evidence of fertility patterns in Sub-Saharan Africa (Morduch, 2000, Rossi and Rouanet, 2015) and sex ratios are considered as exogenous in Burkina Faso (Akresh, 2009). Consistent with the listed papers, focus group discussions revealed that most parents have no ideal gender composition in Tanzania.

between climate shocks and measures of sibling sex composition without conditioning on family size would spuriously attribute these differences to gender composition. Thereby, I simultaneously interact $\text{Shock}_{j(i),t}$ with number of siblings (ranges from 1 to at least 10).[30] In equation 2.2, $\alpha_4$ is the effect of unfavourable agricultural conditions for a child with one sister, the excluded category given simultaneous interactions with sex ratio and number of siblings. This concerns 3.25% of individuals in the sample. Then, $\alpha_6$ captures how the shock differs for an individual with at least 2 siblings. Variations in the estimation stem first from exposures to climatic events which vary across grids and time, then from siblings' characteristics which differ across individuals.

Finally, I reestimate separately equation 2.2 across various demographic groups. In particular, I test whether the impact of income shock diverges according to the young and the old sibship. Equations are the following:

$$\begin{aligned} \text{Dropout}_{i,t} = \beta_0 + \beta_1 \text{Shock}_{j(i),t} + \beta_2 \text{Shock}_{j(i),t} \times \text{Young sex ratio}_i \\ + \beta_3 \text{Shock}_{j(i),t} \times \text{Younger siblings}_i + \beta_4 \text{X}_{i,t} + \theta_i + \vartheta_t + \pi_{i,t} \end{aligned} \tag{2.3}$$

$$\begin{aligned} \text{Dropout}_{i,t} = \beta_5 + \beta_6 \text{Shock}_{j(i),t} + \beta_7 \text{Shock}_{j(i),t} \times \text{Old sex ratio}_i \\ + \beta_8 \text{Shock}_{j(i),t} \times \text{Older siblings}_i + \beta_9 \text{X}_{i,t} + \varrho_i + \chi_t + \sigma_{i,t} \end{aligned} \tag{2.4}$$

where $\beta_0$, - , $\beta_9$, are vectors of parameters to estimate and $\pi_{i,t}$, $\sigma_{i,t}$ are error terms. The identification strategy is identical to the one described above. Young sex ratio represents the number of younger brothers over the total number of younger siblings. Old sex ratio is defined analogously. Younger (older) siblings controls for the number of younger (older) siblings with values from 0 to 7. In equation 2.3, I exclude 15% of individuals who do not have a younger sibling. The reference category is a child with one younger sister (30% of individuals), represented by the coefficient $\beta_1$. Similarly, equation 2.4 excludes children who do not have any older sibling (18% of the sample) and $\beta_6$ shows the effect of $\text{Shock}_{j(i),t}$

[30] Concerns regarding multicolinearity due to the inclusion of both sibling gender and sibling size are mitigated in my context, with a pearson correlation coefficient of -0.08 between the two variables.

on the child with one older sister (24.7% of individuals). To easily interpret the findings, variables reflecting siblings' characteristics are first measured separately per subgroup. Then, I include variables both from younger and older siblings in the same regression to estimate which subgroup has the largest effect.

### 2.4.1 Endogeneity of sibling composition

The main threat to the 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. According to the old-age security hypothesis, fertility choices are influenced by the child role of investment-good or household asset (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. Given this, I address endogeneity concerns in multiple ways. First, sibling demographic variables are constructed from all biological children ever born to a mother, currently alive or deaceased and irrespective of their migration experience. In the main analysis, these variables do not evolve over time and reflect final demographic composition.[31] Secondly, I introduce family and year of birth fixed effects to control for parental and neighborhood characteristics that are correlated with child outcomes. Using only between-sibling variation in school dropout (i.e., holding other family characteristics fixed), I rule out the effect of parental preferences if preferences do not evolve across time. However, as pointed out by Shah and Steinberg (2017), women could delay and/or change fertility patterns in response to drought. For example, mothers may choose to wait out a drought year before having a child. Rural fathers could migrate during drought years in search of work and their absence would result in delayed fertility. In Tanzania, Alam and Pörtner (2018) found that crop shocks decrease the likelihood of pregnancies and childbirth. They argue that the postponement of fertility in households who experienced income shocks is likely the result of the increased contraception use rather than a biological response or sep-

[31] In Table 2.16, I allow variables on sibling composition to vary over time calculating sex ratios and sibling size among ever-born children in a given year. The estimation gives indication on the effect of having an additional sibling of a particular gender when affected by adverse climatic conditions despite introducing additional endogeneity bias related to fertility preferences.

aration of spouses. In this way, I follow Jayachandran and Pande (2017) and investigate the issue of selective fertility using a different sample where mothers have likely completed fertility. I assume that women above 45 years old at the time of shock are not of child-bearing potential.[32] Although endogeneity of sibling composition cannot be completely ruled out, I use interactions between an exogenous term, $\text{Shock}_{j(i),t}$, and potentially endogenous variables, $\text{Sex ratio}_i$ and $\text{Siblings}_i$. Interaction terms can be interpreted as exogenous, once the main effect of the endogenous variable is directly controlled for (Angrist and Krueger, 1999). Valid statistical inference can be performed for the interaction term using OLS particularly in cases where validity of instrumental variables is questionable. The identifying assumption is that the endogenous variable and the outcome variable are jointly independent of the exogenous variable (see also Bun and Harrison (2019)).

Finally, having relatively more brothers might be positively correlated with family's income or mother's wealth. In this case, the selection would most likely bias my results downward; since these children do probably better on health and educational outcomes. Indeed, having grown sons is associated with widowed women who have sufficient means to support themselves (Lambert and Rossi, 2016), and has a positive impact on consumption level among divorcees in Senegal (Lambert, Van de Walle, and Villar, 2017). Especially, having a first-born son improves the mother's nutrition intakes (Li and Wu, 2011) and reduces births, which decrease the risk of child and maternal mortality (Milazzo, 2014). Family size and sex composition might thus be less endogenous in families with a first-born son. In this way, I reproduce the analysis restricting the sample to individuals with a first-born brother, who might share closer characteristics. Results are reported in the robustness part.

### 2.4.2 Other threats to identification

Another concern may arise if the location of the individual's household at the baseline survey 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. Se-

[32]In the KHDS fertility module, administered to women aged between 15-49 years old, the maximum age for women who are currently pregnant is 45 years old.

lective migration introduces bias if high (or low) ability migrant-children are systematically less likely to be surveyed during unusual rainfall. It might be the case if children migrated before 1991 since information on age at school departure from the 2010 survey is available only among respondents who have been interviewed at the baseline. Migration for education, though usually at higher levels of schooling, through child fostering is not uncommon in developing countries (Akresh, 2009). On the one hand, a child living away from his biological parents might be more likely to work, might experience psychological problems, or might suffer due to the disruption of living away from his siblings. On the other hand, it is possible these children could benefit both in the short and long-run from the fostering experience by having access to schools, receiving better nutrition, or exposure to an expanded social network (Akresh, 2004). Serra (2009) demonstrates that being fostered to a better-off household may improve the human capital of a child, even if the child works no less than what he would have done back home. Although the inclusion of fostered children might introduce unclear attenuation biases, migration rates are extremely low. Among 1,897 individuals who complete their schooling by 2010, only 155 individuals (8%) moved out of the original village during teenage years. Lastly, the long-term analysis on years of schooling, including all children irrespective of their presence in the household at the baseline, tells the same story.[33]

Then, the number of siblings is correlated with age and birth order, that may have an independent effect on schooling dropout. Indeed, having a large number of siblings reflects a low birth order. To ensure the results are not driven by an age and birth order effect, I allow direct interactions between age, birth order and shocks in section 7.

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 for instance may be equally distributed among children within the same family.

[33]Given that a family member answers basic questions such as educational attainment, for all individuals listed in the sibling roster questionnaire, children do not need to be interviewed in 2010 to be included in the analysis.

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, measurement errors are possible with the use of retrospective information and reported age at school dropout. This would lead 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).

## 2.5 Results

My main results examine the differential effect of income instability, measured with climate shocks, on school dropout. In a first stage, the analysis is conducted according to the number and gender composition of siblings. Specific results using the composition of younger and then older siblings are presented in Sub-section 5.2. Lastly, sub-section 5.3 analyzes the heterogeneous long-term impacts of climate change on children's human capital.

### 2.5.1 Drought, educational attainment and sibling composition

Table 2.4 reports my first set of results. The dependent variable equals one if the child leaves school in time t and zero otherwise. All estimates include individual and year fixed effects, so that the results are robust to any individual or time-invariant characteristics which may influence educational attainment. The estimations are interpreted as the effect of income shocks within an individual's academic life across time.

As a starting point, the specification in column 1 provides an estimate of the magnitude of the effect of the climate shock on children's education. I reject the hypothesis that this effect is unity since on average, Shock does not have a significant effect on secondary school drop-out. This may be interpreted as indicating that my measure of the local shock does not have the same impact on all children. The rest of Table 2.4 focuses on the extent to which individual-specific smoothing is associated with siblings' characteristics. Introducing the interaction term between exposure to shock and sibling sex ratio confirms this heterogeneity (column 2). Shock$\times$ Sex ratio is negative and highly significant. As family gender composition might not be independent of family size, I simultaneously interact shocks with sibling size and sex composition in column 3. The latter represents my best specification (from equation 2.2). The estimation corroborates the finding, the pattern of the shock still appears to be somewhat different for respondents with relatively more brothers than sisters, controlling for the number of biological siblings. A one-standard-deviation rise in the shock variable (signaling unfavourable agricultural conditions) decreases the likelihood of school dropout in a given year by 5% (0.02-0.07) for children with brothers. The marginal effect is significant at the 1% level. To investigate if the differential effect of drought on children with different sibling gender patterns varies across number of siblings, I need to go one step further exploring the three-way interaction between drought, sex composition and sibling size. The estimation is represented in column 3. The significance of Shock$\times$Sex ratio remains, coefficients on Shock and Shock$\times$Siblings become now significant and the three-way interaction is positive and highly significant. To better interpret the results, I represent marginal results of the climate index for children with different number of siblings and sibling gender patterns in Figure 2.6.[34] A difference estimate is displayed along the number of siblings by their gender composition. The pattern of shock is strongly different for children depending on sibling gender imbalances as well as sibship size. Shock increases the likelihood of school dropout for children with only one daughter (sex ratio equals to 0) by 9%, significant at the 5% level. The magnitude of the effect decreases along the number of biological sisters and becomes non significant from 2 sisters (3.25% of individuals in the sample). In contrast, abnormal climate conditions during the growing seasons decrease the likelihood of school dropout for children with only brothers (sex ratio equals to 1) from 9 (one brother) to 4% (6 brothers).[35] In mixed-gender families, shock increases schooling participation for children but only for individuals with at least 6 siblings.[36]

[34]Figure 2.6 is based on the estimation of Table 2.4, column 3 except that standard errors are clustered at the village level instead of using Conley standard errors. This does not alter the level of significance of my variables and thus the following interpretation.

[35]It concerns around 6% of individuals in the sample.

[36]It concerns 32% of individuals in the sample.

**Table 2.4 –** The effect of drought on girls' school dropout depending on sibling composition

| Dependent variable | School drop-out | | | | | |
|---|---|---|---|---|---|---|
| Shock | During the growing season | | | | Out season | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Shock | -0.025 | 0.010 | 0.027 | 0.089** | -0.001 | 0.011 |
| | (0.020) | (0.022) | (0.029) | (0.041) | (0.017) | (0.037) |
| Shock × Sex ratio | | -0.072*** | -0.072*** | -0.195*** | | -0.012 |
| | | (0.026) | (0.026 | ) (0.054) | | (0.037) |
| Shock × Siblings | | | -0.003 | -0.016** | | -0.001 |
| | | | (0.003) | (0.007) | | (0.004) |
| Shock×Sex ratio×Siblings | | | | 0.027** | | |
| | | | | (0.011) | | |
| Father | 0.012 | 0.013 | 0.012 | 0.013 | 0.012 | 0.012 |
| | (0.024) | (0.024) | (0.024) | (0.024) | (0.024) | (0.024) |
| Mother | -0.087** | -0.089** | -0.088** | -0.087** | -0.087** | -0.087** |
| | (0.042) | (0.042) | (0.041) | (0.042) | (0.041) | (0.041) |
| Observations | 8,377 | 8,377 | 8,377 | 8,377 | 8,377 | 8,377 |
| Number of individuals | 1,105 | 1,105 | 1,105 | 1,105 | 1,105 | 1,105 |
| R-squared | 0.001 | 0.002 | 0.002 | 0.002 | 0.001 | 0.001 |

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses, clustered at the village level. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value from columns (1) to (6). In columns (5) and (6), the climate variable is calculated using only months out of the growing seasons. Sex ratio represents the number of brothers over the total number of siblings. The variable increases with the proportion of biological brothers. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. Father is a dummy variable that equals one if the father is alive. Mother is analogously defined. All estimates include individual and year fixed effects.

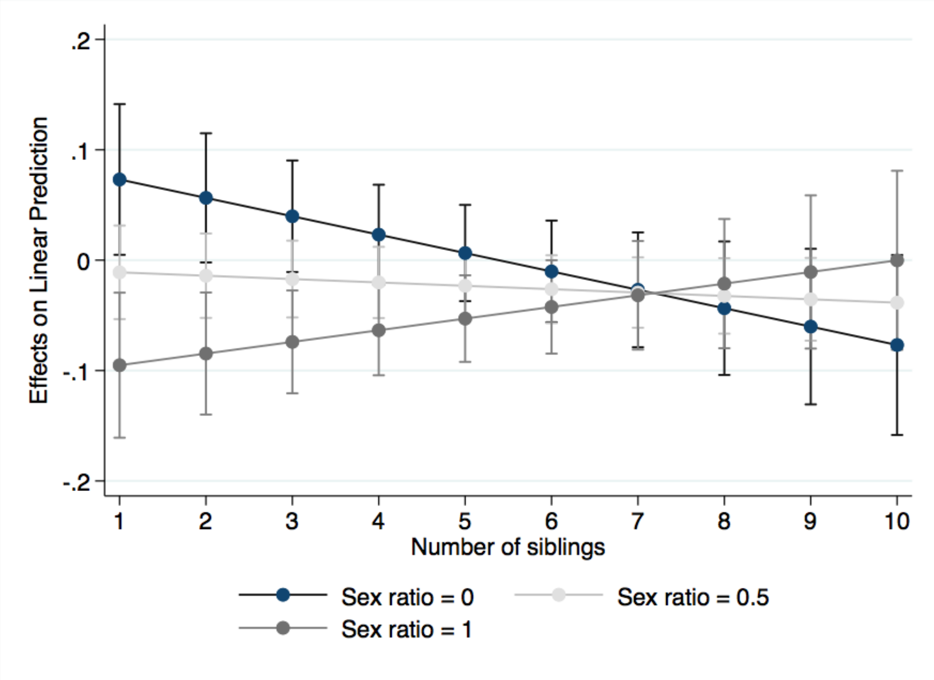



**Figure 2.6** – How the net effect of climate shocks varies with the number of siblings by their sex composition?

Finally, I explore how the results are affected when I measure Shock out of the growing season. Unusual climate conditions out of the growing seasons do not impact school dropout on average (column 5) and do not produce heterogeneous effects according to the sibship composition (column 6). This suggests that the mechanism operates through family income due to low agricultural yields.

According to the predictions based on the pure investment model à la Becker (1992), when market constraints bind, children do fare better when pitted against siblings with fewer intrinsic advantages. In societies with pro-male biases, children with relatively more sisters than brothers would then benefit most. In the current paper, though, relative to brothers, sisters cause earlier-school leaving for children affected by unfavourable dry and wet conditions. Symetrically, in families with only brothers, children do actually better in times of shocks. The results are closer to theories predicting that sons help to overcome economic difficulties in bringing additional resources through higher potential labor income and productivity. Section 6 directly tests the latter explanation.

I pursue my analysis investigating some heterogenous effects of climate shocks on educational attainment across demographic subgroups and by the child's gender.

**Table 2.5** – The effect of drought on school dropout, depending on the younger and older sibling composition

| Dependent variable | School drop-out | | | | | |
|---|---|---|---|---|---|---|
| Shock | During the growing season | | | Out of the growing season | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Shock | 0.009 | -0.023 | 0.001 | 0.030 | 0.020 | 0.042 |
| | (0.028) | (0.022) | (0.032) | (0.026) | (0.023) | (0.035) |
| Shock × Young Sex ratio | -0.080*** | | -0.076*** | -0.037 | | -0.029 |
| | (0.020) | | (0.020) | (0.030) | | (0.025) |
| Shock × Young Siblings | 0.005 | | 0.0046 | -0.001 | | 0.001 |
| | (0.004) | | (0.005) | (0.004) | | (0.005) |
| Shock × Old Sex ratio | | -0.013 | 0.001 | | -0.010 | -0.021 |
| | | (0.020) | (0.021) | | (0.023) | (0.022) |
| Shock × Old Siblings | | -0.005 | -0.002 | | -0.006 | -0.005 |
| | | (0.003) | (0.004) | | (0.003) | (0.004) |
| Father | 0.008 | 0.027 | 0.028 | 0.007 | 0.027 | 0.028 |
| | (0.029) | (0.025) | (0.034) | (0.029) | (0.025) | (0.034) |
| Mother | -0.133*** | -0.077* | -0.131*** | -0.132*** | -0.075* | -0.129*** |
| | (0.043) | (0.042) | (0.043) | (0.043) | (0.042) | (0.043) |
| Observations | 7,040 | 6,819 | 5,482 | 7,040 | 6,819 | 5,482 |
| Individuals | 932 | 901 | 728 | 932 | 901 | 728 |
| R-squared | 0.004 | 0.002 | 0.004 | 0.002 | 0.001 | 0.002 |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Conley standard errors in parentheses. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value from columns (1) to (3). From columns (4) to (6), the climate variable is calculated using only months out of the growing seasons. Young (Old) sex ratio represents the number of younger (older) brothers over the total number of younger (older) siblings. The variable increases with the proportion of biological younger (older) brothers. Young (Old) siblings denotes the number of biological younger (older) siblings, irrespective of their live status or place of residence. Father is a dummy variable that equals one if the father is alive. Mother is analogously defined. All estimates include individual and year fixed effects.

### 2.5.2 Heterogeneous effects across sibling subgroups and by child's gender

Table 2.5 reproduces baseline estimates including the interaction between Shock and demographic characteristics of the young sibship (columns 1, 4 ) and old sibship (columns 2, 5). The focus is on the role of being affected by a negative income shock when having younger or older siblings. Estimations cannot be fully compared to baseline estimates of Table 2.4 since I exclude 15% of individuals who do not have any younger siblings in column 1 and 18% of individuals with no older siblings in column 2.

**Table 2.6 –** The effect of drought on school drop-out, depending on sibling composition and child's gender

| Dependent variable | School drop-out | | | | | | |
|---|---|---|---|---|---|---|---|
| Sibling group | | All | | Young | | Old | |
| | | Boys | Girls | Boys | Girls | Boys | Girls |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Shock | -0.029 | 0.010 | 0.043 | 0.019 | -0.001 | -0.085** | 0.025 |
| | (0.020) | (0.036) | (0.050) | (0.031) | (0.045) | (0.035) | (0.034) |
| Shock × Female | 0.009 | | | | | | |
| | (0.012) | | | | | | |
| Shock × Sex ratio | | -0.014 | -0.132*** | -0.056** | -0.108*** | 0.021 | -0.028 |
| | | (0.036) | (0.046) | (0.028) | (0.030) | (0.032) | (0.033) |
| Shock × Siblings | | -0.004 | -0.001 | 0.001 | 0.010* | 0.003 | -0.012** |
| | | (0.004) | (0.004) | (0.005) | (0.006) | (0.005) | (0.005) |
| Father | 0.012 | 0.025 | 0.010 | 0.040 | -0.018 | 0.032 | 0.037 |
| | (0.024) | (0.039) | (0.031) | (0.046) | (0.037) | (0.040) | (0.034) |
| Mother | -0.087** | -0.062 | -0.129*** | -0.124** | -0.152** | -0.047 | -0.131*** |
| | (0.042) | (0.059) | (0.047) | (0.063) | (0.065) | (0.055) | (0.049) |
| Observations | 8,377 | 4,135 | 4,242 | 3,483 | 3,557 | 3,340 | 3,479 |
| Individuals | 1,105 | 549 | 556 | 463 | 469 | 445 | 456 |
| R-squared | 0.001 | 0.001 | 0.004 | 0.003 | 0.005 | 0.002 | 0.004 |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors in parentheses. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (4), (5) and older siblings only in columns (6), (7). Father is a dummy variable that equals one if the father is alive. Mother is analogously defined. All estimates include individual and year fixed effects.

Column 1 shows that the effect is heterogeneous across individuals with younger siblings. If those children who experienced a shock and had a male-skewed sex ratio among younger siblings are more likely to be enrolled in school in the same year, the interaction between Shock×Young sex ratio should be negative. It is significant at a 1% size effect. The magnitude of the effect is large. A standard deviation increase in the Spei index (in absolute value) decreases the school dropout probability by 8 percentage points for a child with one younger brother (9% of the sample) compared to a child with one younger sister (9.5% of individuals). The magnitude of the effect is larger than the coefficient associated to Shock×Sex ratio in column 3 of table 2.4. This suggests that the heterogeneous effect of drought according to sibling sex composition is mainly driven by the young sibship. The result is robust across specifications and remains qualitatively similar in column 3, which includes interactions with characteristics on the old sibship simultaneously. Analogously, column 2 investigates the heterogeneous effect of climate shock by characteristics of the old sibship. I do not find significant effects. Moreover, I see that the effect is always statistically insignificant when Shock is measured out of the growing seasons (columns 4, 5, 6).

Finally, I reproduce baseline estimates by the child's gender to test one of the main theoretical prediction whereby girls are deeply affected by adverse climate conditions. Table 2.6 shows the results. On average, shock has an insignificant impact on the overall probability of school drop-out and the finding does not differ by the child's gender (column 1). When I interact shock with sex ratio and number of siblings, however, a different picture emerges for girls (column 3), not for boys (column 2). The coefficient associated to Shock×Sex ratio is large and highly significant in column 3: a male-skewed sex ratio largely improves girls' participation in school during adverse climatic conditions compared to girls with one younger sister (the excluded category in column 3). When restricting the analysis on individuals with younger siblings, the young sex ratio strongly alters the effect of shock for both boys (column 4) and girls (column 5), although the magnitude of coefficient is higher among girls. Surprisingly, significant effects emerge on the analysis of the differential effects of shocks according to characteristics of older siblings. In column 6, shock decreases the probability of school dropout among boys with older siblings, at a 5% size test. The number of older siblings produces differential effects on girls in column 7 although the magnitude of the effect is weak.

Overall, the analysis by child's gender and sibling subgroups stresses the role of heterogeneity across children, especially across daughters. Sibling composition exerces stronger influence among girls impacted by negative income shocks. In addition, sibling composition especially the percentage of younger brothers, alters the impact of income shocks on children's likelihood of school dropout. During hard times and in the absence of well-functioning credit or insurance markets, investing differently in children's human capital is a household diversification strategy whereby children who have relatively more younger brothers do particularly better. The results are in line with young males having a relatively higher potential income compared to younger sisters.

Next section adresses the issue whether these period investment shortfalls in children, especially those with younger sisters, are made up for or have permanent effets.

### 2.5.3 Long-run heterogeneous effects of climate shocks

In this part, I investigate the long-term differential effect of drought on children's human capital. I convert the climate variable into a cumulative measure indicating the number of shocks experienced during childhood over different age ranges (i.e., the number of times the Spei index in absolute value exceeds 1). I study the number of shocks experienced from 7 to 14, corresponding to primary school years if enrolled on time. The variable varies between 0 and at least 3 climate shocks. Simultaneously, I include the number of shocks experienced from 15 to 18 years old which corresponds to lower secondary education. The latter varies from 0 to at least 2 shocks. Rather than estimating the probability of school enrollment, the dependent variable reflects educational attainment and is calculated among children who completed their education by 2010. The sample includes individuals who never went to school (who were excluded in the first analysis) and individuals who did not participate to the 2010 survey. Moreover, the outcome allows me to take into account the potential effect of seasonal fluctuations in school attendance as repeated interruptions within the academic year might affect student's performance. I study the heterogeneous effect of cumulative shocks experienced during childhood on the human capital accumulation of a child $i$, living in family $f$, born in year $y$, using the following cross-sectional regression:

$$\begin{aligned}\text{Education}_i &= \iota_0 + \sum_a \iota_1 \text{Shocks at age a}_{j(i)} + \sum_a \iota_2 \text{Shocks at age a}_{j(i)} \times \text{Siblings}_f \\ &+ \sum_a \iota_3 \text{Shocks at age a}_{j(i)} \times \text{Sex ratio}_f + \kappa_f + \phi_y + \mu_i\end{aligned} \tag{2.5}$$

where $\iota_0$, - , $\iota_3$, are vectors of parameters to estimate and $\mu_i$ is the error term. Age a denotes the 7-14, and 15-18 age groups. The specification includes family fixed effects ($\kappa_f$) and year of birth fixed effects ($\phi_y$), to capture time-invariant family characteristics (i.e., preferences in fertility) and time-invariant cohort characteristics (i.e., education reforms and available services in some particular year) that may be related to the child's attainment. The analysis is restricted to 430 families with at least two children. With the use of family fixed effects, variations in rainfall stem from the fact that children are exposed to different shocks within the family. 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. Standard errors are corrected for spatial correlation of shocks (spatial cuf-off of 300 km). In a second stage, I conduct the same analysis focusing on the effect of younger and older siblings when affected by unfavourable agricultural conditions. Given individuals' year of birth, the number and sex composition of younger (older siblings) vary across offsprings and introduce additional variations.[37]

Additional shocks experienced between 15-18 years old decrease educational attainment by 0.2 on average and the effect does not differ by the child's gender (Table 2.7, Column

[37] The basic equation when the analysis is restricted to younger siblings is the following:

$$\begin{aligned}\text{Education}_i = \iota_4 + \sum_a \iota_5 \text{Shocks at age a}_{j(i)} + \iota_6 \text{Younger siblings}_i + \iota_7 \text{Young sex ratio}_i + \\ \sum_a \iota_8 \text{Shocks at age a}_{j(i)} \times \text{Younger siblings}_i + \sum_a \iota_9 \text{Shocks at age a}_{j(i)} \\ \times \text{Young sex ratio}_i + \ \rho_f + \nu_y + \omega_i\end{aligned} \tag{2.6}$$

where $\iota_4$, - , $\iota_9$, are vectors of parameters to estimate and $\omega_i$ is the error term. The equation which focuses on older siblings is defined analogously.

**Table 2.7 –** The effect of drought on the number of hours worked in different activities, depending on sibling composition and child's gender

| Dependent variable | **Years of schooling** | | | | | | |
|---|---|---|---|---|---|---|---|
| Sibling group | | All | | Young | | Old | |
| | | Boys | Girls | Boys | Girls | Boys | Girls |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Shock 7-14 | -0.186 | -0.597* | -0.313 | -0.029 | -1.066*** | -0.946*** | -0.240 |
| | (0.113) | (0.324) | (0.286) | (0.311) | (0.265) | (0.258) | (0.250) |
| Shock 15-18 | -0.236* | -1.068** | 0.262 | -0.750 | -0.795* | -0.824** | -0.111 |
| | (0.139) | (0.538) | (0.459) | (0.465) | (0.475) | (0.391) | (0.429) |
| Shock 7-14 × Female | -0.049 | | | | | | |
| | (0.0920) | | | | | | |
| Shock 15-18 × Female | 0.170 | | | | | | |
| | (0.125) | | | | | | |
| Shock 7-14 × Sex ratio | | -0.022 | -0.346 | 0.246 | 0.397 | 0.398* | -0.228 |
| | | (0.312) | (0.371) | (0.232) | (0.274) | (0.228) | (0.211) |
| Shock 15-18 × Sex ratio | | 0.470 | 0.506 | -0.002 | 1.315*** | 0.607 | 0.269 |
| | | (0.493) | (0.541) | (0.475) | (0.454) | (0.373) | (0.390) |
| Shock 7-14 × Siblings | | 0.033 | -0.003 | -0.076 | 0.119*** | 0.086** | -0.041 |
| | | (0.031) | (0.030) | (0.059) | (0.043) | (0.043) | (0.044) |
| Shock 15-18 × Siblings | | 0.026 | -0.082 | 0.070 | -0.031 | -0.021 | 0.031 |
| | | (0.053) | (0.050) | (0.079) | (0.073) | (0.067) | (0.076) |
| Female | -0.375* | | | | | | |
| | (0.192) | | | | | | |
| Sex ratio | | | | -1.151 | -0.696 | -0.436 | 0.021 |
| | | | | (0.744) | (0.499) | (0.591) | (0.617) |
| Siblings | | | | 0.039 | -0.086 | -0.264* | -0.124 |
| | | | | (0.144) | (0.144) | (0.143) | (0.158) |
| Observations | 1,897 | 917 | 980 | 774 | 854 | 735 | 773 |
| Number of families | 430 | 378 | 383 | 345 | 359 | 336 | 336 |
| R-squared | 0.010 | 0.018 | 0.013 | 0.020 | 0.033 | 0.031 | 0.024 |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors corrected for spatial autocorrelation in parentheses. The dependent variable represents the average years of schooling. Shock is the number of shocks experienced at different age ranges (a shock is defined when the absolute value of the spei index calculated during the six months of the growing seasons exceeds one). Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. All estimates include family and year of birth fixed effects.

1). Following the analysis above, the remaining of Table 2.7 investigates the heterogeneous long-term effects by siblings' characteristics and the child's gender. An additional shock experienced during the primary school years (7-14 years old) decreases boys' educational attainment by 0.6 at the 10% size test. Deeper effects are found for males experiencing shocks during their 15-18 years old. The magnitude of the effect is large and represents more than a complete year of schooling, significant at the 5% level.

Heterogeneous effects are identified only when I interact my measure of shock with the composition of younger siblings for girls (column 5) and older siblings for boys (column 6). Exposure to shocks between 7-14 years old strongly decreases girls' educational attainment by more than one year (significant at the 1% level, column 5). The effect is particularly large among daughters with only one sibling since the presence of additional siblings tend to slightly attenuate the negative effect of these shocks during childhood. For daughters of secondary school-aged, the sex composition of younger sibling is particularly important during income shocks. Indeed, while Shock 15-18 is negative and significant, the variable interacted with the sex ratio among younger siblings becomes highly positive. Quantitatively, both the average effect and its heterogeneity are non-negligible. An additional shock experienced between 15-18 years old decreases educational attainment by more than 0.7 of girls with one younger sister. In contrast, experiencing an additional shock improves educational attainment of girls with one younger brother by 0.5 (-0.75 + 1.31). Finally, I find significant effects on boys' educational attainment when including interactions between shocks and characteristics of the old sibship. In this case, I exclude individuals without any older siblings. Climate shocks experienced at both stages of the human capital process decrease boys' educational attainment. Nevertheless, additional older siblings and in particular older sisters tend to attenuate the negative effects of shocks taking place during primary education (column 6).

The findings are consistent with the analysis on school dropout and indicate that unfavourable agricultural conditions during the growing seasons result in a substantial loss of human capital for both girls and boys on average. In addition, characteristics among younger siblings clearly determine the effect of shocks taking place during secondary school on girls' human capital.

## 2.6 Potential Transmission Mechanisms

The results highlight that individuals with sisters do particularly worst during economic hardship. The negative results on education are particularly driven by the presence of relatively more younger sisters in the family. Nevertheless, little evidence has been provided on the mechanisms involved. How do children, especially those with younger sisters, help the household to absorb the shock? In this part, I quantitatively investigate through KHDS data potential alternatives affecting the education of selected children: school attendance, children's health, labor and marriage, which encompass the migration channel.

### 2.6.1 School attendance

The first answer following the analysis would be to save education costs. Indeed, education 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. During focus group discussions, parents raised the possibility that children cannot go to school during some days as families do not have enough money to respect school requirements (specific shoes, uniforms, furniture of meals, etc). School attendance might be used as a riskcoping instrument when households have difficulties in sustaining consumption.

To test the hypothesis, I use information on the number of hours spent in school during the last week from KHDS 1991-2004 as the dependent variable. For those who are eligible to be in secondary school (aged less than 18 years old) and do not attend school, I assign them the value of zero. The nature of the dependent variable calls for Poisson regressions. One of the drawback with the use of maximum likelihood model is that it does not enable me to use econometric tools to take serial correlation and spatial correlation into account in the

**Table 2.8** – School attendance: alternative measure of climate shocks

| | | Hours in school, last week | | | | | |
|---|---|---|---|---|---|---|---|
| Dependent variable | | Boys | Girls | Boys | Girls | Boys | Girls |
| Sibling group | | All | | Young | | Old | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Shock | -0.051 | 0.209 | -0.467* | -0.054 | -0.487** | 0.454* | 0.204 |
| | (0.074) | (0.196) | (0.256) | (0.210) | (0.206) | (0.239) | (0.176) |
| Shock × Female | 0.200** | | | | | | |
| | (0.097) | | | | | | |
| Shock × Sex ratio | | -0.338 | 0.409 | -0.348 | 0.277 | -0.268 | -0.046 |
| | | (0.239) | (0.328) | (0.219) | (0.220) | (0.231) | (0.211) |
| Shock × Siblings | | -0.022 | 0.067** | 0.047 | 0.139*** | -0.099** | -0.032 |
| | | (0.028) | (0.028) | (0.048) | (0.040) | (0.041) | (0.030) |
| Father in hh | 0.043 | -0.122 | 0.343* | -0.098 | 0.574 | -0.104 | 0.247 |
| | (0.130) | (0.205) | (0.198) | (0.201) | (0.420) | (0.215) | (0.192) |
| Mother in hh | -0.337 | 0.001 | -0.520* | 0.018 | -0.771* | -0.289 | -0.725* |
| | (0.232) | (0.305) | (0.286) | (0.305) | (0.395) | (0.438) | (0.383) |
| Age | -0.146 | -0.255* | -0.089 | -0.308** | -0.122 | -0.189 | -0.134 |
| | (0.102) | (0.139) | (0.149) | (0.156) | (0.163) | (0.170) | (0.178) |
| Season | -0.054 | -0.115 | 0.016 | -0.117 | -0.012 | -0.131 | -0.042 |
| | (0.062) | (0.095) | (0.087) | (0.105) | (0.092) | (0.103) | (0.091) |
| Observations | 1,007 | 492 | 502 | 425 | 449 | 394 | 421 |
| Individuals | 388 | 190 | 192 | 163 | 170 | 153 | 161 |

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Results are estimated using the first five ways of the KHDS survey with Poisson regressions. The dependent variable is the number of hours spent in school during the last week. For those who are eligible to be in secondary school (aged less than 18 years old) and do not attend school, I assign them the value of 0. Shock is the average of the spei index during the month of the survey, in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (4), (5) and older siblings only in columns (6), (7). Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. Season equals 1 if the month of interview belongs to the growing season. All estimates include individual and survey fixed effects.

climate data. A conditional fixed-effects model with robust standard errors to within-group autocorrelation of errors estimates the results presented in Table 2.8. Shock is the average of the Spei drought index during the month of the survey in absolute value, using the 6-month timescale.[38] Controls include a dummy equals to 1 if the interview takes place during the growing season, the time of the survey, the age, the presence of the father and the mother in the household. The results, which are disaggregated by the child's gender, give some indications about the direction of the effects and have to be interpreted with caution given the small sample size.[39]

Column 1 provides an estimate of the magnitude of the effect of the climate shock on school attendance. Shock does not produce any effect on boys while slightly increases girls' hours spent in school. In the following, I investigate heterogenous effects and include the interaction of shocks with the number of siblings and their gender composition to separate the effect of sibling size from sibling gender. Higher values of the shock variable decrease attendance in school for girls with one sister (column 3, coefficient in the first line), while it does not affect girls with relatively more brothers. The effect is significant at the 10% size test. Also, the presence of additional siblings slightly reduces the negative impact associated to unfavourable climate conditions on girls with a female-skewed ratio. The findings are similar in column 5 where I analyze the composition of younger siblings only. Higher values of the climate index decrease the number of hours spent in school for girls with one younger sister (the reference category) and becomes less significant for girls who have additional younger sisters. The magnitude of the impact is large and highly significant, which is consistent with the results identified on girls' school dropout in Table 2.5. For boys, we do not find any significant effects globally, except in column 6. Boys with one older sister actually increase their attendance in school during a shock (coefficient in the first line). The positive effect decreases with additional older siblings. Finally, I test the results using the

[38]Due to insufficient observations, I do not restrict the analysis on individuals interviewed during a month of the growing season. Thereby, Shock encompasses both unusual conditions during a month within or outside the growing seasons in this analysis. Also, I test a cumulative measure indicating the number of droughts (where the continous index, in absolute value, exceeds 1) during the last 3 months, which vary from 0 to at least 2. The results are presented in Appendix, Table B.6.

[39]The sample is restricted to children with information on biological siblings (i.e., listed in the sibling module in 2010) who participated in KHDS-1.

number of shocks experienced during the last 3 months in the Appendix Table B.6. I find similar results: the sex composition of siblings, especially among younger siblings, largely alters the effect of climate shocks on girls' attendance in school.

Saving on education costs may not be sufficient to overcome shocks, in which case other alternatives increasing liquidity may have to be engaged.

### 2.6.2 Child labor

Do children, especially those with younger sisters, help other relatives to share burdens through labor? Before proceeding, one may ask the following question: why increasing child labor would be a rational response to a negative agricultural productivity shock given that most labor consists in agricultural work? When agricultural labor productivity is low, the demand for labor at harvest time reduces, individuals supply less labor and instead borrow or draw on savings to smooth their consumption. This is in line with Shah and Steinberg (2017) who demonstrate that bad rainfall years result in relative paucity of outside options and low wages, which diminishes the use of child labor and increases school attendance. However, in an economy in which individuals are not able to save or borrow, workers might work more in order to meet their consumption needs. According to Jayachandran (2006), workers who are close to subsistence supply labor less elastically to agricultural productivity. In addition, agricultural shocks being local to an area, workers (including children and their parents) might be able to migrate to higher-productivity areas to substitute away from the home labor market. In this setting, 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. In the Kagera region, Beegle et al. (2006) demonstrate that households increase their use of child labor in response to crop shocks. Bandara et al. (2015) extended the results for all Tanzania.

To test the child labor mechanism, I use detailed data on time-use of each household member aged 7 and older, available in the first five ways of the KHDS. Household members were asked the number of hours they worked in the seven days prior to the interview, broken down by specific tasks. Following Beegle et al. (2006), I define the following tasks as house-

**Table 2.9 –** The effect of drought on the number of hours worked in different activities, depending on sibling composition and child's gender

| Dependent variable | Housework | | | Agriculture | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | All | Boys | Girls | All | Boys | Girls | All | Boys | Girls |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Shock | 0.482 | 4.148*** | 0.478 | -1.108* | 1.122 | 5.186*** | -0.557 | 6.345** | 6.028*** |
| | (0.649) | (1.405) | (1.235) | (0.567) | (2.123) | (1.363) | (1.143) | (2.606) | (2.171) |
| Shock × Female | -1.259* | | | 0.371 | | | -0.696 | | |
| | (0.704) | | | (0.668) | | | (1.223) | | |
| Shock × Sex ratio | | -3.255* | -0.842 | | 0.356 | -4.773** | | -4.889 | -4.908 |
| | | (1.865) | (2.022) | | (1.861) | (1.994) | | (3.023) | (3.413) |
| Shock × Siblings | | -0.406*** | -0.0454 | | -0.400 | -0.598** | | -0.830** | -0.715* |
| | | (0.140) | (0.173) | | (0.245) | (0.250) | | (0.333) | (0.379) |
| Father in hh | 0.093 | -3.891*** | 4.648** | -2.620* | -1.915 | -4.223** | -0.489 | -2.700 | 1.626 |
| | (1.481) | (1.409) | (2.004) | (1.412) | (1.679) | (1.804) | (1.975) | (2.365) | (2.880) |
| Mother in hh | -2.224*** | 0.935 | -3.526*** | 0.816 | 0.252 | 0.088 | -0.830 | 1.153 | -2.603 |
| | (0.777) | (1.819) | (0.950) | (1.266) | (1.640) | (1.918) | (1.693) | (3.362) | (1.995) |
| Age | 0.374 | 0.145 | 0.988 | -1.276 | -0.638 | -1.442 | -0.770 | -0.536 | -0.229 |
| | (0.987) | (1.258) | (0.979) | (1.449) | (1.569) | (1.531) | (2.090) | (2.283) | (2.273) |
| Season | -0.941** | -1.335*** | -0.638 | 0.780 | 0.589 | 1.035 | 0.032 | -0.233 | 0.305 |
| | (0.389) | (0.290) | (0.559) | (0.716) | (0.755) | (0.778) | (0.957) | (1.067) | (1.080) |
| Observations | 3,011 | 1,453 | 1,517 | 3,101 | 1,495 | 1,560 | 2,930 | 1,401 | 1,488 |
| Individuals | 1,258 | 599 | 644 | 1,362 | 646 | 696 | 1,243 | 590 | 638 |
| R-squared | 0.006 | 0.014 | 0.011 | 0.005 | 0.006 | 0.017 | 0.001 | 0.009 | 0.007 |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Conley standard errors corrected for spatial autocorrelation in parentheses. Results are estimated using the first five waves of the KHDS with linear regressions. From columns (1) to (3), the dependent variable is the number of working hours in housework during the last 7 days for individuals aged 7-18 years old. Results on agriculture hours are presented from columns (4) to (6) and on total work from columns (7) to (9). Shock is the average of the spei index during the month of the survey, in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. Season equals 1 if the month of interview belongs to the growing season. All estimates include individual and survey fixed effects.

hold work: preparing meals, cleaning the house, shopping for food, collecting firewood and collecting water. Agricultural work pools farm and garden work (i.e., cash crop and food crop production), processing crops for sale, caring for poultry or livestock, collecting and transforming household livestock or animal products for sale. Finally, total work includes household chores, agricultural work, outside employment and labor in non-farm household businesses.[40] More than 90% of children worked at least 1 hour in the previous week.

I investigate the effect of climate shocks on the number of weekly hours spent in household chores, in agriculture and in any work for individuals aged between 7-18 years old. Child labor and education are simultaneous decisions during this age-period given that the upper age range for lower secondary education is 17-18 years old if enrolled on time. Shock is the average of the Spei index, in absolute value, during the month of the survey.[41] Due to insufficient observations, I do not restrict the analysis on individuals interviewed during the growing season.[42] Thereby, my climate variable is computed using months both within and outside the growing season. Nevertheless, all estimates include a dummy equals to 1 if the month of interview belongs to one of the two growing seasons in Kagera. Finally, I include individual and survey fixed effects and I correct for spatial correlation using Conley standard errors (spatial cut-off of 300 km).

Table 3.12 reports the estimates. The number of hours worked in household slightly decreases for girls (column 1), whereas the number of agriculture hours decreases for both gender during a shock (column 4). I do not find significant effects on total labor in column 7. I go beyond average effects and demonstrates the existence of heterogeneous effects in the remaining columns.

[40] My estimates of child labor are based on the number of working hours as a worker or as self-employed taken in combination. This raises serious issues regarding measurement error since the estimation of hours of work in self-employed activities or in household chores is well-known to be very hard. To the extent that difficulties in measurement for a particular individual do not change across waves, these concerns will be somewhat mitigated since measurement error might be likely to be differenced out.

[41] Estimates of the effect of the number of shocks experienced during the last 3 months on child labor, by gender and sibling composition, are reported in Appendix, Table B.5.

[42] The number of observations drops considerably after selecting children belonging to the 2010 sibling module who have been interviewed in KHDS1 and KHDS2.

Boys' labor in household chores strongly varies with the sex composition of siblings in column 2. Unusual climate conditions increase by 4.1 the number of hours worked in household for boys with only one sister (the reference category in the first line), at a 1% size test. Boys' labor increases to a far lesser extent for those with a male-skewed sibling ratio (coefficient in the third line, significant at 10%). Also, the number of siblings slightly reduces the need to increase hours spent in household chores. For girls, I do not find heterogeneous effects depending on sibling composition in column 3. Compared to the sister, the brother help to work in the household mainly by collecting firewood and collecting water in the period of unusual rainfall. This corroborates one of the main theorey exposed in the theoretical part: having relatively more brothers help to overcome climate shocks because boys have a larger work capacity in household chores that increase during an income shock. With regard to agricultural labor, I find significant effects only for girls in column 6. Higher values of the shock variable increase by 5.2 the number of hours spent in agriculture work for girls with one sister, at the 1% size effect. Once again, girls with a male-skewed ratio and additional siblings tend to work less on the farm in times of economic harship. Finally, climate shock increases substantially girls and boys' total labor in columns 8 and 9. For both gender, the total number of worked hours decreases with the presence of additional siblings in the family. The magnitude of the effect and the level of significance are stronger for males than for females.

Estimates using the composition of younger and older siblings by child's gender are reported in Appendix, Table B.4. Overall, I do not observe impacts that are statistically significant except for girls' labor in agriculture. Column 6 shows that a shock raises the number of worked hours for girls with one younger sister. Younger brothers helps to offset the negative effect of Shock on female labor whereas younger sisters cause girls to work more intensively in times of economic hardship.

### 2.6.3 Children's health

Among strategies implemented to cope with shocks, reducing food consumption was one of the most frequent answer mentionned 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.[43] 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 climate shocks on the probability of becoming sick during the last 4 weeks as the outcome.[44] Table 2.10 column 1 shows that higher values of the shock variable increase by 5% the probability of becoming sick on average, at a 5% size effect. The finding does not differ by gender. However, climate shocks are the same across subpopulations regarding nutritional and health status. Indeed, there is almost no difference between children in the probability of becoming sick according to the family structure. Neither the main shock effect nor interactions are significant for girls and boys. Thus, it is unlikely this channel is driving the heterogeneous results identified on school dropout in the main analysis.

**Table 2.10 –** The effect of drought on children's health status, depending on gender and sibling composition

| | | Likelihood of illness, last 4 weeks | | | | | |
|---|---|---|---|---|---|---|---|
| Dependent variable | | Boys | Girls | Boys | Girls | Boys | Girls |
| Sibling group | | All | | Young | | Old | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Shock | 0.053** | -0.020 | 0.055 | -0.009 | -0.04 | 0.115 | 0.004 |
| | (0.024) | (0.082) | (0.064) | (0.059) | (0.124) | (0.088) | (0.066) |
| Shock × Female | -0.013 | | | | | | |
| | (0.033) | | | | | | |
| Shock × Sex ratio | | 0.037 | -0.137 | 0.054 | 0.102 | -0.001 | 0.009 |
| | | (0.072) | (0.104) | (0.058) | (0.100) | (0.074) | (0.104) |
| Shock × Siblings | | 0.013 | 0.003 | 0.021*** | 0.0041 | -0.005 | -0.007 |
| | | (0.012) | (0.009) | (0.007) | (0.022) | (0.018) | (0.012) |
| Father in hh | -0.118 | -0.238 | -0.004 | -0.235 | -0.127 | -0.222 | 0.037 |
| | (0.116) | (0.186) | (0.084) | (0.185) | (0.134) | (0.176) | (0.080) |
| Mother in hh | -0.153** | -0.055 | -0.177*** | -0.112 | -0.169** | 0.125 | -0.149** |
| | (0.063) | (0.114) | (0.066) | (0.115) | (0.078) | (0.134) | (0.073) |
| Age | 0.078* | 0.211*** | -0.058 | 0.217*** | -0.061 | 0.225*** | -0.061 |
| | (0.042) | (0.048) | (0.042) | (0.063) | (0.067) | (0.046) | (0.049) |
| Season | 0.021 | 0.021 | 0.018 | 0.040** | 0.012 | 0.010 | 0.017 |
| | (0.019) | (0.019) | (0.028) | (0.017) | (0.024) | (0.023) | (0.033) |
| Observations | 3,104 | 1,495 | 1,562 | 1,263 | 1,357 | 1,249 | 1,296 |
| Individuals | 1,373 | 652 | 700 | 537 | 581 | 556 | 590 |
| R-squared | 0.005 | 0.013 | 0.006 | 0.016 | 0.008 | 0.014 | 0.004 |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors corrected for spatial autocorrelation in parentheses. Results are estimated using the first five waves of the KHDS with linear regressions. The dependent variable is a dummy variable that equals one if the individual got sick during the last four weeks. Shock is the average of the spei index during the month of the survey, in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (4), (5) and older siblings only in columns (6), (7). Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. Season equals 1 if the month of interview belongs to the growing season. All estimates include individual and survey fixed effects.

**Table 2.11 –** Effect of climate shocks on girls' marriage

| Dependent variable | Probability to be married | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Before 20 years old | | | | Before 25 years old | | | |
| Sibling group | | All | Young | Old | | All | Young | Old |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Shock | -0.001 | 0.032 | -0.022 | 0.055* | -0.009 | 0.011 | 0.003 | -0.016 |
| | (0.009) | (0.040) | (0.047) | (0.029) | (0.016) | (0.031) | (0.032) | (0.028) |
| Shock × Female | 0.009 | | | | -0.006 | | | |
| | (0.011) | | | | (0.014) | | | |
| Shock × Sex ratio | | 0.015 | 0.013 | 0.016 | | -0.021 | -0.068** | 0.045* |
| | | (0.045) | (0.038) | (0.031) | | (0.032) | (0.033) | (0.025) |
| Shock × Siblings | | -0.004 | 0.006 | -0.012** | | -0.003 | 0.003 | -0.007 |
| | | (0.003) | (0.006) | (0.006) | | (0.002) | (0.003) | (0.004) |
| Father | 0.028 | 0.028 | 0.036 | -0.009 | 0.009 | 0.002 | 0.006 | -0.002 |
| | (0.025) | (0.042) | (0.048) | (0.045) | (0.014) | (0.023) | (0.025) | (0.024) |
| Mother | -0.056 | -0.110* | -0.128* | -0.144** | -0.058*** | -0.111** | -0.099* | -0.137** |
| | (0.034) | (0.062) | (0.071) | (0.072) | (0.021) | (0.051) | (0.053) | (0.066) |
| Observations | 5,960 | 3,411 | 3,044 | 2,698 | 10,967 | 5,889 | 5,317 | 4,625 |
| Individuals | 1,243 | 727 | 646 | 577 | 1,244 | 727 | 646 | 577 |
| R-squared | 0.001 | 0.002 | 0.003 | 0.005 | 0.001 | 0.002 | 0.003 | 0.003 |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors in parentheses. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings in columns (1) and (5). The variable is analogously defined using younger siblings in columns (3), (6) and older siblings in columns (4), (7). Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variable is computed using younger (older) siblings only in the second (third) column of each dependent variable. Father is a dummy variable that equals one if the father is alive. Mother is analogously defined. All specifications include individual and year fixed effects.

### 2.6.4 Child marriage

Finally, the composition of households might change during a shock, so that members who were primarily consumers (i.e., children) leave. Do individuals are going to leave the household through marriage to alleviate family expenses? Parents may have incentives to marry their daughter in particular 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, droughts or floods may reduce marriage payments due to lower income of the grooms' family in nearby areas (Gray and Mueller, 2012). The qualitative study has shown that the bride price is usually shared within the community, between all relatives, and is not intended to a specific use such as educational expenditures of remaining children. Nevertheless, individuals can leave the household through marriage to partially alleviate the liquidity shortages weighing on their families at the moment of shock.

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. I complete information on age at first marriage using a question on the length of marriage available in the 2010 sibling module. As the length variable is presented in multiple range scales, I use the median of each interval to generate the variable for missing observations.[45] Similar 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 structure. The outcome is a binary variable coded as 1 in the year the individual gets married, and zero otherwise. Considering that a 15 years-old child may be at risk of early marriage, the duration of interest is the time between

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

[44] The KHDS survey asks to all household members the following question: "During the past 4 weeks, have you had any illness or injury?"

[45] Individuals belonging to the sibling questionnaire were asked "how many years have you been married?" The different intervals are: less than five years, between 5-9 years, between 10-14, between 15-20, more than 20.

his 15 and the age when the individual marries, after which he exits the data.[46] Since I am interested in parents' alternatives to higher education investment, I consider early marriage as a union where one member got married before 20 years old. Thus, individuals married after age 19 are right censored. This concerns 20% of individuals out of 1,244 individuals in the sample. Among them, 86% are females. Due to insufficient cases of early marriage among males, I restrict the analysis on girls. Also, I enlarge the definition of early marriage and investigate the effect of climate shock on the probability of being married before 25 years old (the case of 68% of females). The econometric strategy is similar to the one exposed in sub-section 4, relying on interactions between climate shocks and sibling characteristics.

Table 2.11 presents the results based on the within estimator, using individual and year fixed effects to analyse variations in climate conditions within the life of individuals. In all specifications, I correct for serial correlation within villages and spatial correlation between grids using Conley standard errors. Abnormal climate conditions increase the probability of early marriage by 5 percentage points for girls with only one older sister (reference category in the first line of column 4). The effect is significant at the 10% level. The number of older siblings interacted with the shock variable displays negative and significant coefficients at the 5% level of significance. This suggests that additional older siblings shelter from the negative effect of income shock on girls' early marriage.

Finally, I test the results on the probability of being married before 25 years old from columns 5 to 8. Whereas they are working in opposite direction, both the young and old sibship sex composition determine how shocks will affect girls' probability of marriage. Column 7 demonstrates that younger brothers will mitigate the negative effect of climate change on child marriage, as the coefficient associated to Shock$\times$ (young) Sex ratio is negative and significant at the 5% level. In contrast, relatively more older brothers accelerate a daughter's marriage (positive coefficient in the third line of column 8) probably because parents need some liquidity to pay for the marriage of older sons. In sum, the gender effect coming from

[46] 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 oldest daughters.

**Table 2.12 –** Standard Errors Using Various Spatial and Temporal Cut-offs

| Explanatory variable | Shock | Shock × Sex ratio | Shock × Siblings |
|---|---|---|---|
| Year:0 - Distance:100 | (0.030) | (0.027)*** | (0.003) |
| Year:0 - Distance:300 | (0.028) | (0.026)*** | (0.003) |
| Year:0 - Distance:500 | (0.027) | (0.025)*** | (0.003) |
| Year:2 - Distance:100 | (0.031) | (0.028)*** | (0.003) |
| Year:2 - Distance:300 | (0.029) | (0.027)*** | (0.003) |
| Year:2 - Distance:500 | (0.028) | (0.025)*** | (0.003) |
| Year:4 - Distance:100 | (0.031) | (0.028)*** | (0.003) |
| Year:4 - Distance:300 | (0.029) | (0.027)*** | (0.003) |
| Year:4 - Distance:500 | (0.028) | (0.026)*** | (0.003) |

*** p<0.01, ** p<0.05, * p<0.1. I provide standard errors in parentheses for the estimates of my baseline (Specification 4 in Table 2.4). I use different spatial and temporal cut-offs. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. Coefficient on Shock is 0.027; on Shock×Sex ratio is -0.072; on Shock×Siblings is -0.003. All estimates include individual and year fixed effects.

the sibling composition as well as the size effect from the old sibship pinpointed in the education analysis are identified in the marriage analysis mainly for girls above secondary school age.

## 2.7 Robustness part

In this section, I perform additional estimations to test the robustness of my baseline results.

Baseline estimates use Conley standard errors to take into account the fact that shocks may be persistent over time and be correlated across space. In Table 2.12, I re-estimate my main regression (equation 2.2, shown in Table 2.4 column 4) and test the sensitivity of the standard errors. I consider different cut-offs, from 0 to 4 years for the serial correlation and 100 km / 300 km / 500 km for the spatial correlation. I report only the standard errors of interest variables: Shock, Shock × Sex ratio, and Shock × Siblings. Coefficients are respectively 0.027, -0.072 and -0.003. The results remain similar.[47]

[47] The same exercise has been conducted using the composition of younger and older siblings. I find similar results.

**Table 2.13 –** Alternative definition of school drop-out and specification

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Secondary school drop-out | | | | | | | |
| Dependent variable | ≤ 17 years old | | | | ≤ 22 years old | | | |
| Specification | Individual and year FE | | | | Family and Cohort FE | | | |
| Sibling group | | All | Young | Old | | All | Young | Old |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Shock | -0.012 | 0.056** | 0.030 | -0.001 | -0.003 | 0.056* | 0.045* | -0.009 |
| | (0.020) | (0.028) | (0.028) | (0.023) | (0.013) | (0.029) | (0.024) | (0.028) |
| Shock × Sex ratio | | -0.080*** | -0.075*** | -0.022 | | -0.058** | -0.078*** | 0.007 |
| | | (0.027) | (0.021) | (0.018) | | (0.028) | (0.023) | (0.026) |
| Shock × Siblings | | -0.005 | 0.001 | -0.008** | | -0.005 | -0.003 | -0.001 |
| | | (0.003) | (0.003) | (0.004) | | (0.003) | (0.005) | (0.004) |
| Sex ratio | | | | | 0.028** | 0.060*** | 0.021 | 0.020 |
| | | | | | (0.013) | (0.021) | (0.017) | (0.023) |
| Siblings | | | | | | | -0.006 | 0.004 |
| | | | | | | | (0.004) | (0.003) |
| Father | -0.001 | 0.001 | 0.004 | 0.004 | -0.134*** | -0.134*** | -0.165*** | -0.130*** |
| | (0.022) | (0.022) | (0.030) | (0.024) | (0.024) | (0.024) | (0.029) | (0.024) |
| Mother | -0.030 | -0.031 | -0.053 | -0.039 | -0.187*** | -0.187*** | -0.266*** | -0.182*** |
| | (0.032) | (0.032) | (0.037) | (0.035) | (0.029) | (0.029) | (0.028) | (0.032) |
| Observations | 8,377 | 8,377 | 7,040 | 6,819 | 8,377 | 8,377 | 7,040 | 6,819 |
| Individuals | 1,105 | 1,105 | 932 | 901 | – | – | – | – |
| Families | – | – | – | – | 376 | 376 | 360 | 337 |
| R-squared | 0.000 | 0.001 | 0.002 | 0.001 | 0.011 | 0.012 | 0.016 | 0.011 |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors in parentheses. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (4), (5) and older siblings only in columns (6), (7). Father is a dummy variable that equals one if the father is alive. Mother is analogously defined.

Secondly, I restrict the definition of secondary school-dropout. I consider a drop-out only among secondary school-aged children and exclude individuals aged more than 17 years old at the time of shock. Table 2.13 displays the findings. I directly interpret the heterogeneous estimation presented in column 2. Shock is now positive and significant at the 5% size test and its interaction with sex ratio is negative with larger magnitude (compared to estimation 3 in Table 2.4). Columns 3 and 4 investigate climate shock impacts according to the younger and the older sibship and show results equivalent to Table 2.5.

In addition, I test the education equation using the family instead of the individual as the unit of analysis. I include family and year of birth fixed effects from columns (5) to (8). In so doing, I am able to check the robustness of my baseline gender and size effect to controlling for fixed family characteristics (i.e., preferences in fertility). Since I use the within family variation, I restrict attention to families with at least two children. Shocks affecting children only when they are eligible to be in school, variations between children stem from the distribution of dry and wet climate conditions across time. Additional variation comes from the interaction of Shock with sibling sex ratio, the latter differs between sisters and brothers within the family. Children with relatively more brothers do particularly better during income shocks compared to a child with one sister (column 6, coefficient in the first line). The last two columns investigate the effect of drought between children who are exposed to different climate shocks and have not a similar number of younger (colum 7) and older (column 8) brothers and sisters. The results are similar to baseline findings. Notice that variations in the number and sex ratio of younger and older siblings within the household allow me to observe their direct effect on school enrollment. Sex ratio is positive in columns 5 and 6, suggesting that having relatively more brothers than sisters increases the likelihood of school drop-out on average. The number of younger and older siblings in columns 7 and 8 are insignificant.

Is there a role for birth order in the interpretation of the findings? Later born children only exist in large families, and birth order interacts with sibling sex composition to the extent that parents give birth until the first boy. In this case, earlier-born children is correlated with a female-skewed sibling ratio. To ensure that sibling size and gender composition do not proxy

**Table 2.14 –** Robust effect of drought on school dropout: additional controls and duration analysis

| Method | Linear probability | | | | | | Cox model | | |
|---|---|---|---|---|---|---|---|---|---|
| Sibling group | All | | Young | | Old | | All | Young | Old |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Shock | 0.028 | -0.040 | 0.029 | -0.045 | -0.032 | -0.077 | 0.276 | 0.108 | 0.058 |
| | (0.029) | (0.056) | (0.030) | (0.063) | (0.023) | (0.055) | (0.192) | (0.162) | (0.162) |
| Shock × Sex ratio | -0.074*** | -0.074*** | -0.078*** | -0.080*** | -0.013 | -0.015 | -0.402* | -0.506*** | -0.006 |
| | (0.026) | (0.027) | (0.020) | (0.020) | (0.020) | (0.020) | (0.206) | (0.155) | (0.164) |
| Shock × Siblings | 0.001 | -0.003 | 0.003 | 0.004 | -0.020* | -0.005 | -0.011 | 0.050 | -0.035 |
| | (0.003) | (0.003) | (0.004) | (0.004) | (0.012) | (0.003) | (0.026) | (0.032) | (0.031) |
| Shock × BO | -0.005 | | -0.004 | | 0.013 | | | | |
| | (0.003) | | (0.003) | | (0.009) | | | | |
| Shock × Age | | 0.005 | | 0.004 | | 0.004 | | | |
| | | (0.004) | | (0.004) | | (0.004) | | | |
| Father | 0.012 | 0.012 | 0.007 | 0.008 | 0.027 | 0.027 | 0.130 | 0.077 | 0.149 |
| | (0.024) | (0.024) | (0.030) | (0.029) | (0.025) | (0.025) | (0.092) | (0.113) | (0.093) |
| Mother | -0.089** | -0.089** | -0.132*** | -0.133*** | -0.078* | -0.078* | -0.268*** | -0.421*** | -0.205** |
| | (0.041) | (0.041) | (0.043) | (0.043) | (0.042) | (0.041) | (0.100) | (0.114) | (0.095) |
| Sex ratio | | | | | | | 0.163 | 0.138 | 0.088 |
| | | | | | | | (0.169) | (0.119) | (0.113) |
| Siblings | | | | | | | 0.061 | 0.006 | 0.002 |
| | | | | | | | (0.187) | (0.024) | (0.025) |
| Ind & Year FE | Yes | Yes | Yes | Yes | Yes | Yes | No | No | No |
| Observations | 8,377 | 8,377 | 7,040 | 7,040 | 6,819 | 6,819 | 8,377 | 7,040 | 6,819 |
| R-squared | 0.002 | 0.002 | 0.004 | 0.004 | 0.002 | 0.002 | – | – | – |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors in parentheses from columns (1) to (6). From columns (7) to (9), standard errors are clustered at the village level. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings in columns (1), (2) and (7). The variable is analogously defined using younger siblings in columns (3), (4), (8) and older siblings in columns (5), (6), (9). Siblings denotes the number of biological siblings, irrespective of their live status or place of residence in columns (1), (2) and (7). The variable designates the number of younger and older siblings respectively in columns (3), (4), (8) and (5), (6), (9). Father is a dummy variable that equals one if the father is alive. Mother is analogously defined. Columns (1) to (6) control for the interaction of the spei drought index with individual's birth order and age using a linear probability model with individual and year fixed effects. Columns (7) to (9) report coefficients of the cox model, estimated with maximum likelihood.

for birth order, I include individuals' birth order as an additional control. Child age might also be correlated with siblings' characteristics since a child with many younger siblings is by definition older. Importantly, for each of these covariates, I include their interaction with the climate index in Table 2.14, from columns 1 to 6. The inclusion of these control variables does not reduce the magnitude and significance of Shock×Sex ratio, calculated among the whole family (columns 1, 2) or only among younger siblings (columns 3, 4). Shock decreases early school leaving for girls with additional older siblings. However, the effect disappears when I control for age in column 6.

Some concerns may arise on the use of a linear probability model with a binary dependent variable which is considered as a continuous one. Consequently, I test the results using a duration model with the maximum likelihood estimator. 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 my setting. I analyze time from entry to school until drop-out of school in years with a Cox proportional hazard model (Cox, 1972). The results are reported in Table 2.14, columns 7, 8 and 9. Hazard ratios are interpreted as the proportional change in hazard when the Spei drought index (calculated during the growing seasons in absolute value) increased by one standard deviation compared to the long-term average for a given sibling composition. In column 7, climate shock results in a lower hazard and therefore a higher survival time for children with a male-skewed ratio, given the negative sign associated to the coefficient Shock×Sex ratio. Having relatively more brothers than sisters significantly decreases the risk of being out-of-school when a shock occurs. The magnitude of the coefficients is huge due to the absence of individual and year fixed effects. Column 8 restricts the analysis to the size and sex composition of younger siblings and corroborates previous findings. Abnormal climate conditions have a shallower effect on children with male-skewed sex ratio among younger siblings. Finally, I include old brotherhood variables in column 9. Though, no significant effects are identified.

**Table 2.15 –** Alternative measures of climate shocks

| Dependent variable | Secondary school drop-out | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | SPEI- 1 time-scale | | | SPEI- 3 time-scale | | | Shock in t-1 | | |
| Sibling group | All | Young | Old | All | Young | Old | All | Young | Old |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Shock | -0.038 | -0.053 | -0.054 | 0.041 | 0.014 | -0.001 | 0.031 | -0.002 | 0.032 |
| | (0.051) | (0.042) | (0.040) | (0.033) | (0.026) | (0.024) | (0.029) | (0.026) | (0.025) |
| Shock × Sex ratio | -0.045 | -0.053 | 0.002 | -0.069** | -0.075*** | -0.016 | -0.025 | 0.010 | -0.019 |
| | (0.046) | (0.033) | (0.036) | (0.030) | (0.023) | (0.023) | (0.024) | (0.020) | (0.019) |
| Shock × Siblings | 0.001 | 0.010 | -0.008 | -0.003 | 0.007 | -0.006* | -0.004* | -0.003 | -0.003 |
| | (0.005) | (0.007) | (0.007) | (0.003) | (0.004) | (0.003) | (0.002) | (0.003) | (0.003) |
| Observations | 8,377 | 7,040 | 6,819 | 8,377 | 7,040 | 6,819 | 8,362 | 7,025 | 6,810 |
| R-squared | 0.001 | 0.003 | 0.002 | 0.002 | 0.003 | 0.001 | 0.001 | 0.002 | 0.001 |
| | Shock in t+1 | | | Drought (0/1) | | | Flood (0/1) | | |
| | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) |
| Shock | -0.003 | -0.015 | 0.026 | 0.047 | -0.009 | -0.014 | -0.044 | 0.031 | -0.078*** |
| | (0.029) | (0.026) | (0.024) | (0.056) | (0.047) | (0.039) | (0.037) | (0.023) | (0.021) |
| Shock × Sex ratio | -0.038 | -0.002 | -0.056*** | -0.079 | -0.064 | -0.055** | -0.017 | -0.029 | 0.014 |
| | (0.027) | (0.022) | (0.019) | (0.048) | (0.045) | (0.027) | (0.034) | (0.023) | (0.032) |
| Shock × Siblings | 0.001 | 0.001 | -0.002 | -0.009** | -0.001 | -0.007 | 0.011*** | -0.001 | 0.020*** |
| | (0.002) | (0.004) | (0.003) | (0.004) | (0.005) | (0.005) | (0.004) | (0.003) | (0.004) |
| Observations | 8,372 | 7,036 | 6,816 | 8,377 | 7,040 | 6,819 | 8,377 | 7,040 | 6,819 |
| R-squared | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 | 0.001 | 0.002 | 0.002 |

*** p<0.01, ** p<0.05, * p<0.1. Conley standard errors in parentheses. Alternative measures of drought are used in the following columns: (1)-(3) the absolute value of the average of the spei calculated during the growing seasons (March, April, May and October, November, December) using a one month time-scale; (4)-(6) the absolute value of the average of the spei calculated during the growing seasons using a three month time-scale; (7)-(9) the absolute value of the spei index during the growing season in time t-1 ; (10)-(12) the absolute value of the spei index during the growing season in time t+1; (13)-(15) a dummy variable that equals one for unusual dry conditions (spei index$\leq$-1); (16)-(18) a dummy variable that equals one for unusual wet conditions (spei index $\geq$1). Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are calculated analogously using only younger siblings in columns (2), (5), (8), (11), (14), (17) and older siblings only in columns (3), (6), (9), (12), (15), (18). Controls include dummies equal to one if the father and the mother are alive. All estimates include individual and year fixed effects.

In Table 2.15, I test the results using alternative measures of shocks. I first explore robustness to the choice of the time scale over which water deficits accumulate, which reflects different type of droughts. Shock represents the average of the Spei index, in absolute value, calculted during the growing seasons using a 1 month time-scale in columns 1, 2, 3 and a 3 month time-scale in columns 4, 5 and 6. Meteorological shocks (1 month timescale) do not have (heterogeneous) effects on children's likelihood of enrollment. In contrast, the 3 month timescale indicator gives identical results. The magnitude of coefficients is slightly weaker, though. In addition, I test the sensitivity of the timing of shocks using lags and leads. In columns 7, 8, 9, I explore the persistence of the effects of past shocks. It is interesting to note that the lagged, as opposed to contemporaneous climate shocks, are not significantly associated with dropout in time t. On the one hand, this could simply reflect imprecisions in the timing of events as age at school dropout might suffer from misreporting. On the other hand, it may also suggest that households do not attempt to avoid withdrawing a child from school (or before completing a grade) when a loss of income occurs due to the lack of financial services, assets and savings. In the lack of alternative coping strategies, households may not be able to compensate for temporary income shocks even though they might be reluctant to pull (specific) children out of school in the face of short-term adversity. Also, I test the existence of anticipation effects using leads from columns 10 to 12. I expect coefficients on the lead shock and its interactions to be zero. The effect of exposure to future income shock is generally statistically insignificant. However, column 12 points out a significant coefficient on the interaction of Shock with the sex ratio among older siblings. This suggests the existence of anticipation effects, where children with relatively more older brothers increase their attendance in school before the realization of the climate shock. Finally, I use a dummy variable instead of a continuous index equals to one for a drought of any intensity, identified when the values of the Spei are below -1. The results are qualitatively similar but are generally insignificant. Only the sex composition of older siblings tend to alter how drought affects the outcome. I employ the same strategy for flood (values of the Spei index exceed 1) from columns 16 to 18. Only the specification with the composition of older siblings gives significant coefficients. However, floods decrease the likelihood of secondary school dropout especially for children with few older sisters, which contrasts with previous findings.

Following sub-section 4.1, I allow variables on sibling composition to vary over time and calculate sex ratios and sibling size among ever-born children in a given year (Table 2.16, columns 1, 2, 3). Estimates give some indications on the effect of having an additional sibling of a particular gender when affected by adverse climate conditions. However, it introduces additional endogeneity bias related to fertility preferences. The estimations corroborate previous findings. Besides, it is reasonable to suppose that a climate shock 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, I restrict the analysis on biological siblings eligible to be in school, aged between 6 and 20 years old in time t.[48] In Table 2.16, I investigate the the heterogenous effect of climate shock according to sibling gender composition controlling for sibling size, calculated among all (column 4), younger (column 5) and older (column 6) siblings. Although the results are qualitatively similar, coefficients present large level of significance (almost significant at the 1% size test) but with lower magnitude than in section 5.

From columns 7 to 9, I restrict the analysis to families with complete fertility, that is families where the mother is more than 45 years old at the time of shock. This reduces considerably the sample size. The main effect associated to the interaction between shock and siblings' gender composition disappears except when I calculate sex ratio among younger siblings. Having relatively more younger brothers alleviates the impact of unfavourable agricultural conditions on early school leaving, though the effect is significant at the 10% level.

I need to test for alternative mechanisms behind the differential effects that I find in education. One possibility is that families with a son are less responsive because they tend to be richer, have fewer children, or have more equal gender attitudes. Examining the effect of an income shock on families with a first-born son would allow me to compare families with close characteristics and preferences. Most interestingly, the sex composition of following

[48]More than 91% of individuals in the sample complete their education by 18 years. I extend the age range to 20 years old to consider potential delays in schooling.

**Table 2.16 –** Endogeneity of family composition: test of different alternatives

| | Siblings in t | | | School-aged siblings in t | | | Complete fertility | | | First born son | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sibling group | All | Young | Old | All | Young | Old | All | Young | Old | All | Young | Old |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Shock | 0.016 | -0.016 | -0.019 | 0.018 | -0.013 | 0.004 | -0.071 | -0.073 | -0.215** | 0.012 | -0.017 | -0.083 |
| | (0.027) | (0.033) | (0.025) | (0.023) | (0.035) | (0.025) | (0.090) | (0.087) | (0.099) | (0.042) | (0.035) | (0.061) |
| Shock × Sex ratio | -0.056** | -0.062*** | -0.007 | -0.048*** | -0.051*** | -0.011 | -0.100 | -0.080* | 0.011 | -0.052 | -0.057** | 0.030 |
| | (0.027) | (0.022) | (0.018) | (0.017) | (0.017) | (0.019) | (0.068) | (0.043) | (0.080) | (0.041) | (0.027) | (0.052) |
| Shock × Siblings | -0.002 | 0.011 | -0.007 | -0.003 | 0.017 | -0.012* | 0.002 | 0.012 | 0.015 | -0.002 | 0.008 | -0.003 |
| | (0.003) | (0.008) | (0.006) | (0.004) | (0.013) | (0.007) | (0.007) | (0.018) | (0.010) | (0.004) | (0.005) | (0.008) |
| Father | 0.011 | -0.001 | 0.017 | -0.005 | -0.012 | -0.026 | -0.111 | -0.128 | -0.112 | -0.010 | -0.011 | -0.005 |
| | (0.024) | (0.031) | (0.026) | (0.027) | (0.031) | (0.032) | (0.070) | (0.105) | (0.071) | (0.048) | (0.051) | (0.050) |
| Mother | -0.090** | -0.143*** | -0.094** | -0.089** | -0.150*** | -0.086** | | | | -0.092** | -0.120*** | -0.093** |
| | (0.042) | (0.043) | (0.044) | (0.040) | (0.041) | (0.043) | | | | (0.042) | (0.042) | (0.039) |
| Sex ratio | 0.009 | 0.077* | 0.021 | 0.013 | 0.040 | 0.063* | | | | | | |
| | (0.055) | (0.044) | (0.481) | (0.026) | (0.026) | (0.037) | | | | | | |
| Siblings | 0.011 | -0.013 | | 0.002 | -0.009 | -0.003 | | | | | | |
| | (0.009) | (0.011) | | (0.005) | (0.010) | (0.009) | | | | | | |
| Observations | 8,365 | 7,000 | 6,832 | 8,027 | 6,098 | 6,244 | 2,080 | 1,410 | 1,975 | 4,015 | 3,369 | 3,229 |
| R-squared | 0.002 | 0.004 | 0.002 | 0.002 | 0.004 | 0.003 | 0.005 | 0.006 | 0.007 | 0.002 | 0.003 | 0.003 |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Conley standard errors in parentheses. Columns (1)-(3) reproduce the main analysis using the sibling composition at time t, which vary over time. Columns (4)-(6) use instead school-aged siblings (between 6 and 20 years old) at time t. Columns (7)-(9) estimates the effect using only families who complete their fertility at time of shocks (the mother is above 45 years old). Columns (10)-(12) reproduce the analysis using families with a first born son. SPEI_gs is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (2), (5), (8), (12) and older siblings only in columns (3), (6), (9), (11). Father is a dummy variable that equals one if the father is alive. Mother is analogously defined. All estimates include individual and year fixed effects.

births might be less endogenous in a society of pro-male biases. The results are reported in the last three columns of Table 2.16. Focusing on families with a first-born son reduces the sample size and naturally decreases the level of precision in the coefficient estimates. Nevertheless, the heterogeneous effect of the shock according to the sex composition of younger siblings is also identified in column 11 (coefficient in the second line).

Moreover, I reproduce the estimates using a dummy equals to 1 if children attended school during the last 12 months as the dependant variable in Table 2.17. The latter reflects school attendance instead of permanent dropout. I exploit the KHDS survey and control for age, the presence of the father and the mother in the household. I include individual and survey fixed effects and allow for spatial correlation of shocks in the standard errors. For both girls and boys, sibship gender composition largely alters how shocks impact children's participation in school. Surprisingly, shock increases boys' probability of being in school especially for those with one sister (the reference category in the first line of column 2). Given the negative and highly significant coefficient on Shock$\times$Sex ratio, the positive effect decreases with the presence of biological brothers in the family. This contrasts with the findings so far. Nevertheless, the remaining columns are in line with the elicited effects identified throughout the paper for males and females: a female-skewed sex ratio exacerbates the effect of income shocks on education whereas a male-skewed sex ratio largely improves school participation during shocks. The results are mainly driven by the sex composition of younger siblings.

Finally, to explore the relationship between a child's gender and drought consequences at different stages of the educational process, I estimate individual enrollment regressions broken down by the level of education in time t in Table 2.18. The sample is limited to individuals with less than the primary level in the first four columns and to children with at least primary education in the last four. Climate shocks decrease drastically school attendance for children with at least primary education. The higher the value of the shock variable, the higher the probability of not going to school, given the negative coefficients of Shock in the first line of columns 6, 7 and 8. Notice that women are more negatively affected by unfavourable agricultural conditions since the interaction between shock and being a female

**Table 2.17 –** The effect of drought on school attendance in the last 12 months, depending on sibling composition and child's gender

| Dependent variable | School attendance during the last 12 months | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Boys | Girls | Boys | Girls | Boys | Girls |
| Sibling group | | All | | Young | | Old | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Shock | -0.018 | 0.380** | -0.759*** | -0.522** | -0.833*** | 0.047 | -0.154 |
| | (0.024) | (0.187) | (0.216) | (0.239) | (0.230) | (0.136) | (0.126) |
| Shock × Female | -0.017 | | | | | | |
| | (0.024) | | | | | | |
| Shock × Sex ratio | | -0.328** | 0.872*** | 0.359** | 1.271*** | -0.003 | -0.218 |
| | | (0.133) | (0.328) | (0.169) | (0.109) | (0.153) | (0.236) |
| Shock × Siblings | | -0.068** | 0.021 | 0.087** | 0.033 | -0.101*** | 0.018 |
| | | (0.026) | (0.038) | (0.041) | (0.065) | (0.027) | (0.048) |
| Father in hh | -0.096*** | -0.104*** | -0.272*** | -0.172*** | -0.314*** | -0.086*** | -0.227*** |
| | (0.023) | (0.024) | (0.032) | (0.038) | (0.041) | (0.020) | (0.036) |
| Mother in hh | -0.132*** | -0.234*** | -0.134*** | -0.235*** | -0.198*** | -0.212*** | -0.210*** |
| | (0.019) | (0.028) | (0.027) | (0.039) | (0.045) | (0.027) | (0.035) |
| Age | 0.776 | 0.358 | 1.679** | 0.214 | 1.705* | -0.452 | 1.287 |
| | (0.688) | (0.790) | (0.796) | (0.719) | (0.943) | (0.982) | (1.020) |
| Observations | 4,143 | 891 | 891 | 715 | 732 | 754 | 715 |
| Individuals | 1,790 | 811 | 823 | 660 | 695 | 691 | 681 |
| R-squared | 0.027 | 0.085 | 0.149 | 0.129 | 0.244 | 0.080 | 0.166 |

*** p<0.01, ** p<0.05, * p<0.1. Standard errors corrected for spatial correlation in parentheses. Results are estimated using KHDS-1, 2 and 3 with linear regressions. The dependent variable is a dummy that equals one if the individual attended school during the last 12 months and zero otherwise. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (4), (5) and older siblings only in columns (6), (7). Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. All estimates include individual and survey fixed effects.

is negative, at the 5% size test. The latter finding is consistent with deeper (heterogeneous) effects of droughts on girls' probability of permanent school dropout identified in the main analysis.

## 2.8 Conclusion

This paper studies the existing relationship between climate conditions, sibling composition and children's human capital when markets are incomplete, in a rural region of Tanzania. The study reveals that uninsured households withdraw specific children from school in response to unanticipated income shocks. The results are sensitive to the gender composition of siblings. Children with relatively more sisters than brothers, especially younger sisters, do particularly worse in times of economic hardship. The heterogeneous effect is larger among girls than boys. Globally, the finding indicates that parents adopt strategies decreasing human capital and worsening gender gap within household when the economic environment becomes uncertain. An increase in liquidity to buffer the shock comes mainly from a decrease in girls' school attendance. In addition, boys increase their total work, including household work, and constitute a reliable asset for families affected by income shortfalls. As a result, girls tend to increase their total labor in the absence of younger males during shocks. Finally, the long-term analysis points out permanent effects of climate shocks on children's human capital development and thus future earnings.

This paper shows that mechanisms protecting rural household from agricultural shocks through saving options, are likely to increase educational attainment and reduce gender gap both in education 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 children's educational attainment, gender-equity and thereby economic development.

**Table 2.18 –** The effect of drought on school attendance in the last 12 months, according to educational grade

| Sample restrictions | Child with less than primary in t | | | | Child with at least primary in t | | | |
|---|---|---|---|---|---|---|---|---|
| Sibling group | | All | Young | Old | | All | Young | Old |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Shock | 0.011 | 0.021 | 0.040** | -0.001 | 0.001 | -0.200** | -0.274** | -0.331** |
| | (0.012) | (0.021) | (0.015) | (0.012) | (0.077) | (0.100) | (0.121) | (0.134) |
| Shock × Female | -0.004 | | | | -0.205** | | | |
| | (0.010) | | | | (0.093) | | | |
| Shock × Sex ratio | | -0.007 | -0.008 | 0.005 | | 0.008 | 0.103 | 0.125 |
| | | (0.037) | (0.015) | (0.018) | | (0.162) | (0.143) | (0.141) |
| Shock × Siblings | | -0.001 | -0.007* | 0.003 | | 0.015 | 0.029 | 0.037 |
| | | (0.001) | (0.004) | (0.002) | | (0.009) | (0.020) | (0.023) |
| Father in hh | -0.109*** | -0.107*** | -0.148*** | -0.071** | -0.078 | -0.082 | -0.037 | -0.124 |
| | (0.036) | (0.036) | (0.047) | (0.032) | (0.104) | (0.110) | (0.103) | (0.124) |
| Mother in hh | -0.098*** | -0.100*** | -0.130*** | -0.104*** | -0.028 | -0.050 | -0.030 | -0.044 |
| | (0.025) | (0.025) | (0.032) | (0.027) | (0.075) | (0.081) | (0.077) | (0.118) |
| Age | 0.024*** | 0.025*** | 0.025*** | 0.019*** | -0.001 | -0.011 | -0.009 | -0.019 |
| | (0.007) | (0.007) | (0.009) | (0.006) | (0.057) | (0.058) | (0.055) | (0.062) |
| Observations | 2,448 | 2,411 | 2,028 | 1,979 | 1,022 | 1,013 | 860 | 820 |
| Individuals | 1,267 | 1,248 | 1,006 | 1,068 | 754 | 746 | 596 | 625 |
| R-squared | 0.041 | 0.041 | 0.074 | 0.037 | 0.014 | 0.008 | 0.009 | 0.020 |

*** p<0.01, ** p<0.05, * p<0.1. Standard errors corrected for spatial autocorrelation in parentheses. Results are estimated using KHDS-1, 2 and 3 with linear regressions. The dependent variable is a dummy that equals one if the individual attended school during the last 12 months and zero otherwise. Shock is the average of the spei index in each grid calculated during the six months of the growing seasons (March, April, May and October, November, December), in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (4), (5) and older siblings only in columns (6), (7). Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. All estimates include individual and survey fixed effects.

# 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. (2009). Flexibility of household structure child fostering decisions in burkina faso. *Journal of Human Resources 44*(4), 976–997.

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.

Alam, S. A. and C. C. Pörtner (2018). Income shocks, contraceptive use, and timing of fertility. *Journal of Development Economics 131*, 96–103.

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.

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

Angrist, J. D. and A. B. Krueger (1999). Empirical strategies in labor economics. In *Handbook of labor economics*, Volume 3, pp. 1277–1366. Elsevier.

Arndt, C., W. Farmer, K. Strzepek, and J. Thurlow (2012). *Climate change, agriculture and food security in Tanzania*. The World Bank.

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

Attanasio, O. P. (2015). The determinants of human capital formation during the early years of life: Theory, measurement, and policies. *Journal of the European Economic Association 13*(6), 949–997.

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.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of political Economy 87*(6), 1153–1189.

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

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

Behrman, J. R. (2010). Investment in education—inputs and incentives. In *Handbook of development economics*, Volume 5, pp. 4883–4975. Elsevier.

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.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Boneva, T. and C. Rauh (2018). Parental beliefs about returns to educational investments—the later the better? *Journal of the European Economic Association 16*(6), 1669–1711.

Bun, M. J. and T. D. Harrison (2019). Ols and iv estimation of regression models including endogenous interaction terms. *Econometric Reviews 38*(7), 814–827.

Butcher, K. F. and A. Case (1994). The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics 109*(3), 531–563.

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.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica: Journal of the Econometric Society 56*, 63–90.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy 100*(3), 437–467.

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

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics 92*(1), 1–45.

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.

Dammert, A. C. (2010). Siblings, child labor, and schooling in nicaragua and guatemala. *Journal of Population Economics 23*, 199–224.

De Janvry, A., F. Finan, E. Sadoulet, and R. Vakis (2006). Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks? *Journal of development economics 79*(2), 349–373.

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.

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).

Dube, O. and J. F. Vargas (2013). Commodity price shocks and civil conflict: Evidence from colombia. *The review of economic studies 80*(4), 1384–1421.

Duryea, S., D. Lam, and D. Levison (2007). Effects of economic shocks on children's employment and schooling in brazil. *Journal of Development Economics 84*, 188–214.

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

Ejrnæs, M. and C. C. Pörtner (2004). Birth order and the intrahousehold allocation of time and education. *The Review of Economics and Statistics 86*(4), 1008–1019.

Fafchamps, M. and J. Wahba (2006). Child labor, urban proximity, and household composition. *Journal of development Economics 79*(2), 374–397.

FAO (2019). The united republic of tanzania resilience strategy 2019–2022. Technical report, Rome. 32 pp.

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.

Harari, M. and E. L. Ferrara (2018). Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics 100*(4), 594–608.

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

Homdee, T., K. Pongput, and S. Kanae (2016). A comparative performance analysis of three standardized climatic drought indices in the chi river basin, thailand. *Agriculture and Natural Resources 50*(3), 211–219.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of sciences 107*(35), 15367–15372.

Imbert, C., M. Seror, Y. Zhang, and Y. Zylberberg (2016). Internal migration and firm growth: Evidence from china. Technical report, Mimeo.

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. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy 114*(3), 538–575.

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.

Kruger, D. I. (2007). Coffee production effects on child labor and schooling in rural brazil. *Journal of development Economics 82*(2), 448–463.

Kuepie, M., C. J. Nordman, and F. Roubaud (2009). Education and earnings in urban west africa. *Journal of Comparative Economics 37*(3), 491–515.

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.

Lambert, S. and P. Rossi (2016). Sons as widowhood insurance: Evidence from senegal. *Journal of Development Economics 120*, 113–127.

Lambert, S., D. Van de Walle, and P. Villar (2017). *Marital trajectories and women's well-being in Senegal*. The World Bank.

Larsen, A. F., D. Headey, and W. A. Masters (2019). Misreporting month of birth: Diagnosis and implications for research on nutrition and early childhood in developing countries. *Demography 56*(2), 707–728.

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.

Milazzo, A. (2014). *Son preference, fertility and family structure: Evidence from reproductive behavior among Nigerian women*. The World Bank.

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.

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

Ponczek, V. and A. P. Souza (2012). New evidence of the causal effect of family size on child quality in a developing country. *Journal of Human Resources 47*(1), 64–106.

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 T. P. Schultz (1989). Schooling, information and nonmarket productivity: contraceptive use and its effectiveness. *International Economic Review*, 457–477.

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.

Skoufias, E. and S. W. Parker (2006). Job loss and family adjustments in work and schooling during the mexican peso crisis. *Journal of Population Economics 19*(1), 163–181.

Söderbom, M., F. Teal, A. Wambugu, and G. Kahyarara (2006). The dynamics of returns to education in kenyan and tanzanian manufacturing. *Oxford Bulletin of Economics and Statistics 68*(3), 261–288.

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*.

United Republic of Tanzania (2012). National sample census of agriculture 2007/2008: Regional report–kagera region.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate 23*(7), 1696–1718.

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

Wedgwood, R. (2007). Education and poverty reduction in tanzania. *International Journal of Educational Development 27*(4), 383–396.

World Bank (2004a). Tanzania - secondary education development program project. Technical report, Washington DC: World Bank, http://documents.worldbank.org/curated/en/426501468760789334/Tanzania-Secondary-Education-Development-Program-Project.

World Bank (2004b). User's guide to the kagera health and development survey datasets.

World Bank (2015). Tanzania - agricultural sector risk assessment. agriculture global practice technical assistance paper. Technical report, Washington DC: World Bank Group, http://documents.worldbank.org/curated/en/248961468001158010/Tanzania-Agricultural-sector-risk-assessment.

Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal 125*(588), F347–F371.

# 3 Inter-generational Support and Parental Investment Decisions in Offsprings: Education and Marriage in Indonesia

This paper investigates the effect of parental investment in children's education on their old-age support, by child's gender. Specifically, we explore how a shock on education costs affects children's education and the resulting old-age support transfers from offspring to elderly parents. Firstly, we construct a model where parents anticipate the consequences of their decisions in the first period on monetary support and assistance that they expect to receive in the second period. The model highlights two main channels explaining the relation between children's human capital and assistance to elderly. The first one is the impact of education on expected wages, the labor-market returns to education. The second is the impact of education on the quality of the spouse and on the bargaining power within the newly formed household, that we name the marriage-market returns to education. Then, we test the theoretical predictions using Indonesian data. We exploit human capital variations created by a large school construction program by the Indonesian government during the 70s to estimate the effect of children's education on support to elderly parents, by child's gender. The results are consistent with the model. We find that a substantial fraction of human capital gains, generated by exposure to the educational reform, is shared with the parental generation. Human capital clearly increases financial transfer to old parents. However, contrary to the theoretical predictions, empirical estimates suggest that the strength of the effect does not vary across married sons and married daughters.

***Keywords***: *Family Transfers, Old-Age Support, Education, Marriage.*

## 3.1 Introduction

Most people in low income countries do not have access to formal social security, covering only 25% of the population on average. The social security coverage varies across and within countries. In Indonesia, only 8.6% of all economically active citizens were covered for an old-age monthly pension in 2011, the vast majority of whom were in the formal sector (Ortiz, 2014). As a result, security in old age relies on informal safety nets: the extended family and the clan. According to Cox, Galasso, and Jimenez (2006), transfers from young to old exceed those going from old to young family members in developing countries. A large body of the literature in education assumes that parents' incentives for children's human capital investment are mainly related to transfers for old-age support in the absence of savings and insurance. In particular, the old-age security hypothesis often explains any differential treatment in the allocation of educational resources across gender. In patrilocal society, daughters tend to leave the natal family upon marriage. Since parents invest more in children when they expect greater support (Becker, Murphy, and Spenkuch, 2016), parents might invest less on girls' education anticipating that any returns from these daughters' schooling attainment would accrue to the in-laws or the husband rather than to them. Indeed, there is empirical evidence which strongly indicates that sons and daughters are not treated equally in terms of human capital investment (Strauss and Thomas, 1995, Behrman, 1997, Quisumbing, Estudillo, and Otsuka, 2003).

The objective of the current study is to investigate the effect of parental investment in children's education on their old-age support, by the child's gender. More specifically, we are interested in exploring how a shock on education costs might affect investment in the education of children and the resulting old-age support transfers from offspring to elderly parents.

We first investigate the research question theoretically. We rely on the models of Horowitz and Wang (2004) and Kumar (2013), except that we introduce a marriage dimension. We model a family composed of parents and one child, a daughter or a son.[1] In the first period, parents allocate the time endowment of their child between schooling and child labor, anticipating the consequences of their decisions on monetary support and assistance that they expect to receive in the second period. In particular, parents anticipate that more education to their child is associated with higher wages. Wealthier children are more able to provide financial assistance to family members (Cox, 1987, Lillard and Willis, 1997). We name this channel the labor-market returns to education. We assume here, as in Kumar (2013), that males earn more than females conditional on a given level of education. But, unlike him, we assume that females have higher returns to schooling, as this is supported by empirical evidence on Indonesia (Deolalikar, 1993, Behrman, Deolalikar, et al., 1995), as well as our own estimates. The novelty of our model is that parents also anticipate that higher education enables their child to marry a spouse of better quality and thus to obtain a higher joint income.[2] Additionally, more education impacts the bargaining power of their child within his own household, and thus his control over his household's resources. We name these indirect effects the marriage-market returns to education. This is in line with Lee (2007) who underlines that bargaining power between spouses plays an important role in household resources allocation and thus on monetary transfers to parents, since it is plausible that the wife cares more about her parents than about the husband's parents, and vice-versa (Ham and Song, 2014). Specifically, we model the bargaining power as a function of relative wages of spouses, which indirectly depends on education investments.[3] Thus, in the second period, we

[1] We also provide in the Appendix an extended version of the model in which families are composed of two children, a daughter and a son.

[2] Fafchamps and Quisumbing (2005, 2007) or Anukriti and Dasgupta (2017) indicate that there is ample evidence of the assortative matching hypothesis in developing countries. Moreover, empirical evidence suggests that assortative matching on human capital attributes has increased relative to sorting based on parental wealth and physical capital (Quisumbing and Hallman, 2003, Quisumbing, Behrman, Maluccio, Murphy, and Yount, 2005). Using Philippines data, Boulier and Rosenzweig (1984) find that there are payoffs to spouse search and assortative mating with respect to schooling.

[3] Age at marriage, the difference in age and education between spouses are predicted outcomes for greater household bargaining power in East Africa (Mabsout and Van Staveren, 2010) or Korea (Ham and Song, 2014). Younger bride is associated to less autonomy and are more likely to conform to the wishes of the husband and his family (Anukriti and Dasgupta, 2017, Caldwell, Reddy, and Caldwell,

assume that the adult child is married assortatively in terms of education, and bargains with his spouse on the splitting of the household resources. Depending on this split, he decides on how much to privately consume and how much to redistribute to his old parents.

Results from our model are systematically compared to those derived from a similar model in which marriage-market returns are obliterated. Interestingly, we find that in our setting a marginal change in education positively affects the transfer of daughters more strongly than the one of sons, marriage-market returns to education being taken into account or not. However, when marriage-market returns to education are included, this gender differential is all the stronger. The reason for this result is that when marriage-market returns are discarded, marginal transfers relative to education depend only on marginal wages. In contrast, when marriage-market returns are taken into account, marginal transfers also depend on the marginal bargaining power within the couple, which we show to be always positive for females and negative for males. As a result, observing a significant larger impact of education investment on the transfers of daughters indicates that marriage-market returns matter for old-age support. But, we also find that at similar levels of education, the amount of monetary transfer is always higher for sons, marriage-market returns being taken into account or not. Moreover, we find that a decrease in education costs affects positively more investment in education and transfers of daughters rather than the ones of sons. This latter impact on transfers is all the greater as marriage-market returns to education are taken into account.

Secondly, we conduct an empirical assessment to test the theoretical predictions. We explore one of the largest schooling reforms during the 70s to investigate the impact of increased educational investment on assistance to elderly parents, by the child's gender. The Indonesian government built 56,000 primary schools from 1974-1975 to 1978-1979 budget years.[4] The Inpres reform brought two important changes: it provided a school in virtually

1983, Jensen and Thornton, 2003). However, education attainment contributes to delay the onset of marriage among younger cohorts of women (Garenne, 2004, Gyimah, 2008, Carmichael, 2011).

[4]In the literature, Duflo (2001), Akresh, Halim, and Kleemans (2018) and Ashraf, Bau, Nunn, and Voena (2018) investigate the impact of the INPRES policy reform as well. These papers share a common empiricaly strategy and use the Susenas nationally representative Indonesian data to study the effects of the school construction program. Akresh et al. (2018) found that the program increases years of schooling for both men and women. Positive results among male-wage earners are identified in Duflo (2001) and Ashraf et al. (2018) find positive and significant effect only for girls belonging to a bride-price ethnic groups.

every village, independently of village resources, and it temporarily took over the financing of many of the activities which had formerly been financed by the local community (World Bank, 1989). We use this reform as a natural experiment to identify how children's human capital accumulation affects the assistance to old parents. To do so, we use a large cross-country of individuals born between 1950 and 1973 from the 2000 Indonesia Family Life Survey (IFLS). The data are ideally suited for this research because the survey provides detailed information about monetary transfers and services exchanged between non-coresident parents, children and siblings. Besides data availability, Indonesia provides a suitable case to study parental investments in the form of children with no access to formal old-age insurance for the majority of people, the informal workers. The analysis focuses on married couples so that the decision-making unit is always comparable.

Our identification strategy relies on exogenous variations in exposure to the Inpres reform across cohorts and space to identify the effect of a decrease in education costs on children's educational attainment. We allow the impact to differ by the child's gender, given the preliminary assumptions. Secondly, we use exposure to the reform as an instrumental variable and investigate the causal impact of children's human capital on transfer to old parents. In all cases, we include sub-district and birth year fixed effects to control for omitted time-varying and region-specific unobservables. We find that a decrease in education costs raises children's educational attainment for both males and females. A one unit increase in educational grade raises the likelihood of giving financial assistance by 18%. The intensive marginal effects of education on transfer is more precisely estimated and is more robust. On average, human capital increases by 62% the amount given to elderly parents. Contrary to theoretical findings, we do not identify stronger effects associated to a particular gender.

Finally, we provide empirical evidence on the plausibility of the mechanisms. To this end, we investigate the labor-market returns to education in Indonesia for males and females using information on individual's earnings. The results show that education increases earnings and the estimated marginal return to additional level of education is significantly higher for females than for males. Then, we explore whether education effectively affects old-age transfer through additional bargaining power in the newly formed household. We use exogenous variation in married women's intra-household bargaining across ethnic groups, with different post-marital residence norms, to shed light on how it affects the impact of education on old-age support. Women's intra-household bargaining has an indirect effect. Specifically, the relationship between married women's education and financial transfer to old parents is weaker among ethnic groups where women have lower bargaining power within the household. Empirical results are thus consistent with a model in which both labor- and marriage-markets returns play a role.

Our study relates to the large theoretical literature studying monetary transfers within the family. One strand of this literature specifically considers the old-age security hypothesis which claims that parents invest in children's education to receive old-age transfers. Some articles examine the amount of transfers from children to parents to be exogenously determined (Raut, 1990), whereas others model the decision of both parents and children in an inter-generational strategic game with endogenous fertility (Zhang and Nishimura, 1993, Raut, 1997) or with exogenous fertility (Raut and Tran, 2005). These models systematically focus on labor-market returns to education only. However, taking into account the marriage-market returns to education could arguably change the investment decisions of parents, as well as the resulting transfers to old parents. Our paper thus aims at filling this gap in the literature.

Additional features help distinguish this paper from existing empirical research on old-age transfer. While previous studies generally find positive correlations between children's human capital and old-age support (Lee, Parish, and Willis, 1994, Lillard and Willis, 1997, Lucas and Stark, 1985, Vanwey, 2004), the extent to which these results can have a causal interpretation seems limited. In this paper, we try to overcome the inherent endogeneity of children's education. In this sense, our paper closely relates to De Neve and Fink (2018), where they use the 1974 Tanzania education reform to estimate the effect of children's primary schooling attainment on parental survival.

Also, our study contributes to the theoretical literature on differentiated treatments across children in terms of human capital investment. Horowitz and Wang (2004) consider a family with two children with different human capital accumulation functions. They show that in the presence of intra-household and inter-generational transfers, specialization of heterogeneous children is efficient and may even be beneficial to the disfavored child. Kumar (2013) extends this framework to study the effects of two types of gender biases: the son-preference by parents and the earnings function bias towards male. In a framework with uncertainty, Lilleor (2015) suggests that human capital diversification among children might be the optimal decision for parents in order to minimize risk. Similarly, to our knowledge, this literature has neglected marriage-market returns to education so far.

The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 describes the data used to examine the empirical validity of the model. We explain our identification strategy in Section 4 and present the results in Section 5. Section 6 performs robustness checks and Section 7 gives some empirical evidence on the mechanisms. Finally, Section 8 concludes.

## 3.2 Model

Our theoretical model helps us explore how a shock on education costs when children are young impacts the way parents invest in their education, as well as the resulting consequences in terms of old-age support from adult children to old parents. As stated above, education has both labor-market and marriage-market returns. While most models on investment in education and old age support usually focus on the former, our model aims at disentangling these two channels. To do so, we will systematically compare results and predictions derived from our model with marriage- and labor-market returns to education with those derived from the same model with labor-market returns only. Our goal is to better understand how schooling affects support from children to elderly parents. Moreover, our model helps explore how gender impacts outcomes.

### 3.2.1 The environment

There are two periods. In the first period, parents strategically invest in the education of their child in order to maximize their prospects for old age support. There is no formal credit and insurance markets in our economy, so the only way parents can receive assistance when they get older is with the help of their offspring. In the second period, the adult child is married and provides assistance to his old parents. A grown-up child supports his old parents because of prevalent social norms in Indonesia. Frankenberg, Lillard, and Willis (2002) explain that strong ties between parents and adult children are attributed to social norms that deem children responsible for their elderly parents' welfare in Indonesia.

A family consists of parents $p$ and one child, a daughter $d$ or a son $s$.[5] Parents are treated as a single decision and consumption unit. In the first period, parents supply $L$ units of labor inelastically and receive a wage normalized to one. The young child $k = \{d, s\}$ is endowed with one unit of time that parents allocate between child labor $l^k$ payed at a rate $\omega < 1$, and education $e^k$ at cost $\gamma$.[6] The budget constraint of the family in the first period is thus

$$C + (\omega + \gamma)e^k \leq L + \omega \tag{3.1}$$

with $C$ representing the consumption of the whole family in the first period.

In the second period, parents are old and the grown-up child is married, works, consumes and provides assistance to his elderly parents. The schooling decision made by the parents in the first period impacts both the labor-market and the marriage-market outcomes of their adult child. Let us first consider the labor market side. Assumptions made on the wage functions derive from our own empirical analysis and on empirical papers estimating determinants of wages in Indonesia (Deolalikar, 1993, Behrman and Deolalikar, 1995, Raut and

[5] In the Appendix, we explore an extended version of the model where a family is composed of parents and two children - a daughter and a brother.

[6] Cost of schooling encompasses tuition fees, schooling equipment but also geographic distance and commuting from home to school. In the empirical analysis, we study the impact of schools' construction on investment in education and the resulting old-age support transfers. Schools' construction is interpreted as a positive shock on the cost of education as it reduces geographic distance from home to school for some children.

Tran, 2005, Comola and de Mello, 2013). The wage functions for females and males are respectively denoted $w^f(\cdot)$ and $w^m(\cdot)$. We assume that these wage functions are increasing, concave and twice differentiable functions of education $e^k$.[7] We also assume that wages are strictly positive for individuals who never studied, so $w^f(0) > 0$ and $w^m(0) > 0$. Empirical evidence found in Indonesia on the gender wage gap shows that, controlling for education, females earn less than males but returns to schooling are higher for females than for males.[8] Accordingly, we thus assume that $w^m(e) > w^f(e)$ and $w^{f'}(e) > w^{m'}(e)$.[9] Let us notice here that these assumptions imply that the gender wage gap decreases with education, and that the education-elasticity of female wage, $\epsilon^f_{w,e}(e)$, is greater than education-elasticity of male wage, $\epsilon^m_{w,e}(e)$.

On the marriage market side, two dimensions are to be considered. First, we assume that there is perfect positive assortative matching in terms of education in the marriage market. For simplicity, we assume that an adult child $k$ with education $e^k$ marries a spouse $k'$ with education $e^{k'} = e^k$.[10] Thus the more educated a child, the better the quality of his spouse and the higher their joint income. Second, we allow for bargaining between spouses. In particular, we assume that newly formed households are collective, which means that there exists a stable decision process in the household which leads to Pareto-efficient outcomes.[11] The decision process attributes different Pareto weights to the two members of the household. These weights represent the bargaining power of each spouse in the decision process, and

[7]Actually, our estimated coefficients of $Education^2$ are significant and positive, which suggests that returns to education increase at an increasing rate. However, this coefficient is small. Similar results are found in the empirical section of Raut and Tran (2005). However, in their theoretical section Raut and Tran (2005) model the earnings function as an increasing and concave function of schooling. Most theoretical papers (Horowitz and Wang, 2004, Kumar, 2013) also use an increasing and concave function to model wage as a function of education.

[8]This stylized fact is not confined to Indonesia, it has also been observed in China (Zhang, Zhao, Park, and Song, 2005), rural Bangladesh (Pitt, Rosenzweig, and Hassan, 2012) and also in developed countries (Trostel, Walker, and Woolley, 2002, Dougherty, 2005).

[9]We thus depart here from Kumar (2013) who assumes that $w^m(e) > w^f(e)$ and $w^{m'}(e) > w^{f'}(e)$, without providing empirical evidence though.

[10]We could otherwise assume that a daughter of education $e^k$ marries a spouse of education $e^{k'} = g(e^k)$ with $g'(\cdot) > 0$ to capture the empirical fact that men are on average more educated than women in Indonesia. Our results remain qualitatively similar under this assumption, as it amplifies the wage gap mechanism between spouses already captured with the wage functions.

[11]Collective models of the household have been widely developed in the field of family economics after the seminal work of Chiappori (1988). See Browning, Chiappori, and Weiss (2014) for a recent survey.

depend on their relative wages.[12] So the bargaining powers within the couple of the married child $k$ are

$$\theta^f(e^k) = \frac{[w^f(e^k)]^\mu}{[w^f(e^k)]^\mu + [w^m(e^k)]^\mu} \quad \text{and} \quad \theta^m(e^k) = 1 - \theta^f(e^k) \tag{3.2}$$

where $\theta^f \in [0,1]$ denotes the bargaining power of the wife and $\theta^m$ the bargaining power of the husband. The parameter $\mu$ measures how important relative wages is for bargaining power in the society considered. So $\theta^f$ is equal to one-half when $\mu = 0$ and it approaches unity as $\mu \to \infty$ as soon as $w^f(e^k)$ surpasses $w^m(e^k)$, even very slightly. We assume here that $\mu \in ]0,1[$ in order to introduce a strictly positive lower bound on the bargaining power of the most vulnerable member in the household. Notice that we build on a partial equilibrium model here, as we do not take into account the consequences of optimal education choices by parents on the labor market and the marriage market. Instead, we assume that wages functions are fixed, and that the pool of potential partners is large and given.[13] Given our assumptions on the wage functions, we find that the impact of education on the bargaining power of females is positive, while it is negative for males. It is shown in the Appendix that $\theta^{f'}(e) \geq 0$ if and only if $\epsilon^f_{w,e}(e) \geq \epsilon^m_{w,e}(e)$, which is always true in our setting by assumption.

Within the household of the grown-up child $k$, the preferences of the two spouses are egotistic and there is no public good. In other words, we assume that the married child $k$ cares only about his own consumption $c^k$ and the well-being of his old parents $u^{pk}$, while his spouse cares only about his consumption $c^{k'}$ and the well-being of his own old parents $u^{pk'}$. So child $k$'s and his spouse's utility functions are respectively

$$u^k(c^k, u^{pk}) = \ln c^k + u^{pk} \quad \text{and} \quad u^{k'}(c^{k'}, u^{pk'}) = \ln c^{k'} + u^{pk'} \tag{3.3}$$

[12]Several articles also model Pareto weights as a function of relative wages, for instance Blundell, Chiappori, and Meghir (2005), Iyigun and Walsh (2007) and Baudin, De la Croix, and Gobbi (2015).
[13]In particular, we abstract from difficulties inherent to premarital investment games described for instance in Peters and Siow (2002) and Nöldeke and Samuelson (2015).

Thus, if child $k$ is a daughter, her collective household's utility function is

$$U^k(c^k, c^{k'}, u^{pk}, u^{pk'}) = \theta^f u^k(c^k, u^{pk}) + (1-\theta^f)u^{k'}h(c^{k'}, u^{pk'}) \tag{3.4}$$

In contrast, if child $k$ is a son, his collective household's utility function is

$$U^k(c^k, c^{k'}, u^{pk}, u^{pk'}) = \theta^f u^{k'}(c^{k'}, u^{pk'}) + (1-\theta^f)u^k(c^k, u^{pk}) \tag{3.5}$$

Old parents' utility depends on the monetary transfer $t^k$ provided by their adult child $k$.[14] So the utility function of child $k$'s old parents in the second period is

$$u^{pk}(c^{pk}) = \ln c^{pk} \quad \text{with} \quad c^{pk} = t^k \tag{3.6}$$

Overall, the objective of the parents is to optimally choose $e^k$ in the first period such that it maximizes their intertemporal utility

$$U^p = \ln C + u^{pk}(c^{pk}) \tag{3.7}$$

under the budget constraint defined in equation (3.1).

We solve the model backwards. We first solve for the optimal transfer of the married child, and study how it depends on his gender. Then we solve for the optimal schooling decision of the parents. Finally, we study how a shock on education costs affects investment in education by parents, and the resulting transfers from the adult child to his old parents. Here again we explore how gender impacts results.

[14] In the Appendix, we consider an extended version of the model where assistance from children can take the form of a monetary transfers and informal services. Unlike monetary transfers, informal services are time consuming.

### 3.2.2 Optimal transfers

The model is solved in the Appendix. Solving for the optimization problem of the adult child, we find that optimal transfers conditional on a education level are

$$\begin{aligned} t^{d*}(e^k) =& \theta^f(e^k)[w^f(e^k) + w^m(e^k) + y]/2 \quad \text{if} \quad k = d \\ t^{s*}(e^k) =& \theta^m(e^k)[w^f(e^k) + w^m(e^k) + y]/2 \quad \text{if} \quad k = s \end{aligned} \tag{3.8}$$

As one would expect, the optimal transfer given by an adult child $k$ to his old parents positively depends on his bargaining power within his household and on his couple's joint income, which includes the wages of the two spouses and non-labor income $y$.[15] This result is in contrast with classical models where only labor-market returns to education are considered. In fact, when marriage-market returns are obliterated, child $k$'s transfer only depends on his wage, so that $t^{d*}(e^k) = w^f(e^k)/2$ and $t^{s*}(e^k) = w^m(e^k)/2$. From equations (3.2) and (3.8), we observe that $t^{s*}(e) > t^{d*}(e)$ as we assumed that $w^m(e) > w^f(e)$.

**Testable prediction 1**: *At similar levels of education, the transfer of a son to his old parents is higher than the transfer of a daughter, marriage-market returns to education being taken into account or not.*

How do optimal transfers react to increased education? When marriage-market returns are discarded, optimal transfers obviously increase in education: $dt^{s*}(e)/de = w^{m'}(e)/2 > 0$ and $dt^{d*}(e)/de = w^{f'}(e)/2 > 0$ . In addition, we have that $dt^{d*}(e)/de > dt^{s*}(e)/de$, by assumption. How does the introduction of marriage-market returns to education change results? It is shown in the Appendix that the impact of education is unambiguously positive for females, due to the positive impact of education on their bargaining power described above. For males, the impact is less straightforward but also positive for $y$ not too large with respect to wages. The strength of this impact varies across gender. We show in the Appendix that it depends on education-elasticities of wages. For $\epsilon^f_{w,e}(e)$ high enough relative to $\epsilon^m_{w,e}(e)$, this impact is stronger for daughters than for sons, otherwise this impact is stronger for sons

[15]Non-labor income of the household can be interpreted as economies of scale associated with living as a couple, such as housing or heating costs. In our setting, this parameter rationalizes why individuals marry rather than remaining singles.

than for daughters. Equation (C.1) in the Appendix shows that $\theta^{f'}(e)$ is all the greater (and thus $\theta^{m'}(e)$ all the smaller) as education-elasticity of female wage is high relative to the one of male wage. In other words, the higher the difference $[\epsilon^f_{w,e}(e) - \epsilon^m_{w,e}(e^k)]$, the more women capture a larger part of their couple's income du to increased education. As a consequence, when $[\epsilon^f_{w,e}(e) - \epsilon^m_{w,e}(e^k)]$ is low, increased education has more impact on the transfer of the male than on the female, as husbands benefit from being married with more educated spouses without loosing too much bargaining power. In contrast, when $[\epsilon^f_{w,e}(e) - \epsilon^m_{w,e}(e^k)]$ is high enough, increased education has more impact on the transfer of females, as they obtain a much greater share of the household joint income due to their increased bargaining power.

In order to give clearer predictions, we need to assign functional forms to the wage functions, relying our estimates and the ones reviewed in the literature. To model the fact that the wage is an increasing and concave function of schooling, we use the square root function. We add a positive intercept to these functions to capture the fact that even uneducated individuals can expect a strictly positive wage on the labour-market. Our estimates tell us that for each additional unit of education, the wage of females is multiplied by $1.10$ and the wage of males is multiplied by $1.08$. Thus the ratio of the returns to education of females out of the one males is $1.10/1.08$. We will accordingly choose the parameters such that the derivative of the female wage with respect to education divided by this same derivative for the male wage is equal to this ratio. Our estimates also indicate that on average females earn 67% less than males, which we consider to be a surprisingly large proportion.[16] Our data show that on average, men earn around 600,000 rupiah and women around 318,000 rupiah (see Table 3.1), which reduces this proportion to about 50%. The most recent study estimating wage determinants with Indonesian data is Comola and de Mello (2013), who find that women earn on average around 17% less than men. We decided to strike a balance between their estimate and ours, and choose parameters such than women earn on average 30% less than men. Our chosen wage functions are thus $w^f(e) = 11\sqrt{e} + 1$ and $w^m(e) = 10.8\sqrt{e} + 4$. In Figure C.1

[16]OLS estimates on the effect of education on the log of wages are reported in Appendix, Table C.1. We allow non-linear effects and introduce the square term of years of schooling and educational grades in columns 3 and 4. Labor-market returns to education across gender are investigated in columns 5 and 6.

we plotted these wage functions, their derivatives, elasticities as well as the gender wage gap varying with education. We represent in Figure 3.1 how the difference $[\epsilon^f_{w,e}(e) - \epsilon^m_{w,e}(e^k)]$ and the bargaining powers vary with education for our chosen wage functions.

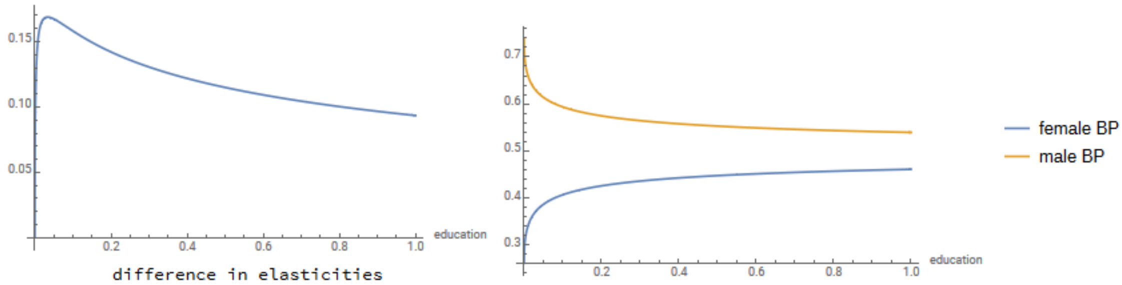



**Figure 3.1** – Elasticities and bargaining powers

We observe that for low levels of education, the difference in elasticities is large and then decreases with education. As a consequence, in low-education marriages, increased education strongly impacts the bargaining power of females, while this effect is lower for high-education marriages. Thus the impact of education on females' bargaining power is always positive but not linear: it is stronger for low-education marriages than for high-education marriages.

Figure C.2 in the Appendix shows optimal transfers and their derivatives for our chosen wage functions. We observe here that differences in elasticities is high enough, so the impact of education on transfers is larger for females than for males whatever the education level. So we find, just like when marriage-market returns are removed, that $dt^{d*}(e)/de > dt^{s*}(e)/de$. However, the amplitude of this difference and its evolution along education differ. Figure 3.2 shows how the difference $[dt^{d*}(e)/de - t^{s*}(e)/de]$ varies with education, distinguishing whether marriage-market returns to education are taken into account or not.

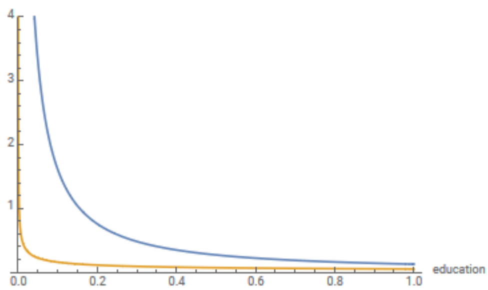



**Figure 3.2** – Gender difference in marginal transfers

When marriage-market returns are not taken into account, we observe that the difference between marginal transfers of males and females is very low at almost all education levels (except for extremely low education levels). In contrast, when marriage-market returns are taken into account this difference is large for low and medium education levels, and decreases for high education levels.

**Testable predictions 2:**

*(1) When marriage-market returns to education are not taken into account, the impact of education on transfers is slightly higher for daughters than for sons. Moreover, this difference does not vary much with education.*

*(2) When marriage-market returns are taken into account, the impact of education on transfers is significantly higher for daughters than for sons. Moreover, this difference varies with education: for low and middle education levels, this difference is large, while it decreases for higher levels of education.*

*(3) Observing a significant higher impact of education on the transfers of daughters than on the transfers of sons is an indication that marriage-market returns matter. Moreover, if marriage-market returns matter, we should observe that this differentiated impact by gender is higher for low and middle levels of education, and decreases for higher education levels.*

### 3.2.3 Optimal investment in education and impact of a shock on costs

Given optimal transfers conditional on an education level, we now study how parents optimally choose investment in education of their child in order to maximize their inter-temporal utility. We solve for optimal investments in education in the Appendix. We find the following expression of optimal investment in education for a daughter

$$\frac{1}{\gamma} = \Phi^d(e^{k*}) = \frac{t^{d*\prime}(e^{k*})e^{k*} + t^{d*}(e^{k*})}{t^{d*\prime}(e^{k*})[L + (1 - e^{k*})\omega] - \omega t^{d*}(e^{k*})} \quad \text{with} \quad k = d \tag{3.9}$$

Similarly we find the expression of the optimal investment in education of a son, depending in particular on the inverse of the cost of education $\gamma$: $1/\gamma = \Phi^s(e^{k*})$. We invert these functions, to obtain the expressions of optimal levels of education as a function of $1/\gamma$, respectively $e^{k*} = \Phi^{d^{-1}}(1/\gamma)$ for daughters and $e^{k*} = \Phi^{s^{-1}}(1/\gamma)$ for sons. We plot their respective curves as well as curves of their derivatives in Figure 3.3.[17]

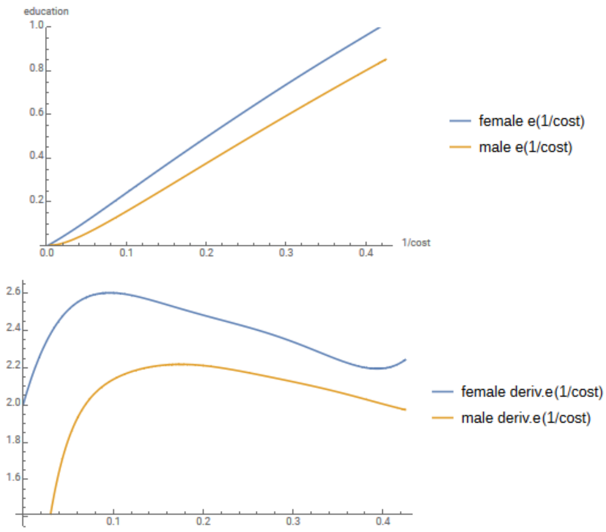



**Figure 3.3** – Optimal investments in education as a function of $1/\gamma$

We find that a decrease in the education costs (so an increase in $1/\gamma$) increases investment in education for daughters and sons, but this impact is much stronger on daughters. Moreover, we observe that this impact increases at an increasing rate for initially high education costs. This indicates that the impact of a decrease in costs on education investments is all the higher for families for whom education costs were initially very high.

Then, we study how a shock on education costs impacts optimal transfers to old parents. Results are displayed in Figure 3.4.[18] We observe here that a decrease in education costs increases the transfers of daughters and sons. Moreover this positive impact is stronger for daughters than for sons, especially for initially high education costs.

Finally, we explore how removing marriage-market returns to education impacts results. As already discussed, when marriage-market returns to education are removed, optimal trans-

[17]Parameters are set to: $y = 5, L = 10, \omega = 0.5, \mu = 0.75$. These parameters are here chosen such that they fit our assumptions : in particular, $y$ is not too high relatives to wages. We chose arbitrarily the other parameters. We should try at a new stage to estimate the parameter $\mu$, which describes the sensitivity of bargaining power to relative wages, as well as $\omega$, the child-to-adult wage ratio, for Indonesia. The parameter $L$ can be considered as the wealth of parents, so varying $L$ would help understand how investment in education depends on wealth.

[18]Locally fuzzy representations of curves in Figures 3.3, 3.4 and 3.5 are due to the approximation procedure to invert $\Phi^d$ and $\Phi^s$. These functions were much more difficult to invert that the equivalent functions without marriage-market returns to education, $\Psi^d$ and $\Psi^s$ presented below. For $\Phi^d$ and $\Phi^s$, we used the *FindFormula* command of Mathematica based on a plot of their curves, while for $\Psi^d$ and $\Psi^s$ we were able to simply use the command *InverseFunction*.

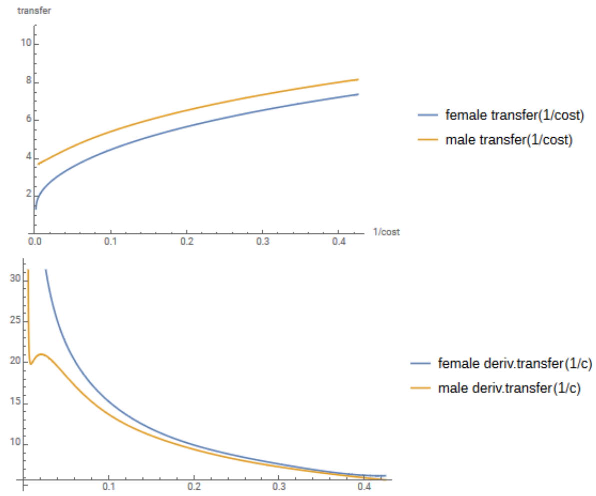



**Figure 3.4** – Optimal transfers as a function of $1/\gamma$

fers from adult children to parents only depend on wages. More precisely: $t^{d*}(e^k) = w^f(e^k)/2$ and $t^{s*}(e^k) = w^m(e^k)/2$. Optimal investment in education of a daughter must here satisfy the following condition

$$\frac{1}{\gamma} = \Psi^d(e^{k*}) = \frac{w^{f'}(e^{k*})e^{k*} + w^f(e^{k*})}{w^{f'}(e^{k*})[L + (1-\omega)e^k] - \omega w^f(e^k)} \quad \text{if} \quad k = d \qquad (3.10)$$

We also obtain the expressions of optimal levels of education when marriage-market returns are discarded as functions of the inverse of education costs, namely $e^{k*} = \Psi^{d^{-1}}(1/\gamma)$ for daughters and $e^{k*} = \Psi^{s^{-1}}(1/\gamma)$ for sons. Similarly, we find that a decrease in education costs increases investments in education for daughters and sons, and that this impact is stronger for daughters. Moreover, the decrease in education costs translates into higher transfers for daughters and sons, this impact being also stronger for daughters. However the amplitude of these effects differs whether marriage-market returns to education are taken into account or not. In Figure 3.5, we plot in the first graph the difference between the impact of costs reduction on education investment for daughters and the one for sons, with and without marriage-market returns. In the second graph of Figure 3.5, we plot the difference between the impact of costs reduction on transfers for daughters and the one for sons, with and without marriage-market returns.

We observe that the differentiated impact of a costs reduction on investment in education by gender is not significantly different whether marriage-market returns to education are

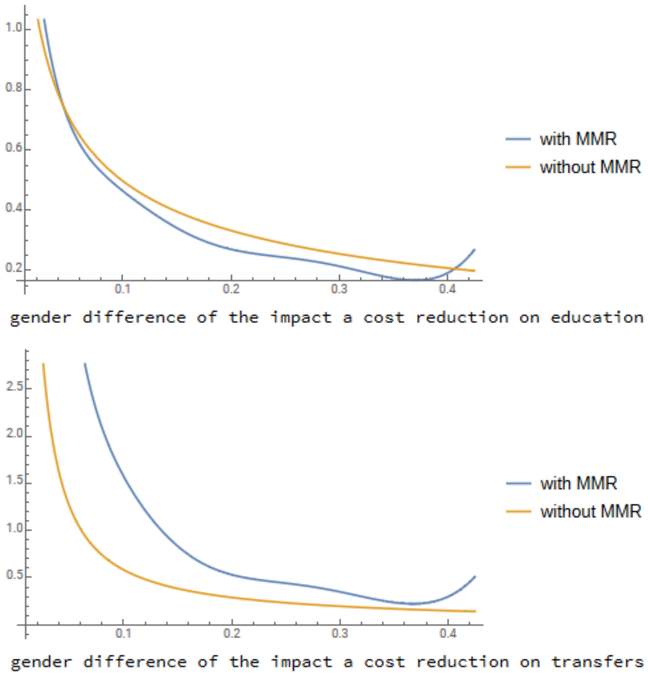



**Figure 3.5** – Gender differences

taken into account or not. However, the differentiated impact of a cost reduction on transfers is significantly greater when marriage-market returns to education are take into account, especially for initially high education costs.

The aforementioned results are summed up in the following set of testable predictions.

**Testable predictions 3:**

*(1) A decrease in education costs increases more the investment in education of daughters than the one of sons. This gender differential is not significantly different whether we take into account marriage-market returns to education or not.*

*(2) A decrease in education costs increase more the transfer of daughters than the one of sons. This gender differential is significantly wider when marriage-market returns are taken into account.*

*(3) Observing a significant stronger impact of cost reduction on the transfer of daughters than on the transfer of sons is an indication that marriage-market returns matter.*

Overall, we derive a number of testable predictions that we sum up here. We empirically estimate prediction 3: the effect of a positive variation in education costs on parental investment in human capital exploiting an exogenous reform in Indonesia that decreases distance to primary school. We allow different treatments by gender, given the preliminary assumptions. Then, we test prediction 2 by investigating the causal effect of human capital on financial assistance to old parents. We use exposure to the program as the first-stage of the 2SLS estimates. In the main analysis, we take as given the channels underlined in the model and directly test the relationship between education and financial transfer. The interpretation of the empirical results will help us to shed light on the competing indirect effects of labor- and marital-returns to education that we do not measure empirically. Subsequently, we will provide empirical evidence on the plausibility of the two mechanisms. First, we will investigate the returns to education in Indonesia for males and females. Secondly, we will explore whether education affects old-age transfer through additional bargaining power within the newly formed household.

## 3.3 Data

The main source of data for this study is the Indonesia Family Life Survey (IFLS), conducted by RAND, UCLA, and the Demographic Institute of the University of Indonesia. The IFLS covered 13 of the 27 provinces in Indonesia with a total of 321 randomly selected enumeration areas included in the survey. The IFLS is a longitudinal survey and represents 83% of the Indonesian population. The first round was collected in 1993/1994, and in 1997, 2000, 2007/2008, 2014/2015 for subsequent rounds. To have sufficient variation in the number of respondents who benefited from the massive primary school construction program starting in 1973, we use the third wave of the survey, done in 2000. In IFLS-1, 7,224 households were interviewed, and detailed individual-level data were collected from over 22,000 individuals. Nearly 91% of IFLS-1 households were interviewed in IFLS-3. Overall, IFLS-3 interviewed 39,000 individuals living in 10,400 households.

### 3.3.1 Variables of interest

A key feature of the IFLS is that the survey collected detailed information about transfers and services, in the form of money, tuition, health care and food, received from and provided to non-coresident parents of adult respondents during the 12 months preceding the survey. Out of the 25,470 adult respondents in the IFLS-3, the module on family transfers was administered only to individuals who did not co-reside with their parents, and for whom the parents were alived during the year before the interview. This module has the advantage of capturing all transfers given by the child to his parents and enables us to compare how transfers depend differently on the education of males and females. Out of the 11,894 individuals who participated in the family transfers module, we exclude individuals who were not married or did not resid with their spouse. Indeed, the analysis in this paper is based on comparing the effect of human capital on the assistance to elderly parents by gender and give some evidence on the labor-market and marital returns to education as potential mechanisms. Restricting the sample to those 8,071 individuals who are married and live with the spouse at the time of survey makes the decision-making unit always comparable. Of these, 5,649 born before 1973 and had information on the place of birth (see the identification strategy that is detailed in the next section). Finally, our analysis sample contains 2,789 males and 2,860 females, born in 794 sub-districts (*kecamatan*).

The dependent variable in the 2SLS equation is a dummy variable equals to one when the respondent provided a monetary transfer to non-coresident elderly parents (in the form of money or food, gifts that are converted into a monetary value) during the last 12 months. As an additional outcome of interest, we focus on the amount of transfer. Education is included in the model using a categorical variable for schooling levels (no education, primary, junior secondary, senior secondary and university). The education system of Indonesia consists of primary school (6 years), junior high school (3 years), senior high school (3 years), and higher education or university (2 or more years). I use years of schooling to test the results in the robustness part. Information on the child's earnings, on his spouse and on key socioeconomic characteristics of non-coresident parents (age, sex, marital status, education and current activity) are also exploited.

Table 3.1 shows summary statistics of main variables, disaggregated by gender. Men complete 7.8 years of education on average, corresponding to junior secondary school, and women achieve the primary level with a mean of 6.4 years of schooling. We observe sex differences in the access to schooling, as only 5% of males are not educated against 9% of females. Gender disparities are larger on the job market, with men being two times as likely to work than women. Indeed, the majority of women in Indonesia are not engaged in waged work. Women generally work in the informal sector as self-employed, casual, or unpaid family workers. The near parity of women's and men's schooling levels despite the low market-participation of married women suggests that women's schooling may be linked in important ways to the marriage market. Indeed, Schaner and Das (2016) shows that married women (with no children) are not less likely to be working compared to their unmarried peers, they are 15 percentage points less likely to be a wage worker, although this is offset by increases in selfemployment and family work. Extensive and intensive margins of women's labor supply are thus important dimensions to consider in our case. This is why our sample includes couples in which women are engaged into the wage labor-market and others are not.[19] Differences in mean of salary (in 10 thousands of rupee), 61 for males and 32 for females, reflect also gender wage gap. In our sample, men are slightly older than women as daughters tend to leave the natal family sooner due to early marriage. Girls are married at a young age, 20 years old on average, against 24 for boys.

Regarding support provided to elderly parents, the probability of giving monetary transfer is similar across men and women. In contrast, the amount of money given is larger for men than for women. As highlighted in the theoretical part, social norms indicate that children are responsible for their elderly parents' welfare in Indonesia (Frankenberg et al., 2002, Geertz, 1961, Wirakartakusumah, 1998). Responsibility for one's aging parents is an im-

[19] Focusing the analysis to couples in which both spouses participate into the wage labor market restricts drastically the sample size to 2,227 individuals. Moreover, Schaner and Das (2016) find evidence that relative to men, women wage workers are positively selected (e.g., more educated) in Indonesia. It is therefore likely that restricting the analysis only on wage workers families would result in bias estimates. Nevertheless, I test the results on this subsample in Section 6.

portant tenet of Islam in particular (Mahmood, 1992). To our knowledge, there is no a clear social norm attributing more importance to sons relative to daughters in taking care of elderly parents. Nevertheless, Table C.2, Panel B (reported in Appendix) shows that the daughter is more frequently reported as the caregiver living with elderly parents among IFLS-2 respondents. 19% of respondents mentionned the youngest daughter, 10% the eldest daughter, 2% the female child in general.[20]

Figure 3.6 represents the evolution of the probability of giving money to old parents according to the level of education and the respondent's gender. Although the probability is fairly steady for males with at least primary education, we observe a U-shaped education pattern for females. Daughters with junior secondary education (14% of women) tend to give money to elderly parents less frequently than daughters with primary (51%) or at least senior secondary (24%).

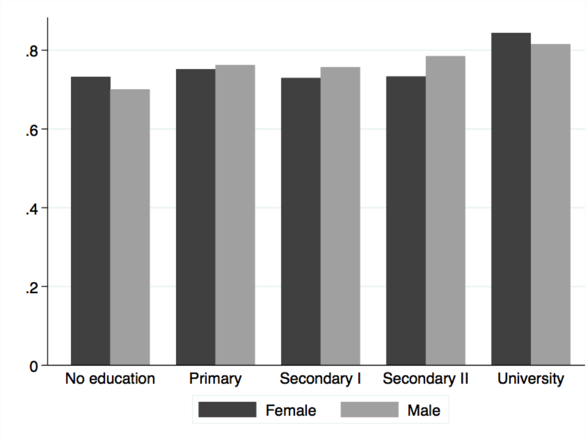



**Figure 3.6** – Probability of monetary transfer

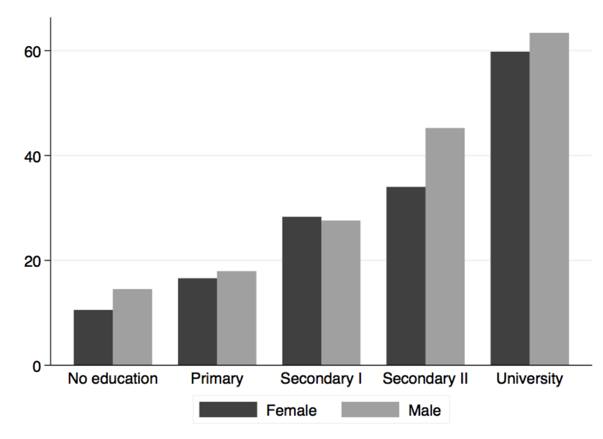



**Figure 3.7** – Amount of monetary transfer

Figure 3.7 clearly demonstrates the repayment hypothesis for investment in human capital, since the most educated give larger amounts. The mean in transfer amounts rises from 14.51 (10.52) for non-educated males (females) to 63.39 (59.80) for the highest educated.

[20] LeBar (1972) finds that social norms differ between sons and daughters across communities and their practices regarding post-marriage residence. For instance in the Madura community, many young couples stay at the wife's parents' house for the initial period of married life. One of the daughters, moreover, remains permanently, with the obligation to care for the parents in their old age. Additional papers stress on the fact that norms on the place of residence after marriage affect care provision by adult children to old parents. See Levine and Kevane (2003), Bau (2019), Jayachandran (2015), Jensen and Miller (2017).

**Table 3.1** – Descriptive statistics

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| Variables | Obs | Mean | SD | Obs | Mean | SD |
| Probability of transfer | 2,789 | 0.75 | 0.42 | 2,860 | 0.74 | 0.43 |
| Amount of transfer* | 2,107 | 30.10 | 54.84 | 2,131 | 21.63 | 40.95 |
| Probability of providing care | 2,789 | 0.10 | 0.30 | 2,860 | 0.10 | 0.30 |
| Frequence of visit | 2,768 | 3.86 | 1.22 | 2,855 | 3.82 | 1.23 |
| Years of education | 2,779 | 7.76 | 4.37 | 2,850 | 6.48 | 4.28 |
| INPRES | 2,123 | 1.76 | 2.97 | 2,181 | 1.82 | 3.06 |
| Father's primary | 2,519 | 0.16 | 0.36 | 2,579 | 0.16 | 0.37 |
| Work status | 2,789 | 0.99 | 0.09 | 2,860 | 0.65 | 0.47 |
| Monthly salary* | 2,714 | 60.86 | 116.40 | 1,462 | 31.82 | 49.36 |
| Spouse's monthly salary* | 1,237 | 34.34 | 57.05 | 2,604 | 65.53 | 342.95 |
| Father's work status | 2,758 | 0.66 | 0.47 | 2,848 | 0.64 | 0.47 |
| Mother's work status | 2,744 | 0.38 | 0.48 | 2,807 | 0.38 | 0.48 |
| Age | 2,768 | 37.08 | 6.85 | 2,855 | 36.22 | 6.65 |
| Spouse's age | 2,764 | 32.69 | 7.20 | 2,848 | 40.81 | 7.99 |
| Age at first marriage | 1,872 | 23.87 | 4.30 | 2,481 | 20.16 | 4.80 |
| Father's age | 1,656 | 65.75 | 10.30 | 1,749 | 64.96 | 9.87 |
| Mother's age | 2,373 | 61.09 | 10.40 | 2,424 | 60.54 | 10.07 |

*per 10 thousands of rupee, which is equivalent to about 0.7$.

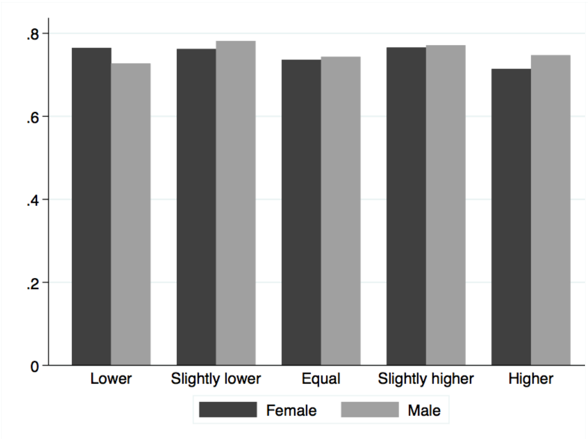



**Figure 3.8** – Probability of monetary transfer and educational gap between spouses

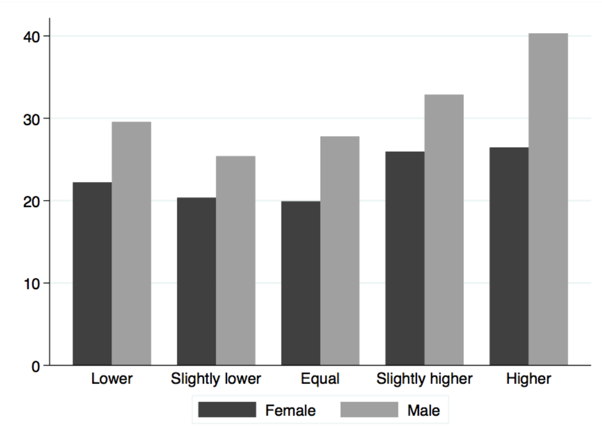



**Figure 3.9** – Amount of monetary transfer and educational gap between spouses

Moreover, educational gap within the household might influence the transfer behaviour of married-children. Whereas differences in education between spouses do not strongly influence the probability of giving to own elderly parents (Figure 3.8), Figure 3.9 shows that larger amounts are sent when the respondent is more educated than his/her spouse. The stylized effect suggests that household bargaining power, reflected here by educational gap between spouses, might affect assistance to elderly parents.

However, the majority of spouses shares close characteristics. Table 3.2 gives Pearson correlation coefficient between husband's and wife's years of education across couples as a measure of sorting. To give insights on the interpretation of our coefficient of 0.69, we refer to Fernandez, Guner, and Knowles (2005). They used household surveys from 34 countries to construct measures of the degree of correlation of spouses' education. On average, across countries the correlation between spouses' years of schooling is around 0.60 with a standard deviation of 0.11. The correlation ranges from 0.322 for Australia to 0.764 for Colombia. By inference, the strong correlation between the education of the wife and that of the husband demonstrates that matching is clearly assortative by education on the marriage market in Indonesia.[21] In addition, we clearly observe assortative matching by age and type of activity. However, the negative correlation associated to the spouses' work status indicates that the decision of participating in the labor market changes at the time of the marriage, with husband's employment acting more often as substitutes to the wife's labor activity.[22]

[21] Instead of using the correlation coefficient between couples' education levels, another strand of the literature measures assortative mating as the proportion of couples who share the same level of schooling (Fernández and Rogerson, 2001). A limitation of these measures is that the conclusions about assortative mating could be confounded by changes in the distribution of men's and women's education. In their paper, Eika, Mogstad, and Zafar (2019) used the observed probability that a husband with education level j is married to a wife with education level i, relative to the probability under random matching with respect to education. Nevertheless, correlation coefficients give some evidence on the martial sorting to support the hypothesis made in our model.

[22] This is in line with Fernandez et al. (2005), in which education measures marital sorting rather than income. They argue that a female's labor force participation decision is often dependent on her spouse's earnings and social norms. In their model, men will still want to marry more educated women as long as skilled women produce higher quality children than their unskilled counterparts.

**Table 3.2 –** Assortative matching on the marriage market

| | Education | Age | Work status | Type of activity |
|---|---|---|---|---|
| Correlation | 0.69* | 0.69* | -0.21* | 0.46* |
| Number of obs | 5,421 | 5,612 | 5,391 | 3,404 |

* p<0.01. Each coefficient represents the Pearson's correlation between spouses. Education measures the number of years of schooling completed. Work status is a categorical variable: housekeeping or no work, unpaid family worker, self-employed, self-employed with employees, private employee or government employee. Type of activity relates the field associated to the main job: agriculture, mining, manufacturing, electricity etc.

### 3.3.2 Primary schools constructed by the INPRES SD reform

In this study, we explore one of the largest schooling reforms to assess the effect of an increase in educational investment on assistance to elderly parents, by the child's gender. Particularly, we construct the variable $INPRES$ which represents the number of school constructed in each sub-districts of birth by the INPRES SD program.

To construct the variable, we first rely on information from the IFLS survey. Indeed, the IFLS collected a school census at the community level, which is two levels more disaggregated than districts (i.e., one level smaller than a sub-district). The household responses were compiled to create a list of schools used by children living in the community with their years of foundation.[23] We assume that schools constructed between 1974 and 1979 result from the INPRES SD program. To test the latter assumption, we link our data with information on schools constructed by the reform in each district (*kabupaten*) of Indonesia from Duflo (2001). Few communities being interviewed in the IFLS within a sub-district, and few sub-districts being interviewed within a district, the main issue is that our data represent only a small share of the schools actually constructed through the program. It is thus difficult to perfectly match our data with those of Duflo (2001). Nevertheless, Table 3.3 represents the correlation between the number of school constructed in 1974-1979 from Duflo (2001) and from the IFLS school census.[24] The coefficients of correlation are given according to the

[23]The list encompasses infrastructures both inside and outside the community, reflecting primary schools available in a sub-district.

[24]Information on schools constructed at the sub-district level (from the IFLS) is aggregated at the district level to conduct the comparison exercise.

**Table 3.3** – Pearson's correlation between data from Duflo (2001) and the IFLS-2

| Percentage of sub-districts interviewed in IFLS-2 | Coefficient | Number of districts (Over 150) |
|---|---|---|
| More than 20% | 0.29** | 53 |
| More than 30% | 0.35* | 32 |
| More than 40% | 0.37* | 23 |
| More than 50% | 0.55** | 17 |
| More than 60% | 0.55* | 12 |

** p<0.05, * p<0.1. Data from Duflo (2001) gives the number of school constructed through the INPRES SD program between 1974 and 1979 per district. In the IFLS 1997, we exploit the number of school constructed between 1974 and 1979 used by members of the communities, aggregated at the district level.

percentage of sub-districts surveyed in a district of the IFLS. The positive and increasing magnitude of the coefficients, around 0.55 significant at the 5% level in districts with more than 50% of the kecamatan interviewed, provide evidence in favour of our hypothesis.

It is important to note that information on schools constructed between 1973-1974 and 1979 from the IFLS is available only for individuals who born in a community interviewed by the survey (57% of the respondents). Indeed, as in Duflo (2001) or Akresh et al. (2018), we use an "individual's region of birth instead of current residence because the latter may be endogenous to the program placement if households move to access schools for their children." For the remaining 47% of respondents, we take the number of primary schools constructed in their district of birth (from Duflo) that we convert into the more disaggregated level of the sub-district. To do so, we multiply the variable with the share of the population aged 7-15 (of primary school age) in 1990. Information on population size at the sub-district level comes from the 1990 Village Potential Statistics (PODES), which provides details about village characteristics for all of Indonesia, not only for villages belonging to the IFLS survey. Unfortunately, information on the population size in each sub-district at the time of the program (1973-1979) is not available. Although population size varies across years, we assume that human population growth rates are constant within districts. Also, we allow the allocation of primary schools within districts to vary according to the size of each sub-district (in km2) in the robustness section.

The way the variable $INPRES$ is constructed has several advantages. Firstly, our measure captures a decrease in the distance to the nearest primary school at the local level (i.e., in the sub-district), which might produce different effects than an aggregate variable.[25] Secondly, it enables to reduce sample selection bias since individuals born in a place that was not part of the IFLS-3 remain in the sample. Focusing only on IFLS respondents who did not migrate from their village of origin might lead to a downward (upward) bias if migration selects children with higher (lower) ability.[26] Figure 3.10 represents the number of schools constructed between 1974 and 1979 in each sub-districts of birth. We see that the intensity of the school construction program varies across space. This is mainly because the number of schools constructed was linked to the regions' primary school enrollment rate in 1972 (prior to the school construction), and areas with low prior enrollment rates had more schools built (Duflo, 2001, Akresh et al., 2018, Ashraf et al., 2018).

Besides geographical variations, the exposure to the program varies across cohorts. Children in Indonesia typically attend primary school between age 7 and 12. Since the program started during the 1973-1974 school year, children who were born before 1961 were 13 years of age or more in 1974 and would not have benefited from schools construction. For these cohorts, $INPRES$ takes the value of 0. In contrast, children younger than 7 in 1974 would have been exposed to the full potential benefits of the reform. Finally, individuals born between 1962 and 1967 (aged between 7 and 12 in 1974) should have been at least partially affected by the reform. The identification strategy, exposed in the next section, relies on both sources of variation.

[25] Exposure to the reform is measured with the number of schools constructed by the INPRES program between 1974 and 1979 per 1,000 children in the individual's birth district in Duflo (2001), Akresh et al. (2018) and Ashraf et al. (2018).

[26] Ashraf et al. (2018) test their empirical results with the IFLS sample and exploit the IFLS school census to measure exposure to the reform. However, their sample is restricted on (i) girls who belongs to a bride-price ethnic group, (ii) who born in an IFLS community.

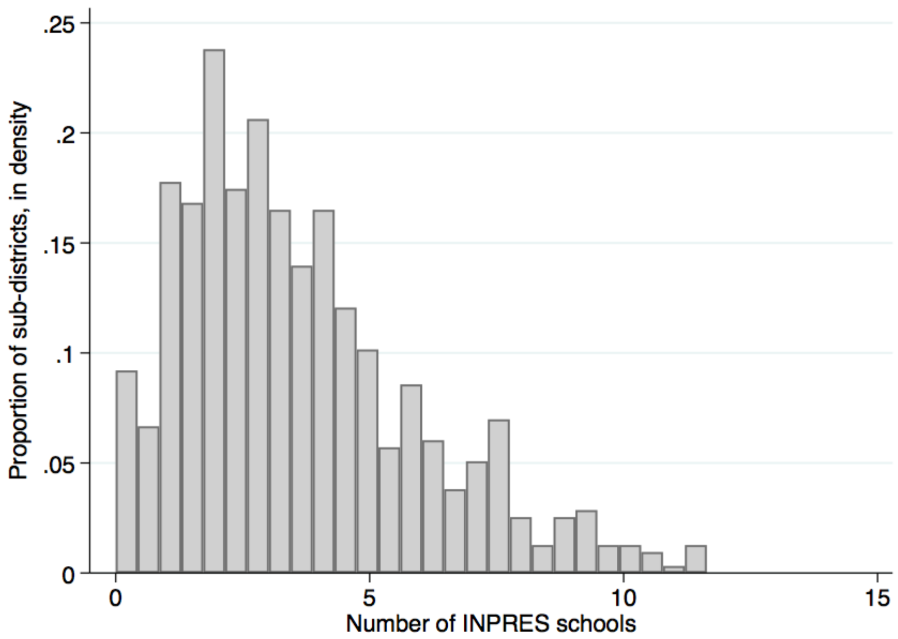



**Figure 3.10** – Distribution of INPRES primary schools constructed between 1973-1979 across sub-districts of birth

## 3.4 Empirical framework

Firstly, we empirically estimate the impact of the Indonesian school construction program on children's human capital, by gender. We exploit variations both across cohorts and across space to identify the effects. In a second step, we investigate the causal impact of human capital on support in old-age. To do so, we use the impact of the reform on children's education as a first-stage of a 2SLS estimation of the effect of education on assitance to elderly.

### 3.4.1 The impact of the educational reform on education

Our empirical work is structured around the massive school construction program during the 70s in Indonesia. As mentioned above, the sub-district of birth and year of birth determined an individual's exposure to the program. Due to overage enrollments, we exclude the cohort aged 12 to 15 in 1974 in the main analysis. In addition, we restrict the analysis on children born before 1973 due to potential endogenous places of birth for subsequent cohorts. The estimation of the impact of the reform on educational attainment, by gender, can be described as

$$\text{Education}_i = \beta_1 \text{INPRES}_{c,k} + \beta_2 \text{ INPRES}_{c,k} \times \text{Female}_i + \beta_3 \text{X}_{c,k} + \gamma_c + \sigma_k + \epsilon_i \tag{3.11}$$

Education is a categorical variable and represents the highest grade completed by individual $i$: no education, primary, junior secondary, senior secondary and university. Years of schooling is also tested in section 6. As mentioned above, $INPRES$ corresponds to the number of schools constructed by the INPRES SD program during individuals' primary school years in the sub-district of birth. The measure reflects a reduction in education costs due to a decrease in the distance to the nearest school at the local level. To investigate heterogeneous effects of the program, we interact $INPRES$ with a dummy variable for being a female. $\beta_2$ captures how the effect of the school program differs by the child's gender.

The identification strategy is valid if there are no omitted time-varying and region-specific effects correlated with the reform. The allocation of school was a function of the enrollment rate in the region prior to the introduction of the program. We follow Duflo (2001) and we control for the enrollment rate in 1971, the allocation of the water and sanitation program, the second largest INPRES program administered during this period, and the population aged 7-14 in 1971. These variables are defined at the district level. To allow variations across cohorts, we interact these controls with year of birth. These variables are added in the vector $X_{c,k}$ of equation 3.11. There might be unobserved village-level heterogeneity. We thus include sub-district fixed effects, denoted by $\sigma_k$, to control for time-invariant specific characteristics and initial differences in the level of economic development. Also, we include year-of-birth fixed effects, $\gamma_c$, to deal with government programs, policies, and other trends that took place during the period. A linear probability model performs the estimation. We report t-statistics that are first clustered at the sub-district level. Later on, we cluster our standard errors both at the cohort and sub-district level.

### 3.4.2 Two-Stage Least-Squares Estimates of the effect of education on old-age transfers

Estimates of equation 3.11 are of intrinsic interest because they provide an assessment of the impact of the program on education. But they also represent the first stage of a Two-Stage Least-Squares (2SLS) estimation of the impact of education on old-age transfers.

Indeed, an important issue concerns the causal interpretation of the effect of education on transfer. Unmeasured individual, household, and community-level resources may affect both education and transfer decisions. Furthermore, education may serve as a proxy for unobservable factors, such as motivation, parental background, which are important determinants of children's transfer choices. Ignoring these factors would lead to biased estimates of the impact of education on assistance within the context of Ordinary Least Square (OLS) estimation. Valid instruments are variables that affect the level of educational attainment but have no direct impact on monetary transfers. If we assume that the INPRES program had no direct effect on transfer, other than through its effect on educational attainment, the number of primary school constructed during the reform in the sub-district of birth can be used as an instrument. Consider the following equation which characterizes the causal effect of education on transfers

$$\begin{aligned}\text{Transfer}_i = {} & \alpha_1 \text{Education}_i + \alpha_2 \text{Female}_i + \alpha_3 \text{Education}_i \times \text{Female}_i \\ & + \alpha_4 \text{X}_{c,k} + \theta_c + \delta_k + \eta_i\end{aligned} \tag{3.12}$$

$$\begin{aligned}\textit{where} \quad \text{Education}_i = {} & \alpha_5 \text{INPRES}_{c,k} + \alpha_6 \ \text{INPRES}_{c,k} \times \text{Female}_i + \alpha_7 \text{X}_{c,k} \\ & + \theta_c + \delta_k + \zeta_i\end{aligned}$$

$$\begin{aligned}\text{Education}_i \times \text{Female}_i = {} & \alpha_8 \text{INPRES}_{c,k} + \alpha_9 \ \text{INPRES}_{c,k} \times \text{Female}_i + \alpha_{10} \text{X}_{c,k} \\ & + \theta_c + \delta_k + \omega_i\end{aligned}$$

$\alpha_1$, - , $\alpha_{10}$, are vectors of parameters to estimate and $\theta_c$, $\delta_k$, $\eta_i$, $\zeta_i$, $\omega_i$ are error terms.

Transfer denotes the probability of giving money and the amount given (conditional on the positive transfer). When analysing the intensive margin of transfers, we implicitly assume that the choice of how much to give is independent of the decision to give. However, this is not necessarily the case and hence respondents who transfer may be a non-random sample of the population. Accordingly, we also test the amount of financial transfers including 0 for those who do not give in Appendix, Table C.3. We interact the level of education with a dummy for being a female. In equation 3.12, $\alpha_3$ captures the heterogeneous impact of the child's education on family transfers by gender. The massive primary school construction program helps to predict the level of education but is not included directly in the equation determining the transfer. In the same way, Education$\times$Female is instrumented with the interaction INPRES$\times$Female. We follow the same strategy than in sub-section 4.1. We include in $\delta_k$ and $\theta_c$ sub-district fixed effects and year of birth fixed effects respectively. $X_{c,k}$ denotes the vector of region characteristics that vary per cohort, exposed previously.

2SLS regression models perform the estimations. One shortcoming of the linear probability model is that it makes extreme assumptions on the distributions of the error terms, which are likely to be violated in the case of a discrete outcome with a mass point at 0. Similarly, estimating a Tobit regression model might be more suitable than a log-linear model of transfer. Indeed, it accounts for the fact that the amount of transfer is a count variable. Also, with the use of the unconditional amount of transfer (included as a robustness check), a large number of observed values equal to zero. However, the main drawback to non-linear regressions is that we have to omit sub-district and cohort fixed effects. The absence of fixed effects might lead to unconvinced estimates. In any case, we believe that the linear model is the best among sub-optimal solutions to study old-age transfers and treat results from Probit and Tobit regressions as indicative. We estimate the probability of giving money with IV-Probit and the amount of transfer with IV-Tobit in Appendix, Table C.3. Finally, standard errors are clustered at the sub-district level and we allow two-way clustering in the robustness section.

## 3.5 Results

In this section, we first investigate the effect of the INPRES SD reform on children's education by gender. Secondly, we examine the causal impact of adult-children's human capital on support to elderly parents, using exposure to the reform as the first-stage of the 2SLS estimation.

**Table 3.4** – The effect of INPRES program on children's education, by gender

| Dependent variable | Years of Schooling | | | Educational Grade | | |
|---|---|---|---|---|---|---|
| *Experiment* | *Individuals Aged 1 to 11 and 15 to 24 in 1974* | | | | | |
| Sample | All | Girls | Boys | All | Girls | Boys |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| INPRES | 0.144*** | 0.205*** | 0.127*** | 0.048*** | 0.070*** | 0.041*** |
| | (0.031) | (0.042) | (0.041) | (0.009) | (0.013) | (0.012) |
| Enrollment in 1971 | -0.275 | -0.403* | -0.443 | -0.046 | -0.063 | -0.109 |
| | (0.212) | (0.241) | (0.311) | (0.066) | (0.078) | (0.095) |
| Water program | 0.103 | 0.111 | 0.066 | 0.006 | 0.009 | -0.006 |
| | (0.083) | (0.090) | (0.132) | (0.023) | (0.028) | (0.036) |
| Children in 1971 | -0.005 | -0.008 | -0.001 | -0.007 | -0.007 | -0.006 |
| | (0.018) | (0.022) | (0.029) | (0.0054) | (0.006) | (0.008) |
| Number of sub-districts | 1,096 | 764 | 755 | 1,096 | 764 | 755 |
| Observations | 4,120 | 2,138 | 1,982 | 4,120 | 2,138 | 1,982 |
| R-squared | 0.110 | 0.176 | 0.126 | 0.095 | 0.154 | 0.114 |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard error in parentheses, clustered at the sub-district level. Educational grade represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. INPRES represents the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15. Population share aged 7-15 in each sub-district is used to disaggregate the variable at the sub-district level. Controls include interactions between year of birth and the allocation of the water and sanitation program in the district between 1974 and 1978. Also, we think about mean reversion by including the enrollment rate in 1971 at the district level interacted with year of birth. We also control for birth year interacted with districts' population aged 7-15 in 1971. All estimates include year of birth and sub-district fixed effects.

### 3.5.1 Reform impact on children's human capital

Table 3.4 shows the estimated impact of the educational reform on years of schooling (from columns 1 to 3) and educational grade (from columns 4 to 6) for adult-married children. We exclude individuals aged 12 to 15 in 1974. The majority of the remaining respondents were of primary school age in 1974 (71% were less than 11 years old in 1974) and 29% were at least 15 years old and did not benefit from the school construction program. INPRES represents the number of school constructed by the reform in the sub-district of birth during primary school age. All OLS estimates include sub-district and year of birth fixed effects.

Columns 1 and 4 report average effects. We find that gaining access to additional primary schools in the community of birth significantly increases educational attainment. The program increases years of schooling by 0.14 years, significant at the 1% size test. In column 4, the magnitude of the effect is more difficult to interpret given the nature of the dependent variable (i.e., a categorical variable). However, qualitatively, the results remain similar.

Secondly, we test the existence of heterogeneous impacts and disaggregate the analysis by the child's gender. Columns 2 and 3 restrict the analysis on females and males respectively. A one unit increase in INPRES results in more than 0.20 years of schooling for women and 0.13 years for men. Our results are close to the findings of Akresh et al. (2018): the program increases average years of schooling for both men and women. Though, we do not find that men benefited more from the reform. On the contrary, larger coefficients are found among the female subsample. In addition, our results contrast with those of Ashraf et al. (2018) since on average, girls are positively affected by the program regardless of their ethnic group. The findings suggest that the Indonesian reform produced large positive effects on boys and girls' human capital, with benefits exceeding the primary level of education as initially targeted. Although the difference between girls and boys is not statistically significant, we observe that women who were exposed to the school construction program largely improved their educational attainment compared to girls who were not affected.[27] The direction of the effect is in line with Prediction 3, in which a decrease in education costs increases more the investment in education of daughters than the one of sons. At that stage, we are not able to shed

[27]We conduct an estimation in which the level of education was interacted with a dummy for being a female. The interaction term was not statistically significant, suggesting that the program does not produce heterogeneous impacts across girls and boys.

light on the competing indirect effects of labor-market returns and marital returns to education though.

Table 3.5 reports alternative specifications and tests. As pointed out in Duflo (2001), the effect of the program should be close to 0 for children 12 or older in 1974 and increasing for younger children, since exposure might be an increasing function of the date of birth. We test the hypothesis in Panel A. Exposure is the interaction between the number of school constructed in the sub-district of birth and a variable reflecting the number of years the child is supposed to benefit from the program. The variable equals to 0 for individuals aged 12 to 24 in 1974, equals to 1 for individuals aged 11 in 1974 and so on. These estimates are very similar, although the magnitude of coefficients associated to the program variable decreases. This might be due to the fact that the child cohort born in 1959-1962, which is now included, might have been partially affected by the program and assigning them a value of 0 might poorly predict their treatment to the reform.

Panel B reproduces the empirical strategy common to the literature on the impact of the INPRES reform (Duflo, 2001, Akresh et al., 2018, Ashraf et al., 2018). In their baseline estimates, the authors compared the cohort born between 1968-1972 (ages 2-6 in 1974) with individuals born after 1963 (ages at least 12 in 1974) who were not exposed to the program. Partial treatment cohorts born between 1963 and 1967 were not included in their sample. The magnitude of INPRES reduces and we are not able to identify a positive effect on boys' years of schooling or educational grades (column 3 and 6).

Placebo estimates are obtained by shifting the reform across time in Panel D and E. In Panel D, we test the results on the cohort aged 12 to 17 and 17 to 24 in 1974. We expect that the reform impact should be very close to 0 since those children should have left primary school before the first INPRES schools were opened. Surprisingly, we still find a significant and positive effect of the school construction program among the treated cohort (aged 12 to 17 in 1974). The effect is driven by men since INPRES is not statistically significant in the subsample of women in column 3 and 5. We reproduce the exercise in Panel D, and exclude individuals aged 12 to 14 in 1974. Indeed, grade repetition and delayed school entry could lead a few of these children to benefit from the program during their last year in schooling. In the IFLS-3, 42% of individuals completed their primary education above the age of 12. Among them, 35% ended their primary school between 13 and 15 years old. Overall, the results are not statistically significant, suggesting that the results are not driven by misspecification in the construction of our school program variable.

**Table 3.5 –** Robust effect of the INPRES program on children's education, by gender

| Dependent variable | Years of Schooling | | | Educational Grade | | |
|---|---|---|---|---|---|---|
| Sample | All | Girls | Boys | All | Girls | Boys |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Panel A: Individuals Aged 1 to 24 in 1974* | | | | | | |
| Exposure | 0.016*** | 0.018*** | 0.011** | 0.003*** | 0.006*** | 0.003** |
| | (0.004) | (0.005) | (0.005) | (0.001) | (0.001) | (0.001) |
| Observations | 3,660 | 1,887 | 1,773 | 3,660 | 1,887 | 1,773 |
| *Panel B: Individuals Aged 2 to 6 and 12 to 24 in 1974* | | | | | | |
| INPRES | 0.083** | 0.154*** | 0.069 | 0.032*** | 0.055*** | 0.023 |
| | (0.037) | (0.057) | (0.055) | (0.011) | (0.017) | (0.017) |
| Observations | 4,957 | 2,566 | 2,391 | 4,957 | 2,566 | 2,391 |
| *Panel C: Individuals Aged 12 to 17 and 17 to 24 in 1974* | | | | | | |
| INPRES | 0.210*** | 0.167 | 0.233** | 0.055** | 0.039 | 0.058* |
| | (0.070) | (0.105) | (0.111) | (0.021) | (0.032) | (0.035) |
| Observations | 2,044 | 992 | 1,052 | 2,044 | 992 | 1,052 |
| *Panel D: Individuals Aged 15 to 17 and 17 to 24 in 1974* | | | | | | |
| INPRES | 0.177* | 0.027 | 0.269 | 0.044 | -0.004 | 0.066 |
| | (0.101) | (0.137) | (0.216) | (0.030) | (0.041) | (0.069) |
| Observations | 1,394 | 658 | 736 | 1,394 | 658 | 736 |
| *Panel E: Individuals Aged 1 to 11 and 15 to 24 in 1974* | | | | | | |
| INPRES, area | 0.103*** | 0.179*** | 0.092* | 0.038*** | 0.060*** | 0.032** |
| | (0.030) | (0.038) | (0.047) | (0.009) | (0.012) | (0.015) |
| Observations | 4,131 | 2,147 | 1,984 | 4,131 | 2,147 | 1,984 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. Educational grade represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. INPRES represents the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15. Controls include the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971. All estimates include year of birth and sub-district fixed effects. In Panel A, Exposure is the interaction between a variable reflecting the number of years each individual were supposed to be affected by the 1974-1979 period and the number of school constructed at the sub-district level. In Panel E, INPRES, area represents the number of school constructed at the sub-district level, where the size of each sub-district is used to convert the variable from the district to the sub-district level.

Finally, Panel E exploits a different indicator to convert the number of schools constructed in a district into the sub-district level: the size of each sub-district, in km2. In this setting, we assume geographic uniformity in the allocation of primary schools within districts. We find similar results except that coefficients are weaker. In addition, the level of significance is lower in the subsample of males (column 3 and 6). The allocation of school was a function of the enrollment rate in the region prior to the introduction of the program. Thereby, we are not surprised that disaggregating the reform at the local level using the share of the population aged 7-15 in the sub-district better predicts exposure to the reform than the geographic size.

### 3.5.2 2SLS findings: the effect of human capital on old-age support

In this section, we investigate the causal effect of human capital on the assistance given to elderly parents. Specifically, we allow different treatment by gender. We test alternative dependent variables: the likelihood of giving and the amount given. In all specifications, we control for regional characteristics that could be correlated with the INPRES SD program. These include the allocation of the water and sanitation program implemented during the period, the primary enrollment rate in 1971 and the children's population in each district in 1971. All controls are interacted with year of birth. All estimates include sub-districts and year of birth fixed effects. We exclude individuals aged between 12 to 15 in 1974 for whom the INPRES indicator might poorly predict their exposure to the reform. We first report OLS estimates where we ignore the concern about the endogeneity of the schooling decision. Then, we report 2SLS estimates in which the effect of the school construction program on education is used as the first-stage.

Table 3.6 explores the effect of human capital on the likelihood of giving monetary transfer (columns 1 to 4). Before displaying heterogeneous effects across married men and women, we begin by estimating average effects. According to prediction 2, we expect a positive sign on Education. Estimates suggest that the level of education strongly affects the probability of giving money to elderly parents during the last 12 months, given the significance of coefficients at the 1% level. Following equation 3.12, column 2 includes the interaction of the level of education with the female dummy, which is not statistically significant.

We pursue our analysis investigating how an exogenous change in education affects the likelihood of giving. Column 3 reports the 2SLS estimate of the average effect. The corresponding first-stage is

**Table 3.6** – Two-Stage Least-Squares Estimates of the effect of education on old-age transfers

| Dependent variable | Monetary transfer (0/1) | | | | Amount of transfer | | | |
|---|---|---|---|---|---|---|---|---|
| Method | OLS | | 2SLS | | OLS | | 2SLS | |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Education | 0.024*** | 0.024*** | 0.157* | 0.180* | 0.345*** | 0.330*** | 0.620*** | 0.666*** |
| | (0.006) | (0.007) | (0.088) | (0.095) | (0.022) | (0.027) | (0.201) | (0.213) |
| Education × Female | | 0.001 | | -0.033 | | 0.031 | | -0.068 |
| | | (0.009) | | (0.047) | | (0.035) | | (0.126) |
| Female | 0.013 | 0.013 | 0.063 | 0.113 | -0.204*** | -0.248*** | -0.043 | 0.057 |
| | (0.012) | (0.018) | (0.038) | (0.080) | (0.043) | (0.064) | (0.099) | (0.203) |
| Water Program | 0.003 | 0.003 | -0.003 | -0.003 | 0.028 | 0.028 | 0.017 | 0.016 |
| | (0.007) | (0.007) | (0.007) | (0.006) | (0.025) | (0.025) | (0.026) | (0.026) |
| Enrollment in 1971 | 0.006 | 0.006 | 0.026 | 0.029 | 0.062 | 0.060 | 0.065 | 0.068 |
| | (0.018) | (0.018) | (0.020) | (0.020) | (0.070) | (0.070) | (0.078) | (0.078) |
| Children in 1971 | 0.002 | 0.002 | 0.003* | 0.003* | 0.010* | 0.010* | 0.009 | 0.009 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.005) | (0.005) | (0.006) | (0.006) |
| F-stat.(KP) | – | – | 23.52 | 11.73 | – | – | 30.16 | 15.05 |
| Observations | 5,370 | 5,370 | 3,566 | 3,566 | 4,004 | 4,004 | 2,604 | 2,604 |
| Sub-districts | 1,359 | 1,359 | 542 | 542 | 1,147 | 1,147 | 434 | 434 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. 2SLS estimates use the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15, as an instrument. Controls include the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971. All estimates include sub-districts and year of birth fixed effects. We remove individuals aged between 12 to 15 in 1974 who might be partially affected by the number of primary schools constructed by the program.

reported in Table 3.4, column 4. It should be noted that the IV strategy relies critically on the strength and validity of the instruments used. As pointed out in Sanderson and Windmeijer (2016), for multiple endogenous variables (Education and Education$\times$Female in our case), inspection of the individual first-stage F-statistics is no longer sufficient. The Cragg and Donald statistic can be used to evaluate the overall strength of the instruments in this case. However, standard Stock-Yogo critical values for weak instruments are calibrated for the case of i.i.d. residuals, and do not apply to the case of clustered standard errors (Bun, De Haan, et al., 2010). Accordingly, we report robust Kleibergen-Paap F-statistics associated to the first-stage regressions of the 2SLS specification. As shown in the previous subsection, exposure to the school construction program has large predictive power of the level of education. The F-statistics (23.5) is above the conventional threshold for weak instruments. It should be noted that avoiding the endogeneity issue underestimatse the positive effect of schooling on old-age support, as IV estimates give larger coefficients. A one unit increase in the level of education increases the likelihood of giving financial support to elderly parents by 15%, although significant at the 10% size test. In addition, we do not find differences between males and females since Education$\times$Female in column 4 is not significant.

We now explore the amount of money given as the dependent variable from columns 5 to 8. The analysis is restricted to individuals who sent money, food or gifts to their parents during the last 12 months and gives intuition about the intensive margin effect of education. We test our findings using OLS based on log transfer in columns 5 and 6. We find similar results. The higher the educational attainment of the child, the higher the transfer amount to parents, at the 1% level. We interpret the results from the preferred specification in column 7, where we think about the endogeneity of human capital. Each additional level of education increases the amount of monetary transfer sent to elderly parents by 62 percentage points on average ($p < .001$). The magnitude of the coefficient slightly increases with the inclusion of the interaction between the level of education and the sex of the respondent. The latter is statistically insignificant, which suggest that responses to the human capital investment do not differ by gender. The larger coefficient of education in the 2SLS specifications with respect to their OLS counterparts suggests that OLS may suffer from an attenuation bias due to omitted variables. We also notice that the values of the Kleibergen-Paap F-statistics are significantly higher than with the previous outcome, in part due to the selection of a different sample.

Overall, the findings are consistent across estimations. Interpreting the coefficient associated to the female dummy gives an empirical estimate of Prediction 1. The coefficient is negative and significant only in columns 5 and 6. Though, these estimates do not represent our preferred specifications since endogeneity is not controlled for. Contrary to Prediction 1, at similar levels of education, the transfer of a son to his old parents is not higher than the transfer of a daughter in general. Moreover, we find that additional level of education increases the assistance provided to parents in the form of money, food or gifts. Human capital increases both the extensive and intensive margin of family transfers. However, the strength of the impact does not differ across gender.

According to Prediction 2, the absence of stronger marginal effects of education on transfers for females suggests that marriage-market returns do not play a significant role. However, the foregoing interpretation might be invalidated if we suppose the existence of compensation effects that are not fully captured by the model. Indeed, predictions 2.2 and 2.3 call for non-linear effects. These predictions suggest that if marriage market returns matter, we should observe a significant higher impact of education on the transfers of daughters than on the transfers of sons among low and middle levels of education. The difference across gender should decrease for higher education levels. If marital returns to education vary in opposite directions among low and high educated children, our effects could be attenuated or even eliminated.

We provide some evidence on the presence of non-linear effects that could mask some heterogeneity across educational grades. To do so, we test the effect of education on family transfers by gender across two subsamples: (1) among individuals who did not complete secondary education; (2) among individuals with at least senior secondary education. In our sample, 64% of males and 75% of females have less than junior secondary education. Table 3.7 reports the findings. Estimates must be interpreted with caution given that the endogeneity of human capital is not considered here. Indeed, exposure to the school construction program predicts poorly the educational attainment for the two subsamples. This is not surprising since the effect of the program on education is less precisely estimated because the sample is smaller. This prevents the use of 2SLS estimates. Similarly, we are not able to conduct reduced-form estimates in which exposure to the reform is used as a proxy for the level of education. Although the results must be interpreted with caution, they are potentially informative.

We observe a higher impact of education on the transfers of daughters than on the transfers of sons among individuals with less than the senior secondary level. The result is identified only on the amount of transfer in column 2. Completing lower secondary is associated to a 24% and a 40%

increase in the probability of giving monetary assistance for males and females respectively. When the analysis is restricted to individuals with at least some senior secondary education, the differential effect across gender disappears. This is in line with Prediction 2.3. Finally, we conduct additional estimates on non-linear effects of education. We add the square term of education to the specifications in columns 5 and 7. In both columns, the coefficient on the linear term of school attainment is positive and statistically significant, but the square term is not significant. However, when we include the interaction of the femaly dummy with both education and its square, some findings emerge in column 8. Although results are difficult to interpret, this configuration of coefficient suggests that the marginal effects of education on amount of transfers are stronger for females than for males (Education and Education$\times$Female are positive and significant, p$<$.10). In addition, coefficients on the square term and its interaction with Female are both significant. This seems to indicate that the relationship between education and the amount of transfers is not monotonic and that the effect varies in opposite directions for males and females. Education leads to higher amount of transfer for the less-educated females, whereas better educated women provide less support for elderly parents. In contrast, educated males tend to give larger transfers compared to low-educated males.

Thereby, the results of Table 3.7 suggest the existence of heterogeneous effects of human capital on the amount of transfer given to elderly parents across gender among the low - middle educated individuals. Although the results cannot be viewed as causal, the direction of effects supports the existence of marital market returns.

Before giving empirical evidence on the mechanisms, we test the results on the relationship between education and transfer in the following section.

## 3.6 Robustness checks on the effect of human capital on transfers

In this section, we perform some robustness checks on our dependent variables: the probability of giving money and how much is given. For each robustness exercise, we report the IV estimates that control for the endogeneity of children's human capital and explore heterogeneous effects across gender.

**Table 3.7 –** Non-linear impacts of education on old-age transfers, by gender

| Individuals with | No more than junior secondary | | At least senior secondary | | All educational levels | | | |
|---|---|---|---|---|---|---|---|---|
| Dependent variable | Give | Amount | Give | Amount | Give | | Amount | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Education | 0.035** | 0.244*** | -0.011 | 0.338** | 0.037* | 0.0531** | 0.255*** | 0.164* |
| | (0.017) | (0.058) | (0.049) | (0.153) | (0.019) | (0.025) | (0.067) | (0.084) |
| Education×Female | -0.005 | 0.168** | 0.063 | 0.122 | -0.001 | -0.030 | 0.032 | 0.210* |
| | (0.022) | (0.075) | (0.058) | (0.214) | (0.009) | (0.031) | (0.035) | (0.110) |
| Education$^2$ | | | | | -0.003 | -0.008 | 0.021 | 0.046** |
| | | | | | (0.004) | (0.006) | (0.016) | (0.022) |
| Education$^2$ ×Female | | | | | | 0.008 | | -0.0501* |
| | | | | | | (0.008) | | (0.030) |
| Female | 0.010 | -0.304*** | -0.185 | -0.702 | 0.012 | 0.023 | -0.252*** | -0.318*** |
| | (0.021) | (0.072) | (0.197) | (0.726) | (0.018) | (0.021) | (0.063) | (0.071) |
| Observations | 3,880 | 2,860 | 1,513 | 1,162 | 5,393 | 5,393 | 4,022 | 4,022 |
| Sub-districts | 1,044 | 869 | 747 | 626 | 1,361 | 1,361 | 1,150 | 1,150 |
| R-squared | 0.009 | 0.058 | 0.027 | 0.080 | 0.009 | 0.009 | 0.122 | 0.123 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. Controls include the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971. All estimates include sub-districts and year of birth fixed effects.

**Table 3.8 –** Reduced-form estimates: effects of the program on the assistance to elderly parents

| Dependent variable | Monetary transfer (0/1) | | Amount of transfer | |
|---|---|---|---|---|
| Variables | (1) | (2) | (3) | (4) |
| INPRES | 0.007* | 0.009** | 0.036*** | 0.043*** |
| | (0.003) | (0.004) | (0.012) | (0.014) |
| INPRES × Female | | -0.004 | | -0.013 |
| | | (0.005) | | (0.019) |
| Female | 0.002 | 0.010 | -0.305*** | -0.279*** |
| | (0.014) | (0.017) | (0.051) | (0.061) |
| Enrollment in 1971 | 0.019 | 0.019 | 0.022 | 0.023 |
| | (0.019) | (0.019) | (0.077) | (0.077) |
| Water program | -0.002 | -0.002 | 0.033 | 0.032 |
| | (0.007) | (0.007) | (0.026) | (0.026) |
| Children in 1971 | 0.002 | 0.002 | 0.005 | 0.004 |
| | (0.001) | (0.001) | (0.006) | (0.006) |
| Observations | 4,129 | 4,129 | 3,100 | 3,100 |
| Sub-districts | 1,096 | 1,096 | 923 | 923 |
| R-squared | 0.007 | 0.007 | 0.032 | 0.032 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. INPRES represents the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15. Population share aged 7-15 in each sub-district is used to disaggregate the variable at the sub-district level. Controls include interactions between year of birth and the allocation of the water and sanitation program in the district between 1974 and 1978. Also, we think about mean reversion by including the enrollment rate in 1971 at the district level interacted with year of birth. We also control for birth year interacted with the population aged 7-15 in 1971 in each district. All estimates include year of birth and sub-district fixed effects.

**Table 3.9 –** Robustness effects: alternative instruments, variables of interests and outcomes

| Instrument | INPRES, area | | Exposure | | INPRES | | INPRES | |
|---|---|---|---|---|---|---|---|---|
| Dependent variable | Give | Amount | Give | Amount | Give | Amount | Give Net | Net Amount |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Education | 0.224* | 0.948*** | 0.260 | 0.729** | | | 0.213** | 0.602*** |
| | (0.124) | (0.289) | (0.169) | (0.302) | | | (0.103) | (0.207) |
| Education × Female | -0.082 | -0.172 | -0.054 | -0.117 | | | -0.055 | 0.104 |
| | (0.072) | (0.182) | (0.054) | (0.146) | | | (0.052) | (0.151) |
| Schooling | | | | | 0.061* | 0.239*** | | |
| | | | | | (0.033) | (0.078) | | |
| Schooling × Female | | | | | -0.010 | -0.027 | | |
| | | | | | (0.014) | (0.040) | | |
| Female | 0.184 | 0.295 | 0.173 | 0.114 | 0.153 | 0.224 | 0.129 | -0.227 |
| | (0.116) | (0.299) | (0.120) | (0.271) | (0.117) | (0.315) | (0.089) | (0.217) |
| F-stat.(KP) | 6.94 | 11.73 | 3.66 | 7.45 | 9.62 | 11.47 | 11.77 | 17.44 |
| Observations | 3,579 | 2,608 | 4,401 | 3,187 | 3,566 | 2,604 | 3,559 | 2,106 |
| Sub-districts | 536 | 430 | 641 | 501 | 542 | 434 | 541 | 392 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. Schooling is the average years of schooling. INPRES represents the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15. INPRES represents the number of school constructed during the reform in the sub-district of birth, ranging from 0 to 15. Population share aged 7-15 in each sub-district is used to disaggregate the variable at the sub-district level. INPRES, area represents the number of school constructed during the reform in the sub-district of birth, using total area in km2 of each sub-district for the variable construction. Exposure is the interaction between a variable reflecting the number of years each individual were supposed to be affected by the 1974-1979 period and the number of school constructed at the sub-district level. All estimate controls for the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971. Sub-districts and year of birth fixed effects are included in all estimates.

Firstly, we study transfers outcomes using OLS in order to capture broad impacts and because the exclusion restriction could be violated if the program caused community-level changes that affect long-term outcomes in ways other than through increased schooling (Akresh et al., 2018). We follow the procedure suggested by Angrist and Pischke (2008). We report the coefficients for excluded instruments in the reduced-form regression of dependent variables on instruments in Table 3.8. The reduced form is proportional to the causal effect of interest and the reduced-form estimates, since they are OLS, are unbiased. Variations of the effect of the program on family transfers parallel those on education. This suggests that the program affects transfers through changes in years of education and comforts our strategy using exposure to the school construction program as an instrumental variable estimate of the effect of education on old-age assistance.

Secondly, we construct exposure to the school construction program in a different way and explore how it affects the second-stage results. In baseline estimates, INPRES represents the number of schools constructed during the reform in the sub-district of birth. The first source relevant for the construction of the variable is the IFLS school census. To complete missing information, we used so far the share of the population aged 7-15 years old in the sub-district of birth to disaggregate information at the local level. In this part instead, we weight the number of primary schools constructed by the size of each sub-districts, in km2, available in the 1991 PODES survey. Table 3.9, columns 1 and 2 reports the 2SLS estimates associated to the first-stage in Table 3.5, Panel E, column 4. For both married men and women, the level of education positively affects financial assistance. However, weighing the number of primary schools constructed by the geographic size instead of the population share aged 7-15 in the sub-district decreases the impact of the educational reform. The lower value of the Kleibergen-Paap F-statistics reported (6.94 and 11.73) supports the argument.

As additional diagnostics, columns 3 and 4 report the 2SLS estimates when Exposure is used to estimate the effect of the school construction program on education in the first-stage (Table 3.5, Panel A, column 4). Exposure is the interaction between the number of school constructed in the sub-district of birth and a variable which reflects the number of years the child is supposed to benefit from the program. We do not find a positive effect of the level of education on the likelihood of giving in column 3. This might be due to the strength of the instrument in the first-stage which decreases drastically (the Kleibergen-Paap F-statistics is 3.6). However, human capital positively affects the amount of transfers. The magnitude of the coefficients is close to IV estimates presented in the main analysis.

In columns 5 and 6, we test the 2SLS estimates using a continuous explanatory variable: years of schooling. The first-stage is reported in Table 3.4, column 1. An additional year of education increases the likelihood of giving by 6%, although significant at the 10% level. Regarding the intensive margin, an additional year of education increases the amount given by 23%, at least at the 1% size test.

So far, financial transfers are measured from adult-married children to parents. But we now take into account financial assistance that the parents can also provide to their adult-married children. Basically, the net flow of financial gifts is the amount given from children to parents minus the amount from parents to children. We investigate the differential effect of education on the probability of net giving to parents, with children who are net recipients or non-givers and therefore, classified as zero. We start by analyzing whether there is giving (coded 1 and 0) in column 7. Then, we analyze how

**Table 3.10 –** Two-Stage Least-Squares Estimates: alternative specifications

| Sample | | | | | Wage-earners couples | | | |
|---|---|---|---|---|---|---|---|---|
| Dependent variable | Give (0/1) | | Amount | | Give (0/1) | | Amount | |
| Method | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Education | 0.024*** | 0.122 | 0.320*** | 0.676*** | 0.021* | 0.188 | 0.354*** | 0.707** |
| | (0.008) | (0.084) | (0.029) | (0.194) | (0.011) | (0.149) | (0.043) | (0.284) |
| Education × Female | 0.001 | -0.019 | 0.041 | -0.158 | -0.009 | -0.065 | -0.022 | -0.179 |
| | (0.010) | (0.046) | (0.037) | (0.128) | (0.013) | (0.063) | (0.053) | (0.175) |
| Female | 0.011 | 0.071 | -0.266*** | 0.157 | 0.043 | 0.179 | -0.082 | 0.367 |
| | (0.020) | (0.070) | (0.070) | (0.199) | (0.030) | (0.126) | (0.111) | (0.325) |
| Year of Birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Sub-District FE | No | No | No | No | Yes | Yes | Yes | Yes |
| Sub-District×Cohort FE | Yes | Yes | Yes | Yes | No | No | No | No |
| F-stat.(KP) | – | 14.92 | – | 20.24 | – | 3.52 | – | 5.18 |
| Observations | 5,370 | 3,310 | 4,004 | 2,399 | 2,178 | 1,244 | 1,684 | 939 |
| Sub-districts | 1,359 | 1,359 | 542 | 542 | 825 | 274 | 699 | 221 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, two-way clustered at the sub-district and cohort level from columns 1 to 4. From columns 5 to 8, standard errors are clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. 2SLS estimates use the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15, as an instrument. Controls include the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971. All estimates include year of birth fixed effects. Columns 1 to 4 include sub-district and cohort fixed effects. Cohort is a dummy variable equals to 1 for the cohort 1962-1973. Sub-district fixed effects are included from columns 5 to 8.

much is given among those who are net donors in column 8. The results on net financial support are consistent with the previous analysis. Human capital increases both the probability of net giving and the amount of net transfers to the same extent.

Moreover, we allow standard errors to be correlated across cohorts and space in Table 3.10. We cluster the standard errors on the interaction between sub-districts of birth and a cohort dummy equals to 1 for the cohort 1962-1973 who was supposed to be affected by the school construction program. In addition, we include sub-districts×cohort fixed effects. The model including sub-districts by cohort fixed effects accounts for any common variations in the school participation rates in the sub-districts across cohorts. We therefore fully absorb any local shocks on the participation rates and identify the model using variations within sub-districts across cohorts. Results are reported from columns 1 to 4. Including local fixed effects as well as controlling for an additional degree of serial correlation in the standard errors improve estimator precision. We observe an increase in the power of the instrumentation given the higher values of the Kleibergen-Paap F-statistics. The magnitude of the coefficient is similar to previous estimates. However, we are not able to find a statistically significant effect of education on the likelihood of giving in the 2SLS estimates in column 2. Column 4 shows that human capital increases transfer amounts and the effect does not differ by the child's gender.

In the last four columns of Table 3.10, we restrict the analysis on couples in which both spouse work. The strategy reduces drastically the sample size since a large proportion of married women do not work. More than 59% of observations are excluded. The estimates should be interpreted with caution given that the school construction instrument is fairly weak (the first-stage F-statistic is 3.04 and 5.18 for the two outcomes), which is not surprising given the small sample size. I directly interpret 2SLS estimates. Education does not have a significant impact on the likelihood of giving among wage-earners couples. In contrast, education increases transfers amount and the magnitude is larger than with previous estimates. However, our results might suffer from selection bias since women who participated into the wage labor market are positively selected.

Finally, we investigate heterogeneous effects according to religion and the practice of the bride price norm. We first restrict the analysis on muslims, since it has been documented that social norms on old-age assistance display more depth in the Islamic religion (Mahmood, 1992). We exclude 20% of our main sample. Table 3.11, from columns 1 to 4, reports the estimates. Consistent with the main analysis, human capital increases old-age assistance and the strength of the effect does not differ by the child's gender. We find weaker effects compared to coefficients in Table 3.6 though, due to a smaller sample size. Also, according to the paper of Ashraf et al. (2018), the practice of bride price allows parents to partake in the returns that their daughters accrue from education and, hence, it helps to complete the intergenerational contract between parents and daughters. Following these statements, we can reasonably suppose that the relationship between education and family transfers differs across ethnic groups. We follow Ashraf et al. (2018) and classify different ethnic groups according to their traditional marriage customs.[28] To empirically test the hypothesis, we interact the level of education with a dummy equals to 1 among ethnic groups that traditionally practice bride price and 0 otherwise. Note that the exercise contrasts with the one exposed in subsection 7.2 since individuals from bride price groups substantially differ from individuals in matrilineal groups, although the two are largely

[28] We are not able to classify 10% of the sample because we do not find the practice of the bride price norm in their ethnicity.

**Table 3.11** – Heterogeneous effects: religion and bride price

| Subsample | Muslim | | | | Girls | | | |
|---|---|---|---|---|---|---|---|---|
| Dependent variable | Give (0/1) | | Amount | | Give (0/1) | | Amount | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Education | 0.021** | 0.199 | 0.331*** | 0.530** | 0.017* | 0.064 | 0.433*** | 0.569*** |
| | (0.008) | (0.123) | (0.028) | (0.229) | (0.010) | (0.069) | (0.037) | (0.203) |
| Education×Female | -0.003 | -0.005 | 0.077** | -0.095 | | | | |
| | (0.011) | (0.052) | (0.035) | (0.133) | | | | |
| Education×Bride price | | | | | -0.014 | 0.088 | -0.235* | -0.701* |
| | | | | | (0.035) | (0.127) | (0.121) | (0.362) |
| Female | 0.021 | 0.088 | -0.318*** | -0.001 | | | | |
| | (0.020) | (0.090) | (0.069) | (0.220) | | | | |
| Bride price | | | | | 0.023 | -0.101 | 0.461 | 1.360 |
| | | | | | (0.098) | (0.278) | (0.340) | (0.902) |
| Observations | 5,719 | 4,055 | 4,378 | 3,069 | 2,969 | 1,980 | 2,189 | 1,424 |
| Sub-districts | 1,254 | 578 | 1,068 | 469 | 875 | 362 | 738 | 290 |
| F-Statistics | – | 6.73 | – | 11.60 | – | 12.17 | – | 10.02 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. 2SLS estimates use the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15, as an instrument. Controls include the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971. All estimates include sub-districts and year of birth fixed effects. We remove individuals aged between 12 to 15 in 1974 who might be partially affected by the number of primary schools constructed by the program.

correlated (Ashraf et al., 2018). Results have to be interpreted with caution because only 12% of the sample belongs to an ethnic group that practice bride price. Findings are reported in Table 3.11, from columns 5 to 8. We report only the estimates on women, since we do not find significant effects for men. An interesting picture emerges from the subsample of women when investigating the effect on the amount of transfers in columns 7 and 8. Interpreting directly the 2SLS estimate, we see that the level of education decreases the amount of transfer given to elderly parents by 13% (0.5 - 0.7 in column 6) among bride-price ethnic groups. The finding supports our expectations.

## 3.7 Empirical evidence on the mechanisms

The theoretical part sheds light on two mechanisms explaining the relationship between educational attainment and assistance given to elderly parents. The first one is the impact of education on expected wages. The second is the impact of education on the bargaining power within the newly formed household. Ideally, one would like to explore the causal effect of education on wages (1) and the causal effect of wages on transfer (2). A similar investigation should be undertaken on household bargaining power. Due to endogeneity issues and the difficulty to find good measures of intra-household bargaining, though, we only provide additional evidence on the plausibility of the mechanisms. In this section, we first investigate relation (1) between education and labor-market opportunities in Indonesia for males and females. Secondly, we explore whether education effectively affects old-age transfer through additional bargaining power in the newly formed household.

### 3.7.1 Labor-market returns to education

Here, we explore how the schooling decision made by the parents impacts the labor-market outcomes of adult-children and whether the relationship is different by gender. We focus on the effect of education on labor-market returns, using the log of wages as the outcome. We take into account two persistent econometric problems when one wants to assess the causal impact of education on earnings (Card, 2001). Indeed, estimating earnings functions using selected populations (i.e., income-earners are only observed when they work) raises concerns over possible sample selection biases. Over half of the households in the sample are agricultural households. Therefore, many workers are self-employed farmers who choose not to enter the wage labor market. A selection bias might be more pronounced among women, where 34% of married women in the sample did not work during the last week. We use Heckman's two-step procedure to address the selection issue, where a probit model estimates the probability to participate to the wage labor-market. In the selection equation, we use the number of children living in the household as the additional variable that is not included in the wage equation.[29] Also, we address the issue of the possible endogeneity of education in the earnings function using an instrumental variable.[30] The number of school constructed by the INPRES SD reform disaggregated

[29] Among 300 women who stopped working, marriage is the main reason for non labor-market participation (40% of women), followed by the child birth (12% of women).

[30] Education might be positively correlated with the earnings residual due to unobserved ability.

**Table 3.12 –** The effect of education on the log of wages, by gender

| Dependent variable | Log of wage | | | | | Log of spouse's wage | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Method | 2SLS | Heckman selection | | IV + Heckman Probit | | 2SLS | Heckman selection | | IV + Heckman Probit | |
| Sample | All | All | Women | All | Women | All | All | Men | All | Men |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Education | 0.435* | 0.305*** | 0.391*** | 0.076 | 0.417 | 0.651*** | 0.312*** | 0.351*** | 0.397** | 0.837* |
| | (0.247) | (0.016) | (0.027) | (0.189) | (0.376) | (0.248) | (0.027) | (0.033) | (0.195) | (0.439) |
| Education × Female | 0.245* | 0.066** | | 0.291** | | -0.209 | -0.018 | | -0.144 | |
| | (0.127) | (0.028) | | (0.123) | | (0.143) | (0.031) | | (0.189) | |
| Female | -0.967*** | -0.758*** | | -1.359** | | 1.352*** | 0.878*** | | 1.538 | |
| | (0.225) | (0.069) | | (0.540) | | (0.290) | (0.074) | | (0.986) | |
| Water Program | -0.454 | -0.241** | 0.036 | -0.467 | -0.943 | | | | | |
| | (0.351) | (0.117) | (0.230) | (0.344) | (1.062) | | | | | |
| Enrollment in 1971 | -1.42 | -0.448 | -1.370* | -1.424 | -1.066 | -0.648 | -0.448 | -0.556 | -0.887 | -1.602 |
| | (0.954) | (0.367) | (0.707) | (0.945) | (1.779) | (1.158) | (0.378) | (0.705) | (1.120) | (2.492) |
| Inverse Mills Ratio | | | | -0.449 | 1.211 | | | | -0.848 | -0.915 |
| | | | | (1.205) | (2.774) | | | | (2.405) | (3.069) |
| Sub-district FE | Yes | No | No | Yes | Yes | Yes | No | No | Yes | Yes |
| Year of Birth FE | Yes | No | No | Yes | Yes | Yes | No | No | Yes | Yes |
| F-stat.(KP) | 5.67 | – | – | 13.10 | 3.05 | 9.66 | – | – | 10.20 | 3.44 |
| Sub-districts (N) | 499 | – | – | 495 | 210 | 444 | – | – | 443 | 189 |
| Obs | 2,804 | 4,575 | 4,575 | 2,778 | 796 | 2,508 | 4,596 | 4,596 | 2,494 | 660 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. Controls include interactions between cohort dummies and the allocation of the water and sanitation program in the district between 1973 and 1978. Also, we think about mean reversion by including the enrollment rate in 1971 at the district level interacted with year of birth dummies. Finally, columns (4), (5), (9) and (10) include the inverse mills ratio, estimated from a first-stage probit equation that models the selection into labor-force participation. In selection equations, we use the number of children living in the household as an additional variable that is not in wage equations. IV estimates are based on the number of school constructed by the INPRES SD reform at the sub-district level. In columns (1), (4), (5), (6), (9) and (10), we include sub-districts and birth year fixed effects. All estimations remove individuals aged between 15-13 in 1973 who might be partially affected by the number of primary schools constructed by the INPRES SD program. The sample is restricted to individuals born before 1973 due to endogenous places of birth following the reform.

at the sub-district level is used as an instrument. We adopt the strategy explained in Semykina and Wooldridge (2010). We first regress the participation to the wage labor-market with a Probit model, using all available exogenous variables (including the INPRES variable) and then obtain the inverse mills ratio. In the second stage, we estimate the earnings function with an IV 2SLS estimation where the entire set of exogenous variables are used as instruments, including the inverse mills ratio. Here again, we explore how gender impacts the results using the interaction term Education×Female. This approach produces consistent estimates of the parameters of interest provided standard conditions for identification hold, and provided the instruments are independent of the inverse mills ratio and uncorrelated with the residual of the earnings function.

The results are presented in Table 3.12. Column (1) controls for the endogeneity of education and ignores the selectivity using 2SLS estimate of the logarithmic monthly wage. The key parameter of interest is the point estimate on Education, the rate of return to an additional completed level of schooling. The marginal rate of return to schooling is 43% for males and 67% for females. However, the size effect is 10%. Turn now to columns (2) and (3) where we set aside the endogeneity issue and present the Heckman selection estimates. We only report the results from the wage equation.[31] We include all married-adult respondents in column (2) and restrict the analysis on women in column (3). In column (2), the coefficient on education is positive and highly significant. Moreover, the interaction between education and being a female has a positive sign, significant at the 5% level. The findings show that returns to female attainment are higher than returns to male attainment. Column (3) corroborates the results. We find in column (4) that, after simultaneously correcting for the selection bias and endogeneity bias, only the coefficient of the interaction (Education×Female) is significant, at a 5% size test. The effect disappears in column (5) when we restrict the sample to working women. The latter finding should be interpreted carefully due to the large variation in the sample size. Nevertheless, the non-significance of the selection correction term (inverse mills ratio) in columns (4) and (5) gives confidence on the standard errors. Accordingly, we do not correct for the generated regressor problem using bootstrapped standard errors.

Overall, the results reveal similar findings: the estimated marginal return to additional grades of completed schooling is positive and is significantly higher for females than for males. The findings are in line with theoretical assumptions and previous literature (Behrman et al., 1995, Dougherty, 2005). Controlling for the level of education, women earn less than men since the female dummy is negative

[31] The probit estimates of the wage work participation are available in Table C.5 in the Appendix.

and highly significant in all models. Total earnings are higher for men than for women, which suggest that the total labor market return to boys is greater than that to girls given the very low proportion of married women participating in the wage labor market. Returns to schooling are higher for females than for males. All the same, the findings confirm the existence of the indirect labor-market channel considering that education is causally linked to income, which in turn affects old-age transfer (Lillard and Willis, 1997).

Finally, we replicate the analysis investigating the effect of the respondent's level of education on the spouse's log of wage from columns (6) to (10). Results will give additional insights on the effect of the child's human capital on his couple's joint income and will directly asses the matching hypothesis on the marriage market. Clearly, regardless of the methodology, the level of education strongly increases the spouse's earnings, at a 5% size test in most equations. However, we do not observe any differential effect by the child's gender.

### 3.7.2 Intra-household bargaining

In this sub-section, we provide an empirical assessment on married-women to investigate the bargaining power channel. Specifically, we explore exogenous variation in women's bargaining power within household across ethnic groups to shed light on how it affects the impact of education on old-age support. Indeed, in Indonesia, patrilineal kinship requires women to move into the homes of their husbands upon marriage. In patriarchal societies, social norms or legal constraints allocate most of the decision power in the husband's favour (Baland and Ziparo, 2017). Economists have provided empirical assessments supporting the hypothesis that matrilocal women, who benefit from higher marriage options, are more empowered than patrilocal ones. Recently, Bargain, Loper, and Ziparo (2019) show that women originating from matrilocal ethnic groups benefit from a larger intrahousehold decision-making power. Accordingly, we follow the same approach and compare the effect of human capital on financial support to elderly parents according to the respondent's traditional cultural practice. We use information from Bargain et al. (2019) in which ethnic groups are associated to a post-marital residence norm, that determines whether the married couple should live with the wife's family or the husband's family after marriage. Then, we match the traditional post-marital residence practice with the respondent's ethnicity, available in IFLS-4.[32] We define a patrilocal dummy variable

[32]For individuals who did not participate in IFLS-4, we use the respondent's place of origin, available in IFLS-3, to define his ethnicity.

equals to one if the individual belongs to an ethnic group which practices patrilocal residence. In our sample, 81% (19%) of married women belong to matrilocal (patrilocal) groups. To estimate how the relationship between married women's education and parental support in old-age differs according to the traditional post-marital residence, we interact the level of education with the patrilocal dummy. As before, we conduct the analysis on the probability of giving money and on the amount transfer. We control for the endogeneity of education using the number of primary schools constructed by the INPRES SD program at the sub-district level.[33] Notice that the educational reform positively affects married-women's education, irrespective of their ethnicity.

Table 3.13 presents the results on married women (from columns 1 to 4) and married men (from columns 5 to 8). The pattern of women's education on the amount transfer appears to be somewhat different across ethnic groups only when the amount of transfer is included in level (not in logs) (column 4). If those women who are educated remit less money due to limited bargaining power in their newly formed household, the interaction between education and the patrilocal dummy should be negative. It is significant at a 5% size effect. The magnitude of the effect is large. Relative to women in matrilocal groups, women from patrilocal ethnicity are associated with lower transfer amount to elderly parents. An additional level of education increases by 14 (ten thousands of rupee) the amount given to parents for women in matrilocal group ($p<.001$). However, human capital does not produce any effect on the amount transfer for women belonging to patrilocal group.

Due to the existence of alternative channels, quantifying with any precision the magnitude of the labor-market returns and the bargaining mechanism separately is more challenging and is beyond the scope of the analysis. Nevertheless, estimates on how education affects financial support to elderly parents among patrilocal groups give some insight on the labor-market returns hypothesis. Similarly, the difference in the impact of education on transfers in matrilocal compared to patrilocal groups might predict the magnitude of the intrahousehold bargaining effect. This suppose that ethnicity is not correlated with earnings or labor-market opportunities, controlling for the level of education. However, Bargain et al. (2019) do not find consistent significant effect of the matrilocal/patrilocal norm on the labor market participation in Indonesia, using the the IFLS data. For married men, the exercise is more difficult given that there is no evidence that their bargaining power differs significantly according to the traditional post-marriage residence norm.

[33] The interaction between education and the patrilocal norm is also instrumented with the interaction of INPRES and the patrilocal dummy.

In this section, we identify a strong indirect effect of women's intra household bargaining on the relationship between education and the amount of transfers. Specifically, we showed that women with lower household bargaining transfer less to elderly parents. Given these heterogenous findings across women, it is surprising not to find any differential effect between men and women in subsection 5.2. We give here potential explanations. First, the overall effect of education on transfer might be driven by the large sample of women for whom higher education translates into higher bargaining power (i.e., matrilocal women) as much as men. The argument makes sense since only 18% of women are from a patrilocal ethnic group in our sample. In addition, if we assume that women are slightly disadvantaged compared to men due to lower household bargaining power for same level of education, there might exist a counter-balancing indirect effect in favour of married-women. In our case, this might come from the labor-market returns to education. The latter is in line with the above sub-section on the differential returns to education for men and women in Indonesia. Finally, if we admit the results of Table 3.7, the absence of heterogeneous effects might be the reflect of non-linear effects of education on transfers among girls with different educational levels.

## 3.8 Conclusion

Our paper contrasts with previous studies on the effect of education on the parental repayment hypothesis. Firstly, we take into account the endogeneity of education and investigate the causal effect of human capital on transfer to old parents using an exogenous education policy in Indonesia. Secondly, the paper models an indirect effect not only based on labor-market returns to education. A channel related to intra-household bargaining power within the newly formed household is included. The results demonstrate that taking into account the marriage dimension could arguably change the investment decisions of the parents. Indeed, we find that a decrease in education costs, measured through expansion of primary schools at the local (sub-district) level, increases educational attainment for both males and females. In addition, we find that human capital strongly affects the amount given to old parents. Whereas our theoretical model identifies differences in the strength of the relationship between education and transfer across gender, our empirical findings do not. Nevertheless, our empirical results clearly confirms the plausibility of marital returns to education modeled theoretically, at least for married women. Indeed, we compare the effect of education on financial transfer to parents across women belonging to different ethnic groups which reflect differences in intra-household

**Table 3.13 –** Effect of education on financial support according to ethnic groups and gender

| Sample | Women | | | | Men | | | |
|---|---|---|---|---|---|---|---|---|
| Estimation | First-stage | Second-stage | | | First-stage | Second-stage | | |
| Dependent variable | Education | Money | Amount (log) | Amount | Education | Money | Amount (log) | Amount |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| INPRES | 0.068*** | | | | 0.033** | | | |
| | (0.018) | | | | (0.014) | | | |
| INPRES×Female | 0.002 | | | | 0.038 | | | |
| | (0.023) | | | | (0.025) | | | |
| Education | | 0.157* | 0.787*** | 14.79*** | | 0.265 | 0.373 | 14.68 |
| | | (0.090) | (0.209) | (5.396) | | (0.201) | (0.447) | (11.93) |
| Education×Patrilocal | | -0.101 | -0.356 | -15.06** | | -0.140 | 0.387 | 10.75 |
| | | (0.087) | (0.292) | (6.729) | | (0.124) | (0.345) | (8.845) |
| Patrilocal | 0.428* | 0.077 | 0.217 | 16.67 | 0.098 | 0.030 | -0.541 | -12.93 |
| | (0.234) | (0.181) | (0.615) | (14.70) | (0.277) | (0.237) | (0.789) | (18.21) |
| Sub-districts (N) | 753 | 334 | 273 | 273 | 747 | 317 | 258 | 258 |
| F-stat.(KP) | – | 10.75 | 12.88 | 12.88 | – | 3.76 | 3.51 | 3.51 |
| Observations | 2,104 | 1,685 | 1,211 | 1,211 | 1,952 | 1,522 | 1,120 | 1,120 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. INPRES represents the number of school constructed by the Indonesian reform at the sub-district level, ranging from 0 to 15. We remove individuals aged between 12 to 15 in 1974 who might be partially affected by the number of primary schools constructed by the program. Patrilocal is a dummy equals to 1 for patrilocal ethnic groups and 0 for matrilocal groups. All estimates include sub-district and year of birth fixed effects and control for the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971.

bargaining. We showed that the optimal transfer given by married women to old parents positively depends on her household bargaining and on the couple's joint income.

# References

Akresh, R., D. Halim, and M. Kleemans (2018). Long-term and intergenerational effects of education: Evidence from school construction in indonesia. Technical report, National Bureau of Economic Research.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Anukriti, S. and S. Dasgupta (2017, August). *Marriage Markets in Developing Countries*, Volume 1. Oxford University Press.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2018). Bride price and female education. *Forthcoming, Journal of Political Economy, (2019)*, 0616_8023R1.

Baland, J.-M. and R. Ziparo (2017). Intra-household bargaining in poor countries. *Towards Gender Equity in Development*, 69.

Bargain, O., J. Loper, and R. Ziparo (2019). Traditional norm, access to divorce and women's empowerment: Evidence from indonesia. *Working Paper*.

Bau, N. (2019). Can policy change culture? government pension plans and traditional kinship pratices. *CEPR Discussion Paper No. DP13486*.

Baudin, T., D. De la Croix, and P. E. Gobbi (2015). Fertility and childlessness in the united states. *American Economic Review*.

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.

Behrman, J. R. (1997, January). Chapter 4 Intrahousehold distribution and the family. In *Handbook of Population and Family Economics*, Volume 1, pp. 125–187. Elsevier.

Behrman, J. R. and A. B. Deolalikar (1995). Are there differential returns to schooling by gender? the case of indonesian labour markets. *Oxford Bulletin of Economics and Statistics 57*(1), 97–117.

Behrman, J. R., A. B. Deolalikar, et al. (1995). Are there differential returns to schooling by gender? the case of indonesian labour markets. *Oxford Bulletin of Economics and Statistics 57*(1), 97–117.

Blundell, R., P.-A. Chiappori, and C. Meghir (2005). Collective labor supply with children. *Journal of Political Economy*.

Boulier, B. L. and M. R. Rosenzweig (1984). Schooling, Search, and Spouse Selection: Testing Economic Theories of Marriage and Household Behavior. *Journal of Political Economy 92*(4), 712–732.

Browning, M., P.-A. Chiappori, and Y. Weiss (2014). *Economics of the Family*. Cambridge University Press.

Bun, M., M. De Haan, et al. (2010). Weak instruments and the first stage f-statistic in iv models with a nonscalar error covariance structure. *University of Amsterdam*.

Caldwell, J. C., P. H. Reddy, and P. Caldwell (1983, November). The causes of marriage change in South India. *Population Studies 37*(3), 343–361.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160.

Carmichael, S. (2011, October). Marriage and power: Age at first marriage and spousal age gap in lesser developed countries. *The History of the Family 16*(4), 416–436.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica 56*(1), 63–89.

Comola, M. and L. de Mello (2013). Salaried employment and earnings in indonesia: new evidence on the selection bias. *Applied Economics 45*(19), 2808–2816.

Cox, D. (1987). Motives for private income transfers. *Journal of political economy 95*(3), 508–546.

Cox, D., E. Galasso, and E. Jimenez (2006). Private Transfers in a Cross Section of Developing Countries. Technical report, Center for Retirement Research.

De Neve, J.-W. and G. Fink (2018). Children's education and parental old age survival–quasi-experimental evidence on the intergenerational effects of human capital investment. *Journal of health economics 58*, 76–89.

Deolalikar, A. B. (1993). Gender differences in the returns to schooling and in school enrollment rates in indonesia. *The Journal of Human Resources 28*(4), 899–932.

Dougherty, C. (2005). Why are the returns to schooling higher for women than for men? *The Journal of Human Resources 40*(4), 969–988.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American economic review 91*(4), 795–813.

Eika, L., M. Mogstad, and B. Zafar (2019). Educational assortative mating and household income inequality. *Journal of Political Economy 127*(6), 2795–2835.

Fafchamps, M. and A. R. Quisumbing (2005). Marriage, Bequest, and Assortative Matching in Rural Ethiopia. *Economic Development and Cultural Change 53*(2), 347–380.

Fafchamps, M. and A. R. Quisumbing (2007, January). Chapter 51 Household Formation and Marriage Markets in Rural Areas. In T. P. Schultz and J. A. Strauss (Eds.), *Handbook of Development Economics*, Volume 4, pp. 3187–3247. Elsevier.

Fernandez, R., N. Guner, and J. Knowles (2005). Love and money: A theoretical and empirical analysis of household sorting and inequality. *The Quarterly Journal of Economics 120*(1), 273–344.

Fernández, R. and R. Rogerson (2001). Sorting and long-run inequality. *The Quarterly Journal of Economics 116*(4), 1305–1341.

Frankenberg, E., L. Lillard, and R. J. Willis (2002). Patterns of intergenerational transfers in southeast asia. *Journal of Marriage and Family 64*(3), 627–641.

Garenne, M. (2004). Age at marriage and modernisation in sub-Saharan Africa. *Southern African Journal of Demography 9*(2), 59–79.

Geertz, H. (1961). *The Javanese Family: a Study of Kinship and Socialization*.

Gyimah, S. O. (2008, October). Cohort Differences in Women's Educational Attainment and the Transition to First Marriage in Ghana. *Population Research and Policy Review 28*(4), 455.

Ham, J. C. and H. Song (2014). The determinants of bargaining power in an empirical model of transfers between adult children, parents, and in-laws for south korea. *Journal of Development Economics 109*(73–86).

Horowitz, A. W. and J. Wang (2004, April). Favorite son? Specialized child laborers and students in poor LDC households. *Journal of Development Economics 73*(2), 631–642.

Iyigun, M. and R. Walsh (2007). Endogenous gender power, household labor supply and the demographic transition. *Journal of Development Economics*.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics 7*(1), 63–88.

Jensen, R. and N. Miller (2017). Keepin' 'em down on the farm: Old age security and strategic underinvestment in children. *NBER working paper 23122*.

Jensen, R. and R. Thornton (2003). Early Female Marriage in the Developing World. *Gender and Development 11*(2,), 9–19.

Kumar, A. (2013, November). Preference based vs. market based discrimination: Implications for gender differentials in child labor and schooling. *Journal of Development Economics 105*, 64–68.

LeBar, F. M. (1972). *Ethnic groups of insular Southeast Asia*.

Lee, J. (2007). Marriage, the sharing rule, and pocket money: The case of south korea. *Economic Development and Cultural Change 55*(3), 557–581.

Lee, Y.-J., W. L. Parish, and R. J. Willis (1994). Sons, daughters, and intergenerational support in taiwan. *American journal of sociology 99*(4), 1010–1041.

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

Lillard, L. A. and R. J. Willis (1997). Motives for intergenerational transfers: Evidence from malaysia. *Demography 34*(1), 115–134.

Lilleor, H. (2015). Human capital diversification among children. *Copenhagen: Rockwool Foundation Research Unit*.

Lucas, R. E. and O. Stark (1985). Motivations to remit: Evidence from botswana. *Journal of political Economy 93*(5), 901–918.

Mabsout, R. and I. Van Staveren (2010). Disentangling bargaining power from individual and household level to institutions: Evidence on women's position in ethiopia. *World Development 38*(5), 783–796.

Mahmood, T. (1992). *Law and the elderly in the Islamic tradition*.

Nöldeke, G. and L. Samuelson (2015). Investment and competitive matching. *Econometrica*.

Ortiz, I. (2014). Social protection for older persons: Key policy trends and statistics. *Geneva: International Labour Office*.

Peters, M. and A. Siow (2002). Competing premarital investments. *Journal of Political Economy*.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan (2012, December). Human capital investment and the gender division of labor in a brawn-based economy. *American Economic Review 102*(7), 3531–60.

Quisumbing, A. R., J. R. Behrman, J. A. Maluccio, A. Murphy, and K. M. Yount (2005, June). Levels, correlates, and differences in human, physical, and financial assets brought into marriages by young Guatemalan adults. *Food and Nutrition Bulletin 26*(2 Suppl 1), S55–67.

Quisumbing, A. R., J. P. Estudillo, and K. Otsuka (2003, November). *Land and Schooling: Transferring Wealth across Generations*. Baltimore: International Food Policy Research Institute.

Quisumbing, A. R. and K. Hallman (2003). Marriage in Transition: Evidence on Age, Education, and Assets from Six Developing Countries. pp. 74.

Raut, L. K. (1990, November). Capital accumulation, income distribution and endogenous fertility in an overlapping generations general equilibrium model. *Journal of Development Economics 34*(1), 123–150.

Raut, L. K. (1997). Learning to Perfect Manipulation: Implications for Fertility, Savings, and Old-Age Social Security. Technical Report 199704, University of Hawaii at Manoa, Department of Economics.

Raut, L. K. and L. H. Tran (2005). Parental human capital investment and old-age transfers from children: Is it a loan contract or reciprocity for indonesian families? *Journal of Development Economics 77*(2), 389–414.

Sanderson, E. and F. Windmeijer (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of Econometrics 190*(2), 212–221.

Schaner, S. and S. Das (2016). Female labor force participation in asia: Indonesia country study. *Asian Development Bank Economics Working Paper Series* (474).

Semykina, A. and J. M. Wooldridge (2010). Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics 157*(2), 375–380.

Strauss, J. and D. Thomas (1995). Human resources: Empirical modeling of household and family decisions. In *Handbook of Development Economics*, Volume 3, Part 1, pp. 1883–2023. Elsevier.

Trostel, P., I. Walker, and P. Woolley (2002). Estimates of the economic return to schooling for 28 countries. *Labour Economics 9*(1), 1 – 16.

Vanwey, L. K. (2004). Altruistic and contractual remittances between male and female migrants and households in rural thailand. *Demography 41*(4), 739–756.

Wirakartakusumah, D. (1998). Household structure and the elderly in indonesia. *Jakarta Demographic Institute, University of Indonesia unpublished manuscript*.

World Bank (1989). Indonesia - basic education study (english). http://documents.worldbank.org/curated/en/963861468044101399/Indonesia-Basic-education-study.

Zhang, J. and K. Nishimura (1993, June). The old-age security hypothesis revisited. *Journal of Development Economics 41*(1), 191–202.

Zhang, J., Y. Zhao, A. Park, and X. Song (2005). Economic returns to schooling in urban china, 1988 to 2001. *Journal of Comparative Economics 33*(4), 730 – 752.

# General Conclusion

The recent increase in the number of out-of-school children in developing countries calls for reexploring and reconsidering the standard educational framework used by the international community. Identifying new determinants of educational choices, studying their origins, and proposing solutions to improve education in developing countries are both major issues of research in economics and major concerns for policy.

This thesis situates in this line of thinking the concept of making education and gender inequality its principal object of analysis. The thesis interrogates the role of intrahousehold allocations and heterogeneous endowments in education in three different contexts and scenarios. Chapter one explores how household members, mothers in particular, have different roles and command over children's educational resources. Chapter two investigates how sibling composition alters the impact of exogenous income shocks on children's education in rural Tanzania. Chapter three analyzes how expectations about old-age support predict investments in girls and boys' education of the parents in Indonesia. From an empirical perspective, this thesis provides the first study that attempts to integrate concerns about intrahousehold allocations with those on intergenerational parent-child exchange when studying education in developing countries. Different roles, endowments, and situations children face are taken into account to highlight new findings. The remaining part of this conclusion presents the main results of the thesis, before discussing its main limitations and possible extensions.

Chapter 1 offers a critical review of the microeconomic approach of women's resources and endowments. This paper reviews the analytical frameworks on children's education. The study explores the roots of the relationship between women's economic rights and children's education at a global level, across 75 developing countries. We establish a positive correlation between women's agency; which reflects how much autonomy women enjoy that would allow them to act on their own in their economic and family environment; and children's education. In this analysis, I argue that women's ability to make choices within the household and in the society is a pre-requisite of improvements in educational attainment. However, this chapter suffers from several limitations. Firstly, this study ignores how values, norms and preferences are formed. Because of data limitations, the paper does not measure the degree of law enforcement. Readers should be aware that a law enrolled in the Constitution does not guarantee that the law is not entirely shaped by asymmetry of information. Indeed, women in rural regions may not be endowed with same information, ignoring particular changes in the legislation that may improve their empowerment and economic status. Also, the chapter is not being able to clearly disentangle the mechanisms at play. Additional data at the individual level will allow for further and more detailed research.

Chapter 2 revisits the relationship between income shocks and children's education in an empirical analysis on rural Tanzania. The paper provides additional insights on one dimension that has received little attention in the literature on family income variations: the role of intra-household allocations and heterogeneous endowments. Heterogeneity is explicitly introduced through measures reflecting the family demographic composition. The study documents a strong causal and heterogeneous impact of income shocks on the likelihood of school dropout for both girls and boys. Income shocks, measured through unusual climate conditions, affect differently children's outcomes according to the gender composition of (younger) siblings. Unfavourable climate shocks for agricultural crops increase the probability of school dropout for children with sisters and decrease the probability for children with relatively more brothers. The quantitative data are contextualized through a qualitative survey that I realized in 2018 in a sub-sample of the KHDS villages. Incorporating the qualitative research helps me to test the hypothesis underlying my quantitative findings: girls have lower educational returns than boys and educational returns are non-linear. Due to data limiations, the paper remains silent on the quality of education though. Introducing additional measures on test scores should provide additional insights on the existence of alternative mechanisms. Moreover, an extended analysis could exploit the empirical framework to study other outcomes such as adolescent pregnancy. Indeed, focus groups discussions among KHDS villages raised the issue of adolescent pregnancy in Tanzania. 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. Girls that fall pregnant are routinely expelled from school and prevented from returning. According to the 2015-2016 Tanzania Demographic and Health Survey, one in four women aged 15-19 are mothers. In such setting, it would be interesting to explore the heterogeneous effect of income shocks on adolescent pregnancy since the relationship between fertility and education has never been so strong.

Chapter 3 explores the role of the old-age insurance in the human capital decisions of the parents in Indonesia, by the child's gender. To correct for the endogeneity of educational attainment, we exploit one of the largest primary school construction program used intensively in the literature: the Inpres program. The key empirical finding is that monetary transfers to parents are increasing in education of children and that there is no significant differences by gender. Two mechanisms are at play: the labor and marital returns to education. However, we can reasonnably suppose that other channels are driving our results. For instance, social norms are important in explaining old-age support from adult children to elderly parents in Indonesia. In this case, social norms could provide an alternative explanation of our results, although daughters are more often associated to be the caregiver. All the same, in an imperfect way, we provide empirical evidence for the existence of household bargaining mechanism. We compare the magnitude of the relationship between education and monetary assistance across demographic groups in which household bargaining power varies in an exogenous way. Another limitation of this study is not being able to take into account the decision to participate into the labor-market. Extensive and intensive margins of labor supply are important dimensions to consider especially for Indonesian women, as a large proportion of married women do not work. In our benchmark model where assistance to elderly parents is modeled by monetary transfers, it would be interesting to endogenize labor supplies by introducing household production. In this case, spouses could decide how to allocate their time between working on the labor market, leisure and time spent at home for domestic production (i.e. doing chores). Finally, another possible extension of this study would be to explore education and family transfers within the extended family. This requires available data on the educational attainment and transfers of all family members, including siblings. Integrating monetary transfers exchanged between siblings who coreside and do not coreside with elderly parents would help to investigate how past educational choices determined the organization of elderly care within poor families.

# A Appendix to the Chapter 1

**Table A.1 –** Countries included in the sample

| | | | | |
|---|---|---|---|---|
| Algeria | Ivory Coast | Kyrgyz Rep. | Niger | Tanzania |
| Argentina | Dominican Rep. | Lao PDR | Pakistan | Thailand |
| Bangladesh | Egypt Arab. Rep. | Lesotho | Papua New Guinea | Togo |
| Benin | Fiji | Liberia | Paraguay | Tunisia |
| Bolivia | Gabon | Malawi | Peru | Turkey |
| Botswana | Ghana | Malaysia | Philippines | Uganda |
| Brazil | Guatemala | Mali | Rwanda | Ukraine |
| Cambodia | Honduras | Mauritania | Senegal | Venezuela RB. |
| Cameroon | India | Mexico | Sierra Leone | Vietnam |
| Central Af. Rep. | Mongolia | Indonesia | South Africa | Yemen Rep. |
| Chile | Iran Islamic. Rep. | Morocco | Sri Lanka | Zambia |
| China | Jamaica | Mozambique | Sudan | Zimbabwe |
| Colombia | Jordan | Namibia | Swaziland | Bulgaria |
| Congo Dem. Rep. | Kazakhstan | Nepal | Syrian Arab. Rep. | |
| Costa Rica | Kenya | Nicaragua | Tajikistan | |

**Table A.2** – Duration of compulsory education in years, by region

| Region | 1998 | 2000 | 2005 | 2010 |
|---|---|---|---|---|
| Sub-Saharan Africa | 7 | 7 | 8 | 8 |
| East Asia and Pacific | 7 | 8 | 9 | 9 |
| Middle East and North Africa | 9 | 9 | 9 | 9 |
| Europe and Central Asia | 9 | 9 | 9 | 9 |
| Latin America and Caribbean | 9 | 9 | 9 | 10 |
| Low Income | 7 | 7 | 7 | 8 |
| Middle Income | 9 | 9 | 9 | 9 |

In each region, high income countries are excluded. Source: Unesco Institute for Statistics: https://data.worldbank.org/indicator/SE.COM.DURS

# B Appendix to the Chapter 2

Over the past decades, farmers in Kagera have diversified both within agriculture (for example by growing cabbages, tomatoes, green peppers and vanilla) and by moving into new non-agricultural income-generating activities, such as Nile perch fishing, mining, trading, and so on (Ellis, 1998, 2000). The most important permanent crop in Kagera region was banana (66%), followed by coffee (23%). The most prevalent staple crops include maize, cassava, rice, sorghum and millet, while the main exported crops are sugar, coffee, cotton, tobacco and tea. The main cropping system is the traditional method, which combines perennial crops (usually bananas and coffee) with annual food crops such as maize, beans, cassava, sweet potatoes and sorghum in certain areas. The area each family cultivates varies with population density but is generally between 1 and 2 hectares. The coffee crop is found in most of the districts and is the main cash-crop. The major economic activity for the people of Kagera is agriculture, for most, subsistence farming. The most important food crops for the area are bananas and beans. Cofee, cotton and tea are the main cash crops and all are grown at subsistence level.

**Table B.1** – The livelihood activities/source of income of the households ranked in order of importance, by district

| District | Livelihood Activity | | | | | | |
|---|---|---|---|---|---|---|---|
| | Annual Crop Farming | Permanent Crop Farming | Livestock Keeping/ Herding | Off Farm Income | Remitt-ances | Fishing/ Hunting & Gathering | Tree/ Forest Resources |
| Karagwe | 2 | 1 | 3 | 4 | 6 | 7 | 5 |
| Bukoba Rural | 2 | 1 | 4 | 3 | 5 | 7 | 6 |
| Muleba | 1 | 2 | 4 | 3 | 7 | 6 | 5 |
| Biharamulo | 1 | 2 | 3 | 4 | 7 | 5 | 6 |
| Ngara | 1 | 2 | 3 | 4 | 6 | 7 | 5 |
| Bukoba Urban | 1 | 2 | 4 | 3 | 5 | 7 | 6 |

Source: . Off-farm income refers to cash generated from non-agricultural activities. This can be either from permanent employment (i.e., government, private sector or other), temporary employment or labourers. It also includes cash generated from working on farms belonging to other farmers.

**Table B.2** – List of selected districts and villages in the qualitative survey

| District | Village | Level of Development |
|---|---|---|
| Bukoba Rural | Omubweha | Remote |
| | Nyakibimbili | Semi Remote |
| | Mubafu | Semi Urban |
| | Igombe | Urban |
| Muleba | Nyakabamo | Remote |
| | Karambi | Semi Remote |
| | Kishamda | Semi Urban |
| | Izigo | Urban |
| Missenyi | Ishumju | Remote |
| | Kanyigo | Semi Remote |
| | Nsumsa | Semi Urban |
| | Kyaka | Urban |
| Bukoba Urban | Ijusanyondo | Green Belt |
| | Kahororo | Green Belt |
| | Hyanga | Green Belt |

No villages were selected in the remaining four administrative units of the Kagera region: Biharamulo, Karagwe, Kyerwa, Ngara.

**Table B.3 –** Number of the KHDS communities which grow the listed crops by order of importance in 1991 and 2004

| Type of crops farmed | 1991 | | | | 2004 | | | |
|---|---|---|---|---|---|---|---|---|
| | First | Second | Third | Fourth | First | Second | Third | Fourth |
| Coffee | 43 | – | – | – | 10 | 16 | 3 | 4 |
| Bananas | 1 | 42 | 6 | 1 | 23 | 12 | 6 | 6 |
| Cotton | 2 | – | – | – | – | – | – | – |
| Cassava | 3 | 2 | 34 | 4 | 1 | 6 | 12 | 7 |
| Tea | – | 2 | – | – | – | – | – | – |
| Sweet Potatoes | – | 2 | 5 | 29 | 3 | 2 | 5 | 8 |
| Maize | – | 1 | 2 | 10 | 2 | – | 9 | 13 |
| Millet | – | – | 1 | 2 | – | – | – | – |
| Sorghum | – | – | – | 1 | 1 | 1 | – | – |
| Rice | – | – | – | 1 | – | – | 1 | – |
| Beans | – | – | – | – | 8 | 5 | 11 | 5 |
| Groundnuts | – | – | – | – | – | – | 1 | 3 |
| Tobacco | – | – | – | – | – | – | – | 1 |
| Other vegetables | – | – | 1 | – | – | – | – | 1 |

Source: KHDS-1 and KHDS-2 Community Questionnaires. The agriculture questionnaire was administered to the agricultural officer in all of the baseline enumeration areas for a total of 49 community interviews. In 2004, for one community the information is missing.

**Table B.4 –** The effect of drought on child labor, depending on the young/old sibling composition and child's gender

| | **Housework** | | | | **Agriculture** | | | | **Total** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dependent variable | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls |
| Sibling group | Young | | Old | | Young | | Old | | Young | | Old | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Shock | 0.313 | 0.867 | 0.969 | 0.670 | 0.331 | 4.284** | -0.999 | -0.693 | 1.143 | 5.231** | -0.140 | 0.642 |
| | (0.955) | (1.085) | (1.255) | (1.482) | (0.747) | (1.736) | (1.321) | (1.830) | (1.696) | (2.570) | (1.906) | (3.032) |
| Shock × Sex ratio | -0.657 | 0.337 | -2.452 | -0.935 | -0.416 | -3.126** | 1.015 | -1.927 | -1.989 | -2.570 | -2.613 | -2.523 |
| | (1.146) | (1.124) | (1.494) | (1.589) | (1.255) | (1.361) | (1.163) | (1.339) | (1.595) | (2.021) | (2.402) | (2.882) |
| Shock × Siblings | -0.002 | -0.259 | -0.106 | -0.162 | -0.484 | -1.021** | -0.165 | 0.187 | -0.529 | -1.249** | -0.139 | -0.106 |
| | (0.321) | (0.222) | (0.183) | (0.209) | (0.305) | (0.410) | (0.158) | (0.266) | (0.549) | (0.584) | (0.249) | (0.359) |
| Father in hh | -3.822*** | 5.513*** | -3.460** | 3.366 | -1.944 | -4.140** | -2.695 | -4.031*** | -2.539 | 3.163 | -2.585 | 0.562 |
| | (1.414) | (1.404) | (1.748) | (2.100) | (1.718) | (2.052) | (1.973) | (1.371) | (2.472) | (2.264) | (2.940) | (2.454) |
| Mother in hh | 0.887 | -3.198*** | -0.025 | -3.839*** | -0.227 | 0.080 | 0.284 | -0.096 | 0.487 | -2.499 | 1.723 | -3.007 |
| | (1.848) | (1.003) | (2.584) | (1.060) | (1.738) | (1.868) | (2.266) | (2.289) | (3.568) | (2.036) | (4.606) | (2.505) |
| Age | 0.122 | 1.498 | -0.547 | 0.702 | -1.480 | -2.187 | -1.648 | -1.803 | -1.605 | -1.840 | -1.769 | -2.471 |
| | (1.156) | (0.980) | (1.172) | (1.072) | (1.577) | (1.365) | (1.788) | (1.627) | (2.164) | (1.882) | (2.168) | (2.254) |
| Season | -1.265*** | -0.587 | -1.510*** | -1.210** | 1.169* | 1.263* | 0.860 | 0.643 | 0.471 | 0.567 | -0.280 | -0.655 |
| | (0.331) | (0.729) | (0.351) | (0.552) | (0.706) | (0.762) | (0.859) | (0.768) | (0.985) | (1.210) | (1.158) | (1.123) |
| Observations | 1,241 | 1,336 | 1,197 | 1,254 | 1,276 | 1,366 | 1,230 | 1,287 | 1,191 | 1,310 | 1,161 | 1,228 |
| Individuals | 498 | 546 | 503 | 542 | 535 | 583 | 545 | 585 | 490 | 542 | 496 | 538 |
| R-squared | 0.008 | 0.011 | 0.013 | 0.014 | 0.010 | 0.022 | 0.007 | 0.008 | 0.004 | 0.010 | 0.005 | 0.003 |

*** p<0.01, ** p<0.05, * p<0.1. Standard errors corrected for spatial autocorrelation in parentheses. Results are estimated using the first five waves of the KHDS with linear regressions. From columns (1) to (4), the dependent variable is the number of working hours for individuals aged 7-18 years old in housework during the last 7 days. Results on agriculture hours are presented from columns (5) to (8) and on total work from columns (9) to (12). Shock is the average of the spei index during the month of the survey, in absolute value. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are computed using younger (older) siblings only in the first (last) two columns of each dependent variable. Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. Season equals 1 if the month of interview belongs to the growing season. All estimates include individual and survey fixed effects.

**Table B.5 –** Alternative measure of drought on child labor

| Dependent variable | Housework | | | Agriculture | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | All | Boys | Girls | All | Boys | Girls | All | Boys | Girls |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Shock | 0.456 | 1.630* | 1.819** | -0.366 | 0.485 | 3.421*** | 0.191 | 2.425* | 5.520*** |
| | (0.284) | (0.924) | (0.793) | (0.439) | (1.006) | (0.816) | (0.623) | (1.428) | (1.139) |
| Shock × Female | -0.578 | | | 0.104 | | | -0.408 | | |
| | (0.532) | | | (0.272) | | | (0.658) | | |
| Shock × Sex ratio | | -1.256 | -1.372 | | 0.259 | -3.929*** | | -1.942 | -5.616*** |
| | | (0.890) | (1.376) | | (0.811) | (1.435) | | (1.334) | (2.120) |
| Shock × Siblings | | -0.123 | -0.153* | | -0.192 | -0.249** | | -0.271 | -0.398*** |
| | | (0.099) | (0.087) | | (0.133) | (0.100) | | (0.175) | (0.145) |
| Father in hh | -0.007 | -3.839*** | 4.713** | -2.601* | -1.660 | -3.908** | -0.563 | -2.524 | 1.872 |
| | (1.480) | (1.380) | (2.033) | (1.396) | (1.724) | (1.744) | (2.044) | (2.518) | (2.977) |
| Mother in hh | -2.143*** | 0.993 | -3.452*** | 0.956 | 0.347 | 0.208 | -0.671 | 1.267 | -2.499 |
| | (0.749) | (1.848) | (0.990) | (1.272) | (1.649) | (1.887) | (1.665) | (3.456) | (1.954) |
| Age | 0.754 | 0.444 | 1.526 | -0.951 | -0.558 | -0.876 | -0.0137 | 0.039 | 0.787 |
| | (0.897) | (1.041) | (1.002) | (1.534) | (1.902) | (1.322) | (2.053) | (2.380) | (2.043) |
| Season | -0.949*** | -1.336*** | -0.671 | 0.820 | 0.691 | 1.041 | 0.053 | -0.110 | 0.258 |
| | (0.364) | (0.284) | (0.515) | (0.697) | (0.723) | (0.746) | (0.917) | (1.021) | (0.991) |
| Observations | 3,011 | 1,453 | 1,517 | 3,101 | 1,495 | 1,560 | 2,930 | 1,401 | 1,488 |
| Individuals | 1,258 | 599 | 644 | 1,362 | 646 | 696 | 1,243 | 590 | 638 |
| R-squared | 0.006 | 0.011 | 0.014 | 0.004 | 0.006 | 0.018 | 0.001 | 0.004 | 0.012 |

*** p<0.01, ** p<0.05, * p<0.1. Standard errors corrected for spatial autocorrelation in parentheses. Results are estimated using the first five waves of the KHDS with linear regressions. From columns (1) to (3), the dependent variable is the number of working hours in housework during the last 7 days for individuals aged 7-18 years old. Results on agriculture hours are presented from columns (4) to (6) and on total work from columns (7) to (9). Shock represents the number of shocks (absolute value of the spei index $\geq 1$) in the past 3 months. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. Season equals 1 if the month of interview belongs to the growing season. All estimates include individual and survey fixed effects.

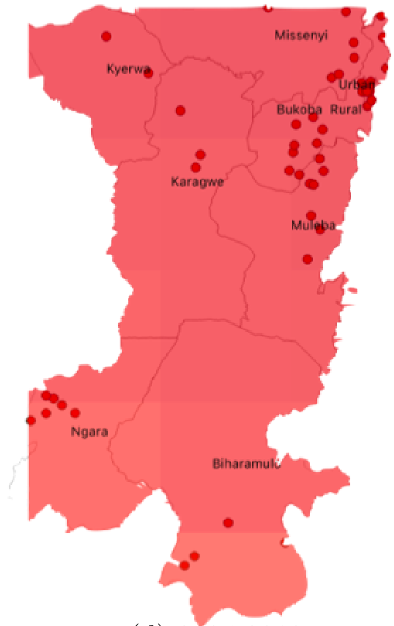
**Table B.6 –** The effect of drought on the number of school hours, depending on sibling composition and child's gender

| | | Hours in school, last week | | | | | |
|---|---|---|---|---|---|---|---|
| Dependent variable | | Boys | Girls | Boys | Girls | Boys | Girls |
| Sibling group | | All | | Young | | Old | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Shock | -0.047 | -0.043 | -0.170 | -0.065 | -0.228* | 0.098 | 0.090 |
| | (0.040) | (0.089) | (0.174) | (0.090) | (0.118) | (0.119) | (0.115) |
| Shock × Female | 0.115** | | | | | | |
| | (0.049) | | | | | | |
| Shock × Sex ratio | | -0.221 | 0.280 | -0.163 | 0.302** | -0.249* | -0.023 |
| | | (0.136) | (0.205) | (0.128) | (0.142) | (0.127) | (0.131) |
| Shock × Siblings | | 0.013 | 0.016 | 0.019 | 0.043** | -0.004 | -0.013 |
| | | (0.010) | (0.014) | (0.021) | (0.022) | (0.022) | (0.015) |
| Father in hh | 0.036 | -0.191 | 0.328** | -0.150 | 0.687* | -0.242 | 0.280 |
| | (0.127) | (0.195) | (0.155) | (0.192) | (0.399) | (0.221) | (0.185) |
| Mother in hh | -0.350 | -0.002 | -0.513* | 0.022 | -0.743* | -0.308 | -0.735** |
| | (0.230) | (0.305) | (0.291) | (0.303) | (0.382) | (0.450) | (0.371) |
| Age | -0.172 | -0.273* | -0.093 | -0.349** | -0.162 | -0.217 | -0.168 |
| | (0.106) | (0.142) | (0.160) | (0.163) | (0.169) | (0.162) | (0.182) |
| Season | -0.062 | -0.122 | 0.007 | -0.137 | -0.016 | -0.132 | -0.050 |
| | (0.062) | (0.095) | (0.085) | (0.106) | (0.093) | (0.105) | (0.090) |
| Observations | 1,007 | 492 | 502 | 425 | 449 | 394 | 421 |
| Individuals | 388 | 190 | 192 | 163 | 170 | 153 | 161 |

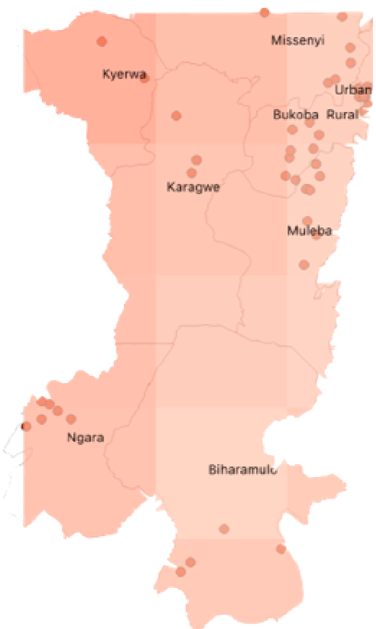
*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses, clustered at the village level. Results are estimated using the first five ways of the KHDS survey with Poisson regressions. The dependent variable is the number of hours spent in school during the last week. For those who are eligible to be in secondary school (aged less than 17 years old) and do not attend school, I assign them the value of 0. Shock represents the number of shocks (absolute value of the spei index $\geq$1) in the past 3 months. Sex ratio represents the number of brothers over the total number of siblings. Siblings denotes the number of biological siblings, irrespective of their live status or place of residence. The variables are analogously computed using younger siblings only in columns (4), (5) and older siblings only in columns (6), (7). Father in hh is a dummy variable that equals one if the father lives in the household. Mother is analogously defined. Season equals 1 if the month of interview belongs to the growing season. All estimates include individual and survey fixed effects.

(a) 1975-1985


(b) 1985-1995




(c) 1955-2005




(d) 2005-2010

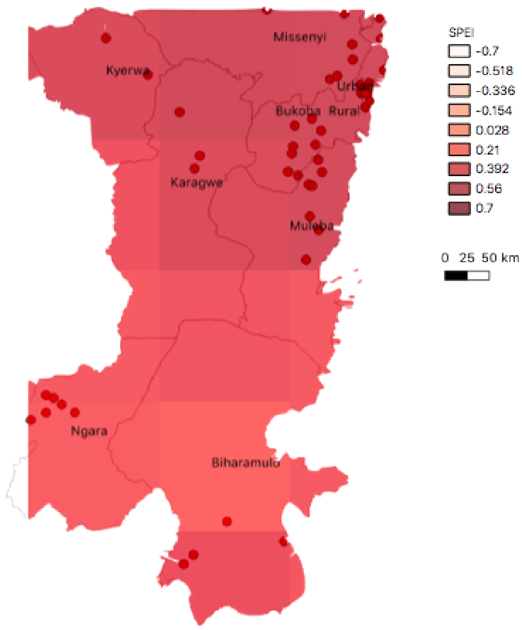



**Figure B.1 –** The evolution of the SPEI over time. Light colours indicate more droughts (a lower Spei level). The figure represents row data of the index (not absolute values). The red dots represent communities interviewed in KHDS-1

# C Appendix to the Chapter 3

## C.1 Formal theoretical proofs

### C.1.1 Proof that the impact of education on the bargaining power of females is positive

We solve for the derivative of the daughter's bargaining power only:

$$\begin{aligned}\frac{\partial \theta^f(e)}{\partial e} &= \frac{\left(\mu \frac{\partial w^f(e)}{\partial e}[w^f(e)]^{\mu-1}([w^f(e)]^{\mu}+[w^m(e)]^{\mu}) - [w^f(e)]^{\mu}(\mu \frac{\partial w^f(e)}{\partial e}[w^f(e)]^{\mu-1} + \mu \frac{\partial w^m(e)}{\partial e}[w^m(e)]^{\mu-1})\right)}{([w^f(e)]^{\mu}+[w^m(e)]^{\mu})^2} \\ &= \frac{\mu \frac{\partial w^f(e)}{\partial e}[w^f(e)]^{\mu-1}[w^m(e)]^{\mu} - \mu [w^f(e)]^{\mu} \frac{\partial w^m(e)}{\partial e}[w^m(e)]^{\mu-1}}{([w^f(e)]^{\mu}+[w^m(e)]^{\mu})^2} \\ &= \left(\frac{\partial w^f(e)}{\partial e} w^m(e) - \frac{\partial w^m(e)}{\partial e} w^f(e)\right) \frac{\mu[(w^f(e) w^m(e)]^{\mu-1}}{([w^f(e)]^{\mu}+[w^m(e)]^{\mu})^2}\end{aligned}$$

So we find that

$$\frac{\partial \theta^f(e^k)}{\partial e^k} = \left(\frac{\partial w^f(e^k)}{\partial e^k} w^m(e^k) - \frac{\partial w^m(e^k)}{\partial e^k} w^f(e^k)\right)\Gamma \tag{C.1}$$

with $\Gamma \equiv \mu[w^f(e^k) w^m(e^k)]^{\mu-1}/([w^f(e^k)]^{\mu}+[w^m(e^k)]^{\mu})^2$, which is always positive. Thus we have that

$$\frac{\partial \theta^f(e^k)}{\partial e^k} \geq 0 \Leftrightarrow \frac{\partial w^f(e^k)/\partial e^k}{w^f(e^k)} \geq \frac{\partial w^m(ek)/\partial e^k}{w^m(e^k)} \Leftrightarrow \epsilon^f_{w,e}(e^k) \geq \epsilon^m_{w,e}(e^k) \tag{C.2}$$

where $\epsilon^f_{w,e}(e^k)$ (resp. $\epsilon^m_{w,e}(e^k)$) denotes the education-elasticity of female (resp. male) wage. Given our assumptions on the wage functions, condition (24) is always satisfied.

### C.1.2 Proof: Solving for Optimal transfers

When married, child $k$ and his spouse choose how much to consume and how much to transfer to their respective old parents. For the sake of space, we only solve the program for child $k$ in the case when she is a female. From equations (3.3), (3.4) and (3.6), the optimization problem for the adult daughter's household is

$$\max_{c^k,c^{k'},t^k,t^{k'}} U^k(c^k,c^h,t^k,t^{k'}) = \theta^f[\ln c^k + \ln t^k] + (1-\theta^f)[\ln c^{k'} + \ln t^{k'}] \quad \text{subject to} \quad c^k + c^{k'} + t^k + t^{k'} = w^f(e^k) + w^m(e^k) + y \tag{C.3}$$

where $w^f(e^k)$ is the wage of the married daughter given an arbitrary level of education $e^k$, $w^m(e^k)$ is the wage of her husband, and $y$ is the non-labor income of the household.

First order conditions yield

$$\frac{\theta^f}{c^k} = \frac{\theta^f}{t^k} = \frac{1-\theta^f}{c^{k'}} = \frac{1-\theta^f}{t^{k'}} \tag{C.4}$$

Along with the constraint, we find the optimal transfer of the adult daughter to her old parents, conditional on her education $e^k$. We find similarly the optimal transfer of the adult child in the case when he is a male. The optimal transfer of the adult child to his old parents is thus

$$\begin{aligned} t^{d*}(e^k) &= \theta^f(e^k)[w^f(e^k) + w^m(e^k) + y]/2 \quad \text{if} \quad k = d \\ t^{s*}(e^k) &= \theta^m(e^k)[w^f(e^k) + w^m(e^k) + y]/2 \quad \text{if} \quad k = s \end{aligned} \tag{C.5}$$

## Proof of the impact of education on optimal transfers

Let us consider the derivative of the optimal transfer of the adult child $k$ with respect to his education level $e^k$ when marriage-market returns to education are taken into account. From equation (3.8), we have

$$\begin{aligned} \frac{\partial t^{d*}(e^k)}{\partial e^k} &= \underbrace{\frac{\partial \theta^f(e^k)}{\partial e^k}}_{\text{(a)}}[w^f(e^k) + w^m(e^k) + y]/2 + \frac{\theta^f(e^k)}{2}\left[\underbrace{\frac{\partial w^f(e^k)}{\partial e^k}}_{\text{(b)}} + \underbrace{\frac{\partial w^m(e^k)}{\partial e^k}}_{\text{(c)}}\right] \\ \frac{\partial t^{s*}(e^k)}{\partial e^k} &= \underbrace{\frac{\partial \theta^m(e^k)}{\partial e^k}}_{\text{(a)}}[w^f(e^k) + w^m(e^k) + y]/2 + \frac{\theta^m(e^k)}{2}\left[\underbrace{\frac{\partial w^f(e^k)}{\partial e^k}}_{\text{(c)}} + \underbrace{\frac{\partial w^m(e^k)}{\partial e^k}}_{\text{(b)}}\right] \end{aligned} \tag{C.6}$$

Equation (C.6) shows that education affects optimal transfers through three channels: (a) the marginal change in education on child $k$'s bargaining power within his household; (b) the marginal change in education on child $k$'s wage; and (c) the marginal change in education on child $k$'s spouse's wage. By assumption, we know that (b) and (c) are positive. By assumption, the sign of (a) is positive for females and negative for males. As a consequence, a marginal change in education on the optimal transfer of a daughter is unambiguously positive. In this configuration, the impact of education on the transfer of boys is less straightforward, but it is proved (below) that for $y$ not too large with respect to $w^f(e^k)$ and $w^m(e^k)$, this impact is also positive for boys.

How does the strength of this impact vary across gender? We find that $\partial t^{d*}(e^k)/\partial e^k \geq \partial t^{s*}(e^k)/\partial e^k$ if and only if

$$\underbrace{[(w^f)^{2\mu} - (w^m)^{2\mu}]}_{<0}\left(\frac{\partial w^f}{\partial e} + \frac{\partial w^m}{\partial e}\right) + \Theta\underbrace{\left(\frac{\partial w^f}{\partial e}w^m - \frac{\partial w^m}{\partial e}w^f\right)}_{\geq 0 \quad \text{or} \quad \leq 0} \geq 0 \tag{C.7}$$

with $\Theta \equiv 2\mu(w^f w^m)^{\mu-1}(w^f + w^m + y)$, which is always positive.[1] So for $\epsilon^f_{w,e}(e^k)$ high enough relative to $\epsilon^m_{w,e}(e^k)$, this impact can be stronger for daughters than for sons.

### C.1.3 Proof that the impact of education on sons' transfers is positive for $y$ not too large

For clarity reasons, we remove the subscript $k$ from equations, as well as the arguments $e$ in wage functions. We solve for the derivative if the daughter and son optimal transfers and then study the sign of their numerators.
We obtain that $\partial t^{d*}/\partial e \geq 0$ if and only if

$$w^f[(w^f)^\mu + (w^m)^\mu]\left(\frac{\partial w^f}{\partial e} + \frac{\partial w^m}{\partial e}\right) + \mu(w^m)^{\mu-1}(w^f + w^m + y)\left(\frac{\partial w^f}{\partial e}w^m - \frac{\partial w^m}{\partial e}w^f\right) \geq 0 \tag{C.8}$$

Similarly, we have that $\frac{\partial t^{s*}}{\partial e} \geq 0$ if and only if

$$w^m[(w^f)^\mu + (w^m)^\mu]\left(\frac{\partial w^f}{\partial e} + \frac{\partial w^m}{\partial e}\right) + \mu(w^f)^{\mu-1}(w^f + w^m + y)\left(\frac{\partial w^m}{\partial e}w^f - \frac{\partial w^f}{\partial e}w^m\right) \geq 0 \tag{C.9}$$

Let us assume that $\frac{\partial w^f}{\partial e}w^m \geq \frac{\partial w^m}{\partial e}w^f$, so that the impact of education on $t^{d*}$ is clearly positive. We need to show that this impact is also positive on $t^{s*}$ for $y$ not too large. We rearrange the condition described in equation (C.9) as follows. We have that $\frac{\partial t^{s*}}{\partial e} \geq 0$ if and only if

$$\begin{aligned}
&\underbrace{(1-\mu)}_{\geq 0} w^m(w^f)^\mu \frac{\partial w^f}{\partial e} + w^m(w^f)^\mu \frac{\partial w^m}{\partial e} + \underbrace{[(w^m)^{\mu+1} - \mu(w^f)^{\mu-1}(w^m)^2]}_{\geq 0} \frac{w^f}{\partial e} + (w^m)^{\mu+1}\frac{\partial w^m}{\partial e} \\
&+ \mu(w^f)^{\mu-1}(w^f + w^m)\frac{\partial w^m}{\partial e}w^f + \mu(w^f)^{\mu-1}\underbrace{\left(\frac{\partial w^m}{\partial e}w^f - \frac{\partial w^f}{\partial e}w^m\right)}_{\leq 0} y \geq 0
\end{aligned} \tag{C.10}$$

We observe in equation (C.10) that all the terms in the LHS are positive, except for the last one which depends on $y$. Thus for $y$ not too large, we have that $\frac{\partial t^{s*}}{\partial e} \geq 0$.

### C.1.4 Proof : Solving for optimal investments in education

We now solve for the optimal investment in education that parents choose in the first period in order to maximize their inter-temporal utility function. According to equations (3.6) and (3.7), parents choose $e^{k*}$ such that

$$\begin{aligned}
&\max_{e^k} \ln C + \ln[t^{d*}(e^k)] \quad \text{if} \quad k = d \\
&\max_{e^k} \ln C + \ln[t^{s*}(e^k)] \quad \text{if} \quad k = s
\end{aligned} \tag{C.11}$$

subject to the budget constraint described in equation (3.1).

[1] For clarity reasons we remove the argument $e^k$ and the subscript $k$ from wage functions in equation (C.7).

If child $k$ is a daughter, and focusing on interior solutions, the first order condition yields

$$\frac{\partial t^{d*}(e^{k*}/\partial e)}{t^{d*}(e^{k*})} = \frac{\omega + \gamma}{L + \omega - (\omega + \gamma)e^{k*}} \tag{C.12}$$

Rearranging equation (C.12), we find an expression of the costs of education $\gamma$ as a function of optimal investment in the education of a daughter:

$$\gamma = \Phi^d(e^{k*}) = \frac{[\partial t^{d*}(e^{k*})/\partial e](L + \omega)}{t^{d*}(e^{k*}) + (\partial t^{d*}(e^{k*})/\partial e)e^{k*}} - \omega \tag{C.13}$$

Similarly, we find an expression of the costs of education $\gamma$ as a function of optimal investment in the education of a son:

$$\gamma = \Phi^s(e^{k*}) = \frac{[\partial t^{s*}(e^{k*})/\partial e](L + \omega)}{t^{s*}(e^{k*}) + (\partial t^{s*}(e^{k*})/\partial e)e^{k*}} - \omega \tag{C.14}$$

## C.2 Figures from simulations

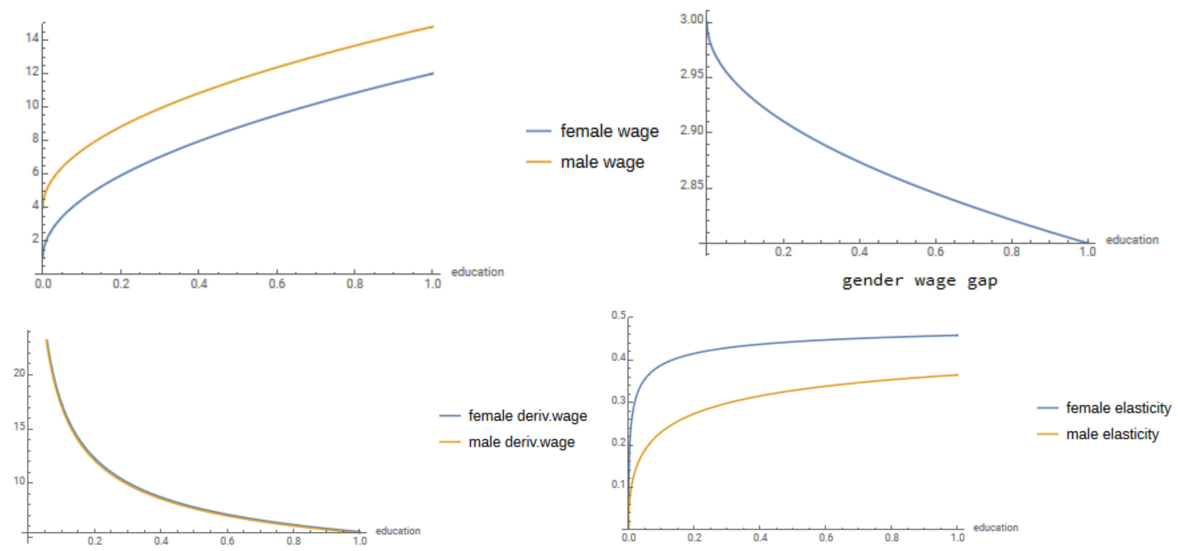



**Figure C.1** – Wages Functions

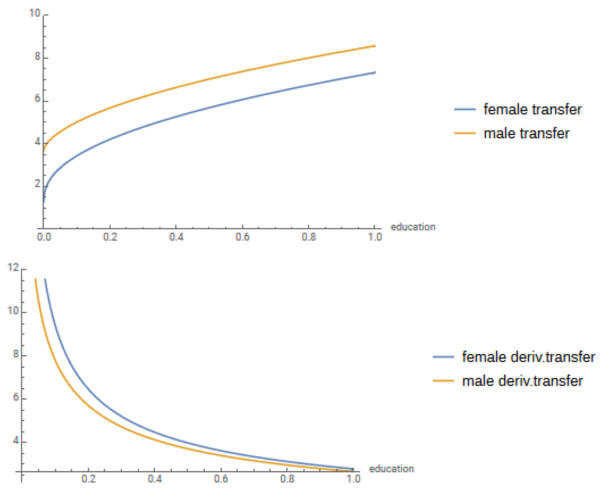



**Figure C.2** – Optimal transfers and derivaties as a function of education

# C.3 Possible theoretical extensions

## C.3.1 Monetary transfers and informal services

Assistance can take the form of monetary transfers $t^k$ and informal services $s^k$. Unlike monetary transfers, informal services are time consuming and their opportunity cost is equal to the wage per unit of time, i.e. $w^f(e^k)$ for a daughter and $w^m(e^k)$ for a son. The utility of the old parents in the second period is

$$u^p(c^p, s^p) = \ln c^p + \ln s^p \tag{C.15}$$

with $c^p = t^k$ and $s^p = s^k$.

Overall, the objective of the parents is to optimally choose $e^k$ in the first period such that it maximizes their intertemporal utility

$$U^p = \ln C + u^p(c^p, s^p) \tag{C.16}$$

under the budget constraint defined in equation (3.1).

Solving for the optimization problem of child $k$'s collective household program in the case where she is a female, we find that optimal transfer and services as functions of education are

$$\begin{aligned} t^{d*}(e^k) =& \theta^f(e^k)[w^f(e^k) + w^m(e^k) + y]/3 \\ s^{d*}(e^k) =& \frac{\theta^f(e^k)}{w^f(e^k)}[w^f(e^k) + w^m(e^k) + y]/3 \end{aligned} \tag{C.17}$$

and equivalently when child $k$ is a male.

## C.3.2 Families of two children

Each family is composed of parents, a daughter $d$ and a son $s$. We assume here that siblings are not strategic when choosing optimal transfers to their old parents. For this extension, the main difference lies in the optimization problem of the parents in the first period which now becomes

$$\max_{e^d, e^s} \ln C + \ln[t^{d*}(e^d) + t^{s*}(e^s)] \tag{C.18}$$

under the constraint $C + (\omega + \gamma)(e^d + e^s) = L + 2\omega$.

FOCS yield the following condition on optimal education investments $e^{d*}$ and $e^{s*}$: $\partial t^{d*}(e^{d*})/\partial e = \partial t^{s*}(e^{s*})/\partial e$.

Using equations (C.6) and (C.1), we find that optimal education decisions of the parents depend on the education-elasticity of female and male wages. In fact, if the education-elasticity of the female wage is low in comparison to the education-elasticity of the male wage, parents invest more in the education of their son than in the education of their daughter. Inversely, if the education-elasticity of the female wage is relatively high, now the parents invest more in the education of their daughter, even though the marginal labor-market return from the time spent in schooling is higher for their son than for their daughter. This result is in stark contrast with classical old-age support models which do not take the marriage-market returns to education. In these models, optimal education investments are such that $\partial w^f(e^{d*})/\partial e = \partial w^m(e^{s*})/\partial e$, which always amounts to the child with the higher marginal labor-market return to education being the more educated.
However, this positive effect on the daughter's education investment disappear when we allow assistance to old parents to be a combination of monetary transfers and time consuming informal services, as daughters have a lower opportunity cost of providing care than sons. Indeed, at similar education levels, daughters provide more informal services to old parents than sons.

### C.3.3 Strategic interaction of adult siblings

We consider the case when assistance to old parents is only with monetary transfers. So each adult sibling chooses an optimal monetary transfer for their old parents taking into account the action of his/her sibling, and depending on the sharing rule determined in their respective collective households. Note that this situation is equivalent to a voluntary contributions game in which each sibling contributes to the public good $c^p$ and then uses any money remaining to buy the private good $c^k$, $k = \{d, s\}$ for himself or herself. Assuming that both goods are normal, this interaction has exactly one Nash equilibrium which can take one of two forms. In the first form, both children contribute to the public good

The program of the adult daughter $d$ becomes

$$\begin{aligned} &\max_{c^d, t^d} \ln c^d + \ln c^p \\ \text{subject to} \quad & c^d + t^d = \rho^d, \quad c^p = t^d + t^s, \quad t^d \geq 0 \quad \text{and} \quad t^s \geq 0 \end{aligned} \tag{C.19}$$

with $\rho^d$ denoting the share of the joint income of her household that will be attributed to her in the bargaining process. We run the same analysis for the son $s$.
Optimal strategic transfers are:
$t^{d*} = (\rho^d - t^{s*})/2$ and $t^{s*} = (\rho^s - t^{d*})/2$, which leads to:
$t^{d*} = (2\rho^d - \rho^s)/3$ and $t^{s*} = (2\rho^s - \rho^d)/3$. So $c^p = (\rho^d + \rho^s)/3$.

In the second form, only one child contributes. For instance, if only the son contributes, then $u^d_{c^p}/u^d_{c^d} = c^{d*}/c^{p*} < 1$ and $t^{d*} = 0$ and $c^{d*} = \rho^d$.

Overall, decisions of the two siblings on transfers are the following: $t^{d*} = \min[\max(0, (2\rho^d - \rho^s)/3), \rho^d/2]$ and $t^{s*} = \min[\max(0, (2\rho^s - \rho^d)/3), \rho^s/2]$

We now derive the optimal split of the couples' income, assuming an interior solution for transfers. For tractability, we assume that the husband of the daughter and the wife of the brother act as in the benchmark model described in C.2 (they are not strategic).

The program of the household of the adult daughter is:

$$
\begin{aligned}
& \max_{\rho^d,\rho^h} \theta^f(\ln c^d + \ln c^p) + (1-\theta^f)(\ln c^h + \ln c^{hp}) \\
\text{subject to} \quad & \rho^d + \rho^h = w^f(e^d) + w^m(e^d) + y \\
& c^d = c^p = (\rho^d + \rho^s)/3, \quad c^h = t^h = \rho^h/2
\end{aligned}
\tag{C.20}
$$

Optimal sharing rule: $\rho^{d*} = \theta^f(e^d)(w^m(e^d) + w^f(e^d) + y) - [1 - \theta^f(e^d)]\rho^{s*}$
$\rho^{s*} = \theta^m(e^s)(w^m(e^s) + w^f(e^s) + y) - [1 - \theta^m(e^s)]\rho^{d*}$
Thus $\rho^{d*} = \frac{\theta^f(e^d)[w^m(e^d)+w^f(e^d)+y]-[1-\theta^f(e^d)]\theta^m(e^s)[w^m(e^s)+w^f(e^s)+y)}{1+[1-\theta^f(e^d)][1-\theta^m(e^s)]}$
$\rho^{s*} = \frac{\theta^m(e^s)[w^m(e^s)+w^f(e^s)+y]-[1-\theta^m(e^s)]\theta^f(e^d)[w^m(e^d)+w^f(e^d)+y]}{1+[1-\theta^f(e^d)][1-\theta^m(e^s)]}$
And $c^p = \frac{\theta^f(e^d)\theta^m(e^s)[w^f(e^d)+w^m(e^d)+w^f(e^s)+w^m(e^s)+2y)}{3[1+(1-\theta^f(e^d))(1-\theta^m(e^s))]}$

## C.4 Additional statistics and empirical results

**Table C.1 –** The effect of education on the log of wages, by gender

| Dependent variable | Log of wage | | | | | |
|---|---|---|---|---|---|---|
| Education | Grade | Years | Grade | Years | Grade | Years |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) |
| Education | 0.310*** | 0.094*** | 0.169*** | 0.021 | 0.276*** | 0.085*** |
| | (0.017) | (0.005) | (0.059) | (0.021) | (0.019) | (0.006) |
| Female | -0.681*** | -0.672*** | -0.694*** | -0.685*** | -0.821*** | -0.834*** |
| | (0.046) | (0.047) | (0.046) | (0.046) | (0.071) | (0.099) |
| Education$^2$ | | | 0.038*** | 0.004*** | | |
| | | | (0.014) | (0.001) | | |
| Education×Female | | | | | 0.095*** | 0.022** |
| | | | | | (0.031) | (0.010) |
| Intercept | 12.37*** | 12.16*** | 12.42*** | 12.33*** | 12.44*** | 12.23*** |
| | (0.137) | (0.138) | (0.141) | (0.154) | (0.137) | (0.138) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Obs | 3,171 | 3,171 | 3,171 | 3,171 | 3,171 | 3,171 |
| R-squared | 0.221 | 0.215 | 0.224 | 0.221 | 0.225 | 0.217 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. Controls include interactions between cohort dummies and the allocation of the water and sanitation program in the district during the 1974-1978 period. Also, we think about mean reversion by including the enrollment rate in 1971 at the district level interacted with year of birth dummies. All estimations remove individuals aged 12 to 15 in 1974 who might be partially affected by the number of primary schools constructed by the INPRES SD program. All estimates include year of birth and sub-district fixed effects.

### C.4.1 Labor-market returns to education: statistics and first-stage selection equations

### C.4.2 Effects of human capital on services provided to elderly parents

Table C.6 shows estimates on the probability of giving health-care, services in the family-farming or the family-house, during the last 12 months. Globally, neither the Education neither its interaction with the female dummy are significant. However, column 4, which reports IV-probit estimates of the effects of the child's education on the probability of giving care and services, shows a negative effect of education regardless of gender. Although the effect is identified at a 10 percent size test, the last finding suggest that additional schooling reduces non-financial assistance to elderly parents.

**Table C.2 –** Social norms on the assistance to elderly parents (statistics)

| | Freq. | Percent |
|---|---|---|
| Panel A: Elderly remain in own home. | | |
| Yes | 440 | 81.18 |
| No | 98 | 18.08 |
| Panel B: Who usually live with elderly. | | |
| Alone | 104 | 23.64 |
| Youngest daughter | 86 | 19.55 |
| Eldest daughter | 45 | 10.23 |
| Youngest son | 36 | 8.18 |
| Eldest son | 30 | 6.82 |
| Youngest child | 23 | 5.23 |
| Female child | 10 | 2.27 |
| Male child | 5 | 1.14 |
| Grandchild | 14 | 3.18 |
| Other answer | 87 | 19.77 |
| Panel C: Elderly prefer to stay with son/daughter. | | |
| Males children | 4 | 3.92 |
| Female children | 71 | 69.61 |
| Same | 20 | 19.61 |
| Panel D: Children are forced to care for elderly. | | |
| Yes | 535 | 98.71 |
| No | 5 | 0.92 |
| Panel E: Other children assist care for elderly. | | |
| Yes | 519 | 97.01 |
| No | 12 | 2.24 |
| Panel F: If the child cares for elders, he inherits more. | | |
| Yes | 180 | 33.21 |
| No | 357 | 65.87 |
| Panel G: If the child cares for elders, he inherits the house. | | |
| Yes | 235 | 43.36 |
| No | 303 | 55.90 |

Statistics are based on the answers provided in the Adat questionnaire in IFLS-2.

**Table C.3 –** Robustness of estimation methods: non-linear models

| Dependent variable | Monetary transfer (0/1) | | | | Amount (non 0) | | | | Amount (with 0) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Method | Probit | | IV Probit | | Tobit | | IV Tobit | | Tobit | | IV Tobit | |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Education | 0.056*** | 0.067*** | 0.056*** | 0.067*** | 10.41*** | 11.11*** | 12.93*** | 19.00*** | 9.766*** | 10.54*** | 11.82*** | 17.97*** |
| | (0.016) | (0.021) | (0.016) | (0.021) | (0.812) | (1.184) | (3.193) | (6.119) | (0.816) | (1.162) | (3.203) | (5.726) |
| Education × Female | | -0.022 | | -0.022 | | -1.438 | | -11.10* | | -1.606 | | -11.68* |
| | | (0.027) | | (0.027) | | (1.497) | | (6.595) | | (1.392) | | (6.681) |
| Female | -0.009 | 0.023 | -0.009 | 0.023 | -3.414** | -1.244 | -1.257 | 15.72 | -3.038** | -0.630 | -0.848 | 16.76 |
| | (0.035) | (0.051) | (0.035) | (0.051) | (1.486) | (1.557) | (1.788) | (10.56) | (1.428) | (1.652) | (1.759) | (10.40) |
| Observations | 5,370 | 5,370 | 5,370 | 5,370 | 4,004 | 4,004 | 3,093 | 3,093 | 5,370 | 5,370 | 4,120 | 4,120 |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard error in parentheses, clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, senior secondary, university. IV estimates use the number of school constructed by the Indonesian reform in the sub-district of birth, ranging from 0 to 15, as an instrument. Controls include the following district variables interacted with year of birth: the allocation of the water and sanitation program, the enrollment rate in 1971, the population aged 7-15 in 1971.

**Table C.4 –** Mean of income and education by gender and employment category

| | Males | | | | Females | | | |
|---|---|---|---|---|---|---|---|---|
| Category | Obs | Mean | Min | Max | Obs | Mean | Min | Max |
| **Panel A: Mean of Income** | | | | | | | | |
| Self-employed | 880 | 398841 | 0 | 8000000 | 601 | 232072 | 0 | 4000000 |
| Self-employed with workers | 694 | 530447 | 0 | 6600000 | 378 | 291172 | 0 | 3000000 |
| Government worker | 362 | 714589 | 0 | 6000000 | 157 | 657600 | 0 | 1500000 |
| Private worker | 1425 | 432590 | 0 | 8500000 | 573 | 225501 | 0 | 5000000 |
| **Panel B: Mean of Education** | | | | | | | | |
| No work | 3 | 6.6 | 3 | 9 | 1681 | 7.26 | 0 | 15 |
| Unpaid family worker | 50 | 6.62 | 0 | 12 | 629 | 5.46 | 0 | 15 |
| Self-employed | 896 | 6.48 | 0 | 15 | 601 | 5.93 | 0 | 15 |
| Self-employed with workers | 711 | 6.62 | 0 | 15 | 383 | 6.10 | 0 | 15 |
| Government worker | 362 | 11.96 | 1 | 16 | 155 | 12.7 | 0 | 15 |
| Private worker | 1425 | 8 | 0 | 16 | 575 | 6.8 | 0 | 15 |

**Table C.5 –** First-stage equations on the probability of working

| Dependent variable | Wage work (0/1) | | | |
|---|---|---|---|---|
| Method | Heckman Selection first-stage | | Heckman Probit | |
| Sample | All | Women | All | Women |
| Variables | (1) | (2) | (3) | (4) |
| Education | 0.190*** | -0.024 | | |
| | (0.031) | (0.016) | | |
| Education × Female | -0.088** | | | |
| | (0.036) | | | |
| Female | -1.312*** | | -2.008*** | |
| | (0.077) | | (0.067) | |
| Water Program | -0.447*** | -0.128 | -0.462*** | -0.534*** |
| | (0.160) | (0.143) | (0.143) | (0.158) |
| Enrollment in 1971 | 0.474 | 0.788* | 0.225 | 0.450 |
| | (0.480) | (0.409) | (0.444) | (0.487) |
| Children in household | -0.002 | 0.037*** | -0.042** | -0.061*** |
| | (0.020) | (0.014) | (0.020) | (0.022) |
| Inpres | | | -0.005 | -0.004 |
| | | | (0.009) | (0.010) |
| Sub-district & Cohort FE | No | No | Yes | Yes |
| Obs | 4,575 | 4,575 | 4,343 | 2,198 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district. Education represents the highest grade completed: no education, primary level, junior secondary, at least senior secondary. Controls include interactions between cohort dummies and the allocation of the water and sanitation program in the district between 1973 and 1978. Also, we think about mean reversion by including the enrollment rate in 1971 at the district level interacted with year of birth dummies. Children in household represents the number of children living in the household, which is the additional variable that is not included in the wage equation. Inpres is the number of school constructed by the INPRES SD reform at the sub-district level. In columns (3) and (4), I include sub-districts and birth year fixed effects. I remove individuals aged between 15-13 in 1973 who might be partially affected by the number of primary schools constructed by the INPRES SD program. The sample is restricted to individuals born before 1973 due to endogenous places of birth following the reform.

**Table C.6** – Effect of education on the probability of giving services to elderly parents

| Dependent variable | Services (0/1) | | | |
|---|---|---|---|---|
| Method | OLS | Probit | 2SLS | IV Probit |
| Variables | (1) | (2) | (3) | (4) |
| Education | 0.007 | 0.012 | -0.082 | -0.283* |
| | (0.005) | (0.027) | (0.085) | (0.168) |
| Education × Female | 0.007 | 0.029 | -0.025 | -0.135 |
| | (0.007) | (0.036) | (0.032) | (0.170) |
| Female | -0.001 | -0.052 | -0.008 | 0.024 |
| | (0.012) | (0.067) | (0.062) | (0.252) |
| Water Program | 0.016 | -0.576*** | 0.055 | -0.665*** |
| | (0.048) | (0.167) | (0.073) | (0.175) |
| Enrollment in 1971 | 0.050 | 2.037*** | -0.072 | 3.614*** |
| | (0.166) | (0.495) | (0.232) | (0.664) |
| Sub-district & Cohort FE | Yes | No | Yes | No |
| Number of sub-districts | 799 | – | 636 | – |
| F-stat.(KP) | | – | 10.85 | – |
| Obs | 5,307 | 5,307 | 4,236 | 4,348 |

*** p<0.01, ** p<0.05, * p<0.1. Standard error in parentheses, clustered at the sub-district level. Education represents the highest grade completed: no education, primary level, junior secondary, at least senior secondary. Controls include interactions between cohort dummies and the allocation of the water and sanitation program in the district between 1973 and 1978. Also, we think about mean reversion by including the enrollment rate in 1971 at the district level interacted with cohort dummies. Columns (3) and (4) use the number of school constructed by the INPRES SD reform at the sub-district level as an instrument. In columns (1) and (3), I include sub-districts and cohort fixed effects. Individuals born before 1973 are excluded due to endogenous place of birth following the reform. I remove individuals aged between 15-13 in 1973 who might be partially affected by the number of primary schools constructed by the INPRES SD program. Also, the sample is restricted to individuals born before 1973 due to endogenous place of birth following the reform.

**Abstract**

In the past decade, millions of children around the world have gained access to educational opportunities. However, three years after the adoption of the Sustainable Development Goals of universal primary and secondary education by 2030, there has been no progress in reducing the global number of out-of-school children. To address this issue, this thesis explores how families strategically invest in their offsprings' education, by gender. Chapter 1 introduces the broad determinants of children's human capital with a focus on family backgrounds such as mothers' endowments. In particular, the empirical study explores the relationship between women's economic rights and children's education in some developing countries. Implicit in many researches on education is the existence of interactions between family members. Accordingly, the second chapter revisits the link between income shocks and educational achievement by considering the role of sibling composition in a rural region of Tanzania. The study shows that children suffer an additional penalty during income shocks the larger the share of girls among (younger) siblings. Finally, families might decide to underinvest in children's education, mostly in their daughters, if they expect that they will not be able to obtain the returns for this education. Thus, the last chapter assesses theoretically and empirically the intergenerational parent-child exchange in Indonesia. We find that a substantial fraction of human capital gains for both girls and boys, generated by exposure to an educational reform, is shared with the parental generation. We show that education positively affects old-age transfers through additional labor and marital market returns for both men and women.



**Keywords:** education; gender; women's economic rights; income shocks; heterogeneity; sibling composition; inter-generational relationship; old-age insurance; labor-market; marriage-market; microeconomic data; empirical analysis; development



**Résumé**

Au cours de la dernière décennie, les chances en matière d'accès à l'éducation ont augmenté considérablement pour des millions d'enfants. Cependant, trois ans après l'adoption des Objectifs de Développement Durable pour une éducation primaire et secondaire universelle à l'horizon 2030, les progrès en matière de décrochage scolaire restent insuffisants. Pour tenter de répondre à ce défi majeur, cette thèse explore comment les familles investissent stratégiquement dans l'éducation de leurs enfants, selon le genre. Le chapitre 1 introduit les principaux déterminants du capital humain en insistant sur les caractéristiques familiales comme les ressources appartenant à la mère. Cette étude empirique examine en particulier la relation entre les droits économiques des femmes et l'éducation des enfants dans plusieurs pays en développement. La plupart des recherches en éducation négligent cependant le rôle des autres membres de la famille dans les choix d'éducation. Le deuxième chapitre revisite alors le lien entre chocs de revenu et éducation en prenant en compte le rôle de la composition des frères et soeurs dans une région rurale de Tanzanie. L'étude montre que le choc négatif affecte davantage les enfants ayant relativement plus de (jeunes) soeurs que de (jeunes) frères. Enfin, les parents peuvent décider de sous-investir dans l'éducation de leurs enfants, notamment des filles, s'ils anticipent qu'ils ne toucheront pas les fruits de leur investissement. Ainsi, le dernier chapitre évalue de façon théorique et empirique les échanges intergénérationnels entre parents-enfants en Indonésie. Nous trouvons qu'une partie des gains d'une hausse du niveau d'éducation à la fois pour les filles et les garçons, générée par une réforme de l'éducation primaire en Indonésie, est partagée avec la génération des parents. Nous montrons que l'éducation affecte positivement les transfers envoyés aux parents à travers davantage de bénéfices sur le marché du travail et sur le marché du mariage.



**Mots-Clés:** Education; genre; droits économiques des femmes; chocs de revenus; hétérogénéité; composition de la fratrie; relation intergénérationnelle; assurance vieillesse; marché du travail; marché du mariage; données microéconomiques; analyses empiriques; développement