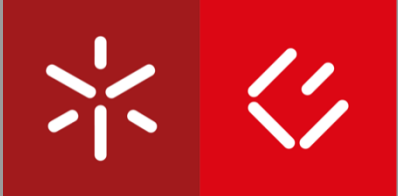




Universidade do Minho
Escola de Economia e Gestão

Emmanuel Temene Ngoy

Essays on Health and Education in
Developing Countries





Universidade do Minho

Escola de Economia e Gestão

Emmanuel Temene Ngoy

Essays on Health and Education in Developing Countries

Doctoral Thesis

Ph.D. in Economics

Work developed under the supervision of:

Professor Carla Sá

and

Professor Paula Veiga

November 2024

Direitos de Autor e Condições de Utilização do Trabalho por Terceiros

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos. Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada. Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição -NãoComercial-SemDerivações

CC-BY-NC-ND

<https://creativecommons.org/licenses/by-nc/4.0/>

Acknowledgements

I wish to extend my sincere gratitude to several people who have provided greater inputs in the execution of this thesis. My special thanks go to my supervisors, Carla Sá and Paula Veiga, for their outstanding mentorship, patience, availability, and valuable advice throughout my PhD journey.

I also wish to acknowledge the support of faculty members and colleagues at the School of Economics and Management of the University of Minho for kindly sharing their knowledge with me. I am particularly thankful to Professor Miguel for his insightful econometric advice, which significantly aided this research.

Furthermore, I am tremendously grateful to my family, without whom this thesis would not have been possible. I express my deepest appreciation to my parents, Robert Ngoy and Noela Kavira, as well as my siblings, Joel Ngoy, Henock Ngoy, Rachel Ngoy, and Christian Ngoy, and my grandmother Leontine Karumb, for their unwavering love, care, moral support, encouragement, and prayers. I am also grateful to friends who made my time in Portugal more enriching, particularly Adebayo Adebisi, Pedro Cadeia, Peter Ayogu, Philip Dennis, Leornado Tadeu, Gilmar Conceição and Songe Ngoma.

Statement of Integrity

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Ensaio sobre Saúde e Educação nos Países em Desenvolvimento

Resumo

Esta tese inclui três ensaios sobre saúde e educação no contexto de países em desenvolvimento. O primeiro ensaio (Capítulo 2) documenta a magnitude da desigualdade socioeconômica nas competências linguísticas em crianças em idades pré-escolar e escolar no Peru. Os resultados indicam a existência de desigualdade socioeconômica nas competências linguísticas que se manifestam desde idades precoces, ainda antes do início da escolaridade formal. O estatuto socioeconômico dos pais é o principal fator explicativo da desigualdade socioeconômica observada. Uma análise mais aprofundada revela que tanto as alterações na desigualdade dos agregados familiares como a elasticidade associada contribuem para a alteração da desigualdade socioeconômica nas competências linguísticas das crianças, no período observado.

O segundo ensaio (Capítulo 3) investiga a relação entre a exposição a conflitos armados durante a infância e os níveis de educação em idade adulta. Os resultados indicam não existir uma relação estatisticamente significativa entre a exposição durante a infância e a probabilidade de concluir o ensino básico e o ensino secundário inferior. Esta conclusão sugere resiliência das comunidades num contexto de conflito prolongado.

Finalmente, o terceiro ensaio (Capítulo 4) estuda o impacto na saúde da criança da exposição pré-natal à seca. Os resultados revelam um efeito prejudicial da exposição, particularmente no que respeita à altura da criança, sobretudo, quando ocorre nos primeiro e últimos trimestres da gravidez. Os cuidados pré-natais surgem como um fator protetor dos riscos climáticos. Estes resultados sugerem a importância de realizar investimentos na proteção contra os choques climáticos, sobretudo em contextos de baixos rendimentos.

Palavras-chave: Capital humano, Desigualdade, Educação, Países em desenvolvimento, Saúde.

Essays on Health and Education in Developing Countries

Abstract

This thesis encompasses three essays on health and education in the context of developing countries. The first essay (Chapter 2) documents the magnitude of socioeconomic-related inequality in language skills in preschool and school-age children in Peru. The results indicate the existence of socioeconomic-related inequality in language skills that manifests itself from an early age, even before the start of formal schooling. Parental socioeconomic status is the main factor explaining the observed socioeconomic-related inequality. Further analysis reveals that both changes in household wealth inequality and the associated elasticity contribute to the change in socioeconomic-related inequality in children's language skills over the observed period.

The second essay (Chapter 3) investigates the relationship between exposure to armed conflict during childhood and schooling outcomes in adulthood. The results indicate no statistically significant relationship between exposure to conflict during childhood and the likelihood of completing primary and lower secondary education. This finding suggests communities resilience to conflict, particularly in the context of protracted armed conflict.

Finally, the third essay (Chapter 4) studies the impact of prenatal exposure to drought on children's health. The results reveal a detrimental effect of prenatal exposure to drought in child health, particularly regarding child height. The effects are stronger during the first and final trimesters of pregnancy. Lower crop yields emerge as the key channel for this negative effect. Prenatal care appears to be a protective factor against climate risks. These findings shed light on the importance of investment in adaptation as a buffer against weather shocks, particularly in low-income settings.

Keywords: Developing countries, Education, Health, Human capital, Inequality.

Contents

1	General Introduction	1
2	Exploring socioeconomic-related inequality in children's cognitive achievement	5
2.1	Introduction	5
2.2	The Peruvian education context	9
2.3	Empirical Methods	11
2.3.1	Measurement and decomposition of socioeconomic-related inequality	11
2.3.2	Estimation of the cognitive achievement production	14
2.3.3	Empirical strategy	15
2.4	Data	16
2.5	Results	19
2.5.1	Socioeconomic-related inequalities in child language performance	19
2.5.2	Cognitive achievement regression	23

2.5.3	Decomposing socioeconomic-related inequalities in child language performance	26
2.5.4	Decomposing change in socioeconomic-related inequality in child language performance	29
2.6	Robustness checks	31
2.6.1	Alternative measures of socioeconomic status	31
2.6.2	Sample restriction	32
2.7	Conclusions and Discussion	33
3	Armed conflict and schooling outcomes in the Congo	35
3.1	Introduction	35
3.2	Background	38
3.2.1	Conflict context	38
3.2.2	Education context	39
3.3	Literature Review	40
3.4	Empirical Approach	43
3.4.1	Data	43
3.4.2	The exposure to conflict intensity index	44
3.4.3	Empirical strategies	45
3.5	Results	50

3.5.1	Armed conflict during childhood and educational outcomes	50
3.5.2	Heterogeneity analysis	51
3.5.3	Mechanisms: Migration selection	53
3.6	Robustness checks	54
3.6.1	Placebo analysis	54
3.6.2	Alternative conflict dataset	55
3.6.3	Alternative specifications	55
3.7	Conclusions	57
4	Prenatal exposure to drought and child health outcomes in Zambia	59
4.1	Introduction	59
4.2	Literature Review: Weather shocks and health outcomes	62
4.3	Data	64
4.3.1	Demographic and Health Survey	64
4.3.2	Weather data: Standard Precipitation Evapotranspiration Index	65
4.3.3	Crop Yields data	67
4.4	Empirical Approach	70
4.5	Results	71
4.5.1	The effect of drought in utero on child health	71

4.5.2	Mechanisms	74
4.6	Robustness and Heterogeneity analysis	80
4.6.1	Robustness checks	80
4.6.2	Results from the heterogeneity analysis	82
4.7	Conclusions	83
5	General Conclusion	85
	Bibliography	88
	Appendix A: Additional tables and figures to chapter 2	109
	Appendix B: Additional tables and figures to chapter 3	119
	Appendix C: Additional tables and figures to chapter 4	133

List of Tables

2.1	Descriptive statistics	18
2.2	Socioeconomic-related inequality in language achievement	22
2.3	Differences in socioeconomic-related inequality in language achievement between preschool and school ages.	22
2.4	Language achievement regression estimates	25
2.5	Decomposition of socioeconomic-related inequality in PPVT scores	28
2.6	Decomposition of socioeconomic-related inequality in PPVT scores at age 15 for the VAM	29
2.7	Changes in socioeconomic-related inequality in PPVT scores between ages 5 and 15	31
3.1	Descriptive statistics	48
3.2	Conflict exposure and Schooling outcomes: Probit marginal effects	52
4.1	Descriptive statistics	68
4.2	The effect of drought in utero on child health outcomes: OLS estimation results	73
4.3	The effect of drought on crop yields: OLS estimation results	75

4.4	The effect of drought in utero on parental investments: OLS estimation results	77
4.5	The effect of drought in utero on child health outcomes- Mitigation factors	79
A1	Concentration index computed with household expenditure per capita	109
A2	Concentration index computed with Parental education	110
A3	Difference in Socioeconomic-related inequality in language achievement between preschool and school ages	110
A4	Differences in Socioeconomic-related inequality in language achievement between preschool and school ages	111
A5	Decomposition socioeconomic-related inequality in PPVT scores–Household expenditures .	112
A6	Decomposition of socioeconomic-related inequality in PPVT scores–Parental education . .	113
A7	Decomposition of value added–with PPVT scores age 8	114
A8	Socioeconomic-related inequality in language achievement with Quechua speaking student	114
A9	Decomposition of socioeconomic-related inequality in PPVT scores with quechua speaking	115
A10	Socioeconomic-related inequality in language achievement with non-attributing children . . .	116
A11	Decomposition of socioeconomic-related inequality in PPVT scores with non-attributing children	117
A12	Correlation between wealth index and other measures of socioeconomic status	118
B1	Violent conflict exposure and schooling outcomes	120
B2	Non-violent conflict exposure and schooling outcomes	121
B3	Conflict exposure and schooling outcomes –gender	122

B4	Conflict exposure and schooling outcomes – location of residence	123
B5	Conflict exposure and probability of migrating	124
B6	Conflict exposure and schooling outcomes–Placebo cohorts	125
B7	Conflict exposure and schooling outcomes–Alternative dataset	126
B8	Conflict exposure and schooling outcomes–Alternative dataset	127
B9	Conflict exposure and Primary schooling outcomes–Alternative specifications	128
B10	Conflict exposure and Lower secondary schooling outcomes–Alternative specifications . .	129
B11	Conflict exposure and Schooling outcomes: LPM	130
C1	The effect of drought in utero on child health outcomes- Alternative specifications	134
C2	The effect of drought in utero on child health outcomes- Alternative definitions of drought .	135
C3	The effect of drought in utero on child health outcomes- Alternative definitions of drought during growing period	135
C4	The effect of drought in utero on child health outcomes-severe drought	136
C5	The effect of lifetime drought on child health outcomes	136
C6	The effect of drought in utero on child health outcomes-different month scales	137
C7	The effect of drought in utero on child health outcomes–Alternative definitions	138
C8	The effect of drought in utero on child health outcomes-Whole sample	138
C9	The effect of drought in utero on mortality and birth interval	139
C10	The effect of drought in utero on child health outcomes–heterogeneity analysis	140

List of Figures

2.1	Socioeconomic inequality in PPVT score performance across all age range	21
3.1	Distribution of conflicts across localities between 1997 and 2014	49
4.1	Number of droughts per district 2001-2018	67
4.2	Percentage of monthly droughts across DHS clusters	69
B1	Conditional marginal effect of completing primary schooling	131
B2	Conditional marginal effect of completing lower secondary schooling	132
C1	DHS clusters at the district level	133

Chapter 1

General Introduction

Health and education are the most critical dimensions of human capital. These dimensions of human capital interact in many ways. For example, good health at an earlier age fosters cognitive development, enabling healthier children to reap a greater benefit from schooling. Moreover, better health implies higher school attendance and educational attainment. There is also an overwhelming evidence on the positive effects of education on health and huge discussions on the mechanisms ([Furnée et al., 2008](#)). This factor makes health and education stronger predictors of an individual's productivity, especially when acquired at the early age. Investment in health and education contributes to economic welfare through a skilled and healthy labor force. In this respect, health and education are strong markers of social development and nations' wealth. Globally, human capital in the form of health and education investment, constitutes the largest component of a nation's wealth, with nearly 70% contribution, particularly from high income countries ([Lange et al., 2018](#)). Moreover, the long-lasting benefits of early health and education inputs are observable across multiple socioeconomic dimensions, including labor market outcomes ([Black et al., 2007](#); [Currie and Almond, 2011](#); [Gertler et al., 2014](#); [Almond et al., 2018](#); [Bütikofer et al., 2019](#); [Rossin-Slater and Wüst, 2020](#); [García et al., 2023](#)).

Improving education and health is thus a primary development goal in its own right. The adoption of the Millennium Development Goals (MDGs) in 2000 and thereafter, the Sustainable Development Goals (SDGs) in 2015 by the United Nations are a strong marker of an international consensus concerning the importance of

health and education, as instrumental for human development ([United Nations, 2024a,b](#)). 5 of the 8 MDGs focus on improving education and health as primary development goals¹. Similarly, the SDGs 3 and 4 entail promoting quality education and good health for all. In this respect, education and health emerge as essential components for the attainment of the other 15 goals. In the last decades, several countries have witnessed progress on some health and education indicators, particularly life expectancy and school enrollments. Life expectancy has witnessed a substantial rise, from 53 to 71 years old in low-and middle-income countries (LMIC) between 1970 and 2018 ([World Bank, 2024](#)). Likewise, primary school enrollments are near universal, with a rate of about 98% in LMIC ([World Bank, 2024](#)). This considerable progress in education and health in developing countries signals a strong policy emphasis on these issues among governments and international organizations.

Yet significant parts of the world still suffer from insufficient education and health outcomes, particularly in developing countries. Approximately two-thirds of the world's school-age children lack basic skills, with the largest share in Sub-Saharan Africa (SSA), accounting for about 94% ([Gust et al., 2024](#)). Likewise, child mortality remains a pressing challenge in the developing world. About 79 children per 1000 live births die before reaching their fifth birthday in SSA ([World Bank, 2024](#)). Moreover, malnutrition during childhood is prevalent, with about 25% stunting rate among under-five children across LMIC ([World Bank, 2024](#)). Poor health and education are the main drivers of poverty and the transmission of disadvantages across generations. Addressing poor health and education is therefore, the same as combating poverty.

This thesis consists of three independent chapters, each of which explores factors that affect health and education outcomes in developing countries. Chapter 2 examines the extent of socioeconomic-related inequality in language skills between preschool and school-age children in Peru. Peru is a country with a higher enrollment rate but poor learning outcomes. The main contribution of this chapter is employing a different approach to inequality in education outcomes through the concentration index and its decomposition techniques—a state-of-the-art in the health economics literature ([Van Doorslaer et al., 1997](#); [Kakwani et al., 1997](#)). The findings reveal that socioeconomic-related inequality in language skills, favoring the better-off, emerges from an earlier age and persists throughout adolescence, albeit dropping. Parental socioeconomic

¹The five goals related to health and education include: eradicating hunger, achieving universal primary education, reducing child mortality, improving maternal health and combating infectious diseases, such as malaria.

status is the main contributor to the observed socioeconomic-related inequality in language skills. The contribution stems from the influence of household wealth with respect to child language skills and household wealth inequality itself. This chapter is forthcoming in the *Journal of Economic Inequality*².

Chapter 3 investigates the relationship between childhood exposure to armed conflicts and schooling outcomes in the Democratic Republic of Congo (DR Congo). The DR Congo has experienced decades of violent conflicts since the 1990s, with higher death tolls and an increased number of internally displaced persons (Coghlan et al., 2007; IDMC, 2022). The study exploits variation in conflict exposure across birth cohorts and the DHS clusters. We combine the geo-coded information from the Demographic and Health Survey (DHS) with the Armed Conflict Location Event dataset (ACLED) to construct a measure of armed conflict exposure at the DHS cluster level, which we refer as the village level. The key finding shows no statistically significant relationship between armed conflict exposure during childhood and the probability of completing primary and lower secondary schooling. These findings suggest a community resilient to conflict, especially in the context of prolonged armed conflict. Moreover, the results indicate that the relationship between conflict exposure and schooling outcomes is context specific, in line with existing evidence (Verwimp and Van Bavel, 2014). For instance, evidence regarding conflict-education nexus is mixed in the literature, and depends on different factors, such as the level of conflict intensity and institutional context of the country³ (Unfried and Kis-Katos, 2023).

Chapter 4 investigates the effect of in utero exposure to drought on child health outcomes in Zambia. Zambia is a country prone to frequent drought, which often results in recurrent crop failure and food insecurity. Additionally, malnutrition remains a pressing challenge, with nearly 34% of stunted children (World Bank, 2024). This chapter merges birth histories from the Zambian DHS with weather information from the global Standardized Precipitation Evapotranspiration Index (SPEI) data. The empirical strategy exploits monthly variations in drought episodes within survey villages. The findings indicate that drought exposure while in utero results in poor health outcomes in early childhood, particularly in terms of height. The effects are homogeneous across child gender and location of residence. A lower crop yield is a key channel through which prenatal drought exposure affects child health outcomes. The result further suggests that prenatal

²The paper is available at: <https://link.springer.com/article/10.1007/s10888-023-09615-4>

³See section 3.2.3 for more details.

care use by a qualified healthcare professional may mitigate the adverse effect of drought on child health outcomes. This finding points to the importance of investment in adaptation, as a potential smoothing factor against adverse weather conditions. This chapter is under review in the *World Development*, since February 2024.

Finally, chapter five outlines the main findings, policy implications and limitations associated with each study.

Chapter 2

Exploring socioeconomic-related inequality in children's cognitive achievement

2.1 Introduction

Socioeconomic inequality is a significant risk factor, as it has negative effects on many outcomes such as health and education outcomes. Across low-and middle-income countries, millions of young children at the bottom of the income distribution are at heightened risk of poor physical and cognitive development ([Grantham-McGregor et al., 2007](#)). Poor cognitive development may translate into lower educational attainment, fewer employment opportunities and lower labor market earnings in adulthood ([Case and Paxson, 2008](#); [Almond and Currie, 2011a](#); [Hanushek et al., 2015](#)). These disadvantages constitute a substantial cost to society, reducing the overall economic welfare and social mobility, hence perpetuating inequities across generations ([Hanushek and Woessmann, 2008](#); [Corak, 2013](#); [Neidhöfer, 2019](#)).

The World Bank's 2018 Development Report, with an emphasis on the learning crisis, highlights the implications of non-school factors, including socioeconomic status, on the education system itself ([World Bank, 2018](#)). Peru is no exception in this regard. Indeed, despite a higher school enrollment than the Latin

American and Caribbean (LAC) countries average, Peruvian children from disadvantaged families perform worse than their peers from less disadvantaged socioeconomic backgrounds (OECD, 2016). As a highly unequal country with a 0.42 Gini coefficient, among the highest in the Latin American region (World Bank, 2021), Peru could attribute a large portion of poor learning outcomes to significant inequality of opportunities, making socioeconomic status an important factor for student achievement. It is estimated that socioeconomic background differences account for more than 20% of variations in PISA (Programme of International Student Assessment) mathematics results, the highest share among LAC countries (OECD, 2016), which suggests an unequal education system. In this context, it is important to understand how human capital accumulation varies by household socioeconomic status.

This paper explores the extent of socioeconomic-related inequality in language skills between preschool and school-age children, as well as the factors explaining such inequalities using longitudinal data from the Peruvian Young Lives survey. Such analysis is important because understanding how these inequalities emerge and evolve from a very young age is crucial for public policies aimed at reducing income and educational inequality observed later in adulthood. The question resonates with the core principle of equal opportunities, by which children's outcomes should not be based on the circumstances in which they have little or no control, namely their parents' socioeconomic status (Roemer and Trannoy, 2016).

Moreover, early cognitive skills are an important determinant of labor market outcomes (Heckman et al., 2006; Sahn and Villa, 2016; Glewwe et al., 2017). On the one hand, recent evidence suggests that numeracy skills are a stronger predictor of wages than literacy skills (Hanushek et al., 2015). But, on the other hand, language skills are more related to home factors than any other cognitive skills. Studies have consistently found that children with more advantaged socioeconomic status (SES) appear to better develop linguistic skills before starting formal schooling, as their families may provide a more stimulating home environment for their children than less disadvantaged ones (Fernald et al., 2013; Weisleder and Fernald, 2013; Pace et al., 2017; Von Stumm et al., 2020). Focusing on language skills is particularly relevant to understanding the impact of family background inequalities on child development.

Since the seminal work of Coleman et al. (1966), the strong association between parental socioeconomic status and their children's educational outcomes has been a well-established fact in the literature (see Björk-

lund and Salvanes, 2011, for a review). Most of the studies focus on how home factors affect children's outcomes, especially at a young age (Cunha et al., 2006; Todd and Wolpin, 2007). Other studies suggest that a more disadvantaged family environment may have a more detrimental effect on young boys than girls' outcomes (Autor et al., 2019, 2023). Early age is a critical period for human capital development when children are particularly malleable to skills acquisition (Cunha and Heckman, 2008, 2009). A growing literature from low-and middle-income countries shows that cognitive outcomes from differences in SES emerge before a child's exposure to formal school, with evidence from children younger than 2 years of age in Peru, Senegal, Indonesia, and India (Fernald et al., 2012), preschool-age children in Ecuador (Paxson and Schady, 2007), Colombia (Rubio-Codina et al., 2015) and Madagascar (Fernald et al., 2011). In the same vein, Cunha and Heckman (2007) show that socioeconomic gap in cognitive skills appears early and remain unchanged during the early stage of schooling in the US.

While there is consistent evidence of a widening SES gap in language ability in early childhood age (Paxson and Schady, 2007; Fernald et al., 2011; Schady et al., 2015; Behrman et al., 2017), the evidence regarding socioeconomic inequality's pattern in language ability during the middle childhood age is mixed. For instance, evidence from longitudinal studies shows that socioeconomic differences in child cognition flatten after the age of 8 in Madagascar (Galasso et al., 2019) and narrow between ages 5 and 12 in Peru and Vietnam, as opposed to the remaining countries from the multi-country Young Lives longitudinal data⁴, Ethiopia and India, where there is no such pattern (Reynolds et al., 2017). In another study, Abufhele et al. (2022) report a persistent socioeconomic gap in children's language performance, with a decreasing magnitude between the ages 2 and 12 in Chile. A mixed evidence emerges also from studies outside developing countries, particularly among primary and secondary student cohorts in the US. An earlier study reports an increasing income-based academic achievement among cohorts born between 1972 and 2001 (Reardon, 2011). On the other hand, Hanushek et al. (2019) show that SES achievement gap among US primary and secondary school age student cohorts has remained unchanged in the past four decades. Recent studies show a decreasing SES gap in academic achievement for the same cohort (Hanushek et al., 2022; Hashim et al., 2023; Nielsen, 2023). However, these studies offer no clear explanation for such a pattern.

⁴See section 4 for further explanation on Young Lives data.

This work extends the existing literature in several ways. First, we provide new evidence on the evolution of the socioeconomic-related inequality in child language skills over an extensive age range, looking not only at early childhood but also at the adolescent years. Indeed, most studies rely on cross-sectional data, and the limited studies on longitudinal data typically focus on early childhood. Second, the previous literature has largely focused on raw test scores as outcomes, which does not take into account that test items may vary in their degree of difficulty. Our work overcomes this limitation by using a common measure of test scores over time, generated through the item response theory (IRT). The IRT allows for the evolution of a child's language ability across different ages while accounting for the difficulty level of the item. Third, we take a different approach to the analysis of inequality in education outcomes by relying on the concentration index and its decomposition, which is commonly used in the health economic literature to examine income-related health inequality (Van Doorslaer et al., 1997; Kakwani et al., 1997). Notable exceptions are Martins and Veiga (2010) and Bouhlila (2017) who apply the same approach in the context of European and Middle East and North Africa (MENA) countries, respectively. Unlike this paper, both studies focus on cross-sectional data and secondary school-aged students. A straightforward approach to assessing socioeconomic-related inequality in language achievement would be an analysis of variation in mean test score achievement across wealth quintiles, for instance. Although suitable, such an approach only provides a half picture of how the language achievement variable may vary across the entire socioeconomic status distribution and cannot fully capture the degree of socioeconomic-related inequality when the socioeconomic variable is continuous (Wagstaff et al., 2007).

Our approach allows us to examine socioeconomic-related inequality in test scores not only across the entire socioeconomic distribution but also in an extensive way. First, we estimate the level of socioeconomic-related inequality in language achievement for children aged 5 years, 8 years, 12 years, and 15 years using a longitudinal sample from Peru. Then, we decompose socioeconomic-related inequalities into several contributing factors for each age. Finally, we explore the changes in the contributing factors of socioeconomic-related inequality in language skills between ages 5 and 15, a pattern under-explored in most existing studies. For each decomposition exercise, the contributing factors are split into elasticity and socioeconomic inequality components. This distinction comes in handy as it may offer insights into public policy, whether the focus should be directed on the elasticity or the socioeconomic inequality component.

We find evidence of socioeconomic-related inequality in the Peabody Picture Vocabulary Test (PPVT) score starting at an early age, with a relatively higher level in preschool than the adolescence age. The regression estimates from language achievement suggest the relevance of family socioeconomic status on child performance in PPVT scores. Similarly, our decomposition analysis strengthens this finding, showing that family socioeconomic status explains a large fraction of the socioeconomic-related inequality in language achievement. After accounting for past achievements, the relative contribution of parental socioeconomic status diminishes but still plays a substantial role. The persistent effect of skills emerges as a major contributor in explaining socioeconomic-related inequality at age 15, consistent with the self-productivity nature of skills, meaning that skills produced at one stage augment the acquisition skills at a later stage as in [Cunha and Heckman \(2007\)](#). Our decomposition analysis also documents that, in most cases, changes in the elasticity of the determinants rather than changes in socioeconomic inequalities in the determinants account for changes in socioeconomic-related inequality in language achievement between ages 5 and 15. However, household wealth shows a different pattern, with changes in household wealth elasticity and wealth inequality explaining differences in socioeconomic inequality PPVT score. The findings are robust to alternative measures of socioeconomic status, namely household expenditures and parental education.

The remainder of the chapter is organized as follows. Section 2.2 provides background on the education context in Peru. The empirical approach is provided in section 2.3. Section 2.4 describes the data. Sections 2.5 and 2.6 present the results, and robustness checks, respectively. Section 2.7 concludes the study.

2.2 The Peruvian education context

Peruvian basic education consists of pre-primary, primary and secondary education. Pre-primary education is provided by two types of institutions: *Centros de Educación Inicial* (CEIS) and *Programas no Escolarizados de Educación Inicial* (PRONOEI). The former is a formal early childhood program for children aged 3 to 5, while the latter is a non-formal community center that typically serves children of the same age in rural and marginal urban areas. Primary education lasts six years and is for children aged 6 to 11. Finally, secondary education is a five-year program for children between 12 and 16 years of age. These levels are mandatory,

as per the General Law of Education 2003 (*Ley General de Educación*).

After decades of remarkable economic performance, the Peruvian education sector has seen considerable progress, with high levels of coverage. Today, more children are enrolled in school, with almost no significant regional and gender differences, indicating greater access to basic education for the marginalized part of the population. The net enrollment⁵ is near universal level, with more than 90% of primary school age students enrolled, and access to preschool, as well as secondary education, have witnessed a substantial participation rate well above the LAC countries average, with a gross enrollment ratio⁶ of around 89.2% and 97.6%, respectively, in 2015, compared to about 74.6% and 94.6% in LAC, for the same level of schooling (UIS, 2022). Similarly, the gross enrollment ratio for pre-primary education is above the LAC average, corresponding to about 89.2 % against 74.7 % in 2015 (UIS, 2022).

Despite this notable improvement, learning outcomes remain relatively poor. According to the Programme of International Student Assessment (PISA) results, Peru was ranked 64th out of 70 countries present in the assessment of educational quality in 2015, making the 15-year-old student group one of the worst performers among the participating countries (OECD, 2016). A similar pattern emerges from the primary school level, with only 27% of students showing satisfactory skills in basic quantitative reasoning from a national student assessment designed for second-grade students (INEI, 2016), reflecting a low level of quality in the education sector.

⁵The student's enrollment of the official age group for a given level of education expressed as a percentage of the population in that age group.

⁶The total enrollment of students in a specific level of education, regardless of their age, expressed as the percentage of the official school-age corresponding to this level of education.

2.3 Empirical Methods

2.3.1 Measurement and decomposition of socioeconomic-related inequality

Our approach to measuring socioeconomic-related inequality in language achievement relies on the concentration index (CI), a standard measure of inequality in the health economics literature ([Kakwani et al., 1997](#); [Van Doorslaer et al., 1997](#)).

Graphically, the concentration curve represents the cumulative share of the outcome variable of interest (e.g. test score) against the cumulative proportion of individuals, ranked by their socioeconomic level, from the poorest to the richest. The concentration index is then defined as twice the area between the concentration curve and the diagonal ([Kakwani et al., 1997](#)), where the diagonal represents the line of equality. If the curve lies above the line of equality, the concentration index is negative, then socioeconomic inequality favors the disadvantaged groups, and when below, socioeconomic inequality is in favor of the rich.

Consider a language test score measure A_i , with $i = 1, 2, \dots, n$ and assign a rank r_i to each child i primarily based on their relative position across the socioeconomic status distribution, with the child from the poorest families ranked first, and that of the richest families ranked last. The concentration index (CI) of language achievement (A_i) can be represented, analogously to [Erreygers and Van Ourti \(2011\)](#), as in equation (2.1):

$$CI = \frac{2}{n^2 \bar{A}} \sum_{i=1}^n R_i A_i \quad (2.1)$$

where A_i denotes the language achievement outcome for child i , $\bar{A}$ is the mean of the language achievement and R_i is the fractional rank of child i , defined as $R_i = r_i - \frac{n+1}{2}$, where r_i is the rank of child i in the socioeconomic distribution. The concentration index is bounded within the interval $[-1, 1]$, with a negative (positive) value, indicating a pro-poor (rich) inequality, that is, the achievement in the test score favors the children from the worst-off (better-off) families. If the concentration index is null, there are no socioeconomic-

related inequalities in language achievement.

Furthermore, from equation (2.1), we can compute the gini coefficient (G) for language achievement (A_i) by using the distribution of achievement test scores as the ranking variable instead of the SES measures to capture overall inequality in language achievement A_i . Then, the relative importance of socioeconomic-related inequality in language achievement in relation to the overall inequality in language achievement can be expressed as the ratio CI/G .

The CI can be estimated through a convenient regression following [Kakwani et al. \(1997\)](#):

$$2\sigma_R^2 \left(\frac{A_i}{\bar{A}} \right) = \rho + \rho_1 R_i + v_i \quad (2.2)$$

where $2\sigma_R^2$ is the variance of the fractional rank R_i , and ρ is the constant term. The coefficient of the rank ρ_1 is an estimate of the CI, which is equivalent to the value obtained from equation (2.1)⁷. v_i is the error term.

The concentration index, as defined above, simply indicates the magnitude of socioeconomic-related inequality, and can be decomposed into the main contributing factors using regression-based decomposition methods, as proposed by [Wagstaff et al. \(2003\)](#).

Let us assume a cognitive production function given by the following linear regression model

$$A_i = \alpha + \sum_k \beta_k x_{ki} + \epsilon_i \quad (2.3)$$

where α is the intercept term, x_{ki} is a vector of the determinants of cognitive production for child i and ϵ_i is the error term. The estimated β_k will be used in the decomposition equations (2.4)-(2.6) to capture the contribution of individual factors x_k to socioeconomic-related inequality in language achievement, in which each contribution is the product of the elasticity of language achievement A_i with respect to factor x_{ki} and the degree of socioeconomic-related inequality in factor x_{ki} .

Following [Wagstaff et al. \(2003\)](#), substituting equation (2.3) into (2.1) yields

⁷See [Kakwani et al. \(1997\)](#) for more details.

$$CI = \sum_{k=1}^K \beta_k \left(\frac{2}{n^2 \bar{A}} \sum_{i=1}^n x_{ki} R_i \right) + \frac{2}{n^2 \bar{A}} \sum_{i=1}^n \epsilon_i R_i \quad (2.4)$$

$$CI = \sum_{k=1}^K \left(\frac{\beta_k}{\bar{A}} \bar{x}_k \right) c_k + \frac{GC\epsilon}{\bar{A}} \quad (2.5)^8$$

where $\frac{\beta_k}{\bar{A}} \bar{x}_k$ captures the elasticity of A_i with respect to x_k evaluated at the sample mean ($\bar{x}_k$ and $\bar{A}$) and c_k is the concentration index for the explanatory variable x_k (defined analogously to CI in equation (2.1)). β_k are the regression coefficients from equation (2.3) to be estimated. x_k is the k -th determinant of child language achievements as defined in equation (2.3). $\bar{x}_k$ is the mean of the x_k variable. $GC\epsilon$, the generalized concentration index for the error term (ϵ_i) as defined in equation (2.3), is computed as $GC\epsilon = \frac{2}{n^2} \sum_{i=1}^n \epsilon_i R_i$ (the second term of equation (2.4)). It represents the residual component of CI decomposition, referring to inequality in child language achievement that cannot be explained by systematic variation across socioeconomic groups in the variables (x_k). Equation (2.5) corresponds to the decomposition of the CI by age.

Complementary to a decomposition analysis by age, is decomposing changes in the socioeconomic-related inequalities between preschool (age 5) and school-age period (age 15). [Wagstaff et al. \(2003\)](#) further suggested an approach to exploring changes in the concentration index over time by applying an Oaxaca decomposition type to equation (2.5)

$$\Delta CI_t = \sum_{i=1}^K \eta_{kt} (c_{kt} - c_{kt-1}) + \sum_{i=1}^K c_{kt-1} (\eta_{kt} - \eta_{kt-1}) + \Delta \frac{GC\epsilon}{\bar{A}_t} \quad (2.6)$$

where $\Delta CI_t = CI_t - CI_{t-1}$ denotes a change in the concentration index, and η_{kt} is the elasticity of the language achievement A_i with respect to a contributing factor (x_k) at age t ($t = 5, 8, 12, 15$), defined as $\eta_{kt} = \frac{\beta_{kt}}{\bar{A}_t} \bar{x}_{kt}$, as in equation (2.5). The first term of equation (2.6) reflects the extent to which changes in inequalities in language achievement are driven by changes in inequality in the determinants of language

⁸Given that $c_k = \frac{2}{n^2 \bar{x}_k} \sum_{i=1}^n R_i x_{ki}$, then $\sum_{i=1}^n R_i x_{ki} = \frac{c_k}{2} n^2 (\bar{x}_k)$. Substituting this expression in equation (2.4) gives equation (2.5).

achievement and the second term of equation refers to the extent to which changes in inequality in language achievement are driven by changes in the elasticities of the determinants of language achievement x_k .

2.3.2 Estimation of the cognitive achievement production

We consider a cognitive production function, in which a child's achievement in the test score depends on the entire history of the family and school inputs, as well as the child's own ability (Todd and Wolpin, 2003, 2007). The estimation strategy relies on two alternative specifications. First, we consider a contemporaneous specification, relating the current achievement to current inputs. The model is estimated separately for each age, as follows:

$$A_i = \alpha + \sum_k \beta_k x_{ki} + \epsilon_i \quad (2.7)$$

where A_i is the language achievement of the child i ; x_{ki} is a set of explanatory variables that are the determinants of cognitive achievement, including family inputs, school inputs as well as child's specific characteristics; β_k are the coefficients, and ϵ_i is the error term containing unobserved and omitted determinants of achievement. The β_k estimated through equation (2.7) will be used in the decomposition, as in equations (2.4)-(2.6), to capture how each determinant separately contributes to socioeconomic-related inequality in language achievement. The model is estimated by Ordinary Least Squares (OLS) separately for different ages, namely ages 5, 8, 12 and 15.

However, OLS imposes strong assumptions and estimates are likely to be biased due to unobserved and omitted variables. For instance, the unobserved ability may affect the performance of children in test achievement. The common way to deal with this caveat is to estimate a value-added specification model, in a way that features the cumulative process of skill acquisition, with the lagged test scores minimizing the omitted bias (Todd and Wolpin, 2003, 2007). The model for panel data is expressed as follows:

$$A_{i,t} = \delta + \gamma A_{i,t-1} + \sum_k \theta_k x_{ki,t} + \mu_{i,t} \quad (2.8)$$

where γ is the persistence coefficient; θ_k includes the input coefficients; δ is the constant term and $\mu_{i,t}$ is the error term. The γ and θ_k estimated through equation (2.8) will be used in the decomposition, as in equations (2.4)-(2.6). The lagged test scores serve as sufficient statistics for the past inputs and unobserved child abilities. While the Generalized Method of Moments (GMM) is often recommended for dynamic panels, particularly for addressing potential endogeneity in lagged dependent variables, we chose to rely on an OLS-based estimation approach under the value-added specification. This decision is based on findings from previous studies that suggest OLS can provide robust and consistent estimates in dynamic settings under certain conditions (Guarino et al., 2015). Specifically, the inclusion of prior achievement controls for unobserved heterogeneity and selection bias related to the grouping and assignment mechanisms, thereby mitigating the risk of omitted variable bias (Andrabi et al., 2016; Singh, 2015). Additionally, recent simulations comparing dynamic model estimators show that, in some cases, OLS is more resilient than GMM to structural assumptions and is less sensitive to potential weak instrument problems that can affect the reliability of GMM (Guarino et al., 2015). Given these considerations, and the consistency of our results with existing empirical literature, we believe the OLS approach adequately captures the relationships in our model while offering greater interpretability and simplicity.

2.3.3 Empirical strategy

The empirical strategy employed in this paper involves the following step and analysis: First, we explore graphically the degree of socioeconomic-related inequality in the child's language skills across age groups through the concentration curve, and we estimate the CI, as in equation (2.2) by age group. We also test for the differences in CI in language performance between preschool and school ages, following the Chow test of equality of coefficients (Chow, 1960).

Second, we estimate the cognitive achievement production function (equations (2.7) and (2.8)), which is used in the decomposition of socioeconomic-related inequalities in child language performance to capture the contribution of each determinant in CI, as in equation (2.5). Each contribution is the product of elasticity and degree of socioeconomic inequality in the variable itself. Finally, we go further in the decomposition analysis

by decomposing the change over time in socioeconomic-related inequality in child language performance (Equation (2.6)).

2.4 Data

This study employs data from the Peruvian Young Lives⁹ sample, a longitudinal survey which initially followed two cohorts of children: 1000 children born in 1994-95 and 2000 children born in 2001-02. The survey comprises five waves, undertaken between 2002 and 2016. This paper focuses only on the younger cohort, for whom a common measure of achievement is available for all the survey rounds. The sample is nationally representative, excluding only 5% of the richest part of the country. The attrition rate is relatively low, with about 90% of the original sample still in the subsequent rounds.

The survey contains rich information on family socioeconomic backgrounds, as well as child development measures, such as cognitive assessment tests. We use the Peabody Picture Vocabulary Test (PPVT), as a measure of language skills. It was first given to the younger cohort children in the second round, and the children were asked to select the correct picture that best matches the word presented orally to them by the examiner.

The test contains 125 items and was administered in Spanish and some indigenous languages of Peru, such as Quechua. Each student answers to a different subset of those items, which requires an appropriate technique to compute each student's test score. The test score is then generated using the item response theory (IRT)¹⁰, an approach widely used in international comparison assessment tests, such as Trends in International Mathematics and Science Study (TIMSS). The key advantage of IRT is not only does it generate a comparable test score across time for a common item but also captures the probability of a child answering a given item correctly based on the difficulty level. As the test item cannot be equated across different languages due to the effect of cultural variation, only the tests administered in Spanish were generated using

⁹Young Lives is a longitudinal study of child poverty, which tracked two cohorts of children in four low-and middle-income countries: Ethiopia, India (only the state of Andhra Pradesh), Peru and Vietnam.

¹⁰See [Das and Zajonc \(2010\)](#) for further details. The PPVT IRT is based on the maximum likelihood estimate, generated using the OpenIRT command in Stata, written by [Das and Zajonc \(2010\)](#).

IRT. Accordingly, we restrict our sample to individuals who took the test in Spanish, which represents about 90 percent of the sample. We use the IRT to generate a test score based on the 125 items, normalized to have a mean of 500 and a standard deviation of 100. We have performed robustness checks using the whole sample and the results are similar to that of the restricted sample.

Our measure of socioeconomic status is the household wealth index, which is a simple average of three sub-indices: the housing quality index, the access to basic service index and the consumer durable index. The housing quality sub-index is a simple average of the quality of materials used in the construction of the family home such as floor, wall, and roof. The access to service sub-index is computed as a simple average of access to basic services such as electricity, safe drinking water, sanitation, and adequate fuel for cooking. The consumer durable sub-index is the simple average of a set of dummy variables, indicating whether households own common consumer durables such as mobile phones, refrigerators, televisions, automobiles, motorcycles, and so on.

While there is no consensus on how to measure socioeconomic status, an asset-based wealth index is thought to capture the long-term household economic status in a similar way to consumption expenditures ([Filmer and Pritchett, 2001](#)). Likewise, parental education is closely related to permanent income as it is stable over time ([Schulz, 2005](#)). Furthermore, it has been shown that permanent income has a stronger effect on child development than current income ([Blau, 1999](#)). Despite these similarities, the ranking of households within the distribution may differ as a result of a specific component of SES ([Filmer and Scott, 2008](#)). The household expenditure adjusted by household size and parental education will be used for robustness checks.

There are other variables in the data that enter the model specification. Its definition is provided in detail in Table 2.1.

Table 2.1: Descriptive statistics

	Definition (1)	Age 5 (2)	Age 8 (3)	Age 12 (4)	Age 15 (5)	Pooled (6)
PPVT	Peabody Picture Vocabulary test score standardized mean=500 and SD=100	500.0 (100.00)	500.0 (100.0)	500.0 (100.00)	500.0 (100.00)	500.0 (100.00)
Male	=1 if child is male, otherwise 0	0.502	0.506	0.504	0.507	0.505
First born	=1 if child is the firstborn, otherwise 0	0.383	0.368	0.367	0.363	0.370
Child stunted	=1 if child is stunted, otherwise 0	0.287	0.187	0.180	0.156	0.202
Child attended preschool	=1 if child ever attended preschool , otherwise 0	0.936	0.933	0.928	0.928	0.931
Mother age	Age of mother in years	31.01 (6.662)	33.09 (6.734)	37.81 (6.906)	40.79 (6.799)	35.92 (7.732)
Mother schooling	Years of education of the mother	7.791 (4.365)	7.647 (4.435)	7.676 (4.531)	7.971 (4.660)	7.772 (4.501)
Father schooling	Years of education of the father	8.828 (4.072)	8.585 (4.047)	8.734 (4.139)	9.155 (4.102)	8.830 (4.101)
Parental average education (years)	Average years of mother and father schooling	8.228 (3.866)	8.075 (3.936)	8.129 (3.990)	8.453 (4.026)	8.222 (3.958)
Both parents alive	=1 if both parents are alive	0.986	0.982	0.996	0.996	0.990
Household size	Number of persons in the household	5.422 (2.075)	5.378 (1.888)	5.219 (1.181)	5.235 (1.887)	5.312 (1.919)
Small household size	=1 if the household size if below the average, otherwise 0	0.608	0.614	0.647	0.639	0.627
Head of household female	=1 if the head of household is female, otherwise 0	0.108	0.132	0.172	0.191	0.151
Wealth index	household wealth based on simple average of housing quality, access to basic services and consumer durables	0.496 (0.228)	0.556 (0.206)	0.597 (0.193)	0.633 (0.177)	0.571 (0.208)
Expenditures	Total monthly expenditures per capita in real 2006 Soles	196.3 (162.1)	211.1 (153.2)	302.4 (312.0)	309.7 (801.3)	255.1 (445.0)
Urban	=1 if child lives in urban area, otherwise 0	0.768	0.757	0.748	0.755	0.757
N		1742	1805	1828	1811	7186

Notes: PPVT measures the language skill and is internally standardized with mean=500 and Standard deviation (SD)=100. SD are in parentheses.
Source: Peruvian Young Lives survey; own computations.

2.5 Results

2.5.1 Socioeconomic-related inequalities in child language performance

We begin our analysis by exploring the degree of socioeconomic-related inequality in the child's language skills across ages. Figure 2.1 shows the concentration curve for different ages. We plot the cumulative proportion of children's performance in PPVT scores across the wealth distribution ranking. The diagonal captures the line of equality. The concentration curve lies below the diagonal for each age range, indicating that socioeconomic inequality in child performance in PPVT favors those from better-off families. However, socioeconomic inequality in test score performance seems to be more pronounced in earlier than later ages.

Table 2.2 shows the estimation results for the concentration index as given by equation (2.2), which is estimated separately for each age group. The concentration index is positive and statistically significant across all age groups, implying the existence of socioeconomic inequality in language achievement, favoring children from well-off families. However, we observe a relatively high level of socioeconomic-related inequality at age 5 with an estimate of 0.069 (Table 2.2(1)). This finding suggests that the socioeconomic gradient emerges before the child's entrance into formal school. On the other hand, we can notice little change in socioeconomic-related inequality in test score performance during the school years, with a lower CI size at the adolescent age, as shown by the CI of 0.054 in the last row of column 1 of Table 2.2. Such a pattern may imply that socioeconomic-related inequality in language achievement changes less as children get older, albeit dropping.

On average, socioeconomic-related inequality in language achievement accounts for more than 50% of the total inequality in language achievement, with about 67% at age 5 (Table 2.2(3)). Moreover, we can notice that overall inequality in language achievement is quite stable across ages, with a Gini coefficient of about 0.10 (Table 2(2)).

Although Table 2.2 provides some insight into socioeconomic-related inequality in children's language performance by age groups, it does not indicate whether differences in the CI across different ages are

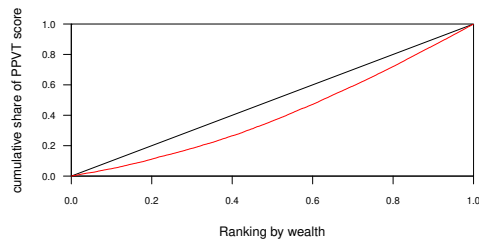
statistically significant. To provide more information, we conduct a Chow test of the coefficient equality (Chow, 1960). The test for the difference in the CI in Table 2.3 indicates a statistically significant reduction in socioeconomic-related inequality in language performance between the ages of 5 and 15, although differences are small (-0.015). However, we do not find any significant difference in the CI between the ages of 5 and 8.

This finding is similar to previous studies from both the developing countries and the US, showing that the socioeconomic gap in language appears very early and remains unchanged during the early stage of schooling (Cunha and Heckman, 2007; Rubio-Codina et al., 2015; Schady et al., 2015). In addition, we observe a decreasing socioeconomic-related inequality in language performance at age 15. This finding complement previous studies that typically focus on younger aged children (Reynolds et al., 2017). Family characteristics may impact more language performance during the early age, which could explain why socioeconomic-related inequality in language performance remains unchanged between the ages of 5 and 8. As children progress through late childhood, the impact of home factors on language development tends to become stable and other factors, such as school may come into play. This might decrease, to some extent socioeconomic-related inequality in language performance during late childhood. Additionally, a decline in socioeconomic inequality itself may reduce socioeconomic-related inequality in language performance over time. This pattern suggests that the influence of family inputs in socioeconomic-related inequality in children's language performance may have different effects at different stages of childhood.

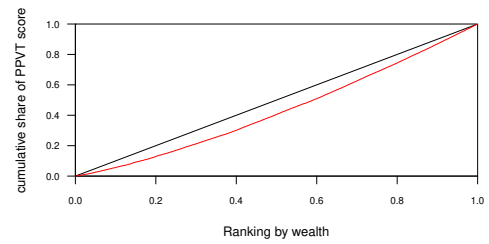
Moreover, our findings could be relevant in other settings, particularly in developing countries where a greater majority of children are exposed to poor home environments during the early ages. In such a setting, socioeconomic-related inequality in language performance is more likely to emerge strongly from an early age. However, a declining pattern in socioeconomic-related inequality in language performance observed in this study could be different from other developing countries, as there is some decline in socioeconomic inequality in Peru over time, despite being high (World Bank, 2021)¹¹. Galasso et al. (2019) shows that socioeconomic differences in children's language ability emerge at ages 3-4 years and flatten after the age of 8 in Madagascar. In Chile, Abufhele et al. (2022) report a persistent socioeconomic gap in children's language performance, starting from an early age, with a decreasing magnitude between the ages of 2 and 12.

¹¹The gini coefficient has decreased from around 0.54 to 0.43 between 2002 and 2016, which corresponds to the observed period in this study (World Bank, 2021).

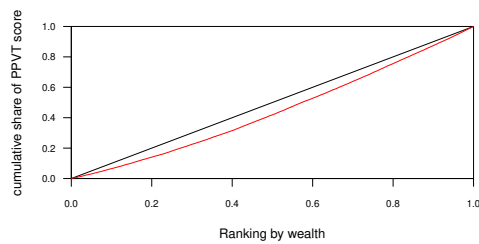
Overall, the results suggest that children from lower socioeconomic groups enter formal school with a significantly lower level of language skills than their peers from higher socioeconomic groups, and these disadvantages change less throughout the school years, albeit declining. Note that, we have performed robustness checks with a balanced panel and the results are similar.



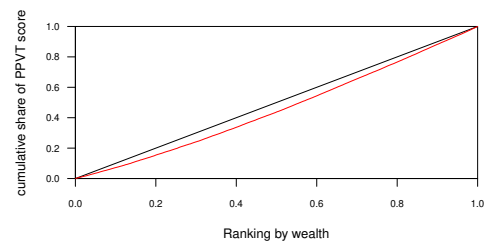
(a) concentration curve age 5



(b) concentration curve age 8



(c) concentration curve age 12



(d) concentration curve age 15

Figure 2.1: Socioeconomic inequality in PPVT score performance across all age range

Table 2.2: Socioeconomic-related inequality in language achievement

	<i>CI</i> (1)	G (2)	<i>CI/G</i> (3)
Age 5	0.069*** (0.002)	0.103*** (0.001)	0.670
Age 8	0.067*** (0.002)	0.107*** (0.001)	0.626
Age 12	0.062*** (0.002)	0.106*** (0.001)	0.585
Age 15	0.054*** (0.002)	0.107*** (0.001)	0.505

Notes: CI is estimated based on equation (2.2), that is, the point estimates for the CI and the standard errors are obtained using a convenient regression approach (Kakwani et al., 1997). Robust standard errors in parenthesis. Significance levels: *, 10%; **, 5%; ***, 1%. The language skill is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is the household wealth. G is the overall inequality in language achievement, and it is based on equation(2) with PPVT score as the ranking variable. Source: Peruvian Young Lives survey; own computations.

Table 2.3: Differences in socioeconomic-related inequality in language achievement between preschool and school ages.

	Differences in <i>CI</i>
Age 8 and Age 5	-0.001 (0.003)
Age 12 and Age 5	-0.007*** (0.003)
Age 15 and Age 5	-0.015*** (0.003)

Notes: CI is estimated based on equation (2.2), that is, the point estimates for the CI and the standard errors are obtained using a convenient regression approach (Kakwani et al., 1997). Differences in CI are tested by means of a Chow test (Chow, 1960) of equality of coefficients. Robust standard errors in parenthesis. Significance levels: *, 10%; **, 5%; ***, 1%. The language skill is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is household wealth. Source: Peruvian Young Lives survey; own computations.

2.5.2 Cognitive achievement regression

The estimation results for the cognitive achievement production (equations (2.7) and (2.8)) are provided in Table 2.4. Each column reports the estimate for the respective ages of children. columns (1)-(4) refer to the contemporaneous specification, while the value-added specification is shown in column (5). The estimates do not differ much in terms of significance and sign but the size of the coefficient drops significantly after controlling for the lagged test scores.

The results suggest a positive association between household wealth and child performance in language achievement across all ages. The strongest association, however, is observed in column 2, with household wealth associated with a 2.12 SD increase in language test scores performance at age 8. With the inclusion of the lagged test score, the size of the household wealth coefficient diminishes to 0.75 SD (Table 2.4(5)). It may be the case that important omitted factors, including school characteristics, may overestimate the association between household wealth and test scores in the contemporaneous specifications. On the other hand, 1 SD increase in the lagged achievement implies a 0.39 SD increase in test scores performance at age 15 (Table 2.4(5)). The result shows that lagged achievement, reflecting the accumulative effect of skills, is an important predictor of achievement at age 15.

There seems not to be gender differences in the test score performance at age 5 (Table 2.4(1)), and yet boys tend to perform better in the PPVT than girls during the school years. The coefficient from the contemporaneous specification shows a higher magnitude than the one obtained from column 5. The test performance is 0.18 SD higher for boys (Table 2.4(5)).

As expected, the performance in the test scores is lower for children with poor nutrition. However, it is not clear whether this association gets stronger with age. Nonetheless, the lowest association is observed in the value-added model, where the test performance is 0.19 SD lower for children who are stunted (Table 2.4(5)).

While the contemporaneous specification (Table 2.4 (1)-(4)) seems to suggest evidence of an association between preschool participation and performance in the PPVT, the significance disappears after controlling for the lagged test scores at age 5 (Table 2.4(5)). Preschool participation appears to have a stronger association

during the earlier age, as we would expect. The performance in the PPVT score is 0.37 SD higher for children who attended preschool education (Table 2.4(1)).

We can also note important differences in test performance between children living in urban and rural areas, with a larger magnitude at age 5. For instance, Table 2.4(1) shows that the performance in the test scores is 0.41 SD higher for those in urban areas.

Although the results yield useful information on the association between child test score performance and socioeconomic status, they must be interpreted with caution as they do not capture causality. This finding suggests the relevance of socioeconomic status for child performance in test score achievement.

Table 2.4: Language achievement regression estimates

	Contemporaneous specification				VAM
	Age 5 (1)	Age 8 (2)	Age 12 (3)	Age 15 (4)	
Wealth index	183.919*** (10.997)	212.770*** (12.402)	195.411*** (11.770)	181.579*** (13.536)	75.470*** (14.938)
Male	2.620 (3.706)	7.577** (3.830)	19.270*** (3.660)	19.266*** (4.151)	18.026*** (4.147)
Child stunted	-26.017*** (5.078)	-23.717*** (5.670)	-26.741*** (4.815)	-25.136*** (4.852)	-19.711*** (4.992)
First born	10.238** (4.615)	15.306*** (4.871)	15.180*** (4.910)	19.510*** (5.358)	11.339** (5.349)
Child attended preschool	37.864*** (9.947)	20.372** (10.138)	23.007*** (6.354)	21.506*** (6.604)	8.200 (7.543)
Mother age	0.957*** (0.339)	0.236 (0.332)	0.768** (0.316)	0.829** (0.334)	0.424 (0.339)
Small household size	6.948* (4.003)	15.802*** (4.565)	21.027*** (4.481)	14.391*** (4.536)	11.224** (4.656)
Head of household female	3.247 (6.174)	6.634 (4.834)	3.274 (5.777)	5.635 (5.565)	-2.259 (5.467)
Urban	41.294*** (6.784)	33.807*** (5.958)	33.724*** (4.731)	33.677*** (5.795)	9.228 (6.475)
Both parents alive	9.340 (17.880)	33.188* (18.185)	23.623 (25.040)	-3.617 (41.755)	-14.289 (41.235)
Lagged PPVT (age 5)					0.392*** (0.031)
<i>N</i>	1679	1694	1737	1722	1531
<i>R</i> ²	0.421	0.394	0.372	0.270	0.317

Notes: The model is estimated by OLS using equations (2.7) and (2.8). Robust standard errors in parenthesis (clustered at the child level). Significance levels: *, 10%; **, 5%; ***, 1%. The dependent variable is PPVT standardized with a mean of 500 and SD 100. The socioeconomic status measure is the household wealth VAM stands for the value-added model. Source: Peruvian Young Lives survey; own computations.

2.5.3 Decomposing socioeconomic-related inequalities in child language performance

We now turn to the decomposition of various factors associated with child language achievement. Table 2.5 provides the sources of the socioeconomic-related inequality in children's performance in the PPVT for the respective age, as given by equation (2.5).

By and large, household wealth explains more than half of the socioeconomic-related inequality in child performance across all age groups, ranging from 68.1% to 73.5% (Table 2.5). For each age, the contribution stems from the combined effect of the concentration index c_k and elasticity effects η_k . The relatively larger size of the elasticity effect (0.23) at age 8 (Table 2.5(4)) may point to a strong positive association between household wealth and a child's test performance in line with the regression estimates in Table 2.4. This means that children from wealthy families tend to perform better in test scores performance.

The socioeconomic disparities between rural and urban areas account for substantial socioeconomic-related inequality in language achievement, contributing to around 16% of the socioeconomic-related inequality in test score performance (Table 2.5(3)). After age 8, we observe a widening in the socioeconomic inequalities between rural and urban areas, favorable to urban residents, which explains the potential increase to around 18.5% in the size of the contribution to socioeconomic-related inequality in test scores performance at age 15 (Table 2.5(12)).

There is a great deal of socioeconomic inequality in poor nutrition that is disproportionately concentrated among the worse-off children, as given by the c_k , ranging from 0.35 to 0.32 (Table 2.5). However, the small association between a child who is stunted and test scores reduces the overall contribution, making it modest in explaining socioeconomic-related inequality in test score performance, ranging from 4.7% to 7.4% (Table 2.5). Living in a small household, being the first born and participating in preschool play a little role in the observed socioeconomic-related inequality in language achievement, contributing less than 5% of this inequality (Table 2.5).

Overall, the decomposition results suggest the importance of the combination of wealth inequality in Peru

with family socioeconomic background in explaining socioeconomic-related inequality in the language ability between preschool and school-age (from age 5 to age 15).

However, the results change when accounting for lagged achievement at age 5. Table 2.6 displays the decomposition of the socioeconomic inequality in language performance at age 15. After accounting for the lagged achievement, the relative contribution of household wealth diminishes from 0.68 to 0.32 (compare column 3 of Tables 2.5 and 2.6). Nonetheless, it still plays a non-negligible role in explaining socioeconomic-related inequality in child language performance at age 15. The high socioeconomic inequality (0.14) rather than the elasticity of language achievement with respect to family wealth (0.09) is the most important observable contributor to the socioeconomic-related inequality in test scores (Table 2.6).

The past achievement in language performance emerges as a major contributor to the observed socioeconomic-related inequality in language achievement at age 15, with a contribution of 53.4% (Table 2.6(3)). This finding is consistent even when using the test scores at age 8 (Table A7 in the Appendix A). On the other hand, the results seem to suggest that schools reproduce the existing pattern of socioeconomic inequality in test scores, favoring the well-off since children from higher socioeconomic groups are more likely to achieve better scores than those from lower socioeconomic groups. This is in line with previous evidence showing that disadvantaged children learn less in school than their peers in Peru ([Glewwe et al., 2017](#)).

Table 2.5: Decomposition of socioeconomic-related inequality in PPVT scores

	Age 5			Age 8			Age 12			Age 15		
	η_k (1)	c_k (2)	cont (3)	η_k (4)	c_k (5)	cont (6)	η_k (7)	c_k (8)	cont (9)	η_k (10)	c_k (11)	cont (12)
Wealth index	0.1827	0.2622	0.6980	0.2368	0.2091	0.7348	0.2334	0.1813	0.6814	0.2300	0.1587	0.6814
Male	0.0026	0.0093	0.0004	0.0077	-0.0033	-0.0004	0.0193	0.0107	0.0033	0.0195	0.0067	0.0024
Child stunted	-0.0015	-0.3488	0.0743	-0.0087	-0.3853	0.0500	-0.0095	-0.4029	0.0614	-0.0078	-0.3203	0.0468
First born	0.0078	0.1148	0.0131	0.0109	0.1116	0.0181	0.0109	0.1150	0.022	0.0141	0.1064	0.0281
Child attended preschool	0.0707	0.0274	0.0282	0.0381	0.0320	0.0181	0.0426	0.0293	0.0201	0.0401	0.0289	0.0216
Mother age	0.0592	0.0049	0.0042	0.0160	0.0037	0.0009	0.0578	-0.0014	-0.0013	0.0677	-0.0014	-0.0017
Small household size	0.0085	0.0674	0.0083	0.0193	0.0974	0.0279	0.0271	0.0817	0.0356	0.0182	0.0549	0.0187
Head of household female	0.0007	0.0446	0.0005	0.0017	0.0519	0.0013	0.0101	-0.0214	-0.0004	0.0021	-0.0244	-0.0010
Urban	0.0636	0.1818	0.1683	0.0513	0.1845	0.1404	0.0507	0.1913	0.1563	0.0507	0.1945	0.1851
Both parents alive	0.0185	0.0021	0.0006	0.0654	0.0048	0.0046	0.0469	-0.0008	-0.0006	-0.0072	-0.0002	0.000

Notes: The decomposition of the CI is computed based on Equation (2.5). η_k = elasticity defined as $\frac{\partial_k}{A}(\bar{x}_k)$; c_k = concentration index for x_k ; cont = contribution to CI. The outcome variable is language achievement, measured by PPVT with mean 500 and 100 SD. Source: Peruvian Young Lives survey; Own computations.

Table 2.6: Decomposition of socioeconomic-related inequality in PPVT scores at age 15 for the VAM

	η_k (1)	c_k (2)	Cont (3)
Wealth index	0.0967	0.1459	0.3180
Lagged PPVT (age 5)	0.3849	0.0615	0.5337
Male	0.0178	0.0099	0.0040
First born	0.0085	0.0892	0.0171
Child stunted	-0.0053	-0.3054	0.0366
Child attended preschool	0.0152	0.0236	0.0081
Mother age	0.0339	0.0007	0.0005
Small household size	0.0142	0.0460	0.0148
Head of household female	-0.0008	-0.0548	0.0010
Urban	0.0146	0.1495	0.0492
Both parents alive	-0.0279	-0.0001	0.0000

Notes: The decomposition of CI using equation (2.5), accounting for the lagged PPVT at age 5.

PPVT measures the language achievement and is normalized with mean 500 and SD 100.

η_k =elasticity defined as $\frac{\beta_k}{A}(x_k)$, c_k = the concentration index for x_k . cont= contribution to CI .

Source: Peruvian Young Lives survey; own computations.

2.5.4 Decomposing change in socioeconomic-related inequality in child language performance

Simply relying on the decomposition by age does not allow for exploring the source of changes in socioeconomic-related inequality over time. Table 2.7 presents the sources of changes in the socioeconomic-related inequality in language achievement between preschool and school age, as given by the Oaxaca-like decomposition (equation (2.6)). The first component, Δc_k , represents the variation in socioeconomic inequality attributable to the change in the concentration index for each factor x_k , given the constant elasticity, and the second component, $\Delta \eta_k$, is the change in elasticity, reflecting the change in the relative importance of each factor x_k in explaining changes in the distribution of test score performance in the language achievement. The final column, which is a combination of these two components, indicates how much change in the concentration index can be explained by changes in each factor x_k .

In most cases, changes in elasticity appear far more important in explaining the change in socioeconomic-

related inequality associated with performance in language achievement between the preschool and school years. For preschool participation, for instance, the change in socioeconomic-related inequality is negligible (0.0001), and it is attributable to changes in the relative importance of participation in preschool education on test scores performance (Table 2.7(2)).

Changes in household wealth emerge as an important driver of the observed change in socioeconomic-related inequality in language performance between the ages of 5 and 15, with changes in household wealth inequality and its associated change in elasticity contributing almost equally to the observed socioeconomic-related inequality in language performance, as shown by 0.02 and 0.01, respectively (Table 2.7(1)-(2)). Furthermore, we can notice a decline in the overall wealth inequality as shown in Table 2.7(1). This decreased in inequality coincide with higher level of wealth, as shown by a wealth index of 0.63 at age 15 (see Table 2.1(5)).

Furthermore, the change in elasticity with respect to stunted children appears to explain the change in socioeconomic-related inequality in test score performance between preschool and school age, as given by 0.0024 (Table 2.7(2)). Such a result seems to point out that policy should be directed at affecting the elasticity rather than the socioeconomic inequality component itself. An example could be policies that affect the nutritional status of children, such as school feeding programs.

Table 2.7: Changes in socioeconomic-related inequality in PPVT scores between ages 5 and 15

	Δc_k (1)	$\Delta \eta_k$ (2)	total (3)
Wealth index	-0.0238	0.0124	-0.0114
Male	-0.0001	0.0002	0.0001
First born	-0.0001	0.0007	0.0006
Child stunted	-0.0002	-0.0024	-0.0026
Child attended preschool	0.0001	-0.0008	-0.0008
Mother's age	-0.0004	0.0000	-0.0004
Small household size	-0.0002	0.0007	0.0004
Head of household female	-0.0001	0.0001	-0.0001
Urban	0.0007	-0.0023	-0.0016
Both parents alive	0.0000	-0.0001	-0.0000

Notes: The change in decomposition to CI is computed based on Equation (2.6).

PPVT measures the language achievement and is normalized with mean 500 and SD 100. Δc_k = change in c_k ; $\Delta \eta_k$ =change in elasticity.

$$\eta_k = \frac{\beta_k}{A}(\bar{x}_k).$$

Source: Peruvian Young Lives survey; own computations.

2.6 Robustness checks

2.6.1 Alternative measures of socioeconomic status

Although there is no single way to measure socioeconomic status, there is some general agreement that components of socioeconomic status such as parental education, household expenditures, and wealth better reflect the household's permanent income and have a strong association with child development outcomes (Brooks-Gunn and Duncan, 1997; Blau, 1999; Filmer and Pritchett, 2001). Furthermore, proxies of socioeconomic status are closely related (See Table A12 in the Appendix A). In this section, we assess the degree to which our finding is sensitive to the use of parental education and household expenditures per capita as measures of SES instead of the wealth index. The results are reported in the Appendix A.

Tables A1 and A2 provide the results of the concentration index using household expenditure and parental education, respectively. The findings show that the estimate of the concentration index is slightly similar to

the baseline results obtained from the household wealth as a measure of socioeconomic status (Table 2.2). The only difference is that the pattern of socioeconomic-related inequality, as measured by the CI is relatively stable when using parental education instead of household wealth. Nevertheless, the findings point to the existence of socioeconomic inequality in language performance in favor of children from better-off families. A similar pattern emerges from the decomposition analyses, which exhibit a closer finding to the baseline result from Table 2.5, as shown in the Appendix A (Tables A5 and A6).

Tables A3 and A4 present the test for the difference in the CI based on the Chow test of equality of coefficients ([Chow, 1960](#)) using household expenditures and parental education, respectively. The results are consistent with those in Table 2.3. The exception is that we do not find any significant difference in the concentration indices when using parental education for all the age groups (see Table A4).

Overall, the alternative measures of socioeconomic status have little influence on the main findings of this analysis.

2.6.2 Sample restriction

The restriction of the sample to children who took the PPVT in Spanish may introduce sample selection bias, and therefore affect the estimated results of our study. We check whether our findings are robust to the inclusion of children who took the test assessment in the Quechua language. The results are provided in the Appendix A.

Tables A8 and A9 show that the inclusion in the sample of children who took the PPVT in the Quechua language does not affect our main results. In table A8, the concentration index presents a similar feature of socioeconomic-related inequality in test score performance, with little change across all age ranges as shown in Table 2.2. Furthermore, the decomposition results in Table A8 display a very similar result to that of the baseline result in Table 2.5, showing that household wealth itself accounts for the bulk of socioeconomic-related inequality in test score performance across all age ranges.

An additional concern may be the attrition of the sample and missing observations across some periods

since this could introduce heterogeneity in the distribution of wealth and affect the stability of our estimated results. We address this issue by restricting the sample to the same observations across all rounds. Again, Tables A10 and A11 are similar to the baseline result in terms of the degree and source of socioeconomic inequality in test scores (See Tables 2.1 and 2.2).

2.7 Conclusions and Discussion

Differences in cognitive outcomes that emerge from early ages explain substantial variation in several later-life outcomes, including performance in school, employability, and earnings. In this chapter, we investigated the evolution of socioeconomic-related inequality in language skills for children between the ages of 5 and 15 years old in Peru. First, we employ the concentration index to estimate the extent of socioeconomic-related inequality in language achievement. Then, we estimate the achievement regression, which is used in the decomposition analysis as contributing factor to socioeconomic-related inequality. The decomposition of socioeconomic-related inequality is then divided in two main components: the elasticity of language performance to each factors and the degree of socioeconomic-related inequality in that factor itself. Lastly, we decompose the source of changes in socioeconomic-related inequality in language performance between ages 5 and 15.

This study complements evidence from both developed and developing countries. The results suggest that socioeconomic inequality in language achievement, favoring the well-off, emerges from the early ages and persists until adolescence, albeit dropping. This finding is consistent with the previous evidence showing that disparities in child cognition change little during middle childhood ([Galasso et al., 2019](#)). The regression estimate from the language achievement shows that household wealth is positively associated with child better performance in test scores. The decomposition exercises mirror this finding, indicating that family socioeconomic status combined with a high level of socioeconomic inequality accounts for a significant portion of the observed socioeconomic-related inequality in language achievement. However, after controlling for lagged achievement, the persistence of past achievement emerges as a major contributor at age 15. The relative importance of family wealth reduces but still plays a critical role in explaining socioeconomic-related inequality in performance in test scores. The results remain unchanged even when using the achievement

in the PPVT scores at age 8. The high persistence of test scores may imply the effect of school influence in explaining the difference in test score performance to some extent ([Castro and Rolleston, 2018](#)) although we lack data to further explore such mechanism.

These findings highlight important policy implications. First, our results from the concentration index emphasize the importance of intervening at an earlier age. For instance, [Sánchez et al. \(2020\)](#) document cognitive gains from children exposed to Juntos conditional cash transfer during their first four years of life in Peru. These interventions can narrow socioeconomic-related inequality between students at the start of the school. Second, the decomposition results seem to point out that reducing socioeconomic-related inequality in language achievement would imply policies that address either the elasticity or the socioeconomic inequality component of family wealth. However, some trade-offs need to be considered. For instance, reducing the association between family wealth and child language achievement via the elasticity component may decrease the average test score performance. Therefore, policies seeking to promote equity would require in-depth assessment of cost and benefit analysis associated with each component, which is beyond the scope of this paper.

This study is not without limitations. The exclusion of 5% of the richest part of the country from the data sampling approach may underestimate socioeconomic-related inequality in the estimated results. Likewise, the measurement errors in the household wealth index may either overestimate or underestimate the estimated results, as it is self-reported. Another limitation is that the result should not be interpreted as a causality. Furthermore, we lack sufficient information on school characteristics. Further research should explore the interaction between the schools and family factors in relation to socioeconomic inequality in child cognitive development, including its evolution.

Chapter 3

Armed conflict and schooling outcomes in the Congo

3.1 Introduction

Developing countries have become a hotspot of armed conflicts in recent years. For example, virtually one-third of Sub-Saharan African (SSA) countries have experienced some form of violent conflicts since the 1990s ([Gleditsch et al., 2002](#); [Raleigh et al., 2010](#); [SIPRI, 2022](#)). This proliferation of armed conflicts poses a significant development retardation, with deleterious consequences for the civilian population. It is no surprise that the 2011 World Development Report named armed conflict as the major development issue of our time ([World Bank, 2011](#)). Indeed, households in conflict-affected areas face multiple risks, including death, injury, malnutrition, disease, asset loss, and forced displacement. The negative consequences of armed conflict may be severe for children. Well-established studies show that children are particularly vulnerable to armed conflict, as exposure during this critical development period may result in a substantial human capital loss ([Alderman et al., 2006a](#); [Currie and Vogl, 2013](#); [Cuesta and Leone, 2020](#)).

This chapter investigates the relationship between exposure to armed conflict during childhood and school-

ing outcomes in the Democratic Republic of Congo (DR Congo). The DR Congo has suffered from decades of violent conflict since the mid-1990s, especially in the eastern provinces. The second Congo war was particularly brutal, causing severe humanitarian crises. During this period, the country recorded one of the highest death tolls since World War II, with more than 3.8 million fatalities between 1998 and 2004 ([Coghlan et al., 2007](#)). Similarly, the number of internally displaced people is the highest on the African continent, estimated at 5.4 million since the beginning of the conflict in 1996 ([IDMC, 2022](#)). Later, the Congo War gave rise to a series of smaller conflicts in the eastern regions, which continued to severely impair the local population's livelihoods.

Our empirical approach exploits variation in conflict intensity across birth cohorts and DHS cluster levels, combining the Demographic and Health Survey (DHS) with the Armed Conflict Location Event Dataset (ACLED). We use the geographical coordinate information from both datasets to construct our measure of conflict exposure. Then, we interact this measure of conflict exposure with the birth cohorts following the standard approach in the literature ([Chamarbagwala and Morán, 2011](#); [Swee, 2015](#)). We define the cohort groups based on the age of exposure at the onset of the conflict, splitting them into affected and unaffected cohorts. The affected cohorts are individuals between the ages of 6 and 11 at the start of the conflict for the likelihood of primary school completion as an outcome. For the lower secondary completion outcome, we consider individuals between the ages of 12 and 13 at the beginning of the conflict. The unaffected cohorts are individuals between the ages of 12 and 24 at the beginning of the conflict for primary schooling. Individuals between the ages of 14 and 24 are the unaffected groups for lower secondary schooling. These cohort groups are beyond their school age at the onset of the conflict. We use an additional cohort, those above the age of 24, at the beginning of the conflict, for the placebo analysis. We also include a set of fixed effects to control for unobserved shocks. Province fixed effects account for differences in socioeconomic characteristics across survey villages. Year fixed effects to control for year-specific shocks that may confound the relationship between armed conflict during childhood and educational outcomes at later ages.

This study makes two contributions to the literature on the legacy of violent conflicts. First, our work is one of the few studies focusing on protracted armed conflict, an under-explored question in the literature. Most papers on armed conflict impacts on human capital tend to focus on short-term armed conflicts ([Shemyakina,](#)

2011; Akresh et al., 2012; Minoiu and Shemyakina, 2014; Akresh et al., 2022, 2023). Short-term armed conflicts are often characterized by a single conflict, and this may not be the case in the context of protracted armed conflicts. Protracted armed conflicts may have a series of distinct conflicts with varying levels of intensity. Moreover, in the context of protracted armed conflicts, some public services may still operate, albeit below the optimal level (Verwimp et al., 2019).

Second, this study uses conflict exposure measures at a finer geographical level. Several studies rely on conflict exposure at higher administrative units, such as region or district levels (Chamarbagwala and Morán, 2011; Minoiu and Shemyakina, 2012; Justino, 2012; León, 2012; Swee, 2015; Bundervoet and Fransen, 2018). Such an approach may capture an imprecise exposure to conflicts, especially if households reside in larger districts. In this study, we compute conflict exposure as the weighted average of the number of conflicts, using the inverse distance between the household location and the conflict site as weights. This measure accounts not only for households living close to conflict zones but also for those living far away, by assigning a higher weight to those living close to conflict zones. This is relevant as armed conflicts may have spillover effects on those living far away from the conflict zone.

Our results reveal no evidence of a negative relationship between exposure to armed conflict during childhood and schooling outcomes. The interaction terms between exposure to conflict and affected cohorts are statistically insignificant for both the probability of completing primary and lower secondary schooling. This finding suggests a community resilience to conflict, particularly in the context of prolonged armed conflict (Utsumi, 2022). Furthermore, the result indicates that the relationship between conflict exposure and schooling outcomes is context-specific, consistent with existing empirical evidence (Verwimp and Van Bavel, 2014). For example, evidence regarding conflict-education nexus is mixed in the literature and depends on different factors, such as the level of conflict intensity and institutional context of the country (Unfried and Kis-Katos, 2023). The findings are robust with alternative definitions of conflict exposures and specifications.

The rest of the chapter proceeds as follows. Section 3.2 provides background on the conflict and the education system in the Democratic Republic of the Congo. Section 3.3 discusses the literature on armed conflict exposure and human capital. Section 3.4 describes the empirical approach. Section 3.5 discusses the results on the association between childhood exposure to armed conflict and schooling outcomes. Section

3.6 presents the robustness checks for the main results. Section 3.7 concludes the study.

3.2 Background

3.2.1 Conflict context

The DR Congo is a fragile state plagued with protracted armed conflicts. Since the early years of independence, the country quickly fell into political chaos, followed by years of dictatorship under the Mobutu regime. However, the dramatic turning point occurred in the mid-1990s with two major civil wars: the first Congo war (1996–97) and the second Congo war (1998–2003), the precursors to today's ongoing conflict, particularly in the eastern regions.

The origin of recent Congolese conflicts is not clear-cut. Nevertheless, the dominant narrative is a spillover effect from the 1994 Rwandan genocide ([Autesserre, 2010](#)). In the wake of the Rwandan genocide, a massive inflow of Hutu refugees fled to the Congo, fearing the persecution of the new Tutsi-led government. Most of the Hutus settled in refugee camps in North and South Kivu, which the former Hutu militia used as their base to launch attacks on the Tutsi Rwandan government and the local Tutsi population, residing in South Kivu.

As ethnic tensions grew, Uganda and Rwanda, together with the *Alliances des Forces Démocratiques pour la Liberation du Congo* (AFDL), launched an anti-Mobutu rebellion in the eastern Congo. The rebellion quickly erupted into a full-fledged war, becoming what is now known as the first Congo War (1996-1997). The war ended with the fall of the long-standing Mobutu dictatorship and the control of Kinshasa by the AFDL and its leader, Kabila, in May 1997. Under intense pressure from the local population, Kabila ordered Rwandan and Ugandan troops to vacate the country. Soon, disputes arose between Kabila and his former allies—Rwanda and Uganda. A new rebel movement, the Rally for the Congolese Democracy (RCD), emerged in eastern Congo in August 1998, setting the stage for the Second Congo War (1998-2003).

The conflict intensified as other parties joined the war. The Hutu Interahamwe and local Mayi Mayi

militias, as well as troops from Angola, Chad, Namibia, and Zimbabwe, fought alongside the DR Congo forces against the rebel forces during the war. Unlike the first Congo war, the 1998 war was one of the most brutal conflicts in the African continent, killing more than 3 million people between 1998 and 2004 (Coghlan et al., 2007). The conflict started in the Kivu region and quickly spread into other parts of the country. As a result, the country was split into different sections. The Congolese government controlled the southern and western parts of the country. The Ugandans and their rebel-backed movements, the Movement for Liberation of the Congo (MLC) and Rally for Congolese Democracy-Kisangani (RCD-K), occupied a large portion of the northeast and northwest parts of the countries, particularly the Equator and Orientale provinces. Finally, the Rally for Congolese Democracy-Goma (RCD-G) seized the central-eastern parts of the country, including the Kivu regions, the northern part of Katanga, and parts of the Kasai-Oriental provinces. The areas controlled by the government and MLC experienced little or no fighting, whereas intense military confrontation occurred in the eastern part, particularly in Nord-Kivu, Sud-Kivu and Ituri (Turner, 2007).

One year after the conflict, the first peace agreement started in July 1999, the Lusaka peace agreement. The Lusaka ceasefire agreement called for the withdrawal of foreign troops from the fighting zone and the deployment of UN peacekeeping operations in the Congo. However, fighting continued in the northeastern part of the country and territories held by the RCD-Goma. The war officially ended in July 2003 after a series of peace agreements in Pretoria, which led to a transitional government from 2003 to 2006. However, conflict is ongoing in the eastern regions, particularly in the Kivu provinces, the Ituri district, and northern Katanga, causing massive humanitarian crises in the country.

3.2.2 Education context

Modern education in the DR Congo emerged in the early 1900s with the establishment of schools by faith-based organizations, laying the basis for the education system in the post-independence period. The DR Congo education system consists of primary, secondary and tertiary education. Primary education encompasses six years of schooling for children aged 6 to 11 and is compulsory. In the final grade of primary education, children need to pass a national examination, the *Test National de Fins d'Etude Primaires* (TENAFEP), which

gives access to secondary education. Secondary education consists of two years of lower secondary level (age 12-13) and four years of upper secondary level (age 13-17). At the end of secondary education, the student takes a national examination, which gives access to higher education. Public schools are the major providers of primary and secondary education. The Congolese public school system consists of state schools (*Ecole non-conventionnée*) and faith-based schools (*Ecole Conventionnée*).

During the 1970s and 1980s, the education sector in the DR Congo was expanding rapidly. In the same period, the primary enrollment ratio stood above 80 percent, even with the widespread use of school fees in the 1980s, following a drastic reduction of public financing in the education sector ([World Bank, 2015](#)). In the 1990s, including the years of the onset of the war, the education system suffers from a decline in the enrollment rate at the basic level, with a gross enrollment ratio of 53% compared to two decades ago of 95% enrollment ratio ([World Bank, 2022](#)). In the same period, over 60% of primary school-age children were out of the education system, higher than the SSA average (41%) ([World Bank, 2022](#)).

After a decade of conflict, overall education seems to have improved. For instance, in recent years, the main indicators of basic education are now converging to the pre-conflict level, with an enrollment level in primary education of over 90% as of 2010 ([World Bank, 2022](#)). However, several factors restrict the provision and access to education, particularly in conflict-affected areas. The destruction of school facilities and displacement have kept many children out of school. Such a factor constitutes a development challenge for a country with over 40% of its population under the age of 15 ([World Bank, 2022](#)).

3.3 Literature Review

This study is related to a growing literature on the link between armed conflict and human capital accumulation ([Justino, 2012](#)). While earlier literature primarily focuses on the causes of armed conflict with an emphasis on greed and grievance ([Collier and Hoeffler, 1998, 2004b; Collier et al., 2009](#)), attention has shifted to the consequence of armed conflict in recent years ([Blattman and Miguel, 2010](#)).

The literature on the consequences of armed conflict encompasses both macroeconomic and microe-

conomic aspects. Empirical evidence on the effect of armed conflict at the aggregate level predicts a quick recovery to the pre-war trend a few years after the end of the conflict ([Blattman and Miguel, 2010](#); [Hodler, 2019](#)). For instance, [Davis and Weinstein \(2002\)](#) show a rapid recovery from heavy bombing in Japanese cities two decades after World War II. [Miguel and Roland \(2011\)](#) provide similar evidence in the context of the Vietnam war, with no evidence of long-term impact on poverty. The potential explanation behind such a quick recovery is that physical capital can be remediated through infrastructure reconstruction and aid program ([Collier and Hoeffler, 2004a](#); [Blattman and Miguel, 2010](#)).

At the microeconomic level, however, findings of recent growing literature suggest a robust negative effect of armed conflict on educational outcomes ([Cuesta and Leone, 2020](#)). Furthermore, the gendered impact of conflict on education may be context-specific. For instance, the adverse effect of armed conflicts on education may be stronger for boys as a result of military enrollment ([Blattman and Annan, 2010](#); [Swee, 2015](#)), labor force participation ([Rodríguez and Sánchez, 2012](#); [Justino et al., 2014](#)) and existing pre-conflict gender inequalities in educational access ([Akresh and de Walque, 2008](#); [Verwimp and Van Bavel, 2014](#); [Bertoni et al., 2019](#)). On the other hand, girls suffered greater loss in educational outcomes due to parents' concern for their safety and marriageability, as in the case of the Tajikistan armed conflict ([Shemyakina, 2011](#)). In some instances, armed conflict may have a positive effect or no effect at all, particularly in the context of a low level of conflict intensity ([Valente, 2014](#); [Justino et al., 2014](#); [Pivovarova and Swee, 2015](#)).

In other cases, disadvantaged and vulnerable demographic groups may suffer the most from civil conflict. For example, [Chamarbagwala and Morán \(2011\)](#) show rural Mayan exposed to Guatemalan civil conflict during their school-age complete few years of schooling. [León \(2012\)](#) shows that children exposed to the Peruvian civil conflict during the preschool age are the most affected in terms of their educational attainment compared to those exposed during their primary school age. In a similar vein, [Bundervoet and Fransen \(2018\)](#) provide evidence of the persistent effect of violent conflict exposure in utero on educational attainment in the context of the Rwandan genocide.

The literature has identified several channels through which armed conflict impacts educational outcomes. Armed conflict may affect the provision of schooling through the destruction of infrastructure. For example, [Akbulut-Yuksel \(2014\)](#) reports a loss of educational attainment through the destruction of schools due to

intense bombing in Germany during World War II. Similarly, displacement induced by conflict may affect children's educational outcomes and disrupt schooling supply ([Justino, 2016](#)). A change in distributional effects, such as socioeconomic inequality, may affect the provision of education, leading to inequality in education access in conflict-affected zones. On the demand side, armed conflict may result in household asset losses. Under this circumstance, households may face a higher trade-off between sending children to school and participating in the labor market ([Rodríguez and Sánchez, 2012](#)). Furthermore, adverse health shocks induced by armed conflict may affect the demand for schooling. For example, [Alderman et al. \(2006b\)](#) documents that children who suffered from malnutrition due to Zimbabwean civil conflict completed a few years of schooling. Armed conflict may result in significant trauma and fear, which may result in lower school enrollment. For example, [Brück et al. \(2019\)](#) show that Israel-Palestinian conflict reduces schooling outcomes through adverse psychosocial well-being.

Several of the existing studies attempt to identify the causal effect of conflict using variation across birth cohorts and conflict intensity in difference-in-differences (DiD) settings combined with fixed effects. For instance, [Shemyakina \(2011\)](#) combines DiD with household fixed effects, [Verwimp and Van Bavel \(2014\)](#) use province-level fixed effects, and [Chamarbagwala and Morán \(2011\)](#) consider department and year of birth fixed effects. [Bertoni et al. \(2019\)](#) specify two different models: a DiD to study the number of years of education and a fixed effects panel data model for school enrollment, accounting for individual unobserved heterogeneity. Similarly, some studies benefit from panel data availability, such as [Justino et al. \(2014\)](#), [Valente \(2015\)](#), and [Pivovarova and Swee \(2015\)](#). The study of [Rodríguez and Sánchez \(2012\)](#) combines IV techniques with detailed conflict data, using violence incidence data as an instrument to address the endogeneity of conflict exposure. Other studies, like [León \(2012\)](#), compare children within households exposed to conflict at different ages, using birth cohort and household fixed effects to assess the impacts on educational outcomes.

3.4 Empirical Approach

3.4.1 Data

The analysis combines data from two sources. The individual-level data comes from two waves of the Demographic and Health Survey (DHS) from the DR Congo, conducted in 2007 and 2014. These two waves consist of different samples of individuals since they are observed at one point in time. The survey was implemented by the Ministry of Planning and funded by the United States Agency for International Development (USAID).

The DHS is a nationally representative cross-sectional survey which provides information on health, education and other socioeconomic outcomes. The data contain information on the location where the survey took place, with the Global Positioning System (GPS) coordinates, also referred to as clusters. However, the geo-coordinates have been randomly displaced up to a 10 km radius to protect the anonymity of survey respondents. The direction of displacement is randomly chosen, with the caveat that coordinates are not displaced outside of the region. This displacement introduces a classical measurement error and thus has the potential to bias our estimates downward. The survey was administered in 828 clusters.

The sample consists of reproductive age individuals: women (aged 15-49) and men (aged 15-59) at the time of the survey. The outcome variables for the analysis are schooling outcomes at the time of the survey. We consider the probability of completing primary and lower secondary schooling. Primary schooling completion is defined as a dummy variable taking the value 1 if the individuals have completed at least 6 years of schooling. Lower secondary completion is defined as a dummy variable if the individuals have completed at least 8 years of schooling. We also have information about the years of birth of the individuals in the sample. Women in the sample are born between 1957 and 1999, and men are born between 1947 and 1999. We use this information to identify exposure to conflict at their time of school-going age, whether primary or lower secondary schooling at the time when the conflict started, in 1996.

The caveat is that the DHS does not provide information on migration history. We do not know whether an individual moved from a high conflict place to a safer area, which could affect the estimated results.

Fortunately, the DHS 2007 contains information on the number of years that individuals have been living in the residence. We take advantage of this information for the robustness analysis.

The second data source comes from the Armed Conflict Location Event Dataset (ACLED). ACLED provides information on the location and timing of conflict events at a disaggregated level, with point coordinates. ACLED draws its sources from local and international news as well as reports from humanitarian agencies. We focus on the period 1997-2014, encompassing the Congo war (1997-2004) and the ongoing conflict in the eastern part of the country (2004-2014). The dataset includes a broader aspect of conflict, including the actors involved. Conflict events combine violent and non-violent conflicts. Violent conflicts refer to battles, violence against civilians and remote violent events. Non-violent conflicts encompass protests and riots.

The main information on conflict intensity comes from the number of conflict events. The dataset contains 8257 total conflict events between 1997 and 2014. ACLED data is not without limitations. First, it suffers from under-reporting bias ([Van der Windt and Humphreys, 2016](#)). Second, the precision of the geolocation of conflict events has been questioned¹² ([Eck, 2012](#); [De la Sierra, 2020](#)). We use an alternative conflict dataset, the Uppsala Conflict Data Program Georeferenced Data (UCDP GED)¹³ for robustness checks. This dataset contains a more precise localization of conflict events since 1989. Unlike ACLED, the UCDP GED has less information on conflict events. This is because UCDP GED defines an armed conflict event as the use of armed forces between two parties, of which at least one is the government of a state, resulting in at least 25 battle-related deaths in a calendar year ([Sundberg and Melander, 2013](#)).

3.4.2 The exposure to conflict intensity index

One of the key challenges in the study of conflict impact is being able to identify the conflict at a finer geographical level. Most studies rely on examining the intensity of conflict at a very high administrative level, such as the district level ([Justino, 2012](#); [León, 2012](#); [Swee, 2015](#); [Bundervoet and Fransen, 2018](#)). The limi-

¹²We found about 58 locations wrongly assigned to conflict events. We assign the wrong location to the territory level and drop the remaining wrong location for which we do not have information in the dataset's notes or source sections of the dataset.

¹³UCDP-GED is developed by the Department of Peace and Conflict Research at the Uppsala University, and we use version 23.

tation of such an approach may be a less precise way of capturing individual exposure to conflict, especially when households reside in larger districts. A more accurate way of measuring conflict intensity would imply connecting the conflict zone to household information, such as in the case of [Bertoni et al. \(2019\)](#).

In this study, we construct a measure of conflict exposure, closely related to [Maystadt and Verwimp \(2014\)](#). The measure of conflict exposure intensity is defined as the inverse weighted distance from the DHS cluster to the conflict zones, accounting for the number of conflict events during the wartime period, that is, between 1997 and 2014. The idea behind such an approach is that DHS clusters closer to conflict zones are more likely to be affected by the conflict than those farther away. This approach is also related to that of [Akresh et al. \(2022\)](#), but it allows for spatial variation in the intensity of conflict while accounting for indirect effects on individuals living far away from the conflict zone since it assigns less weight to them.

The exposure to conflict intensity is defined as follows:

$$conflict\ exposure_c = \sum_j \frac{1}{dist_{cj}} * number\ conflict_c \quad (3.1)$$

where $dist_{cj}$ is a geodesic distance between conflict location c and the DHS cluster j ; $number\ conflict_c$ is the number of total conflicts in a given location. For the robustness checks, we separate the number of total conflicts into violent and non-violent conflicts. Furthermore, we use the number of deaths from the UCDP GED to capture conflict exposure.

3.4.3 Empirical strategies

We follow the standard approach in the literature by exploiting variation in conflict exposure across birth cohorts and DHS clusters ([Chamarbagwala and Morán, 2011](#); [Swee, 2015](#); [Bertoni et al., 2019](#))¹⁴. Our approach is similar to a continuous DID akin to [Duflo \(2001\)](#), comparing the outcomes of younger cohorts (treated group)

¹⁴We consider school-age individuals at the start of the conflict as the affected cohort in line with those studies ([Chamarbagwala and Morán, 2011](#); [Swee, 2015](#); [Bertoni et al., 2019](#)). For robustness checks, We consider the age group 0-5 as an additional affected cohort for the primary schooling. The individuals aged 0-5 and 6-11 age group 6-11 are additional affected cohorts for lower secondary schooling.

whose schooling is more likely to be affected by conflict than the older cohorts (control group) for a given schooling level. We divide the individuals into affected and unaffected cohorts based on their school-going age at the onset of conflict, in 1996. This option follows from the fact we do not observe the same individuals in two moments of time. Instead, we have two different waves, each with a different sample of individuals.

The affected cohorts are those between the ages of 6 and 13 at the beginning of the conflict in 1996, since their schooling may be directly affected by the conflict. Specifically, these are the affected cohorts for primary schooling completion. Individuals in this affected group are born between 1985 and 1990. For lower secondary schooling completion, the affected group includes individuals between the ages of 12 and 13 at the onset of the conflict. Individuals in this affected group are born between 1983 and 1984. Finally, the unaffected cohorts are individuals between the ages of 12 and 24 at the onset of the conflict for primary schooling. For lower secondary schooling, the unaffected cohorts are individuals between the ages of 14 and 24 at the beginning of the conflict. Individuals in this cohort group are assumed to have completed either primary or lower secondary schooling at the onset of the conflict. We use an additional cohort, those above the age of 24, at the beginning of the conflict, for the placebo analysis.

To examine the relationship between exposure to armed conflict during childhood and schooling outcomes, we estimate probit models on the probabilities of completing primary and lower secondary education.¹⁵ Consider that the propensity of a given individual to complete primary/lower secondary education is given by

$$Y_{itcd}^* = \mathbf{W}\boldsymbol{\beta} = \beta_0 + \beta_1(\text{conflict exposure}_c * \pi_t) + \mathbf{X}\boldsymbol{\beta}_k + \epsilon_{itcd} \quad (3.2)$$

$$Y_{itcd} = \begin{cases} 1 & \text{if } Y_{itcd}^* > 0 \\ 0 & \text{if } Y_{itcd}^* \leq 0 \end{cases}$$

where Y_{itcd}^* is the (latent) propensity for completing primary or lower secondary schooling of an individual i , born in the year t , residing in a cluster c and in a province d . Y_{itcd} is a binary variable equal to 1 if the individual completed primary/lower secondary education. The main explanatory variable is the interaction

¹⁵We also estimate Linear Probability models as robustness checks. The results are similar to those of the baseline estimates.

term ($conflict\ exposure_c * \pi_t$), where π_t is a binary indicator for the affected cohort. $\mathbf{X}$ includes several controls such as the DHS year of survey dummy, a set of dummies for the province, a set of dummies for the birth cohorts, and individuals' characteristics, such as household's wealth, whether the household lives in urban areas and gender. The wealth index is an asset-based index that captures the socioeconomic characteristics of the household. The birth cohort dummies account for common characteristics of all individuals born in a given year. Province dummies account for all (un)observable time-invariant characteristics across regions. The DHS years of survey dummy accounts for macroeconomic shocks. $\mathbf{W}$ stands for the matrix with all variables, the interaction as the main variable and the controls. β_k is the vector of the control variables' coefficients with $k = 2, \dots, K$, whereas β includes the full set of parameters to be estimated. ϵ_{itcd} is the error term.

The implicit identifying assumption is that differences in schooling attainment between the affected and unaffected cohorts in a cluster experiencing low conflict intensity is a valid counterfactual for what would have happened to school attainment between the cohorts in high-conflict clusters in the absence of conflict, conditional on the control variables and unobserved factors, such as province and time fixed effects. Then, the probability of completing primary or lower secondary education is given by :

$$P(Y_{itcd} = 1 \mid \mathbf{W}) = P(Y_{itcd}^* > 0 \mid \mathbf{W}) = P(\mathbf{W}\beta + \epsilon_{itcd} > 0) = \Phi(\mathbf{W}\beta) \quad (3.3)$$

where ϵ_{itcd} is assumed to be normally distributed. The model is estimated based on the maximum likelihood method. Furthermore, observations are assumed to be independent, and all relevant explanatory variables are included, implying that the model is correctly specified.

Note that the estimated coefficients are not directly interpretable since they are parameters of the latent model. To provide a more intuitive interpretation, we computed the average marginal effects to capture how a one-unit change in the independent continuous variables affects the probability of primary/lower secondary schooling completion. For dummy independent variables, the marginal effects are computed as the difference in the estimated probabilities between the two categories.

Table 3.1 presents the descriptive statistics for the sample. More than half of the individuals in the

sample have completed at least primary schooling (57%). Only a handful of individuals have completed lower secondary education (42%). About 58% of individuals live in rural areas. About 28% of the individuals in the sample are male, which may not be representative of the male population. It is likely that women may be oversampled, given the focus of the survey on women. We apply sampling weights to mitigate this gender imbalance.

Figure 3.1 shows the distribution of armed conflicts between 1997 and 2014 across localities. We can observe a high number of conflict events in the eastern part of the country, particularly in the Kivu and Ituri districts, which are also the regions with a higher conflict intensity. We can also note a sizable number of conflicts in the western part of the country, where non-violent conflicts tend to predominate.

Table 3.1: Descriptive statistics

	Mean	SD
Primary schooling completion	0.574	
Lower secondary completion	0.417	
Conflict exposure	23.517	40.283
Violent conflict exposure	16.897	27.168
Non-violent conflict exposure	6.620	17.428
Conflict exposure (UCDP)	233572.6	500002.4
Conflict exposure 1996 (UCDP)	171.250	251.918
Wealth index	0.219	1.158
Urban	0.420	
Male	0.284	
N		22551

Conflict exposure includes non-violent and violent conflicts. Violent conflict exposure refers to battles, violence against civilians and remote Non-violent conflict exposure refers to protests and riots. Conflict exposure 1996 is the number of deaths in 1996 from the UCDP. Conflict exposure is the number of deaths during the wartime from the UCDP.

Source: DHS, own computations.

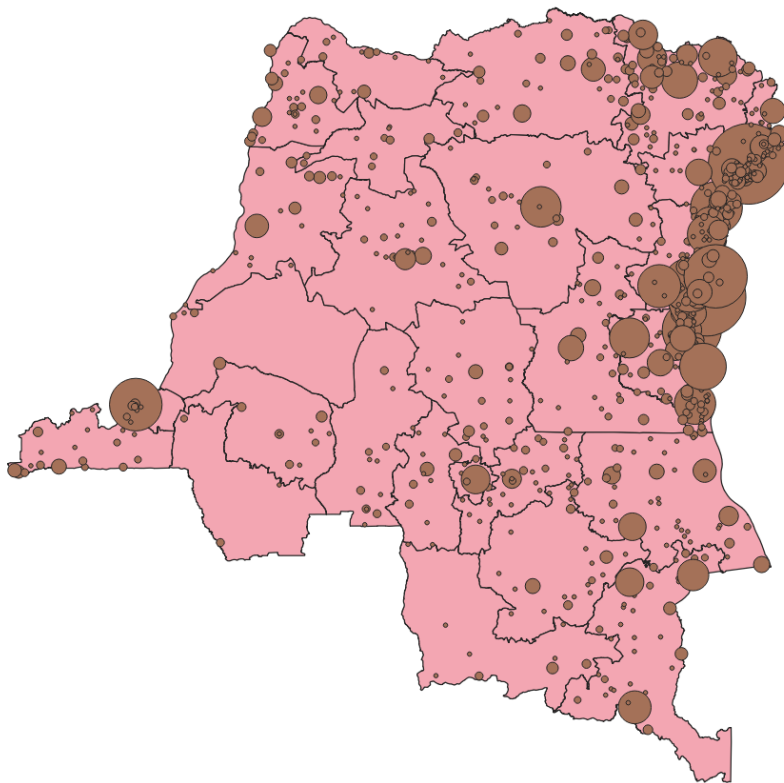


Figure 3.1: Distribution of conflicts across localities between 1997 and 2014

Source: ACLED, Own computations.

3.5 Results

In this section, we present the empirical results for the relationship between childhood exposure to armed conflicts and schooling outcomes.

3.5.1 Armed conflict during childhood and educational outcomes

Table 3.2 presents the estimated results for the relationship between childhood exposure to armed conflict and schooling outcomes based on equation (3.1). Columns (1)-(2) capture the probability of completing primary schooling, and the likelihood of lower secondary schooling completion is the outcome in the last two columns. We consider individuals aged 6-11 at the beginning of the conflict in 1996 as the affected group for the likelihood of primary school completion. Individuals aged 12-13 at the onset of the conflict are the affected group for the probability of lower secondary school completion. The unaffected groups are individuals between the ages of 12 and 24 and those aged 14-24 at the beginning of the conflict. For both outcomes, the regression is estimated with(out) controlling for the province fixed effects and year of survey dummies.

The results indicate no statistically significant relationship between armed conflict during childhood and schooling outcomes. The interaction between conflict exposure and the affected groups is not statistically significant in relation to both primary and lower secondary school completion. While one might expect conflict exposure during childhood to reduce the probability of primary and lower secondary school completion, evidence is mixed in the literature ([Verwimp and Van Bavel, 2014](#)). For instance, a recent study conducted in the SSA shows that armed conflict may have no effect on educational attainment in the context of low-intense conflict ([Unfried and Kis-Katos, 2023](#)). This highlights that the relationship between conflict and education is highly context-specific ([Justino, 2016](#)). For example, evidence regarding the conflict-education nexus is mixed in the literature and depends on different factors, including the institutional context of the country([Unfried and Kis-Katos, 2023](#)) Moreover, the lack of statistical significance in the relationship between conflict exposure and schooling attainment may reflect community resilience, particularly in the context of prolonged armed conflict. [Utsumi \(2022\)](#) provides evidence of community resilience in the face of conflict in the context of

Afghanistan. This observation implies that some public services may be provided in the context of protracted armed conflict, although below the optimal level (Verwimp et al., 2019). The size of the estimates is relatively lower compared to previous studies closer to ours (Swee, 2015). For instance, Swee (2015) finds no evidence of the effect of Bosnian conflict on primary schooling attainment, with an estimate of 0.3, which is higher than the estimates we obtain in this study (See Table 3.2).

Furthermore, on average, an increase in household wealth is associated with a higher probability of primary and lower secondary schooling completion. Likewise, males are more likely to complete primary and lower secondary schooling, on average, compared to females. We can also observe, on average, that areas with high conflict intensity are associated with a decreased probability of completing primary and lower secondary schooling, but the conflict did not reinforce that. It may be that, on average, areas with more conflict exposure are less likely to have higher schooling attainment because they are poor by themselves.

Overall, the estimated results suggest no statistically significant relationship between armed conflict exposure during childhood and schooling outcomes. Moreover, the results are stable across specifications.

3.5.2 Heterogeneity analysis

3.5.2.1 Violent versus non-violent conflict exposure

In this section, we consider alternative measures for conflict exposure by differentiating between violent and non-violent conflicts. The construction of conflict exposure is akin to equation (3.1), with the total number of (non-)violent conflicts between 1997 and 2014 in each conflict location.

Results in Table B1 in Appendix B refer to the relationship between childhood exposure to violent conflict and schooling outcomes. The model is estimated with(out) the province fixed effects and the year of the survey dummies. The estimated coefficients are similar in terms of significance to those in the baseline results. Table B2 (Appendix B) repeats the same exercise focusing on the relationship between non-violent conflict exposure and schooling outcomes. The results are consistent with the baseline estimates regarding

Table 3.2: Conflict exposure and Schooling outcomes: Probit marginal effects

	Primary schooling		Lower secondary	
	(1)	(2)	(3)	(4)
Conflict exposure	-0.033*** (0.008)	-0.014 (0.010)	-0.026*** (0.009)	-0.003 (0.007)
Total conflict exposure*age 6-11	0.000 (0.006)	0.003 (0.008)		
Urban	0.130*** (0.023)	0.121*** (0.020)	0.108*** (0.027)	0.095*** (0.020)
Wealth index	0.171*** (0.011)	0.190*** (0.011)	0.161*** (0.010)	0.189*** (0.009)
Male	0.241*** (0.010)	0.243*** (0.010)	0.261*** (0.011)	0.264*** (0.011)
Conflict exposure*age 12-13			-0.005 (0.012)	-0.013 (0.013)
Province FE		Yes		Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy		Yes		Yes
<i>N</i>	22551	22551	13793	13793

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: DHS, own computations.

the likelihood of primary school completion.

However, the results show a negative relationship between non-violent conflict exposure and the likelihood of completing lower secondary schooling. There may be several explanations for this finding. It may be that areas with non-violent conflicts are more widespread than those with violent conflicts in terms of geographical extent. Furthermore, areas with violent conflicts are more likely to receive aid than non-violent ones, which may explain the lack of statistical significance between violent conflict exposure and schooling outcomes. Finally, areas with more violent conflicts may induce higher mortality, and this may underestimate the relationship between conflict exposure and schooling outcomes.

3.5.2.2 Sub-sample analysis

In this section, we estimate whether the relationship between armed conflict exposure and schooling attainment differs by gender and residence location. Table B3 in Appendix B considers the relationship between childhood exposure to armed conflict and educational outcomes using a separate regression by gender. The results are indistinguishable from the baseline estimates, yielding a statistically insignificant relationship between conflict exposure during childhood and schooling outcomes for both regressions. This finding indicates that the use of different samples has little impact on the baseline estimates.

Table B4 in Appendix B repeats the same exercise using a separate regression by the residence location. Again, the results are similar to the baseline estimate in terms of significance.

3.5.3 Mechanisms: Migration selection

In a time of great conflict intensity, individuals may move to a relatively safer place. If such an effect exists, this may affect the estimated results. For example, if the new location where the individuals have moved results in a lower educational outcome, we may overestimate the relationship between armed conflict and educational outcomes. However, if the individuals fleeing the conflict zone present characteristics likely to lead to higher

educational attainment due to higher ability or socioeconomic status, our results may underestimate the association between armed conflict and education. While we do not have information about the migration history, the DHS 2007 contains information about the number of years that an individual has been living to the residence up to the time of the survey. We use this information to construct a variable for migration status. Migration status is defined as a binary variable, taking the value 1 if the individual has moved to a different location during the wartime period since 1996, and 0 otherwise.

We test selective migration by estimating the relationship between exposure to armed conflict and the probability of migrating following previous studies ([Akbulut-Yuksel, 2014](#); [Akresh et al., 2023](#)). Table B5 in Appendix B reveals a non-significant relationship between conflict exposure and the probability of migrating. This finding mitigates the concern for selective migration. It is possible that a poor road network makes it difficult to move far away from the conflict zones. Moreover, the country's Internal Displacement Monitor Center (IDMC) report indicates that a greater majority of internally displaced persons remain close to their place of origin, making them prone to the same risk that caused them to flee in the first place ([IDMC, 2022](#)).

3.6 Robustness checks

3.6.1 Placebo analysis

In this section, we apply a placebo analysis. The placebo analysis aims to test the implicit identifying assumption that in the absence of conflict exposure, schooling attainment between cohorts would have followed a similar trend between high and low-conflict intense zones. by estimating the relationship between conflict exposure and schooling outcomes for the older cohorts. The conflict needs not to affect these groups as they were old enough to complete their basic schooling, namely secondary education. We consider individuals above the age of 24 at the start of the conflict as an additional unaffected cohort. The unaffected cohorts are individuals between the ages of 12 and 24, and those aged 14-24 at the onset of the conflict. Table B6 in Appendix B indicates that there is no statistically significant relationship between conflict exposure and

schooling attainment for the placebo cohort, which validates the baseline estimates.

3.6.2 Alternative conflict dataset

In this section, we use the UCDP GED dataset as an alternative to ACLED. UCDP GED provides a more accurate geo-localization of conflicts than the ACLED ([Eck, 2012](#)). On the other hand, ACLED contains more comprehensive information on conflict events than UCDP GED. This difference stems from the definition of armed conflict. UCDP considers an event as an armed conflict if it results in at least 25 battle-related deaths in a calendar year. On the other hand, ACLED may have measurement errors due to the inaccurate location of conflict events. We replicate the estimates using conflict exposure in a similar way to equation (3.1), accounting for the number of deaths due to conflict.

Table B7 reports the estimation results for the relationship between conflict exposure during childhood and schooling outcomes, focusing on the number of deaths in 1996. The choice of this year is motivated by the fact that the conflict started in 1996. The results are similar to the baseline estimates in Table 3.2 in terms of significance. However, areas with higher conflict exposure tend to have more schooling. This result should be interpreted with caution since there may be some measurement errors in the conflict exposure variable, and it does not account for the distribution of populations in a given area. Table B8 (Appendix B) shows the estimation results for the relationship between armed conflict and schooling outcomes, considering the number of deaths during wartime. Again, the results are indistinguishable from the baseline estimates in terms of significance.

3.6.3 Alternative specifications

In this section, we check whether the baseline results are sensitive to different specifications. We consider an additional age group as an affected cohort, particularly individuals between the ages of 0 and 5 at the onset of the conflict in 1996 for the likelihood of primary schooling completion. The individual in this age group may be affected by the conflict at their later ages. Likewise, we consider individuals between the ages

of 0 and 5, and those aged 6-11 as additional affected cohorts for the likelihood of lower secondary schooling completion. The unaffected groups are individuals between the ages of 12 and 24, and those aged 14-24 at the beginning of the conflict in 1996. Table B9 (Appendix B) reports the relationship between conflict exposure during childhood and the likelihood of completing primary schooling, using different cohort groups. The findings are similar to the baseline estimate in terms of significance.

Figure B1 (Appendix B) presents the marginal effect of completing primary schooling for different cohort groups exposed to conflict. The likelihood of completing primary schooling is increasing with higher conflict exposure for individuals aged 0-5 at the onset of the conflict. Nonetheless, the results show no statistically significant relationship between childhood exposure to conflict and the probability of completing primary schooling across cohort groups.

Table B10 (Appendix B) reports the relationship between conflict exposure during childhood and the likelihood of completing lower secondary schooling, using different cohort groups. The findings show a statistically significant positive relationship between conflict exposure and the likelihood of completing lower secondary schooling for individuals aged 0-5 at the onset of conflict. This finding suggests that the preschool cohort at the onset of the conflict may have escaped more intense conflict than the other school-aged groups. The Congolese conflict was brutal during the late 1990s, particularly at the onset of the conflict. However, this result should be interpreted with caution since it does not account for the heterogeneity in conflict exposure over time. Nevertheless, Figure B2 show no statistically significant relationship between exposure to conflict and the probability of completing lower secondary schooling at different intensities of conflicts across cohort groups exposed to conflict.

We also use the linear probability model to estimate the relationship between exposure to armed conflict and schooling outcomes as additional robustness checks. Table B11 (Appendix B) shows a qualitatively similar result to that of the baseline estimates. The results show no statistically significant association between armed conflict exposure and the likelihood of primary and lower secondary schooling completion.

3.7 Conclusions

This chapter investigates the association between exposure to armed conflict during childhood and schooling outcomes in the context of Congolese armed conflicts. The study exploits variation in conflict exposure across DHS clusters and birth cohorts, combining the DHS and ACLED data.

This study contributes to the existing literature by using a conflict exposure measure at a more accurate level. To capture conflict exposure, we use the weighted average of the number of conflicts taking the inverse distance between household clusters and conflict locations as weights. Such a measure accounts not only for households living close to conflict zones but also for those living far away, as they may be indirectly affected through spillover effects ([Alfano and Cornelissen, 2021](#)). Our key findings show no statistically significant relationship between exposure to armed conflict during childhood and schooling outcomes. These findings suggest that a community may build resilience to conflict, particularly in the context of prolonged armed conflict. [Utsumi \(2022\)](#) shows no impact of conflict exposure on educational outcomes in Afghanistan due to community resilience. It may be the case that communities exposed to conflict see education as a means of restoring normalcy in the midst of chaos ([Buckland et al., 2005](#)). Furthermore, the community may also play a crucial role in preserving the safety of schools in the context of conflict ([Justino, 2016](#)).

Additionally, the finding indicates that the provision of public services may still continue, particularly in the context of prolonged armed conflict, albeit below the optimal level ([Verwimp et al., 2019](#)). In the context of Congolese armed conflict, qualitative studies indicate that the education sector is resilient even in the context of crises and state failure ([Trefon, 2009](#); [Titeca and De Herdt, 2011](#)). This may be explained by household involvement in the education sector through financial contributions, allowing the education sector to survive despite the crises ([De Herdt and Titeca, 2016](#)). The resilience of the education sector could also reflect the perceived return to education related to the alternative labor market, particularly the formal sector ([World Bank, 2005](#)). The finding also highlights that the relationship between armed conflict exposure and educational outcomes is strongly context-specific ([Verwimp and Van Bavel, 2014](#)). For instance, [Unfried and Kis-Katos \(2023\)](#) show that conflict exposure has no impact on education, particularly when its intensity is

lower.

Our analysis calls for further research. While community resilience may mute the negative relationship between conflict exposure and schooling attainment, its effects on the quality of schooling may be different. Further research should investigate the links between armed conflict exposure and the quality of schooling, such as learning outcomes in the context of protracted armed conflict. Moreover, future research should explore specific mechanisms connecting conflict exposure and schooling outcomes, particularly in the context of community resilience.

The study is not without limitations. Data quality issues through a potential measurement error in self-reported measures may affect the estimated results, leading either to a downward or upward bias. Selective mortality could underestimate the relationship between armed conflict exposure and schooling attainment, as the sample consists of survival individuals. Moreover, the cross-sectional data restricts us from capturing the heterogeneity of conflict intensity over time. Future research may use longitudinal data and account for the heterogeneity of conflict intensity in the analysis. Limitations in the data make it challenging to address such caveats. Still, this study shed some light on the legacy of conflict, focusing on protracted armed conflicts.

Chapter 4

Prenatal exposure to drought and child health outcomes in Zambia

4.1 Introduction

Drought remains a disaster type that affects more people than any other, particularly in low-income settings ([CRED et al., 2023](#)). Over the last century, the world has recorded 771 drought episodes, with nearly 44% of incidences in the African continent ([EM-DAT, 2023](#)). Drought is likely to increase in frequency in the coming decades, as a result of climate change ([IPCC, 2023](#)). Low-income countries are susceptible to drought due to a greater dependence on rain-fed agriculture for their livelihoods. Furthermore, households in low-income countries have lower adaptive capacity to weather shocks, owing to poor infrastructures and limited access to social safety nets ([Deschenes, 2014](#); [Meierrieks, 2021](#)). In such a setting, drought can reduce food availability through crop failure and local price increases. Pregnant women and young children are particularly vulnerable to food insecurity. Poor nutritional intake in utero may affect the offspring growth and overall development ([Barker, 1990](#); [Almond and Currie, 2011b](#); [Beshir and Maystadt, 2020](#)).

This chapter examines the effect of prenatal exposure to drought on child health outcomes in Zambia.

Zambia is an interesting setting for several reasons. It is an area prone to frequent drought and is expected to increase in the next decades. Historically, the country has recorded a series of drought episodes with increased severity and frequency in recent years. The years 1991/92, 2015/16 and 2018/19 witnessed a severe dry spell that affected millions of people and resulted in widespread crop failure and food insecurity in major parts of the country ([World Food Programme, 2023](#)). Moreover, the country has one of the highest rates of child malnutrition in the African continent. About 34% of children under the age of five suffer from stunting ([World Bank, 2024](#)). The focus on the fetal period resonates with the "early origin literature", which shows that the adverse effects of the in-utero shocks can persist over the lifecycle ([Barker, 1990, 1995](#)). Likewise, poor health in early childhood leads to adverse socioeconomic outcomes later in life ([Currie and Almond, 2011](#)).

The empirical strategy exploits variations in drought episodes over time within villages. We link information from birth histories from the *Zambian Demographic and Health Survey (DHS)* with weather information from the global *Standardized Precipitation Evapotranspiration Index (SPEI)* data. We combine this information to capture the prevailing drought conditions during pregnancy. The empirical approach then compares the health outcomes of children who were prenatally exposed to intense drought episodes with those who experienced less intense drought conditions in utero. The finer geographical nature of the dataset allows for the inclusion of village-by-month fixed effects and district-specific time trends. They absorb unobserved sources of heterogeneity. For instance, village-by-month fixed effects account for seasonality that may confound the estimates due to a possible correlation of seasonality with birth outcomes ([Buckles and Hungerman, 2013](#)). We also explore some potential channels and other mediating factors that may explain the relationship between drought and child health, namely agricultural yields and parental investment responses.

Our analysis makes two primary contributions to the literature that relates weather shocks to health outcomes. First, we use the SPEI as a measure of drought conditions, consistent with recent climate literature¹⁶ ([Dai et al., 2018](#); [Spinoni et al., 2019](#); [Song et al., 2020](#)). SPEI combines different aspects of drought, including precipitation and evapotranspiration¹⁷ in a unified way ([Vicente-Serrano et al., 2010](#)). In this regard,

¹⁶These studies show that drought results primarily in increased actual evapotranspiration, particularly in subtropical and semi-arid region of Sub-Saharan Africa.

¹⁷Evapotranspiration refers to the amount of available water transferred to the atmosphere from soil surface and plants.

our work differs from previous literature¹⁸ that typically relies on precipitation alone (Maccini and Yang, 2009; Rocha and Soares, 2015; Dinkelman, 2017; Carrillo, 2020; Le and Nguyen, 2021). This is relevant since drought reflects not only a deficit in rainfall, but also in the amount of available soil water for crop yields.

Second, our study contributes to recent discussions on the growing importance of adaptation mechanisms to climate change (Fankhauser, 2017). Indeed, adaptation is now a hot topic at the forefront of public policies. And yet, evidence remains scarce (Deschenes, 2014). We identify the role of prenatal care by a qualified healthcare professional as a possible mitigating factor to extreme weather events in low-income settings. Although we do not exploit a quasi-experimental design in prenatal care use, this study complements the growing literature that points to access to healthcare services as an adaptation mechanism against the adverse effect of climate change and other environmental conditions on health outcomes (Barreca et al., 2016; Mullins and White, 2020; Banerjee and Maharaj, 2020; Cohen and Dechezleprêtre, 2022; Sarmiento, 2023).

Our key finding suggests that children exposed to drought while in utero have short stature for their age and poor physical development. Specifically, an additional month of drought during pregnancy reduces child height-for-age and physical growth by 0.07 standard deviation (SD) and 0.04 SD, respectively. However, the effect of prenatal exposure to drought on child weight outcomes is less robust. Weight captures short-term nutritional status and may be more related to shocks in recent events. In a supplementary analysis, we observe the negative effect of drought among younger children born two years before the survey. This finding is consistent with the fact that height is more latent to in utero shocks than weight (Le and Nguyen, 2021; Bratti et al., 2021). Furthermore, we find that a drought episode during pregnancy results in greater parental investments. This finding highlights suggestive evidence of parental compensatory investments, consistent with the evidence that parents react to health shocks during childhood (Almond and Mazumder, 2013).

The negative effects of prenatal drought are larger during the first and last trimesters of pregnancy. This finding is consistent with medical evidence showing that the first and third trimesters are particularly sensitive to fetal growth (Harding and Bocking, 2001; Altshuler et al., 2003). Additionally, our heterogeneity analysis indicates a similar effect of prenatal drought exposure across child gender and residence of location. We

¹⁸While Dimitrova (2021) use SPEI to examine the contemporaneous effect of drought on child health in Ethiopia, our study looks at the prenatal exposure to drought and provide specific channels connecting drought and child health outcomes

identify lower crop yields as a key channel connecting prenatal drought exposure and child health outcomes. We find that exposure to drought during the growing season reduces the major crop yields. We further show that prenatal care use by a qualified healthcare professional could help mitigate the negative effects of drought during pregnancy on child health outcomes. The finding underscores the importance of investment in adaptation as a potential smoothing factor to adverse weather shocks, particularly in low-income countries.

The remainder of the chapter is organized as follows. Section 4.2 provides related literature on extreme weather impacts on child health. Section 4.3 presents the data used in the study. Section 4.4 outlines the empirical methods. Section 4.5 presents the results. Section 4.6 includes the robustness checks and heterogeneity analysis. Finally, Section 4.7 concludes the study.

4.2 Literature Review: Weather shocks and health outcomes

The study builds on the economic and epidemiological literature on the link between prenatal conditions and health outcomes, also known as the fetal origin hypothesis ([Barker, 1990, 1995](#); [Bloomfield et al., 2006](#); [Almond and Currie, 2011b](#)). These studies suggest that prenatal shocks of different kinds, including violence ([Currie et al., 2022](#); [Mansour and Rees, 2012](#)), famine ([Neelsen and Stratmann, 2011](#)), disease ([Almond, 2006](#); [Barreca, 2010](#); [Schwandt, 2017](#)), natural disasters ([Currie and Rossin-Slater, 2013](#); [Torche, 2018](#); [De Oliveira et al., 2023](#)), and economic crises ([Bozzoli and Quintana-Domeque, 2014](#); [Bharadwaj et al., 2019](#)) have detrimental effects on offspring outcomes.

While prior studies typically focus on the effect of more severe shocks, recent growing literature is now turning attention to more recurrent shocks, such as weather fluctuations. Existing studies use various identification approaches, with several of them considering time-space weather variations. [Andalón et al. \(2016\)](#) use variation exposure to temperature shock by municipality and by year and month of conception in a fixed effects model, whereas the inter-annual variation in the temperature within a given municipality and month of the year, considering that these variations are largely unpredictable, is the basis for [Molina and Saldarriaga \(2017\)](#) empirical strategy. [Le and Nguyen \(2021\)](#) utilize the exogenous variation in deviations from long-term

local rainfall averages to estimate the impact of rainfall variability experienced during the prenatal period on health outcomes for children under five. By exploiting exogenous variation in weather over time within a given spatial area, these studies identify the causal effect of climate shocks on several outcomes, including health.

Within this strand of literature, studies document an adverse effect of exposure to heat during pregnancy on birth weight ([Deschênes et al., 2009](#); [Grace et al., 2015](#); [Andalón et al., 2016](#); [Molina and Saldarriaga, 2017](#); [Chen et al., 2020](#)), gestational length ([Dadvand et al., 2011](#); [Auger et al., 2014](#); [Barreca and Schaller, 2020](#)), probability of stillbirth ([Ha et al., 2017](#)), infant mortality ([Geruso and Spears, 2018](#); [Banerjee and Maharaj, 2020](#)) and children health outcomes ([Bratti et al., 2021](#)).

Similar evidence emerges from studies focusing on drought exposure during pregnancy. Along this research line, evidence points out that drought exposure during pregnancy has a negative effect both on birth ([Persson et al., 2012](#); [Abiona and Ajefu, 2023](#)) and children health outcomes ([Kumar et al., 2016](#); [Le and Nguyen, 2021](#)). However, evidence is mixed regarding excess rainfall. Some studies reveal the beneficial effect of excess rainfall ([Maccini and Yang, 2009](#); [Rocha and Soares, 2015](#)) while others report a negative effect on health outcomes ([Persson et al., 2012](#)). These contrasting findings emerge from different mechanisms at play. For instance, [Maccini and Yang \(2009\)](#) show a positive effect of higher rainfall on health, education and socioeconomic outcomes of rural Indonesian women via better agricultural production. On the other hand, [Persson et al. \(2012\)](#) reveal a negative effect of higher rainfall on infant mortality in Africa, due to a higher incidence of exposure to infectious diseases.

Exposure to drought during pregnancy may affect offspring outcomes through two potential channels, namely the biological and nutritional channels. Regarding the biological channel, drought may exert substantial psychological distress among pregnant women, which can affect fetal development ([Ha, 2022](#)). A study by [Aizer et al. \(2016\)](#) indicates that stress through high levels of cortisol during pregnancy has an adverse effect on offspring cognition and health outcomes. Additionally, drought may result in increased disease exposure, such as diarrhoea through reduced safe drinking water availability ([Bandyopadhyay et al., 2012](#); [Rocha and Soares, 2015](#); [Epstein et al., 2020](#)). Evidence shows that disease exposure during pregnancy may have a detrimental effect on offspring health outcomes ([Barreca, 2010](#)). On the nutritional channel, drought may result in crop failure and food shortage, especially in low-income settings dependent on rain-fed agriculture

(Akpalu et al., 2009; Lobell et al., 2011; Vogel et al., 2019). Maternal undernutrition during pregnancy may result in intrauterine growth restrictions and poor offspring outcomes (Almond and Mazumder, 2013; Miller, 2017).

4.3 Data

This study combines different data sources: the individual level data from the Demographic Health Survey (DHS) and weather data from the global gridded Standard Precipitation Evapotranspiration Index (SPEI) database. We also use the crop yield data from the Central Statistical Office of Zambia for our mechanisms analysis.

4.3.1 Demographic and Health Survey

Our individual-level data comes from three waves of the Zambian DHS, conducted in 2007, 2013 and 2018, respectively. The DHS is a nationally representative cross-sectional survey, which is funded by the United States Agency for International Development (USAID). The DHS provides information on health, education and other socioeconomic outcomes. The DHS follows a two-stage stratified sampling design. In the first stage, enumeration areas are randomly selected from the population census, stratified by regions and urban and rural areas. Then, households are selected randomly for the interview within the selected enumeration areas.

The survey contains retrospective information on birth history from women aged 15 to 49 at the time of the survey. This birth history encompasses the month and year of birth of each child in the past five years before the survey. Children are aged 0 to 59 months at the time of the survey. The sample consists of 14128 children born between 2002 and 2018. For the main outcome variables, we use the 2006 WHO growth standards for anthropometric measures (De Onis and Onyango, 2008). We consider two indicators: the height-for-age z-score (HAZ) and the weight-for-height z-score (WHZ). These anthropometric measures capture

different dimensions of a child nutritional status. The HAZ captures children's long-term nutritional status while the WHZ capture the short-term nutritional status of children. From the sample, we excluded children whose anthropometric measures are implausible from a biological perspective (Mei and Grummer-Strawn, 2007). We also aggregate the two measures to compute the physical growth as the average between HAZ and WHZ in the spirit of Armand and Kim Taveras (2021).

The DHS includes also information on the location where the survey took place, with geographical coordinates, referred to as clusters. Clusters refer to villages (cities) in rural (urban) areas. To preserve the anonymity of respondents, DHS randomly displaces the coordinates of the clusters up to 2 km in urban areas and up to 5 km in rural areas, with a further 1% of rural clusters displaced up to 10 km. The direction of displacement is randomly chosen, with the caveat that coordinates are not displaced outside of the region. This displacement introduces a classical measurement error and thus has the potential to bias our estimates downward.

Another caveat is the lack of migration history in the dataset. Nevertheless, DHS contains information on the number of years the household has been living in the same place up to the time of the survey. We use this information to restrict the sample to women who spent the gestational period in the same residence up to the time of the survey (Grace et al., 2015; Davenport et al., 2020; Bratti et al., 2021). Table 4.1 reports the descriptive statistics of key variables of the study. On average, a child's HAZ and physical growth are 1.5 and 0.7 SD shorter, respectively. The negative HAZ reflects the prevalence of child malnutrition in the country. Regarding maternal characteristics, the average age of mothers at birth in the sample is 27 years. Also, mothers have, on average 5 years, which corresponds to incomplete primary education. Figure C1 in Appendix C shows the location of DHS clusters in each three waves of the survey.

4.3.2 Weather data: Standard Precipitation Evapotranspiration Index

Our weather data comes from the global gridded SPEI database¹⁹, developed by Vicente-Serrano et al. (2010). The SPEI dataset is based on the monthly precipitation and potential evapotranspiration from the Climatic

¹⁹Available at: <http://spei.csic.es/database.html>

Research Unit of East Anglia (CRU) at 0.5 x 0.5 degree spatial resolution. The methodology behind the construction of the SPEI is described in detail by [Vicente-Serrano et al. \(2010\)](#), [Vicente-Serrano et al. \(2012\)](#) and [Beguería et al. \(2014\)](#).

The SPEI corresponds to the difference between precipitation and evapotranspiration. This difference is referred to as the climatic water balance. The main advantage of SPEI is that it captures other factors than precipitation, such as temperature, wind speed, and soil moisture which may play an important role in drought conditions. The index can be computed at different time scales, ranging from 1 to 48-month scales. In this study, we use the 3-month scale, which accounts for the effect of seasonal climatic conditions in Zambia. The country experiences a subtropical climate with three distinct seasons: a hot and dry season (mid-August to mid November), a wet rainy season (mid-November to April) and a cool dry season (May to mid-August). Additionally, This scale is commonly used to capture the agricultural drought, which takes longer to develop than the meteorological drought ([Vicente-Serrano et al., 2010](#); [Homdee et al., 2016](#)).

The index is standardized to have a mean of 0 and 1 standard deviation within each grid cell over the entire historical period (1900 to 2021). Negative values represent conditions more likely to lead to drought relative to the historical average. For each year and month in a given cluster, we define drought as a binary variable, taking the value of 1 if the SPEI is less than -1 SD ([McKee et al., 1993](#)). Then, we count the number of droughts during the nine months preceding the childbirth. Figure 4.1 maps the distribution of the number of droughts in Zambia districts between 2001-2018. The figure reveals that drought is a recurrent phenomenon in Zambia, with a greater majority of districts experiencing drought. However, drought incidences vary substantially across districts. Figure 4.2 displays the percentage of DHS clusters with monthly drought episodes ($\text{SPEI} < -1 \text{ SD}$) between 2000 and 2018. The figure shows that monthly drought episodes vary substantially across Zambia clusters. For instance, episodes of drought occur, on average in about 60% of the clusters within a given month. And yet, there are months with no drought episode in a given cluster. This pattern indicates that drought incidences in a given period are not homogeneous.

4.3.3 Crop Yields data

The crop yield data comes from the Statistical Office (CSO) of Zambia. We use the Crop forecast survey, which contains information on yearly crop yields during the agricultural seasons at the district level. This data contains information on agricultural yields, production, and area cropped from 1996 to 2018. For this study, we use the information on major crop yields, including maize, sorghum and beans in the years 2001 and 2018. The outcome variables are expressed in yields, which is defined as the crop production in tonnes metric divided by the area cultivated in hectare (ha). We use this information to uncover the pathways linking prenatal exposure to drought and child health outcomes.

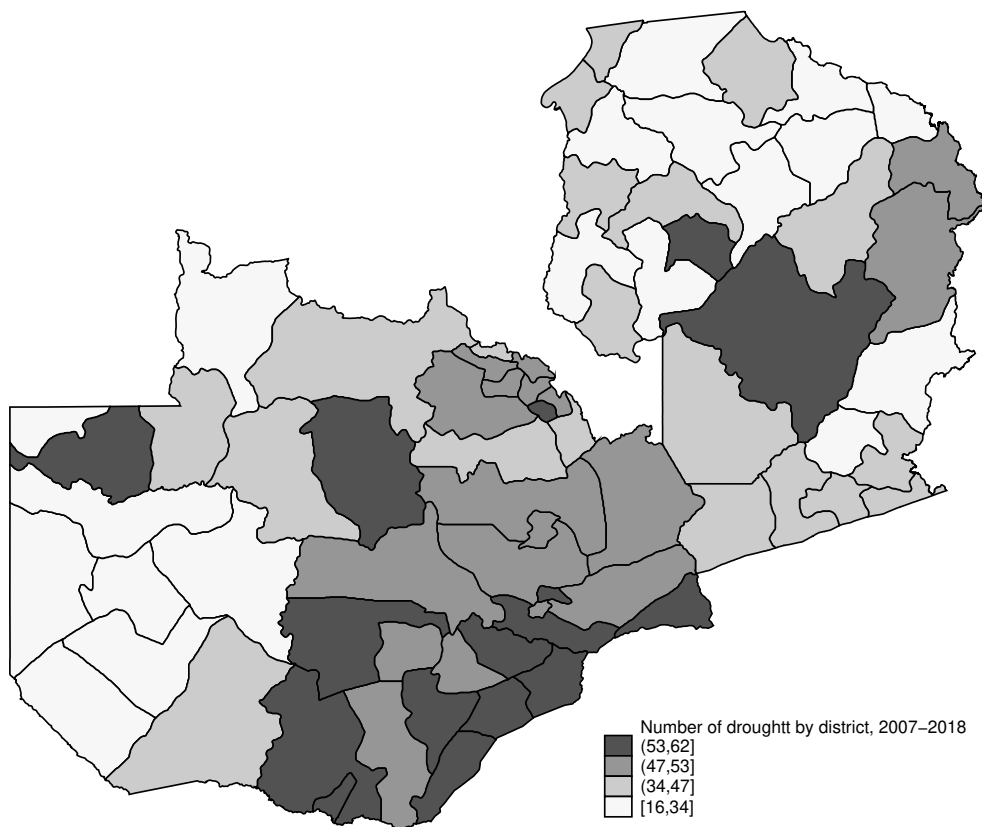


Figure 4.1: Number of droughts per district 2001-2018

Source: Global SPEI dataset, Own computations.

Table 4.1: Descriptive statistics

	Definitions	Mean
HAZ	Child height-for age z-score standardized by age and gender	-1.592 (1.584)
WHZ	Child weight for height standardized by age and gender	0.0435 (1.296)
Physical growth	Average of HAZ and WHZ	-0.775 (0.901)
Stunted	=1 if the child HAZ is less than -2 SD, 0 otherwise	0.401
Wasted	=1 if the child WHZ is less than -2 SD, 0 otherwise	0.0539
Physical growth deficit	1 if the child physical growth is less than -2 SD, 0 otherwise	0.0781
Months drought utero	Number of months with SPEI < -1SD during pregnancy	1.780 (1.943)
Drought utero (dummy)	=1 if child experiences at least one drought episode while in utero	0.670
Months drought utero (1st trim)	Number of drought months during the first trimester of pregnancy	0.609 (0.937)
Months drought utero (2nd trim)	Number of drought months during the second trimester of pregnancy	0.592 (0.931)
Months drought utero (3rd trim)	Number of drought months during the third trimester of pregnancy	0.582 (0.922)
Child male	=1 if child is boy, 0 otherwise	0.491
Child age	Age of child in months	29.53 (17.20)
Birth order	A child's order of birth	4.110 (2.508)
Mother age at birth	Age of mother at child birth	27.27 (7.029)
Mother schooling	Respondent years of education	5.587 (3.491)
Head HH male	=1 if the head of household is male, 0 Otherwise	1.186
Urban	=1 if household lives in urban area, 0 otherwise	0.257
Observations		14128

Standard deviation (SD) in parentheses.

Source: DHS and global SPEI data, own computations.

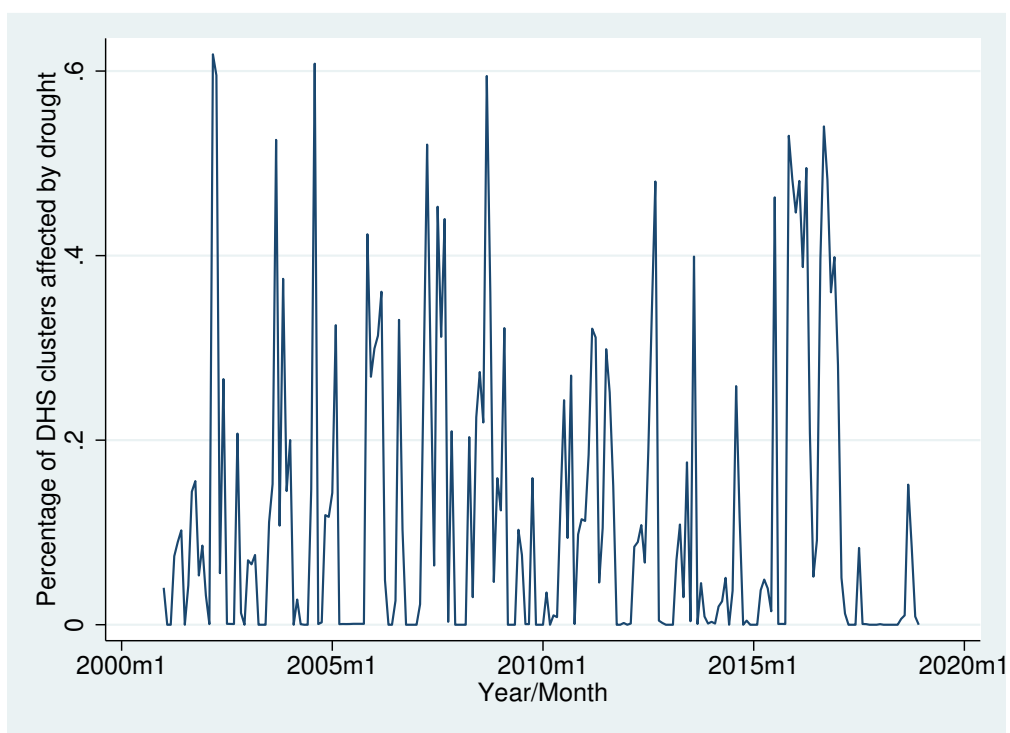


Figure 4.2: Percentage of monthly droughts across DHS clusters

Source: DHS and global SPEI dataset, Own computations.

4.4 Empirical Approach

To examine the effects of exposure to drought during pregnancy on child health outcomes, we follow the existing literature, which uses the following equation, exploiting variation in weather events over time within a given spatial area (Bratti et al., 2021; Le and Nguyen, 2021). This approach effectively controls for observed and time-invariant unobserved confounders, which adequately addresses potential sources of bias in estimating the causal impact of weather shocks on health outcomes. As misspecification risks are low, this approach allows for more simplicity and efficiency than other alternative models (e.g. Inverse Probability Weighted Regression Adjustment). The model equation we estimate by the Ordinary Least Squares (OLS) is:

$$H_{icdmy} = \alpha + \beta \text{drought}_{icdmy} + \psi X_{icdmy} \phi + \theta_c m + \eta_{dy} + \epsilon_{icdmy} \quad (4.1)$$

where H_{icdmy} is an outcome indicator for child i , residing in cluster c of district d , born in month m and in year y . We refer to cluster as village for shorthand. The outcomes considered in this study are children HAZ, WHZ, and physical growth. drought_{icdmy} counts the number of drought months preceding the nine months before the child's birth. To put it clearly, it captures the number of months with SPEI below -1 SD from the long-run average. β is the coefficient of interest, which measures the effect of an additional month of drought shocks during the gestational period. X_{icdmy} is a vector of controls for child and mother level characteristics, such as child's age, child birth order, whether the child is female, mother's age at birth, mother's years of schooling, and whether the household resides in urban areas.

Our empirical strategy compares the health outcomes of children exposed to intense drought against the outcomes of children who are exposed to less intense drought shock during their prenatal period. The number of drought months during pregnancy reflect the intensity of drought, as more months suggest a prolonged exposure to adverse drought conditions. Our identification strategy rests on the assumption that the occurrence of drought shocks is a plausibly random event within a given village over time. Under this assumption, exposure to drought during pregnancy is as good as random.

We also include a set of fixed effects to control for unobserved factors. Village-by-month of birth fixed

effects (θ_{cm}) account for seasonal patterns that may vary across survey villages. For example, children born in the rainy season may have different characteristics than children born during the dry season. Moreover, Studies have shown that a season of birth may confound the relationship between weather shocks and health outcomes (Strand et al., 2011; Buckles and Hungerman, 2013; Isen et al., 2017).

Second, district-specific linear time trend (η_{dy}) accounts for time-specific shocks such as economic crises that may vary over time by districts and affect future health outcomes. ϵ_{icdmy} are the error term. The error term is assumed to be identically and independently distributed (iid) across survey villages but correlated within each village due to spatial correlation of drought incidences for households in the same area. We therefore cluster standard errors at the village level.

Our empirical approach may face several challenges. First, a concern may arise if parents postpone fertility given a particular season of the year or in anticipation of weather shocks, resulting in selective fertility. We address this issue in the robustness checks by estimating the relation between drought exposure in utero and the birth interval. Likewise, a bias may occur if mortality happens during pregnancy or after birth since the sample consists of surviving children. In this case, our estimate would be biased toward zero. We address this selection bias in the robustness check by estimating the effect of drought in utero on child mortality.

4.5 Results

4.5.1 The effect of drought in utero on child health

Table 4.2 provides the OLS estimates of Equation (4.1) for the effect of drought exposure during pregnancy on child health outcomes. In the first two columns, the outcome variable is HAZ. WHZ and physical growth are the outcome variables in columns (3)-(4) and (4)-(5), respectively. For each outcome, the regression estimates include village-by-month fixed effects and district-specific linear time trend.

The results indicate a statistically significant reduction in child health from the exposure to drought during

pregnancy. Regarding the interpretation of estimated coefficients, column 2 of Table 4.2 (Panel A) suggests that one additional month of drought during pregnancy reduces a child's physical growth by 0.04 SD. This reduction of physical growth stems from the reduction in a child HAZ. For example, an additional month of drought in utero reduces a child's height by 0.07 SD. By contrast, we do not find a statistically significant effect on a child weight-based outcomes.

These findings are consistent with existing literature that documents a robust negative effect of weather shocks on a child long-term nutritional outcomes ([Bratti et al., 2021](#)). There are two plausible explanations for the lack of significance on weight-related outcomes. First, weight-for height captures short-term nutritional outcomes, and hence less sensitive to experience in utero. Second, parents may respond to exposure to shocks in utero and the child initial poor endowment via compensatory investments.

Taken together, these results suggest that exposure to drought during pregnancy has primarily a negative effect on a child long-term health outcomes. Furthermore, our findings underscores the prenatal period as a critical period where adverse conditions can have long-lasting effects on health outcomes ([Barker, 1990, 1995](#)).

Panel B of Table 4.2 reports the results for the effect of drought exposure by trimesters of pregnancy. The results indicate that the negative effect of drought occurs during the whole stage of pregnancy with regard to HAZ and physical growth. For WHZ, the negative effect of drought shocks appears only in the first trimester of pregnancy. For instance, Column (2) shows that an additional month of drought during the first trimester of pregnancy reduces a child physical growth by 0.05 SD, driven by both a decrease in a child HAZ and WHZ. This finding corroborates medical evidence of the sensitive period of prenatal development ([Glynn et al., 2001](#); [Lee et al., 2003](#); [Mulder et al., 2002](#)). On the one hand, the early stage of pregnancy is critical for the embryo development where major organ formation takes place ([Altshuler et al., 2003](#)). On the other hand, the later stage of pregnancy is sensitive to fetal development since the fetus experiences a rapid growth during this phase ([Harding and Bocking, 2001](#)).

Table 4.2: The effect of drought in utero on child health outcomes: OLS estimation results

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A			Main Results			
Months drought utero	-0.058*** (0.010)	-0.065*** (0.010)	-0.007 (0.008)	-0.008 (0.008)	-0.031*** (0.005)	-0.036*** (0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
R^2	0.370	0.380	0.336	0.345	0.368	0.378
PANEL B			Trimester analysis			
Months drought utero (1st trim)	-0.092*** (0.021)	-0.100*** (0.021)	-0.024 (0.016)	-0.028* (0.016)	-0.055*** (0.011)	-0.062*** (0.011)
Months drought utero (2nd trim)	-0.090*** (0.019)	-0.096*** (0.019)	-0.016 (0.017)	-0.017 (0.017)	-0.052*** (0.011)	-0.057*** (0.011)
Months drought utero (3rd trim)	-0.096*** (0.021)	-0.103*** (0.021)	0.008 (0.017)	0.008 (0.017)	-0.042*** (0.011)	-0.046*** (0.011)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
R^2	0.369	0.378	0.336	0.345	0.367	0.377
N	14106	14039	14090	14024	14128	14061

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero is the the number of months of drought in utero. Physical growth is the average of HAZ and WHZ. 1st trim is the first trimester of pregnancy. 2nd trim is the second trimester of pregnancy. 3rd trim is the third trimester of pregnancy. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth, child age in months, birth order. Source: own computations.

4.5.2 Mechanisms

4.5.2.1 The role of agricultural yields

While there may be several mechanisms at play, limitations in data prevent us from uncovering all the possible channels in more detail. As a result, this study focuses on crop yields as a key channel. Drought represents a significant threat to crop yields especially in settings where agriculture is predominantly rain-fed. In this context, lower crop yields may result in food shortage and thus reduced nutritional intake.

To examine the relationship between drought exposure and crop yields, we consider the following regression:

$$C_{dy} = \alpha + \varpi \text{drought}_{dy} + \lambda_d + \gamma_y + \epsilon_{dy} \quad (4.2)$$

Where C_{dy} is the logarithm of yield of a particular crop (Maize, beans, sorghum) during the growing season in a district d and in year y . λ_d are district fixed effects and γ_y are year fixed effects. drought_{dy} is the key explanatory variable, measured as the number of droughts during the growing period in a district d and year y . ϖ captures the effect of an additional month of drought during the growing period on the crop yields.

District fixed effects account for (unobserved) differences in crop yields at the district level, such as land suitability or quality. Year fixed effects account for all year-specific changes in crop yields common across district, such as availability of agricultural subsidy during a particular year. We do not need to control explicitly for seasonality as the focus is on a specific cropping period, as mentioned earlier. ϵ_{dy} is the error term. Standard errors are clustered at the district level.

Table 4.3 provides the regression estimates of months drought during growing season on major crop yields using equation (4.2). The results show that one additional month of drought during the growing period has a negative effect on crop yields. Column 1 indicates that an additional month of drought during the growing season is associated with 7% loss in maize yield. Likewise, an additional month of drought during the growing seasons reduces beans and sorghum yields by 6% loss, respectively (Columns (3) and (4)). These findings are in line with previous literature showing a detrimental effect of drought on agricultural productivity

in the context of Zambia (Ngoma et al., 2021; Matchaya et al., 2022). Moreover, our result suggests that the adverse effect of exposure to drought on health outcomes operate via crop failure, which may lead to malnutrition on a pregnant women and thereby, affecting child health outcomes.

Table 4.3: The effect of drought on crop yields: OLS estimation results

	Maize yield (1)	Bean Yield (2)	Sorghum yield (3)
Months drought in year	-0.069*** (0.014)	-0.056* (0.031)	-0.056** (0.027)
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
<i>N</i>	1208	997	888
<i>R</i> ²	0.738	0.293	0.343

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Zambia CSO. Robust standard errors clustered at the district level in parenthesis. The dependent variables are maize, bean and sorghum yields. Months drought in year is the number of droughts months episode during the growing period. Source: own computations.

4.5.2.2 The role of parental investments responses

While section 4.5.2.1 focuses on crop yields as the pathway through which drought can affect child health outcomes, it is possible that drought affect parental investment. These mediating factors may have an indirect effect on child health outcomes. Several studies point out that parents responds to fetal shocks and children's initial endowment²⁰ either through compensatory or reinforcing investments (Almond and Mazumder, 2013; Adhvaryu and Nyshadham, 2016).

On the one hand, drought exposure may result in income shocks, affecting parents ability to invest in their children health outcomes, especially in the context of credit constraints. Under these circumstances, parental investment would reinforce the initial effect of drought exposure during pregnancy. On the other hand, income shocks resulting from drought shocks may reduce the opportunity cost of parental time investment.

²⁰We also estimated the effect of prenatal drought on birth weight. The effect is negative but the estimates are less precise due to measurement errors and smaller samples. The results are not reported here, and are available upon request.

For instance, if parents anticipate a lower return to crop yields due to adverse weather conditions, they may engage in more child care activities, such as breastfeeding (Miller and Urdinola, 2010). In this regard, parental investment would result in compensating the initial negative effect of drought shocks on offspring outcomes.

To explore the mediating role of parental investment response, we estimate a version of equation (4.1) in which we replace the outcome variable with prenatal and postnatal investments, namely, prenatal checks by health professional, vaccination take-up and breastfeeding. Empirical evidence shows that prenatal care and vaccination are important health inputs, especially in a setting with limited access to basic health services (Okeke and Abubakar, 2020). Additionally, breastfeeding may represent an important input in the context of a limited food supply.

Table 4.4 captures the impact of drought in utero on prenatal checks, vaccination take-up and breastfeeding duration. We define vaccination take-up as the number of basic vaccination that a child needs to take during the early ages. Breastfeeding duration corresponds to the number of months the children were breast-fed by mother. Prenatal check is defined as a dummy variable whether for the mother receives prenatal care from a certified health professional. The results show that drought exposure during pregnancy is associated with increased parental investments (Panel A of Table 4.4). This finding provide a suggestive evidence of parental compensatory investments in the context of Zambia ²¹.

Panel B of Table 4.4 provides the results of gender heterogeneity analysis for the relationship between drought in utero and parental investment. Our finding shows no evidence of a heterogeneous effect in parental investment. This finding suggest that parents do not discriminate against a child gender in terms of resource allocation. The finding corroborate with empirical evidence, showing no evidence of gender preference in parental resource allocation in Zambia (Rossi and Rouanet, 2015).

²¹Our results provide suggestive evidence of compensatory investment, as limitations in data prevent us to disentangle whether parents respond specifically to prenatal drought shock that affects the child or to the observed child health after birth, such as low birth weight.

Table 4.4: The effect of drought in utero on parental investments: OLS estimation results

	Prenatal check (1)	Vaccination take-up (2)	Breastfeeding (3)
Panel A	Main estimates		
Months drought utero	0.003* (0.002)	0.299*** (0.021)	0.579*** (0.089)
Controls	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes
District-specific time trend	Yes	Yes	Yes
R^2	0.442	0.391	0.527
Panel B	Heterogeneity by gender		
Child male x months drought utero	0.000 (0.003)	0.007 (0.030)	0.103 (0.135)
Controls	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes
District-specific time trend	Yes	Yes	Yes
R^2	0.442	0.391	0.527
N	9599	15235	10849

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are prenatal check, vaccination take-up and breastfeeding. Prenatal check is a dummy of whether mothers receive prenatal care by a skilled health care professional. Vaccination take-up is the number of essential vaccines taken by children during the early ages. Breastfeeding is the number of months the child was breastfed by mothers. Month droughts utero is the number of months of drought in utero. The controls include a dummy for child child male, mother schooling, a dummy for urban, a dummy for the head for the the household male, mother's age at birth, child age in months, birth order.

Source: own computations.

4.5.2.3 The role of prenatal health care as mitigating factors

We further explore the role of prenatal care during pregnancy as a potential mitigating factor. We interact the month of drought episodes during pregnancy with a dummy variable for whether a woman receives prenatal care by a qualified healthcare professional. Table 4.5 provides the result after accounting for prenatal checks by a qualified healthcare professional. The results indicate that prenatal care use by a qualified healthcare professional may offset the adverse effect of drought shocks during pregnancy on a child health outcomes.

While prenatal care use may not be exogenous as it may result from parental responses (see Table 4.4), our finding highlights suggestive evidence on the importance of healthcare access during pregnancy, particularly in a resource-constrained country. Furthermore, our finding complements recent studies that document access to prenatal care as effective in mitigating the negative effects of weather shocks ([Banerjee and Maharaj, 2020](#); [Cohen and Dechezleprêtre, 2022](#)). It may be the case that women receive necessary treatment during prenatal care for drought induced physiological stress coupled with important information on preventing behavior that might attenuate the detrimental effect of drought on the yet unborn children.

Table 4.5: The effect of drought in utero on child health outcomes- Mitigation factors

	HAZ (1)	WAZ (2)	Physical growth (3)
Months drought utero	-0.148*** (0.042)	-0.021 (0.031)	-0.083*** (0.023)
Prenatal check	0.070 (0.115)	-0.038 (0.090)	0.017 (0.060)
Prenatal check x months drought utero	0.076* (0.042)	0.006 (0.031)	0.040* (0.024)
Controls	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes
District-specific time trend	Yes	Yes	Yes
<i>N</i>	8776	8761	8794
<i>R</i> ²	0.425	0.385	0.429

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero is the number of months of drought in utero. Physical growth is the average of HAZ and WHZ. Prenatal check is a dummy whether mother receives prenatal for a skilled health professional, 0 otherwise. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of household male, mother's age at birth, child age in months, birth order. Source: own computations.

4.6 Robustness and Heterogeneity analysis

4.6.1 Robustness checks

4.6.1.1 Alternative specifications

This section explores the degree to which the results are sensitive to different specifications. In Panel A of Table C1 (Appendix C), we use district-by-month fixed effects instead of village-by-month fixed effects to account for seasonal patterns. Additionally, we use a different level of clustering, that is, by the district-level. The results are similar to that of the baseline results, with drought exposure in utero yielding a statistically negative effect on a child HAZ and physical growth. Panel B of Table C1 repeats the same exercise using the mother fixed effect. The findings are indistinguishable from the baseline results in terms of significance and sign.

4.6.1.2 Alternative definitions of drought

In this section, we conduct a series of robustness checks to test that our baseline results are not sensitive to alternative definitions of drought. In Appendix C, Table C2, we show that an alternative definition of utero drought exposure yields similar findings to that of the baseline results in terms of sign and significance. We define drought exposure in utero as a dummy variable, taking the value 1 if a child experiences at least a drought episode during the gestational period. The findings indicate that drought exposure during pregnancy is associated with lower HAZ and physical growth. The effect is not statistically significant for a child weight outcomes. Table C3 repeats the same exercises using the number of drought months exposure during the growing period. The coefficient estimates remain stable in terms of magnitude and significance.

Table C4 considers the month of severe droughts exposure during pregnancy as the key explanatory variable. We define a severe drought as the number of months with SPEI ≤ -1.5 during pregnancy. The results are similar in terms of significance and magnitude. In Table C5, we replace drought exposure in utero

with the lifetime drought exposure, from pregnancy to the time of the survey. The measure is a dummy variable, taking the value of 1 if a child has experienced at least one drought episode over the course of his lifetime. The results hold in term of significance and magnitude.

Finally, in Table C6 we replicate the main results using drought with different SPEI month of scales: 1-month of scale and 12-month scale. Again, the sign and statistical significance are similar to the baseline results.

4.6.1.3 Additional checks

We perform other additional robustness checks in Appendix C. The sign and magnitude are very similar to the baseline when we replace the outcome variables using alternative definitions of child nutritional status (Table C7, Appendix C). The outcome variables are stunting, wasting and physical growth deficits, respectively. The findings show that exposure to drought in utero is associated with an increased prevalence of stunting and physical growth deficits among under-five age children. For example, column 2 of Table C7 indicates that one additional month of drought during pregnancy increases the prevalence of stunting by a 1.5 percentage point.

Table C8 in the Appendix reports the effect of drought in utero with the full sample. The results yield similar results to that of baseline estimate in terms of magnitude and significance. This finding implies that the effect of drought on those unable to migrate is not different from that of the whole sample.

4.6.1.4 Mortality and Fertility selection

Two major forms of sample selection may affect the baseline estimated results: mortality and fertility selection. Regarding mortality selection, some children may have died during the gestational period when exposed to drought shocks and others during the early years of life. If such an effect exists, our results may underestimate the true effect of drought exposure in utero on a child nutritional status, since our sample consists of survival children. We address the issue by estimating the relationship between prenatal drought exposure and child mortality both in utero and after birth. Table C9 in Appendix C indicates no evidence of drought exposure in

utero on pregnancy loss and postnatal mortality (Columns (1)-(3) of Table C10 in Appendix C). This finding implies that mortality selection does not contaminate the main results of this study.

Furthermore, adverse weather shocks may influence the decision of fertility ([Lam and Miron, 1996](#); [Hajdu and Hajdu, 2022](#)). For instance, a woman may decide to delay her childbirth in anticipation of adverse weather events, and weather events may be associated with a season of birth. Because of the close connection between drought and season, individuals may anticipate drought incidence during particular season of a year, such as dry season, where such phenomenon is likely to occur and adjust their fertility decision accordingly. In this regard, this selectivity may affect the estimated results. To test for the fertility selection, we estimate the relation between drought in utero and the preceding birth interval months. We find no evidence of association between prenatal drought exposure and the timing of conception (Column (4) of Table C9). This finding mitigates the concern of fertility selection as a potential bias to our main results.

4.6.2 Results from the heterogeneity analysis

In this section, we explore the heterogeneous effect of drought exposure during pregnancy by child gender, residence of location and recent birth. Panel A of Table C10 (Appendix C) presents the heterogeneous effect of drought by interacting the months of drought in utero with a dummy variable for whether a child is a boy. The results suggest no heterogeneous effect of drought exposure during pregnancy by child gender, even though boys appear to have poor nutrition relative to girls. The finding implies no evidence of parental gender bias in the context of Zambia ([Rossi and Rouanet, 2015](#)).

Panel B of Table C10 (Appendix C) reports the heterogeneous results of drought exposure in utero by residence location. Again, the results show no evidence of heterogeneous effect of drought exposure in utero on child health outcomes by residence of location. Taken together, the heterogeneity analysis suggests that the effects of drought exposure during pregnancy on child health are similar across child gender and residence of location.

Finally, Panel C of Table C10 (Appendix C) presents the heterogeneity result of prenatal drought exposure

by recent birth. We define recent birth as a dummy variable for whether the child is at most 24 months of age at the time of the survey. The results indicate that the negative effect of drought on child health outcomes is pronounced among younger children.

4.7 Conclusions

Drought poses an existential threat, especially in resource-constrained areas where rain-fed agriculture remains a primary source of subsistence. This study exploits spatio-temporal weather variation across the Zambian DHS cluster to capture the effect of drought exposure during pregnancy on child health outcomes.

The results show compelling evidence of the detrimental effect of exposure to drought during pregnancy on child health outcomes, particularly height. The effect of drought exposure in utero is less robust to child weight outcomes. This finding implies that exposure to drought during pregnancy has a latent effect on long-term nutritional outcomes. Our findings corroborate recent evidence that document a robust negative effect of weather shock on child height ([Le and Nguyen, 2021](#); [Bratti et al., 2021](#)). The negative effects are larger during the first and third trimesters of pregnancy. These findings cohere with medical evidence that documents the first and last trimesters as a highly sensitive period to fetal development ([Harding and Bocking, 2001](#); [Altshuler et al., 2003](#)).

Our heterogeneity analysis indicates that the effect of drought shocks on child health is similar across gender and residence location. We further show that drought during growing periods have a negative impact on major crop yields. In low-income settings heavily reliant on rain-fed agriculture, drought can increase the incidence of food insecurity ([Ngcamu and Chari, 2020](#)). This insight suggests that lower crop yields are the main pathway through which prenatal drought exposure affects child health outcomes. Moreover, our analysis indicates that drought during pregnancy result into greater parental investments. This finding provides a suggestive evidence of parental compensatory investment, consistent with empirical evidence of parental responses to health shock during early childhood ([Almond and Mazumder, 2013](#)).

Our analysis of potential mitigating factors suggests that prenatal care use by a qualified healthcare pro-

fessional may attenuate the adverse effect of in utero shocks. While this finding is suggestive, it provides a rationale for incorporating the health needs of pregnant women and children into climate adaptation and disaster-related policies. Investment in adaptation is critical to protect populations from the detrimental effects of extreme weather events. This insight is consistent with the findings that document better access to healthcare as an adaptation strategy to the adverse effect of weather shocks on health outcomes ([Banerjee and Maharaj, 2020](#); [Mullins and White, 2020](#); [Cohen and Dechezleprêtre, 2022](#)). Our study calls for further research on the role of adaptation strategy to adverse weather shocks on other human capital dimensions such as cognitive outcomes.

Chapter 5

General Conclusion

In the last decades, several countries have witnessed substantial progress in some health and education indicators, most especially rising life expectancy and school enrollment ([World Bank, 2024](#)). Yet, significant parts of the world are far from quality education and good health for all, particularly in low-income countries ([Grantham-McGregor et al., 2007](#); [Gust et al., 2024](#)). This thesis examines factors that affect health and education in the context of developing countries, which yields three distinct empirical topics.

Chapter 2 explores the extent of socioeconomic-related inequality in language skills among Peruvian children aged 5-15. This chapter complements evidence from existing literature from both developed and developing countries. The results indicate that socioeconomic-related inequality in children's cognitive skills is observable during the early ages and persists throughout adolescence, though declining. Parental socioeconomic status emerges as the main contributing factor to this observed socioeconomic-related inequality in language skills. This contribution is the product of inequality in wealth and parental wealth's influence on their children's language skills. This finding yields important policy implications. First, it highlights the importance of intervening early. Early public interventions can narrow socioeconomic inequality in language skills at the beginning of formal schooling ([Sánchez et al., 2020](#)). Second, the finding suggests that reducing socioeconomic inequality entails either reducing wealth inequality or the elasticity of parental wealth on children's language skills. However, there are some important trade-offs. For instance, reducing the elasticity of wealth

would decrease the average performance in language skills. Therefore, policy aimed at reducing socioeconomic inequality should consider the costs and benefits associated with each component of socioeconomic inequality.

Chapter 3 examines the relationship between childhood exposure to armed conflicts and schooling outcomes. This chapter extends the existing literature by using a conflict exposure measure at a finer geographical level. The result reveals a statistically insignificant relationship between armed conflict exposure during childhood and the probability of primary and lower secondary school completion. This finding suggests a community's resilience to conflict, particularly in the context of protracted armed conflict. Additionally, this finding indicates that the relationship between conflict exposure and educational outcomes is context specific, in line with existing empirical evidence ([Verwimp and Van Bavel, 2014](#)).

Chapter 4 focuses on the links between prenatal exposure to drought and child health outcomes. The results show an adverse effect of drought exposure during pregnancy on child health, particularly regarding height. The effects are similar across child gender and residence location. Lower crop yields emerge as the main pathway connecting drought and child health outcomes. Further results show that prenatal care use by a qualified healthcare professional may attenuate the adverse effect of drought on child health outcomes. This finding highlights the importance of considering the health needs of pregnant women and children into climate adaptation and disaster-related policies. Investment in adaptation is crucial to protect populations from the adverse effects of extreme weather events.

This thesis is not without limitations. The main limitation is related to data constraints. In chapter 2, the focus on the family environment may convey only a partial picture of the evolution of socioeconomic inequality across different childhood ages. Further research can incorporate the schooling environment into the analysis and disentangle the contribution of schools from family factors in explaining socioeconomic-related inequality in children's cognition.

In chapter 3, measurement errors in self-reported data may affect the estimated results, leading either to a downward or upward bias estimate. Moreover, the cross-sectional nature of the data restrict us from capturing conflict heterogeneity over time. Conflict intensity may vary over time across cohorts, especially in the context

of protracted armed conflicts. Future research should account for this timing effect using longitudinal data, if available, with better identification strategies. Moreover, accounting for the timing effect in conflict intensity would shed light on the heterogeneous effect of armed conflict on adult outcomes according to the age of exposure. Additionally, there is a need for further research to uncover specific channels connecting conflict exposure and educational outcomes, especially in the context of community resilience. While community resilience may mute the negative relationship between conflict exposure and educational attainment, its effects on the quality of schooling may be different. Further research should investigate the links between armed conflict exposure and the quality of schooling, such as learning outcomes in the context of protracted armed conflict.

Regarding chapter 4, future research should extend the analysis of prenatal exposure to drought by considering other dimensions of human capital, such as cognitive and socio-emotional development. There is also a need for further research on understanding the implications of public policy interventions in the relationship between weather shocks and human capital.

Bibliography

- Abiona, O. and J. B. Ajefu (2023). The impact of timing of in utero drought shocks on birth outcomes in rural households: Evidence from Sierra Leone. *Journal of Population Economics* 36(3), 1333–1362.
- Abufhele, A., D. Contreras, E. Puentes, A. Telias, and N. Valdebenito (2022). Socioeconomic gradients in child development: Evidence from a Chilean longitudinal study 2010–2017. *Advances in Life Course Research* 52, 100451.
- Adhvaryu, A. and A. Nyshadham (2016). Endowments at birth and parents' investments in children. *The Economic Journal* 126(593), 781–820.
- Aizer, A., L. Stroud, and S. Buka (2016). Maternal stress and child outcomes: Evidence from siblings. *Journal of Human Resources* 51(3), 523–555.
- Akbulut-Yuksel, M. (2014). Children of war: The long-run effects of large-scale physical destruction and warfare on children. *Journal of Human Resources* 49(3), 634–662.
- Akpalu, W., R. M. Hassan, et al. (2009). Climate variability and maize yield in South Africa: Results from GME and MELE methods. Discussion Paper 00843, IFPRI, Washington, DC.
- Akresh, R., S. Bhalotra, M. Leone, and U. Osili (2023). First-and second-generation impacts of the Biafran war. *Journal of Human Resources* 58(2), 488–531.
- Akresh, R., S. Bhalotra, M. Leone, and U. O. Osili (2012). War and stature: Growing up during the Nigerian civil war. *American Economic Review* 102(3), 273–277.

- Akresh, R., G. D. Caruso, and H. Thirumurthy (2022). Detailed geographic information, conflict exposure, and health impacts. *World Development* 155, 105–890.
- Akresh, R. and D. de Walque (2008). Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide. Technical report, IZA, Bonn.
- Alderman, H., J. Hoddinott, and B. Kinsey (2006a). Long Term consequences of early childhood malnutrition. *Oxford Economic Papers* 58(3), 450–474.
- Alderman, H., J. Hoddinott, and B. Kinsey (2006b). Long term consequences of early childhood malnutrition. *Oxford economic papers* 58(3), 450–474.
- Alfano, M. and T. Cornelissen (2021). Spatial Spillover Effects of Conflict: Propagation through Food Prices in Somalia. Technical report, Department of Economics University, University of Strathclyde, Glasgow.
- Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. *Journal of Political Economy* 114(4), 672–712.
- Almond, D. and J. Currie (2011a). Human Capital Development Before Age Five. In *Handbook of Labor Economics*, Number 4, pp. 1315–1486.
- Almond, D. and J. Currie (2011b). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives* 25(3), 153–172.
- Almond, D., J. Currie, and V. Duque (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature* 56(4), 1360–1446.
- Almond, D. and B. Mazumder (2013). Fetal origins and parental responses. *Annual Reviews of Economics* 5(1), 37–56.
- Altshuler, K., M. Berg, L. M. Frazier, J. Laurensen, J. Longstreth, W. Mendez, and C. A. Molgaard (2003). Critical periods in development. Ochp paper series, OCHP, Kansas.
- Andalón, M., J. P. Azevedo, C. Rodríguez-Castelán, V. Sanfelice, and D. Valderrama-González (2016). Weather shocks and health at birth in Colombia. *World Development* 82, 69–82.

- Andrabi, B. T., J. Das, A. I. Khwaja, and T. Zajonc (2016). Do Value-Added Estimates Add Value ? Accounting for Learning Dynamics. *American Economic Journal: Applied Economics* 3(3), 29–54.
- Armand, A. and I. Kim Taveras (2021). The Effect of Nature's Wealth on Human Development: Evidence from Renewable Resources. CEPR Discussion Paper 15680, CEPR, London and Paris.
- Auger, N., A. I. Naimi, A. Smargiassi, E. Lo, and T. Kosatsky (2014). Extreme heat and risk of early delivery among preterm and term pregnancies. *Epidemiology*, 344–350.
- Autesserre, S. (2010). *The Trouble with the Congo: Local Violence and the Failure of International Peacebuilding*. Cambridge Studies in International Relations. Cambridge: Cambridge University Press.
- Autor, D., D. Figlio, K. Karbownik, J. Roth, and M. Wasserman (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics* 11(3), 338–381.
- Autor, D., D. Figlio, K. Karbownik, J. Roth, and M. Wasserman (2023). Males at the tails: How socioeconomic status shapes the gender gap. *The Economic Journal* 133(656), 3136–3152.
- Bandyopadhyay, S., S. Kanji, and L. Wang (2012). The impact of rainfall and temperature variation on diarrheal prevalence in Sub-Saharan Africa. *Applied Geography* 33, 63–72.
- Banerjee, R. and R. Maharaj (2020). Heat, infant mortality, and adaptation: Evidence from India. *Journal of Development Economics* 143, 102378.
- Barker, D. J. (1990). The fetal and infant origins of adult disease. *BMJ: British Medical Journal* 301(6761), 1111.
- Barker, D. J. (1995). Fetal origins of coronary heart disease. *BMJ* 311(6998), 171–174.
- Barreca, A., K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro (2016). Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century. *Journal of Political Economy* 124(1), 105–159.

- Barreca, A. and J. Schaller (2020). The impact of high ambient temperatures on delivery timing and gestational lengths. *Nature Climate Change* 10(1), 77–82.
- Barreca, A. I. (2010). The long-term economic impact of in utero and postnatal exposure to malaria. *Journal of Human Resources* 45(4), 865–892.
- Beguería, S., S. M. Vicente-Serrano, F. Reig, and B. Latorre (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: Parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology* 34(10), 3001–3023.
- Behrman, J., D. Contreras, I. Palma, and E. Puentes (2017). Wealth Disparities for Early Childhood Anthropometrics and Skills: Evidence from Chilean Longitudinal Data. Working Paper 17-019, Penn Institute for Economic Research, Philadelphia.
- Bertoni, E., M. Di Maio, V. Molini, and R. Nisticò (2019). Education is forbidden: The effect of the Boko Haram conflict on education in North-East Nigeria. *Journal of Development Economics* 141, 102–249.
- Beshir, H. A. and J.-F. Maystadt (2020). In utero seasonal food insecurity and cognitive development: Evidence on gender imbalances from Ethiopia. *Journal of African Economies* 29(4), 412–431.
- Bharadwaj, P., J. Bietenbeck, P. Lundborg, and D.-O. Rooth (2019). Birth weight and vulnerability to a macroeconomic crisis. *Journal of Health Economics* 66, 136–144.
- Björklund, A. and K. G. Salvanes (2011). Education and Family Background. Mechanisms and Policies. In *Handbook of the Economics of Education*, Volume 3, pp. 201–247.
- Black, S. E., P. J. Devereux, and K. G. Salvanes (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics* 122(1), 409–439.
- Blattman, C. and J. Annan (2010). The Consequences of Child Soldiering. *The Review of Economics and Statistics* 92(4), 882–898.
- Blattman, C. and E. Miguel (2010). Civil War. *Journal of Economic Literature* 48(1), 3–57.

- Blau, D. M. (1999). The Effect of Income on Child Development. *The Review of Economics and Statistics* 81(2), 261–276.
- Bloomfield, F. H., M. H. Oliver, and J. E. Harding (2006). The late effects of fetal growth patterns. *Archives of Disease in Childhood - Fetal and Neonatal Edition* 91(4), F299–F304.
- Bouhlila, D. S. (2017). Parents' education and literacy skills: Evidence on inequality of socioeconomic status in Arab countries. *World Development Perspectives* 5, 34–43.
- Bozzoli, C. and C. Quintana-Domeque (2014). The weight of the crisis: Evidence from newborns in Argentina. *Review of Economics and Statistics* 96(3), 550–562.
- Bratti, M., P. B. Frimpong, and S. Russo (2021). Prenatal exposure to heat waves and child health in Sub-Saharan Africa. Discussion Paper 14424, IZA, Bonn.
- Brooks-Gunn, J. and G. J. Duncan (1997). The effects of Poverty on Children. *Future of Children* 7(2), 55–71.
- Brück, T., M. Di Maio, and S. H. Miaari (2019). Learning the hard way: The effect of violent conflict on student academic achievement. *Journal of the European Economic Association* 17(5), 1502–1537.
- Buckland, P. et al. (2005). Reshaping the future: Education and postconflict reconstruction. Technical report, The World Bank, Washington, DC. Available at <https://documents1.worldbank.org/curated/en/860731468138857154/pdf/344190PAPER0Re1010fficial0use0only1.pdf>.
- Buckles, K. S. and D. M. Hungerman (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics* 95(3), 711–724.
- Bundervoet, T. and S. Fransen (2018). The educational impact of shocks in utero: Evidence from Rwanda. *Economics and Human Biology* 29(2018), 88–101.
- Bütikofer, A., K. V. Løken, and K. G. Salvanes (2019). Infant health care and long-term outcomes. *Review of Economics and Statistics* 101(2), 341–354.
- Carrillo, B. (2020). Early rainfall shocks and later-life outcomes: Evidence from Colombia. *The World Bank Economic Review* 34(1), 179–209.

- Case, A. and C. Paxson (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy* 116(3), 499–532.
- Castro, J. F. and C. Rolleston (2018). The contribution of early childhood and schools to cognitive gaps: New evidence from Peru. *Economics of Education Review* 64, 144–164.
- Chamarbagwala, R. and H. E. Morán (2011). The human capital consequences of civil war: Evidence from Guatemala. *Journal of Development Economics* 94(1), 41–61.
- Chen, X., C. M. Tan, X. Zhang, and X. Zhang (2020). The effects of prenatal exposure to temperature extremes on birth outcomes: The case of China. *Journal of Population Economics* 33, 1263–1302.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society* 28(3), 591–605.
- Coghlan, B., T. Stewart, P. Ngoy, F. Mulumba, C. Hardy, J. Lewis, and R. Brennan (2007). Mortality in the Democratic Republic of Congo: An ongoing crisis. Technical report, International Rescue Committee, New York.
- Cohen, F. and A. Dechezleprêtre (2022). Mortality, temperature, and public health provision: Evidence from Mexico. *American Economic Journal: Economic Policy* 14(2), 161–192.
- Coleman, J. S., E. Campbell, C. Hobson, J. McParland, and Others (1966). Equality of educational opportunity. Technical report, Office of Education, U.S. Department of Health, Education, and Welfare, Washington, D.C. Available at <https://eric.ed.gov/?id=ED012275>.
- Collier, P. and A. Hoeffler (1998). On economic causes of civil war. *Oxford Economic Papers* 50(4), 563–573.
- Collier, P. and A. Hoeffler (2004a). Aid, policy and growth in post-conflict societies. *European Economic Review* 48(5), 1125–1145.
- Collier, P. and A. Hoeffler (2004b). Greed and grievance in civil war. *Oxford Economic Papers* 56(4), 563–595.
- Collier, P., A. Hoeffler, and D. Rohner (2009). Beyond greed and grievance: feasibility and civil war. *Oxford Economic Papers* 61(1), 1–27.

- Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives* 27(3), 79–102.
- CRED, UCLouvain, and USAID (2023). 2022 disasters in numbers. Technical report, Centre for Research on the Epidemiology of Disasters (CRED), Brussels. Available at <https://cred.be/disasters-numbers-2022>.
- Cuesta, J. and M. Leone (2020). Humanitarian Crises and Adolescent Well-Being: Knowledge, Gaps, and Prospects. *Journal of Economic Surveys* 34(1), 3–34.
- Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review* 97(2), 31–47.
- Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources* 43(4), 738–782.
- Cunha, F. and J. J. Heckman (2009). The Economics and Psychology of Inequality and Human Development. *Journal of the European Economic Association* 7(2), 320–364.
- Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). Chapter 12 Interpreting the Evidence on Life Cycle Skill Formation. In *Handbook of the Economics of Education*, Volume 1, pp. 697–812.
- Currie, J. and D. Almond (2011). Human capital development before age five. In *Handbook of Labor Economics*, Volume 4, pp. 1315–1486. Elsevier.
- Currie, J., M. Mueller-Smith, and M. Rossin-Slater (2022). Violence while in utero: The impact of assaults during pregnancy on birth outcomes. *Review of Economics and Statistics* 104(3), 525–540.
- Currie, J. and M. Rossin-Slater (2013). Weathering the storm: Hurricanes and birth outcomes. *Journal of Health Economics* 32(3), 487–503.
- Currie, J. and T. Vogl (2013). Early-life health and adult circumstance in developing countries. *Annual Review of Economics* 5, 1–36.

- Dadvand, P., X. Basagana, C. Sartini, F. Figueras, M. Vrijheid, A. De Nazelle, J. Sunyer, and M. J. Nieuwenhuijsen (2011). Climate extremes and the length of gestation. *Environmental Health Perspectives* 119(10), 1449–1453.
- Dai, A., T. Zhao, and J. Chen (2018). Climate change and drought: A precipitation and evaporation perspective. *Current Climate Change Reports* 4, 301–312.
- Das, J. and T. Zajonc (2010). India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics* 92(2), 175–187.
- Davenport, F., A. Dorélien, and K. Grace (2020). Investigating the linkages between pregnancy outcomes and climate in sub-Saharan Africa. *Population and Environment* 41, 397–421.
- Davis, D. R. and D. E. Weinstein (2002). Bones, Bombs, and Break Points: The geography of Economic Activity. *American Economic Review* 92(5), 1269–1289.
- De Herdt, T. and K. Titeca (2016). Governance with empty pockets: The education sector in the Democratic Republic of Congo. *Development and Change* 47(3), 472–494.
- De la Sierra, R. S. (2020). On the origins of the state: Stationary bandits and taxation in Eastern Congo. *Journal of Political Economy* 128(1), 32–74.
- De Oliveira, V. H., I. Lee, and C. Quintana-Domeque (2023). Natural disasters and early human development: Hurricane Catarina and infant health in Brazil. *Journal of Human Resources* 58(3), 819–851.
- De Onis, M. and A. W. Onyango (2008). WHO child growth standards. *The Lancet* 371(9608), 204.
- Deschenes, O. (2014). Temperature, human health, and adaptation: A review of the empirical literature. *Energy Economics* 46, 606–619.
- Deschênes, O., M. Greenstone, and J. Guryan (2009). Climate change and birth weight. *American Economic Review* 99(2), 211–217.
- Dimitrova, A. (2021). Seasonal droughts and the risk of childhood undernutrition in Ethiopia. *World Development* 141, 105417.

- Dinkelman, T. (2017). Long-run health repercussions of drought shocks: Evidence from South African homelands. *The Economic Journal* 127(604), 1906–1939.
- 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.
- Eck, K. (2012). In data we trust? A comparison of UCDP GED and ACLED conflict events datasets. *Cooperation and Conflict* 47(1), 124–141.
- EM-DAT (2023). Emergency Events Database. Available at <https://public.emdat.be/> (Accessed 2024.01.05).
- Epstein, A., T. Benmarhnia, and S. D. Weiser (2020). Drought and illness among young children in Uganda, 2009–2012. *The American Journal of Tropical Medicine and Hygiene* 102(3), 644.
- Erreygers, G. and T. Van Ourti (2011). Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice. *Journal of Health Economics* 30(4), 685–694.
- Fankhauser, S. (2017). Adaptation to climate change. *Annual Review of Resource Economics* 9, 209–230.
- Fernald, A., V. A. Marchman, and A. Weisleder (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science* 16(2), 234–248.
- Fernald, L. C., P. Kariger, M. Hidrobo, and P. J. Gertler (2012). Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal. *Proceedings of the National Academy of Sciences* 109(2), 17273–17280.
- Fernald, L. C., A. Weber, E. Galasso, and L. Ratsifandrihamanana (2011). Socioeconomic gradients and child development in a very low income population: Evidence from Madagascar. *Developmental Science* 14(4), 832–847.
- Filmer, D. and L. H. Pritchett (2001). Estimating Wealth Effects without Expenditure Data-or Tears : An Application to Educational Enrollments in States of India. *Demography* 38(1), 115–132.

- Filmer, D. and K. Scott (2008). Assessing Asset Indices. Policy Working Paper 4605, The World Bank Development Research Group, Washington, DC.
- Furnée, C. A., W. Groot, and H. M. van Den Brink (2008). The health effects of education: a meta-analysis. *European Journal of Public Health* 18(4), 417–421.
- Galasso, E., A. Weber, and L. C. Fernald (2019). Dynamics of Child Development: Analysis of a Longitudinal Cohort in a Very Low Income Country. *World Bank Economic Review* 33(1), 140–159.
- García, J. L., J. J. Heckman, and V. Ronda (2023). The Lasting Effects of Early-Childhood Education on Promoting the Skills and Social Mobility of Disadvantaged African Americans and Their Children. *Journal of Political Economy* 131(6), 1477–1506.
- Gertler, P., J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* 344(6187), 998–1001.
- Geruso, M. and D. Spears (2018). Heat, humidity, and infant mortality in the developing world. Working Paper 24870, NBER, Cambridge.
- Gleditsch, N. P., M. Eriksson, and H. Strand (2002). Armed Conflict 1946-2001 : A New Dataset. *Journal of Peace Research* 39(5), 615–637.
- Glewwe, P., Q. Huang, and A. Park (2017). Cognitive skills, noncognitive skills, and school-to-work transitions in rural China. *Journal of Economic Behavior & Organization* 134, 141–164.
- Glewwe, P., S. Krutikova, and C. Rolleston (2017). Do Schools Reinforce or Reduce Learning Gaps between Advantaged and Disadvantaged Students? Evidence from Vietnam and Peru. *Economic Development and Cultural Change* 65(4), 699–739.
- Glynn, L. M., P. D. Wadhwa, C. Dunkel-Schetter, A. Chiciz-DeMet, and C. A. Sandman (2001). When stress happens matters: Effects of earthquake timing on stress responsivity in pregnancy. *American Journal of Obstetrics and Gynecology* 184(4), 637–642.

- Grace, K., F. Davenport, H. Hanson, C. Funk, and S. Shukla (2015). Linking climate change and health outcomes: Examining the relationship between temperature, precipitation and birth weight in Africa. *Global Environmental Change* 35, 125–137.
- Grantham-McGregor, S., Y. B. Cheung, S. Cueto, P. Glewwe, L. Richter, and B. Strupp (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet* 369(9555), 60–70.
- Guarino, C. M., M. D. Reckase, and J. M. Wooldridge (2015). Can value-added measures of teacher performance be trusted? *Education Finance and Policy* 10(1), 117–156.
- Gust, S., E. A. Hanushek, and L. Woessmann (2024). Global universal basic skills: Current deficits and implications for world development. *Journal of Development Economics* 166, 103205.
- Ha, S. (2022). The changing climate and pregnancy health. *Current Environmental Health Reports* 9(2), 263–275.
- Ha, S., D. Liu, Y. Zhu, S. Soo Kim, S. Sherman, K. L. Grantz, and P. Mendola (2017). Ambient temperature and stillbirth: A multi-center retrospective cohort study. *Environmental Health Perspectives* 125(6), 067011.
- Hajdu, T. and G. Hajdu (2022). Temperature, climate change, and fertility. *Handbook of Labor, Human Resources and Population Economics*, 1–25.
- Hanushek, E. A., J. D. Light, P. E. Peterson, L. M. Talpey, and L. Woessmann (2022). Long-run Trends in the US SES–Achievement Gap. *Education Finance and Policy* 17(4), 608–640.
- Hanushek, E. A., P. E. Peterson, L. M. Talpey, and L. Woessmann (2019). The unwavering SES achievement gap: Trends in US student performance. Working Paper 25648, National Bureau of Economic Research, Cambridge, DC.
- Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review* 73, 103–130.
- Hanushek, E. A. and L. Woessmann (2008). Role of Cognitive Skills in Economic Development. *Journal of Economic Literature* 46(3), 607–668.

- Harding, R. and A. D. Bocking (2001). *Fetal growth and development*. Cambridge University Press.
- Hashim, S. A., T. J. Kane, T. Kelley-Kemple, M. E. Laski, and D. O. Staiger (2023). Have income-based achievement gaps widened or narrowed? Working Paper 27714, National Bureau of Economic Research, Cambridge, DC.
- Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24(3), 411–482.
- Hodler, R. (2019). The economic effects of genocide: Evidence from Rwanda. *Journal of African Economies* 28(1), 1–17.
- 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.
- IDMC (2022). Global Report on Internal Displacement 2022. Technical report, Internal Displacement Monitoring Centre. Available at <https://www.internal-displacement.org/publications/2022-global-report-on-internal-displacement>.
- INEI (2016). Perú: Síntesis Estadística 2016. Technical report, Instituto Nacional de Estadística e Informática, Lima. Available at https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1391/libro.pdf.
- IPCC (2023). IPCC Climate Change 2023: Synthesis Report, Summary for Policymakers. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Intergovernmental Panel on Climate Change (IPCC), Geneva.
- Isen, A., M. Rossin-Slater, and R. Walker (2017). Relationship between season of birth, temperature exposure, and later life wellbeing. *Proceedings of the National Academy of Sciences* 114(51), 13447–13452.
- Justino, P. (2012). Violent conflict and human capital accumulation. In G. K. Brown and A. Langer (Eds.), *Elgar Handbook of Civil War and Fragile States*, Chapter 13, pp. 187–200. Cheltenham: Edward Elgar.

- Justino, P. (2016). Supply and demand restrictions to education in conflict-affected countries: New research and future agendas. *International Journal of Educational Development* 47, 76–85.
- Justino, P., M. Leone, and P. Salardi (2014). Short-and long-term impact of violence on education: The case of Timor Leste. *World Bank Economic Review* 28(2), 320–353.
- Kakwani, N., A. Wagstaff, and E. van Doorslaer (1997). Socioeconomic inequalities in health: Measurement, computation, and statistical inference. *Journal of Econometrics* 77(1), 87–103.
- Kumar, S., R. Molitor, and S. Vollmer (2016). Drought and early child health in rural India. *Population and Development Review*, 53–68.
- Lam, D. A. and J. A. Miron (1996). The effects of temperature on human fertility. *Demography* 33(3), 291–305.
- Lange, G.-M., Q. Wodon, and K. Carey (2018). *The changing wealth of nations 2018: Building a sustainable future*. Washington, DC: World Bank Publications.
- Le, K. and M. Nguyen (2021). In-utero exposure to rainfall variability and early childhood health. *World Development* 144, 105485.
- Lee, B. E., E. H. Ha, H. S. Park, Y. J. Kim, Y. C. Hong, H. Kim, and J. T. Lee (2003). Exposure to air pollution during different gestational phases contributes to risks of low birth weight. *Human Reproduction* 18(3), 638–643.
- León, G. (2012). Civil conflict and human capital accumulation: The long-term effects of political violence in Perú. *Journal of Human Resources* 47(4), 991–1022.
- Lobell, D. B., W. Schlenker, and J. Costa-Roberts (2011). Climate trends and global crop production since 1980. *Science* 333(6042), 616–620.
- Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review* 99(3), 1006–1026.

- Mansour, H. and D. I. Rees (2012). Armed conflict and birth weight: Evidence from the al-aqsa intifada. *Journal of development Economics* 99(1), 190–199.
- Martins, L. and P. Veiga (2010). Do inequalities in parents' education play an important role in PISA students' mathematics achievement test score disparities? *Economics of Education Review* 29(6), 1016–1033.
- Matchaya, G. C., G. Tadesse, and A. N. Kuteya (2022). Rainfall shocks and crop productivity in Zambia: Implication for agricultural water risk management. *Agricultural Water Management* 269, 107648.
- Maystadt, J. F. and P. Verwimp (2014). Winners and losers among a refugee-hosting population. *Economic Development and Cultural Change* 62(4), 769–809.
- McKee, T. B., N. J. Doesken, J. Kleist, et al. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, Volume 17, pp. 179–183. California.
- Mei, Z. and L. M. Grummer-Strawn (2007). Standard deviation of anthropometric Z-scores as a data quality assessment tool using the 2006 WHO growth standards: A cross country analysis. *Bulletin of the World Health Organization* 85, 441–448.
- Meierrieks, D. (2021). Weather shocks, climate change and human health. *World Development* 138, 105228.
- Miguel, E. and G. Roland (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics* 96(1), 1–15.
- Miller, G. and B. P. Urdinola (2010). Cyclicalities, mortality, and the value of time: The case of coffee price fluctuations and child survival in Colombia. *Journal of Political Economy* 118(1), 113–155.
- Miller, R. (2017). Childhood health and prenatal exposure to seasonal food scarcity in Ethiopia. *World Development* 99, 350–376.
- Minoiu, B. C. and O. Shemyakina (2012). Child Health and Conflict in Côte d'Ivoire. *American Economic Review* 102(3), 294–299.
- Minoiu, C. and O. N. Shemyakina (2014). Armed conflict, household victimization, and child health in Côte d'Ivoire. *Journal of Development Economics* 108, 237–255.

- Molina, O. and V. Saldarriaga (2017). The perils of climate change: In utero exposure to temperature variability and birth outcomes in the andean region. *Economics & Human Biology* 24, 111–124.
- Mulder, E. J., P. R. De Medina, A. C. Huizink, B. R. Van den Bergh, J. K. Buitelaar, and G. H. Visser (2002). Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Human Development* 70(1-2), 3–14.
- Mullins, J. T. and C. White (2020). Can access to health care mitigate the effects of temperature on mortality? *Journal of Public Economics* 191, 104259.
- Neelsen, S. and T. Stratmann (2011). Effects of prenatal and early life malnutrition: Evidence from the Greek famine. *Journal of Health Economics* 30(3), 479–488.
- Neidhöfer, G. (2019). Intergenerational mobility and the rise and fall of inequality: Lessons from latin america. *The Journal of Economic Inequality* 17(4), 499–520.
- Ngcamu, B. S. and F. Chari (2020). Drought influences on food insecurity in Africa: A Systematic literature review. *International Journal of Environmental Research and Public Health* 17(16), 5897.
- Ngoma, H., P. Lupiya, M. Kabisa, and F. Hartley (2021). Impacts of climate change on agriculture and household welfare in Zambia: An economy-wide analysis. *Climatic Change* 167(3-4), 55.
- Nielsen, E. (2023). The Income-Achievement Gap and Adult Outcome Inequality. *Journal of Human Resources* 58(1), 0519–10220R3.
- OECD (2016). PISA 2015 Results (Volume I): Excellence and Equity in Education. Technical report, OECD, Paris. Available at https://www.oecd-ilibrary.org/education/pisa-2015-results-volume-i_9789264266490-en.
- Okeke, E. N. and I. S. Abubakar (2020). Healthcare at the beginning of life and child survival: Evidence from a cash transfer experiment in Nigeria. *Journal of Development Economics* 143, 102426.
- Pace, A., R. Luo, K. Hirsh-Pasek, and R. M. Golinkoff (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics* 3, 285–308.

- Paxson, C. and N. Schady (2007). Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health, and Parenting. *Journal of Human Resources* 42(1), 49–84.
- Persson, T., D. Strömberg, and M. Kudamatsu (2012). Weather and infant mortality in africa. Discussion Paper, CEPR, London and Paris.
- Pivovarova, M. and E. L. Swee (2015). Quantifying the microeconomic effects of war using panel data: Evidence from Nepal. *World Development* 66(2014), 308–321.
- Raleigh, C., A. Linke, H. Hegre, and J. Karlsen (2010). Introducing ACLED: An armed Conflict Location and Event Dataset. *Journal of Peace Research* 47(5), 651–660.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In *Whither opportunity?: Rising Inequality, Schools, and Children's Life Chances*, pp. 91–116. Russell Sage Foundation.
- Reynolds, S. A., C. Andersen, J. Behrman, A. Singh, A. D. Stein, L. Benny, B. T. Crookston, S. Cueto, K. Dear-den, A. Georgiadis, S. Krutikova, and L. C. Fernald (2017). Disparities in children's vocabulary and height in relation to household wealth and parental schooling: A longitudinal study in four low- and middle-income countries. *SSM - Population Health* 3, 767–786.
- Rocha, R. and R. R. Soares (2015). Water scarcity and birth outcomes in the Brazilian semiarid. *Journal of Development Economics* 112, 72–91.
- Rodríguez, C. and F. Sánchez (2012). Armed Conflict Exposure, Human Capital Investments, And Child Labor: Evidence From Colombia. *Defence and Peace Economics* 23(2), 161–184.
- Roemer, J. E. and A. Trannoy (2016). Equality of opportunity: Theory and Measurement. *Journal of Economic Literature* 54(4), 1288–1332.
- Rossi, P. and L. Rouanet (2015). Gender preferences in Africa: A comparative analysis of fertility choices. *World Development* 72, 326–345.

- Rossin-Slater, M. and M. Wüst (2020). What is the added value of preschool for poor children? Long-term and intergenerational impacts and interactions with an infant health intervention. *American Economic Journal: Applied Economics* 12(3), 255–286.
- Rubio-Codina, M., O. Attanasio, C. Meghir, N. Varela, and S. Grantham-McGregor (2015). The socioeconomic gradient of child development: Cross-sectional evidence from children 6-42 months in Bogota. *Journal of Human Resources* 50(2), 464–483.
- Sahn, D. E. and K. Villa (2016). Labor outcomes during the transition from adolescence to adulthood: The role of personality, cognition, and shocks in Madagascar. Discussion Paper 10359, IZA, Bonn.
- Sánchez, A., G. Meléndez, and J. R. Behrman (2020). Impact of the Juntos conditional cash transfer program on nutritional and cognitive outcomes in Peru: comparison between younger and older initial exposure. *Economic Development and Cultural Change* 68(3), 865–897.
- Sarmiento, J. H. (2023). Into the tropics: Temperature, mortality, and access to health care in Colombia. *Journal of Environmental Economics and Management* 119, 102796.
- Schady, N., J. Behrman, M. C. Araujo, R. Azuero, R. Bernal, D. Bravo, F. Lopez-Boo, K. Macours, D. Marshall, C. Paxson, and R. Vakis (2015). Wealth gradients in early childhood cognitive development in five Latin American countries. *Journal of Human Resources* 50(2), 446–463.
- Schulz, W. (2005). Measuring the Socio-Economic Background of Students and Its Effect on Achievement on PISA 2000 and PISA 2003. Available at <https://files.eric.ed.gov/fulltext/ED493510.pdf>. Paper prepared for the Annual Meetings of the American Educational Research Association, San Francisco.
- Schwandt, H. (2017). The lasting legacy of seasonal influenza: In-utero exposure and labor market outcomes. Discussion Paper 10589, IZA, Bonn.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics* 95(2), 186–200.

- Singh, A. (2015). Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics* 113, 16–32.
- SIPRI (2022). SIPRI Yearbook 2022: Armaments, Disarmament and International Security. Technical report, Stockholm International Peace Research Institute, Oxford.
- Song, X., Y. Song, and Y. Chen (2020). Secular trend of global drought since 1950. *Environmental Research Letters* 15(9), 094073.
- Spinoni, J., P. Barbosa, A. De Jager, N. McCormick, G. Naumann, J. V. Vogt, D. Magni, D. Masante, and M. Mazzeschi (2019). A new global database of meteorological drought events from 1951 to 2016. *Journal of Hydrology: Regional Studies* 22, 100593.
- Strand, L. B., A. G. Barnett, and S. Tong (2011). The influence of season and ambient temperature on birth outcomes: A review of the epidemiological literature. *Environmental Research* 111(3), 451–462.
- Sundberg, R. and E. Melander (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research* 50(4), 523–532.
- Swee, E. L. (2015). On war intensity and schooling attainment: The case of Bosnia and Herzegovina. *European Journal of Political Economy* 40(2015), 158–172.
- Titeca, K. and T. De Herdt (2011). Real governance beyond the ‘failed state’: Negotiating education in the Democratic Republic of the Congo. *African affairs* 110(439), 213–231.
- Todd, P. and K. I. Wolpin (2007). The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps. *Journal of Human Capital* 1(1), 91–136.
- Todd, P. E. and K. I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. *Economic Journal* 113(485), 3–33.
- Torche, F. (2018). Prenatal exposure to an acute stressor and children’s cognitive outcomes. *Demography* 55(5), 1611–1639.

- Trefon, T. (2009). Public service provision in a failed state: looking beyond predation in the Democratic Republic of Congo. *Review of African Political Economy* 36(119), 9–21.
- Turner, T. (2007). *The Congo wars: Conflict, myth and reality*. New York: Zed Books.
- UIS (2022). UNESCO Institute for Statistics. Available at <http://data.uis.unesco.org/?lang=fr> (Accessed 2022.05.12).
- Unfried, K. and K. Kis-Katos (2023). The heterogeneous effects of conflict on education: A spatial analysis in Sub-Saharan Africa. *Journal of Peace Research* 60(6), 968–984.
- United Nations (2024a). The 2030 Agenda for Sustainable Development. Available at <https://sdgs.un.org/goals>.
- United Nations (2024b). The 2030 Agenda for Sustainable Development. Available at <https://www.un.org/millenniumgoals/>.
- Utsumi, Y. (2022). Armed conflict, education access, and community resilience: Evidence from the afghanistan nrva survey 2005 and 2007. *International Journal of Educational Development* 88, 102512.
- Valente, C. (2014). Education and civil conflict in Nepal. *World Bank Economic Review* 28(2), 354–383.
- Valente, C. (2015). Civil conflict, gender-specific fetal loss, and selection: a new test of the trivers–willard hypothesis. *Journal of health economics* 39, 31–50.
- Van der Windt, P. and M. Humphreys (2016). Crowdsourcing in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time. *Journal of Conflict Resolution* 60(4), 748–781.
- Van Doorslaer, E., A. Wagstaff, H. Bleichrodt, S. Calonge, U. G. Gerdtham, M. Gerfin, J. Geurts, L. Gross, U. Häkkinen, R. E. Leu, O. O'Donnell, C. Propper, F. Puffer, M. Rodríguez, G. Sundberg, and O. Winkelhake (1997). Income-related inequalities in health: Some international comparisons. *Journal of Health Economics* 16(1), 93–112.
- Verwimp, P., P. Justino, and T. Brück (2019). The microeconomics of violent conflict. *Journal of Development Economics* 141(October 2018), 102–297.

- Verwimp, P. and J. Van Bavel (2014). Schooling, violent conflict, and gender in burundi. *The World Bank Economic Review* 28(2), 384–411.
- 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.
- Vicente-Serrano, S. M., S. Beguería, J. Lorenzo-Lacruz, J. J. Camarero, J. I. López-Moreno, C. Azorin-Molina, J. Revuelto, E. Morán-Tejeda, and A. Sanchez-Lorenzo (2012). Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interactions* 16(10), 1–27.
- Vogel, E., M. G. Donat, L. V. Alexander, M. Meinshausen, D. K. Ray, D. Karoly, N. Meinshausen, and K. Frieler (2019). The effects of climate extremes on global agricultural yields. *Environmental Research Letters* 14(5), 054010.
- Von Stumm, S., K. Rimfeld, P. S. Dale, and R. Plomin (2020). Preschool verbal and nonverbal ability mediate the association between socioeconomic status and school performance. *Child Development* 91(3), 705–714.
- Wagstaff, A., O. O'Donnell, E. Van Doorslaer, and M. Lindelow (2007). *Analyzing health equity using household survey data: A guide to techniques and their implementation*. Washington DC: World Bank Publications.
- Wagstaff, A., E. Van Doorslaer, and N. Watanabe (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics* 112(1), 207–223.
- Weisleder, A. and A. Fernald (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science* 24(11), 2143–2152.
- World Bank (2005). *Education in the Democratic Republic of Congo: priorities and options for regeneration*. Washington, DC: The World Bank. Available at <https://openknowledge.worldbank.org/entities/publication/1200c247-66f7-56aa-b338-3ccf7f24f9e0>.

- World Bank (2011). World development report 2011 : Conflict, security, and development. Technical report, The World Bank, Washington, D.C. Available at <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/806531468161369474/world-development-report-2011-conflict-security-and-development-overview>.
- World Bank (2015). Public expenditure review of the education sector in the Democratic Republic of Congo. An efficiency, effectiveness and equity analysis. Technical report, The World Bank, Washington, D.C. Available at <https://openknowledge.worldbank.org/entities/publication/4ec86aec-ab69-5970-ae27-e2b9521c21b9>.
- World Bank (2018). World Development Report 2018 : Learning to Realize Education's Promise. Technical report, The World Bank, Washington, DC. Available at <https://www.worldbank.org/en/publication/wdr2018>.
- World Bank (2021). World Development Indicators. Available at <https://databank.worldbank.org/source/world-development-indicators> (Accessed 2021.06.01).
- World Bank (2022). World Development Indicators. Available at <https://databank.worldbank.org/source/world-development-indicators> (Accessed 2022.11.11).
- World Bank (2024). World Development Indicators. Available at <https://databank.worldbank.org/source/world-development-indicators> (Accessed 2024.02.05).
- World Food Programme (2023). WFP Zambia Country Brief. Available at <https://www.wfp.org/countries/zambia>.

Appendix A: Additional tables to chapter 2

Table A1: Concentration index computed with household expenditure per capita

	<i>CI</i>
Age 5	0.052*** (0.002)
Age 8	0.051*** (0.003)
Age 12	0.043*** (0.002)
Age 15	0.036*** (0.003)

Notes: CI is estimated based on equation (2.2), that is the point estimates for the CI and the standard errors are obtained using a convenient regression approach ([Kakwani et al., 1997](#)). Robust standard errors in parenthesis. Significance levels: *, 10%; **, 5%; ***, 1%. The language achievement is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is the household expenditures per capita in log form. Source: Peruvian Young Lives survey; own computations.

Table A2: Concentration index computed with Parental education

	<i>CI</i>
Age 5	0.064*** (0.002)
Age 8	0.067*** (0.002)
Age 12	0.066*** (0.002)
Age 15	0.061*** (0.002)

Notes: CI is estimated based on equation (2.2), that is the point estimates for the CI and the standard errors are obtained using a convenient regression approach (Kakwani et al., 1997). Robust standard errors in parenthesis. Significance levels: *, 10%; **, 5%; ***, 1%. The language achievement is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is the average parental years of education. Source: Peruvian Young Lives survey; own computations.

Table A3: Difference in Socioeconomic-related inequality in language achievement between preschool and school ages

	Differences in <i>CI</i>
Age 8 and Age 5	-0.002 (0.004)
Age 12 and Age 5	-0.009*** (0.003)
Age 15 and Age 5	-0.017*** (0.004)

Notes: CI is estimated based on equation (2.2), that is, the point estimates for the CI and the standard errors are obtained using a convenient regression approach (Kakwani et al., 1997). Differences in CI are tested based on the chow test of equality of coefficients (Chow, 1960). Robust standard errors in parenthesis. Significance levels: *, 10%; **, 5%; ***, 1%. The language achievement is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is household expenditures in log form. Source: Peruvian Young Lives survey; own computations.

Table A4: Differences in Socioeconomic-related inequality in language achievement between preschool and school ages

	Differences in <i>CI</i>
Age 8 and Age 5	0.003 (0.004)
Age 12 and Age 5	0.002 (0.003)
Age 15 and Age 5	0.003 (0.003)

Notes: CI is estimated based on equation (2.2), that is, the point estimates for the CI and the standard errors are obtained using a convenient regression approach (Kakwani et al., 1997). Differences in CI are tested based on the chow test of equality of coefficients (Chow, 1960). Robust standard errors in parenthesis
Significance levels: *, 10%; **, 5%; ***, 1%. The language achievement is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is the average of parental education.
Source: Peruvian Young Lives survey; own computations.

Table A5: Decomposition socioeconomic-related inequality in PPVT scores–Household expenditures

	Age 5			Age 8			Age 12			Age 15		
	η_k (1)	c_k (2)	cont (3)	η_k (4)	c_k (5)	cont (6)	η_k (7)	c_k (8)	cont (9)	η_k (10)	c_k (11)	cont (12)
Log expenditures	0.4382	0.0736	0.6172	0.5082	0.0652	0.6547	0.3594	0.0648	0.5389	0.3254	0.0617	0.5634
Male	0.0029	0.0057	0.0003	0.0086	-0.0157	-0.0027	0.0198	0.0039	0.0018	0.0215	-0.0067	-0.0040
Child stunted	-0.0182	-0.2741	0.0954	-0.0128	-0.2582	0.0655	-0.0129	-0.2665	0.0795	-0.0087	-0.2538	0.0620
First born	0.0081	0.1556	0.0242	0.0099	0.1732	0.0340	0.0133	0.1194	0.0369	0.0174	0.1007	0.0492
Child attended preschool=1	0.0865	0.0251	0.0416	0.0588	0.0259	0.0301	0.0537	0.0261	0.0324	0.0629	0.0178	0.0314
Mother age	0.0751	-0.0055	-0.0079	0.0338	-0.0055	-0.0037	0.0748	-0.0030	-0.0052	0.0879	-0.0028	-0.0068
Small household size	-0.0091	0.1642	-0.0285	0.0032	0.1848	0.0117	0.0173	0.1699	0.0680	0.0088	0.1443	0.0357
Head of household female	-0.0000	0.0807	-0.0000	0.0019	0.0440	0.0016	-0.0010	0.0737	-0.0017	0.0009	0.0495	0.0013
Urban	0.1234	0.0965	0.2280	0.1156	0.0864	0.1974	0.1103	0.0862	0.2200	0.1036	0.0647	0.1881
Both parents alive	0.0106	0.0021	0.0004	0.0684	0.0039	0.0053	0.0076	-0.0000	-0.0000	-0.0314	-0.0000	-0.0000

Notes: The decomposition of the CI is computed based on Eq (2.5). η_k = elasticity defined as $\frac{\partial k}{\partial A}(\bar{x}_k)$; c_k =concentration index for x_k ; cont= contribution to CI. The outcome variable is language achievement, measured by PPVT with mean 500 and 100 SD. Source: Peruvian Young Lives survey; Own computations.

Table A6: Decomposition of socioeconomic-related inequality in PPVT scores–Parental education

	Age 5			Age 8			Age 12			Age 15		
	η_k (1)	$\hat{c}_k$ (2)	cont (3)	η_k (4)	c_k (5)	cont (6)	η_k (7)	c_k (8)	cont (9)	η_k (10)	c_k (11)	cont (12)
Parental average education	0.1410	0.2677	0.5917	0.1524	0.2795	0.6386	0.1568	0.2795	0.6650	0.1547	0.2718	0.6870
Male	0.0021	0.0134	0.0004	0.0078	0.0012	0.0001	0.0201	0.0060	0.0018	0.0204	0.0061	0.0020
Child stunted	-0.0163	-0.3292	0.0841	-0.0103	-0.3517	0.0544	-0.0099	-0.3530	0.0531	-0.0067	-0.3346	0.0365
First born	0.0061	0.1784	0.0172	0.0076	0.1967	0.0223	0.0061	0.1978	0.0182	0.0079	0.1965	0.0254
Child attended preschool	0.0692	0.0308	0.0334	0.0433	0.0337	0.0219	0.0338	0.0347	0.0178	0.0317	0.0342	0.0177
Mother age	0.0828	-0.0074	-0.0096	0.0555	-0.0108	-0.0090	0.0836	-0.0117	-0.0148	0.0894	-0.0103	-0.0151
Small household size	0.0026	0.1054	0.0043	0.0167	0.1201	0.0300	0.0215	0.1062	0.0347	0.0117	0.0847	0.0162
Head of household female	0.0007	0.0585	0.0007	0.0026	0.0233	0.0009	0.0003	0.0409	0.0002	0.0025	-0.0023	-0.0001
Urban	0.1001	0.1385	0.2172	0.0834	0.1487	0.1858	0.0737	0.1497	0.1674	0.0665	0.1497	0.1628
Both parents alive	0.0350	0.0012	0.0006	0.0774	0.0041	0.0048	0.0049	-0.0001	-0.0000	0.0195	-0.0006	-0.0002

Notes: The decomposition of the CI is computed based on Equation (2.5). η_k = elasticity defined as $\frac{\partial x_k}{\partial A}(\bar{x}_k)$; c_k = concentration index for x_k ; cont = contribution to CI. The outcome variable is language achievement, measured by PPVT with mean 500 and 100 SD. Source: Peruvian Young Lives survey; Own computations.

Table A7: Decomposition of value added—with PPVT scores age 8

	η_k (1)	c_k (2)	Cont (3)
Wealth index	0.0490	0.1503	0.1589
PPVT age 8	0.5840	0.0588	0.7412
Male	0.0156	0.0010	0.0004
First born	0.0048	0.0955	0.0099
Child stunted	-0.0034	-0.3018	0.0222
Child attended preschool	0.0151	0.0260	0.0085
Mother's age	0.0374	0.0004	0.0003
Small household size	0.0083	0.0447	0.0080
Head of household female	-0.0010	-0.0431	0.0009
Urban	0.0084	0.1631	0.0296
Both parents alive	-0.0015	-0.0001	0.0000

Notes: The decomposition to CI based on equation (2.5), accounting for the PPVT at age 8.

PPVT measures the language achievement and is normalized with mean 500 and SD 100.

η_k =elasticity defined as $\frac{\beta_k}{A}(\bar{x}_k)$, c_k = the concentration index for x_k . cont= contribution to *CI*. Source: Peruvian Young Lives survey; own computations.

Table A8: Socioeconomic-related inequality in language achievement with Quechua speaking student

	<i>CI</i>
Age 5	0.074*** (0.002)
Age 8	0.066*** (0.002)
Age 12	0.062*** (0.002)
Age 15	0.055*** (0.002)

Notes: CI is based on equation (2.2). The point estimates for the CI and the standard errors are obtained using a convenient regression approach (Kakwani et al., 1997). Robust standard errors in parenthesis Significance levels: *, 10%; **, 5%; ***, 1%. The language achievement is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is household wealth. The sample includes children assessed in Quechua.

Source: Peruvian Young Lives survey; own computations.

Table A9: Decomposition of socioeconomic-related inequality in PPVT scores with quechua speaking

	Age 5			Age 8			Age 12			Age 15		
	η_k (1)	c_k (2)	cont (3)	η_k (4)	c_k (5)	cont (6)	η_k (7)	c_k (8)	cont (9)	η_k (10)	c_k (11)	cont (12)
Wealth index	0.2023	0.2791	0.7582	0.2150	0.2174	0.7050	0.2299	0.1835	0.6595	0.2376	0.1603	0.6922
Male	0.0030	0.0120	0.0005	0.0089	-0.0003	-0.0000	0.0181	0.0108	0.0030	0.0152	0.0081	0.0022
Child stunted	-0.0103	-0.3450	0.0475	-0.0098	-0.3738	0.0552	-0.0101	-0.4001	0.0631	-0.0084	-0.3170	0.0483
First born	0.0083	0.1271	0.0142	0.0111	0.1183	0.0199	0.0110	0.1198	0.0206	0.0145	0.1103	0.0290
Child attended preschool=1	0.0376	0.0316	0.0160	0.0509	0.0292	0.0224	0.0459	0.0286	0.0205	0.0533	0.0280	0.0272
Mother age	0.0521	0.0026	0.0018	0.0212	0.0021	0.0007	0.0415	-0.0017	-0.0011	0.0533	-0.0023	-0.0023
Small household size=1	0.0152	0.0877	0.0179	0.0139	0.1115	0.0233	0.0256	0.0857	0.0344	0.0223	0.0567	0.0230
Head of household female=1	0.0017	0.0536	0.0012	0.0018	0.0577	0.0016	0.0013	-0.0116	-0.0002	0.0026	-0.0271	-0.0013
Urban	0.0416	0.2411	0.1348	0.0496	0.2184	0.1635	0.057	0.2078	0.1776	0.0499	0.2007	0.1822
Both parents alive	-0.0054	-0.0067	0.0005	0.0635	0.0046	0.0044	0.0443	-0.0009	-0.0006	-0.0086	0.0004	-0.0001

Notes: The decomposition of the CI is computed based on Equation (2.5). η_k = elasticity defined as $\frac{\beta_k}{A}(\bar{x}_k)$; c_k =concentration index for x_k ; cont= contribution to CI. The outcome variable is language achievement, measured by PPVT with mean 500 and 100 SD. The sample includes children assessed in Quechua. Source: Peruvian Young Lives survey; Own computations.

Table A10: Socioeconomic-related inequality in language achievement with non-attributing children

	<i>CI</i>
Age 5	0.071*** (0.002)
Age 8	0.066*** (0.003)
Age 12	0.059*** (0.003)
Age 15	0.044*** (0.003)

Notes: CI is based on equation (2.2). The point estimates for the CI and the standard errors are obtained using a convenient regression approach ([Kakwani et al., 1997](#)). Robust standard errors in parenthesis .Significance levels: *, 10%; **, 5%; ***, 1%. The language achievement is measured by PPVT standardized with mean 500 and SD 100. The socioeconomic status measure is household wealth. The sample is a balanced panel of children aged 5 to 15.

Source: Peruvian Young Lives survey; own computations.

Table A11: Decomposition of socioeconomic-related inequality in PPVT scores with non-attriting children

	Age 5			Age 8			Age 12			Age 15		
	η_k	$\hat{c}_k$	cont	η_k	$\hat{c}_k$	cont	η_k	c_k	cont	η_k	c_k	cont
Wealth index	0.2041	0.2537	0.7328	0.2709	0.1942	0.7953	0.2796	0.1611	0.7614	0.2308	0.1411	0.7354
Male	0.0033	0.0137	0.0006	0.0079	0.0062	0.0007	0.0211	0.0138	0.0049	0.0195	0.0082	0.0036
Child stunted	-0.0120	-0.3702	0.0628	-0.0068	-0.3976	0.0411	-0.0061	-0.4000	0.0412	-0.0062	-0.2819	0.0397
First born	0.0118	0.1162	0.0194	0.0116	0.0881	0.0155	0.0126	0.1036	0.0220	0.0147	0.0962	0.0320
Child attended preschool=1	0.0343	0.0245	0.0119	0.0157	0.0250	0.0059	0.0616	0.0163	0.0170	0.0665	0.0197	0.0296
Mother age	0.0831	0.0063	0.0074	0.0288	0.0060	0.0026	0.0799	0.0016	0.0022	0.0815	0.0008	0.0015
Small household size	0.0135	0.0714	0.0137	0.0218	0.0809	0.0267	0.0241	0.0675	0.0276	0.0161	0.0471	0.0171
Head of household female	0.0004	0.0450	0.0002	0.0009	0.0505	0.0007	0.0031	-0.0438	-0.0023	0.0022	-0.0613	-0.0031
Urban	0.0612	0.1667	0.1445	0.0489	0.1444	0.1067	0.0480	0.1349	0.1094	0.0347	0.1355	0.1063
Both parents alive	-0.0071	0.0023	-0.0002	0.0222	0.0037	0.0012	0.1162	-0.0013	-0.0025	-0.0549	-0.0005	0.0006

Notes: The decomposition of the CI is computed based on equation (2.5). η_k = elasticity defined as $\frac{\beta_k}{A}(\bar{x}_k)$; c_k =concentration index for x_k ; cont= contribution to CI. The outcome variable is language achievement, measured by PPVT with mean 500 and 100 SD. The sample is a balanced panel of children aged 5 to 15. Source: Peruvian Young Lives survey; Own computations.

Table A12: Correlation between wealth index and other measures of socioeconomic status

	(1) Wealth Index
Parental average education	0.593***
Log expenditure	0.568***
<i>N</i>	7183

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Peruvian Young Lives survey; Own computations.

Appendix B: Additional tables and figures to chapter 3

Table B1: Violent conflict exposure and schooling outcomes

	Primary schooling		Lower secondary	
	(1)	(2)	(3)	(4)
Violent conflict exposure	-0.032*** (0.009)	-0.008 (0.010)	-0.030*** (0.008)	-0.001 (0.007)
Violent conflict exposure*age 6-11	0.002 (0.007)	0.002 (0.007)		
Urban	0.132*** (0.023)	0.121*** (0.020)	0.110*** (0.027)	0.095*** (0.021)
Wealth index	0.167*** (0.010)	0.189*** (0.011)	0.158*** (0.010)	0.187*** (0.009)
gender	0.241*** (0.010)	0.243*** (0.010)	0.261*** (0.011)	0.264*** (0.011)
Violent conflict exposure*age 12-13			0.001 (0.014)	-0.004 (0.013)
Province FE		Yes		Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy		Yes		Yes
<i>N</i>	22551	22551	13793	13793

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Violent conflict exposure is the weighted average number of violent conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: DHS, Own computations.

Table B2: Non-violent conflict exposure and schooling outcomes

	Primary schooling		Lower secondary	
	(1)	(2)	(3)	(4)
Non-violent conflict exposure	-0.024*** (0.006)	-0.017*** (0.004)	-0.013* (0.007)	-0.003 (0.004)
Non-violent conflict exposure*age 6-11	0.000 (0.006)	0.003 (0.008)		
Urban	0.128*** (0.024)	0.119*** (0.020)	0.107*** (0.028)	0.094*** (0.020)
Wealth index	0.169*** (0.011)	0.191*** (0.011)	0.158*** (0.010)	0.189*** (0.009)
Male	0.241*** (0.009)	0.243*** (0.010)	0.261*** (0.011)	0.264*** (0.011)
Non-violent conflict exposure*age 12-13			-0.013** (0.006)	-0.018*** (0.005)
Province FE		Yes		Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy		Yes		Yes
<i>N</i>	22551	22551	13793	13793

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Non-violent conflict exposure is the weighted average number of non-violent conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: DHS, Own computations.

Table B3: Conflict exposure and schooling outcomes –gender

	Primary schooling		Lower secondary	
	Male (1)	Female (2)	Male (3)	Female (4)
Conflict exposure	-0.001 (0.017)	-0.018** (0.008)	-0.001 (0.019)	-0.003 (0.006)
Conflict exposure*age 6-11	-0.008 (0.009)	0.008 (0.009)		
Urban	0.088*** (0.024)	0.134*** (0.021)	0.076** (0.031)	0.102*** (0.021)
Wealth index	0.129*** (0.021)	0.207*** (0.011)	0.188*** (0.017)	0.184*** (0.010)
Conflict exposure*age 12-13			0.000 (0.025)	-0.015 (0.014)
Province FE	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy	Yes	Yes	Yes	Yes
<i>N</i>	6413	16138	3998	9795

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. DID specification. Source: DHS, Own computations.

Table B4: Conflict exposure and schooling outcomes – location of residence

	Primary schooling		Lower secondary	
	Urban (1)	Rural (2)	Urban (3)	Rural (4)
Conflict exposure	-0.011* (0.006)	0.005 (0.020)	-0.007 (0.005)	0.016** (0.006)
Conflict exposure*age 6-11	0.010 (0.006)	-0.016 (0.013)		
Wealth index	0.138*** (0.009)	0.297*** (0.029)	0.194*** (0.008)	0.231*** (0.028)
Male	0.126*** (0.013)	0.316*** (0.012)	0.223*** (0.016)	0.281*** (0.014)
Conflict exposure*age 12-13			-0.006	-0.047
Province FE	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy	Yes	Yes	Yes	Yes
<i>N</i>	9478	13073	5664	8129

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: DHS, Own computations.

Table B5: Conflict exposure and probability of migrating

	migration (1)	migration (2)
Conflict exposure	0.003 (0.011)	0.015 (0.011)
Conflict exposure*age 6-11	0.013 (0.020)	
Urban	0.025 (0.038)	0.011 (0.041)
Wealth index	-0.010 (0.015)	-0.013 (0.018)
gender	-0.032* (0.017)	-0.014 (0.019)
Conflict exposure*age 12-13		-0.029*** (0.011)
Province FE	Yes	Yes
Birth cohort FE	Yes	Yes
<i>N</i>	9537	5844

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates. Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variable is the likelihood of migrating. The omitted groups are those ages 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: Own computations.

Table B6: Conflict exposure and schooling outcomes–Placebo cohorts

	Primary schooling		Lower secondary	
	(1)	(2)	(3)	(4)
Conflict exposure	-0.045*** (0.013)	-0.024* (0.012)	-0.028*** (0.011)	-0.005 (0.010)
Conflict exposure*age 25+	0.010 (0.011)	0.010 (0.013)		
Urban	0.109*** (0.019)	0.100*** (0.017)	0.097*** (0.019)	0.088*** (0.016)
Wealth index	0.184*** (0.010)	0.199*** (0.010)	0.166*** (0.007)	0.187*** (0.008)
Male	0.279*** (0.009)	0.279*** (0.008)	0.279*** (0.009)	0.280*** (0.009)
Conflict exposure*age 25+			0.001 (0.009) (0.010)	0.000 (0.011) (0.010)
Province FE		Yes		Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy		Yes		Yes
<i>N</i>	21646	21646	18946	18946

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS cluster and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: DHS, Own computations.

Table B7: Conflict exposure and schooling outcomes–Alternative dataset

	Primary schooling		Lower secondary	
	(1)	(2)	(3)	(4)
Conflict exposure 1996	-0.004 (0.008)	0.014* (0.007)	-0.004 (0.009)	0.018*** (0.007)
Conflict exposure 1996*age 6-11	-0.003 (0.005)	-0.005 (0.005)		
Urban	0.157*** (0.031)	0.115*** (0.026)	0.135*** (0.041)	0.080*** (0.028)
Wealth index	0.146*** (0.011)	0.160*** (0.011)	0.144*** (0.012)	0.167*** (0.010)
Male	0.230*** (0.011)	0.232*** (0.011)	0.258*** (0.013)	0.261*** (0.013)
Conflict exposure 1996*age 12-13			-0.006 (0.008)	-0.008 (0.008)
Province FE		Yes		Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy		Yes		Yes
<i>N</i>	17992	17992	10985	10985

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Conflict exposure 1996 is the weighted average number of deaths in 1996 using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: DHS, Own computations.

Table B8: Conflict exposure and schooling outcomes–Alternative dataset

	Primary schooling		Lower secondary	
	(1)	(2)	(3)	(4)
Conflict exposure	-0.061*** (0.013)	-0.006 (0.022)	-0.056*** (0.012)	0.037* (0.020)
Conflict exposure*age 6-11	0.005 (0.009)	0.005 (0.009)		
Urban	0.123*** (0.023)	0.119*** (0.020)	0.105*** (0.026)	0.090*** (0.020)
Wealth index	0.161*** (0.010)	0.187*** (0.010)	0.150*** (0.010)	0.186*** (0.008)
Male	0.241*** (0.010)	0.243*** (0.010)	0.262*** (0.011)	0.264*** (0.011)
Conflict exposure*age 12-13			0.001 (0.016)	-0.005 (0.017)
Province FE		Yes		Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy		Yes		Yes
<i>N</i>	22551	22551	13793	13793

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates based on equation (3.2). Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of deaths using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: Own computations.

Table B9: Conflict exposure and Primary schooling outcomes-Alternative specifications

	Primary schooling (1)	Primary schooling (2)
Conflict exposure	-0.097*** (0.025)	-0.047 (0.034)
Conflict exposure*age 0-5	0.026 (0.036)	0.057 (0.043)
Conflict exposure*age 6-11	0.002 (0.020)	0.010 (0.026)
Urban	0.422*** (0.071)	0.434*** (0.065)
Wealth index	0.529*** (0.038)	0.596*** (0.046)
Male	0.732*** (0.035)	0.769*** (0.036)
Province FE		Yes
Birth cohort FE	Yes	Yes
DHS dummy		Yes
<i>N</i>	29727	29727

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates. Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variable the likelihood of completing primary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 . Source: DHS, Own computations.

Table B10: Conflict exposure and Lower secondary schooling outcomes–Alternative specifications

	Lower secondary (1)	Lower secondary (2)
Conflict exposure	-0.075*** (0.028)	-0.028 (0.028)
Conflict exposure*age 0-5	0.077** (0.037)	0.102*** (0.036)
Conflict exposure*age 6-11	0.017 (0.026)	0.021 (0.030)
Conflict exposure*age 12-13	-0.018 (0.040)	-0.039 (0.044)
Urban	0.362*** (0.078)	0.379*** (0.067)
Wealth index	0.499*** (0.028)	0.589*** (0.031)
Male	0.714*** (0.033)	0.753*** (0.035)
Province FE		Yes
Birth cohort FE	Yes	Yes
DHS dummy		Yes
<i>N</i>	29727	29727

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Probit Estimates. Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variable is the likelihood of completing lower secondary schooling. The omitted groups are those aged 14-24 at the beginning of the conflict in 1996.

Source: DHS, Own computations.

Table B11: Conflict exposure and Schooling outcomes: LPM

	Primary schooling		Lower secondary	
	(1)	(2)	(3)	(4)
Conflict exposure	-0.025*** (0.006)	-0.009 (0.007)	-0.025*** (0.009)	-0.004 (0.006)
Conflict exposure*age 6-11	0.001 (0.005)	0.002 (0.005)		
Urban	0.192*** (0.027)	0.185*** (0.023)	0.143*** (0.031)	0.128*** (0.024)
Wealth index	0.125*** (0.009)	0.153*** (0.009)	0.163*** (0.010)	0.191*** (0.009)
Male	0.246*** (0.011)	0.248*** (0.012)	0.279*** (0.014)	0.281*** (0.014)
Conflict exposure*age 12-13			-0.004 (0.012)	-0.009 (0.012)
Province FE		Yes		Yes
Birth cohort FE	Yes	Yes	Yes	Yes
DHS dummy		Yes		Yes
<i>N</i>	22551	22551	13793	13793

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. Estimates based on LPM. Clustered standard errors in parenthesis at the village level. Conflict exposure is the weighted average number of conflicts using the inverse distance between DHS clusters and conflict zones as weights. The dependent variables are the likelihood of completing primary and lower secondary schooling. The omitted groups are those aged 12-24 at the beginning of the conflict in 1996 for primary schooling and individuals aged 14-24 for lower secondary schooling. Source: DHS, own computations.

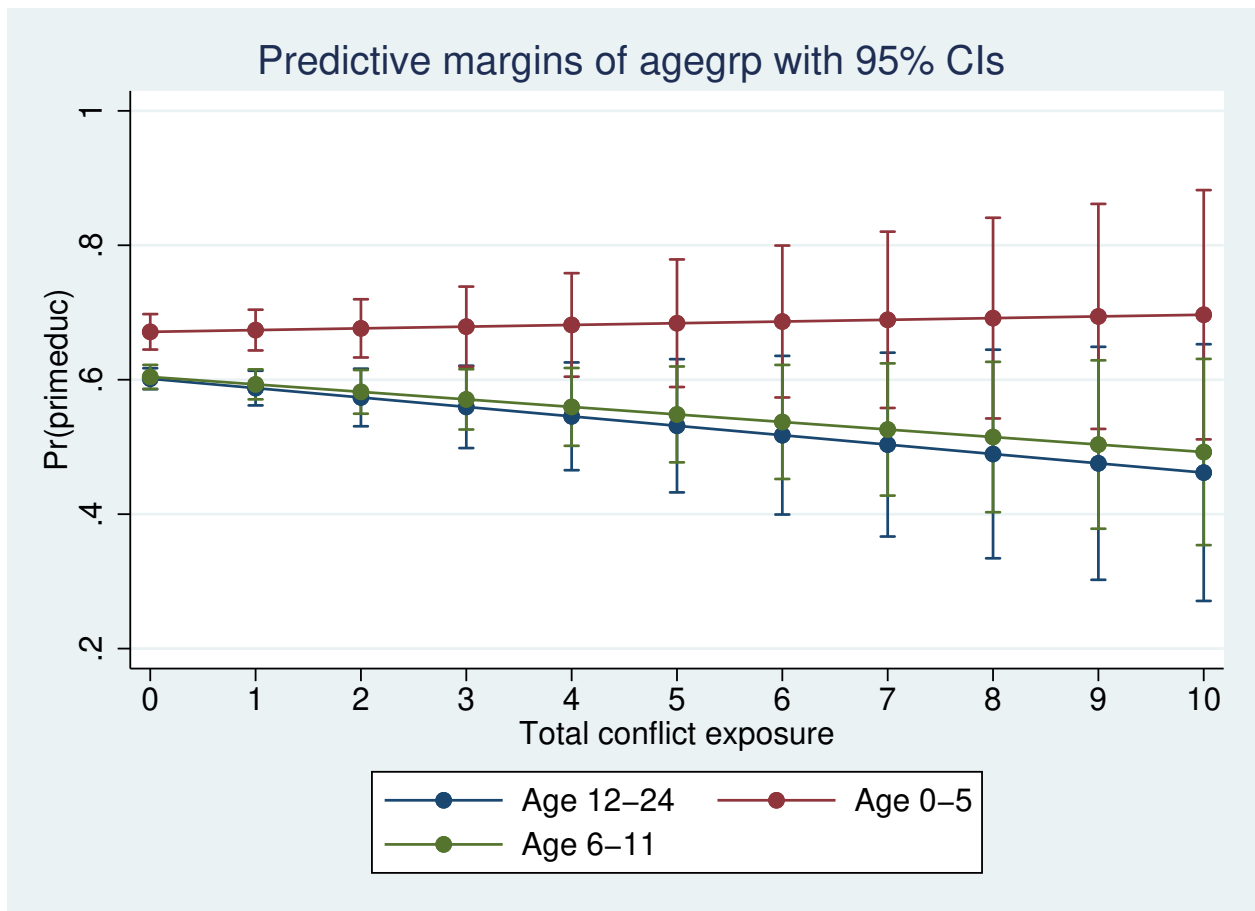


Figure B1: Conditional marginal effect of completing primary schooling

Source: DHS, Own computations.

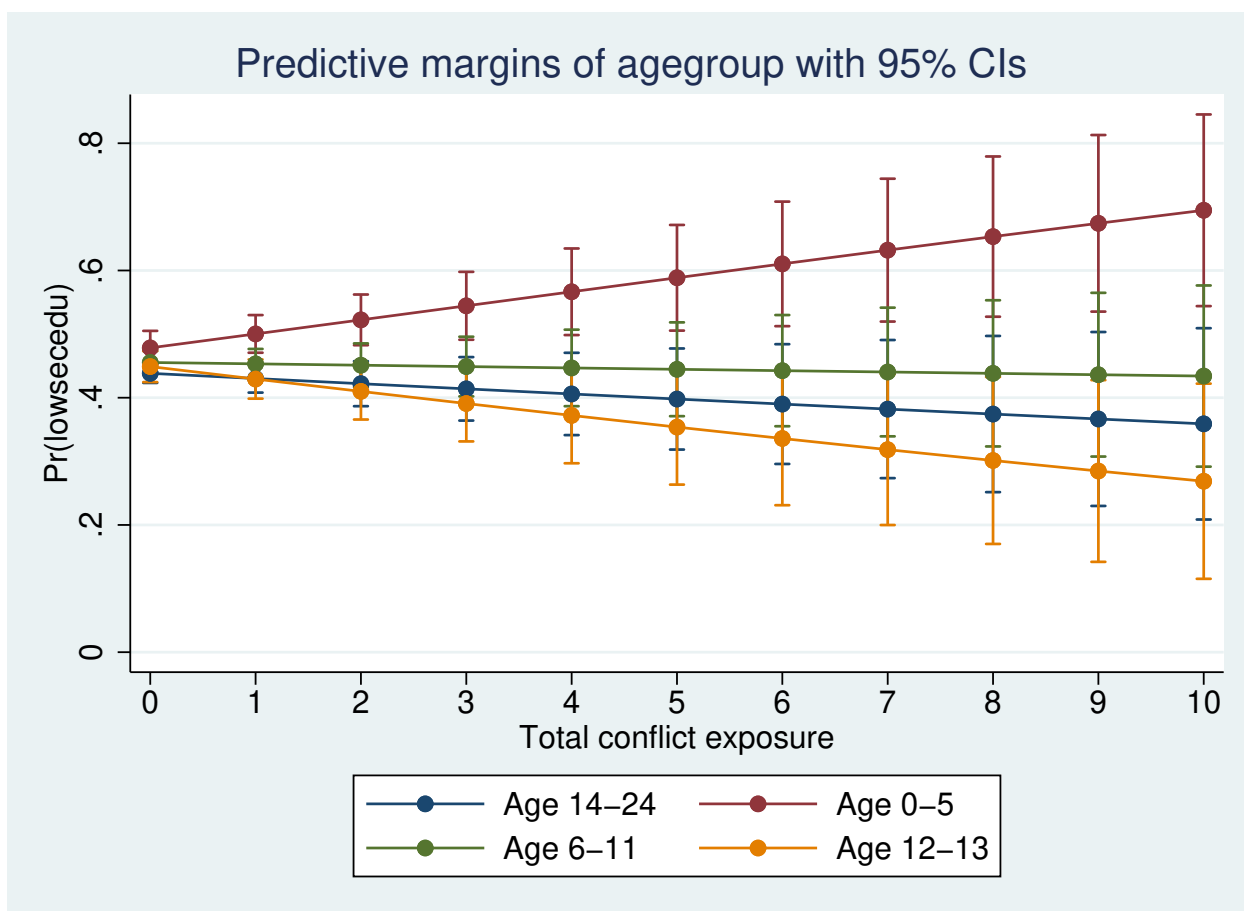


Figure B2: Conditional marginal effect of completing lower secondary schooling

Source: DHS, Own computations.

Appendix C: Additional tables and figures to chapter 4

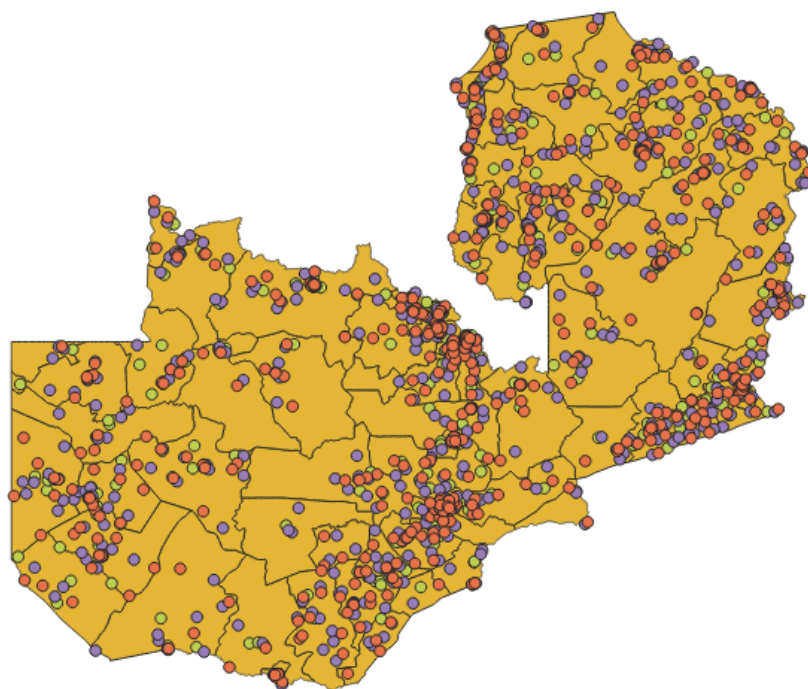


Figure C1: DHS clusters at the district level

Source: DHS, Own computations.

Table C1: The effect of drought in utero on child health outcomes- Alternative specifications

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A	District x month birth FE					
Months drought utero	-0.055*** (0.009)	-0.052*** (0.010)	-0.002 (0.006)	-0.004 (0.007)	-0.027*** (0.004)	-0.027*** (0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
<i>N</i>	16372	16372	16356	16356	16392	16392
<i>R</i> ²	0.109	0.118	0.070	0.081	0.106	0.112
PANEL B	Mother FE					
Months drought utero	-0.086*** (0.015)	-0.084*** (0.016)	-0.007 (0.011)	-0.006 (0.011)	-0.046*** (0.008)	-0.044*** (0.008)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
District x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
<i>N</i>	8348	8348	8341	8341	8367	8367
<i>R</i> ²	0.694	0.703	0.674	0.681	0.708	0.716

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero is the number of months of drought in utero. Physical growth is the average of HAZ and WHZ. Whole sample. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth, child age in months, birth order. Source: own computations.

Table C2: The effect of drought in utero on child health outcomes- Alternative definitions of drought

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
Drought in utero (dummy)	-0.158*** (0.039)	-0.174*** (0.039)	-0.010 (0.031)	-0.017 (0.031)	-0.082*** (0.021)	-0.096*** (0.022)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
<i>N</i>	14106	14039	14090	14024	14128	14061
<i>R</i> ²	0.368	0.378	0.336	0.345	0.366	0.376

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Drought in utero (dummy) is a dummy if a child has experienced at least one drought episode in utero. Physical growth is the average of HAZ and WHZ
Source: own computations

Table C3: The effect of drought in utero on child health outcomes- Alternative definitions of drought during growing period

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
Months drought utero (growperiod)	-0.072*** (0.015)	-0.081*** (0.016)	-0.006 (0.013)	-0.007 (0.013)	-0.038*** (0.009)	-0.043*** (0.009)
Control	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District linear trend		Yes		Yes		Yes
<i>N</i>	14106	14039	14090	14024	14128	14061
<i>R</i> ²	0.369	0.379	0.336	0.345	0.367	0.377

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero growing period in utero Physical growth is the average of HAZ and WHZ. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth, child age in months, birth order. Source: own computations.

Table C4: The effect of drought in utero on child health outcomes-severe drought

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
Months drought utero (severe)	-0.056*** (0.012)	-0.063*** (0.012)	0.002 (0.010)	-0.000 (0.010)	-0.026*** (0.007)	-0.031*** (0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
<i>N</i>	14106	14039	14090	14024	14128	14061
<i>R</i> ²	0.369	0.379	0.336	0.345	0.366	0.376

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero is the number of months of a severe drought in utero. Physical growth is the average of HAZ and WHZ. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth, child age in months, birth order. Source: own computations.

Table C5: The effect of lifetime drought on child health outcomes

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
Lifetime drought	-0.274*** (0.090)	-0.325*** (0.091)	-0.100 (0.062)	-0.135** (0.066)	-0.175*** (0.051)	-0.227*** (0.051)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
<i>N</i>	14106	14039	14090	14024	14128	14061
<i>R</i> ²	0.368	0.378	0.336	0.345	0.366	0.377

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Lifetime drought is a dummy variable for whether a child has experienced drought in his lifetime. Physical growth is the average of HAZ and WHZ. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the household male, mother's age at birth, child age in months, birth order. Source: own computations.

Table C6: The effect of drought in utero on child health outcomes-different month scales

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A		1-month scale				
Months drought utero	-0.066*** (0.012)	-0.074*** (0.012)	-0.009 (0.010)	-0.013 (0.010)	-0.035*** (0.006)	-0.042*** (0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
R^2	0.370	0.379	0.336	0.345	0.367	0.378
PANEL B		12-month scale				
Months drought utero	-0.016** (0.008)	-0.021*** (0.008)	-0.008 (0.006)	-0.010 (0.006)	-0.011** (0.004)	-0.015*** (0.004)
Control	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
R^2	0.368	0.377	0.336	0.345	0.366	0.376
N	14106	14039	14090	14024	14128	14061

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero is the number of months of drought in utero. Physical growth is the average of HAZ and WHZ. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth child age in months, birth order. Source: own computations.

Table C7: The effect of drought in utero on child health outcomes–Alternative definitions

	Stunting		Wasted		Physical growth deficit	
	(1)	(2)	(3)	(4)	(5)	(6)
Months drought utero	0.016*** (0.003)	0.015*** (0.003)	0.000 (0.001)	-0.001 (0.001)	0.005*** (0.002)	0.005*** (0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
<i>N</i>	14106	14039	14090	14024	14128	14061
<i>R</i> ²	0.352	0.365	0.329	0.337	0.332	0.338

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child stunted, Wasted and physical growth deficit. Months drought utero is the number of months of drought in utero. Physical growth is the average of HAZ and WHZ. stunted is a dummy whether $HAZ < -2SD$. Physical growth deficit is a dummy whether physical growth $< -2SD$. Wasted is a dummy whether $WHZ < -2SD$. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth, child age in months, birth order.

Source: own computations.

Table C8: The effect of drought in utero on child health outcomes-Whole sample

	HAZ		WHZ		Physical growth	
	(1)	(2)	(3)	(4)	(5)	(6)
Months drought utero	-0.059*** (0.007)	-0.064*** (0.007)	-0.004 (0.006)	-0.007 (0.006)	-0.031*** (0.004)	-0.035*** (0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District-specific time trend		Yes		Yes		Yes
<i>N</i>	23446	23363	23415	23333	23480	23397
<i>R</i> ²	0.318	0.327	0.279	0.286	0.315	0.323

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero is the number of months of drought in utero. Physical growth is the average of HAZ and WHZ. Whole sample. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth, child age in months, birth order. Source: own computations.

Table C9: The effect of drought in utero on mortality and birth interval

	in utero mortality (1)	Infant mortality (2)	Child mortality (3)	Birth interval (4)
Months drought utero	-0.000 (0.002)	0.001 (0.001)	0.001 (0.001)	0.081 (0.117)
Controls	Yes	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes	Yes
District linear trend	Yes	Yes	Yes	Yes
<i>N</i>	16837	16838	16838	11992
<i>R</i> ²	0.347	0.310	0.307	0.522

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are in utero mortality, infant mortality, child mortality and birth interval. In utero mortality is a dummy for whether a woman has ever had terminated pregnancy. Infant mortality is a dummy of whether the child died within 12 months. Child mortality is a dummy of whether the child died before reaching five years old. Birth interval is the number of months of preceding birth intervals. Months drought utero is the number of droughts in utero. The controls include a dummy for child male, mother schooling, a dummy for urban, a dummy for the head of the household male, mother's age at birth, child age in months, birth order. Source: own computations.

Table C10: The effect of drought in utero on child health outcomes–heterogeneity analysis

	HAZ (1)	WHZ (2)	physical growth (3)
PANEL A	Heterogeneity by child's gender		
Child male x months drought utero	-0.005 (0.018)	-0.001 (0.013)	-0.004 (0.010)
Controls	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes
District-specific time trend	Yes	Yes	Yes
R^2	0.380	0.345	0.378
PANEL B	Heterogeneity by residence of location		
Urban x months drought utero	-0.000 (0.022)	0.026 (0.017)	0.014 (0.012)
Controls	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes
District-specific time trend	Yes	Yes	Yes
R^2	0.380	0.345	0.378
PANEL C	Heterogeneity by recent birth		
Recent birth x months drought utero	-0.053*** (0.020)	-0.025* (0.015)	-0.039*** (0.010)
Control	Yes	Yes	Yes
Village x Month birth FE	Yes	Yes	Yes
District linear trend	Yes	Yes	Yes
R^2	0.379	0.344	0.373
N	14039	14024	14061

Notes: Significance levels: *, 10%; **, 5%; ***, 1%. DHS 2007, 2013/14 and 2018. Robust standard errors clustered at the village level in parenthesis. The dependent variables are child HAZ, WHZ and physical growth. Months drought utero is the number of months of drought in utero. Physical growth is the average of HAZ and WHZ. Recent birth is a dummy whether the child is 24 months of age. The control includes a dummy for child male, age in months, birth order. Source: own computations.