Doctoral Thesis

Inter-generational and Within-generation Spillovers in Human Capital Formation

Gerald McQuade

This thesis is submitted for the degree of
Doctor of Philosophy.

Lancaster University Management School
Department of Economics
April 2025

Abstract

Inter-generational and Within-generation Spillovers in Human Capital Formation

Gerald McQuade

This thesis consists of three pieces of applied work, each combining multiple data sources to investigate the short, medium and long-term determinants of health and human capital formation in Peru.

Chapter 2 examines the multigenerational effects of maternal grandmothers' exposure to drought during pregnancy. Matching longitudinal data on the birthdate and location of the mother of Young Lives respondents in Peru with external climate data, I find that exposure to drought has a persistent negative impact on the health stock of the daughter and grandchild. Grandchildren display lower height-for-age, beginning in early childhood and persisting through adolescence, with the height gap widening as they enter puberty. Additionally, grandchildren have lower early-life weight-for-age, however this effect diminishes with age. The effect is strongest for grandsons, and isolated to grandmothers living in rural areas during exposure, with exposure in early pregnancy having the largest impact. The first generation are also affected, with mothers being shorter in stature in adulthood. Estimating the average controlled direct effect (ACDE), I find that the mother's long-term health is the primary mediator for transmission between the first and second generation, although I cannot fully rule out other unobserved mechanisms.

Chapter 3 also matches Young Lives data to historical climate data to assess how exposure to rainfall shocks in early life can have a persistent influence on the development of personality traits in adolescence and adulthood. I find high rainfall exposure in the second and third years of life negatively affects scores. Additionally, high prenatal rainfall, specifically in the 3rd trimester, has a positive impact on scores, driven by girls and those in the poorest households. Examining underlying mechanisms, I find that parents increase labour supply in response to shocks, which has a negative impact on parent-child interaction in early life, with no effects on measures of material investments or child physical development.

Chapter 4 combines school administrative data with national standardised tests to identify siblings and estimate spillover effects in attainment arising from a public school-

day extension reform. Using a regression discontinuity design based on school eligibility criteria, we estimate a positive effect of older siblings' schooling on younger sibling attainment in reading and mathematics. Positive spillover effects are driven by the effect on girls, with the largest effect amongst sister-sister pairs, compared to null effects for younger brothers. Our results indicate that evaluations which consider only the benefits for the targeted child could systematically understate the benefit-cost ratio of educational reforms.

Declaration of Authorship

I, Gerald McQuade, declare that this thesis titled, "Inter-generational and Within-generation Spillovers in Human Capital Formation" and the work presented in it are my own and has not been submitted in any other form for the award of a higher degree elsewhere.

Funding for my studies was provided by the ESRC through a 1+3 studentship with the North West Social Science Doctoral Training Partnership (NWSSDTP, grant number ES/P000665/1), including additional funding for a 6-month internship placement and a 3-month overseas institutional visit to Peru. I declare that the funders bear no responsibility for the contents of this thesis.

Chapter 4, titled "Sibling Spillover Effects in Education: Evidence From an Extended School-day Reform in Peru" is a joint authored piece of work with myself, Catherine Porter and Alan Sanchez. A signed declaration of joint authorship can be found in Appendix A.

An approximate estimate of the word count is: 40835

Gerald McQuade
April 2025

Acknowledgements

While completing a thesis can often feel like a very lonely task, it was only through the support and encouragement of many that it was possible. I am not sure I can ever find the words to express my gratitude to you all, but I'll give it a go.

First, I would like to thank my supervisors Catherine Porter and Saurabh Singhal. Cath, thank you for your unwavering support from undergraduate, to masters, and through to the PhD – I would not have reached this stage without it. Your patience, ever-relevant advice, and excellent sense of humour have been invaluable in helping me understand the nuanced and often confusing world of international development. Saurabh, your insight and sharp eye for detail always pushed me to be a better researcher, and helped me see the bigger picture when I was stuck in the weeds. The quality of this thesis would no doubt be much lower without your advice. I will sorely miss the guidance and support of you both, and I hope we can collaborate together again soon. I also want to thank Robbie Mochrie, for your supervision, guidance, and numerous reference letters that helped facilitate my path to a PhD.

I wish to thank the Department of Economics at Lancaster, for all the support I received in my PhD studies, and for providing me a teaching job that supported me while I reached the finish line. Thank you to all my colleagues for your advice and kindness. In particular, I owe a lot of thanks (and blue tack, marker pens, and notebooks) to everyone in the PG and UG offices for their patience and understanding, especially Caren Wareing, who would always go above and beyond to find an answer to my endless questions and requests.

I could not have got through each day without my PhD colleagues. First and foremost to the Dev Econ gang: Alessandra Hidalgo, mi causita, you always provided the best feedback, advice, and plenty of laughs along the way; and Utsa Banerjee, thank you for your kindness, patience, and thoughtful advice as I ranted about every minor disaster. You are both inspiring and I can't wait to see what you achieve. To Tanisha (Tanisha), Rubab Ahmed, and the rest of the Sulby Drive crew, for opening up your home and providing many great memories. To Matteo Caruso, who I would walk 500 miles for, and would walk 500 more. To Joseph Richardson, for great discussions that always tested my knowledge of economics, and to Jinghui Yu, for always being there to spill the tea. I owe a lot of my early knowledge about PhD life to my former office-mates,

Aristotelis Margaris and Amanda De Pirro, and I could not have adjusted to Lancaster without Nathan Nabil and Chiara Costi, who were the life of the party. Lastly, Thank you to Harry Rolls, Jerome Santoul, Emma Pooley, and the rest of the EconHub team, for making teaching fun and for providing impromptu group therapy sessions.

During my PhD, I worked closely with many colleagues at Young Lives, both in Oxford and across the different study countries. I am grateful to Marta Favara, who graciously hosted me at ODID as a visiting student, and who trusted me enough to let me dive head first into helping deliver round 7 in Oxford, Addis Ababa, and Hyderabad. Thank you also to Alan Sanchez, for supporting my work all the way through my PhD, for facilitating my eventual fieldwork in Peru, and for being an excellent co-author. To Angeles and Georgina Fensom, for dealing with my never-ending data requests, and to Dani Shahzadi for providing support and great company too. Finally, I want to thank Laura Ahlborn and Sophie Von Russdorf, who were superstars and the best fieldwork teammates (and co-authors) I could have asked for. I hope I can be lucky enough to work together again in the near future.

I was lucky enough to be able to visit GRADE in Peru, an invaluable experience which helped me grow as a researcher. While I wish I could thank everyone for their kindness, I especially want to thank Dianela Espinoza and Jennifer Lopez, for your friendship, great dance moves, and excellent taste in food; Santiago Cueto and Antonio Campos, for being such accommodating hosts; Monica Lizama, for my daily Spanish lessons; and the rest of the Niños del Milenio team for such a warm welcome. I am indebted to the hard-work of the Monitoring and Evaluation Unit team at MINEDU, without which my final chapter would not have been possible. I also want to thank Sergio, and the rest of the Hidalgo family, without your kindness and hospitality, Lima would have felt a lot more lonely.

Just as important was the support I received from outside of academia. Thanks especially to Robbie Binnie and Ross Scott for being great flatmates and even greater friends. To Tara Matthews, for always having time to listen to my woes and offer timely advice over a pint. To Sophie Baird, I will always be grateful for your patience and support through these tough years. And to the Carluke boys and the Cardiff gang, for always having my back, although you are all going to need some new patter once I graduate.

Finally, all of this would not have been possible without the love and support of my family. Fran and Ashleigh, your bravery, kindness, and resilience are an inspiration for me every day. Wee baby Innes, you have been a source of so much joy these last few months. This world is yours and I can't wait to see what you make of it, although I may have some words of warning if you ever consider doing a PhD. Dad, I am so lucky to be able to pursue my dreams knowing I have your unwavering love and support behind me. Finally Mum, you told me long ago, when I felt I wasn't cut out for university, that you

knew that wasn't true, and that you could see me with a PhD one day. You were right, and your faith, encouragement, and boundless love has carried me through it all. I love you all so much.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

While it is clear that shocks and investments experienced by an individual in the current period can have an impact on their cognitive, non-cognitive, and health outcomes in the short run, a growing body of evidence explores how shocks and investments in previous periods, often many years prior, can have persistent impacts on long-term human capital formation. In particular there is a well established literature exploring how circumstances in early life or perinatal period, often seen a a crucial stage for development, can shape future abilities and health trajectories, with effects persisting into adolescence and adulthood, even after remaining latent for many years (Almond & Currie, 2011). An important theoretical implication of this literature is the potential for complementarities between shocks or investments in this early period and subsequent investments in later periods, highlighting the importance of a good start (Cunha & Heckman, 2008; Cunha et al., 2010; Heckman, 2007).

Building on the evidence that early life experiences have consequences for life-long human capital, a further strand of literature has emerged examining how effects may be transmitted to subsequent generations, potentially impacting on the health and human capital stock of the next generation (Almond et al., 2018; East & Page, 2020). This has important policy implications, as it suggests that without accounting for the potential for multigenerational transmission, policy makers likely under-estimate the true cost of an environmental or economic shock, with subsequent interventions being insufficient to mitigate the longer-term impacts (Doyle & Jernström, 2023), leading to a vicious cycle of entrenched disadvantage.

Furthermore, recent evidence indicates that investments and shocks experienced by one's peers, in particular close family members such as siblings, can also act as important determinants of an individual's human capital outcomes (Black et al., 2021; Sacerdote, 2014). Such spillovers are important to quantify, especially if they have unintended consequences for individuals who were not the intended target of an intervention or reform (Figlio et al., 2023). However it can be difficult to causally identify these spillovers due to the "reflection problem" that plagues empirical studies of peer effects (Manski,

1993).

The purpose of this thesis is to explore a range of short, medium and long-term determinants of health and human capital. It consists of three self-contained chapters, each focusing on shocks and investments which occur in early life and childhood, including across generations, or between individuals. I employ a range of data matching strategies and "quasi-experimental" econometric techniques to causally identify the effects on respondents' health, attainment, and personality trait formation. As the exact transmission mechanism for these impacts is ambiguous *a priori*, where possible I conduct in-depth analyses of the underlying mechanisms. Additionally, I assess how these effects may propagate heterogenously across sub-groups, interacting with previous or subsequent investments or shocks (Almond et al., 2018). Each chapter is outlined in greater detail below.

In Chapter 2 I examine the multigenerational effects of maternal grandmothers' exposure to drought while pregnant on the health and human capital outcomes of her daughter and grandchild. Using detailed data on the birthdate and location of the mothers of respondents to the Young Lives longitudinal study in Peru, I match external data which measures local crop-growing conditions to identify prenatal exposure to drought during the grandmother's pregnancy. I find that exposure to a drought shock has a persistent negative impact on the long-term health stock of the daughter (first generation) and the grandchild (second generation). Grandchildren display lower height-for-age, my measure of health stock, first becoming apparent in early childhood (age 5) and persisting through adolescence (up to age 15), with the height gap widening as they enter puberty. Additionally, grandchildren display a lower early-life weight-for-age, used as a measure of health flow, with the largest impact shortly after birth, however this effect diminishes as children age. These impacts are present at the extensive margin, as measured by indicators of stunting and wasting. The first generation are also affected, with mothers being shorter in stature in adulthood, but with no effect found for weight. Given weight is more sensitive to current health inputs, while height better captures the long-term cumulative health stock of an individual, this indicates that multigenerational effects can have a lasting impact beyond affecting the initial health endowment.

Considering heterogeneities, impacts are strongest for grandsons. Effects are driven by grandmothers living in rural areas at time of exposure, with shocks experienced in the early stages of pregnancy having the largest impact. To assess potential transmission mechanisms, I estimate the average controlled direct effect (ACDE) for a range of biological and economic mediators. I find strong evidence that the mother's long-term health, as measured by height, is the primary mediator for transmission across generations, while socioeconomic mediators such as household wealth and maternal education do not explain the effect on the second generation.

While Chapter 2 contributes to the early "second generation" literature (Doyle &

Jernström, 2023), Chapter 3 contributes to the more established "first generation" literature on how an individual's early life circumstances can influence long-term human capital formation (Almond & Currie, 2011). This strand has a theoretical foundation in the "foetal origins" hypothesis (Barker, 1990), which focused narrowly on the intra-uterine environment and later life disease, but has widened to consider a range of outcomes across health, education, and labour market skills, and has grown to accommodate the early life environment experienced just after birth, which is also seen as a crucial period for development (Currie & Vogl, 2013). This chapter contributes to the literature by exploring the relationship between early-life circumstances and the formation of personality traits (often referred to as non-cognitive, socio-emotional, or 'soft' skills), which is still largely unexplored (Almond et al., 2018). Understanding this relationship is salient given to the important role these traits and skills can play in determining future socioeconomic success, especially given the theoretical complementarity of early shocks and investments (Almlund et al., 2011; Cunha et al., 2010).

Specifically, I assess the impact of exposure to early life rainfall shocks on measures of respondents' appraisal of their own self-worth, competence, and capability, as measured by their core self-evaluation (CSE), a construct strongly associated with socioeconomic success (Chang et al., 2012; Judge et al., 1998). I find exposure to a positive rainfall shock (+ 1.5 S.D. from the long-run, location specific mean) in the second and third years of life negatively affects scores measured in adolescence and adulthood (age 14-23). Additionally, exposure to a positive rainfall shock in-utero has a positive impact on scores, however an analysis of heterogenous effects indicates that this is isolated to girls and those living in the poorest households. Exploring the underlying mechanisms, I find that parents increase labour supply in response to higher rainfall, which during a crucial period for social interaction and development, has a negative impact on measures of parent-child bonding. No effects are found for measures of material investments or children's physical development, suggesting that the substitution effect of reduced parental availability dominates the potential positive income effects from increased labour supply.

Finally, Chapter 4, co-authored with Catherine Porter and Alan Sanchez, assesses the potential for spillover effects between siblings. It is generally difficult to disentangle the causal effect siblings have on each other from their shared background (Manski, 1993). We therefore exploit an exogenous change in the schooling of the older sibling, arising from a national reform that extended the length of the school day and improved other inputs within 1,000 public secondary schools in Peru, to estimate spillovers on the educational attainment of a younger sibling.

Assignment to the program was not random, however we are able to exploit initial arbitrary eligibility criteria which were based on the number of home-rooms in the school (*secciones*, or sections), using a fuzzy regression continuity design to identify the local

average treatment effect for compliers. We estimate a positive effect on the attainment of primary school-aged younger siblings, increasing test scores in both reading and mathematics. Considering how spillover effects may propagate differently across sibling pairs, we find effects are concentrated amongst younger sisters, with the largest effect found for sister-sister pairs, compared to null effects for younger brothers, regardless of older sibling gender. Our results indicate that evaluations which consider only the benefits for the targeted child could systematically understate the benefit-cost ratio of educational reforms.

Overall, the findings of this thesis contribute to understanding the complex process of health and human capital formation, and how shocks and investments experienced in the crucial stages of life can have an important effect on an individual's human capital formation, including effects which are transmitted across generations and between peers.

Chapter 2

Grandmothers and Grandsons: Multigenerational Effects of Drought Exposure in Peru

2.1 Introduction

There is a well established literature addressing the effect of prenatal shocks on life-long health and human capital outcomes (See Almond & Currie, 2011; Almond et al., 2018; Currie & Vogl, 2013, for reviews of the literature). It builds on the "foetal origins" hypothesis (Barker, 1990), which posits that the intrauterine environment is critical for long-term development, with shocks and investments experienced during this period having effects which persist long after birth, through "programming" the expression of parts of the genome crucial for healthy growth and cognitive function (Petronis, 2010). Furthermore, there is a growing cross-discipline literature that posits that effects are not limited to the generation exposed to these insults, but can echo down to subsequent generations (Aiken & Ozanne, 2014; Doyle & Jernström, 2023; Drake & Liu, 2010; East & Page, 2020).

Although there is an established body of evidence derived from lab-based animal studies that this developmental programming can affect subsequent non-exposed generations (Aiken & Ozanne, 2014), as well as ample evidence of strong correlations in health, educational attainment, and socio-economic outcomes across generations (Almond et al., 2012; Behrman & Rosenzweig, 2002; Bevis & Barrett, 2015; Bhalotra & Rawlings, 2013; Currie & Moretti, 2003, 2007; Emanuel et al., 1992), causal evidence of multigenerational effects within humans is very limited. Although this is in-part due to practical data limitations (Almond et al., 2018), it is also due to the difficulty of disentangling a causal effect from other confounding factors across generations (East & Page, 2020). Furthermore it can be difficult to identify the mechanisms underpinning

the transmission of these effects across generations, given many commonly experienced shocks are correlated with background characteristics.

Understanding the potential for multigenerational effects has important policy implications. If a negative effect of a shock experienced by one generation has a lasting impact on the next generation, even in the absence of further shocks, it is likely that policy-makers do not fully account for these consequences for later generations, underestimating the true cost of shocks, as well as the cost-benefit ratio of any subsequent policy interventions aimed at mitigation (Doyle & Jernström, 2023). It is therefore of benefit to quantify the presence and magnitude of cross-generational effects, as well as identify the likely underlying channels of transmission, to inform future mitigation strategies.

Using novel data from the Peruvian sample of the Young lives study including the birth location and date of the mother of respondents, I match external climate data to identify their prenatal exposure to drought. I find that the exposure of a gestating grandmother to a drought shock has a negative impact on the long-term health of her daughter, who is 0.75cm shorter on average in adulthood than non-exposed individuals. This effect is also transmitted to her grandchild, who is also less healthy. Considering the dynamic effects, exposure is associated with grandchildren having a lower height-for-age z-score (HAZ), an effect that is persistent from early childhood into late adolescence (between -0.076 and -0.173 S.D. across ages 5-15). While the grandchildren of exposed grandmothers also display a lower weight-for-age z-score (WAZ) in early childhood (-0.179 S.D. and -0.109 S.D. at age 1 and 5, respectively), this impact diminishes by age 8. Effects are also realised at the extensive margin, with shock exposure being associated with a higher incidence of stunting in mid-childhood and adolescence, and a higher probability of being classed as underweight in early childhood. Given weight is more sensitive to current health inputs, while height better captures the long-term health stock of an individual, this indicates that prenatal drought exposure can have a lasting multigenerational impact beyond the initial health endowment. This suggests the effects on subsequent generations may not be easily addressed by post-exposure investments, highlighting the importance of mitigating the initial shock exposure. In contrast to health outcomes, I find little evidence of a multigenerational effect of drought exposure on cognitive ability or educational attainment.

This chapter contributes to the nascent literature documenting causal multigenerational effects in three key ways. First, it exploits exogenous spatial and temporal variation in exposure to drought conditions experienced by the grandmother (zero generation) while pregnant, to establish the multigenerational effect of prenatal drought exposure on the health and cognitive outcomes of i) their offspring, who was exposed in-utero (first generation) and ii) of their grand-offspring (second generation).[1] Second,

[1]This follows commonly used notation from the epidemiology literature, set out by Skinner (2008),

this study examines the dynamics of how these multigenerational effects manifest in the second generation from birth into adolescence, as well as exploring the potential for heterogeneity in the transmission of effects. Third, through formal mediation analysis, I provide evidence that effects are predominantly transmitted through a biological pathway, impacting the long-term health of the exposed first generation, however results may suggest that either a direct effect (via an effect on the second generation germline) or the role of other unobserved environmental/economic pathways cannot be fully discounted.

These multigenerational effects on health outcomes are driven by the impact of exposure of grandmothers living in rural areas, with a large and significant negative effect on the HAZ of their grandchild, first appearing at age 5 and remaining into late adolescence, compared to a null and insignificant effect for grandmothers living in urban areas at time of exposure. Early impacts on WAZ are also isolated to the offspring of rural grandmothers. This suggests that the direct effect of exposure is larger in rural areas, where a higher proportion of households would be reliant on local food sources and agriculture-related income, likely directly impacting resources available for the first generation in-utero or immediately after birth. Additionally, I assess if the effects of exposure during the grandmothers' pregnancy differ when shock exposure is disaggregated by trimester. Results indicate that both the effects on grandchildren HAZ and early years WAZ are strongest for exposure to a shock during the first trimester, consistent with existing evidence that exposure earlier in the pregnancy has the largest impact on second generation outcomes (Khan, 2021; Stein & Lumey, 2000).

Considering heterogeneity amongst the second generation, I look at potential sex-specific differences in the transmission of effects between boys and girls, finding that the second generation effects are primarily exhibited in grandsons of exposed grandmothers, compared with small, insignificant effects for granddaughters. This is consistent with previous second generation findings where effects on HAZ and WAZ are isolated to the grandsons of shock-exposed maternal grandmothers (Fung & Ha, 2010). Additionally, using self-reported data on indicators of puberty, I assess how shock exposure interacts with entering in to pubertal growth in adolescent years. Results suggest that the gap in height-for-age between the grandchildren of unexposed and exposed grandmothers widens once in the pubertal stage of growth, with a large negative effect on HAZ estimated for those reporting signs of puberty at age 12, compared with a smaller insignificant impact for those not reporting signs of pubertal growth.

which describes a gestating female (F0 or zero generation) being exposed to an environmental insult, resulting in the embryo (F1 or first generation) and potentially the germ-line/reproductive cells (F2 or second generation) being exposed in-utero. Following Skinner (2008) and Drake and Liu (2010), I refer to effects on the F1 and F2 generation as "multigenerational", rather than "transgenerational", which is reserved for the impacts on the F3 generation (i.e. the great-grandchild of the F0 generation) and beyond.

Finally, I conduct formal mediation analysis, estimating the average controlled direct effect (Acharya et al., 2016; Joffe & Greene, 2009; VanderWeele, 2009) of shock exposure, net of the effect of the shock operating indirectly through some mediator. Results suggest that measures of the household environment experienced by the second generation seem to account for very little of the effect of shock exposure, while the long-term health of the first generation accounts for the whole of the baseline effect for outcomes at almost all ages, supporting a biological transmission of health across generations as the primary mechanism.

The rest of this study is as follows: Section 2.2 provides a summary of the evidence for multigenerational effects of prenatal shocks, as well examining the likely biological and environmental mechanisms which account for these effects. Section 2.3 summarises the data sources, defining the key variables and providing sample descriptive statistics, with the empirical strategy described in Section 2.4. Results, including additional analyses of heterogeneous effects, sensitivity, and robustness checks, are provided in Section 2.5. Finally, results from the mediation analysis are presented in Section 2.6, and Section 2.7 concludes.

2.2 Background

2.2.1 Literature Review

An extensive literature of "first generation" studies links prenatal and early life shocks with later life outcomes within a single generation (Almond & Currie, 2011; Almond et al., 2018; Currie & Vogl, 2013). Given evidence from animal studies that the impacts of early life shocks can echo across generations (Aiken & Ozanne, 2014), there is a clear incentive to assess the potential for this phenomenon in human studies, leading to the emergence of a "second generation" literature. Recent reviews of this burgeoning literature are provided by East and Page (2020) and Doyle and Jernström (2023) for the impacts on health and education/labour market outcomes respectively, therefore I will summarise only those most relevant to this analysis.

While a large body of cross-disciplinary evidence correlates maternal birth-weight and disease exposure with offspring birth-weight and educational outcomes (Almond et al., 2012; Bhalotra & Rawlings, 2013; Bhalotra & Rawlings, 2011; Currie & Moretti, 2007; Drake & Walker, 2004; Emanuel et al., 1992), these likely reflect a wide range of causal mechanisms, and cannot separate environmental or epigenetic effects from cross-child variation in growth due to inherited genetic endowments. Additionally, some studies exploit twins or adoptees to control for genetic inheritance and isolate the impact of early life shocks (Royer, 2009; Thompson, 2014), however it is likely that these studies may have limited external validity, and are difficult to conduct outside of high-income,

data-rich contexts.

Alternatively, several works attempt to identify a causal relationship between conditions experienced by the first generation in-utero (that is, during the grandmother's pregnancy) and their children's outcomes.[2] Early contributions exploit differences between cohorts exposed to famine and starvation prenatally and surrounding cohorts who were not exposed. Lumey (1992) studies the inter-generational effects of the Dutch Hunger Winter using hospital records, finding the children of mothers exposed to war-induced famine in-utero in the first and second trimester were more likely to be low birth-weight (LBW), with exposed mothers also more likely to be low birth-weight. However, a subsequent study using the same data-set found no significant relationship between maternal birth-weight and offspring birth-weight after adjusting for confounders, compared with a positive relationship for non-exposed mothers (Stein & Lumey, 2000). Painter et al. (2008) also study the Dutch hunger winter, finding women exposed in-utero become mothers at a younger age, give birth to more offspring, and have more twins than those not exposed. In contrast, male reproductivity was unaffected.

Similarly, a number of studies assess the impact of in-utero exposure to the 1959-1961 Great Chinese Famine. Almond et al. (2007), while mainly focused on the cohort exposed to famine, find that women exposed to the famine in early years and whose parents subsequently migrated to Hong Kong in 1962 had a higher ratio of female to male births than unexposed native-born mothers. Distinguishing exposure for mothers and fathers, Fung and Ha (2010) find that the children of mothers exposed in-utero have lower weight- and height-for-age (HAZ and WAZ), with no significant effect of fathers' in-utero exposure. Kim et al. (2014) also find a gendered effect of exposure, with the children of mothers exposed in-utero 5-7 percentage points less likely to enter middle school. These results may indicate that maternal shock exposure is more important than that of the father, however, given the extent of the famine, maternal and paternal exposure is highly correlated, making it difficult to disentangle effects.[3] Fung and Ha (2010) also find sex-specific differences for second generation outcomes, with the effect on HAZ and WAZ of mother's in-utero exposure limited to boys, compared with a null effect for girls. A major limitation with the studies above are that famines are extreme events.[4] The estimated effects are therefore likely to suffer from selection bias, often with only survivors of extreme malnutrition, starvation, or sickness observed (Royer & Witman, 2014). Additionally, the estimates obtained by the above studies of the Dutch winter famine may be confounded by other effects of the Second World War and its

[2] A broader literature studies the causal relationship between parent and child income and educational attainment (See Black & Devereux, 2011), however these studies focus only on the effect of shocks or investments experienced post-birth, in early childhood.

[3] Indeed, considering intensity of exposure for both parents, Li and An (2015) find negative effects of more intensive exposure to the famine on childrens' HAZ, regardless of which parent is exposed.

[4] The Great Chinese Famine for example, is estimated to have caused between 16.5-30 million deaths (Li & Yang, 2005).

aftermath.

More relevant for this study is a strand of literature exploring the impact of short-run unexpected deviations in climate conditions, including drought, experienced by the grandmother while pregnant or in the early years of the parents' life. In an early contribution, Venkataramani (2011) assesses the inter-generational transmission of health in Vietnam, as measured by height, including using an instrumental variable approach, using early life rainfall, grandparent socioeconomic status and regional fixed effects to capture non-genetic components of parent height variation. They find a strong relationship between maternal height variation and child height, while the association with paternal height is near zero under the instrumental variables approach.[5]

Using the India Household Development Survey, Khan (2021) assesses the impact of a rainfall deficit during the grandmother's pregnancy on her grandchildren's health and cognitive outcomes. They find that the grandchildren of exposed grandmothers have a lower HAZ amongst a pooled sample aged 0-5. Similarly, Hyland and Russ (2019) match DHS data in 19 Sub-Saharan countries to historical temperature and precipitation time series, finding that the children of mothers exposed in-utero to extreme drought are more likely to be born with low birth weight. A shortcoming of all previously mentioned studies is that they only give a snapshot of the effects, either using at-birth outcomes or by pooling respondents across a wide range of ages. This likely hides the potential for transmitted effects to alter as offspring age, either due to biological growth faltering or catch-up, or perhaps due to subsequent investments during their lifetime.

To the best of my knowledge, only two studies provide age-disaggregated results. Tafere (2017) uses the 1983-1985 Ethiopian famine as a natural experiment in a subsample of households located in famine-affected clusters of the Young Lives study. using panel data methods, they find that mothers exposed to the famine either in-utero or within the first three years of life are shorter on average and complete less schooling. Their children are also more likely to be likely to be shorter (5% less than average), with the intensity of exposure also being important. Additionally, pooled OLS results using a triple interaction between a famine-cohort dummy, number of months of early life exposure, and the survey round for each observation, show a U-shaped relationship between famine exposure duration and height as children age, although only statistically significant at ages 1 and 12.

Most closely related to this analysis, Bevis and Villa (2022) use an instrumental variable approach to estimate the potential transmission of health between mothers and children on Cebu island in the Philippines. They instrument variation in health with an array of early-life weather conditions, using a novel dimensionality reduction technique to derive a single value instrument for early-life weather variation. They find an early-life

[5]However, this study suffers from a weak multiple instruments problem, while it is also unclear if several instruments satisfy the exclusion restrictions required for causal interpretation.

weather-induced 1cm increase in mother's adult height is associated with a persistent effect on their child's health stock (the long-term cumulative health of an individual), measured by HAZ across childhood from age 1 until age 15. They also find an effect on birthweight and early age WAZ, their measure of health flow (which remains more sensitive to current inputs, such as maternal health), however this effect diminishes and disappears by age 8.

This study expands the evidence base for the multigenerational impacts of early life shock exposure in three key ways. First, I provide evidence of the multigenerational effects of shocks experienced specifically in the place of birth of the mother. A key limitation of the previous studies is that their identification strategy relies on defining shock exposure using the place of residence of the household during interview, often assuming zero migration between the zero and second generation, or restricting analysis to never-migrating households.[6] In the context of Peru, where significant rural-urban migration has occurred in the later half of the 20th century, this would likely introduce significant measurement error.[7] Combined with data on the month and year of birth of the mother, I am able to accurately identify shock exposure of mothers in-utero, and provide further analysis of heterogeneities across urban and rural-born mothers.

Second, I will expand the limited evidence on the potential dynamic effects of transmission. Using data from a rich longitudinal cohort study, I am able to identify how the transmission of effects to the second generation presents at specific ages, rather than at a single snapshot (Hyland & Russ, 2019; Lumey, 1992), or using a pooled sample of respondents of different ages (Fung & Ha, 2010; Khan, 2021; Venkataramani, 2011). This allows me to assess how exposure impacts postnatal growth trajectories and if effects grow or diminish as offspring age. Third, I exploit rich longitudinal data to conduct an in-depth analysis of potential mediator variables (Acharya et al., 2016), providing evidence for a primary mechanism through which effects of shock exposure are transmitted from the zero to second generation. These potential mechanisms are discussed further in the following section.

2.2.2 Transmission Mechanisms

In estimating the multigenerational effects of early life shock exposure, I also explore the potential mechanism channels through which these effects are transmitted across

[6]For example, both Tafere (2017) and Khan (2021) assume no migration in Ethiopia and India, respectively. This is unlikely to hold, particularly in India, where a large proportion of women migrate for marriage (Rosenzweig & Stark, 1989). Hyland and Russ (2019) restrict their analysis to only never-migrated households, which represents less than half of their full pooled sample. Alternatively, Bevis and Villa (2022) do not have information on where mothers are born within Cebu island, therefore weather shocks are defined for the entire study area, limiting spatial variation in shock exposure.

[7]In Peru rural-to-urban migration is generally associated with a lower prevalence of stunting and improved HAZ (Escobal & Flores, 2009) but greater overweight prevalence (Rougeaux et al., 2022) for offspring, however it is unclear how migration interacts with early life shock exposure.

generations. Following Doyle and Jernström (2023), I consider two broad channels for which shock exposure during pregnancy can transmit to future generations: i) directly via a "biological" pathway, for example through epigenetic inheritance or by affecting germ-lines; and ii) indirectly, through an "environmental" pathway, by impacting the household environment experienced by the second generation.

Within these broad channels, the exact mechanism through which the effects are primarily transmitted may vary. For the environmental channel, it is possible that in-utero shock exposure impacts the physical and cognitive development of the mother, leading to lower educational attainment and labour market outcomes (Almond et al., 2018; Black & Devereux, 2011). This could directly affect the mother's capacity for child care through reduced parenting knowledge/ability (Mani et al., 2013), or by limiting the resources available to invest in her offspring's development (Cunha & Heckman, 2007; Del Bono et al., 2016; Todd & Wolpin, 2007). Additionally, if an in-utero shock impacts maternal adulthood health, educational attainment, or socio-economic status (SES), then it may impact the quality of her chosen partner's human capital (e.g. her partner may also have poorer health or lower skills/parenting knowledge), further limiting resources available to invest in subsequent generations. For example, Behrman and Rosenzweig (2002) and Akresh et al. (2023) attribute a significant portion of the relationship between mother and child schooling outcomes to be driven by assortative matching. Furthermore, evidence from the economics literature shows a clear association between the attainment and SES of parents and the health and cognitive outcomes of their children (Almond et al., 2012; Behrman et al., 2017; Bevis & Barrett, 2015; Bhalotra & Rawlings, 2013; Black & Devereux, 2011; Black et al., 2005; Currie & Moretti, 2003, 2007; Royer, 2009). If this mechanism acts as the primary pathway for transmission then it is likely that effects persist across generations, by perpetuating economic or environmental disadvantage, suggesting that post-shock interventions are likely an effective way to mitigate intergenerational effects.

For the biological channel, it is possible that an in-utero shock can have a permanent effect on maternal physiology and metabolism, by altering or "programming" gene expression (Bale, 2015; Skinner, 2014). This could impact maternal health during pregnancy, by either creating an abnormal intra-uterine environment and/or altering her ability to transfer vital nutrients to her offspring (Gluckman & Hanson, 2004; Godfrey & Barker, 2000). Alternatively, it is possible that exposure could directly impact the germ cells (the gametes/reproductive cells), present within the first generation as a foetus while in-utero, and from which the second generation will be formed (Drake & Liu, 2010; Skinner, 2008).[8]

In support of this mechanism, a large body of research within epidemiology based on animal studies finds a persistence of in-utero nutritional shocks that last several

[8]See Figure B.1 for a visual representation.

generations, even if all subsequent generations are fed a normal diet and the mother is returned to a normal diet after the birth of her offspring (see Aiken & Ozanne, 2014; Drake & Liu, 2010, for details). For observational studies of humans, while it is hard to disentangle biological factors from environmental, an emerging body of work also provides suggestive evidence that the biological mechanism explains a large proportion of the transmission of health between generations (Bevis & Villa, 2022; Dabelea et al., 2000; Hyland & Russ, 2019; Ibáñez et al., 2000; Klebanoff et al., 1999; Van Den Berg & Pinger, 2016; Venkataramani, 2011). Therefore if this mechanism acts as the primary pathway for transmission then it suggests that post-shock interventions may have a limited ability to mitigate multigenerational effects, and that interventions that protect the zero and first generation individuals from the initial exposure to shocks should be prioritised.

2.3 Data

2.3.1 Young Lives

I use data from Young Lives (YL), a longitudinal cohort study of around 12,000 children and their families in four low- and middle-income countries (Ethiopia, India, Peru, and Vietnam) examining the causes and consequences of poverty (Favara et al., 2022). It consists of two cohorts: the younger cohort, born in 2000-2002, and the older cohort, born in 1994-1996. This analysis focuses on the younger cohort of the Peru survey, who were first interviewed in 2002 and revisited in 2006, 2009, 2013, and 2016 – at ages 1, 5, 8, 12, and 15 respectively.[9] The Young Lives Peru study employs a multi-stage, cluster-stratified, random sampling technique, and was evaluated to be suitable for analysing the causal determinants of child welfare and their longitudinal dynamics (Escobal & Flores, 2008).

The younger cohort consists of 2052 respondent children and their households in the first round. Attrition is low given extensive tracking: by round 5 (2016) attrition due to respondent refusal, moving abroad, death, or being untraceable was 9.36%, with 1860 respondents present in round 5. My analytical sample is restricted further as information on the place of birth and birth month of the mother, required to identify her exposure to early life drought, was collected in round 4 only if the mother is still alive and residing in the same household as the child. The place and date of birth was derived for a total of 1734 mothers. I focus only on the drought exposure of maternal grandmothers, rather than paternal grandmothers, due to the practical limitation that birth location

[9] Additionally, 5 rounds of phone surveys were conducted throughout 2020-2021 during the global COVID-19 pandemic when younger cohort respondents were aged between 18-20 (Favara et al., 2022). As no physical health measures or cognitive ability tests could be administered, these survey waves are not considered in this analysis.

is only available for a subsample of fathers present during interview and therefore is missing for significant number of respondents in a likely non-random pattern, which could produce biased estimates. However, current evidence from the literature indicates that matrilineal transmission is generally more important for the effects of early life shock exposure (Caruso, 2015; Fung & Ha, 2010; Painter et al., 2008; Venkataramani, 2011). A further theoretical argument is that there are clearer potential transmission channels between maternal and offspring health than for paternal health if transmission occurs predominantly through a biological pathway.

Table 2.1: Summary Statistics: Comparison of Baseline Characteristics

	Omitted			Attrition/unbalanced observation		
	(1)	(2)	(3)	(4)	(5)	(6)
	R1 Sample	Omitted	Diff. (2)-(1)	Balanced	Omitted	Diff. (5)-(4)
Child outcomes						
HAZ	-1.29	-1.34	-0.05	-1.26	-1.76	-0.50***
	(1.27)	(1.42)	[0.07]	(1.26)	(1.35)	[0.12]
Stunted	0.28	0.31	0.04	0.27	0.40	0.13***
	(0.45)	(0.46)	[0.03]	(0.44)	(0.49)	[0.04]
WAZ	-0.21	-0.16	0.05	-0.19	-0.51	-0.32***
	(1.17)	(1.32)	[0.07]	(1.16)	(1.28)	[0.11]
Underweight	0.07	0.09	0.02	0.06	0.15	0.09***
	(0.25)	(0.29)	[0.01]	(0.24)	(0.36)	[0.02]
Female	0.49	0.53	0.03	0.49	0.47	-0.02
	(0.50)	(0.50)	[0.03]	(0.50)	(0.50)	[0.05]
Mother outcomes						
Height (in cm)	149.96	150.29	0.33	150.01	149.08	-0.94
	(5.56)	(5.51)	[0.37]	(5.50)	(6.35)	[0.57]
Weight (in Kg)	58.70	58.35	-0.36	58.77	57.50	-1.28
	(9.94)	(9.99)	[0.66]	(9.99)	(9.17)	[1.08]
Grade attainment	7.13	7.17	0.04	7.21	6.01	-1.20***
	(4.51)	(4.73)	[0.29]	(4.52)	(4.29)	[0.45]
Age in years	27.02	25.98	-1.04***	27.08	26.10	-0.98
	(6.70)	(7.00)	[0.39]	(6.65)	(7.41)	[0.67]
HH outcomes						
HH size	5.70	5.74	0.05	5.69	5.72	0.02
	(2.36)	(2.23)	[0.13]	(2.34)	(2.57)	[0.23]
Wealth index	0.43	0.41	-0.01	0.43	0.35	-0.08***
	(0.24)	(0.24)	[0.01]	(0.23)	(0.24)	[0.02]
N	1670	382	2052	1560	110	1670

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Columns (1)-(2) and (4)-(5) provide mean values, with standard deviations in parentheses. (1) Provides baseline summary statistics for the analytical sample at round 1. column (2) presents values for those observations present in the full cohort that are omitted from the analytical sample. columns (3) and (6) provide the difference in means from a 2-sample t-test, with standard errors in square brackets.

Once missing outcomes and covariates are accounted for and singleton observations

are dropped the analytical sample consists of 1670 mother-child pairings present in round 1 (R1). Age 1 sample summary statistics for those included in round 1 and those omitted are provided in columns (1) and (2) of Table 2.1. The difference in means are reported with standard errors from two-sample t-tests in brackets in column (3), and indicates that those omitted from the R1 sample do not differ significantly in terms of baseline child, mother, or household characteristics, with exception that mothers in my sample are roughly 1 year older on average than those omitted.

Notably, my sample size varies slightly across rounds as some respondents were absent during one or more interim survey rounds. Additionally, observations are set to missing in a round if their measure for either height-for-age or weight-for-age are flagged as biologically implausible following WHO standards (See Briones, 2018, for details). In my sample this explains all differences between round 1 and round 4 given data for mother's birth location and month are first collected in round 4, as discussed above. therefore further attrition from the sample occurs only between round 4 and round 5. Column (4) and (5) provide age 1 summary statistics for those who appear in all rounds (balanced panel) compared with those observations which are either missing in at least one interim round or attrited between rounds 4 and 5. Column (6) indicates that there are large difference between those in the balanced panel and those dropped, in particular respondents are significantly more likely to be stunted and underweight, with significantly lower height- and weight-for age. Households have lower wealth scores and mothers have lower educational attainment (highest grade achieved). Additionally, while the difference is not statistically different from zero, mothers are also almost a 1cm shorter, 1.28Kg lighter, and 1 year younger than those in the balanced panel. If shock exposure is negatively related to second generation outcomes and to either mothers health or socioeconomic outcomes, which may play an important role as potential pathways for transmission, this suggests that excluding these observations from this analysis may downwards bias estimates of the effect of shock exposure on the outcomes of interest. Therefore for my primary specification I use the full unbalanced panel in each round. This decision is consistent with other work which provides estimates of second generation effects at different ages (Bevis & Villa, 2022; Khan, 2021). However, as shown in subsection 2.5.4, results remain robust to using the balanced panel.

2.3.2 Outcomes

2.3.2.1 Second Generation

The impact of exposure of mothers (first generation) to drought while in-utero during the grandmother's (zero generation) pregnancy is measured on the outcomes of the second generation children in two dimensions of human capital: health and cognitive ability.

Health is measured using anthropometric outcomes related to an individual's height

and weight. Health stock is measured using child height-for-age z-scores (HAZ).[10] Child growth is seen as a high-quality indicator of child health, capturing the cumulative effect of health shocks/investments, nutrition, and environmental factors (Case & Paxson, 2008; De Onis, 2017; Martorell & Habicht, 1986). Child health flow in early years is captured by weight-for-age (WAZ), which is more susceptible to current health inputs (Bevis & Villa, 2022; WHO, 1995) than health stock. WHO reference tables for weight-for-age are provided only up to age 10, as it is considered inadequate for monitoring growth beyond childhood (De Onis et al., 2007). As such, Young Lives provide weight-for-age scores only up to round 3, when children were aged 7.5-8.5 years old (Briones, 2018). For rounds 4-5 (roughly ages 12 and 15 respectively) health flow is therefore measured as BMI-for-age z-scores (BMIAZ). However, this is an imperfect measure as it is constructed using both weight and height, and therefore may mask changes in weight if accompanied by changes in height in the same direction. Z-scores for child growth are preferred to using raw measures as they provide an indication of how a child's growth compares with that of a healthy individual of the same age and gender. Observations which are flagged as biologically implausible based on WHO standards are dropped.

Mean height-for-age (age 1-15), weight-for-age (1-8), and BMI-for-age (available for ages 1-15) are presented in Table 2.2. Indicator variables are also defined for a child being classed as stunted, underweight, or wasted if their age-specific z-score is less than 2 standard deviations (S.D.) from the mean for HAZ, WAZ, and BMIAZ respectively, based on WHO reference tables (De Onis, 2017; De Onis & Habicht, 1996). The sample is relatively short, with a mean score of less than -1 S.D. below the reference average height for a child of the same gender at all ages. The incidence of stunting (HAZ $\leq$ -2 S.D.) in early years is relatively high, with approximately 28% stunted at age 1. Stunting peaks at age 5 at 33%, before falling to 16% by age 15. In contrast, the incidence of underweight (WAZ $\leq$ -2 S.D.) is relatively low, at only 7% at age 1, decreasing slightly to 5% for ages 5 and 8, with average WAZ remaining around -0.20 to -0.53 S.D. from reference values. Interestingly, the rate of wasting is very low (0.3% to 2%), with the average BMI-for-age being positive. This potentially reflects the well documented "double burden" of malnutrition in Peru (and amongst many other middle-income countries), a recent trend which has seen the simultaneous coexistence of high levels of childhood stunting (or under-nutrition), and an increasingly high prevalence of child and teenage overweight/obesity, particularly amongst girls and women in rural, poor areas (Perez-Escamilla et al., 2018; Santos et al., 2021; WHO, 2017). As such, adolescents with short stature and relatively high weight will display relatively higher BMI scores (calculated as weight in kilograms divided by height in metres squared, kg/m^2), however may still exhibit poor health. This highlights a potential limitation of using anthropometrics, particularly BMI, as a measure of overall health.

[10] For age 1 this is measured as length-for-age using a board.

Table 2.2: Summary Statistics: Time-Varying Child Outcomes

	Age 1	Age 5	Age 8	Age 12	Age15
Anthropometrics					
Height-for-age	-1.28	-1.53	-1.15	-1.02	-1.15
	(1.27)	(1.11)	(1.04)	(1.10)	(0.88)
Stunted	0.28	0.33	0.20	0.18	0.16
	(0.45)	(0.47)	(0.40)	(0.39)	(0.37)
Weight-for-age	-0.20	-0.53	-0.33		
	(1.18)	(1.01)	(1.17)		
Underweight	0.07	0.05	0.05		
	(0.25)	(0.22)	(0.23)		
BMI-for-age	0.73	0.66	0.51	0.55	0.42
	(1.18)	(0.95)	(1.04)	(1.07)	(0.97)
Wasted	0.02	0.00	0.01	0.01	0.01
	(0.15)	(0.05)	(0.09)	(0.10)	(0.09)
Cogntive scores					
Vocabulary z-score		0.01	0.03	0.01	0.01
		(1.00)	(0.99)	(1.00)	(1.00)
Mathematics z-score		0.03	0.03	0.03	0.02
		(0.99)	(1.00)	(0.99)	(1.00)

Notes: Sample mean values, with standard deviations in parentheses.

Cognitive ability is measured across two sub-dimensions. First, using the Spanish version of the revised Peabody Picture Vocubulary Test (PPVT) (Dunn et al., 1986), a widely-used and well-validated assessment of vocabulary acquisition.[11] The test is administered orally, is un-timed, and norm-referenced. While the 125 items in the test are the same for all ages, not all are administered in each test, with only a subset of questions administered after a basal item and ceiling item are established, depending on the number of consecutive correct/incorrect responses (Espinoza Revollo & Scott, 2022; Leon, 2020). This measure is first made available in round 2, when younger cohort respondents were aged 5, and was administered until round 5 when respondents are aged 15.

Second, respondents' quantitative skills are assessed. In round 2, the Cognitive Development Assessment (CDA), developed by the International Evaluation Association to study the effect of preschool attendance on cognitive development in children, was administered to the younger cohort. Given the long administration time for the full assessment, only the quantity subscale (15 items) was administered (Espinoza Revollo & Scott, 2022). Beginning in round 3, mathematics tests were administered to respondents, based on previously validated items from the Trends in International Mathematics

[11] *Test de Vocabulario en Imagenes Peabody* (TVIP) in Spanish. This test is adapted and validated for use in Latin America. Additionally, it has been further translated and validated by the Young Lives team to make it available for children whose primary language is Quechua.

and Science Study (TIMSS) and the Programme for International Student Assessment (PISA). The contents of maths tests differ slightly across rounds, and differ from the CDA administered in round 2. Therefore to provide a more relevant measure, raw scores are age-standardised, to provide a measure of relative performance within cohort.[12] For further details, see Espinoza Revollo and Scott (2022) and Leon (2020).

2.3.2.2 First Generation

While the primary focus of this study is the impact on grandchildren's outcomes, I also assess the relevancy of the shock on the adult age outcomes of mothers who are exposed while in-utero. In doing so, this may provide an insight in to the potential channels through which effects are transmitted to the second generation, however it does not provide information on the relevance or importance of this variable as a mediator for the multigenerational effect, and I cannot rule out that there are unobserved variables which determine both mother and grandchild outcomes. Therefore, to address the potential mechanism channels formally, I conduct a mediation analysis in section 2.6. Following from subsection 2.2.2, I hypothesis two potential channels, the biological and environmental mechanism. If effects are predominantly transmitted to grandchildren through the biological channel, then it could be expected that mother's exposed to a shock in-utero would display signs of poorer adult health. I use height (cm) and weight (kg) as proxies of mothers' health stock and flow, respectively. Alternatively, if effects are transmitted through the environmental channel by directly impacting maternal cognitive development, then it could be expected that shock exposure has some impact on measures of adult human capital accumulation or socioeconomic status. I use mother's educational attainment, measured as years of completed schooling, as well as her household wealth index to capture these dimensions.[13] Finally, if effects are transmitted indirectly through the environmental channel by impacting mothers' mating/marriage prospects, then it is expected that they may match with lower quality partners. To capture this potential indirect effect, I also regress the educational attainment and height of the father of the child, where available, on mother's shock exposure.

2.3.3 Standardised Precipitation-Evapotranspiration Index

The first generation mothers were born between 1950 and 1988 in 381 districts across Peru, with median year of birth being 1976, and the youngest mother being aged 13 at the

[12] Leon (2020) derive cross-round comparable scores for PPVT and Math tests (not including the CDA subscale) based on item response theory (IRT) – Results remain unchanged when using IRT scores rather than age-standardised scores.

[13] This is a country-specific composite measure of household socioeconomic status, measuring households' access to services such as water and sanitation, their ownership of consumer durables, and the quality of materials used in their dwelling. See Briones (2017) for details.

birth of the YL child (second generation).[14] To identify the exposure of the grandmother (zero generation) to drought shocks during her pregnancy, I match data on historical drought exposure from the Standardised Precipitation-Evapotranspiration Index (SPEI) (Begueria et al., 2010; Vicente-Serrano et al., 2010). It provides a multi-scalar drought measure, which accounts for the effects of temperature and potential-evapotranspiration (PET, i.e. the amount of water that is used by plants, or evaporates from the surface, under local normal conditions) on the intensity and duration of droughts. It has been shown to perform better in predicting changes in crop yields and local weather conditions over other common drought indices (Vicente-Serrano et al., 2012), while retaining the simplicity of calculation and multi-temporal nature of probabilistic measures such as the Standardised Precipitation Index (McKee et al., 1993). It has seen increased use in recent economics literature, most notably Harari and Ferrara (2018).

I use monthly gridded data ($0.5°$ resolution) derived as the across a 12-month rolling window time-scale from the SPEIbase (v2.9) database (Begueria et al., 2023), which covers the period 1901-2022. Cell values are aggregated to the district level (3rd level administrative area) as the area-weighted mean value of all overlapping grid cells, to provide a monthly district-level time series. The index is normalised with mean zero and standard deviation (S.D.) one. Following the drought classification system of McKee et al. (1993), I define a drought shock experienced during the grandmother's pregnancy:

$$Drought_{m,d,t} = SPEI_{d,t} \leq -1S.D. \tag{2.1}$$

Where $SPEI_{d,t}$ is the SPEI value for the 12 months preceding the month of birth t of mother m, in the district of birth d, such that $Drought_{m,d,t}$ takes a value of one when conditions are less than or equal to one standard deviation worse than the long-run location specific mean conditions, and zero otherwise.[15]

2.4 Empirical Strategy

2.4.1 Second Generation Effects

To examine the multigenerational effect of exposure of the grandmother while pregnant to drought on the outcomes of her grandchildren, I estimate the following equation using OLS:

$$Y^a_{c,m,v} = \beta_0 + \beta^a_1 Drought_{m,d,t} + \beta_2 female_{c,m} + \gamma_y + \delta_p + \rho_0 + \sigma^a_v + \varepsilon_p \tag{2.2}$$

[14] See Figure B.2.

[15] That is, if the mother was born in July 1979, the SPEI value refers to the deviation in conditions between August 1978 and July 1979 from the long-term average.

Where $Y^a_{c,m,v}$ are the outcomes of child c, of mother m, at each age a. Health outcomes are estimated for each age $a \in \{1, 5, 8, 12, 15\}$ separately. As discussed above, health stock is proxied by height-for-age z-scores from ages 1-15. Similarly, health flow is measured using weight-for-age and BMI-for-age z-scores for ages 1-8 and 12-15, respectively. Additionally, I assess the effects on cognitive ability, estimating for ages $a \in \{5, 8, 12, 15\}$ the impact on age-standardised PPVT and maths scores (age 5 maths is measured using the CDA quantity subscale, while age 8-15 is measured using YL mathematics tests (Espinoza Revollo & Scott, 2022)).

An indicator that the child c is female is included, as well as a fixed effect (σ^a_v) for child cluster of residence v at age a (for $a = 1$, this is their cluster of birth) and month of birth cohort (ρ_0). Fixed effects for mothers' year-of-birth (γ_y) and province-of-birth (δ_p) are also included. Standard errors are clustered at the level of mother's province of birth (ϵ_p), to account for localised spatial correlation in shock exposure, which varies at the district level.[16]

2.4.2 Heterogeneous Treatment Effects

A common finding within the literature is that the effects of early life investments or shocks can be sex-specific (Almond & Currie, 2011; Almond et al., 2018), and previous evidence suggests this may extend to multigenerational effects (Fung & Ha, 2010; Venkataramani, 2011). To explore whether the multigenerational effect is different for male and female grandchildren, I expand on Equation 2.2, including an interaction between grandmothers exposure to drought during pregnancy with if their grandchild is female.[17]

Additionally, it is unclear *a priori* if effects found in childhood and early adolescence represent a permanent impact on growth, or simply represent slow growth during childhood, with a subsequent catch-up once children enter into puberty. Therefore I assess how the effects of grandmother's exposure to drought interact with respondents entering into pubertal growth, as measured using reported physical indicators of puberty, discussed further below.

It is expected that drought shocks impact the zero and first generation either directly, through impacting local crop yields or food prices, impacting food availability, or by impacting agricultural income and hence resources or nutrition available to the grandmother during pregnancy or for the mother immediately after her birth. As such

[16] Results are also robust to clustering standard errors at the cluster of birth, which is the sampling level. Results are presented in Table B.1.

[17] While previous work has noted effects of extreme famine on sex ratios of the second generation (Almond et al., 2012), which would suggest the sex of the grandchild may be endogenous, it is unclear if this effect extends to this context, where drought exposure is less intense and is less likely to present issues of selective mortality. While not directly testable in this context, drought exposure does not predict the sex of grandchildren within the sample, as shown in column 6 of Table B.12.

it is expected that the effect is either driven exclusively by, or is at least strongest in, rural areas, where a higher proportion of households are likely to be reliant on local food sources and agriculture-related income. Therefore, I also interact the drought exposure indicator with an indicator of if a mother was born in an urban or rural district. Finally, it is common within the early life shocks literature that in-utero shock exposure in a specific trimester may have a stronger effect than in other periods (Almond & Mazumder, 2013). Indeed, within the second generation literature Khan (2021) and Stein and Lumey (2000) find exposure earlier in the pregnancy during the first and second trimester has the largest effect on second generation outcomes. I therefore assess if effects of drought exposure differ by exposure in specific trimesters of pregnancy.

2.4.3 First Generation Effects

I also assess the direct impact of in-utero shock exposure on the adult age outcomes for the first generation (the mother), estimating the following equation:

$$M_m = \alpha_0 + \alpha_1 Drought_{m,d,t} + \gamma_y + \gamma_p + \nu_p \tag{2.3}$$

Where M_m are the health and human capital outcomes for the mother m, measured by her adult height and weight, educational attainment, and household socioeconomic status. Additionally, to capture potential effects on mother's mating/marriage market prospects, I regress the height and educational attainment of the father of the child on mother's shock exposure. Maternal fixed effects for year and province of birth are as above. While this exercise may provide suggestive evidence of potential mechanisms, it does not provide information on the relevance or importance of this variable as a mediator for the multigenerational effect on grandchildren, therefore I conduct a formal mediation analysis in section 2.6.

2.5 Results

2.5.1 First Generation Effects: Mother Outcomes

Before addressing the impacts of drought on the outcomes of the grandchildren, I first assess the evidence of effects being present in the first generation, who were in-utero during the shock exposure. Panel A of Table 2.3 provides estimates of the impact of an in-utero exposure to drought on their individual adult age outcomes, while Panel B provides estimates on the potential indirect effect of shock exposure on household and partner outcomes, as described in subsubsection 2.3.2.2. The only dimension for which a significant effect is estimated is in mother's adult height, my measure of long-run health stock. Exposure to drought in-utero is associated with a -0.752cm lower height

in adulthood (-0.5% decrease compared to the sample mean). There are no significant effects found for measures of educational attainment, household socio-economic status, or on the outcomes of their partner. Additionally, there is no significant effect found on adult age health flow, which is more susceptible to current inputs. This is suggestive that shock exposure may have an effect on mother's long-run physical health, and with no evidence of an effect on household socio-economic status, attainment, or selective pair-bonding, could indicate that any transmitted effects may operate predominantly through a biological channel. However, while much of the variation in adult height in developing countries is thought to be due to negative shocks experienced in early life, rather than genetic potential (Beard & Blaser, 2002; Silventoinen, 2003; WHO, 1995), it is an imperfect measure. It is likely the result of all unobserved cumulative health shocks and investments experienced in childhood and early adulthood, which may be positively or negatively correlated with the prenatal shock of interest, potentially biasing estimates of the effect. As such, with this method I cannot rule out that there is some unobserved mediator which acts as an intermediate confounder between in-utero shock exposure, maternal height, and second generation outcomes. Therefore inferences about the causal mechanisms for the zero to second generation effects are restricted to those found using formal mediation analysis in section 2.6.

Table 2.3: Effect of Shock Exposure on First Generation Outcomes

	Height	Weight	Education
Panel A: Mother outcomes			
In-utero shock	-0.752	-0.590	0.201
	(0.324)**	(0.580)	(0.247)
Mean	149.97	58.78	7.15
N	1656	1632	1671
	Father		Household
	Height	Education	Wealth
Panel B: Father/Household outcomes			
In-utero shock	0.325	-0.067	-0.000
	(0.315)	(0.294)	(0.010)
Mean	162.18	8.25	0.43
N	1168	1425	1671

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Cluster robust standard errors in parentheses. Fixed effects for mother year- and province- of birth are suppressed. Sample mean values for dependent variables are reported in the foot of each panel.

2.5.2 Second Generation Effects: Grandchild Outcomes

This section presents the estimated multigenerational effects of a grandmother's exposure to a drought shock while pregnant on the outcomes of her grandchild. I first present the impact on anthropometric outcomes, followed by cognitive outcomes.

2.5.2.1 Anthropometry

Panel A of Table 2.4 and Figure 2.1 present estimates the impact on second gerneration health stock, measured as height-for-age z-scores for ages 1-15. Panel B of Table 2.4 and Figure 2.2 provides estimates of the effect on health flow, measured as weight-for-age for ages 1-8, BMI-for-age for ages 12-15.

Table 2.4: Effect of Shock Exposure on Second Generation Outcomes: Anthropometric Z-scores

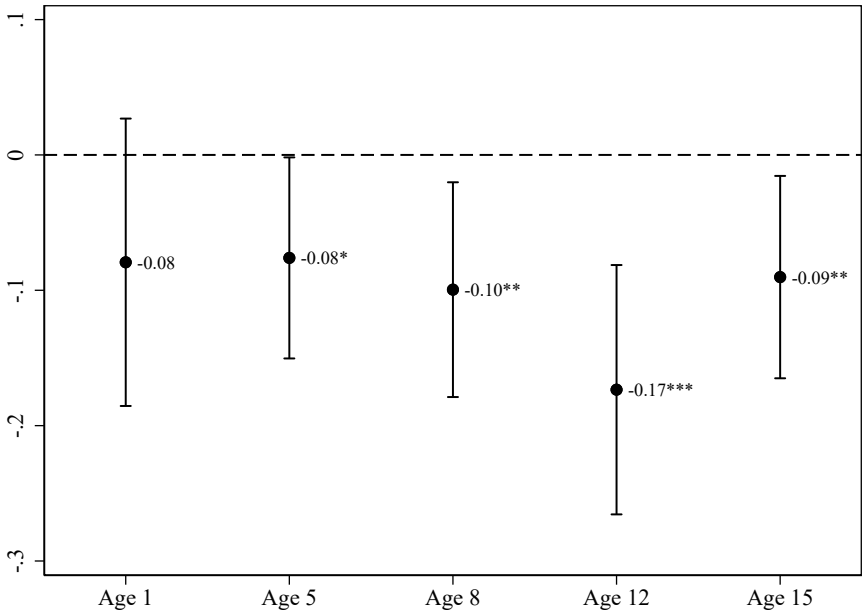
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
In-utero shock	-0.079	-0.076	-0.100	-0.173	-0.090
	(0.064)	(0.045)*	(0.048)**	(0.055)***	(0.045)**
Controls	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
In-utero shock	-0.179	-0.109	-0.059	0.045	-0.021
	(0.063)***	(0.047)**	(0.056)	(0.049)	(0.059)
Controls	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

A persistent negative effect is estimated for HAZ at each age (-0.076 to -0.173 S.D.). While I cannot reject the null hypothesis $H_0 : \hat{\beta}_2^{a=1} = 0$, for age 1 HAZ, estimated effects beginning age 5 and persisting to age 15 are statistically significant at conventional levels, with the largest difference measured at age 12. Although there is a peak at age 12, given overlapping confidence intervals, it is unclear if this effect is increasing in age until early adolescence, or simply persistent.[18] The result of a persistent effect on child height are consistent with the findings of Bevis and Villa (2022) who find a lasting relationship

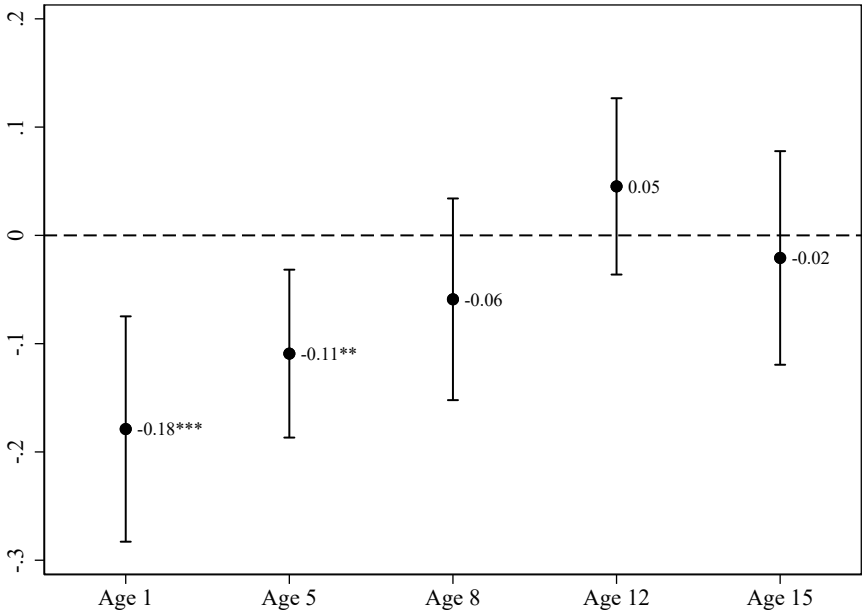
[18] Additionally, confidence intervals are widest for age 1 HAZ, for which measurement was conducted using a length board rather than height scale. This may indicate greater measurement error, as is common with length measurements where it can be difficult and stressful to ensure a struggling/crying infant is laying fully stretched out for measurement (WHO, 2006).

Figure 2.1: Effect of Shock Exposure on Second Generation Outcomes: HAZ



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. 90% Confidence intervals. Coefficients reported in Table 2.4. P-values calculated using cluster robust standard errors. Controls include an indicator for if the child is female and fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Figure 2.2: Effect of Shock Exposure on Second Generation Outcomes: WAZ/BMIAZ



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. 90% Confidence intervals. Coefficients reported in Table 2.4. P-values calculated using cluster robust standard errors. Controls include an indicator for if the child is female and fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

between a 1cm increase in mother's height due to early life weather shocks (using an novel instrumental variable approach) and child height-for-age measured across their whole childhood. However, this contrasts with the findings of Khan (2021) in India, who finds that while in-utero shock exposure is associated with a negative impact on HAZ for a pooled sample aged 0-5, there is no significant effect on HAZ in the age 8-11 sample.

A large negative effect is estimated for weight-for-age at ages 1 and 5 (-0.179 and -0.109 S.D.), significant at the 1% and 5% level respectively. At older ages the estimated effect size diminishes towards zero, with no significant effect estimated for weight-for-age at age 8, or for BMI-for-age at ages 12 and 15. This finding, and the effect on maternal health stock noted above, complements the literature indicating the strong association between maternal health and children's weight in early years (Currie & Moretti, 2007; Hyland & Russ, 2019). However, if the impact of shock is transmitted between mother and child through its effect on the biological channel, then it also provides a similar conclusion to that of Bevis and Villa (2022), that while (early-life shock-induced) variation in mother's health is important for child health flow in early years, the significance and magnitude diminishes as children age. This is intuitive, given that health flow is more variable with current period inputs, the importance of maternal health endowment becomes less important as children age. Alternatively, the estimated impact on health stock suggests that the transmitted effect of the shock seems to have a more permanent effect on long-term cumulative health. I explore this possibility further in section 2.6.

The average effect sizes are relatively small (-0.076 to -0.179 S.D.) compared to the large deficits that categorise stunting and wasting (z-scores $\leq$ -2 S.D.), however as previously documented in Table 2.2, the sample mean is generally below average height and weight at each age (with the exception of having a slightly positive BMI in adolescent years). Given the importance of these categories for global health targets and policy decisions (De Onis, 2017), I consider the multigenerational effect of shock exposure on the extensive margins of growth. Estimating a linear probability model by OLS, Table 2.5 provides estimates of the effect of grandmother's shock exposure on probability of stunting (HAZ $\leq$ -2 S.D.) in Panel A, and underweight/wasting (WAZ and BMIAZ $\leq$ -2 S.D., respectively) in Panel B. Coefficients are reported in percentage points (p.p.). While some estimates are not significant (notably age 12 stunting and age 1 underweight) the pattern remains similar to that seen in the intensive margin, with a positive impact of the shock on the probability of stunting of 6.4 p.p. (a 32.0% increase relative to the sample mean) and 5.5 p.p. (34.4% relative increase) for ages 8 and 15, respectively. For health flow, shock exposure is associated with a 3.4 p.p. (68%) increase in probability of being underweight at age 5, all significant at the 5% level. This suggests that while the magnitude of the effect on the intensive margin appears relatively small, there is a notable multigenerational effect of drought exposure on commonly used health targets, adding relevance to these findings for future health policy. I now turn to

the results for child cognitive outcomes.

Table 2.5: Effect of Shock Exposure on Second Generation Outcomes: Stunting & Wasting

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Stunted					
In-utero shock	0.031	0.014	0.064	0.033	0.055
	(0.028)	(0.023)	(0.026)**	(0.027)	(0.023)**
Controls	Yes	Yes	Yes	Yes	Yes
Mean	0.28	0.33	0.20	0.18	0.16
N	1670	1657	1665	1671	1620
	Underweight			Wasted	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Underweight/Wasted					
In-utero shock	0.004	0.034	0.026	-0.001	0.001
	(0.016)	(0.015)**	(0.016)	(0.006)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes
Mean	0.07	0.05	0.05	0.01	0.01
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in percentage points. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed. Sample mean values for dependent variables are reported at the foot of each panel.

2.5.2.2 Cognitive Ability

Considering the implications of shock exposure on grandchild cognitive ability, Table 2.6 shows the estimated effect of exposure on age-standardised scores for receptive vocabulary (Panel A) and quantitative skills (Panel B). Receptive vocabulary is measured using the Spanish-version of the PPVT (Dunn et al., 1986) for all ages, while quantitative skills are measured by the CDA quantity sub-scale for age 5 and using Young Lives mathematics tests for ages 8-15. All scores are age standardised. There is little evidence of an effect on cognitive ability as measured by these indicators, with exception of a negative impact of shock exposure on performance in the CDA quantity sub-scale at age 5, however this effect is only marginally significant at 10%. Given limited evidence of an effect on cognitive outcomes, the remainder of this chapter will focus on the multigenerational impacts on health. Next I assess the potential for heterogeneous effects.

Table 2.6: Effect of Shock Exposure on Second Generation Outcomes: Cognitive Scores

	Age 5	Age 8	Age 12	Age 15
Panel A: Vocabulary score				
In-utero shock	-0.041	0.027	-0.011	-0.017
	(0.046)	(0.057)	(0.060)	(0.057)
Controls	Yes	Yes	Yes	Yes
N	1620	1562	1624	1580
	CDA	Young Lives tests		
	Age 5	Age 8	Age 12	Age 15
Panel B: Maths score				
In-utero shock	-0.115	-0.045	0.015	0.006
	(0.058)*	(0.050)	(0.062)	(0.051)
Controls	Yes	Yes	Yes	Yes
N	1620	1562	1624	1580

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age-standardised sample mean. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

2.5.3 Heterogeneous Effects

2.5.3.1 Sex Specific Differences

Panel A of Table 2.7 and Figure 2.3 show the estimated average marginal effects of grandmothers' shock exposure on HAZ of male and female grandchildren. Coefficients plotted are the average marginal effects for a discrete change in the indicator of grandchild sex. Regression coefficients for the level and interaction terms are reported in Table B.2. Results suggest that the majority of the impact of shock exposure on HAZ is driven by the effect on boys. This is consistent with Venkataramani (2011), who finds that shock-induced variation in maternal height is most strongly associated with boys height against a null effect for girls, and Fung and Ha (2010), who find that the second generation effects on HAZ and WAZ of maternal in-utero exposure to the great Chinese famine is isolated to boys. Additionally for age 1 length-for-age, there is a large negative effect of shock exposure for boys, compared with a small insignificant positive effect on girls. This could suggest that, at least for boys, the effect on height could be present across their whole childhood from age 1, however this effect remains imprecisely estimated with wide confidence intervals, and is only significant at the 10% level.[19] That the effect of shock exposure on health stock is isolated to boys means that it is not clear whether the multigenerational shock could persist beyond the second generation. Given

[19]This imprecision may reflect greater measurement error for infant length-for-age, given reported difficulties with keeping infants still and stretched out.

limited information on the birth location of fathers in my sample, and that I do not currently observe the third generation in this dataset, I am not able to assess how effects may persists patrilineally, however evidence from the wider literature provides evidence that matrilineal transmission is more important for second generation effects (Caruso, 2015; Fung & Ha, 2010; Painter et al., 2008; Venkataramani, 2011).

Panel B of Table 2.7 and Figure 2.4 show estimates for WAZ and BMIAZ. While effects are consistently more negative for boys, in contrast to HAZ, infant WAZ is significantly different from zero for both boys and girls. Overall results indicate that there is a sex-specific difference in multigenerational effects, with the largest effects found for boys.

Table 2.7: Marginal Effect of Shock Exposure on Second Generation Outcomes: By Sex

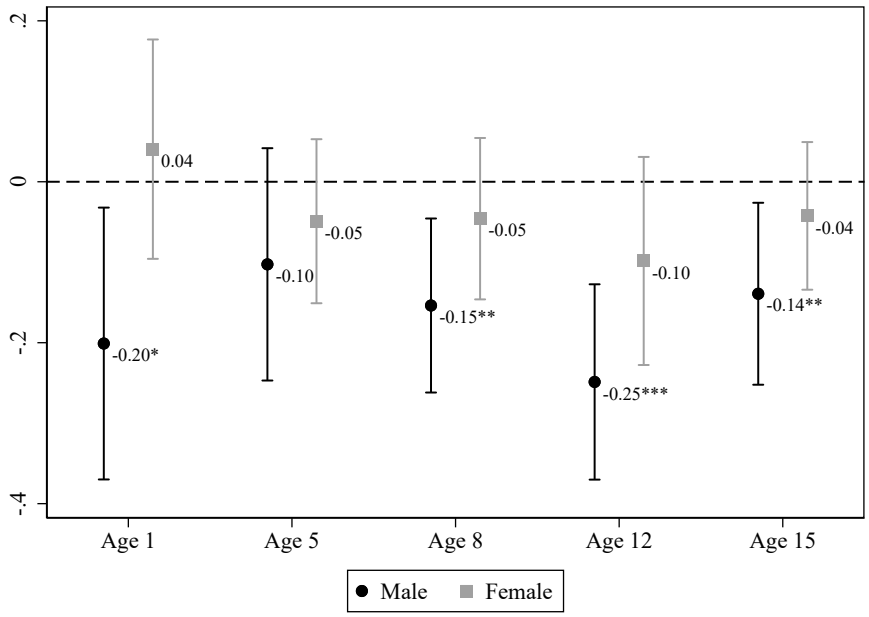
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
In-utero shock × Female = 0	-0.201	-0.103	-0.154	-0.249	-0.139
	(0.103)*	(0.088)	(0.066)**	(0.074)***	(0.069)**
In-utero shock × Female = 1	0.040	-0.049	-0.046	-0.098	-0.042
	(0.083)	(0.062)	(0.061)	(0.079)	(0.056)
N	1670	1657	1665	1671	1620
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
In-utero shock × Female = 0	-0.230	-0.167	-0.113	0.010	-0.093
	(0.083)***	(0.058)***	(0.077)	(0.063)	(0.075)
In-utero shock × Female = 1	-0.128	-0.050	-0.006	0.080	0.050
	(0.071)*	(0.067)	(0.077)	(0.087)	(0.092)
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Table reports the marginal effects of shock exposure on second generation HAZ (Panel A) and WAZ/BMIAZ (Panel B). Relevant regression coefficients are reported in Table B.2. Cluster robust standard errors in parentheses. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

2.5.3.2 Urban/Rural Differences

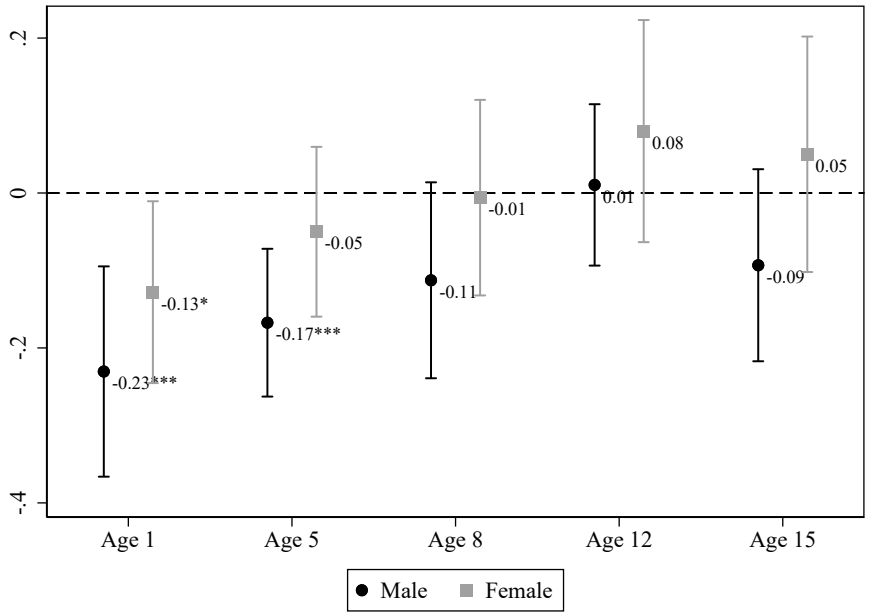
Panel A of Table 2.8 and Figure 2.5 show the marginal effects for HAZ of the second generation child for mothers who were born in urban areas compared with rural areas. While a large and significant effect is estimated for rural-born mothers (urban-born = 0) for ages 5-15, the marginal effect for the children of urban-born mothers is insignificant across all rounds, with inconsistent sign and diminished magnitude. Similarly, Panel B of the same table and Figure 2.6 show large and statistically significant effects for age 1-5 WAZ for a baseline effect, with an insignificant marginal effect estimated for the children of urban-born mothers. This evidence suggests that the overall effect seems to

Figure 2.3: Marginal Effect of Shock Exposure on Second Generation Outcomes: HAZ, by Sex



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. 90% Confidence intervals. Marginal effects reported in Table 2.7. P-values calculated using cluster robust standard errors. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Figure 2.4: Marginal Effect of Shock Exposure on Second Generation Outcomes: WAZ/BMIAZ, by Sex



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. 90% Confidence intervals. Marginal effects reported in Table 2.7. P-values calculated using cluster robust standard errors. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

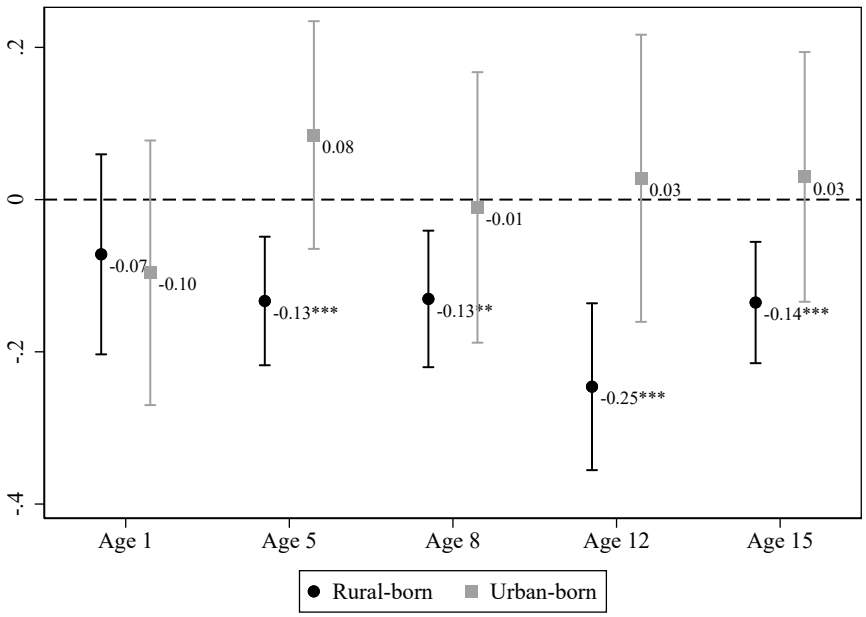
be driven by the large and significant effect for the children of mothers born outside of urban areas.

Table 2.8: Marginal Effect of Shock Exposure on Second Generation Outcomes: By Mother Birth-Location

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
In-utero shock × Urban-born = 0	-0.072	-0.133	-0.131	-0.246	-0.135
	(0.080)	(0.051)***	(0.055)**	(0.067)***	(0.048)***
In-utero shock × Urban-born = 1	-0.096	0.085	-0.010	0.028	0.030
	(0.106)	(0.091)	(0.108)	(0.115)	(0.100)
N	1670	1657	1665	1671	1620
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
In-utero shock × Urban-born = 0	-0.213	-0.144	-0.094	0.061	-0.036
	(0.067)***	(0.054)***	(0.069)	(0.063)	(0.064)
In-utero shock × Urban-born = 1	-0.084	-0.010	0.041	0.002	0.023
	(0.127)	(0.108)	(0.116)	(0.070)	(0.098)
N	1670	1657	1665	1671	1620

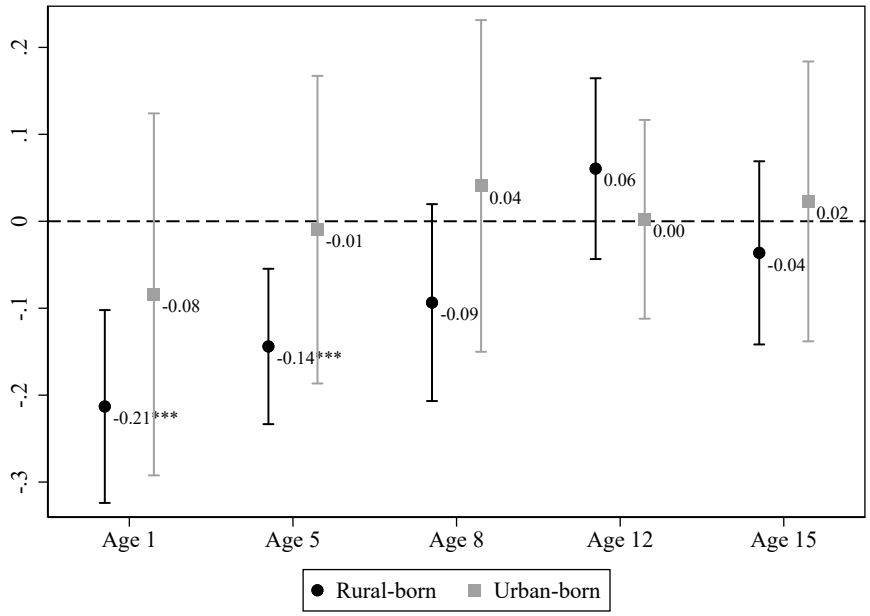
Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Table reports the marginal effects of shock exposure on second generation HAZ (Panel A) and WAZ/BMIAZ (Panel B). Relevant regression coefficients are reported in Table B.3. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Figure 2.5: Marginal Effect of Shock Exposure on Second Generation Outcomes: HAZ, by Mother Birth-Location



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. 90% Confidence intervals. Marginal effects reported in Table 2.8. P-values calculated using cluster robust standard errors. Indicator for if the child is female and fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Figure 2.6: Marginal Effect of Shock Exposure on Second Generation Outcomes: WAZ/BMIAZ, by Mother Birth-Location



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. 90% Confidence intervals. Marginal effects reported in Table 2.8. P-values calculated using cluster robust standard errors. Indicator for if the child is female and fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

2.5.3.3 Growth Stage Differences

Notably, the results in Table 2.4 suggest that second generation effects for height-for-age are most pronounced between age 8 and 15, with the largest effect at age 12. This peak may suggest that impacts are largest around when grandchildren may be transitioning between childhood growth and pubertal growth, but it is unclear if the gap between exposed and unexposed individuals remains stable once they enter into pubertal growth. It is possible the height differential widens once in puberty, providing at least suggestive evidence that effects may impact final adulthood growth potential and indicating a persistent multigenerational gap that is potentially difficult to remediate. Alternatively, if the gap narrows in puberty, it could suggest that the multigenerational effect of drought exposure may impact only on second generation childhood growth velocity, but not necessarily final adult height potential, indicating a greater opportunity for catch up or remediation with post-exposure intervention.

To assess potential heterogeneity in the transmission of effects by adolescent or childhood growth stage, I use information on the timing of physical signs of pubertal onset to construct an indicator for likely pubertal growth. beginning in around 4 (age 12) respondents are asked about when they first noticed certain physical traits associated with puberty. For girls, respondents are asked if they experience menstruation, and at what age did they experience their first period. Similarly for boys, respondents were

asked if they have noticed hair on their chin, and at what age they first noticed hair growing on their chin. In round 5 (age 15), girls who had not yet experienced their period and boys who had not reported/had visible facial hair were asked these questions again.[20] Pooling information between rounds 4 and 5, I construct an indicators of if a respondent has reported first experiencing their period or having noticed hair on their chin by age 12 or 15. For a small number of observations, this information is missing across both rounds, due to refusal or non-response.[21] A limitation of this data is that it is self reported, however previous research using the Young Lives dataset has shown their reliability as an indicator, as the most important determinant of growth velocity in adolescence (Duc & Tam, 2015).[22]

Table 2.9 and Figure 2.7 show the estimated average marginal effects of grandmother's shock exposure on HAZ and BMIAZ of grandchildren, interacted with the likely pubertal growth indicator for ages 12 and 15, respectively. Interestingly, the marginal effects estimated for HAZ at ages 12 and 15 for those reporting signs of puberty (-0.244 S.D. and -0.099 S.D.) are larger than those found in the main specification, significant at the 1% and 10% level respectively. While negative effects are estimated for those not yet reporting signs of puberty, these effects are not statistically different from zero. This indicates that effects estimated in adolescence are primarily driven by those who have reported signs of puberty, suggesting that the gap between those exposed and unexposed may widen as respondents begin pubertal growth. In contrast, there is no significant effect estimated for BMI-for-age, regardless of growth stage, consistent with my primary findings. Compared to a large effect estimated for HAZ at age 12, the effect at age 15 is diminished. This could suggest that the effect, while initially wider in early puberty, narrows as children reach young adulthood, however the effect is imprecisely estimated, with overlapping confidence intervals, therefore it is not possible to draw clear conclusions in this aspect.

These effects are in contrast with those found by Bevis and Villa (2022), who provide suggestive evidence that the transmitted effect of early life weather induced height variation between first and second generation peaks at the average age of puberty onset for boys (age 11) and girls (age 8), however they do not find a statistical difference in transmission effects between those likely in pubertal growth and child growth at the next round (age 15 and age 11 for boys and girls, respectively).

[20] Notably, this question was incorrectly coded in round 5 for the Peru survey, but was subsequently collected either by in-person follow up or via an additional phone survey. This data is not currently available in the public release and must be requested directly from Young Lives.

[21] A concern is that the age of pubertal onset may be endogenous to shock exposure. I do not observe the eventual age of onset for respondents who do not report signs, therefore I cannot directly test this relationship, however shock exposure is not predictive of onset of puberty by age 12.

[22] A further limitation is that these measures, in particular for boys are not directly relatable to well established measures of pubertal growth, such as the Tanner stages/Sexual Maturity Rating scale (Marshall & Tanner, 1969, 1970).

Relevant regression coefficients are reported in Table B.4, with large positive coefficients estimated for the level term for puberty for both HAZ and BMIAZ as the outcome, providing suggestive evidence that indicator does seem to explain growth associated with actual puberty. However, given this indicator is constructed using self-reported data, it is likely there is significant measurement error arising from recall error and misreporting. Therefore results are interpreted with caution.

Table 2.9: Marginal Effect of Shock Exposure on Second Generation Outcomes: By Growth Stage

	Height-for-age		BMI-for-age	
	Age 12	Age 15	Age 12	Age 15
pubertal growth = 0 × In-utero shock	-0.102	-0.029	0.048	-0.128
	(0.075)	(0.118)	(0.052)	(0.118)
pubertal growth = 1 × In-utero shock	-0.244	-0.099	0.047	0.018
	(0.081)***	(0.053)*	(0.091)	(0.072)
N	1665	1617	1665	1617

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Table reports the marginal effects of shock exposure on second generation adolescent HAZ and WAZ/BMIAZ. Relevant regression coefficients are reported in Table B.4. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

2.5.3.4 Trimester of Exposure

Finally, I assess if exposure to drought within a specific trimester of the grandmother's pregnancy is important for grandchild effects. Estimates for the second generation effects of exposure to a SPEI shock ≤ -1 S.D. in each of the approximate trimesters of the grandmother's pregnancy are shown in Table 2.10. Results indicate that both the effects on HAZ and early years WAZ are strongest for exposure to a shock in the first trimester. A significant effect on second generation HAZ of 1st trimester exposure is estimated between ages 5 and 12, although the effect found at age 15 for the main analysis is no longer significant at conventional levels. These results are consistent with those in other multigenerational studies, which find exposure to negative shocks earlier in the pregnancy during the first and second trimester has the largest effect on second generation outcomes (Khan, 2021; Stein & Lumey, 2000). A limitation however, given that birth date is only available at the month and year level, is that it is not possible to precisely define if a respondent was exposed to a shock in a specific trimester, or if the defined intervals include periods prior to conception or post-birth, therefore these results may be subject to some measurement error.[23]

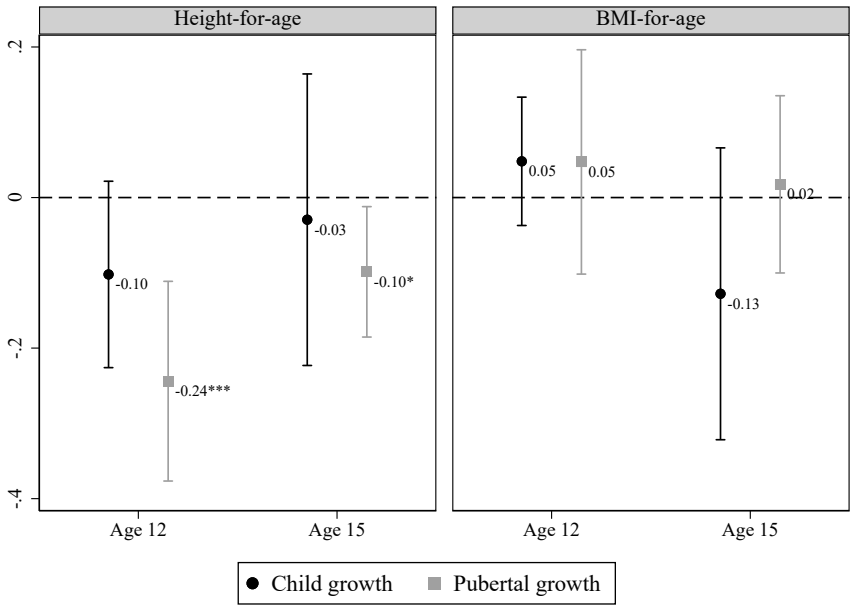
[23]Additionally, SPEI shocks for trimesters are defined over a different time period than the main specification shock, therefore they may not represent the same intensity of drought as a shock defined

Table 2.10: Effect of Shock Exposure on Second Generation Outcomes: By Trimester

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
1st trimester	-0.048	-0.151	-0.103	-0.125	-0.077
	(0.065)	(0.056)***	(0.046)**	(0.054)**	(0.047)
2nd trimester	-0.027	-0.013	-0.016	-0.008	0.019
	(0.067)	(0.051)	(0.052)	(0.057)	(0.048)
3rd trimester	-0.062	-0.023	-0.001	-0.069	-0.042
	(0.063)	(0.053)	(0.053)	(0.060)	(0.058)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
Observations	1670	1657	1665	1671	1620
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
1st trimester	-0.145	-0.124	-0.103	-0.056	-0.045
	(0.071)**	(0.060)**	(0.054)*	(0.052)	(0.051)
2nd trimester	-0.070	-0.050	-0.013	-0.013	-0.034
	(0.063)	(0.058)	(0.071)	(0.053)	(0.074)
3rd trimester	-0.068	-0.026	0.009	0.006	0.004
	(0.059)	(0.057)	(0.064)	(0.057)	(0.058)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
Observations	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Figure 2.7: Marginal Effect of Shock Exposure on Second Generation Outcomes: HAZ & BMIAZ, by Growth Stage



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Marginal effects reported in Table 2.9. 90% Confidence intervals. P-values calculated using cluster robust standard errors. Controls include an indicator for if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

2.5.4 Robustness Checks

To assess the sensitivity of results to shock definition, in Table B.5 and Table B.6 I re-estimate results for height and weight outcomes respectively using alternative exposure indicators. Panel A uses a lower cutoff point of less than or equal -0.8 S.D. for 12-month SPEI values proceeding the month of mother's birth. Panel B defines a drought shock as an average monthly SPEI value $\leq$ -1 S.D. throughout the growing season prior to the mother's month of birth of the primary crop in each department. The primary crop for each department of Peru is defined based on the annual sown area in hectares, using data from the Peruvian Ministry of Agriculture (MINAGRI).[24] The crop-growing season for each primary crop is defined based on a global gridded crop calendar data from the University of Wisconsin-Madison's Nelson Institute (Sacks et al., 2010), which provides estimates of planting and harvesting days for 19 crops on a 0.5x0.5° global grid, based on national and sub-national agricultural censuses.[25] I aggregate this grid-level data to the department level to obtain the mean planting and harvesting date, rounded to the

over a longer period, which indicate a more sustained period of below average rainfall.

[24]This data is extracted from the *cropdatape* R package, available: https://github.com/omarbenites/cropdatape. Crops included are rice, quinoa, potato, sweet potato, tomato and wheat. Tomato was excluded as it is a perennial crop.

[25]Available from: https://sage.nelson.wisc.edu/data-and-models/datasets/crop-calendar-dataset/

month-of-year level. The pattern of results remains for both specifications.

The sample size varies slightly across rounds due to both attrition and interrupted observations, as discussed in subsection 2.3.1. Table B.7 re-estimates Equation 2.2 for anthropometric outcomes on a balanced sample with results remaining robust (small variations in sample size remain due to singleton observations, which are dropped during estimation).

This analysis focuses on in-utero shock exposure, however it is possible that the multigenerational effects of early life shocks are not exclusive to exposure in the prenatal period. To test this, in Table B.8 and Table B.9 I estimate the impact of shock exposure in each year from 3 years prior to birth until 5 years after birth on HAZ and WAZ/BMIAZ, respectively. Shock exposure in any 12 month period outwith the 12 months prior to birth (the in-utero shock) has no significant impact on HAZ at any age. Interestingly for BMIAZ at age 15, a positive effect is estimated for shock exposure two or three years before birth, as well as 5 years after birth, although only the effect at 2 years before birth is significant at the 5% level or above. Otherwise There is no significant impact for any other outcome outside of the year prior to birth.[26]

While there is no evidence of an effect on the socioeconomic outcomes of the mother, a potential threat is if there is non-random selection in to treatment, with poorer families being more likely to be exposed. To check if there is an endogenous relationship between socioeconomic status of the grandparents' household and shock exposure, I regress in-utero shock exposure on the education level of the grandparents, measured by an indicator of at least one grandparent completing secondary education, finding no evidence of grandparent educational attainment as being predictive of shock exposure. However, data on grandparent education is only available for a sub-sample of households where at least one grandparent is present in the household roster at any point between round 1 and 5, likely leading to a selection bias. Therefore I use as an alternative measure an indicator of whether the maternal grandmother's mother tongue was Spanish, which is available for the majority of the sample. This is a rough proxy of long-term SES, with large inequalities present in education and SES between Spanish and indigenous language speaking households (Leon et al., 2021). I find no significant relationship with shock exposure. Results are reported in Table B.12.

Alternatively a further threat to identification is if shock exposure leads to selection into the sample, for example by influencing the subsequent migration choices of the mother's family after her birth. To assess this, I use an indicator of if the mother moved prior to age 5, estimating an insignificant null effect of exposure to shock on migration choices, as reported in column 1 of Table B.13. Shock exposure is also not predictive of ever-migrating (prior to birth of the YL child), migration to a departmental capital, or to

[26] Additionally, I estimate all these periods separately, with similar results reported in Table B.10 and Table B.11.

Lima/Callao, as shown in columns 2-4. Finally, shock exposure may impact the health of mothers born if shock exposure has an effect on infant mortality, leading to survivorship bias amongst exposed mothers. Alternatively, if some grandmothers can react to shock exposure by choosing to delay having children, then shock exposure may also affect the composition of cohorts.[27] To attempt to address the latter, in column 5 of Table B.13 I regress shock exposure on the month of birth (January to December) of the mother, finding no significant relationship.[28] For the former, I am unable to assess the impact of shock exposure on mortality within my dataset. However, if selective mortality occurred it would be expected that the surviving mothers are on average healthier, therefore my estimates of the negative effect of shock exposure on maternal height would likely represent a lower bound of the effect of shock exposure. In the next section I explore the potential transmission channels for multigenerational effects using mediation analysis.

2.6 Mechanisms

2.6.1 Mediation Analysis

To explore the potential mechanism channels I conduct a mediation analysis, estimating the average controlled direct effect (ACDE)(Joffe & Greene, 2009; VanderWeele, 2009), which is the effect of changing treatment status with the mediator held at a fixed value for all units. The ACDE therefore provides an estimate of the direct effect of treatment that does not operate through the specified mediator (Acharya et al., 2016). If the effect of treatment is completely mediated by some variable M and a set of other mediator variables W, then a non-zero ACDE for mediator M implies that the effect of treatment does not exclusively operate through that channel M, allowing alternative mechanisms to be ruled out (VanderWeele, 2011). Additionally, if the null hypothesis that the ACDE is not different from zero cannot be rejected at conventional levels, then that mediator M is likely the main mechanism through which the treatment causes the outcome (Bellemare et al., 2021), provided identifying assumptions hold. See Appendix B.1 for greater detail.

To measure socioeconomic status of the household, therefore capturing the home environment and ability of parents to invest in children (Khan, 2021), I use the Young Lives wealth index (Briones, 2017). To measure mother's human capital, I use mother's educational attainment (highest grade/level achieved), reflecting her cognitive ability, parenting skills, and earnings potential, which are important determinants of child health and human capital (Van Den Berg & Pinger, 2016). Finally, to capture maternal health I use mother's adult height (cm). As discussed above, adult height is a measure of

[27] Notably, in the 1986 DHS survey 65.5% of married women aged 15-49 reported having ever used contraception, although 86.7% of those only reported using traditional methods, e.g. withdrawal and rhythm (Goldman et al., 1989).

[28] the distribution of month of birth of the mother is shown in Figure B.3.

cumulative health and a good proxy of morbidity risk (Case & Paxson, 2008).

If multigenerational impacts are transmitted predominantly through the environmental channel then it is expected that the ACDE for measures of SES and parent human capital will be close to zero (Acharya et al., 2016). Alternatively, if effects are transmitted predominantly through the biological channel then it is expected that the ACDE for environmental mediators are non-zero and do not differ significantly from the baseline estimate, while the ACDE for maternal health will be close to zero.

Under strong assumptions, the use of standard regression analysis using a single equation with the mediator as an additional regressor could be a valid way of testing a mechanism only if there is no omitted variables for the effect of treatment on both mediator and outcome, nor for the effect of mediator on outcome, and importantly, only if all relevant confounders are pre-treatment (Bellemare et al., 2021). However, this is insufficient if there exists some post-treatment covariate Z, which is influenced by treatment D, influences the mediator M, and is independently associated with the outcome Y (Robins, 1986). The exclusion of these "intermediate" confounders if they exist could induce a spurious relationships between treatment and the outcome when including mediators in the regression equation (Rosenbaum, 1984), while conversely including them as regressors could introduce intermediate variable bias to the estimate of the direct effect (Acharya et al., 2016).

An important advantage therefore of the ACDE for mediators is that it can be identified in the the face of these intermediate confounders when estimated using "sequential g-estimation" (or reverse sequential two-stage (RS2S) parametric estimation), as set out by VanderWeele (2009) and Joffe and Greene (2009). Full details of identifying assumptions and implementation are provided in Appendix B.1. I identify a number of socioeconomic controls and measures of parental investments in the child which are likely related to the mediator and outcome, and could potentially be caused by the treatment. These vary across rounds due to relevance for that stage of development, and are listed in Table B.14. Furthermore, I condition on further "pretreatment" variables, that is variables which can affect the treatment, outcome or even the mediator, but are not determined by the treatment and do not come between the mediator and outcome. For this, I include an indicator for if the child is female, as well as all the fixed effects included in the baseline specification.[29]

Finally, Acharya et al. (2016) note that the demediation function generally identifies the ACDE when the mediator is set to zero, which may be nonsensical in context, and that in these cases it is suitable to recentre the mediator around a specific value. For my analysis, a maternal height of 0 cm is not plausible, therefore I recentre the height at the sample mean, creating a normalised index of mean 0 and standard deviation 1,

[29]Within my sample grandmother's shock exposure during pregnancy does not predict the sex of her grandchild, as shown in column 6 of Table B.12.

however results are robust to use of the raw value in centimetres.

2.6.2 Mediation Results

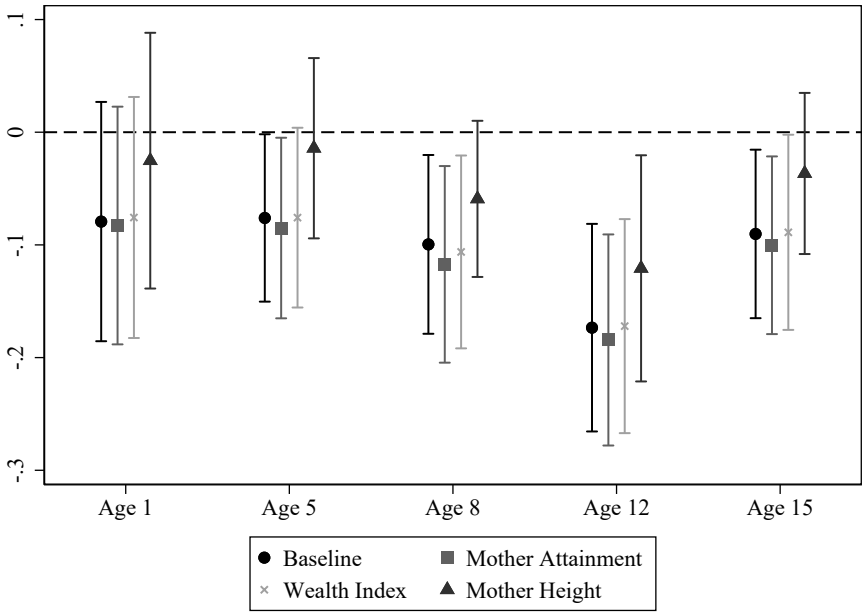
Table 2.11 and Figure 2.8 provide comparisons of baseline results for HAZ with the ACDE estimated separately after accounting for three different potential mediators: i) mother's attainment (highest grade/qualification achieved) and ii) household wealth index, measures representing the hypothesised environmental channel, through which the shock impacts the household environment, parenting ability, or resource constraint for investments in children; and iii) mother's height, representing the biological channel, wherein exposure to an in-utero shock has permanent effects on maternal physiology, metabolism and ability to transfer nutrients to her offspring in-utero.

Table 2.11: Mediation Analysis: Comparison of Baseline Results with ACDE, HAZ

		ACDE		
	Baseline Result	Mother Attainment	HH Wealth Index	Mother Height
Panel A: Age 1				
In-utero shock	-0.079	-0.083	-0.076	-0.025
	(0.064)	(0.064)	(0.065)	(0.069)
Observations	1670	1670	1670	1655
Panel B: Age 5				
In-utero shock	-0.076	-0.085	-0.076	-0.014
	(0.045)*	(0.049)*	(0.048)	(0.049)
Observations	1657	1657	1657	1649
Panel C: Age 8				
In-utero shock	-0.100	-0.117	-0.106	-0.059
	(0.048)**	(0.053)**	(0.052)**	(0.042)
Observations	1665	1665	1665	1650
Panel D: Age 12				
In-utero shock	-0.173	-0.184	-0.172	-0.121
	(0.055)***	(0.057)***	(0.058)***	(0.061)**
Observations	1671	1671	1671	1655
Panel E: Age 15				
In-utero shock	-0.090	-0.100	-0.089	-0.037
	(0.045)**	(0.048)**	(0.053)*	(0.043)
Observations	1620	1620	1620	1609

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Pretreatment controls include an indicator of if the child is female. Intermediate confounders for each age are listed in Table B.14. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Figure 2.8: Mediation Analysis: Comparison of Baseline Results with ACDE, HAZ



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Coefficients reported in Table 2.11. 90% Confidence intervals. P-values calculated using cluster robust standard errors. Pretreatment controls include an indicator for if the child is female. Intermediate confounders for each age are listed in Table B.14. Fixed effects include child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth. Controls and fixed effects are suppressed.

Estimates of the ACDE net of the indirect effect of maternal attainment and household wealth are of similar magnitude as the baseline estimated effect, and remain significantly different from zero, with exception of the ACDE for age 5 HAZ net of household wealth, for which the same effect size is only marginally significant under the baseline specification. This suggests that a significant effect of shock exposure that is either transmitted directly or through an alternative set of mediators remains, and that these environmental channel measures do not seem to play a role in the causal pathway between shock exposure and grandchildren outcomes.

In contrast, all estimates of the ACDE net of the indirect effect of mother's health stock, measured by her adult height, are considerably diminished. The majority of estimates are also not significantly different from zero, with exception of the age 12 coefficient, which remains significantly different from zero above the 5% level. A similar pattern emerges for health flow, as shown in Table 2.12 and Figure 2.9, where estimates for the ACDE net of environmental mediators do not differ from the baseline results. The direct effect with mother's height as the mediator is diminished and is not significantly different from zero at age 5, however while diminished for age 1, the ACDE remains non-zero (significant at the 5% level), suggesting that mother's health is not the sole mediator for the effect on infant WAZ.

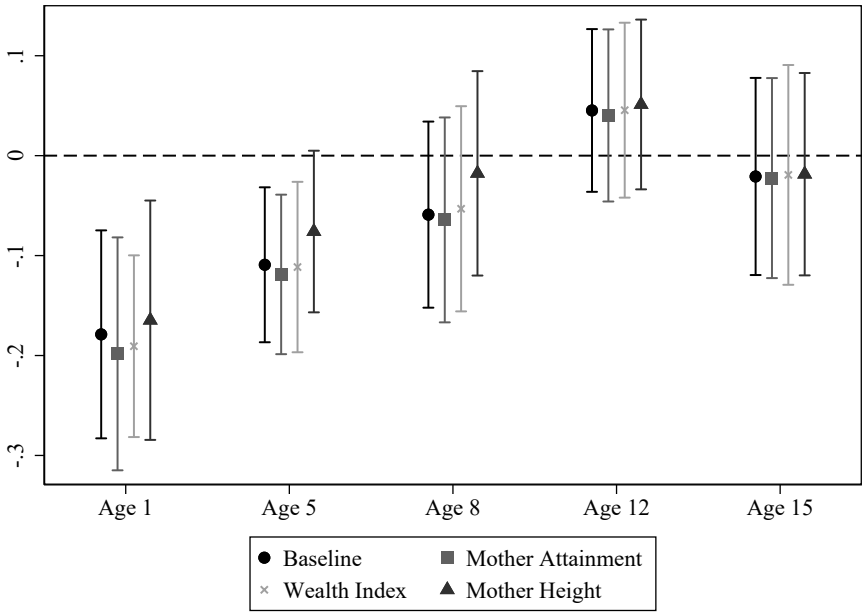
While the evidence above suggests that maternal health, as measured by height, is

Table 2.12: Mediation Analysis: Comparison of Baseline Results with ACDE, WAZ/ BMIAZ

		ACDE		
	Baseline Result	Mother Attainment	HH Wealth Index	Mother Height
Panel A: Age 1				
In-utero shock	-0.179	-0.198	-0.191	-0.165
	(0.063)***	(0.071)***	(0.055)***	(0.073)**
Observations	1670	1670	1670	1655
Panel B: Age 5				
In-utero shock	-0.109	-0.119	-0.111	-0.076
	(0.047)**	(0.049)**	(0.052)**	(0.049)
Observations	1657	1657	1657	1649
Panel C: Age 8				
In-utero shock	-0.059	-0.064	-0.053	-0.018
	(0.056)	(0.062)	(0.062)	(0.062)
Observations	1665	1665	1665	1650
Panel D: Age 12				
In-utero shock	0.045	0.040	0.046	0.051
	(0.049)	(0.052)	(0.053)	(0.052)
Observations	1671	1671	1671	1655
Panel E: Age 15				
In-utero shock	-0.021	-0.022	-0.019	-0.019
	(0.059)	(0.061)	(0.067)	(0.062)
Observations	1620	1620	1620	1609

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Predetermined controls include an indicator of if the child is female. Intermediate confounders for each age are listed in Table B.14. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Figure 2.9: Mediation Analysis: Comparison of Baseline Results with ACDE, WAZ/ BMIAZ



Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Coefficients reported in Table 2.12. 90% Confidence intervals. P-values calculated using cluster robust standard errors. Pretreatment controls include an indicator for if the child is female. Intermediate confounders for each age are listed in Table B.14. Fixed effects include child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth. Controls and fixed effects are suppressed.

not the sole mediator for the effect, several issues which may impact the estimated ACDE must be considered. First, direct effects are likely biased if the mediator is measured with error, or the measure of a mediator does not capture fully all the effect of the treatment on all dimensions of the mediating variable (VanderWeele, 2012). To illustrate, consider an example presented by Huber (2019), where there exists a measure of a mediator that only captures the extensive margin (e.g. a dummy indicator of employment) but not the intensive margin (actual hours worked). If the effect of treatment on the mediator induces change in both the intensive and extensive margins, the estimate of the indirect effect will only account for the proportion of the actual indirect effect related to treatment-induced changes at the extensive margin, while the treatment-induced changes to the mediator at the intensive margin will be wrongly attributed to the direct effect (that is, as not operating through that mediator). Given adult height represents a relatively imperfect measure of maternal health, which may not capture all the dimensions of health which are affected by in-utero shock exposure (alterations to maternal physiology which impact maternal prenatal health but not stature), it is likely that the indirect effect transmitted through the biological channel is underestimated, biasing the ACDE estimate upwards from the true population value.

Second, there could exist an alternative biological transmission channel which operates not through impacting maternal health of the first generation, but is transmitted directly to the second generation if shock exposure in-utero impacts the germ cells (the gametes/reproductive cells), present within the first generation as a foetus while in-utero, and from which the second generation will be formed (so-called "gametic epigenetic effects" (Youngson & Whitelaw, 2008)). However, while this potential channel has been discussed in epidemiology, the evidence from animal studies is limited and how exactly this potential channel operates is not well understood. Therefore it is outside the scope of this study to attempt to address it directly here. For a discussion of the theory and some evidence from studies of rats, see Drake and Liu (2010).

Overall the results suggest that maternal health seems to play a considerable role in explaining the mechanism of transmission, providing support to the biological channel as the primary mechanisms for the multigenerational effects, however I cannot conclusively rule out the potential role of other mediating variables.

2.7 Conclusions

In this chapter, I contribute to the growing "second generation" literature on early childhood shocks. Using high-quality data from the Young Lives Peru study I estimate the multigenerational effects of prenatal shock exposure on the outcomes of the first and second generations. To identify exposure to exogenous variation in drought experienced by a grandmother while pregnant, I link gridded time series data from SPEIbase (Begueria

et al., 2023) to the date and location of birth of the mother of the Young Lives child. Using longitudinal data tracking a cohort of children and their family from birth into adulthood allows for a detailed assessment of how effects in the second generation manifest and vary as the child grows into adulthood, rather than providing just a snapshot at one age.

This chapter provides three contributions to the literature. First, I present evidence that the exposure of the grandmother to drought while pregnant has a negative multigenerational effect on the long-term health stock of both her child and her grandchildren, with the first generation being shorter on average in adulthood, and with significant impacts on the height and weight of the second generation. Exposure to drought in the first trimester of pregnancy in particular is associated with a negative impact on the growth outcomes of her grandchildren. Furthermore, evidence suggests that effects are isolated to those grandmothers located in rural areas, where households may have been more reliant on local food supply and agricultural income.

Second, exploring the dynamics of this multigenerational effect in the second generation shows this impact becomes clear early in life, with a negative impact on infant and early childhood weight-for-age. However while impacts on early life health flow diminish, a persistent effect on height-for-age is evident from early/mid-childhood and remains significant into adolescence, suggesting a permanent effect on long-term health stock. Using self-reported data on signs of pubertal growth, I find evidence that this height differential remains significant and appears to widen for those reporting pubertal growth at either age 12 or 15. Additionally, I find evidence of sex-specific effects, with a disproportionate impact of grandmother's exposure on the height-for-age of boys, consistent with other work in this literature (Fung & Ha, 2010; Venkataramani, 2011). That the effects for grandchildren are primarily seen in boys makes it unclear whether this multigenerational effect would persist beyond the second generation, a limitation which cannot be addressed within this current study. Furthermore, current data limitations means it is not yet possible with this dataset to fully assess if effects persist into adulthood, affecting final adult height potential. In contrast to significant impacts on physical growth, there is little evidence of an effect on measures of cognitive ability. Results remain robust to alternative shock definitions, and I provide evidence that the multigenerational effect of shock exposure is limited to in-utero exposure of the first generation only.

Third, results from mediation analysis suggest the primary channel which facilitates this transmission of health effects across generations is biological, with the majority of the baseline effect operating indirectly through an impact on the maternal health of the first generation, while there is little to no indirect effect operating through the environmental pathway, capturing maternal human capital accumulation and socioeconomic status. However for some outcomes a non-zero direct effect remains after accounting for maternal health as well as a range of potential intermediate confounders, suggesting that

the causal effect on the first generation is not fully captured by my measure of mothers' health, or alternatively that there potentially exists another unobserved transmission channel, for example through "gametic epigenetic effects".

These findings provide some policy implications. First, in-utero exposure to a negative drought shock has a lasting impact across generations. As such, if policy is designed without accounting for the potential of multigenerational consequences, it will likely underestimate i) the full cost of exposure of individuals to a shock; and ii) the true cost-benefit ratio of any policy aimed at mitigation (Doyle & Jernström, 2023). Second, that effects remain persistent in to adulthood for the second generation and are not mediated by the human capital accumulation or socioeconomic status of the first generation also suggests that effects are not easily remediated after exposure. Therefore an emphasis should be placed on the importance of early intervention for the timing and targeting of future policy.

Chapter 3

Some People Feel the Rain; Others Just Get Wet: Early-life Shocks and Personality Trait Formation in Peru

3.1 Introduction

The importance of early life experiences in determining a range of later life outcomes is well established, from anthropometric and health indicators to employment, cognitive ability and educational attainment (see Almond & Currie, 2011; Almond et al., 2018; Currie & Vogl, 2013, for reviews). A developed literature demonstrates that not only genetics, but also early life environment and shocks play an important role in determining human capital formation (Almond & Currie, 2011). A growing strand of this evidence addresses how exposure to weather shocks during the perinatal phase can have effects that may be felt for many years, with early contributions from Pathania (2007) and Maccini and Yang (2009). However, while the relationships between early life rainfall and the health and cognitive dimensions of human capital are well explored (Carrillo, 2020; Nübler et al., 2021; Rosales-Rueda, 2018; Shah & Steinberg, 2017; Thai & Falaris, 2014; Zimmermann, 2020), less studied is the relationship between early life shocks and the evolution of later-life personality traits (often referred to as non-cognitive, socio-emotional, or 'soft' skills). Understanding this relationship is important given the evidence of the potential key role they play an in determining future labour market outcomes and socioeconomic success (Almlund et al., 2011; Caliendo et al., 2015; Fletcher, 2013; Heckman & Kautz, 2012; Hilger et al., 2022; Mueller & Plug, 2006; Nordman et al., 2019), and the theoretical dynamic complementarity of early shocks and investments on later period outcomes (Cunha & Heckman, 2008; Cunha et al., 2010; Heckman, 2007).

This chapter contributes to this literature by assessing the impact of exposure to

rainfall shocks around the time of birth on personality trait formation in adolescence and adulthood. Using the Peruvian sample of Young Lives, a cohort study of childhood poverty following respondents from infancy to adulthood, I estimate the impact of early life exposure to rainfall shocks on measures of respondents' appraisal of their own self-worth, competence, and capabilities, as measured by their core self-evaluation (CSE) (Chang et al., 2012; Judge et al., 1998). I construct a measure of rainfall shocks as unexpected deviations from the long-term community-specific mean, using monthly rainfall data from The University of Delaware's Terrestrial Precipitation Gridded Time Series (UDEL-TS)(Matsuura & Willmott, 2018), matching this with detailed data on the respondents' location and date of birth to identify exposure to rainfall shocks in early life.

I find a 1-month exposure to a positive rainfall shock (total monthly rainfall $\geq +1.5$ S.D. from the long-term mean) in the prenatal period is associated with a 0.068 S.D. higher standardised CSE score in adolescence and adulthood, while exposure to the same type of shock in the 2^{nd} and 3^{rd} year of life is associated with a 0.090 S.D. and 0.105 S.D. lower score, respectively. There is no significant effect for 1st year exposure, and I find no effect of exposure to a negative shock ($\leq$ -1.5 S.D.) in any period. Considering heterogeneities, I find that prenatal exposure to a positive rainfall shock has a positive impact on female adolescent and adulthood CSE, compared with a null effect for males, with the largest effects for those in the poorest households.

These results contribute to the literature which identifies the importance of early life circumstances in determining future human capital, expanding the limited evidence base for the effects on personality trait and socio-emotional skill formation (Brando & Santos, 2015; Leight et al., 2015; Moorthy, 2021; Shoji, 2023; Webb, 2024). In this literature, the methodology varies significantly in terms of shock type and exposure period considered; age at follow up; and outcome measures used. Most closely related to this study are two recent contributions by Chang et al. (2022) and Krutikova and Lilleør (2015). Krutikova and Lilleør (2015) find that exposure in-utero to a 10% increase in total rain season rainfall compared to the 10-year average is associated with a 0.08 S.D. increase in CSE scores at age 17-28 in a sample of rural Tanzanian households. In contrast Chang et al. (2022), find a -0.161 S.D. decrease in CSE at age 15 associated with prenatal exposure to a 1-month rainfall shock in India.

I expand on these studies in three ways. First, I provide an extensive assessment of potential mechanisms. I find that a positive rainfall shock exposure is positively associated with current period household and parental labour supply, suggesting parents are less available in the household at a key period of socio-emotional development. This impacts the time parents can spend interacting with their child in the early years, affecting the socio-emotional bond developed through parent-child interaction, consistent with the experimental literature exploring the role of early life psycho-social stimulation on later-

life personality trait formation (Attanasio et al., 2020; Heckman et al., 2013; Sevim et al., 2023; Walker et al., 2022). All household adults work more in response, while non-adult household members do not alter their time use, suggesting this reduction in interaction is not offset by others.

Second, I consider not just the prenatal period but also exposure during a sensitive period after-birth, which has seen an increasing focus in the recent literature (Almond et al., 2018), and through which I can address the experimental literature discussed above. Third, I offer a detailed exploration of the robustness of my estimation strategy, including how the construction of both outcome and treatment variables may lead to significant measurement error and attenuation of estimates, including a robust assessment of the suitability of a single latent factor model, testing alternative shock variable construction, and accounting for several potential sources of bias (Anderson, 2008; Cameron et al., 2008; Conley, 1999; Dell et al., 2014) which are often unaddressed in similar studies using climate data.

The rest of the chapter is as follows: the study setting and data are described in section 3.2 and section 3.3. The empirical strategy is outlined in section 3.4. The main results, analysis of heterogeneous effects and robustness checks are presented in section 3.5, with the potential mechanisms underlying these results explored in section 3.6. Finally, concluding remarks are provided in section 3.7.

3.2 Context

Peru experiences a complex climate with significant variation in rainfall across its geographically diverse regions, from the warm and wet tropical Amazonian jungle and lowlands in the east to the semi-arid Pacific coast in the west, both separated by the drought- and frost-prone Andean highlands which run from north to south. Since the 1960s, rainfall patterns in the region have changed drastically (Haylock et al., 2006), with an increase in the frequency and intensity of precipitation-related extreme weather events, such as rainstorms, floods, mudslides and forest fires (Gloor et al., 2013; USAID, 2011).[1] Within a wider regional context, Peru is located in a climate-sensitive Andean South American region, prone to quasi-periodic extreme precipitation and temperature anomalies associated with the El Niño-Southern Oscillation (ENSO)(Ramírez & Briones, 2017). As a middle-income country with a high degree of inequality, individuals are often less able to shield from the effects of such anomalies than in high-income contexts, particularly those in the poorest households.

There have been several studies within the wider northern South America region

[1] https://climateknowledgeportal.worldbank.org/country/peru/climate-data-historical: The number of intense rainstorms, mudflows and forest fires has more than doubled in the past 10 years and floods have increased by 60% since the 1970s.

which assess the impacts of early life exposure to rainfall shocks on educational attainment, health, and cognitive ability (Brando & Santos, 2015; Carrillo, 2020; Duque et al., 2019; Rosales-Rueda, 2018). Within Peru specifically, Danysh et al. (2014) report an increasing trend over time in height-for-age of 0.09 S.D./year for cohorts born between 1991-1997 in Tumbes, a region on the far north coast that is particularly prone to the effects of the El Niño. They find this rate is reduced to 0.04 S.D./year for cohorts born during or after the 1997-1998 El Niño event, with the subset of children in the most flood prone households subject to negative growth rates. Considering specifically the impact of rainfall shocks on socio-emotional development, Brando and Santos (2015) assess the effect of exposure to the 2010-2011 La Niña in Colombia, finding exposure to high rainfall is associated with an increased incidence of socio-emotional problems by age 5 (0.19 S.D.). However to the best of my knowledge, this study is the first in the region to assess the longer term impacts of early life shocks on socio-emotional skill formation in adolescence and adulthood.

3.3 Data

3.3.1 Young Lives

Young Lives (YL) is a longitudinal study of 12,000 children and their families across four developing countries (Ethiopia, India, Peru, and Vietnam) examining the causes and consequences of poverty (Boyden et al., 2018). The younger cohort of 2052 respondents were born in 2000-2002 and were tracked at age 6-18 months beginning in 2002, being revisited in 2006, 2009, 2013, and 2016 at ages 5, 8, 12, and 15 respectively. An older cohort (714 respondents), born in 1994-1996, were interviewed concurrently at ages 8, 12, 15, 19 and 22. This analysis focuses on the Peruvian sample, including both cohorts in the sample.

In Peru, the study employs a multi-stage, cluster-stratified, random sampling technique. Although a deliberate choice is made to oversample poor households by excluding the top 5% wealthiest districts prior to randomisation, a comparison with nationally representative surveys shows that households were broadly similar to the average household, although with slightly better access to health and education services, indicating the sample is generally suitable for analysing causal relations and modelling child welfare (Escobal & Flores, 2008).

At round 1 (2002) the total sample consists of 2766 children. Attrition is low given extensive tracking: by round 5 (2016) attrition due to respondent refusal, death, or being untraceable was 8.2% and 14.1% respectively, with 2468 respondents present in all rounds. Beginning in round 2, GPS coordinates are collected for the centre of a

community with 3 or more respondents.[2] This GPS dataset was cleaned, validated, and matched to climate data, allowing for the identification of potential exposure to rainfall shocks using respondent's date of birth for 2386 of the respondents in 118 communities (McQuade & Favara, 2024). Accounting for missing responses for outcomes and control variables, a final sample of 2089 respondents is derived.

Children were first tracked just after birth, therefore an issue in attributing exposure to rainfall shocks is pinpointing if the mother resided in the relevant community from the date of conception throughout the period considered. To address this, in addition to the full sample, I specify an 'in-community' sub-sample of those that can be identified as definitely conceived in the community (N=1675, 80.2% of the final sample). From round 2, mothers were asked how many years they have lived in the community. Subtracting this from the date of interview, I calculate the approximated date of community move-in for the mother. Specifying a gestational period of 40 weeks prior to their child's date of birth to determine the likely date of conception, mothers for whom this date occurs after the move-in date are considered to have conceived their child in the community.[3] However, this indicator is restrictive and problematic if by round 2, when the oldest respondents in sample are aged 12-13 years old, shock exposure had systematically impacted post-exposure migration choices – representing a confounding factor, with affected families self-selecting out of the sample. Therefore I continue to conduct the analysis using both samples.

3.3.2 Core Self-Evaluation

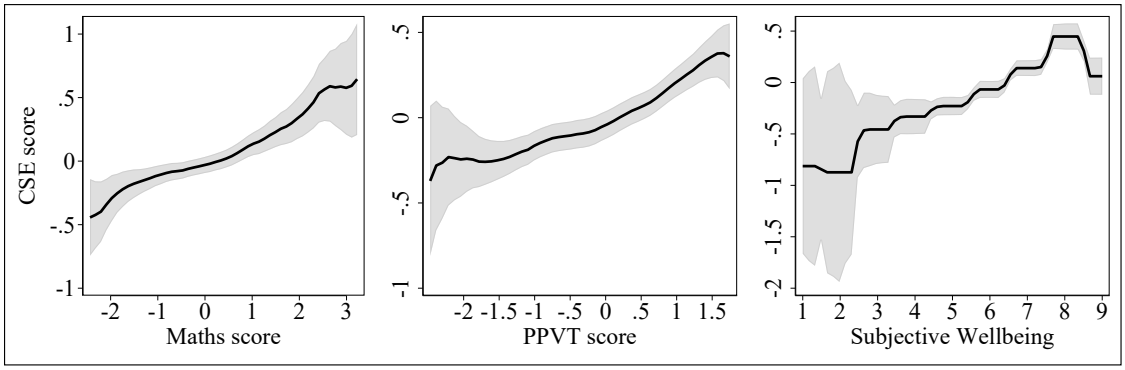
The outcome of interest is an individual's core self-evaluation (CSE) (Judge et al., 1998), measured in round 5 when respondents are aged between 14-23 years old. CSE has been shown to be strongly associated with life and job satisfaction, earnings, and educational attainment (Chang et al., 2012; Judge & Hurst, 2007). It reflects an individual's confidence in their own abilities and self-control, with a high score indicating a person has a positive and proactive view of themselves and their relationship to the world (Almlund et al., 2011). In the absence of a dedicated CSE questionnaire in the Young Lives study, an indirect approach (Chang et al., 2012) is used, drawing responses from the self-esteem, self-efficacy, and agency scales.[4] Self-esteem is derived from the Marsh (1990) self-description questionnaire II, and is a widely used measure in longitudinal studies (Laajaj

[2]For round 2, some larger round 1 communities are split into smaller communities with separate identifiers. For these cases the round 2 community is used as the location of birth, assuming no movement between disaggregated communities between rounds 1 and 2.

[3]While full term pregnancies commonly occur across a range of 37-42 weeks, given limited information about length of pregnancy in the YL survey (self-reported prematurity is only available for the younger cohort), a specific, relatively conservative cut-off is defined for simplicity.

[4]The big five inventory (BFI) neuroticism/emotional stability scale (Costa & McCrae, 2008; John, 1990) can also be included in the construction of CSE measures, however this scale is not administered amongst the younger cohort.

Figure 3.1: Within-Sample Association Between CSE and Cognitive and Wellbeing Outcomes



Polynomial line fit using Epanechnikov kernel, plotted with 95% CIs. CSE, PPVT and Maths scores are age-standardised based on age at R5. Subjective wellbeing is measured using a 9-point Cantril's ladder scale.

& Macours, 2021). Self-efficacy (an individual's belief in their ability to cope with adversity and succeed) is measured using a scale developed by Schwarzer and Jerusalem (1995), and is well validated in a range of low- and middle-income countries. The agency scale was developed specifically for Young Lives to be administered to children in developing countries, and is closely related to Rotter's 'locus of control' concept (Rotter, 1966). These measures have been validated and display high internal consistency and reliability in YL samples (Yorke & Ogando Portela, 2018). Within sample associations between CSE, age-standardised cognitive test scores and subjective wellbeing (Cantril, 1965) at outcome are shown in Figure 3.1.

I specify a latent factor model, similar to Cunha et al. (2010), using exploratory factor analysis, a method commonly used to assess the psychometric properties of scale items, as well as for dimension reduction (Osborne, 2015). An advantage of this method over a simple average composite score is that it accounts for the disproportionate contribution of each item to the CSE construct to be recognised, allowing the shared variance between items and each item's unique variance to be distinguished. Negatively phrased items were first reverse-coded to ensure unidirectionality.[5] All item scores were then standardised by age in years. In total 22 items are included in the initial model (see Table C.2), with latent factors estimated using principal factors. Results strongly support the *a priori* assumption of a one factor model, with the first factor explaining 95.3% of shared variance, and a range of criteria, presented in Appendix C.1, indicating that a single factor should be extracted. To assess the robustness of results to an alternative method

[5]Two items from the agency scale were negatively worded such that a higher score reflected lower self-agency (e.g. *"I have no choice about the work I do - I must do this sort of work"*). Additionally, one item from the agency scale, *"If I study hard at school, I will be rewarded by a better job in the future"* was missing for a third of respondents, the majority of whom are no longer in school, and therefore was excluded for non-relevancy and to preserve sample size.

of outcome construction, I follow Laajaj and Macours (2021), defining a 'naïve' factor score – the standardised average score of all items in the 3 constituent scales.

3.3.3 Controls

Child level controls include indicators for gender and if their mother tongue is Spanish. Maternal controls for age and an indicator of if she has completed primary education are also included. Household characteristics are captured by an indicator of if the household is in a rural area and the round 1 household wealth index, a country-specific measure of household socioeconomic status (Briones, 2017).[6] Fixed effects are included for the child's year-month birth cohort and district of birth. Summary statistics for the main sample are provided in Table 3.1. The wealth index is a continuous measure with values ranging between 0 and 1 for the poorest and wealthiest household respectively, as such the sample is slightly pro-poor. Additionally, while only 29% report the household being rural, 61% of the sample live in communities where the most important economic activity is agriculture, as reported in the community questionnaire, and 56% of households reported being actively engaged in agricultural work in round 1.

Table 3.1: Summary Statistics

	Mean	S.D.	Min	Max
Child characteristics				
EFA 1st factor CSE score	-0.00	(1.00)	-4.70	3.14
Naïve CSE score	0.00	(1.00)	-3.90	3.48
Age in years at outcome	16.67	(3.03)	14.08	22.83
Female	0.49	(0.50)	0.00	1.00
Spanish first language	0.86	(0.35)	0.00	1.00
Mother characteristics				
Completed primary	0.63	(0.48)	0.00	1.00
Age at child birth	26.56	(6.80)	13.00	48.00
Household characteristics				
Wealth index	0.44	(0.24)	0.00	0.92
House location is rural	0.29	(0.46)	0.00	1.00
Agricultural community	0.61	(0.49)	0.00	1.00
Engaged in Agricultural work	0.56	(0.50)	0.00	1.00
N	2089			

Notes: Sample means are reported with standard deviations in parentheses.

[6]While the wealth index is potentially measured contemporaneously or post-shock exposure, it captures longer-term indicators of household wealth, such as housing quality, access to services, and durable goods, therefore is less likely to be sensitive to short run deviations in climate conditions.

3.3.4 Climate Data

I exploit spatial and temporal variation in precipitation to identify exposure to abnormal amounts of rainfall. Identification relies on short-run fluctuations from the long-run month-specific mean rainfall for a location being unpredictable and plausibly exogenous. While YL households provide self-reports of their experience of recent climate shocks, this data is likely endogenous and subject to significant measurement error (Bound et al., 2001), depending on respondent recall and their perceived impact of the shock, for which they may systematically over- or under-report exposure (Nguyen & Nguyen, 2020). Additionally, it is difficult to verify the timing and intensity of self-reported shocks.

Data on rainfall comes from the Terrestrial Precipitation Gridded Time Series (v5.01, 1901-2017) (Matsuura & Willmott, 2018) from the University of Delaware (UDEL-TS). This data provides estimates of monthly total precipitation on a 0.5x0.5° grid, derived from a range of publicly available station records. I match this data to the location of Young Lives communities to create community-level estimates of monthly total rainfall (McQuade & Favara, 2024).[7]

To identify exposure to an abnormal rainfall shock, I derive a standardised precipitation index (SPI) (McKee et al., 1993). The SPI is a widely used drought index which benefits from its simplicity in calculation, requiring only precipitation data. It is used to identify the duration and/or severity of a drought or high level of rainfall on a relative scale (Hayes et al., 1999). Rainfall is non-negative and typically positively skewed in the short run, therefore non-zero estimates of community rainfall across 1988-2017 were fitted to a two-parameter gamma distribution, to approximate the long-term distribution of rainfall for each month of the year at each community location. Resulting distributions are transformed to standard normal with mean 0 and standard deviation 1 (S.D.), following Abramowitz and Stegun (1968). Following the drought classifications defined by McKee et al. (1993), I consider a monthly SPI value of ≤ -1.5 S.D. from the long-term mean as an indicator of severe drought-like conditions, (herein, a "negative" rainfall shock) and similarly rainfall $\geq +1.5$ S.D. as a "positive" shock (corresponding to the "severely wet" category). For greater detail, see McQuade and Favara (2024) and Lloyd-Hughes and Saunders (2002).

Almost all children were exposed to at least one mild shock of 1 S.D. in some periods, while very few were exposed to any extreme shock >2 S.D., as such these cut off points are unsuitable for use. The distribution of shock exposure (of at least one month) within the perinatal period, estimated separately for the prenatal phase (9 months prior to birth) and each of the first three years of life (up to the month of the child's 3rd birthday) for the postnatal phase, across the full and restricted in-community sample are provided in

[7] Community estimates are calculated as the inverse distance-weighted average of the four nearest grid points to the community centre point, which is defined as the main square, or in their absence, an important landmark such as a church, school, or post office.

Table 3.2: Exposure to (±)1.5 S.D. Shocks, by Period and Sample

	% Exposed		Mean exposure	
	Full	In-comm.	Full	In-comm.
Positive shocks				
Prenatal	53.9	53.5	0.66	0.64
1st year	66.7	67.5	0.80	0.81
2nd year	61.9	61.1	0.95	0.94
3rd year	38.4	36.4	0.63	0.60
Negative shocks				
Prenatal	19.3	19.2	0.25	0.25
1st year	35.1	32.7	0.47	0.43
2nd year	28.4	28.5	0.31	0.31
3rd year	22.5	21.6	0.26	0.25
N	2089	1680	2089	1680

Notes: % Exposure is the share of sample exposed to at least 1 monthly shock in each period between conception and 3rd Birthday. Mean exposure captures the mean number of months of exposure experienced. "In-comm." refers to the the restricted in-community sample, consisting only respondents who are definitely resident in the community from conception until round 2.

columns 1 and 2 of Table 3.2, respectively. The mean number of months of exposure in each period is provided in columns 3 and 4.

3.4 Empirical Strategy

To assess the effects of early life rainfall shock exposure on personality trait formation in adolescence and adulthood, I estimate the following equation using ordinary least squares (OLS):

$$CSE_{ijgr} = \beta_0 + \beta_1^{'} P_{gt} + \beta_2^{'} N_{gt} + \beta_3^{'} H_{ij} + V_g + B_{t_0} + \varepsilon_r \tag{3.1}$$

Where CSE_{ijgr} is age-standardised CSE measured at outcome in round 5 (age 14-23), for child i, born in household j, in community g, located in district r. P_{gt} and N_{gt} are vectors of community-level rainfall shock indicators for each of the 4 periods, t: 9 months of gestation (prenatal phase) and the first, second, and third year of life (postnatal phase). These are defined:

$$P_{gt} = \sum_{n=1}^{m} \mathbb{1}(SPI_{gn} \geq 1.5), \qquad N_{gt} = \sum_{n=1}^{m} \mathbb{1}(SPI_{gn} \leq -1.5) \tag{3.2}$$

where P_{gt} and N_{gt} capture the magnitude of positive and negative shocks experienced in community g in period t, respectively. These are measured in units of months m, where the function $\mathbb{1}(\cdot)$ takes a value of one if the SPI value for community g in month n =

$1, 2, ..., m$ is equal to or more extreme than the cutoff values defined in subsection 3.3.4. $\mathbf{H_{ij}}$ is a vector of child- and household-specific controls, as described in subsection 3.3.3. V_g is a time-invariant community fixed effect, and B_{t_0} is a year-month birth cohort fixed effect.

Given potential similarities in weather patterns across nearby communities, estimates of standard errors at the treatment level are likely biased (Auffhammer et al., 2013). Therefore, I cluster standard errors in the base specification by district, r, a higher administrative level than the community, to allow for local spatial correlation across communities within the same district area, as recommended by Dell et al. (2014). However, this yields a relatively small number of clusters (38) of unequal size. Asymptotic justification for cluster robust standard errors assumes many clusters, generally exceeding 40-50 groups, of equal size. In the presence of too few clusters, standard errors are biased towards zero and inference based on standard asymptotic tests will lead to an over-rejection of the null hypothesis. As such, I implement a cluster wild bootstrap procedure, as recommended by Cameron et al. (2008), to derive adjusted p-values, based on 10,000 iterations.

Additionally, to allow for arbitrary spatial correlation over space regardless of administrative boundaries, I compute standard errors adjusting for spatial correlation between nearby units, as proposed by Conley (1999), using a Bartlett kernel decay which allows for a spatial-weighted covariance matrix with weights declining linearly from one to zero over a distance of 50km from the community. Finally, I assess the robustness of results to adjustments for multiple hypothesis testing, deriving adjusted q-values following Anderson (2008), reported separately in Table C.11.

3.5 Results

3.5.1 Main Results

Results for the impact of rainfall shocks for each period of the perinatal phase on full and in-community sample age-standardised CSE scores are listed in Table 3.3. For the main results, three p-values are reported: those derived from a) the potentially downwards biased cluster robust standard errors, reported in parentheses; b) cluster wild bootstrap procedure using 10,000 replications, reported in square brackets, and c) the spatial correlation robust standard errors (Conley, 1999), reported in curled brackets. As expected, the p-values for wild bootstrap specifications are generally more conservative than standard cluster robust p-values. Interestingly, the p-value derived from Conley spatially-robust standard errors are generally smaller than the cluster robust values. this is likely due to many communities in the sample being located within 50km of others,

given the clustered nature of sampling in YL.[8] This leads to there being relatively few independent clusters, and potentially fewer than by clustering at district level in cases where the distance from a community to the district border is smaller than a 50km radius. As this methodology is also asymptotically justified, standard errors will be downwards biased.[9] Therefore, this method likely does not represent a refinement over cluster robust standard errors. For the rest of the analysis the more conservative wild bootstrap approach is the preferred specification with p-values reported in all subsequent tables, and alternative p-values reported in the relevant appendix tables.

A pattern is clear across all specifications, that a positive rainfall shock (+1.5 S.D.) experienced in the 2nd and 3rd year of life is associated with a lower age-standardised CSE score in adolescence and young adulthood. Exposure to a similar positive shock in the prenatal phase is associated with a higher later-life standardised CSE score, with exception of the full sample naïve score estimated effect, which is marginally insignificant under the wild bootstrap procedure. There is no significant effect estimated for exposure to a positive shock in the 1st year of life. Similarly, there are no statistically significant effects estimated for exposure to negative shocks in any period. Results are consistent between the full and in-community samples, suggesting shock exposure does not seem to influence migration choices.[10]

Estimates are more precise for the preferred EFA-derived score compared to the naïve score. As noted above, it is intuitive that EFA provides a more precise estimate, as the technique extracts as the 1st factor the dimension which explains the greatest amount of variation, reducing dimensionality and noise. Additionally, the naïve score treats all three scales as contributing with equal weighting towards the higher order construct of CSE, even though there is strong evidence that the different constituent scales contribute asymmetrically (Chang et al., 2012).[11]

In the primary specification, exposure to a positive rainfall shock during the prenatal phase is associated with a 0.068 S.D. higher CSE score (cluster wild bootstrap p-value $p = 0.055$). This result is consistent with Krutikova and Lilleør (2015), who find an 0.083S.D. higher CSE score associated with a 10% increase in the natural log deviation of rainfall in the year preceding birth in Tanzania. Considering the postnatal period, exposure to a similar shock in the 2nd and 3rd year of life is instead associated with a -0.090S.D. and -0.105S.D. lower score respectively, significant at the 1% level. This reverse of direction compared to the in-utero period is again consistent with the findings

[8] The mean number of communities within 50km is 8.

[9] For an example of this, see: https://blogs.worldbank.org/impactevaluations/randomly-drawn-equators

[10] Exposure to these short-run, relatively mild shocks is likely not severe or long-lasting enough to elicit a migratory response.

[11] Table C.3 estimates the impact of shock exposure on the three component scales separately. While the pattern of effects are similar, the magnitude of effect is different across scales, with the largest effect sizes estimated for the Marsh self-esteem measure.

Table 3.3: Impact of (±)1.5 S.D Shocks on CSE Scores, by Measure and Sample

	EFA 1st Factor		**Naive z-score**	
	Full	In-comm.	Full	In-comm.
Positive shock				
Prenatal	0.068	0.096	0.052	0.081
	(0.036)**	(0.013)**	(0.096)*	(0.037)**
	[0.055]*	[0.041]**	[0.123]	[0.074]*
	{0.019}**	{0.002}***	{0.084}*	{0.017}**
1st year	0.043	0.051	0.027	0.043
	(0.175)	(0.262)	(0.430)	(0.338)
	[0.162]	[0.280]	[0.424]	[0.368]
	{0.124}	{0.216}	{0.339}	{0.260}
2nd year	-0.090	-0.093	-0.091	-0.095
	(0.007)***	(0.016)**	(0.007)***	(0.012)**
	[0.006]***	[0.016]**	[0.007]***	[0.014]**
	{0.005}***	{0.008}***	{0.004}***	{0.005}***
3rd year	-0.105	-0.129	-0.097	-0.115
	(0.001)***	(0.004)***	(0.003)***	(0.009)***
	[0.003]***	[0.009]***	[0.006]***	[0.020]**
	{0.000}***	{0.001}***	{0.001}***	{0.004}***
Negative shock				
Prenatal	-0.030	-0.062	-0.036	-0.073
	(0.434)	(0.190)	(0.346)	(0.147)
	[0.434]	[0.257]	[0.364]	[0.246]
	{0.424}	{0.169}	{0.351}	{0.139}
1st year	0.066	0.075	0.036	0.043
	(0.170)	(0.232)	(0.423)	(0.510)
	[0.180]	[0.255]	[0.420]	[0.543]
	{0.149}	{0.199}	{0.405}	{0.492}
2nd year	-0.056	-0.048	-0.056	-0.033
	(0.478)	(0.517)	(0.457)	(0.637)
	[0.571]	[0.550]	[0.556]	[0.662]
	{0.470}	{0.520}	{0.451}	{0.642}
3rd year	-0.084	-0.069	-0.071	-0.045
	(0.120)	(0.258)	(0.170)	(0.399)
	[0.168]	[0.287]	[0.235]	[0.414]
	{0.119}	{0.209}	{0.170}	{0.355}
Controls	Yes	Yes	Yes	Yes
N	2089	1675	2089	1675

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Full sample refers to children geolocated in round 1. In-community restricts sample to those whose mother lived in the same community from conception until round 2. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

of Krutikova and Lilleør (2015) who, in an additional analysis, estimate a small negative effect on CSE scores of exposure to increased rainfall in the 2nd year of life (they do not assess the impacts for a 3rd year of life), however their result is not significant. That a significant effect is estimated for the 2nd to 3rd year of life, may reflect the pattern of brain development and plasticity, with the number of synapses per neuron in the brain growing from approximately 2500 at birth to a peak of around 15,000 between age 2 and 3 (Gopnik et al., 1999). However, the potential channels through which rainfall impacts CSE and how they may have a heterogeneous effect in different periods of exposure are not clear *a priori*, and are discussed in more detail in section 3.6. Next, I assess the potential for heterogeneity across different sub-groups.

3.5.2 Heterogeneous Effects

A common finding in the literature is that early-life shocks impact outcomes heterogeneously across different sub-groups, particularly across gender and socio-economic standing (Almond et al., 2018; Currie & Vogl, 2013), with results often being driven predominantly by the impact on boys or girls, and the strongest effects generally found amongst the poorest or least-educated households. To assess this, shock variables are interacted with indicators for: a) if a respondent child is female; b) if the mother has completed primary education (achieved grade 6 or higher); and c) if the household is in the bottom quartile of the wealth index (Or "poorest" category, see Briones, 2017). Lastly, it is hypothesised that rainfall shocks may disproportionately impact agricultural communities which are reliant on rainfall for crop production. A community is considered agricultural if, for all households in that community, 40% or more report household members being actively engaged in agricultural work. Estimated effects for EFA 1st factor CSE scores are shown in Table 3.4.

Notably, there is a positive effect estimated for the interaction between a prenatal shock and if the respondent is female or from the poorest households, both significant at the 5% level, accompanied with an insignificant baseline effect. The linear combinations of the shock exposure and the interaction terms are both statistically significant, with cluster wild bootstrap p-values of $p < 0.001$. This suggests that the mainline positive effect estimated for prenatal shock exposure is driven predominantly by the effect higher rainfall has on girls and those from the lowest wealth households. In contrast, for post-birth shock exposure there are no evident heterogeneous effects by gender. Interestingly, a large and significant baseline effect is estimated along with an almost equally large and oppositely signed interaction term for the poorest households, suggesting the negative effect of exposure in the 3rd year is nullified for the poorest households (I fail to reject the null hypothesis that the linear combination is statistically different from zero; cluster wild bootstrap p-value $p = 0.887$). This suggests that any potential benefit from increased

Table 3.4: Heterogeneous Effects of +1.5 S.D Shocks on CSE Scores

	Female	Poorest	Mother's education	Agricultural
Level term	0.087	-0.193	0.226	-
	[0.519]	[0.350]	[0.004]***	
Positive shock				
Prenatal	0.024	0.046	0.073	0.038
	[0.650]	[0.222]	[0.189]	[0.405]
**Interaction*	0.100	0.173	-0.004	0.086
	[0.080]*	[0.024]**	[0.949]	[0.183]
1st year	0.057	0.062	0.081	0.078
	[0.354]	[0.136]	[0.246]	[0.328]
**Interaction*	-0.029	-0.040	-0.059	-0.041
	[0.760]	[0.782]	[0.446]	[0.666]
2nd year	-0.106	-0.081	-0.055	-0.125
	[0.010]**	[0.021]**	[0.224]	[0.154]
**Interaction*	0.032	-0.046	-0.055	0.034
	[0.601]	[0.426]	[0.493]	[0.735]
3rd year	-0.100	-0.142	-0.116	-0.114
	[0.073]*	[0.000]***	[0.001]***	[0.049]**
**Interaction*	-0.013	0.133	0.010	0.038
	[0.825]	[0.024]**	[0.890]	[0.586]
Negative shock				
Prenatal	0.029	-0.051	0.019	-0.022
	[0.576]	[0.177]	[0.850]	[0.597]
**Interaction*	-0.122	0.153	-0.071	0.006
	[0.118]	[0.267]	[0.551]	[0.960]
1st year	0.082	0.071	0.066	-0.062
	[0.256]	[0.255]	[0.307]	[0.670]
**Interaction*	-0.039	0.015	-0.004	0.143
	[0.656]	[0.848]	[0.952]	[0.360]
2nd year	0.009	-0.046	0.020	0.085
	[0.955]	[0.473]	[0.759]	[0.445]
**Interaction*	-0.122	-0.041	-0.123	-0.190
	[0.345]	[0.743]	[0.096]*	[0.181]
3rd year	-0.071	-0.080	-0.062	-0.147
	[0.388]	[0.297]	[0.347]	[0.283]
**Interaction*	-0.035	-0.037	-0.042	0.120
	[0.726]	[0.796]	[0.681]	[0.316]
Controls	Yes	Yes	Yes	Yes
N	2089	2089	2089	2089

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed. Alternative p-values are reported in Table C.4.

rainfall for the poorest households may offset the negative effect seen in more affluent households.

In contrast, whilst mother's level of education, as measured by an indicator of if she completed primary education, is an important factor in determining CSE scores, there is no strong evidence of any interaction with shock exposure. Lastly, it was hypothesised that agricultural communities may be more vulnerable to the impacts of rainfall shocks, but there is no clear evidence of a differentiated experience of shocks.[12] The next section reports the results of a range of robustness and validation checks.

3.5.3 Robustness Checks

The results of calculating an SPI measure can be influenced by the choice of distribution used. Most commonly short interval data (1- or 3-month SPIs) are best fitted to a gamma distribution, however it can also be fitted as a lognormal, Weibull, or exponential distribution, and the optimal distribution differs based on local climate characteristics (Mishra & Singh, 2010). I define an alternative shock measure, fitting values to a lognormal distribution, as shown in Table C.5. Results indicate a similar pattern, however the magnitude and significance of some results differ, likely impacted by outliers and poor fit of data. Figure C.4 and Figure C.5 show the multi-density plot of monthly SPI values for each community in blue, in comparison with a standard normal density plot (of mean zero and standard deviation one) overlayed in red, for the transformed gamma-fitted and lognormal-fitted SPI measures respectively. SPI values should be approximately normally distributed if the underlying rainfall data is well-fitted to the theoretical distribution chosen. With exception of a few outliers, gamma-fitted distributions for each month generally follow an approximately normal distribution around mean zero (although often with a greater peakedness, suggesting extreme values may be less common than under the theoretical normal distribution). In contrast the lognormal-fitted SPI distributions for every month display a significant negative skew, and a large peak above zero, but with very few extreme positive values, indicating that lognormal provides a poor fit for the relative distribution of rainfall in YL communities in Peru.[13]

Although I address arbitrary spatial correlation in estimates, it is possible that there is a temporal auto-correlative component impacting results when the effects of shock exposure in different periods are estimated jointly. Table C.6 shows the results obtained when age-standardised CSE scores are regressed on shock exposure in each period indi-

[12] This does not result from the indicator being a poor measure of if a community is agricultural, as alternative specifications (at household level, if an the individual HH reports a member of the household being engaged in agriculture as a primary activity or the location type of a household is rural; or at community level if a community leader reports arable crop or livestock farming as the primary activity for the community) do not yield qualitatively different results.

[13] Additionally, Couttenier and Soubeyran (2014) show that the SPI and other related indices, when well-defined, are more efficient than other commonly used linear measures.

vidually. The overall pattern of results remains, with exception that the prenatal effect is diminished in magnitude and not significant at conventional levels (wild bootstrap $p = 0.185$), which could suggest that the effect estimated for the prenatal period may be correlated with subsequent exposure in the postnatal period. However, column 1 of Table C.7 shows that when controlling for the cumulative number of positive or negative shocks, the estimated effect of cumulative shock exposure is non-significant, suggesting the link with subsequent exposure is likely weak. Under this specification, the prenatal effect remains significant at the 10% level, and the pattern of results remains similar. Columns 3 and 4 of Table C.7 report the effects of shock exposure when estimated separately by shock type (positive or negative), with results similar in sign, magnitude, and significance to the mainline results.

Furthermore, I test if other periods outside the window of gestation and the first 3 years of life have an impact on later-life CSE. Column 2 Table C.7 estimates the effect of exposure to shocks across all base periods as well as for the year prior to conception and 4^{th} year of life. To minimise measurement error, this is estimated for the in-community sample only, for which it is certain all respondent mothers are resident in the community across the whole period from before conception up until age 4 (the age of the youngest individuals interviewed in round 2). The pattern of results remains, with no significant effect found for exposure to shocks in the year prior to conception, or in the 4^{th} year of life. Overall these robustness checks suggest that results are unlikely to be spurious and are robust across specifications, although there may be some correlation between exposure in different periods, which may inflate the effect size when estimated jointly.

Auffhammer et al. (2013) show that precipitation and temperature can be correlated, with the sign of this correlation dependent on the region, and suggest that not controlling for temperature may lead to omitted variable bias. Column 1 of Table C.8 estimates the main specification controlling for community-specific average temperature across each period, with results remaining unchanged.

A potential transmission mechanism for early life rainfall shocks is that the shock impacts households directly through an agricultural channel. This can operate by affecting household agricultural yields and local food prices, or by increasing/decreasing the amount of time household adults spend working in agriculture-related employment. The YL sample includes several communities which are located within or on the outskirts of Lima, a large highly-urbanised and globally connected metropolitan area. It is likely that these communities would be the least affected by shocks if effects operate through this channel. I therefore re-estimate the main results on a sub-sample excluding these urban communities, as is common in studies of the effects of climate on human capital (e.g. Krutikova and Lilleør (2015) and Maccini and Yang (2009)). Results are reported in column 3 of Table C.8 with estimates of the coefficients of interest being 8-21% larger in magnitude.

Additionally, if results are predominantly driven by the effect on local agriculture, then it is possible that rainfall shocks which occur during the growing season of primary crops are most salient (similar to Krutikova & Lilleør, 2015). As a climatically and geographically diverse country, the primary crop grown and timing of planting and harvesting varies across the YL communities. Using data from the Peruvian Ministry of Agriculture (MINAGRI) on department-level yields for 6 different crops in 2010, I identify the primary crop by area sown in each department, as shown in Figure C.6.[14]

Following a similar procedure to Webb (2024) and Auffhammer et al. (2013) I estimate the department-level mean planting and harvesting dates using gridded crop calendar data (Sacks et al., 2010). Estimates of the average planting and harvesting days for 19 crops are provided on a 0.5x0.5° global grid, based on the nearest agricultural census data from 2000.[15] I aggregate this grid-level data to the department level to obtain the mean planting and harvesting date for the primary crop, rounding to the month level.

A growing season shock is defined if, for a given month, that month-of-year falls between the estimated planting and harvesting month (inclusive) for the primary crop of the department in which the YL community is located, using the same measure of shock as in the main results. The effects of early life growing season shock exposure on CSE scores at outcome are estimated in column 2 of Table C.8. The pattern of results remain consistent, however the magnitude of estimates for prenatal and 2nd year exposure are diminished and not significant at conventional levels. Due to data limitations, the exact crops which are the most important for YL households cannot be accurately identified, and the crops sown and activities carried out by households may differ significantly from the department level trend on a year-to-year basis.

The Young Lives study is structured as a cohort study, with the primary, younger cohort, born between 2000-2002, and a smaller, older cohort, born between 1994-1996.[16] While the main specification controls for time-invariant characteristics specific to every month-year birth group, there may be wider time-invariant differences in characteristics and response patterns between the older and younger cohorts. Column 4 of Table C.8 reports estimates after including a cohort fixed effect, with results remaining consistent and unchanged.

Another threat to my design is that early life shocks may lead to selective mortality, potentially altering sex ratios (E.g. Hoffmann, 2014). To assess potential sex selection within the sample conditional on shock exposure, I regress the sex of the child on prenatal shock exposure (as well as postnatal exposure, to test for any potential spurious

[14]The data used was obtained as part of the "cropdatape" R package, available at: https://github.com/omarbenites/cropdatape. Crops included are rice, quinoa, potato, sweet potato, tomato and wheat. Tomato was excluded as a perennial crop.

[15]available from the Center for Sustainability and the Global Environment at the University of Wisconsin-Madison: https://sage.nelson.wisc.edu/data-and-models/datasets/crop-calendar-dataset/

[16]An additional panel of the nearest younger sibling in age to the younger cohort child is also defined, however not all required outcomes were administered in this sample.

relationship between climate shocks and sex). Results are reported in column 1 of Table C.9, with no evidence of impacts of shock exposure on sex ratio within sample. This is expected, as short run deviations in rainfall are mild compared to extreme shocks that could influence mortality (Almond et al., 2018).

Furthermore, results could be biased if shock exposure is predictive of selection into the sample. In particular if shock exposure influences migration decisions, those exposed pre-birth may migrate out of shock-prone communities. While this is not directly testable within this analysis, I test if shock exposure is predictive of the household's choice to migrate between birth and round 2 (when respondents are aged 4-13). Column 2 of Table C.9 shows that shock exposure in the prenatal period and years 1-3 of life does not predict migration.

Given the finding of a positive effect on later life CSE of prenatal exposure to positive shocks, it is of interest to understand if the effect of exposure is isolated to a specific trimester of gestation. Both Krutikova and Lilleør (2015) and Chang et al. (2022) assess the effect of exposure to precipitation shocks by trimester in-utero on later life CSE, with both finding no differential effect of shocks by trimester. In contrast, I find that the effects of prenatal shock exposure are isolated to the 3rd trimester of pregnancy, with exposure to a positive shock in this period associated with 0.115 S.D. higher adolescent and adulthood CSE, as shown in Table C.10. Additionally, a negative shock in the 3rd trimester is associated with a -0.133 S.D. lower score. Overall, this is suggestive that exposure in the final trimester is particularly important for future personality trait formation, whether due to a direct effect on the mother and child, or indirectly by impacting the home environment and resources available within the household immediately after-birth.

Lastly, given a large number of hypotheses being tested in the main results, a concern is the potential for over-rejecting the null hypothesis (type I error) due to multiple inference. In Table C.11 I control for the false discovery rate (FDR), the expected proportion of rejections that are type I errors, computing sharpened q-values as described by Anderson (2008). This procedure presents a trade-off between preserving statistical power and reducing false positives by vastly reducing the penalty of additional hypotheses.[17] The outcomes of interest remain significant when controlling FDR at q=0.10, with exception of the prenatal effect in the full sample naïve z-score measure (q-value = 0.203).

3.6 Mechanisms

The exact causal channel through which early life rainfall shocks impact adolescent and adulthood personality trait formation is unclear *a priori*. This section explores several

[17]For comparison, the Bonferroni method is generally overly conservative (Romano & Wolf, 2005), providing low statistical power.

potential mechanisms and presents a body of evidence across a range of additional analyses.

3.6.1 Nutrition

Rainfall shocks may impact personality trait formation in low- or middle-income settings by affecting nutrition. This could occur directly, influencing local food availability, impacting the child in-utero through intrauterine exposure to maternal under-nutrition, or immediately post-birth by affecting breastfeeding and/or early nutrition when weaning (Krutikova & Lilleør, 2015; Rosales-Rueda, 2018). I assess the role of nutrition by regressing early life shock exposure on an indicator of being stunted at age 8. Stunting, a commonly used indicator of chronic under-nutrition or poor health, is defined as height $\leq$ -2 S.D. from the age and gender specific mean (height-for-age), following World Health Organisation (WHO) child growth standards (WHO, 1995). Using Young Lives data, Dercon and Sanchez (2013) show a positive association between age 8 height-for-age and self-esteem, a component of my CSE measure, in Peru. Results are reported in column 1 of Table C.12. No significant effect is found on the probability of being stunted, suggesting that nutrition is not a primary transmission mechanism between early life shock exposure and personality trait.

3.6.2 Child Health

Additionally, rainfall variation can impact child health beyond nutrition. Increased rainfall may disrupt sewage or drainage, contaminating water supplies, damaging crops, and leading to bodies of standing water, which can contribute to the incidence of water-borne diseases (Rocha & Soares, 2015). To assess the role of poor health, I assess the impact of shock exposure on 3 binary indicators of child health and disease burden: a) self-reported good/very good health at outcome, a measure of overall health; b) if a child has suffered a serious illness between birth and age 8; and c) if a child reports a long-term disability at outcome. Results are reported in columns 2-4 of Table C.12. I find no evidence of a link between child health and positive rainfall shock exposure.

3.6.3 Caregiver Stress

Alternatively the effects of early life shock exposure may be transmitted indirectly by impacting parental mental wellbeing. Pressures caused by shocks on finances, food availability, and employment may increase parental stress – affecting their parenting practices, availability, or temperament (Duque, 2017; Shoji, 2023; Trinh et al., 2021). Primary caregiver mental wellbeing is assessed for the younger cohort in each round using the WHO self-reported questionnaire (SRQ-20), a screening tool measuring symptoms

of non-clinical anxiety and depression that caregivers experienced in the 30 days prior to interview (Tuan et al., 2004). Using data from rounds 1-5, I construct a measure of the total number of symptoms reported (0-20), as well as a 'caseness' score for respondents reporting 7 or more symptoms. Results are reported in Table C.13.[18] I find no evidence of an effect of positive shocks in the 12 months prior interview on caregiver mental wellbeing.

Following Favara (2018), I construct an index of early parenting practices using questions administered to caregivers of the younger cohort in the 1st round regarding what actions they take in response to their child crying. A positive score indicates the respondent reports more good practices (such as cradling the child or singing to them, coded as $+1$), and a negative score indicates they report more detrimental practices (such as ignoring, shaking, or spanking the child, coded as -1). A list of reported practices is available in Table C.14. Results are reported in column 4 of Table C.13 and do not provide evidence of an association between positive prenatal rainfall shocks and a change in early life parenting practices.

3.6.4 Parent-Child Relationship

An important mechanism through which personality traits may be affected by early life shocks is by the influence they have on both the time and resources available for parents to invest in nurturing and interacting with children in the early years of development. A positive rainfall shock could have either an income effect, where higher rainfall may impact household income from agricultural work, leading to greater material investments, or a substitution effect, with parents increasing their labour supply in response to increased labour demand, reducing the time they are available in the household to care for the child. Evidence from neurobiology shows that a strong positive attachment with parents in early life promotes healthy brain development (Schore, 2001). Additionally, experimental literature on early life investments show that increased social stimulation and interaction between parents and children in early years has a lasting benefit on socio-emotional skills and personality traits, even when effects on other outcomes, such as cognitive ability, diminish as children age (Attanasio et al., 2020; Heckman et al., 2013; Walker et al., 2022). However, it is unclear *a priori* which effect dominates in the case of positive rainfall exposure (Kochar, 1999; Nordman et al., 2022). Therefore, I assess how rainfall shocks in the year prior to interview may impact the reported working hours of parents, and subsequently how this may drive a change in parents' investments, both materially and psycho-socially, in their children.

In rounds 2 and 4, adult household members were asked about their average daily

[18] As a robustness check I construct an alternative caseness score for 8 or more symptoms (Beusenberg et al., 1994), reported in column 3 of Table C.13.

hours spent working in up to three economic activities, and the relative importance of each activity in terms of income. In column 1 of panel A of Table 3.5 I assess the effect of a positive rainfall shock in the 12 months preceding interview on the daily working hours for the most important (main) activity. Controls for household rural location; household wealth index; respondent gender, age and age-squared; and fixed effects for survey round, month of interview and community are also included. I estimate that for each month of positive rainfall shock experienced in the previous 12 months, a parent works an additional 11 minutes per day in their main activity. This suggests parents are working more and spending less time in the household, although this effect may not be substantial in practical terms.

However, it is possible that main activity labour supply is inelastic to shock exposure if that activity is contracted/salaried with fixed hours. 46% of respondents report more than one activity, with the most important task economically not always the task for which respondents spend the most time working on. Therefore, changes in labour may be masked if adults respond by working more in other activities. In panel B I assess the effect of shock exposure on the average sum of hours worked across all reported paid activity. I find that each month of exposure to a positive rainfall shock is associated with an additional 26 minutes per day of work, indicating that parents respond by working more in all reported activities, for which labour supply may be more elastic.

It is expected that mothers and fathers may respond differently to shock exposure, particularly if there is an uneven distribution of childcare and domestic work. As such, column 2 reports the impact of a positive shock including an interaction term for if the respondent is female (e.g. the mother) for the main activity (panel A) and all paid activity (panel B). Looking at the differences between fathers and mothers, results differ between the main activity and all paid work. For the primary economic task, while mothers report working fewer hours generally, there is little difference by sex in the response to positive shocks, with both men and women increasing work in their main job. the p-value for the linear hypothesis suggests that the total estimated effect for women of 10 minutes per day is significant at the 10% level. In contrast in Panel B, while the effect for the base group (fathers) is large (+55 minutes per day for each month of exposure), a large negative additional effect is estimated for mothers, and I cannot reject the null hypothesis for the combined effect at conventional levels.

Additionally, other household members, such as adult older siblings, aunts and uncles, or grandparents, who may assist with caring for the child, may also alter working hours in response to rainfall shocks, substituting for more work or more childcare to accommodate the parent's responsibilities. Therefore, I report the same specifications for all working age (15-64) household members present in the household roster and reporting economic activity in columns 3 and 4. Results for the main effect indicate a very similar impact of shock exposure on the main activity working hours of all household members, with the

total effect in column 3 estimated at 11 minutes, and the combined effect for women of 10 minutes, both significant at the 5% level. Similar to parents, for all paid activity, a base group effect for men of 45 minutes is estimated, but I cannot reject the null of a zero effect on total working hours for women.

Table 3.5: Impact of +1.5 S.D Shocks in Previous Year on Adult Hours Worked

	Parents		All HH adults	
	(1)	(2)	(3)	(4)
Panel A: Main activity				
Female	-1.808	-1.837	-1.541	-1.548
	[0.000]***	[0.000]***	[0.000]***	[0.000]***
Positive Shock	0.190	0.208	0.189	0.202
	[0.073]*	[0.096]*	[0.047]**	[0.045]**
**interaction*		-0.037		-0.031
		[0.786]		[0.754]
$H_0 : \beta_2 + \beta_3 = 0$ p-val.		0.057		0.027
N	5324	5324	7341	7341
Panel B: All paid activity				
Female	-4.384	-3.534	-3.479	-2.793
	[0.000]***	[0.000]***	[0.000]***	[0.000]***
Positive Shock	0.432	0.917	0.356	0.743
	[0.038]**	[0.000]***	[0.005]***	[0.000]***
**interaction*		-1.160		-0.983
		[0.020]**		[0.012]**
$H_0 : \beta_2 + \beta_3 = 0$ p-val.		0.372		0.233
N	5394	5394	7438	7438

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on wild bootstrap procedure (10,000 replications) provided in "[.]" brackets; Controls include: if HH is rural and wealth index (R1); respondent is female; age and age-squared. Fixed effects for survey year, month of interview, and community are suppressed. Alternative p-values are reported in Table C.15.

These results could indicate that while men, in particular fathers, experience a large increase in all activity, women may shift labour supply between economic activities such that although they increase hours slightly in their primary activity, there is no overall effect in total working hours. It may also indicate that women carry out set of additional economic tasks compared to men, which are not impacted by changes in rainfall. However, a limitation of the Young Lives dataset is that data on specific activity or industry code is missing for a large proportion of the sample, likely reflecting a high level of informal work.[19]

[19]Additional analysis using the 2015-2017 waves the Encuesta Nacional de Hogares (ENAHO), a

In summary, across all specifications, exposure to a positive rainfall shock is associated with an increase in working hours in paid work. The effect for all paid activity is slightly larger for parents, but the pattern of results shows that all household adults work more hours in response to recent positive rainfall shocks across both the main activity and all paid economic activity.

If household adult members respond to positive rainfall shocks by increasing labour supply, it is possible the burden of childcare and domestic work within the family home, farm, or business may shift to non-adult household members, such as older siblings. Young Lives collects time use data for all household members between the ages of 5 and 17 between rounds 2 and 5, asking how many hours are allocated across several categories on a normal weekday. In Table C.17, for older siblings of the YL child, I regress the average daily number of hours spent in each time use category on exposure to rainfall shocks in the previous 12 months. No category is associated with any statistically significant or meaningful changes in response to shock exposure, suggesting that the increased hours in work is not substituted for by other adult household members, who also work more, or by older sibling children, who do not alter their routines. As a result, a child may be experiencing less quality time with parents, negatively impacting opportunities for psycho-social stimulation and play.

However, it remains unclear theoretically whether the substitution or income effect of a rainfall shock dominates. For example, it may be that increased income leads to greater material investments in children, outweighing a reduction in availability of parents. I measure material investment in children in two ways. First, using questions administered to caregivers in round 3 regarding their investment in household reading resources (number of books, dictionary ownership, and child reading habits), I construct an index of "reading encouragement". For each question in Table C.18, if the caregiver responds affirmatively to these questions, they score a 1, otherwise scoring 0. A z-score is derived of the mean item score. Additionally, I construct a measure of educational expenditure on the YL child in round 3, including on clothing, fees, materials and transport. Results are reported in columns 3 and 4 of Table 3.6.

If however, the substitution effect of a positive rainfall shock dominates, then, as hypothesised above, we would expect to see some negative impact on measures of the social relationship between parents and children. to test this I construct two measures for the parent child relationship, from the perspective of both parties. First I proxy the caregiver's perception of the relationship by constructing an index measuring the level of involvement and knowledge of their child's life, based on questions administered in round 3, listed in Table C.18. Construction follows the same procedure as the reading

nationally representative cross-sectional household survey, suggests a 1-month exposure to a positive rainfall shock at district level is associated with a moderate increase in hours worked per week by respondents working specifically in an agriculture related occupation (based on ISIC Rev.4 4-number occupation code). Results from ENAHO are reported in Table C.16.

encouragement scale. Results for the impact of early life rainfall shock exposure on this index are reported in column 1 of Table 3.6. I measure the child-parent relationship from the perspective of the child (column 2), using the parent relations scale of the Marsh Self description questionnaire-II (Marsh, 1990), administered as a self-reported measure for both cohorts in round 4. A higher score indicates a child has a positive relationship with their parents. Similar to the main CSE outcome, I construct an EFA 1st factor score using age-standardised item responses. The scale shows high unidirectionality and internal consistency (Yorke & Ogando Portela, 2018), with all items loading highly on the 1st factor only. A list of the included scale items and 1st factor loadings is given in Table C.19.[20]

Table 3.6: Impact of +1.5 S.D Shocks on Parent-Child Relationship Measures

	Parent involvement	Parent Relations	Reading encouragement	Education expenditure
Prenatal	-0.017	0.085	-0.053	-0.044
	[0.641]	[0.042]**	[0.137]	[0.172]
1st year	-0.047	0.013	-0.066	-0.070
	[0.408]	[0.840]	[0.170]	[0.122]
2nd year	0.036	-0.006	0.081	0.030
	[0.339]	[0.875]	[0.112]	[0.682]
3rd year	-0.109	-0.062	0.018	0.016
	[0.003]***	[0.028]**	[0.647]	[0.727]
Controls	Yes	Yes	Yes	Yes
N	2089	1995	2089	2089

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on wild bootstrapped procedures (10,000 replications) provided in "[.]" brackets. Controls: HH is rural and HH wealth index; mother age and education; child gender, mother tongue, age, and if they were enrolled in pre-school. Fixed effects for birth month cohort and community are suppressed. Alternative p-values reported in appendix Table C.20.

Effects estimated for material investments are insignificant at conventional levels for positive shock exposure in any period. In comparison, prenatal exposure to a positive shock is associated with a 0.085 S.D. increase in the child's perception of the parent-child social relationship, while exposure in the 3rd year of life is associated with a -0.062 S.D. decrease, both significant at the 5% level. This pattern is similar to that found for CSE and is partially mirrored by the variation in the caregiver-reported parent involvement index, with a significant -0.109 S.D. decrease associated with shock exposure in the 3rd year, although no significant relationship for this measure is estimated for prenatal exposure.

This supports the hypothesis that the negative substitution effect of the shock dom-

[20]The sample size is smaller than the full sample as i) scale items were only administered to respondents if at least one parent is alive at the time of the round 4 interview, and ii) the scale was not constructed if respondents were missing at least one item.

inates a positive income effect, at least in the postnatal period. However, it remains unclear how positive shock exposure in the prenatal period translates to a positive impact on adolescent and adulthood personality trait formation. That shock exposure affects an individual's relationship with their parent in the same pattern observed for later life CSE scores may suggest that the effect of shocks pre-birth have a differential effect on parental labour supply than exposure post-birth. For example, exposure to positive shocks during pregnancy could have a positive effect, by increasing the hours worked before the child is born, allowing for consumption smoothing and for parents to be able to reduce labour supply in the following period after the child is born.

However as a limitation of this study, I am unable to assess how parental labour supply responds to shocks before and just after the birth of a child. As households are first tracked shortly after birth, no data was collected on parental labour supply prior to the birth of the young lives child. Additionally, as Young Lives is a cohort study, it would be difficult to disentangle the effects of shock exposure for households from the general effect on labour supply of having a newborn in the household. I am therefore unable to provide further insight into the underlying mechanisms for effects of shock exposure prior to birth.

3.7 Conclusion

I contribute to the literature identifying the importance of early life circumstances in determining later-life human capital stock, expanding the limited evidence on the effects of early life rainfall shocks on personality trait formation, and offer extensive analysis of the transmission mechanisms. I find prenatal exposure to a positive rainfall shock is associated with a higher core self-evaluation in adolescence and adulthood. In contrast, a similar exposure to a positive shock in the 2^{nd} and 3^{rd} year of life is associated with a lower later-life CSE score.

Considering mechanisms, there is no evidence that this effect operates through child nutrition or health, parental mental health, or through influencing material investments in children. In contrast, I find that households respond to a positive rainfall shock by increasing labour supply, particularly for the father of the child. Evidence suggests that this affects the emotional and social bond developed through parent-child interaction, with both parent and child perceptions of their long-term relationship being impacted by postnatal exposure to rainfall shocks. However it remains unclear how the positive effects of a prenatal shock exposure are transmitted.

Assessing heterogeneity, I find that the positive effect on CSE of a prenatal shock is isolated to girls and those born in the poorest households, findings common in the early life circumstances literature (Almond et al., 2018). Results are robust to asymptotic refinements which adjust for too few clusters, providing more conservative test statistics,

and remain robust after adjusting for multiple hypothesis testing.

That there are differential impacts of rainfall shocks pre- and post-birth indicates that the timing of exposure is important, suggesting that policy interventions that allow households to smooth consumption over periods, and facilitate greater early childhood social stimulation between children and parents, could be the most effective at improving later-life socio-emotional skills – for example, child benefit payments targeted at the early years of childhood. Furthermore, results suggest future evaluations of the effects of climate shocks should incorporate the potential long-term impacts on non-cognitive outcomes, with a growing literature showing that, even when cognitive differences diminish over time, socio-emotional effects often persist (Attanasio et al., 2020; Heckman et al., 2013; Sevim et al., 2023; Walker et al., 2022). This will likely become increasingly important as abnormal climate shocks become more frequent due to climate change.

Chapter 4

Sibling Spillover Effects in Education: Evidence from an Extended School-day Reform in Peru

4.1 Introduction

Educational reforms generally target specific individuals, with subsequent policy evaluations commonly focused only on the direct effect on those individuals. However this practice does not capture any potential spillover effects the policy has on others in their family, including siblings. Siblings grow up in the same home, belong to many of the same social groups, experience the same social interactions, and strongly influence each others' choices and outcomes (Black et al., 2021). Importantly, if policies targeting one child have an unanticipated and unmeasured impact on their siblings' outcomes, such evaluations will systematically under- or over-estimate the cost-benefit ratio (Altmejd et al., 2021; Figlio et al., 2023). It is therefore salient for policymakers to quantify potential spillovers, particularly when they affect those not eligible for the reform. However, it can be difficult to causally identify the influence of siblings on each other, given the "reflection problem" which is prevalent in the study of peer effects (Manski, 1993). This is particularly true outside of data-rich high-income contexts, where it remains unclear how sibling spillovers propagate through households which may be subject to credit constraints, and in which the cost of decisions, such as additional schooling and work, may be greater.

This chapter examines the potential for sibling spillovers on younger siblings' outcomes in Peru, exploiting exogenous variation in older siblings' schooling as a result of *Jornada Escolar Completa* (or JEC), a nationwide extended school-day reform which added two pedagogical hours per day and improved school resources in 1,000 public secondary schools. This reform aimed to bring the weekly number of instructional hours

and the quality of inputs and resources in treated public secondary schools into line with those offered by high-quality private schools in the country, which often have better learning outcomes than their public counterparts (Agüero et al., 2021). The reform was designed to be comprehensive, with increased pedagogical resources and pay for teachers to reflect the additional responsibilities and contact time, as well as providing students with access to psychological support and improved IT infrastructure. Assignment of schools to the program was not random, however we are able to exploit initial arbitrary eligibility criteria which were based on the number of home-rooms in the school (*secciones*, or sections). The final selection of schools was likely based on several unobservable decisions, therefore we use a fuzzy regression discontinuity design (FRDD) motivated under the assumptions of the local randomisation framework (Cattaneo et al., 2015, 2017, 2024).

Previous work has identified robust moderate to large positive effects of JEC on the learning outcomes of the targeted child (Agüero, 2016; Agüero et al., 2021). This chapter expands this evidence by considering for the first time the potential for sibling spillover effects of the JEC reform. Specifically, we assess the spillover effect of an exogenous change in older siblings' schooling on the learning outcomes of their younger, primary school-aged siblings, who are not exposed to the treatment.

Our findings indicate that primary school-aged younger siblings of a child attending a JEC high school experience a positive spillover effect on their educational attainment, increasing test scores by 0.12 S.D. and 0.14 S.D. in reading and mathematics respectively. While this positive effect on test scores does not significantly impact the likelihood of younger siblings attaining the top grade classification ("at grade", in line with expected outcomes for their grade), it does have a positive and statistically significant effect on the likelihood of being classed as "in progress" (the second highest classification) or above in mathematics (5.4 percentage points (p.p.)). This suggests that spillovers have a greater effect for younger siblings who are lower in the grade distribution, partially reducing the attainment gap. Our findings are consistent with other work which addresses sibling spillovers related to the quality or amount of schooling experienced by one sibling (Figlio et al., 2023; Qureshi, 2018a).

Considering heterogeneities, we find that spillover effects are isolated to younger sisters, regardless of older sibling gender, although the largest effects are found for sister-sister pairs. In contrast, no statistically significant effect is found for boys, regardless of the gender of their older sibling. This finding informs the limited evidence on how spillovers may propagate differently across sibling pair gender mix (Dahl et al., 2023; Nicoletti & Rabe, 2019; Qureshi, 2018a), and complements the literature on the gendered responses to household inputs in education (Autor et al., 2019). Finally, we show that our results are robust to a range of validation and falsification tests.

This study contributes to the growing literature assessing the potential for sibling

spillovers in education (Altmejd et al., 2021; Dustan, 2018; Goodman et al., 2015; Gurantz et al., 2020; Joensen & Nielsen, 2018; Karbownik & Özek, 2023; Nicoletti & Rabe, 2019; Oettinger, 2000; Qureshi, 2018a, 2018b), in particular expanding the very limited evidence outside of high-income, data rich contexts. In doing so we provide, to the best of our knowledge, the first evidence of sibling spillovers resulting from a schoolday extension policy, widening the scope of benefits beyond just the targeted child. Our results are salient for policy evaluation, highlighting the importance of accounting for within-family externalities when determining the benefits and costs of educational reforms.

The rest of this chapter is set out as follows: section 4.2 provides a summary of the existing literature on sibling spillovers in education, and on school day extension reforms (subsection 4.2.1), as well as providing a conceptual discussion of how sibling spillovers may propagate (subsection 4.2.4 and subsection 4.2.5), and summarising the JEC reform and its selection criteria (subsection 4.2.2 and subsection 4.2.3). Section 4.3 describes the datasets and data matching process, including descriptive statistics. Our identification and estimation strategy, as well as pre-analysis checks are discussed in section 4.4. All results and discussion follow in section 4.5, and section 4.6 concludes.

4.2 Context

4.2.1 Literature Review

Our work relates to two strands of literature. First, it contributes to the evidence on the effects of school-day extension programs, which aim to address poor student performance with a focus on the importance of instructional time in learning outcomes (Ben-Porath, 1967; Carroll, 1963; Figlio et al., 2018). However, the evidence for its effectiveness is mixed. A US-focused review of literature suggests positive effects of school-day extension programs, in particular amongst students at risk of failing, however the weak research design of many studies makes it difficult to disentangle the effect of increased instructional time from other inputs, limiting the strength of any causal inference (Patall et al., 2010). Outside of a high-income setting, there may be a greater appeal for a transition from part-day to full-day schools, providing increased childcare and even school lunch programs (if offered), acting as a safety net for families (Pablo et al., 2015). However, there is also a higher opportunity cost, given smaller educational budgets and the lower quality of other complementary inputs (Agüero et al., 2021), although the design of JEC as a comprehensive reform with increased school resources and improved infrastructure may mitigate this issue. Furthermore, increased school time comes with a trade-off, reducing students' ability to work in family businesses, help with chores, or care for relatives, which may prove important in credit constrained and

low-income contexts.

Within Latin America a number of countries have implemented school-day extension programs, including Mexico (Cabrera-Hernández, 2020), Argentina (Edo & Nistal, 2022; Llach et al., 2009), Dominican Republic (Garganta et al., 2022), Chile (Barrios-Fernández & Bovini, 2021; Bellei, 2009), Uruguay (Cerdan-Infantes & Vermeersch, 2007), Brazil (Almeida et al., 2016), and Peru. While the evidence suggests a mostly positive impact on learning, amongst other outcomes, there is a great deal of heterogeneity between studies, and a recent cost-benefit exercise suggests there are likely other more cost-effective policies as alternatives (see Pablo et al., 2015, for a review and in-depth discussion).

Second, it contributes to the literature on within-family spillovers. Although there is a developed literature addressing intergenerational spillovers (e.g. parent-to-child; see Black & Devereux, 2011), less studied is the potential for spillovers between siblings on each other's educational choices and human capital outcomes. As siblings grow up together and make choices concurrently, estimates of cross-sibling correlations in outcomes likely suffer from the same "reflection problem" identified in peer effects by Manski (1993). That is, it is difficult to infer from observed outcomes whether peer-group behaviour affects individual behaviour, or if group behaviour is simply the average of individual behaviour – even after controlling for shared characteristics. A further practical issue is that studying sibling spillovers generally requires access to high quality administrative data on students' enrolment and attainment, which can also be successfully linked to those of their siblings and/or wider household (Dahl et al., 2023). As such, the evidence on sibling spillovers has been very limited until recently. In an early contribution, Oettinger (2000) uses two-stage least squares to estimate the impact of having an older sibling graduate high school on younger siblings' graduation rates in the US using the National Longitudinal Survey of Youth, finding a positive effect on younger siblings' graduation probability. However, in the absence of exogenous variation (e.g. a policy reform, or natural experiment) their instrument validity (sibling-specific background characteristics) relies on strong assumptions.

Several papers take advantage of country/state-wide school and health administrative data, matching healthcare records of siblings to an individual's attainment data, as well as data on school-peer abilities, to assess the impact on learning outcomes of having a sibling with a physical, mental, or learning disability (Black et al., 2021; Breining, 2014; Persson et al., 2021). More closely related to this chapter are studies which focus on spillovers across siblings resulting from educational choices or inputs. Generally, these exploit quasi-random variation created by a policy reform, providing exogenous variation in the treatment status of one sibling. A number of these studies focus on college or school course choice, exploiting pre-determined entry cut-off scores to assess the impact of an older sibling being accepted to a selective school, college, or major on the choices of younger siblings. Goodman et al. (2015) estimate the relationship between the college

choices of siblings, finding younger siblings are more likely than their peers to enroll in 4-year and highly-competitive colleges when their older siblings do so first. Altmejd et al. (2021) assess spillovers in college and major choices in Chile, Croatia, Sweden and the US, using a regression discontinuity design with multiple college-specific thresholds for admissions, finding younger siblings are more likely to enrol in the same college and college/major combination if their older sibling was marginally accepted and attended, especially amongst social groups with lower college enrolment rates.[1]

Looking at high school level decisions, Dustan (2018) explores how an older sibling scoring above the threshold on a placement exam for high-quality schools in Mexico city positively influences the probability of younger siblings applying for those schools, while Gurantz et al. (2020) find that younger siblings in the US are more likely to take Advanced Placement courses if their older sibling marginally passes an exam. Similarly, Joensen and Nielsen (2018) exploit a pilot program in Denmark which provides variation in the requirements for older siblings to take advanced courses in secondary school, but not for younger siblings, to identify positive spillovers on the probability that younger siblings choose the same courses later on. They find younger siblings are more likely to choose advanced mathematics and science classes, especially amongst brothers who are close in age, suggesting the sex-mix of the sibling dyad and age-spacing matter for spillovers. This finding is echoed by Dahl et al. (2023), who exploit admission thresholds based on GPA in Sweden for oversubscribed college majors. Their results suggest the magnitude and direction of sibling spillovers are dependent on the sex-mix of the dyad, with same-sex younger siblings more likely to apply for the same major as their older siblings'. Younger brothers are particularly likely to follow their brother if there is a larger than 3 year gap in age (therefore they were not enrolled in high school at the same time). In contrast, younger brothers are less likely to go for the same course as an older sister, more so when there is a less than 3-year gap in age.

Finally, the strand of work closest to our analysis is that which assesses spillover effects on younger sibling attainment, rather than course or college choice, by exploiting variation in the quality or amount of schooling experienced by older siblings. Nicoletti and Rabe (2019) estimate the spillover effect of having a high-achieving older sibling on younger siblings' achievement, using a fixed effects value-added model, and by instrumenting the test scores of older siblings with the mean scores of their peers, finding a significant positive spillover from older to younger siblings. Similarly, using student level data which links siblings via birth records, and further linking to school level data in North Carolina, Qureshi (2018a) estimates the spillover effect of a child being taught by a more experienced teacher on their older or younger siblings, finding a positive spillover

[1] Additionally, Aguirre and Matta (2021) also assess spillovers in college and major choices in Chile, finding a large effect of older siblings' college choices on younger siblings choices, however do not find any effect on the choice of major.

to the younger child. While these solutions are useful as they don't rely on a specific policy reform for variation, leading to greater generalisability of results, they may require strict and potentially unrealistic assumptions (Sacerdote, 2014; Todd & Wolpin, 2003), and access to rich panel data which can be linked across schools and siblings, limiting their viability outside of data-rich high income contexts.

Another identification strategy exploits differences in the school starting age of older siblings, exploiting either sharp or fuzzy cutoffs in the date for school eligibility. Age at first entry to school has previously been shown to have a significant impact on the focal child, with children who are older relative to their peers showing better attainment at the same stages (Bedard & Dhuey, 2006; McEwan & Shapiro, 2008). Karbownik and Özek (2023) identify spillovers from having an older sibling who is born after the cutoff (hence one of the oldest in their class) on younger siblings in Florida, with effects concentrated in low socio-economic status households, who score higher on standardised tests. These results are similar to those found by Zang et al. (2023) in North Carolina.

The closest work to ours is Figlio et al. (2023), who assess the impact of an older sibling marginally missing a minimum reading score threshold in 3rd grade, leading to grade retention and being provided additional targeted support, on younger sibling outcomes in Florida. They find a large spillover effect of an older sibling's increased schooling and extra support on younger sibling scores in the same reading test.

The majority of the literature focuses on high-income contexts, where the mechanisms through which spillovers transmit likely differ from those in low- and middle-income contexts, where older siblings can play a large role in the care of younger siblings, including helping with homework or tutoring. Qureshi (2018b) estimates how a more educated eldest sister impacts younger brothers' educational attainment in rural Pakistan. They use an instrumental variables approach, exploiting cultural norms for chaperones which create significant disparities in girls' access to schooling based on their distance to the nearest school, finding that a more educated older sister is associated with higher literacy, numeracy, and years of schooling for younger brothers.

This study will expand the literature on sibling spillovers in low- and middle-income countries by investigating how variation in the amount and quality of public schooling experienced by an older sibling may have a causal spillover effect on the outcomes of their younger siblings. Specifically we will assess the potential for externalities on the academic achievement of the younger sibling of a child who attends a Jornada Escolar Completa (JEC) public school in Peru. Our approach differs from other similar studies as assignment to a JEC school is not based on measures of the older siblings' academic ability, but rather on school level selection criteria driven primarily by budget constraints, discussed in detail below. Additionally, our study is the first to consider the potential for sibling spillovers as a result of a school-day extension reform, described in detail in the following section.

4.2.2 Jornada Escolar Completa

There has been a significant expansion in access to education in Peru since the beginning of the 21st century, with high rates of enrolment in both primary and secondary school, and low rates of children out-of-school, however there is significant heterogeneity in the quality of education and learning (Saavedra & Gutierrez, 2020). This is typified by Peru's performance in PISA (Programme for International Student Assessment), where it has consistently ranked near the bottom in mathematics, reading, and science, and ranked last among the 65 participating countries in 2012.

In light of this, Peru has enacted several reforms to the education system, including introducing the *Jornada Escolar Completa* (JEC) program in 2015. It was designed as a comprehensive reform, planned to improve the quality of public secondary education and close learning gaps (Escobar & Sanchez Castro, 2021). As a result of increased school enrolment in the 1970s, many Peruvian secondary schools operated with separate morning and afternoon shifts to allow for a greater capacity (Saavedra & Gutierrez, 2020). A major component of JEC was to introduce a full-day model, increasing the number of pedagogical hours from 35 to 45 hours per week (equating to 2 hours extra per school day) in 1,000 public secondary schools nationwide.[2] A breakdown of differences with regular shift-based public schools by subject is provided in Table D.1. However, the reform was designed to be comprehensive, therefore the program aimed to also improve complementary inputs and resources, focusing on 3 components: 1) improved pedagogical support; 2) improved school management and organisational practices; and 3) improved physical infrastructure and increased IT resources. The pedagogical component includes a support programme for teachers (*Acompañamiento pedagogico*), as well the provision of psychologists to meet with students at least twice annually. Online support and training for school management was offered, and the salaries of teachers and principals were increased in line with the additional workload required. Finally, the number of computers and laptops available to classrooms was increased, with additional IT maintenance support being provided. These changes were designed to mimic the contact time and resources provided by high-quality private schools (Alcázar, 2016; Escobar & Sanchez Castro, 2021). Given valid concerns about the potential for low quality complementary inputs to limit the potential positive impacts of schooling reforms in low- and middle-income contexts (Kerwin & Thornton, 2021; Mbiti et al., 2019; Pablo et al., 2015), it is likely that improving these resources will lead to a more effective impact of increased instructional time (Agüero et al., 2021).

Previous research has identified large to moderate positive effects on a range of outcomes for the targeted child. Using a fuzzy regression discontinuity design, Agüero et al. (2021) find significant increases in scores for mathematics (0.23 S.D.) and read-

[2] A pedagogical hour is 45 minutes long.

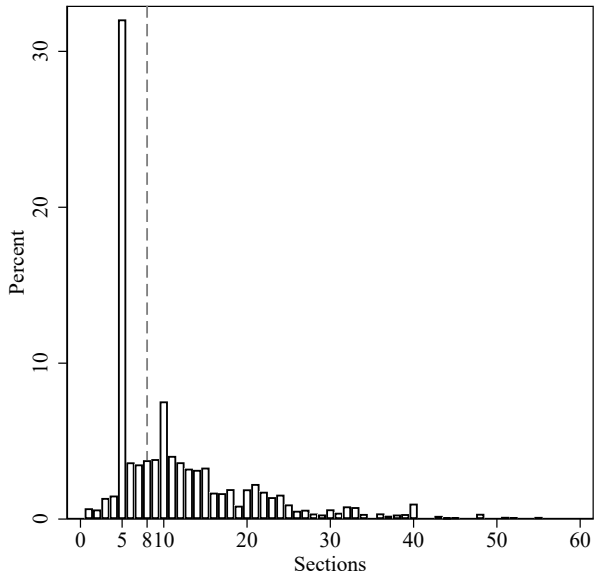
ing (0.19 S.D.) for children who attend a school which was part of the first round of JEC implementation in 2015, as well as improvements in socio-emotional competencies and technical/digital skills. Using the Young Lives Peruvian survey, Sanchez and Favara (2019) also find effects on non-school based tests of mathematics and reading comprehension (0.13 and 0.19 S.D., respectively) and higher self-reported self-esteem and self-efficacy. Rodrigues and Campos Flores (2021) use a mixed research design to evaluate the impacts of JEC beyond the initial sample of schools. The expansion of the program in 2016-2017 followed different eligibility criteria than in 2015, widening the types of schools eligible, therefore the outcomes for students in schools made eligible in these years were assessed using propensity score matching and a difference-in-differences approach. While Fuzzy RDD estimates for schools made eligible in 2015 were similarly as large as those of Agüero et al. (2021), the effect for students in schools made eligible in 2016-2017 were smaller in magnitude (0.04-0.07 S.D.). This may reflect that as the list of eligible schools widens, additional schools may not benefit as much from changes, or may simply reflect potential omitted variable bias in estimates. Looking at the long-term impacts of JEC on higher education outcomes using the Young Lives COVID-19 phone surveys, Hidalgo Arestegui (2021) finds students in JEC schools are more likely to completely secondary school (11.3 p.p.). While the estimated effect for students' likelihood of accessing university was not statistically significant at conventional levels, students of JEC schools were more likely to enter into a STEM major/stream at university (16.5 p.p.). Sanchez and Favara (2019) provide some support for these findings, seeing a 7.3 p.p. increase in students' aspirations to complete university education.

There is so far limited evidence of the wider impacts of JEC outside of the educational outcomes of the targeted child. Ortega (2018) and Sanchez and Favara (2019) assess how extended school-days may impact teenage fertility and sexual behaviour. Ortega (2018) estimate a moderate reduction in adolescent pregnancy (0.6 p.p.) using an instrumental variables approach. The findings of Sanchez and Favara (2019) provide some support for this, with male students having improved sexual health and contraceptive knowledge and female students showing increased pride, self-esteem, and sense of agency over their lives. Finally, and most closely related to our work, Ersoy and Forshaw (2023) assess the potential spillover effects of a child attending a JEC school on the labour market outcomes of parents, using a novel approach to match respondent households to schools via a geo-spatial algorithm. While they find weak evidence of a positive effect on fathers' individual income, this finding is not robust. Their dataset does not record the school of children in the household, relying instead on matching publicly-schooled children of secondary school age to the nearest public school by walking distance. Therefore their results by design are subject to measurement error, which may introduce significant noise or bias in estimates.

4.2.3 Selection Criteria

Eligibility for JEC was based on a number of criteria, summarised in Table D.2. It applies only to public schools which operate a morning shift (allowing for the expansion of school hours without impacting afternoon shift students). Schools were also required to be large enough to accommodate for additional resources and infrastructure. Specifically, schools were required to have a minimum of 8 *secciones* (or sections, equivalent to home-rooms or form classes in the US or UK). This choice of 8 sections is arbitrary and primarily driven by budgetary constraints (Agüero, 2016). Given secondary education consists of five grades, schools most commonly have five sections, one per grade, and the density of schools remains relatively smooth around 8 sections as shown in Figure 4.1 (this is tested empirically in subsection 4.4.3).

Figure 4.1: Distribution of Public Morning-Shift Schools by Number of Sections



Prior to implementing JEC, personnel from the Peruvian Ministry of Education (MINEDU) used data on school characteristics, infrastructure and enrolment from the 2013 school census (*Censo escolar*) to identify 1,360 eligible secondary schools which match these criteria. A further 52 "emblematic" schools, which may not have meet all requirements but were believed to benefit greatly from inclusion, were added to this list. The list of 1,412 potential schools was sent to local coordinators to validate that schools met the requirements, with several schools being removed and added, to select a total of 1,343 schools. MINEDU then hired evaluators to reduce this list down to 1,000 schools, which were then included in the September 2014 government directive announcing the JEC reform (RM N°451-2014-MINEDU). Finally, this list was amended once more in February 2015 by replacing 6 schools (RM N°062-2015-MINEDU), prior to the implementation of JEC for the school year beginning March 2015.

The process of moving from the original 1,360 eligible schools to the final selection of 1,000 schools is driven by unobserved characteristics, reflecting the potential influence of local coordinators bargaining with administrators (Agüero, 2016), and endogenous selection of "emblematic schools" and replacement schools. As such a simple comparison of sibling pairs attending included schools with those attending excluded schools would not be suitable. We instead exploit the initial criteria used to identify the first list of 1,360 schools. Specifically, we restrict our sample to schools which are publicly administrated and have only a morning shift registered, and exploit the discontinuous jump in participation in JEC at 8 or more sections.

Notably, while the JEC reform was expanded in 2016 and 2017 to include further schools, the eligibility rules changed to allow schools with different characteristics to join, therefore extending our identification strategy would not be feasible. As such we focus only on those schools eligible under the 2015 criteria, excluding from our analysis schools which joined JEC in subsequent years.

4.2.4 Spillover Transmission Mechanisms

Sibling spillovers can arise through two channels. First, spillover effects could occur through a within-family peer effects channel (Manski, 1993; Sacerdote, 2014) based on the interactions between siblings. This could be due to direct interaction, with older siblings sharing knowledge and influencing behaviours by helping with homework or teaching skills. This could be of particular importance in low- and middle-income contexts, where often older siblings play an important role in caring for their younger siblings and may have more formal education than their parents (Qureshi, 2018b).

However, in this context the direction of effect is unclear *a priori*. In attending a JEC school, older siblings may have less time to spend in the household, directly impacting interaction by reducing the time they can spend interacting with their sibling, or indirectly by passing on these household responsibilities (e.g. household chores, unpaid labour or caring for others) to the younger sibling, negatively impacting younger sibling time use for schoolwork (substitution effect). Alternatively, given robust findings of the positive effects of increased instructional time and improved resources resulting from JEC on students' educational outcomes (Agüero et al., 2021; Rodrigues & Campos Flores, 2021), older siblings attending a JEC school may have greater knowledge, mastery of topics, and ability to provide more effective help and academic mentoring to younger siblings, leading to a greater return from time spent learning together (productivity/ability effect).

Furthermore this channel may lead to indirect spillovers. For example an older sibling attending a JEC school and benefitting from the reform may represent a role model, or encourage sibling rivalry and competitiveness, influencing the behaviours, aspirations, and choices of their younger siblings. This indirect channel is prevalent in the literature

addressing how older sibling choices of school, college, or major impacts younger sibling educational choices (Aguirre & Matta, 2021; Altmejd et al., 2021; Dustan, 2018). Generally, spillovers occurring through this indirect channel are expected to be positive, although it could be possible that younger siblings respond negatively to older sibling achievement.

Second, an indirect consequence of a policy may be that parents reallocate limited household resources to take advantage of any policy benefits. Given the positive impact of increased instructional time and better school inputs, they may shift resources towards the targeted sibling (reinforcing behaviour; Becker & Tomes, 1986; Grätz & Torche, 2016), shifting resources away from other siblings and potentially negatively impacting their outcomes. Or they may shift resources toward the untreated younger sibling (compensatory behaviour; Fan & Porter, 2020; Pitt et al., 1990) leading to positive spillover effects due to benefitting from increased household inputs. Alternatively, parents could respond in a more complex manner, focusing on equalising inputs across siblings (potentially positive spillover effects; Berry et al., 2020), or by investing differently across differing dimensions of human capital (leading to an ambiguous overall effect; Yi et al., 2015).

Finally, sibling spillovers may operate differently across sibling pair characteristics (Black et al., 2021; Karbownik & Özek, 2023), in particular by sibling pair gender mix, with potentially stronger effects for same gender pairs (Karbownik & Özek, 2023; Nicoletti & Rabe, 2019; Qureshi, 2018a; Zang et al., 2023) compared with mixed-gender sibling pairs, although such effects can vary across country contexts (Altmejd et al., 2021), and may not be as clear in low- and middle-income contexts, where for example older sisters may have disproportionately more responsibility for caring for and educating younger siblings (Qureshi, 2018b).

Unfortunately a limitation of this study is that we cannot directly address the potential underlying mechanism due to data limitations, however we do provide some context for how spillovers are transmitted by exploring how effects may propagate differently across sibling characteristics.

4.2.5 The Reflection Problem

As noted above, simple estimates of correlations between siblings' educational outcomes likely suffer from the "reflection problem". This problem is pervasive in studies of peer effects, where it is hard to disentangle from observed outcomes whether peer-group behaviour affects an individual's behaviour, or if group behaviour is the aggregate of individual behaviour. Manski (1993) identifies three component reasons for why peers may exhibit similar outcomes: 1) correlated effects, where individuals tend to behave similarly due to having similar individual characteristics or environmental influences; 2)

exogenous effects, where individual behaviour varies with the exogenous characteristics of the group; and 3) endogenous effects, where the behaviour of an individual is influenced by the behaviour of others in the group. The endogenous effect between group members (e.g. siblings in a household) is the effect of interest in this study, however even if one can adequately control for exogenous aspects at the group level (e.g. socioeconomic status, urban/rural location, etc.), simple estimates of group-member correlations will likely still be contaminated with unobserved correlated effects. For siblings, the issue of correlated effects is even greater than with larger peer groups, where it is possible that random assignment to the group can be exploited. We therefore attempt to isolate the endogenous effect of siblings' schooling on each other by exploiting exogenous variation in the amount of schooling experienced by one sibling, but not the other in the pair, due to the Jornada Escolar Completa reform.

4.3 Data

Data for this analysis comes from three sources: i) the *Evaluación Censal de Estudiantes* (ECE), a national assessment of student learning; ii) *Sistema de Información de Apoyo a la Gestión de la Institución Educativa* (SIAGIE), a digital platform used by schools to manage administrative student data, and iii) the *Censo Escolar*, an annual school level census. Student-level data from ECE was matched to corresponding records in SIAGIE. As this data includes confidential, identifying data, this matching was carried out by MINEDU, based on the request and specification of the authors. Further School level data was matched from the *Censo Escolar* by authors. Details on each dataset and the matching process are discussed below.

First, data on attainment for primary school children was collected from the 2015 and 2016 waves of the *Evaluación Censal de Estudiantes* (ECE), a national assessment of student learning administered in every region of the country to children in all primary schools, both public and private, with at least 5 students registered in the grade evaluated. The assessments are carried out in either Spanish or the native language for students whose first language is not Spanish. Primary school students are assessed in reading and mathematics.[3] For each topic, students are assigned a score based on the dichotomous Rasch model with mean 500 and standard deviation 100, and using cut-off points are categorised in one of three classifications: *satisfactorio*, indicating the child displays the expected ability for this grade and is likely prepared to face the next grade; *en proceso*, indicating only partial achievement of learning outcomes for this grade, but that they are likely still on track to achieve this; and *en inicio*, indicating the child has not

[3] Assessments are also carried out on a wider range of subjects for students in the 2nd grade of secondary school, but for primary students only reading and mathematics are assessed.

displayed expected ability, managing only tasks below what is required for this grade.[4] In this study we refer to these as "at grade", "in progress" and "initial", with scores standardised by grade. Descriptives are provided in Table 4.1.

Table 4.1: Outcome Descriptives

	(1) All	(2) Matched sibling	(3) + Public + morning	(4) + 7/8 Sections
Mathematics				
Z-score	-0.00	-0.01	-0.13	-0.18
	(1.00)	(1.00)	(1.01)	(0.98)
Initial or lower	0.29	0.30	0.35	0.36
	(0.46)	(0.46)	(0.48)	(0.48)
In-progress	0.40	0.40	0.39	0.40
	(0.49)	(0.49)	(0.49)	(0.49)
At-grade	0.31	0.30	0.27	0.25
	(0.46)	(0.46)	(0.44)	(0.43)
Reading				
Z-score	-0.00	-0.07	-0.29	-0.35
	(1.00)	(0.98)	(0.96)	(0.92)
Initial or lower	0.15	0.16	0.21	0.20
	(0.36)	(0.36)	(0.40)	(0.40)
In-progress	0.41	0.43	0.47	0.50
	(0.49)	(0.50)	(0.50)	(0.50)
At-grade	0.44	0.41	0.33	0.30
	(0.50)	(0.49)	(0.47)	(0.46)
Observations	1533067	152761	59966	4368

Notes: Column (1) presents the mean scores and proportion of students per score classification for the pooled ECE sample of all students. column (2) provides statistics for those students with at least one matched sibling attending a secondary school in 2015, with column (3) restricted further to only those students with a sibling attending a public morning shift-only school. Column (4) includes only those matched sibling pairs within the window $W = [x_{-1}, c] = [7, 8]$ used for local randomisation analyses, with all previous sample restrictions.

For the 2015 wave, only students in the 2nd grade of primary were assessed, while in 2016 students in both the 2nd and 4th grade were assessed. As such our dataset encompasses three consecutive cohorts of students (who, following normal progression, would have been in grades 1, 2, and 3 when JEC was first implemented in 2015), although assessment is conducted at two different stages. Tests were conducted in primary schools towards the end of the school year (March-December), on November 10th-11th in 2015, and in 2016 on November 29th-30th and December 1st-2nd for the 2nd and 4th grades

[4] for 4th grade students there is an additional category *previo al inicio*, for students below even *en inicio*. To remain comparable across waves and grades, these two categories are merged.

respectively. Coverage of schools and students for surveys is provided in Table D.3, with surveys administered to between 94-96.9% of students countrywide in their relevant grades.

As well as test scores, ECE provides basic information on school (school ID code, location, public/privately administered) and child characteristics (gender, age, first language, and parent education level). Summary statistics for the pooled sample are provided in column (1) of Table 4.2.

Second, Student-level data from ECE is matched to corresponding records in the *Sistema de Información de Apoyo a la Gestión de la Institución Educativa* (SIAGIE), a digital platform developed by the Peruvian ministry of education (MINEDU) which is used by schools to manage student registration, log enrolment and attendance, and record student class performance. A common student ID is shared across ECE and SIAGIE, allowing respondents in ECE to be matched to their record. When enrolling students, a parent or guardian must provide their national identity card number (*Documento Nacional de Identidad* or DNI). Using this unique identifier for a parent or guardian, potential siblings can be identified as students who are enrolled under the same parent DNI number. All records of children registered with the same DNI number and who were enrolled in a secondary school in 2015, the year JEC was implemented, were identified and matched by MINEDU.

This method of matching is likely imperfect, for example two siblings would not be matched if an older child was registered by the mother, while the younger child is registered by the father, leading to different associated DNI numbers.[5] We cannot account for this in our dataset, however we expect that those missing due to this limitation are missing at random. Notably, many other studies in this literature also suffer from imperfect matching, and are often based on string-based matching using home address and shared last names (for example: Dustan, 2018; Goodman et al., 2015; Gurantz et al., 2020; Karbownik & Özek, 2023; Nicoletti & Rabe, 2019), which could potentially lead to a greater number of false-positive matches, or based on birth records, which do not allow for matching of siblings who co-reside together but were born to different mothers, or born out of state/country (Dahl et al., 2023; Figlio et al., 2023; Qureshi, 2018a; Zang et al., 2023).

Finally, using publicly available data from the 2013 *Censo Escolar*, a school census which was used by MINEDU personnel to define the original eligibility criteria for JEC, we match school level data on school type (public or private), shift pattern, and the number of sections, which is used to identify and determine treatment status, as well as other predetermined characteristics for testing covariate balance across the discontinuity

[5] Communication with MINEDU staff suggest that this is likely uncommon, as generally mothers register children at school. Our data reflects this, with 78.9% of representatives who register a child being female.

Table 4.2: Summary Statistics

	(1) All	(2) Matched sibling	(3) + Public + morning	(4) + 7/8 Sections
Household characteristics				
Mother completed primary	0.85	0.81	0.70	0.66
	(0.36)	(0.39)	(0.46)	(0.47)
Indigenous language	0.06	0.08	0.16	0.22
	(0.24)	(0.28)	(0.36)	(0.41)
Macro region				
Costa	0.56	0.52	0.37	0.32
	(0.50)	(0.50)	(0.48)	(0.46)
Sierra	0.32	0.37	0.49	0.55
	(0.47)	(0.48)	(0.50)	(0.50)
Selva	0.12	0.11	0.14	0.13
	(0.33)	(0.32)	(0.35)	(0.34)
Younger sibling characteristics				
Public school	0.72	0.82	0.96	0.97
	(0.45)	(0.39)	(0.20)	(0.16)
School is urban	0.87	0.83	0.68	0.69
	(0.34)	(0.38)	(0.47)	(0.46)
Child is female	0.49	0.49	0.49	0.49
	(0.50)	(0.50)	(0.50)	(0.50)
Older sibling characteristics				
Attends JEC school		0.16	0.34	0.39
		(0.37)	(0.47)	(0.49)
Public school		0.77	1.00	1.00
		(0.42)	(0.00)	(0.00)
School is urban		0.83	0.66	0.67
		(0.37)	(0.47)	(0.47)
Child is female		0.50	0.49	0.48
		(0.50)	(0.50)	(0.50)
Observations	1533067	152761	59966	4368

Notes: Column (1) presents summary statistics for the pooled ECE sample of all students. column (2) provides statistics for those students with at least one matched sibling attending a secondary school in 2015, with column (3) restricted further to only those students with a sibling attending a public morning shift-only school. Column (4) includes only those matched sibling pairs within the window $W = [x_{-1}, c] = [7, 8]$ used for local randomisation analyses, with all previous sample restrictions.

(see below). These are matched to older siblings using the school code provided within SIAGIE for the institution they were registered to in 2015. Summary statistics for the sub-sample of matched siblings are provided in column (2) of Table 4.2.

We exclude matches to siblings who are not school age, or who are listed as the same age or younger than the younger sibling. This matching method allows for multiple older siblings attending secondary school to be identified, with 15.8% of matches having more than one older sibling identified. Where applicable, we choose over other matched siblings a sibling who: 1) attends a JEC school (as this indicates the younger sibling is exposed to our treatment, i.e. having at least one older sibling attending a JEC school); or 2) otherwise attends a public, morning shift only school that is not a JEC school. For multiples that remain, we select the first matched sibling. As discussed in subsection 4.2.3, the JEC reform was expanded in 2016 and 2017. As inclusion criteria changed for these schools, they are no longer comparable with our initial selection criteria. To avoid potential bias in the estimated effect of treatment, we exclude sibling pairs for which the older sibling attended a school that subsequently joined JEC after 2015. Summary statistics for our analytical sample are provided in column (3) of Table 4.2.

4.4 Empirical Strategy

4.4.1 Local Randomisation

We exploit an initial eligibility criteria, which required schools to have eight or more classes to be eligible for JEC, to assess the potential for spillover effects arising due to exogenous variation in the schooling of an older sibling, on the educational outcomes of a younger sibling. This rule provides a clearly defined threshold at which the conditional probability of being assigned to JEC changes discontinuously (Lee & Lemieux, 2010). As the final selection of schools for the JEC treatment was driven by other unobservable decisions, the discontinuous jump in treatment probability is less than one, resulting in imperfect compliance (Imbens & Lemieux, 2008). We therefore implement a fuzzy regression discontinuity design (RDD), instrumenting participation in JEC with this initial eligibility rule.

All RDDs consist of three fundamental aspects: a score, which all units receive, a cutoff point, and a treatment rule that assigns units to treatment at values above the cutoff point. Conventional inference relies on the assumption that the conditional expectations of potential outcomes given a score are continuous (and differentiable) at the cutoff point, as formalised by Hahn et al. (2001), wherein the difference between the treatment and control average observed outcomes is equal to the average treatment effect in the limit (Cattaneo et al., 2020). This justifies the fitting of local polynomial regressions to approximate the unknown regression functions above and below the cutoff

and extrapolating these towards the cutoff point. This framework is referred to as the continuity-based RDD framework. Notably however, this framework assumes that the score variable that determines treatment assignment is a continuous random variable which can take any value. While continuity based approaches are still valid in the presence of discrete values when the number of values is sufficiently large, it can be inadequate when a discrete score variable consists of very few "mass points" (that is that many observations may share the same, relatively few values of a score) or that the mass points are sparse, such that most values are far from the cutoff (Cattaneo et al., 2024).

Our score variable is the number of sections in each school, and represents one such case. A suitable alternative is to consider a local randomisation approach to regression discontinuity design (LRRDD), formalised by Cattaneo et al. (2015) following work by Lee (2008). Under this framework, the validity of comparisons between treated and control units stems from assuming that, at least for a small window surrounding the cutoff, treatment is "as-if randomly" assigned. While this assumption can be seen as relatively stronger than the continuity assumption (at least when the score variable is continuous), it justifies the use of methods from experimental literature for estimation and inference, interpreting RDD as a local randomised experiment near the cutoff (Cattaneo et al., 2017). This approach is intuitive and closely-aligned with the justifications provided by Thistlethwaite and Campbell (1960) in first introducing RDD. The advantages of this method are that it avoids modelling assumptions, instead relying on assumptions on the assignment mechanism for units near the cutoff, and that inference is robust to the use of discrete score variables with few mass points (Kolesár & Rothe, 2018), indeed as little as one mass point either side of the cutoff (Cattaneo et al., 2024). Furthermore, this approach is easily extendable for the analysis of regression discontinuity designs with imperfect compliance (fuzzy RDD). The basic local randomisation framework for regression discontinuity, extension to fuzzy RD designs, and the required assumptions for identification and estimation, are set out in Appendix D.1.

Under the relevant assumptions, estimation of the fuzzy LRRDD local average treatment effect can proceed through the standard two-stage least squares procedure. In a first stage, we estimate for a sibling pair i the effect of the cutoff rule, $T_i = \mathbb{1}(X_i \geq 8)$, on the probability receiving treatment as the difference-in-means of those observations just below and just above the cutoff within a small window $W = [c - w, c + w]$ where local randomisation holds:

$$D_i^W = \alpha_0 + \alpha_1 T_i + \nu_i$$

Where D_i^W is the treatment received by sibling pair i with an older sibling attending a secondary school with a number of sections $X_i \in W$. We restrict our sample to include

only pairs with an older sibling attending a morning shift-only public secondary school, aligning with the other initial eligibility criteria as listed in Table D.2, which provides better comparison between the treatment and control groups. Given a discrete integer running variable our window is defined as the smallest possible window $W = [x_{-1}, c] = [7, 8]$, consisting of schools with 7 sections and 8, minimising the required extrapolation. Additionally, if local randomisation holds, it must hold within the narrowest window, therefore a window selection procedure, such as that proposed by Cattaneo et al. (2015) is not necessary, however we conduct this as a robustness check in subsection 4.5.3. Summary statistics for sibling pairs within this window are provided in column (4) of Table 4.2.

The second stage estimates the local average treatment effect (LATE) for compliers of JEC on younger sibling outcomes Y_i, using the fitted values from the first stage $\widehat{D}_i^W$, again as the difference-in-means between those sibling pairs just below and just above the cutoff within the same window:

$$Y_i = \beta_0 + \beta_1 \widehat{D}_i^W + \varepsilon_i$$

Where Y_i are our measures of younger sibling educational attainment. Specifically, we assess the grade-standardised score in reading and mathematics, as well as the probability of being classified as "at grade", or as "in progress" or higher. We make the assumption that potential outcomes are independent of the score variable X_i (LR 1 and LR 2 in Appendix D.1) within the window $W = [7, 8]$, therefore the regression functions are flat and the effects are estimated as the vertical distance between average observed outcomes.[6] Given our relatively large local randomisation sample, inferences rely on standard Gaussian large-sample approximations based on a heteroskedastic-robust covariance estimator of variance (Cattaneo et al., 2024).[7]

4.4.2 Two-Stage Least Squares

In addition to our local randomisation approach to RDD, we also apply a parametric global polynomial approach, estimating the effect of a sibling pair receiving the JEC treatment, instrumented by the initial eligibility threshold, using conventional two-stage least squares (2SLS). We estimate as a second stage the following:

$$Y_{ij} = \gamma_0 + \gamma_1 \widehat{D_{ij}} + \gamma_2(X_{ij} < 8) + \gamma_3(X_{ij} \geq 8) + \gamma_4 \mathbf{Z_i} + \mu_j$$

Where $\gamma_2(X_{ij} < 8)$ and $\gamma_2(X_{ij} \geq 8)$ are polynomial functions of the running variable

[6] While it is possible to relax this assumption (given other less restrictive assumptions (Cattaneo et al., 2017)) to allow for polynomial adjustments as in standard parametric RDD designs, it is neither necessary, nor applicable given our window consists of only one mass point either side of the cutoff.

[7] Alternatively, inference for small samples can be obtained under Fisherian finite-sample methods (Cattaneo et al., 2015).

X_{ij}, for sibling pair i, for which the older sibling attends school j, interacted with an indicator of being above or below the threshold. Our primary results are estimated using both a linear and quadratic form.[8] $\mathbf{Z_{ij}}$ is a vector of controls: sibling pair characteristics are captured by an indicator of being the same gender and age difference measured in years, as well as mother's level of education; fixed effects for survey year and older sibling's grade are also included. While controls are not required for identification of the RDD effect, they can provide asymptotic efficiency gains in large samples (Cattaneo et al., 2023), however we show in our robustness checks that 2SLS results are robust to estimation without covariates. $\widehat{D_{ij}}$ represents the residuals from the first stage measuring the impact of the cutoff rule $T_{ij} = \mathbb{1}(X_{ij} \geq 8)$ on the probability of receiving treatment, with all other terms included as listed in the second stage. Standard errors μ_{ij} are cluster robust at the level of the older sibling's school j. Cattaneo et al. (2020) note that global polynomial approximations, while providing a good approximation of unknown regression functions overall, likely provide poor approximations of the conditional expectations at the boundary point, and can be influenced by outliers far from the cutoff, leading to unreliable RDD estimates. Therefore, this specification is provided primarily for comparability, and we present the LRRDD results as our primary results.

4.4.3 Validation

Similar to a conventional IV setting, we require evidence of a non-zero and sufficient first stage. First stage results are reported in Table 4.3 for our local randomisation RDD effect, as well as for our linear and quadratic 2SLS specifications. A large and statistically significant effect is estimated for all three specifications, with a discontinuous jump in older sibling participation in JEC of between 60.3-68.0%. Graphical evidence of this discontinuity is provided in Figure 4.2. The relative strength of our first stage relationship can be measured by the effective F-statistic, with Cattaneo et al. (2024) recommending a higher rule-of-thumb threshold for RDD contexts of 20 or more. All three specifications provide evidence of a strong and relevant first stage.

Additionally, we assess the validity of our local randomisation RDD estimates by testing for systematic differences between treated and control groups near the cutoff. For the implementation of all validity tests we focus on intention-to-treat effects, as recommended by Cattaneo et al. (2024), as these tests aim to assess how similar observations are just below and above the cutoff, rather than the difference between those who receive treatment and those who do not. Balance tests for a list of predetermined sibling pair and school level characteristics are provided in Table 4.4. These tests provide evidence of the validity of our LRRDD estimates, as well as suggestive evidence that our

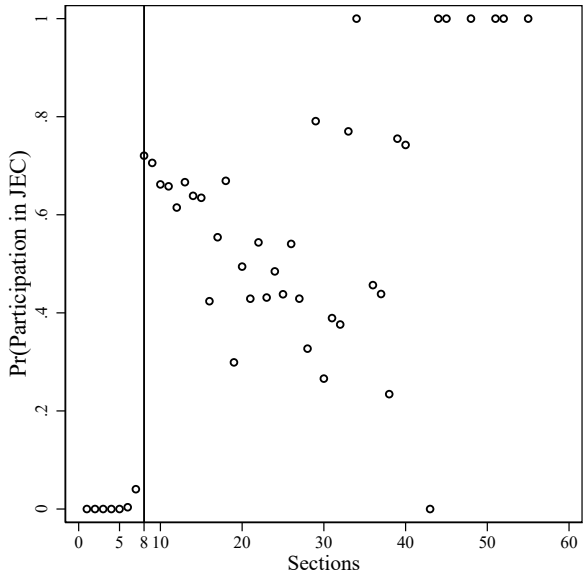
[8] higher order polynomials are not recommended, as they likely lead to noisy estimates and poor coverage of confidence intervals (Gelman & Imbens, 2019).

Table 4.3: First Stage: Participation in JEC

	Local rand.	Two-stage least squares (2SLS)	
	(1)	(2)	(3)
Sections≥8	0.680	0.604	0.661
	[0.000]***	-	-
	-	(0.024)***	(0.036)***
Spline	-	Linear	Quadratic
eff. F-stat.	4307.00	651.32	339.43
N	4368	59192	59192

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. 2SLS Cluster robust standard errors at the older sibling school level are reported in parentheses. Local randomisation results modelled without polynomial adjustment for the smallest possible window. For 2SLS, linear and quadratic splines of the running variable are specified. 2SLS additional covariates include sibling pair age difference in years, an indicator of being the same gender, and mother's educational attainment in years. Fixed effects for older sibling grade and survey year are also included.

Figure 4.2: Discontinuity in Participation in JEC at Cutoff



assumption of as good as random assignment within the window W holds, although we cannot account for imbalances in unobserved characteristics. Results suggest covariates are well balanced within our window, with non-significant differences estimated for those below and above the cutoff, with exception that sibling pairs are slightly less likely to report speaking an indigenuous language (-3.0% above the threshold) and schools are less likely to be located in districts which eligible for the Crecer welfare program. Balance tests for our 2SLS sample are provided in Table D.4, as well as graphical evidence of covariate smoothness in Figure D.1 for household/sibling dyad characteristics and school level characteristics in Figure D.2.

Table 4.4: Local Randomisation: Predetermined Covariates Balance Test

	School district receives program		Proportion of students		Pass rate
	Juntos	Crecer	Girls	Indigenous	Total
Panel A: School level					
Sections≥8	0.009	-0.053	0.003	-0.004	-0.004
	[0.532]	[0.000]***	[0.260]	[0.754]	[0.359]
Control mean	0.425	0.728	0.471	0.216	0.670
N	4368	4368	4368	4368	4368
	Sibling Dyad			Household	
	Same gender	Age diff.	Grade diff.	Indig. lang.	Parent educ.
Panel B: Sibling dyad level					
Sections≥8	-0.015	0.047	0.041	-0.030	0.006
	[0.314]	[0.415]	[0.366]	[0.017]**	[0.698]
Control mean	0.509	6.533	6.383	0.235	0.662
N	4368	4368	4368	4368	4320

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Additionally, we assess the density of observations. If units lack the ability to control precisely the value of cutoff score, placement of units below and above the cutoff should be as if random in the window W around the cutoff (Cattaneo et al., 2017). Agüero et al. (2021) provide evidence that students did not systematically select in to JEC schools.[9] Table 4.5 provides the results of a binomial test with null hypothesis $H_0 : Pr(k = 0.5)$. While we reject the null that the success probability is exactly equal to a half, our observed probability (k=0.52) indicates placement above the cutoff is close to that

[9]Additionally, the final list of JEC schools was only published in February 2015 (RM No 062-2015-MINEDU). With the school year running from March to December, it is unlikely that many parents could select in to treatment.

under a simple unbiased coin flip, suggesting sorting is unlikely near the cutoff. This result does not imply that the local randomisation assumptions are violated, and results are consistent with the fact that we would naturally expect slightly more students to be placed in 8 sections schools, given the additional class likely meaning more students per school on average, and with an assignment process using a bernoulli trail with a true probability of success k slightly more than one half (Cattaneo et al., 2024).[10] We carry out further RDD validation and falsification checks as part of our robustness checks in subsection 4.5.3.

Table 4.5: Local Randomisation: Binomial Test

	Observations below cutoff	Observations above cutoff	Observed probability	p-value
$H_0 : Pr(k = 0.5)$	2103	2265	0.52	0.015

Notes: Expected probability $Pr = 0.5$. Sample includes sibling pairs for which the focal child attends a public secondary school with morning only shift and either 7 (below cutoff) or 8 sections (above cutoff).

4.5 Results

4.5.1 Main Results

LRRDD Results are presented in Table 4.6. columns (1) and (4) provide estimates of the spillover effect for compliers of JEC (the LATE) on younger sibling scores, with effects of 0.120 S.D. and 0.135 S.D. estimated for reading and mathematics. Estimates for the probability of being graded as "at grade" (the highest group) are reported in columns (2) and (5), with estimated effects on the probability of being "in progress" or above reported in columns (3) and (6), for reading and mathematics respectively. While our estimates are positive, the effect sizes for reading grade probability are close to null, suggesting that although the spillover has a relatively large effect on the intensive margin, the effect on younger sibling reading does not have a significant impact at the extensive margin. The control mean score in reading for our analytical sample is -0.393 S.D. lower than the average calculated for all students in their grade, with siblings in our sample skewing towards the lower end of the distribution of attainment. For mathematics, while the effect estimated for the highest group ("at grade") is not statistically significant at conventional levels, an increase of 5.4 p.p. is estimated for the probability of being classed as "in progress" and above, significant at the 5% level.

[10] The conventional McCrary (2008) density test used in the continuity framework for non-parametric and parametric RDD is not suitable for discrete running variables with few mass points, providing misleading results (Frandsen, 2017). Therefore we do not provide an equivalent test for 2SLS results.

This effect at the extensive margin is modest, representing an 8% increase from the control mean, however performance in mathematics is generally poorer than reading for all students in ECE (see Table 4.1), therefore this represents a tangible improvement for treated siblings in our sample.

Table 4.6: Effects of JEC on Younger Sibling Outcomes: Local Randomisation Inference

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Sections≥8	0.120 [0.003]***	0.006 [0.760]	0.011 [0.530]	0.135 [0.002]***	0.030 [0.124]	0.054 [0.011]**
Control mean	-0.393	0.293	0.795	-0.226	0.238	0.626
N	4368	4368	4368	4368	4368	4368

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

That we see a stronger impact on younger sibling mathematics than on reading, in particular at the extensive margin, is consistent with the effects found by (Agüero et al., 2021), who find a larger and more robust effect on targeted child mathematics, while impacts for reading are smaller in magnitude and less robust, likely reflecting that students attending JEC schools benefitted from a 2-hour increase in mathematics, compared with a 1-hour increase in reading. This may indicate that our spillover effects operate through the direct within-family peer effect channel, with older siblings seeing greater improvements in their own mathematics ability, and therefore likely providing higher quality help to younger siblings (e.g. if they help their younger sibling with homework, or with studying generally).

Reduced form intention-to-treat effects are provided in Table D.5, with findings consistent with the fuzzy RDD estimates, although smaller in magnitude, reflecting downwards bias as a result of non-compliance. Additionally we re-estimate our main results using a parametric 2SLS specification as discussed above; results based on a linear and quadratic functional form are presented in Table 4.7, in panels A and B respectively. Findings under the linear specification are consistent with our primary LRRDD results: the older sibling's attendance of a JEC school is associated with a positive spillover effect on younger sibling reading and mathematics scores of 0.159 S.D. and 0.135 S.D.. In contrast with the LRRDD effects, while there is a positive effect estimated for both the probability of being "at grade" and for being "in progress" or higher, the effect for "at grade" is large (7.3 p.p. and 6.1 p.p.) and significant at the 1% level, while the effects for "in progress" and higher are no longer significant at conventional levels, although

they remain positive (3.2 p.p. and 4.3 p.p., respectively).[11]

Table 4.7: Effects of JEC on Younger Sibling Outcomes: 2SLS

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Panel A: Linear specification						
Sections≥8	0.159	0.073	0.032	0.135	0.061	0.043
	(0.059)***	(0.024)***	(0.022)	(0.063)**	(0.023)***	(0.028)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	59192	59192	59192	59192	59192	59192
Panel B: Quadratic specification						
Sections≥8	-0.056	-0.038	-0.011	-0.057	-0.013	-0.035
	(0.095)	(0.041)	(0.036)	(0.099)	(0.039)	(0.045)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	59192	59192	59192	59192	59192	59192

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) at the cutoff, including a linear (panel A) and quadratic (panel B) spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses. Additional covariates include sibling pair age difference in years, an indicator of being the same gender, and mother's educational attainment in years. Fixed effects for older sibling grade and survey year are also included.

Interestingly, estimates under the quadratic form are estimated as slightly negative, although none are statistically significant, with large standard errors greater in magnitude than the effect size. This suggests the global quadratic specification likely provides imprecise estimates of our effect. As previously noted, global parametric specifications tend to provide a poor approximation of the RDD effect at the boundary point (Cattaneo et al., 2020), in particular those of higher order which can provide noisy estimates and poor coverage of confidence intervals (Gelman & Imbens, 2019). Although the estimates, at least under a linear spline, are consistent with our LRRDD results, they may also be subject to the effects of large outliers, particularly far away from the cutoff; therefore we consider our findings under local randomisation to be more robust. An additional concern is the inclusion of so-called "bad controls", which may bias the effect upwards from zero. Given the inclusion of additional covariates is not required for the identification of RDD effects, we re-estimate our 2SLS results without the inclusion of controls in Table D.6, and find our estimates remain robust, although with slightly reduced magnitude and increased variance, as expected (Cattaneo et al., 2023; Noack et al., 2023).

4.5.2 Heterogeneity

In this section, we consider the potential for heterogeneity in effects across sibling pair gender mix. First we assess differences in the spillover effect across the gender of the

[11] graphical evidence of the discontinuity in outcomes at the cutoff, based on a linear fit, is presented in Figure D.3.

younger sibling. Results are presented in columns (1) and (4) of Table 4.8. A large spillover effect is estimated for younger sisters of compliers of JEC, with smaller non-significant effects estimated for younger brothers. Although our results are consistent with the literature documenting gender differences in sensitivity to family inputs (Autor et al., 2019), indicating that girls respond more to their older sibling's increased schooling than boys, and with the average female advantage seen in education (Autor et al., 2016, 2023) generally, our effects are the opposite of those found by Figlio et al. (2023), who find spillover effects of an older sibling being flagged for 3rd grade retention are concentrated amongst younger brothers, rather than sisters.

Table 4.8: Heterogeneous Effects of JEC on Younger Sibling Outcomes: By Gender Mix

	Female			Male		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Older sister	Older brother	All	Older sister	Older brother
Reading						
Z-score	0.189	0.216	0.163	0.057	0.025	0.076
	[0.002]***	[0.015]**	[0.050]*	[0.297]	[0.754]	[0.320]
At grade	0.023	0.048	0.001	-0.009	-0.024	-0.000
	[0.438]	[0.274]	[0.986]	[0.751]	[0.559]	[0.997]
≥ In prog.	0.053	0.025	0.081	-0.027	-0.041	-0.018
	[0.041]**	[0.525]	[0.023]**	[0.275]	[0.237]	[0.613]
Mathematics						
Z-score	0.227	0.315	0.146	0.049	0.051	0.037
	[0.000]***	[0.000]***	[0.095]*	[0.416]	[0.553]	[0.650]
At grade	0.055	0.077	0.035	0.005	0.012	-0.003
	[0.044]**	[0.050]*	[0.354]	[0.851]	[0.755]	[0.944]
≥ In prog.	0.105	0.130	0.082	0.007	-0.008	0.018
	[0.001]***	[0.005]***	[0.058]*	[0.798]	[0.850]	[0.649]
N	2136	1033	1103	2232	1070	1162

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Similar to other work in this literature, we further assess if there are heterogeneities based on the sibling-pair gender mix, which may reveal the role of older sibling gender or gender matching/difference in the transmission of spillover effects. In general, evidence for gender differences in sibling spillovers is mixed (Dahl et al., 2023; Steelman et al., 2002). However, recent studies find stronger effects for same-gender siblings (Karbownik & Özek, 2023; Nicoletti & Rabe, 2019; Qureshi, 2018a; Zang et al., 2023) compared with mixed-gender sibling pairs, although some studies find mixed gender effects, such as (Qureshi, 2018b) who shows in Pakistan that older sister education matters for younger

brothers educational outcomes.[12] Furthermore in a multi-country study, Altmejd et al. (2021) find that the importance of gender composition varies across countries.

Breaking down same and mixed gender effects by specific genders, results suggest that brother pairs are more likely to make similar choices in college or majors (Dahl et al., 2023; Joensen & Nielsen, 2018), suggesting boys are more likely to be influenced by their older brother's choices, but effects on younger sibling attainment are less clear (Karbownik & Özek, 2023; Zang et al., 2023). Columns (2) and (5) show the effects of having an older sister attend a JEC school, for girls and boys respectively. Likewise, columns (3) and (6) show the effects for having an older brother. As with our analysis above, results are concentrated amongst female younger siblings. Positive effects are estimated for female younger siblings regardless of having an older sister or brother, however the magnitude of effect sizes estimated for sister pairs are larger than those for the mixed younger sister-older brother pairs, suggesting that while girls generally respond more to spillovers, the response is greatest for same gender spillovers from their older sister. This would be consistent with both the female advantage in sensitivity to family inputs, as well as the evidence that same-gender pairs are more responsive, likely due to greater direct interaction, a higher likelihood to share similar interests, and to experience stronger role model effects (Benin & Johnson, 1984). Results estimated under the linear form for two-stage least squares are reported in Table D.7, and are consistent with the findings under our primary specification.

Another potentially important characteristic for heterogeneities is sibling age or grade-spacing, with effects generally larger for siblings who are close in age or grade (Dahl et al., 2023; Figlio et al., 2023; Zang et al., 2023) Closely spaced siblings are more likely to share peer networks and interact more, particularly if they attended the same primary school together. Given our policy reform and grade specific testing, we see limited variation in age-spacing in our sample. Additionally, as our dataset provides a snapshot, we do not observe siblings' educational histories, therefore we are unable to identify if students attended primary school together for any period of time prior to 2015, or if they attended the same primary school.[13] Therefore we are unable to accurately address the potential for heterogeneity by age or grade difference.

4.5.3 Robustness

This section presents additional analyses to test the robustness and validity of our primary findings. We provide three falsification tests. First, we assess if our estimated effects are spurious by conducting a placebo test with an alternative group of sibling pairs which

[12] They do not estimate effects on younger sisters, as they generally attend the same schools as their older sister.

[13] Grade progression in Peru is less clear-cut than in high-income contexts, with almost 30% of Peruvian respondents to the 2009 PISA reporting having repeated at least one grade (OECD, 2011).

were not eligible for the JEC reform, where the older sibling attends a private school. It is expected that there should be no systematic differences in outcomes between those units within the window below and above the cutoff. Results are presented in Table D.8, with no significant effects estimated across our outcomes.

Second, we assess the potential for spurious results by selecting placebo cutoffs where the probability of treatment assignment is not expected to change, keeping our window width the same (Imbens & Lemieux, 2008). We select one placebo threshold below the real cutoff at 4 sections, with the window $W = [3, 4]$, and one above with $W = [11, 12]$. These cutoffs are selected to avoid contamination from the actual treatment effect, such that no observations from the original window are present. Additionally, we choose these windows as they include similar numbers of observations either side of the cutoff, whereas windows including either 5 or 10 sections may show systematic differences as these are the most common mass points (see Figure 4.1). Results are presented for both placebo cutoffs in Table D.9, with no significant effects estimated, with exception of the positive effect for "at grade" reading in the $W = [11, 12]$ window (2.9 p.p.), significant at the 5% level, against a null effect in our main results. These tests together provide good evidence that our results at the actual threshold are unlikely to be spurious.

Third, we assess the robustness of our results to the choice of window. Following Cattaneo et al. (2015), we conduct a data-driven window selection procedure, based on the balance of a set of pre-determined covariates, Z_i, which iteratively tests the null hypothesis, $H_0 : Z_i(1) = Z_i(0) \quad \forall i$, within a widening window. We set a conservative threshold significance level for rejection, set at $\alpha \geq 0.15$ as recommended by Cattaneo et al. (2024), tolerating a higher probability of type I error to lower the chance of failing to reject a false null of balanced covariates.[14] Table D.10 displays the results for 5 increasingly widening windows, beginning with the original narrowest window, with a minimum step of 1 between windows, given our discrete running variable. The second and third columns provide the name and p-value of the variable with the minimum difference-in-means p-value within that window. the suggested window in this case is $W = [x_{-2}, c_{+1}] = [6, 9]$, after which the minimum p-value falls below our significance level threshold. Table D.11 presents results for our main outcomes, re-estimated using this window, showing results consistent with those estimated in our primary specification.

Our analysis uses data collected in the 2015 and 2016 waves of ECE. Unfortunately, due to disruptions and damage caused by flooding and heavy rains related to the 2017 El Niño Southern Oscillation (ENSO), the 2017 ECE was not conducted. Furthermore the 2019 ECE survey was conducted only in secondary schools, with further planned surveys disrupted by the COVID-19 pandemic. An ECE survey was conducted in primary schools

[14] As we have already shown two of our pre-determined covariates are not fully balanced in Table 4.4, these are excluded from our set Z_i. We therefore use this window only for testing the sensitivity of results, with all other inference being based on the original narrowest window.

for 2018 amongst 4th grade students only. Given our sample includes 2nd grade students tested in 2016, it is likely that the 2018 survey consists of a large overlap of respondents, however in our anonymised datasets we are unable to identify students across surveys. Therefore it is not possible to account for potential repeated observation with a panel data structure in our analysis if we were to include the 2018 survey. Additionally, the disruption of the 2017 ENSO, the worst to hit Peru since 1925 (Ramírez & Briones, 2017), may have impacted learning outcomes significantly, overtaking the effect of JEC. Therefore we did not consider the 2018 ECE survey as suitable for our primary analysis. However as a robustness check, we re-run our primary analysis on a pooled sample including the 2018 survey wave in Table D.12, with consistent results estimated, although slightly reduced in magnitude.

Finally, we estimate our results separately by survey year and wave, to assess if effects are similar across different groups. Results are presented in Table D.13. While generally consistent, relatively larger effect sizes are estimated for the 2015 2nd grade group, as well as for the 2016 4th grade group (who would generally be in 3rd grade in 2015), while smaller effects are estimated for the 2016 2nd grade group (1st grade in 2015), which are not significant at conventional levels. this may indicate that the spillover effect is greater for older children who are, on average, closer in age to their older siblings at the time of the reform in 2015.

4.6 Discussion and Conclusions

This study documents the potential for sibling spillover effects from a school day extension reform in Peru. Specifically, we exploit an arbitrary cutoff rule based on the number of classes (sections), which was used to define the selection of schools into the JEC program, to provide exogenous variation in the amount and quality of schooling experienced by an older sibling. We measure the impact of having an older sibling attending a school just above this cutoff on the reading and mathematics scores of younger siblings who are not yet eligible for the program. Using a local randomisation approach to regression discontinuity design we estimate the effect for compliers above the cutoff, finding evidence of a positive spillover effect on both standardised reading and mathematics scores of having an older sibling attend a JEC school, compared with those below the cutoff. In addition to impacts on scores, we also find some evidence of a spillover effect on the extensive margin, with a positive effect on the probability of being categorised as "in progress" or above for mathematics. However, we find no significant effect at the extensive margin for reading. Assessing heterogeneity in spillover effects, we find that effects are concentrated amongst girls, with the largest effect found for older sister-younger sister sibling pairs. Our results are robust to a range of falsification tests, indicating that estimates are valid and unlikely to be spurious.

There are however, several limitations that cannot be addressed by this study. While our novel matching strategy allows us to link siblings across schools to assess the impacts on educational outcomes, our dataset does not allow for an in-depth exploration of the specific mechanisms through which these spillovers are transmitted between siblings. We are also unable to conduct detailed analysis of how spillover effects may differ across household background characteristics, and we can only provide limited evidence of how effects differ across sibling pair characteristics. Additionally, given data limitations, we cannot produce robust evidence of how persistent effects are for siblings as they age. Finally, we focus primarily on the potential for sibling spillovers for younger siblings, in particular siblings who were still attending primary school, and therefore too young to be exposed to JEC. However, we are unable to provide any insight into the potential implications for the older siblings of the targeted child, or on other family members, who may be positively or negatively impacted by the focal child's increased time spent in school.

This study contributes to the growing literature for sibling spillovers in education, in particular, it expands the very limited literature assessing educational spillovers outside of high income contexts. Additionally, to our best knowledge, we provide the first evidence of sibling spillovers arising from a school day extension policy. We therefore also contribute to the literature on the efficacy of school day extension policies. Specifically, it provides evidence that there are positive spillover effects of the JEC reform, beyond the previously established impacts for the targeted child. Importantly, our findings of significant spillovers between siblings suggest that policy evaluations which fail to account for such externalities within families will likely provide misleading conclusions. It is therefore salient that policymakers consider the potential for spillovers when determining the benefits and costs of an educational reform.

Chapter 5

Conclusion

The chapters presented in this thesis provide evidence from Peru of how shocks and investments in different periods and across individuals can influence human capital formation. Chapter 2 investigated the multigenerational effects of prenatal exposure to drought. Chapter 3 examined how exposure to rainfall shocks in early life can have persistent impacts on personality trait formation, as measured in adolescent and young adulthood. Chapter 4 estimated spillover effects in educational attainment between siblings, arising from a nationwide school-day reform in public secondary schools. This final chapter now concludes with a brief review of the findings of each chapter, their implications, and suggestions for future areas of research.

Chapter 2 examines the multigenerational effects of maternal grandmothers' exposure to drought while pregnant, using a novel dataset from the Young Lives study in Peru. Results indicate that drought exposure during the pregnancy has a persistent impact on the health stock of their daughter (first generation) and grandchild (second generation). The first generation experience lasting impacts on their long-term health stock, being shorter in stature in adulthood. Looking at the dynamic effects for the second generation, I see persistent negative impacts on health stock, measured by height-for-age, from early childhood into adolescence, while early negative effects on weight-for-age, measuring health flow, fade as the child ages into mid-childhood. Effects are largest for the descendants of grandmothers living in rural areas during exposure, with exposure in the final of trimester having the strongest impact. Considering transmission mechanisms, I find evidence that the biological pathway of maternal long-term health acts as the primary channel, while measures of socioeconomic environment pathways do not have a significant mediating effect.

This chapter contributes to a small but growing literature identifying causal effects for the multigenerational transmission of health and human capital. The findings highlight the importance of quantifying and accounting for the multigenerational impacts of shocks in designing subsequent policy interventions. However, several questions remain. In particular, while effects persist into late adolescence, and seem to widen as children

enter pubertal growth, it is unclear if second generation effects are permanent, such that there is an impact on final adulthood growth potential. A future extension of this research could aim to address this issue, using an upcoming round of data collected in adulthood, which was previously delayed as a result of the global COVID-19 pandemic. Furthermore, that effects are driven by the impact on male grandchildren raises the question of how, if at all, effects may be passed on to further generations through the paternal line – a pathway that could not be explored in this analysis.

Chapter 3 assesses the impact of early life exposure to rainfall shocks on respondents' appraisal of their own self-worth, competence, and capabilities, as measured by their core self-evaluation (CSE) in adolescence and adulthood. Results indicate that exposure to positive rainfall shocks in the second and third year of life is associated with a lower CSE in later life. Alternatively, exposure to a positive rainfall shock in the prenatal period is associated with a higher CSE, although this effect is driven by female respondents and those in the poorest households. No effect is found for exposure to negative rainfall shocks in any early life period. Examining underlying mechanisms, results indicate that parents (and all working-age household members) increase labour supply in response to higher rainfall, particularly fathers, which has a negative impact on early-life social interaction and parent-child bonding. No effects are found on material investments in the child, or on indicators of the child's physical health and nutrition, suggesting that the substitution effect of reduced parental availability outweighs potential positive income effects from increased labour supply.

These results contribute to the literature which identifies the importance of early life circumstances in determining future human capital, expanding the very limited evidence on the effects on personality and socio-emotional skill formation. These dimensions of human capital are of growing interest, given strong associations with life-long academic and socioeconomic success, and these results suggest that the early life period plays an important role in their formation. While I attempt to address the underlying mechanisms for these effects, providing suggestive evidence for the importance of parent availability and social interaction in the early years of life after birth, it remains unclear how rainfall shocks experienced in-utero can positively influence later life personality traits. As measures of socio-emotional skills and personality traits are increasingly included in longitudinal and household surveys, a future line of work could shed some further light on how these effects are mediated prior to or just after birth.

Finally, Chapter 4 explores how siblings can influence each other's human capital. Specifically, we estimate sibling spillover effects on attainment from a national reform that extended the length of the school day and improved other inputs within public secondary schools. findings indicate that there are positive spillovers from having an older sibling attend a JEC school for younger sibling outcomes, as measured by their maths and reading scores in a national survey of student attainment. While this does

not impact the probability of attaining the highest grade classification, there is a positive effect on scoring the second highest grade of "in-progress" or above in mathematics, suggesting effects may be greatest for those in the middle or lower end of the grade distribution. Additionally, effects are driven by the impact on younger sisters, especially for sister-sister pairs, while there is little evidence of an effect on younger brothers. Given data limitations, it was not possible to provide evidence of how effects may differ across family background characteristics, such as the wealth, nor was it possible to provide an in-depth assessment of the transmission pathway for these positive spillover effects. Finally, given the design of our study, we cannot comment on younger-to-older sibling spillovers, or on the wider impact on parent time use and labour supply. This suggests a more comprehensive investigation of potential within-family spillovers as a future avenue for research, which would provide important implications for future education reform design.

Appendix A

Joint Authorship Statement

We certify that Gerald McQuade was involved in the conception, design, and procurement of data for the work in Chapter 3 titled: "Sibling Spillover Effects in Education: Evidence From an Extended School-day Reform in Peru". All data cleaning and data analysis was carried out by Gerald, and he was the primary contributor to the empirical design and drafting of the chapter.

Prof. Catherine Porter

Dr. Alan Sanchez

Appendix B

Chapter 2 Appendix

B.1 Estimating ACDE Using Sequential G-estimation

B.1.1 Average Controlled Direct Effect

Let Y_i be the *observed* outcome for unit i and $Y_i(a)$ the *potential* outcome if treatment was set to a. Following the potential outcomes framework (Rubin, 1974), the causal effect of treatment is the difference between the two potential outcomes in which unit i switched from treatment level a' to a:

$$\tau_i(a, a') = Y_i(a) - Y_i(a').$$

As we only observe one of these potential outcomes for a given unit, we focus on the average treatment effect (or total effect), defined as the difference in means between two different potential outcomes:

$$ATE(a, a') \equiv \tau(a, a') = E[Y_i(a) - Y_i(a')],$$

where $E[\cdot]$ is the expectation over units in the population of interest. Given some mediator M for the effect of treatment on the outcome, the controlled direct effect (CDE) can be defined as the effect of changing treatment value from a' to a while holding fixed the value of the mediator as m. As such, $Y_i(a, m)$ is the potential outcome for unit i, for a set level of treatment a, and mediator m. The potential value of the mediator may also be defined similarly as $M_i(a)$, the level the mediator takes on given treatment level a. The controlled direct effect is therefore expressed:

$$CDE_i(a, a', m) = Y_i(a, m) - Y_i(a', m).$$

As above, we focus on the average of CDE, defining based on the expectation over units in the population the average controlled direct effect (ACDE) as given by:

$$ACDE(a, a', m) = E[Y_i(a, m) - Y_i(a', m),] \tag{B.1}$$

which describes the average direct effect of treatment if the mediator is fixed at value m for all units in the population. See Acharya et al. (2016), Joffe and Greene (2009), and VanderWeele (2009) for in-depth discussion.

B.1.2 Assumptions

Following Acharya et al. (2016) I estimate the average controlled direct effect (ACDE) using "sequential g-estimation" (or reverse sequential two-stage (RS2S) parametric estimation), as set out by VanderWeele (2009) and Joffe and Greene (2009). The ACDE is identified under the following assumptions: 1) Sequential unconfoundedness; and 2) No intermediate interactions.

Assumption 1: Sequential unconfoundedness:

$$\{Y_i(a, m), M_i(a)\} \perp\!\!\!\perp A_i \mid X_i = x, \tag{B.2}$$

$$Y_i(a, m) \perp\!\!\!\perp M_i \mid A_i = a, X_i = x, Z_i = z, \tag{B.3}$$

For which the following conditional probabilities must be non-zero:

$$P(A_i = a \mid X_i = x) > 0,$$

$$P(M_i = m \mid A_i = a, X_i = x, Z_i = z) > 0,$$

For all possible treatment values $a \in \mathcal{A}$, mediator values $m \in \mathcal{M}$, covariates $x \in \mathcal{X}$, and intermediate confounders $z \in \mathcal{Z}$.

Equation B.2 states that the potential outcome $Y_i(a, m)$ (that unit i takes if treatment is set at value a and mediator at value m) and potential mediator value $M_i(a)$ (that the mediator would take under treatment level a) are conditionally independent of the observed treatment status A_i given covariates $X_i = x$. That is, there are no omitted relevant variables (U_{i1}) for the effect of treatment on the outcome or mediator conditional on pretreatment covariates. This is assumed to hold in this study given exogenous and random exposure to climate shocks in-utero. Equation B.3 further states that conditional on set levels of treatment, covariates and post-treatment (intermediate) confounders, the potential outcome is independent of observed value of mediator, e.g. there are no omitted relevant variables (U_{i2}) for the effect of the mediator on outcomes.

Under strong assumptions, these conditions could justify the use of standard regression analysis using a single equation, however this unlikely to be sufficient if there exists some post-treatment covariate Z which is influenced by treatment A, influences the mediator M, and is independently associated with the outcome Y (Robins, 1986). The

ACDE can still be identified non-parametrically under the above assumption alone in the presence of these intermediate confounders z_i, however this requires this requires the distribution of these confounders (conditional on A_i and X_i) to be known and correctly specified (Acharya et al., 2016; Joffe & Greene, 2009), therefore a further assumption is made.

Asumption 2: No intermediate interactions

$$E[Y_i(a,m) - Y_i(a,m') \mid A_i = a, X_i = x, Z_i = z] = E[Y_i(a,m) - Y_i(a,m') \mid A_i = a, X_i = x], \tag{B.4}$$

For all values $a \in \mathcal{A}$, $m, m' \in \mathcal{M}$, $z \in \mathcal{Z}$, and $x \in \mathcal{X}$.

This states that the effect of the mediator on the outcome must be conditionally independent of any intermediate confounders.

B.1.3 Identification

To derive the ACDE of treatment on outcome, we define a demediation function:

$$\gamma(a,m,x) = E[Y_i(a,m) - Y_i(a,0) \mid X_i = x]. \tag{B.5}$$

This function describes the average difference between outcomes with mediator set at level m and zero, and does not depend on the levels of intermediate confounders if Equation B.4 holds. By subtracting the demediation function from the observed outcome $Y_i = Y_i(A_i, M_i)$, variation in the outcome due to the mediator is removed:

$$E[Y_i - \gamma(a, M_i, x) \mid A_i = a, X_i = x] = E[Y_i(a,0) \mid X_i], \tag{B.6}$$

provided **assumption 1** is met, the effect of the mediator on the outcome is identified. The ACDE,

$$E[Y_i(a,0) - Y_i(0,0) \mid X_i = x],$$

conditional on pretreatment covariates X_i, is therefore identified as the difference in means of the demediated outcome:

$$E[Y_i - \gamma(a, M_i, x) \mid A_i = a, X_i = x] - E[Y_i - \gamma(0, M_i, x) \mid A_i = 0, X_i = x]. \tag{B.7}$$

B.1.4 Estimation

The ACDE is estimated parametrically using sequential g-estimation in a two-stage process.

B.1.4.1 First Stage

Under **assumption 1**, the demediation function (Equation B.5) can be estimated from the data as the difference in means estimator, conditioning on both the pretreatment covariates X_i and intermediate confounders Z_i. I therefore first regress the outcome on the treatment, mediator, and all covariates to obtain an estimate of the effect of the mediator on the outcome, from which I can derive the demediation function. In the simplest specification:

$$Y_i = \delta_0 + \delta_1 A_i + \delta_2 M_i + \delta_i X_i + \delta_i Z_i + \nu_i, \tag{B.8}$$

where there is no interaction between the mediator and outcome or covariates (and by **assumption 2** no interaction with intermediate confounders), the coefficient of interest is δ_2. The sample version of the demediation function is expressed as:

$$\hat{\gamma}(A_i, M_i, X_i; \hat{\alpha}) = \hat{\delta_2} M_i. \tag{B.9}$$

B.1.4.2 Second Stage

First, the outcome is adjusted using the estimated demediation function:

$$\tilde{Y_i} = Y_i - \hat{\delta_2} M_i. \tag{B.10}$$

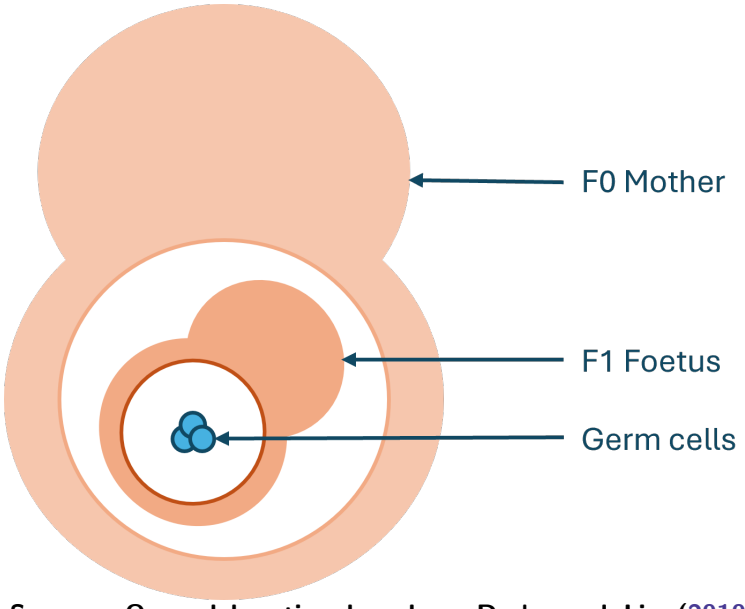
The demediated outcome is then regressed on the treatment (A_i) and pretreatment covariates (X_i), as outlined in Equation B.7:

$$\tilde{Y_i} = \beta_0 + \beta_1 A_i + \beta_2 X_i + \varepsilon_i, \tag{B.11}$$

where the least squares estimator $\hat{\beta_1}$ is the consistent estimate of the ACDE. Given this is a two-step process, standard errors on $\hat{\beta_1}$ are biased, therefore bootstrap standard errors are obtained.

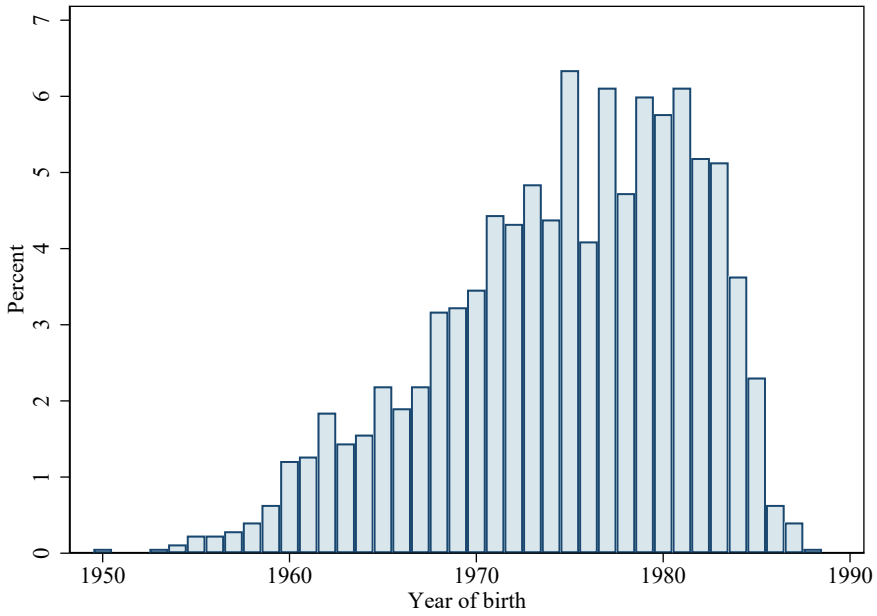
B.2 Additional Tables and Figures

Figure B.1: Multigenerational Exposure to an Environmental Shock In-Utero



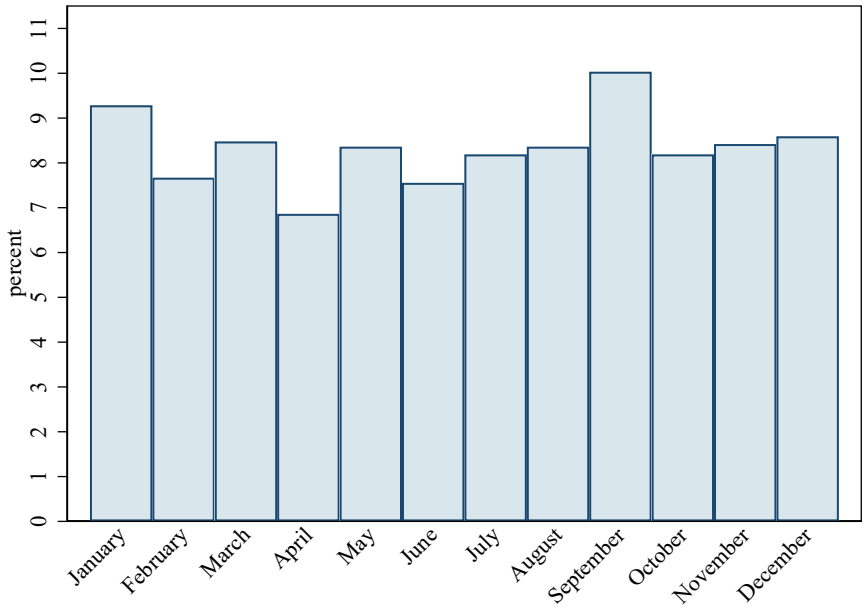
Source: Own elaboration based on Drake and Liu (2010).
Notes: An environmental insult during pregnancy to a mother (F0 generation) might affect not only the developing foetus (F1 generation) but also the germ cells which will go on to form the F2 generation.

Figure B.2: Distribution of Mother Year of Birth



Notes: Year of birth of mother based on reported age in years in R4 household roster. Bins are discrete, representing one year.

Figure B.3: Distribution of Mother Month of Birth



Notes: Month of birth of mother as reported in R4 household roster. Bins are discrete, representing one month-of-year.

Table B.1: Effect of Shock Exposure on Second Generation Outcomes: Alternative Cluster Group

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
In-utero shock	-0.079	-0.076	-0.100	-0.173	-0.090
	[0.076]	[0.040]*	[0.047]**	[0.058]***	[0.051]*
Controls	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
In-utero shock	-0.179	-0.109	-0.059	0.045	-0.021
	[0.064]**	[0.043]**	[0.058]	[0.064]	[0.059]
Controls	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors at the child cluster of birth presented in square brackets. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.2: Effects of Shock Exposure on Second Generation Outcomes: By Sex, Regression Coefficients

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
In-utero shock	-0.201	-0.103	-0.154	-0.249	-0.139
	(0.103)*	(0.088)	(0.066)**	(0.074)***	(0.069)**
Female	0.131	-0.039	0.020	-0.102	-0.259
	(0.091)	(0.048)	(0.051)	(0.061)*	(0.070)***
Shock*Female	0.242	0.054	0.108	0.150	0.097
	(0.134)*	(0.123)	(0.084)	(0.104)	(0.087)
N	1670	1657	1665	1671	1620
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
In-utero shock	-0.230	-0.167	-0.113	0.010	-0.093
	(0.083)***	(0.058)***	(0.077)	(0.063)	(0.075)
Female	0.152	-0.195	-0.113	-0.256	0.115
	(0.068)**	(0.053)***	(0.053)**	(0.053)***	(0.068)*
Shock*Female	0.102	0.117	0.107	0.069	0.143
	(0.088)	(0.081)	(0.104)	(0.116)	(0.117)
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.3: Effects of Shock Exposure on Second Generation Outcomes: By Mother Birth-Location, Regression Coefficients

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
In-utero shock	-0.072	-0.133	-0.131	-0.246	-0.135
	(0.080)	(0.051)**	(0.055)**	(0.067)***	(0.048)***
Urban-born	-0.102	-0.197	-0.161	-0.301	-0.110
	(0.206)	(0.147)	(0.137)	(0.161)*	(0.145)
Shock*Urban-born	-0.024	0.218	0.120	0.274	0.165
	(0.132)	(0.115)*	(0.125)	(0.142)*	(0.109)
Controls	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
In-utero shock	-0.213	-0.144	-0.094	0.061	-0.036
	(0.067)***	(0.054)***	(0.069)	(0.063)	(0.064)
Urban-born	-0.073	-0.165	-0.186	0.076	-0.168
	(0.156)	(0.124)	(0.141)	(0.130)	(0.119)
Shock*Urban-born	0.129	0.134	0.134	-0.058	0.059
	(0.142)	(0.130)	(0.148)	(0.094)	(0.107)
Controls	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.4: Effects of Shock Exposure on Second Generation Outcomes: By Growth Stage, Regression Coefficients

	Height-for-age		BMI-for-age	
	Age 12	Age 15	Age 12	Age 15
In-utero shock	-0.102	-0.029	0.048	-0.128
	(0.075)	(0.118)	(0.052)	(0.118)
Pubertal growth	0.639	0.152	0.396	0.092
	(0.065)***	(0.069)**	(0.063)***	(0.076)
Pubertal growth = 1 × In-utero shock	-0.142	-0.069	-0.001	0.145
	(0.112)	(0.139)	(0.105)	(0.142)
Controls	Yes	Yes	Yes	Yes
N	1665	1617	1665	1617

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.5: Effects of Shock Exposure on Second Generation Outcomes: HAZ, Alternative Specifications

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: SPEI ≤ -0.8 S.D.					
SPEI ≤ -0.8 S.D.	-0.082	-0.039	-0.086	-0.127	-0.066
	(0.074)	(0.046)	(0.048)*	(0.057)**	(0.045)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620
Panel B: Growing Season SPEI ≤ -1 S.D.					
Growing SPEI ≤ -1 S.D.	-0.126	-0.143	-0.145	-0.207	-0.156
	(0.087)	(0.078)*	(0.099)	(0.098)**	(0.088)*
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.6: Effects of Shock Exposure on Second Generation Outcomes: WAZ/BMIAZ, Alternative Specifications

	WAZ			BMIAZ	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: SPEI ≤ -0.8 S.D.					
SPEI ≤ -0.8 S.D.	-0.180	-0.091	-0.090	0.080	-0.009
	(0.072)**	(0.044)**	(0.056)	(0.053)	(0.071)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620
Panel B: Growing Season SPEI ≤ -1 S.D.					
Growing SPEI ≤ -1 S.D.	-0.126	-0.123	-0.150	-0.007	-0.049
	(0.070)*	(0.081)	(0.081)*	(0.068)	(0.085)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.7: Effects of Shock Exposure on Second Generation Outcomes: Balanced Panel Sample

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: Height-for-age					
In-utero shock	-0.067	-0.057	-0.080	-0.147	-0.077
	(0.070)	(0.045)	(0.045)*	(0.053)***	(0.046)*
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1563	1563	1563	1561	1562
	Weight-for-age			BMI-for-age	
	Age 1	Age 5	Age 8	Age 12	Age 15
Panel B: Weight-/BMI-for-age					
In-utero shock	-0.179	-0.099	-0.027	0.094	-0.003
	(0.066)***	(0.043)**	(0.052)	(0.046)**	(0.060)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1568	1568	1568	1565	1568

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.8: Effects of Shock Exposure on Second Generation HAZ: Joint Estimation with Additional Periods

	Age 1	Age 5	Age 8	Age 12	Age 15
3 years before birth	-0.002	0.016	0.027	0.012	-0.061
	(0.086)	(0.080)	(0.072)	(0.077)	(0.060)
2 years before birth	0.066	0.055	0.033	0.028	0.037
	(0.057)	(0.059)	(0.058)	(0.066)	(0.065)
In-utero shock	-0.067	-0.072	-0.095	-0.161	-0.088
	(0.062)	(0.048)	(0.049)*	(0.055)***	(0.046)*
1 years after birth	0.097	0.080	0.094	0.103	0.035
	(0.081)	(0.077)	(0.062)	(0.066)	(0.079)
2 years after birth	-0.093	-0.025	-0.064	-0.048	-0.038
	(0.081)	(0.048)	(0.052)	(0.051)	(0.043)
3 years after birth	-0.012	-0.021	-0.015	0.019	-0.008
	(0.087)	(0.067)	(0.057)	(0.059)	(0.061)
4 years after birth	-0.011	0.018	-0.040	0.071	0.014
	(0.068)	(0.056)	(0.054)	(0.065)	(0.065)
5 years after birth	0.016	-0.064	-0.019	-0.002	-0.034
	(0.056)	(0.055)	(0.048)	(0.059)	(0.046)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.9: Effects of Shock Exposure on Second Generation WAZ/BMIAZ: Joint Estimation with Additional Periods

	Age 1	Age 5	Age 8	Age 12	Age 15
3 years before birth	0.012	-0.028	0.028	0.058	0.099
	(0.067)	(0.067)	(0.081)	(0.053)	(0.056)*
2 years before birth	0.022	0.088	0.121	0.080	0.152
	(0.083)	(0.061)	(0.081)	(0.077)	(0.056)***
In-utero shock	-0.181	-0.089	-0.039	0.053	0.005
	(0.066)***	(0.050)*	(0.063)	(0.052)	(0.063)
1 years after birth	-0.003	0.071	0.095	-0.022	0.049
	(0.062)	(0.069)	(0.073)	(0.052)	(0.060)
2 years after birth	-0.062	-0.070	-0.012	-0.025	-0.037
	(0.060)	(0.046)	(0.051)	(0.059)	(0.066)
3 years after birth	0.037	0.026	0.029	0.044	0.027
	(0.087)	(0.056)	(0.054)	(0.062)	(0.063)
4 years after birth	0.100	0.049	-0.000	0.052	0.059
	(0.088)	(0.060)	(0.066)	(0.073)	(0.082)
5 years after birth	-0.042	0.050	0.008	0.025	0.081
	(0.050)	(0.041)	(0.049)	(0.052)	(0.044)*
Controls & FEs	Yes	Yes	Yes	Yes	Yes
N	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.10: Effects of Shock Exposure on Second Generation HAZ: Separate Estimation with Additional Periods

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: 3 years before birth					
SPEI ≤ -1 S.D.	-0.010	0.007	0.020	-0.005	-0.067
	(0.082)	(0.074)	(0.071)	(0.075)	(0.055)
Panel B: 2 years before birth					
SPEI ≤ -1 S.D.	0.063	0.057	0.036	0.027	0.043
	(0.059)	(0.059)	(0.061)	(0.068)	(0.065)
Panel C: In-utero					
In-utero shock	-0.079	-0.076	-0.100	-0.173	-0.090
	(0.064)	(0.045)*	(0.048)**	(0.055)***	(0.045)**
Panel D: 1 year after birth					
SPEI ≤ -1 S.D.	0.110	0.086	0.105	0.121	0.050
	(0.078)	(0.073)	(0.062)*	(0.063)*	(0.079)
Panel E: 2 years after birth					
SPEI ≤ -1 S.D.	-0.093	-0.023	-0.061	-0.056	-0.034
	(0.075)	(0.050)	(0.054)	(0.049)	(0.042)
Panel F: 3 years after birth					
SPEI ≤ -1 S.D.	-0.002	-0.020	-0.007	0.016	0.004
	(0.079)	(0.059)	(0.056)	(0.060)	(0.058)
Panel G: 4 years after birth					
SPEI ≤ -1 S.D.	-0.008	0.027	-0.032	0.077	0.023
	(0.074)	(0.053)	(0.056)	(0.064)	(0.063)
Panel H: 5 years after birth					
SPEI ≤ -1 S.D.	0.027	-0.055	-0.003	0.009	-0.025
	(0.056)	(0.053)	(0.047)	(0.057)	(0.044)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
Observations	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.11: Effects of Shock Exposure on Second Generation WAZ/BMIAZ: Separate Estimation with Additional Periods

	Age 1	Age 5	Age 8	Age 12	Age 15
Panel A: 3 years before birth					
SPEI ≤ -1 S.D.	0.001	-0.043	0.010	0.049	0.082
	(0.058)	(0.064)	(0.083)	(0.055)	(0.056)
Panel B: 2 years before birth					
SPEI ≤ -1 S.D.	0.023	0.084	0.119	0.066	0.135
	(0.077)	(0.064)	(0.081)	(0.074)	(0.055)**
Panel C: In-utero					
In-utero shock	-0.179	-0.109	-0.059	0.045	-0.021
	(0.063)***	(0.047)**	(0.056)	(0.049)	(0.059)
Panel D: 1 year after birth					
SPEI ≤ -1 S.D.	0.017	0.088	0.094	-0.033	0.041
	(0.058)	(0.065)	(0.072)	(0.050)	(0.064)
Panel E: 2 years after birth					
SPEI ≤ -1 S.D.	-0.066	-0.074	-0.015	-0.031	-0.041
	(0.055)	(0.045)	(0.053)	(0.056)	(0.061)
Panel F: 3 years after birth					
SPEI ≤ -1 S.D.	0.041	0.029	0.023	0.034	0.008
	(0.078)	(0.053)	(0.056)	(0.062)	(0.062)
Panel G: 4 years after birth					
SPEI ≤ -1 S.D.	0.108	0.043	-0.010	0.037	0.037
	(0.085)	(0.061)	(0.068)	(0.071)	(0.077)
Panel H: 5 years after birth					
SPEI ≤ -1 S.D.	-0.036	0.053	0.010	0.010	0.073
	(0.047)	(0.039)	(0.047)	(0.048)	(0.045)
Controls & FEs	Yes	Yes	Yes	Yes	Yes
Observations	1670	1657	1665	1671	1620

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are reported in standard deviations from the age- and gender-specific mean value. Cluster robust standard errors in parentheses. Controls include an indicator of if the child is female. Fixed effects for child cluster-of-residence, year-month birth cohort, and mother year- and province-of-birth are suppressed.

Table B.12: Effect of Zero Generation Socioeconomic Status on Probability of Shock Exposure

	In-utero shock	In-utero shock
Grandmother speaks Spanish	-0.072	
	(0.046)	
Grandparent completed secondary		0.044
		(0.063)
Controls	No	Yes
N	1670	522

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Cluster robust standard errors in parentheses. Controls include if the grandparent reporting educational attainment if female, and their age in years. Fixed effects for mother year- and province-of-birth are suppressed.

Table B.13: Effect of Shock Exposure on Zero Generation Migration/Fertility Choices

	Migration				Fertility	
	Before age 5	Ever migrate	Rural-urban	Lima/ Callao	Mother birth month	Grand-child gender
In-utero shock	-0.004	0.013	0.021	-0.009	-0.410	0.028
	(0.008)	(0.028)	(0.021)	(0.016)	(0.277)	(0.032)
N	1632	1632	1632	1632	1670	1670

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Cluster robust standard errors in parentheses. Fixed effects for mother year- and province-of-birth are suppressed.

Table B.14: Socieoeconomic Status Controls and Parent Investments at Each Survey Round

Age	Intermediate confounders
1	HH size, family own house, attended antenatal classes, attended birth, hospital birth, parenting skills index
5	HH size, family own house, p/c food expenditure, child food diversity/frequency, HH food security, health and education expenditure, preschool
8	HH size, family own house, p/c food expenditure, child food diversity/frequency, HH food security, health and education expenditure, # of books in HH, caregiver involvement index
12	HH size, family own house, p/c food expenditure, child food diversity/frequency, HH food security, health and education expenditure, # of books in HH, caregiver involvement index
15	HH size, family own house, p/c food expenditure, child food diversity/frequency, HH food security, health and education expenditure, # of books in HH

Notes: The list of intermediate variables varies over rounds as not all questions are asked in each round. *Abbreviations:* HH - household; p/c - per capita.

Appendix C

Chapter 3 Appendix

C.1 Criteria for EFA Factor Selection

The number of latent factors to extract is assessed using several criteria, following Osborne et al. (2014):

1. **Theory:** the literature on core self-evaluations points to a single highly internally consistent construct. Therefore this provides an a priori assumption about the number of factors to extract, however this may not always be supported by EFA results.

2. **Kaiser criterion:** Kaiser (1960, 1970) suggests a rule of thumb of any eigenvalues greater than 1, as a theoretical lower bound for a true component in a principle components analysis (PCA) with an infinite sample size (Guttman, 1954). However this is often an inaccurate method, particularly as the number of items analysed increases (Costello & Osborne, 2005). Similar to Webb (2024), I also consider a less conservative 0.7 threshold.

3. **Screeplot:** Graphical assessment of the eigenvalue scree plot for evident 'elbows' in the plot, where an obvious change of slope occurs, with the number of points prior to the elbow considered a good estimate. This is not considered sufficient alone for determining the number of factors to extract.

4. **Parallel Analysis:** Observing that the eigenvalues from PCA would be greater than one in a finite sample due to sample-error and least squares bias, Horn (1965) suggests adjusting the eigenvalues of each factor by subtracting the mean sample error from many iterations of uncorrelated data sets, retaining components with adjusted eigenvalues greater than one (Dinno, 2009). Therefore using a Monte-Carlo procedure I simulate uncorrelated data of the same dimension as my sample with 5,000 replications, keeping eigenvalues greater than the 95^{th} percentile value of simulated eigenvalues.

5. **Minimum Average Partial criterion:** In the context of PCA, Velicer (1976) proposes partialling out the shared variance as each component is created sequentially, to the point at which common variance is at a minimum. The number of components for which a minimum is reached represents the number to extract.

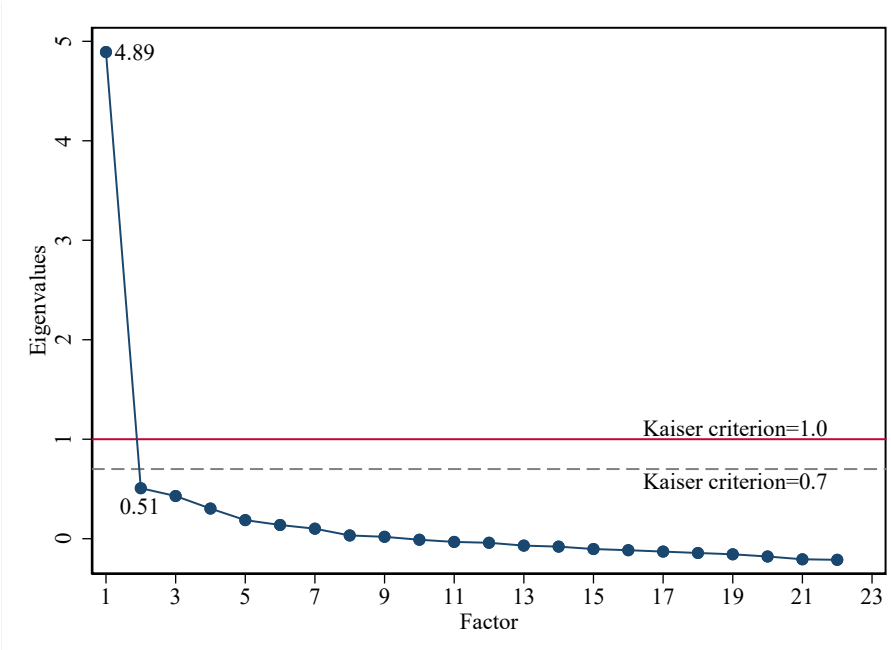
A summary of the number of factors to extract is given by Table C.1. The criteria described in 2 and 3 are shown by the scree plot of eigenvalues, Figure C.1.

Table C.1: Number of Factors to Extract, by Method

Method	# of Factors
Kaiser criterion > 1	1
Kaiser criterion > 0.7	1
Screeplot 'elbow'	1
Parallel analysis	3
Minimum average partial	1
Extracted	1

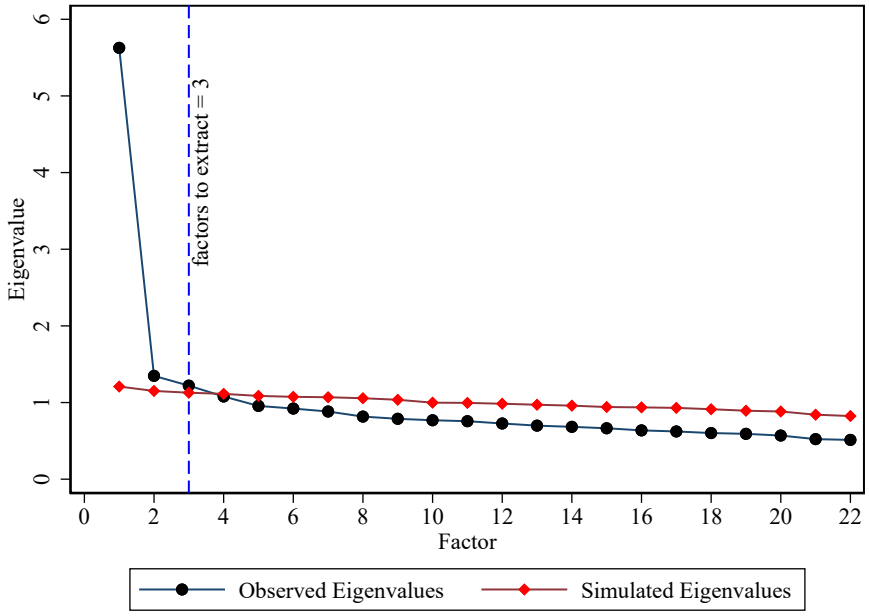
The 1st factor eigenvalue is 4.89, and explains 95.3% of the shared variance in the latent factor model. All other factors displayed an eigenvalue significantly below the threshold of 1 (and the more conservative 0.7 cut off), with the 2nd factor eigenvalue of 0.51. There is an evident change of slope at the second factor, suggesting an 'elbow' above which one factor lies. Although there are other changes in slopes between further factors, this is minimal in comparison to the drastic change in slope at the identified elbow.

Figure C.1: Screeplot of Eigenvalues from EFA Latent Factor Model



Observed eigenvalues from principle components analysis (unadjusted) of the latent factor model are plotted in Figure C.2, with the 95th percentile of simulated eigenvalues from 5000 replications plotted in red. Three eigenvalues lie above the simulated eigenvalues, suggesting, in contrast with all other criteria, a three factor model. However there is a clear distinction of the 1st factor, while factors 2 and 3 lie marginally above their relevant threshold. As discussed by Osborne et al. (2014), with large sample sizes parallel analysis may not prove to be as useful as other criteria, with only small deviations from 1 estimated over many iterations. Finally, Figure C.3, provides a graphical plot of the average partial correlations for the factor partialled out. Evidence suggests that the average partial is minimised when the 1st factor is partialled out.

Figure C.2: Horn's Parallel Analysis



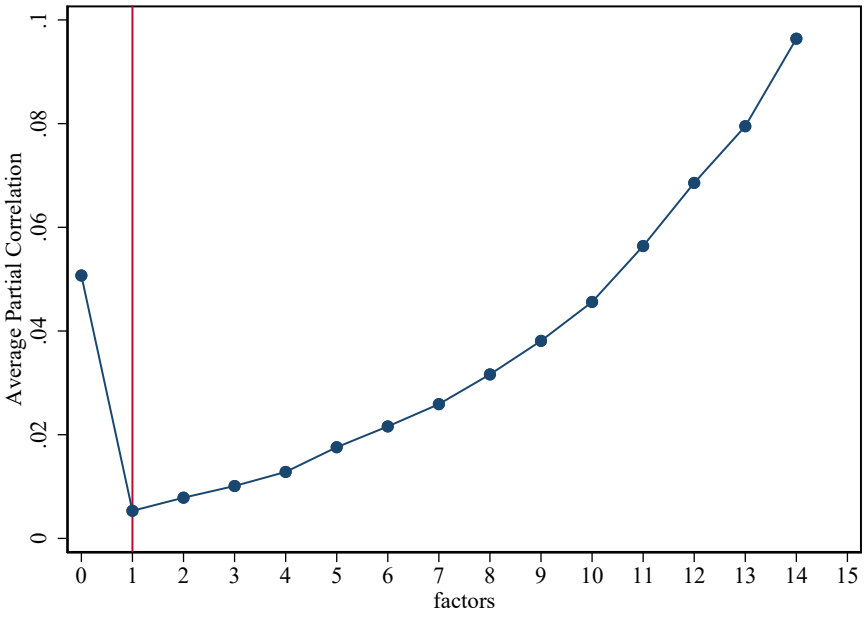
Overall the majority of criteria are aligned with the a priori assumption of a one factor model, therefore no other factors were retained. The factor loadings of each item on the 1st factor are shown in Table C.2, alongside the share of item unique variance, Ψ. Following Attanasio et al. (2020) and Webb (2024), I discount low factor loadings below a threshold of 0.3 in constructing the 1st factor score (A lower cut off of <0.25, as used by Krutikova and Lilleør (2015) does not alter results). In total 19 of 22 items load above the cut off within the range of 0.420 and 0.576. A factor score is constructed as a loading-weighted mean of these items. Finally, the factor score is standardised as a z-score with mean 0 and standard deviation 1.

Table C.2: 1st Factor Loadings for CSE

	Loading	Ψ
YL – Agency		
If I try hard, I can improve my situation in life.	0.457	0.791
I like to make plans for my future studies and work.	0.441	0.805
I have no choice about the work I do - I must do this sort of work.	-0.033	0.999
Other people in my family make all the decisions about how I spend my time.	0.000	1.000
Self-efficacy		
I can always manage to solve difficult problems if I try hard enough.	0.573	0.671
If someone opposes me, I can find the means and ways to get what I want.	0.221	0.951
It is easy for me to stick to my aims and accomplish my goals.	0.518	0.732
I am confident that I could deal efficiently with unexpected events.	0.422	0.822
Thanks to my resourcefulness, I know how to handle unforeseen situations.	0.579	0.664
I can solve most problems if I invest the necessary effort.	0.594	0.647
I can remain calm when facing difficulties because I can rely on my coping abilities.	0.569	0.677
When I am confronted with a problem, I can usually find several solutions.	0.494	0.756
If I am in trouble, I can usually think of a solution.	0.518	0.732
I can usually handle whatever comes my way.	0.492	0.758
SDQ – Self-esteem		
I do lots of important things.	0.479	0.771
In general, I like being the way I am.	0.537	0.712
Overall, I have a lot to be proud of.	0.497	0.753
I can do things as well as most people.	0.510	0.740
Other people think I am a good person.	0.417	0.826
A lot of things about me are good.	0.529	0.720
I'm as good as most other people.	0.424	0.820
When I do something, I do it well.	0.489	0.761

Notes: Factor loadings ≥ 0.3 are highlighted in green. Ψ is the share of item unique variance.

Figure C.3: Mini Average Partial Correlation Analysis



C.2 Additional Tables and Figures

Figure C.4: Multi-Density Plot of Community-Level Gamma-Fitted SPI Values, by Month

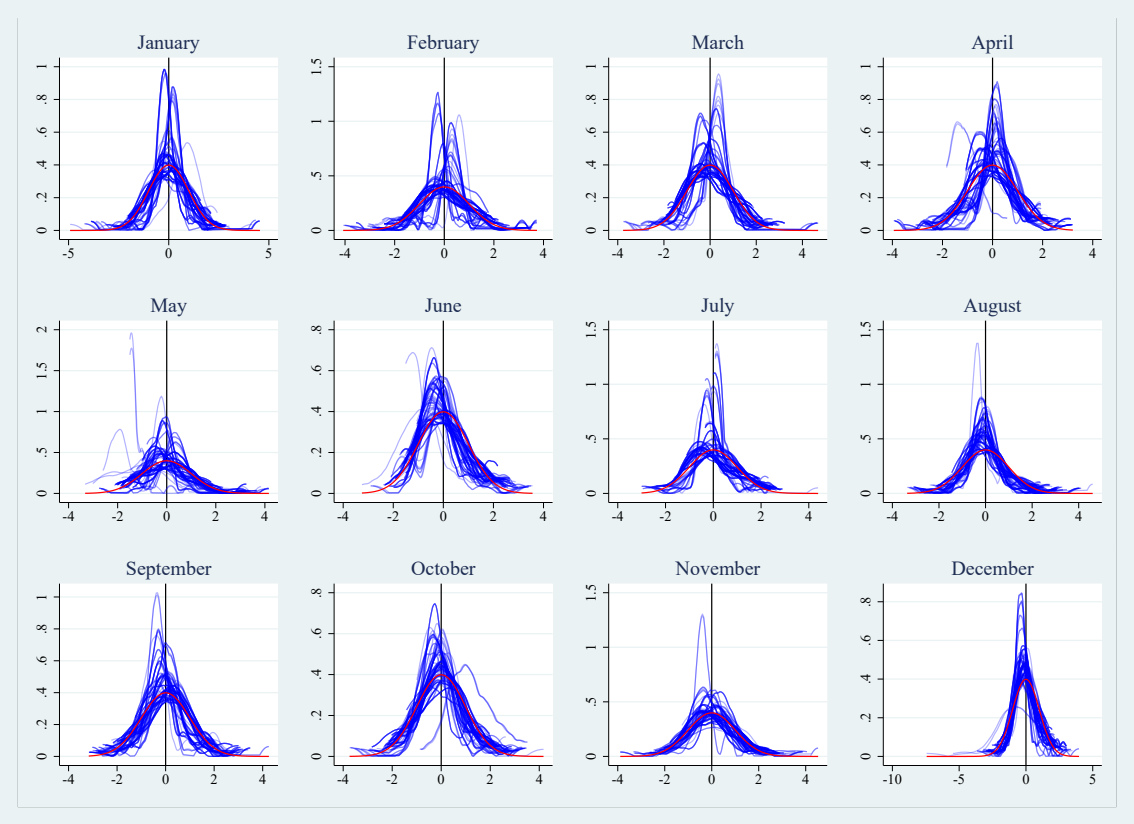


Figure C.5: Multi-Density Plot of Community-Level Lognormal-Fitted SPI Values, by Month

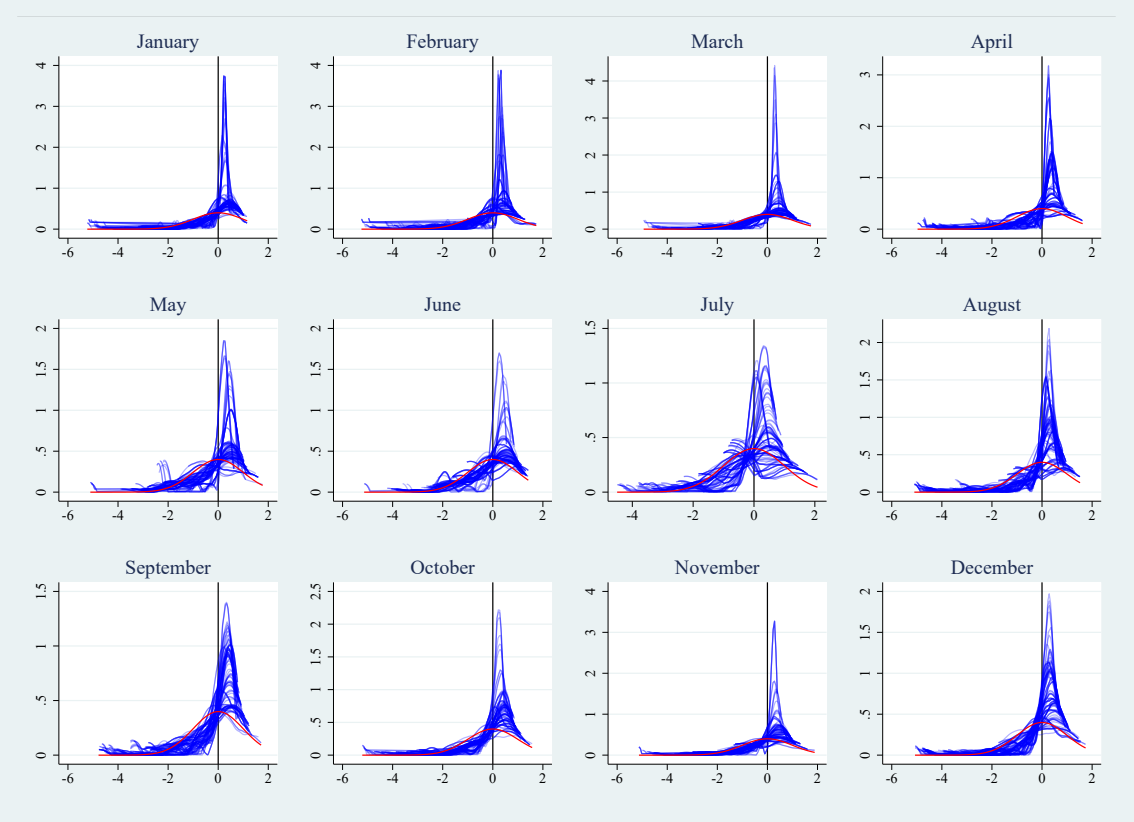


Figure C.6: Primary Crop (Hectares Sown) in 2010, by Department

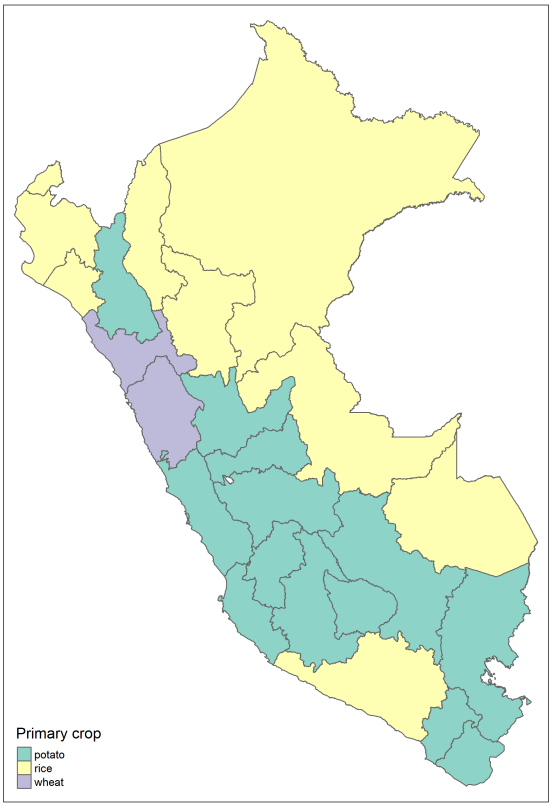


Table C.3: Impact of (±)1.5 S.D Shocks on Constituent Scale Naïve Scores

	Agency	Self-esteem	Self-efficacy
Positive shock			
Prenatal	0.027	0.073	0.024
	(0.365)	(0.047)**	(0.511)
	[0.399]	[0.103]	[0.538]
	{0.339}	{0.010}***	{0.501}
1st year	0.018	0.004	0.033
	(0.680)	(0.919)	(0.352)
	[0.710]	[0.923]	[0.364]
	{0.622}	{0.850}	{0.288}
2nd year	-0.049	-0.094	-0.065
	(0.141)	(0.014)**	(0.065)*
	[0.146]	[0.020]**	[0.080]*
	{0.077}*	{0.009}***	{0.047}**
3rd year	-0.046	-0.087	-0.084
	(0.224)	(0.008)***	(0.006)***
	[0.272]	[0.015]**	[0.018]**
	{0.209}	{0.008}***	{0.002}***
Negative shock			
Prenatal	-0.051	-0.047	0.002
	(0.238)	(0.249)	(0.972)
	[0.350]	[0.408]	[0.975]
	{0.213}	{0.220}	{0.971}
1st year	-0.063	0.013	0.094
	(0.238)	(0.771)	(0.102)
	[0.267]	[0.787]	[0.131]
	{0.227}	{0.786}	{0.078}*
2nd year	0.086	-0.097	-0.074
	(0.126)	(0.186)	(0.298)
	[0.200]	[0.268]	[0.389]
	{0.114}	{0.293}	{0.273}
3rd year	0.058	-0.108	-0.083
	(0.379)	(0.044)**	(0.200)
	[0.468]	[0.098]*	[0.271]
	{0.348}	{0.018}**	{0.172}
Controls	Yes	Yes	Yes
N	2089	2089	2089

Notes: p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.4: Heterogeneous Effects of (±)1.5 S.D shocks on CSE Scores

	Female	Poorest	Mother's education	Agricultural
Level term	0.087	-0.193	0.226	-
	(0.469)	(0.299)	(0.004)***	(-)
	{0.366}	{0.320}	{0.001}***	
Positive shock				
Prenatal	0.024	0.046	0.073	0.038
	(0.615)	(0.185)	(0.172)	(0.396)
	{0.581}	{0.156}	{0.151}	{0.331}
**Interaction*	0.100	0.173	-0.004	0.086
	(0.070)*	(0.017)**	(0.948)	(0.127)
	{0.043}**	{0.012}**	{0.948}	{0.081}*
1st year	0.057	0.062	0.081	0.078
	(0.315)	(0.133)	(0.237)	(0.243)
	{0.222}	{0.074}*	{0.190}	{0.104}
**Interaction*	-0.029	-0.040	-0.059	-0.041
	(0.734)	(0.764)	(0.416)	(0.624)
	{0.707}	{0.764}	{0.395}	{0.551}
2nd year	-0.106	-0.081	-0.055	-0.125
	(0.004)***	(0.013)**	(0.219)	(0.101)
	{0.002}***	{0.010}**	{0.203}	{0.061}*
**Interaction*	0.032	-0.046	-0.055	0.034
	(0.557)	(0.415)	(0.456)	(0.706)
	{0.555}	{0.412}	{0.458}	{0.667}
3rd year	-0.100	-0.142	-0.116	-0.114
	(0.030)**	(0.000)***	(0.004)***	(0.005)***
	{0.032}**	{0.000}***	{0.001}***	{0.000}***
**Interaction*	-0.013	0.133	0.010	0.038
	(0.796)	(0.017)**	(0.866)	(0.550)
	{0.809}	{0.005}***	{0.867}	{0.505}
Negative shock				
Prenatal	0.029	-0.051	0.019	-0.022
	(0.574)	(0.169)	(0.828)	(0.578)
	{0.566}	{0.173}	{0.804}	{0.567}
**Interaction*	-0.122	0.153	-0.071	0.006
	(0.081)*	(0.171)	(0.510)	(0.952)
	{0.093}*	{0.123}	{0.473}	{0.945}
1st year	0.082	0.071	0.066	-0.062
	(0.209)	(0.239)	(0.280)	(0.592)
	{0.157}	{0.221}	{0.263}	{0.549}
**Interaction*	-0.039	0.015	-0.004	0.143
	(0.605)	(0.848)	(0.950)	(0.294)
	{0.573}	{0.849}	{0.951}	{0.230}
2nd year	0.009	-0.046	0.020	0.085
	(0.938)	(0.470)	(0.746)	(0.376)

	{0.934}	{0.468}	{0.730}	{0.295}
**Interaction*	-0.122	-0.041	-0.123	-0.190
	(0.234)	(0.690)	(0.084)*	(0.102)
	{0.189}	{0.670}	{0.065}*	{0.091}*
3rd year	-0.071	-0.080	-0.062	-0.147
	(0.351)	(0.246)	(0.323)	(0.061)*
	{0.355}	{0.232}	{0.322}	{0.049}**
**Interaction*	-0.035	-0.037	-0.042	0.120
	(0.715)	(0.769)	(0.665)	(0.183)
	{0.660}	{0.765}	{0.658}	{0.175}
Controls	Yes	Yes	Yes	Yes
N	2089	2089	2089	2089

Notes: Extension of Table 3.4. * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.5: Impact of Lognormal SPI (±)1.5 S.D Shocks on CSE Scores

	EFA 1st factor	% Exposed
Positive shock		
Prenatal	0.220	7.6
	(0.004)***	
	[0.034]**	
	{0.001}***	
1st year	0.123	7.7
	(0.406)	
	[0.457]	
	{0.284}	
2nd year	-0.077	3.4
	(0.488)	
	[0.489]	
	{0.434}	
3rd year	-0.356	0.8
	(0.015)**	
	[0.354]	
	{0.008}***	
Negative shock		
Prenatal	-0.070	32.3
	(0.083)*	
	[0.236]	
	{0.061}*	
1st year	-0.027	43.6
	(0.577)	
	[0.603]	
	{0.553}	
2nd year	-0.059	49.0
	(0.178)	
	[0.243]	
	{0.167}	
3rd year	0.053	47.7
	(0.213)	
	[0.233]	
	{0.148}	
Controls	Yes	
N	2089	

Notes: p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed. % Exposed refers to the share of sample exposed to at least 1 monthly shock in each period between conception and 3rd Birthday.

Table C.6: Impact of (±)1.5 S.D. Shocks on CSE Scores, by Period of Exposure

	Prenatal	1st year	2nd year	3rd year
Positive shock	0.049	-0.000	-0.066	-0.067
	(0.141)	(0.988)	(0.033)**	(0.033)**
	[0.185]	[0.989]	[0.044]**	[0.044]**
	{0.134}	{0.987}	{0.029}**	{0.020}**
Negative shock	-0.055	0.053	-0.069	-0.048
	(0.232)	(0.332)	(0.300)	(0.224)
	[0.272]	[0.377]	[0.335]	[0.237]
	{0.222}	{0.289}	{0.283}	{0.220}
Controls	Yes	Yes	Yes	Yes
N	2089	2089	2089	2089

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. P-values based on cluster robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Cumulative shocks refers to the total number of periods a respondent experience at least one of that shock type. Controls include child gender and if Spanish is their mother tongue; mothers age and if mother completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.7: Impact of (±)1.5 S.D. shocks on EFA CSE Scores, Robustness Checks

	Cumulative	**Other periods**	**Shock type**	
			Positive	Negative
Positive shock				
-2nd Year		0.054		
		(0.299)		
		[0.315]		
		{0.252}		
Prenatal	0.081	0.114	0.077	
	(0.072)*	(0.018)**	(0.023)**	
	[0.079]*	[0.034]**	[0.051]*	
	{0.049}**	{0.006}***	{0.011}**	
1st Year	0.059	0.060	0.054	
	(0.129)	(0.214)	(0.099)*	
	[0.127]	[0.245]	[0.101]	
	{0.059}*	{0.166}	{0.059}*	
2nd Year	-0.071	-0.095	-0.090	
	(0.076)*	(0.017)**	(0.009)***	
	[0.118]	[0.029]**	[0.010]**	
	{0.075}*	{0.007}***	{0.008}***	
3rd Year	-0.091	-0.136	-0.099	
	(0.011)**	(0.004)***	(0.003)***	
	[0.033]**	[0.009]***	[0.005]***	
	{0.007}***	{0.001}***	{0.001}***	
4th Year		0.000		
		(0.996)		
		[0.996]		
		{0.995}		
Negative shock				
-2nd Year		-0.021		
		(0.633)		
		[0.687]		
		{0.619}		
Prenatal	-0.069	-0.056		-0.055
	(0.307)	(0.266)		(0.215)
	[0.417]	[0.367]		[0.278]
	{0.215}	{0.252}		{0.210}
1st Year	0.034	0.085		0.050
	(0.596)	(0.141)		(0.344)
	[0.652]	[0.139]		[0.359]
	{0.579}	{0.114}		{0.301}
2nd Year	-0.097	-0.031		-0.037
	(0.493)	(0.683)		(0.575)
	[0.662]	[0.724]		[0.593]
	{0.467}	{0.685}		{0.559}
3rd Year	-0.132	-0.038		-0.076

	(0.243)	(0.549)		(0.098)*
	[0.367]	[0.580]		[0.122]
	{0.210}	{0.522}		{0.113}
4th Year		0.102		
		(0.181)		
		[0.215]		
		{0.077}*		
Cumulative shocks				
Positive	-0.032			
	(0.469)			
	[0.508]			
	{0.431}			
Negative	0.071			
	(0.566)			
	[0.632]			
	{0.532}			
Controls	Yes	Yes	Yes	Yes
N	2089	1675	2089	2089

* p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on cluster robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Cumulative shocks refers to the total number of periods a respondent experience at least one of that shock type. Controls include child gender and if Spanish is their mother tongue; mothers age and if mother completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.8: Impact of (±)1.5 S.D. Shocks on CSE Scores, Additional Specifications

	Avgerage temperature	Growing season	Exlcude Lima	YC fixed effect
Positive shocks				
Prenatal	0.068	0.040	0.081	0.068
	(0.037)**	(0.442)	(0.024)**	(0.036)**
	[0.053]*	[0.453]	[0.068]*	[0.058]*
	{0.022}**	{0.416}	{0.015}**	{0.019}**
1st Year	0.046	0.035	0.048	0.043
	(0.169)	(0.419)	(0.164)	(0.175)
	[0.172]	[0.428]	[0.134]	[0.159]
	{0.102}	{0.379}	{0.150}	{0.124}
2nd Year	-0.087	-0.051	-0.109	-0.090
	(0.011)**	(0.307)	(0.005)***	(0.007)***
	[0.016]**	[0.318]	[0.004]***	[0.006]***
	{0.008}***	{0.312}	{0.002}***	{0.005}***
3rd Year	-0.104	-0.094	-0.113	-0.105
	(0.003)***	(0.032)**	(0.002)***	(0.001)***
	[0.019]**	[0.044]**	[0.011]**	[0.003]***
	{0.002}***	{0.029}**	{0.001}***	{0.000}***
Negative shocks				
Prenatal	-0.027	-0.032	-0.039	-0.030
	(0.467)	(0.525)	(0.374)	(0.434)
	[0.469]	[0.564]	[0.391]	[0.433]
	{0.452}	{0.485}	{0.358}	{0.424}
1st Year	0.067	0.081	0.088	0.066
	(0.168)	(0.250)	(0.063)*	(0.170)
	[0.192]	[0.317]	[0.085]*	[0.177]
	{0.150}	{0.200}	{0.051}*	{0.149}
2nd Year	-0.051	-0.040	-0.058	-0.056
	(0.541)	(0.595)	(0.456)	(0.478)
	[0.639]	[0.612]	[0.550]	[0.559]
	{0.531}	{0.541}	{0.446}	{0.470}
3rd Year	-0.097	-0.077	-0.103	-0.084
	(0.064)*	(0.309)	(0.055)*	(0.120)
	[0.110]	[0.407]	[0.097]*	[0.171]
	{0.070}*	{0.326}	{0.049}**	{0.119}
Controls	Yes	Yes	Yes	Yes
N	2089	2089	1754	2089

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on cluster robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and if Spanish is their mother tongue; mothers age and if mother completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.9: Impact of (±)1.5 S.D. Shocks on Migration and Sex of Child

	Female	Migration
Positive shocks		
Prenatal	-0.002	0.007
	(0.929)	(0.561)
	[0.938]	[0.558]
	{0.934}	{0.546}
1st Year	-0.016	0.014
	(0.498)	(0.414)
	[0.571]	[0.462]
	{0.466}	{0.383}
2nd Year	0.008	0.013
	(0.683)	(0.294)
	[0.699]	[0.339]
	{0.709}	{0.271}
3rd Year	0.008	0.016
	(0.741)	(0.290)
	[0.771]	[0.348]
	{0.729}	{0.237}
Negative shocks		
Prenatal	0.002	-0.003
	(0.945)	(0.905)
	[0.943]	[0.934]
	{0.946}	{0.903}
1st Year	0.002	-0.026
	(0.943)	(0.174)
	[0.941]	[0.199]
	{0.940}	{0.125}
2nd Year	-0.007	0.010
	(0.863)	(0.653)
	[0.891]	[0.674]
	{0.856}	{0.624}
3rd Year	0.021	-0.015
	(0.418)	(0.300)
	[0.456]	[0.318]
	{0.431}	{0.241}
Controls	Yes	Yes
N	2089	2089

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on cluster robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and if Spanish is their mother tongue; mothers age and if mother completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.10: Impact of Prenatal (±)1.5 S.D Shocks on CSE Scores, by Trimester

	EFA 1st Factor	% Exposure	Mean exposure
Positive shock			
1st trimester	-0.017	23.5	0.25
	(0.711)		
	[0.705]		
	{0.722}		
2nd trimester	0.079	20.2	0.21
	(0.235)		
	[0.269]		
	{0.196}		
3rd trimester	0.115	19.2	0.20
	(0.052)*		
	[0.079]*		
	{0.048}**		
Negative shock			
1st trimester	0.025	9.0	0.10
	(0.772)		
	[0.801]		
	{0.767}		
2nd trimester	-0.065	6.9	0.07
	(0.700)		
	[0.743]		
	{0.689}		
3rd trimester	-0.133	7.4	0.08
	(0.029)**		
	[0.040]**		
	{0.025}**		
Controls	Yes		
N	2089		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. P-values based on cluster robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed. % Exposure is the share of sample exposed to at least 1 monthly shock in each trimester. Mean exposure captures the mean number of months of exposure experienced.

Table C.11: Impact of (±)1.5 S.D Shocks on CSE Scores, Adjusted Q-values

	EFA 1st Factor		**Naive z-score**	
	Full	In-comm.	Full	In-comm.
Positive shock				
Prenatal	0.068	0.096	0.052	0.081
	[0.085]*	[0.049]**	[0.203]	[0.085]*
1st year	0.043	0.051	0.027	0.043
	[0.277]	[0.345]	[0.440]	[0.422]
2nd year	-0.090	-0.093	-0.091	-0.095
	[0.044]**	[0.049]**	[0.044]**	[0.049]**
3rd year	-0.105	-0.129	-0.097	-0.115
	[0.030]**	[0.038]**	[0.038]**	[0.044]**
Negative shock				
Prenatal	-0.030	-0.062	-0.036	-0.073
	[0.440]	[0.286]	[0.422]	[0.277]
1st year	0.066	0.075	0.036	0.043
	[0.277]	[0.345]	[0.440]	[0.464]
2nd year	-0.056	-0.048	-0.056	-0.033
	[0.457]	[0.464]	[0.451]	[0.494]
3rd year	-0.084	-0.069	-0.071	-0.045
	[0.242]	[0.345]	[0.277]	[0.440]
Controls	Yes	Yes	Yes	Yes
N	2089	1675	2089	1675

Notes: * q < 0.10, ** q < 0.05, *** q < 0.01. Sharpened q-values provided in "[.]" brackets. Full sample refers to children geolocated in round 1. In-community restricts sample to those whose mother lived in the same community from conception until round 2. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.12: Impact of +1.5 S.D Shocks on Health and Nutrition Mechanisms

	Stunting	Good health	Serious illness	LR health
Prenatal	0.003	-0.002	0.010	0.004
	(0.841)	(0.913)	(0.502)	(0.737)
	[0.844]	[0.919]	[0.541]	[0.737]
	{0.844}	{0.918}	{0.444}	{0.680}
1st year	0.007	0.007	-0.003	0.004
	(0.627)	(0.728)	(0.849)	(0.706)
	[0.650]	[0.741]	[0.865]	[0.713]
	{0.608}	{0.691}	{0.838}	{0.677}
2nd year	-0.001	0.036	-0.024	0.001
	(0.965)	(0.108)	(0.150)	(0.964)
	[0.966]	[0.147]	[0.190]	[0.969]
	{0.962}	{0.086}*	{0.109}	{0.958}
3rd year	-0.009	-0.001	0.014	-0.011
	(0.484)	(0.968)	(0.332)	(0.174)
	[0.483]	[0.973]	[0.377]	[0.166]
	{0.410}	{0.967}	{0.291}	{0.164}
Controls	Yes	Yes	Yes	Yes
N	2072	2089	2089	2085

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.13: Impact of +1.5S.D. Shocks on Caregiver Stress and Parenting Practices

	Stress (SRQ20)			**Practices**
	Total score	Score=>7	Score=>8	z-score
Positive shock	0.003	-0.008	-0.000	-0.033
	(0.979)	(0.419)	(0.978)	(0.461)
	[0.981]	[0.485]	[0.982]	[0.520]
	{0.729}	{0.842}	{0.713}	{0.482}
Controls	Yes	Yes	Yes	Yes
N	7044	7044	7044	1503

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include child gender and indicator for if they speak Spanish as their mother tongue; mothers age and indicator for if they completed primary; household wealth index (R1) and if in a rural location. Fixed effects for community and month of birth cohort are suppressed.

Table C.14: Parenting Practices as Reported in Round 1

Good practices
Carry him/her (on front or on back).
Soothe him/her, sing to him/her.
Rock him/her, walk around with child in arms.
Give him/her water to calm him/her.
Breast or bottle feed him/her.
Swaddle him/her in blanket, tightly so he/she is quiet.
Bad practices
Smack him/her.
Shake him/her.
Threaten him/her.
Pinch him/her, squeeze him/her tightly.
Put him/her face down on bed so he/she cries into mattress.
Nothing - let him/her cry until he/she falls asleep.

Notes: Each category is coded as 1 if the caregiver reports a good practice, and -1 if they report a bad practice, as a response to their infant child crying.

Table C.15: Impact of +1.5 S.D Shocks in Previous Year on Adult Hours Worked

	Parents		All HH adults	
	(1)	(2)	(3)	(4)
Panel A: Main activity				
Female	-1.808	-1.837	-1.541	-1.548
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
	{0.000}***	{0.000}***	{0.000}***	{0.000}***
Positive Shock	0.190	0.208	0.189	0.202
	(0.014)**	(0.038)**	(0.003)***	(0.010)**
	{0.005}***	{0.025}**	{0.001}***	{0.007}***
**interaction*		-0.037		-0.031
		(0.746)		(0.729)
		{0.749}		{0.723}
$H_0 : \beta_2 + \beta_3 = 0$ p-val.		0.057		0.027
N	5324	5324	7341	7341
Panel B: All paid activity				
Female	-4.384	-3.534	-3.479	-2.793
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
	***	{0.000}***	{0.000}***	{0.000}***
Positive Shock	0.432	0.917	0.356	0.743
	(0.001)***	(0.000)***	(0.000)***	(0.000)***
	***	{0.000}***	{0.000}***	{0.000}***
**interaction*		-1.160		-0.983
		(0.002)***		(0.002)***
		{0.003}***		{0.002}***
$H_0 : \beta_2 + \beta_3 = 0$ p-val.		0.372		0.233
N	5394	5394	7438	7438

Notes: Extension of Table 3.5. * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; p-values for SHAC robust SEs provided in "{.}" brackets. Controls include: if HH is rural and wealth index (R1); respondent is female; age and age-squared. Fixed effects for survey year, month of interview, and community are suppressed.

Table C.16: Impact of +1.5 S.D Shocks in Previous Year, ENAHO 2015-2017

	Hours worked	
Positive shock	0.028	-0.021
	(0.601)	(0.722)
**Agricultural work*		0.212
		(0.035)**
Agricultural work	-10.638	-11.076
	(0.000)***	(0.000)***
Controls	Yes	Yes
N	144713	144713

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. Data sourced from ENAHO annual waves 2015-2017. Dependent variable is total hours (usually) worked in previous week. Model (1) is the base model for all working age respondents (15-64). (2) interacts shock exposure with if the respondent reports working in an agricultural occupation (ISIC rev4. 4-number code 0100-0199). Standard errors are clustered at the district level, with p-values reported in parentheses "(.)". Controls for respondent: is female, mother tongue is Spanish, married or cohabitating, completed primary education, age and age-squared, and if works in agricultural occupation (column (1)), as well as a Rural/urban community indicator. District, month interview and year of survey fixed effects are suppressed.

Table C.17: Impact of +1.5S.D. Shocks on Older Sibling Time Use

	Unpaid work	Paid work	Housework	Childcare	School	Study	Play	Sleep
Positive Shock	0.052	0.064	-0.030	-0.013	-0.045	0.008	-0.019	0.015
	(0.189)	(0.212)	(0.094)*	(0.622)	(0.532)	(0.779)	(0.627)	(0.554)
	[0.297]	[0.256]	[0.173]	[0.656]	[0.603]	[0.797]	[0.689]	[0.588]
	{0.032}**	{0.718}	{0.160}	{0.629}	{0.978}	{0.689}	{0.537}	{0.729}
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5180	5180	5180	5180	5180	5180	5180	5180

* p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; wild bootstrapped (10,000 replications) p-values provided in "[.]" brackets; p-values for SHAC robust SEs provided in "{.}" brackets. Controls: HH is rural, respondent gender and age. Fixed effects for survey round, month of interview, and community are suppressed.

Table C.18: Home Environment Measures, Summary Statistics

	Mean	S.D.	Min	Max
Parent-child relationship				
Marsh SDQ Parent relations scale	0.00	(1.00)	-5.12	2.23
Parental involvement index	0.00	(1.00)	-3.87	1.21
Parental involvement				
Know friend's names	0.84	(0.37)	0.00	1.00
Know friend's parents	0.73	(0.44)	0.00	1.00
Know teacher's name	0.96	(0.19)	0.00	1.00
Know child's after-school activity	0.95	(0.23)	0.00	1.00
Feel close with their child	0.94	(0.23)	0.00	1.00
Talk to child about politics	0.22	(0.41)	0.00	1.00
Reading index	-0.00	(1.00)	-2.38	1.65
Reading encouragement				
Encourage to read	0.52	(0.50)	0.00	1.00
Child reads for fun	0.62	(0.49)	0.00	1.00
HH has dictionary	0.88	(0.32)	0.00	1.00
Child uses dictionary	0.78	(0.41)	0.00	1.00
HH has more than 20 books	0.18	(0.39)	0.00	1.00
Education expenditure				
ln(Education expenditure on child)	5.40	(1.23)	0.00	8.56

Notes: Sample means are reported with standard deviations in parentheses.

Table C.19: 1st Factor Loadings for Marsh SDQ Parent Relations Scale

	Loading	Ψ
I like my parents.	0.529	0.720
My parents like me.	0.510	0.740
My parents and I spend a lot of time together.	0.540	0.709
I get along well with my parents.	0.662	0.562
My parents understand me.	0.655	0.571
If I have children of my own, I want to bring them up like my parents raised me.	0.559	0.687
My parents are easy to talk to.	0.530	0.719
My parents and I have a lot of fun together.	0.605	0.634

Notes: Factor loadings ≥ 0.3 are highlighted in green. Ψ is the share of item unique variance.

Table C.20: Impact of +1.5S.D. Shocks on Parent-Child Relationship Measures

	Parent involvement	Parent Relations	Reading encouragement	Education expenditure
Prenatal	-0.017	0.085	-0.053	-0.044
	(0.622)	(0.011)**	(0.132)	(0.156)
	{0.620}	{0.010}***	{0.129}	{0.152}
1st year	-0.047	0.013	-0.066	-0.070
	(0.340)	(0.816)	(0.103)	(0.107)
	{0.318}	{0.823}	{0.076}*	{0.096}*
2nd year	0.036	-0.006	0.081	0.030
	(0.335)	(0.873)	(0.068)*	(0.637)
	{0.314}	{0.871}	{0.032}**	{0.615}
3rd year	-0.109	-0.062	0.018	0.016
	(0.003)***	(0.026)**	(0.621)	(0.703)
	{0.004}***	{0.041}**	{0.594}	{0.691}
Controls	Yes	Yes	Yes	Yes
N	2089	1995	2089	2089

Notes: Extension of Table 3.6. * p < 0.10, ** p < 0.05, *** p < 0.01. P-values based on clustered robust SEs at district level are in parentheses "(.)"; p-values for SHAC robust SEs provided in "{.}" brackets. Controls: HH is rural and HH wealth index; mother age and education; child gender, mother tongue, age, and if they were enrolled in pre-school. Fixed effects for birth month cohort and community are suppressed.

Appendix D

Chapter 4 Appendix

D.1 Local Randomisation Framework

D.1.1 Sharp Local Randomisation RDD

Given a score, X_i, treatment assignment is given by $T_i = \mathbb{1}(X_i \geq c)$, where c is the cutoff point. Y_i is the *observed* outcome for unit i, with *potential* outcomes $Y_i(0)$ and $Y_i(1)$ under control and treatment. We assume for some scalar $w > 0$ there exists a window $W = [c-w, c+w]$ containing the cutoff c, such that the assumptions described below hold. Let $\mathbf{X}_W$ be the vector of scores and $\mathbf{T}_W$ be the vector of treatment statuses of all units i within that window for which $X_i \in W$, and let their equivalent potential outcomes be given by $\mathbf{Y}_W(0)$ and $\mathbf{Y}_W(1)$. Following Cattaneo et al. (2017, 2024), we define the conditions required for the basic local randomisation framework, assuming random potential outcomes drawn from a super-population. We first assume that there exists some window $W = [c-w, c+w]$, for which the potential outcomes of a unit i are statistically independent of their score X_i:

LR 1 (Independence of Scores from Potential outcomes):

$$(Y_i(0), Y_i(1)) \perp\!\!\!\perp X_i | X_i \in W$$

Alternatively, This can be stated in terms of probability distribution functions:

$$\mathbb{P}[X_i \leq x | Y_i(0), Y_i(1), X_i \in W] = \mathbb{P}[X_i \leq x | X_i \in W] \tag{D.1}$$

This ensures that, for all units $X_i \in W = [c-w, c+w]$, placement above and below the cutoff is not related to potential outcomes, and potential outcomes are not related to the score[1]. This implies $\mathbb{E}[Y_i(d)|X_i, X_i \in W] = \mathbb{E}[Y_i(d)|X_i \in W]$ for $d = 0, 1$, meaning the

[1]Alternatively a less strict assumption can be made that potential outcomes are independent from treatment assignment. This assumption however does not imply that potential outcomes are unrelated to the score variable (Sekhon & Titiunik, 2017), requiring an additional exclusion restriction assumption

conditional expectations are constant functions of the score inside the window (although they may have non-zero slopes outwith).

We can therefore define the parameter of interest. defining $\mathbb{E}_W[\cdot]$ as the conditional expectation with respect to the probability $\mathbb{P}_W[\cdot]$, computed conditionally for units with $X_i \in W$. Defining N_W as the number of units in window W, the local randomisation sharp RD treatment effect is given by:

$$\theta_{SRD} = \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{T_i Y_i}{\mathbb{P}_W[T_i = 1]} \right] - \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{(1 - T_i) Y_i}{1 - \mathbb{P}_W[T_i = 1]} \right] \tag{D.2}$$

Which expresses the RD as the difference in average observed outcomes just below and just above the cutoff within the window. Under LR 1, this is equivalent to our parameter of interest:

$$\theta_{SRD} \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W [Y_i(1) - Y_i(0)] \tag{D.3}$$

The average treatment effect of T_i on Y_i within the window. The next section will extend the local randomisation framework for cases of imperfect compliance (Fuzzy RDD).

D.1.2 Fuzzy Local Randomisation RDD

In Fuzzy RDD, with assignment of treatment $T_i = \mathbb{1}(X_i \geq c)$, either some units with $X_i \geq c$ fail to receive the treatment or $X_i < c$ receive the treatment anyway. Therefore, the change in probability of receiving the treatment at the cutoff is not 0 to 1. Following Cattaneo et al. (2024), we define D_i as the indicator of treatment received, which has two potential values: $D_i(1)$ is the treatment received by unit i when assigned to the treatment condition $(T_i = \mathbb{1}(X_i \geq c) = 1)$, and $D_i(0)$ treatment received when assigned to control $(T_i = 0)$, with $D_i(1), D_i(0) \in \{0, 1\}$. We can write, for example, $D_i(0) = 1$ if a unit receives treatment even though it is assigned to the control condition (non-compliance), or $D_i(0) = 0$ if it is assigned control assignment and does not receive treatment, and similarly with $D_i(1)$. These are defined as our *potential* treatments.

Given noncompliance, we redefine potential outcomes as $Y_i(T_i, D_i(T_i))$, which includes arguments for both treatment assigned and *potential* treatment status. There are now four potential outcomes. The potential outcome for unit i assigned to treatment is given by $Y_i(1, D_i(1)) = D_i(1)Y_i(1, 1) + (1 - D_i(1))Y_i(1, 0)$, resulting in $Y_i(1, 1)$ if $D_i(1) = 1$ and $Y_i(1, 0)$ if $D_i(1) = 0$, with the potential outcomes for unit i assigned to

that potential outcomes are not a function of the score(Cattaneo et al., 2017). While our assumption is stricter, given that our narrow window this is reasonable, and ensures this exclusion holds implicitly.

control defined analogously. To interpret the below parameters of interest we must, in addition to assuming potential outcomes are not a function of the score (LR 1), assume that potential treatments $D_i(1)$ and $D_i(0)$ are also unaffected by the score X_i within W:

LR 2 (Independence of Score from Potential Treatments):

$$(D_i(0), D_i(1)) \perp\!\!\!\perp X_i | X_i \in W$$

under LR 2, as the potential treatments $D_i(1), D_i(0)$ are not related to the score, we can say that our augmented potential outcome $Y_i(T_i, D_i(T_i))$ is also not related to the score, fulfilling LR 1. This allows for the definition of an assumption which is analogous to the exclusion restriction evoked in traditional instrumental variable (IV) settings. For a given value of treatment received $D_i = d$, for all units i with $X_i \in W$:

LR 3 (Exclusion Restriction on the Treatment Assignment):

$$Y_i(T_i, d) = Y_i(d) \qquad \forall d$$

This implies that the treatment assignment affects potential outcomes and potential treatments only through the treatment received D_i, not but not directly. Given assignment $T_i = \mathbb{1}(X_i \geq c)$ is a function of X_i, assuming that our augmented outcomes $Y_i(T_i, 0)$ and $Y_i(T_i, 1)$ are not related to the score implies that, given a particular value of treatment received $D_i = d$, potential outcomes do not depend on treatment assignment T_i.

However, LR 3 alone is not sufficient to recover the effects of treatment in fuzzy RDD. This is due to the decision to comply with treatment assignment still being unrestricted. To be able obtain our treatment effects of interest, we must make additional assumptions. Following the standard set up of Imbens and Angrist (1994), we define four different groups based on their compliance decisions: *Compliers*, whose treatment received matches their assigned treatment status (whose potential treatments are such that $D_i(1) = 1$ and $D_i(0) = 0$); *Always-takers*, who always take up treatment whether assigned or not $(D_i(1) = D_i(0) = 1)$; *Never-takers*, who always refuse treatment regardless of assignment $(D_i(1) = D_i(0) = 0$; and *defiers*, who would always receive the opposite treatment from the one assigned $(D_i(1) = 0$ and $D_i(0) = 1$. We assume that, within $W = [c - w, c + w]$:

LR 4 (Monotonicity):

$$D_i(1) \geq D_i(0) \qquad \forall i, X_i \in W$$

This condition ensures that assignment to treatment affects the treatment received in a monotone way. That is, if respondents on average are more likely to receive treatment

given $D_i(1)$ than $D_i(0)$, then anyone who would receive treatment under $D_i(0)$ must also receive treatment under $D_i(1)$. As such, this condition rules out the potential for defiers.

Let B_i be a binary variable denoting compliance status, $B_i = 1$ for compliance with treatment assignment and $B_i = 0$ for non-compliance. Then under assumptions LR 1 – LR 4 it can be shown that the effect of the treatment received for the sub-population of units that comply with treatment assignment within the window W, commonly referred to as the Local Average Treatment Effect (LATE) is given by:

$$\theta_{FRD} = \frac{1}{N_W} \sum_{i:B_i, X_i \in W} \mathbb{E}\left[Y_{1,1} - Y_{1,0} | B_i = 1\right] \tag{D.4}$$

When the above assumptions are met, we can show that the ratio of the effect of the treatment assignment on the outcome, θ_{ITT}, and the effect of the treatment assignment on treatment received, θ_{FS} is equivalent to the LATE for compliers, θ_{FRD}:

$$\theta_{FRD} \equiv \frac{\theta_{ITT}}{\theta_{FS}} \tag{D.5}$$

From this we can proceed with identifying the required parameters given our assumptions.

D.1.3 Intention-to-Treat Effects

Define the effect of being *assigned* to treatment, whether treatment is received or not, on the outcome Y_i as θ_Y. Applying the fuzzy RDD context to Equation D.2 we can estimate the following parameter:

$$\theta_Y = \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{T_i Y_i(1, D_i(1))}{\mathbb{P}_W[T_i = 1]} \right] - \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{(1 - T_i) Y_i(0, D_i(0))}{1 - \mathbb{P}_W[T_i = 1]} \right] \tag{D.6}$$

The difference in average observed outcomes just below and just above the cutoff within the window. Under LR 1 and LR 2:

$$\theta_Y = \theta_{ITT}, \qquad \theta_{ITT} \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W[Y_i(1, D_i(1)) - Y_i(0, D_i(0))] \tag{D.7}$$

such that the estimated difference in average observed outcomes above and below the cutoff is equivalent to the causal effect of T_i on Y_i, commonly referred to as the "intention-to-treat effect" (ITT), θ_{ITT}, within the window W.

D.1.4 First Stage

We also define the effect of assignment to treatment T_i on the probability of receiving D_i, which reveals information regarding compliance, as θ_D, treating D_i as the outcome. We define the following parameter:

$$\theta_D \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{T_i D_i}{\mathbb{P}_W[T_i = 1]} \right] - \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{(1 - T_i) D_i}{1 - \mathbb{P}_W[T_i = 1]} \right] \tag{D.8}$$

θ_D captures the difference in the probability of *receiving* the treatment between units *assigned* to treatment and units *assigned* to control, within the window W. Under assumptions LR 1 and LR 2, we can interpret this parameter as the causal impact of T_i on D_i. Following the IV literature, we refer to this as the first stage effect, θ_{FS}:

$$\theta_D = \theta_{FS}, \qquad \theta_{FS} \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W[D_i(1) - D_i(0)] \tag{D.9}$$

D.1.5 Estimation of LATE

In the local randomization framework, θ_{ITT} and θ_{FS}, can be estimated by calculating sample difference-in-means between units above (with subscript $W+$) and below (subscript $W-$) the cutoff for units with scores in W:

$$\widehat{\theta}_{ITT} = \bar{Y}_{W+} - \bar{Y}_{W-} \quad \text{and} \quad \widehat{\theta}_{FS} = \bar{D}_{W+} - \bar{D}_{W-}, \quad \widehat{\theta}_{FRD} = \frac{\bar{Y}_{W+} - \bar{Y}_{W-}}{\bar{D}_{W+} - \bar{D}_{W-}} \tag{D.10}$$

where

$$\bar{Y}_{W+} = \frac{1}{N_{W+}} \sum_{i:X_i \in W} \omega_i T_i Y_i, \quad \bar{Y}_{W-} = \frac{1}{N_{W-}} \sum_{i:X_i \in W} \omega_i (1 - T_i) Y_i \tag{D.11}$$

and

$$\bar{D}_{W+} = \frac{1}{N_{W+}} \sum_{i:X_i \in W} \omega_i T_i D_i, \quad \bar{D}_{W-} = \frac{1}{N_{W-}} \sum_{i:X_i \in W} \omega_i (1 - T_i) D_i, \tag{D.12}$$

,

where ω is a weighting scheme for unit i. Inference is based on standard IV large sample approximations using the Delta method, applied to observations within the window W. As θ_{FRD} is a ratio, it will be undefined when the denominator is zero, thus a further assumption must be made that the first stage exists and is non-zero:

LR 5 (Instrument Relevance):

$$\theta_{FS} \neq 0, \quad or \quad \mathbb{P}[D_i = 1 | X_i \geq c, X_i \in W] > \mathbb{P}[D_i = 1 | X_i < c, X_i \in W]$$

Where the second definition is similar to that stated by Cattaneo et al. (2017). Furthermore, similar to IV settings, when the cutoff rule has a non-zero but small effect on the probability of treatment ("weak instruments"), standard Gaussian approximations of the distributions of test statistics are not reliable. The strength of the first stage can be assessed by the size of the F-statistic in the first stage regression, with Cattaneo et al. (2024) recommending a higher rule of thumb threshold of an F-statistic of 20 or more.

D.2 Additional Tables and Figures

Figure D.1: Covariate Smoothness Across Cutoff: Sibling-Pair Characteristics

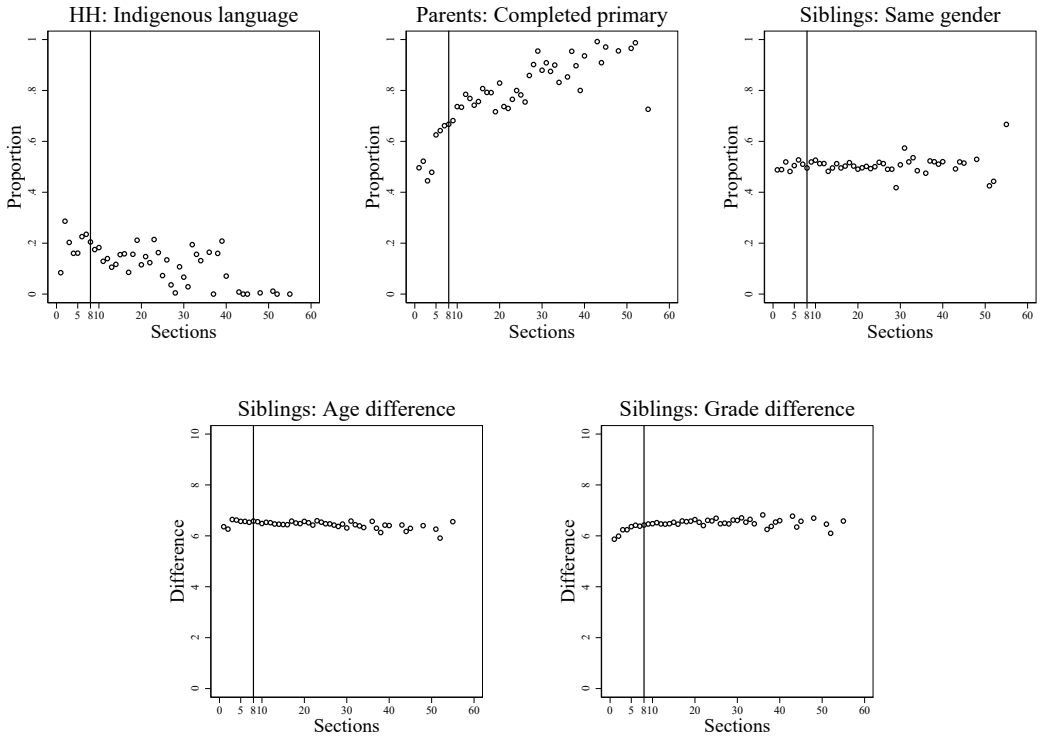


Figure D.2: Covariate Smoothness Across Cutoff: School-Level Characteristics

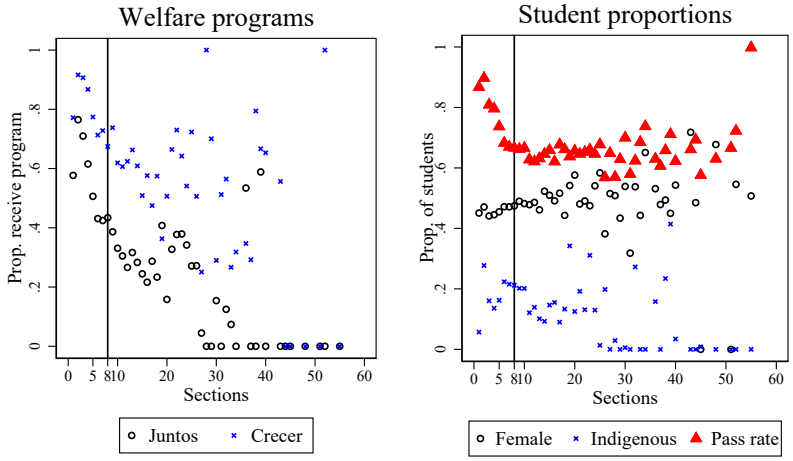


Figure D.3: Discontinuity in Outcomes at Cutoff

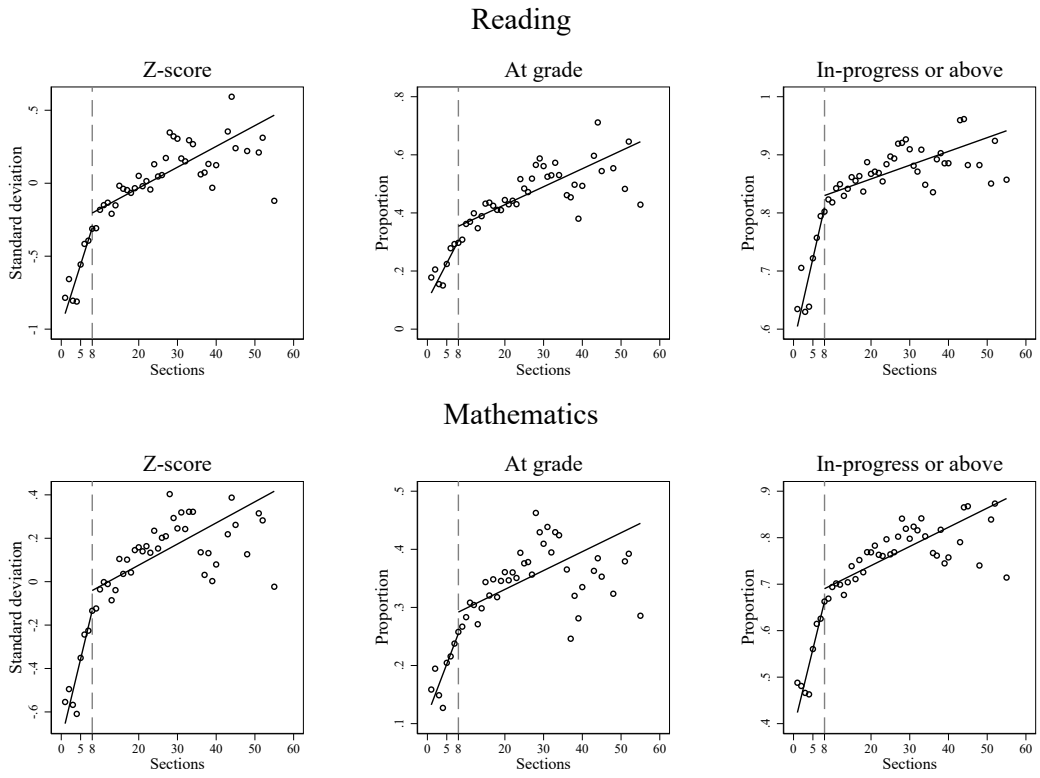


Table D.1: Distribution of Weekly Hours by Public School Type

Subject	Basic	JEC
Mathematics	4	6
Reading	4	5
English	2	5
Science	3	5
History	3	3
Work education	2	3
Civics	2	3
Person, family & community	2	2
Physical education	2	2
Art	2	2
Religion	2	2
Tutoring	1	2
Free	6	5
Total	35	45

Notes: A pedagogical hours is 45 minutes long.
Source: Peru's Ministry of Education.

Table D.2: Selection Process for JEC

	Steps	**Schools**
1.	**Initial eligibility:** – Public secondary schools – Eight or more sections – Registered as morning shift only – School premises only used in the morning – Sufficient space for additional resources	1362
2.	52 "emblematic" schools added.	1412
3.	Local coordinators carry out validation process. Remaining schools required to provide additional information.	1343
4.	External assessors use information to select 1000 schools.	1000
5.	JEC reform announced in September 2014, along with list of schools (RM N° 451-2014-MINEDU)	1000
6.	On February 10th list is modified, with six schools replaced (RM N° 062-2015-MINEDU)	1000

Source: Alcázar (2016) and Agüero et al. (2021).

Table D.3: ECE Coverage, by Year and Grade

Year	Grade	% Schools evaluated	% Students evaluated
2015	2nd Grade	99.7	94
2016	2nd Grade	99.8	96.5
	4th Grade	99.8	96.9

Notes: Coverage of primary schools is provided only as total for all grades evaluated. *Source:* MINEDU.

Table D.4: TSLS Predetermined Covariate Smoothness Tests

	School district receives program		Proportion of students		Pass rate
	Juntos	Crecer	Girls	Indigenous	Total
Panel A: School level					
Sections≥8	0.007	-0.029	0.014	-0.047	0.030
	(0.036)	(0.035)	(0.011)	(0.027)*	(0.012)**
Control mean	0.508	0.775	0.457	0.171	0.736
N	59338	59338	59338	59338	59338
	Sibling Dyad			Household	
	Same gender	Age diff.	Grade diff.	Indig. lang.	Parent educ.
Panel B: Sibling dyad level					
Sections≥8	-0.010	0.041	-0.029	-0.054	-0.014
	(0.012)	(0.040)	(0.027)	(0.024)**	(0.021)
Control mean	0.506	6.563	5.702	0.174	0.616
N	59338	59338	59338	59338	59338

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the reduced form intention-to-treat effect at the cutoff, including a linear spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses.

Table D.5: Effects of JEC on Younger Sibling Outcomes: ITT Results

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Sections≥8	0.082	0.004	0.008	0.092	0.020	0.037
	[0.003]***	[0.760]	[0.530]	[0.002]***	[0.124]	[0.011]**
Control mean	-0.393	0.293	0.795	-0.226	0.238	0.626
N	4368	4368	4368	4368	4368	4368

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table D.6: Effects of JEC on Younger Sibling Outcomes: 2SLS, No Controls

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Panel A: Linear specification						
Sections≥8	0.133	0.060	0.017	0.128	0.057	0.039
	(0.063)**	(0.026)**	(0.023)	(0.064)**	(0.023)**	(0.029)
Controls	No	No	No	No	No	No
N	59954	59954	59954	59954	59954	59954
Panel B: Quadratic specification						
Sections≥8	-0.085	-0.064	-0.060	-0.045	-0.012	-0.045
	(0.102)	(0.043)	(0.038)	(0.102)	(0.039)	(0.046)
Controls	No	No	No	No	No	No
N	59954	59954	59954	59954	59954	59954

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) at the cutoff, including a linear (panel A) and quadratic (panel B) spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses.

Table D.7: Heterogeneous Effects, by Gender, 2SLS

	Female			Male		
	(1) All	(2) Older sister	(3) Older brother	(4) All	(5) Older sister	(6) Older brother
Reading						
Z-score	0.298	0.235	0.351	0.030	0.019	0.033
	(0.069)***	(0.084)***	(0.086)***	(0.065)	(0.076)	(0.081)
At grade	0.113	0.124	0.101	0.039	0.037	0.038
	(0.030)***	(0.037)***	(0.038)***	(0.028)	(0.035)	(0.034)
≥ In prog.	0.067	0.032	0.097	-0.003	-0.012	0.006
	(0.027)**	(0.035)	(0.034)***	(0.025)	(0.033)	(0.033)
Mathematics						
Z-score	0.277	0.251	0.297	0.003	-0.025	0.026
	(0.071)***	(0.086)***	(0.089)***	(0.071)	(0.084)	(0.088)
At grade	0.102	0.092	0.112	0.023	0.019	0.027
	(0.026)***	(0.033)***	(0.034)***	(0.029)	(0.035)	(0.037)
≥ In prog.	0.118	0.091	0.139	-0.026	-0.034	-0.019
	(0.033)***	(0.043)**	(0.041)***	(0.032)	(0.038)	(0.041)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	29237	14546	14691	30101	14598	15503

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) at the cutoff, including a linear spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses. Additional covariates include sibling pair age difference in years, an indicator of being the same gender, and mother's educational attainment in years. Fixed effects for older sibling grade and survey year are also included.

Table D.8: Local Randomisation Placebo Test, Private Schools

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Sections≥8	-0.050	-0.085	-0.034	-0.032	0.020	-0.037
	[0.640]	[0.173]	[0.320]	[0.780]	[0.722]	[0.508]
Mean Dep.	0.130	0.532	0.924	-0.003	0.278	0.722
N	424	424	424	424	424	424

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table D.9: Local Randomisation Placebo Test, Placebo Cutoffs

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Panel A: Below cutoff						
Sections≥4	-0.006	-0.005	0.009	-0.042	-0.022	-0.003
	[0.894]	[0.787]	[0.702]	[0.371]	[0.192]	[0.895]
Mean Dep.	-0.805	0.155	0.630	-0.568	0.149	0.466
Window	[3,4]	[3,4]	[3,4]	[3,4]	[3,4]	[3,4]
N	1718	1718	1718	1718	1718	1718
Panel B: Above cutoff						
Sections≥12	0.016	0.029	0.007	-0.009	-0.004	-0.003
	[0.547]	[0.041]**	[0.504]	[0.758]	[0.767]	[0.844]
Mean Dep.	-0.150	0.370	0.842	-0.002	0.308	0.702
Window	[11,12]	[11,12]	[11,12]	[11,12]	[11,12]	[11,12]
N	4616	4616	4616	4616	4616	4616

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table D.10: Alternative Window Selection Procedure

Window	Variable	Min. p-value	Obs.<8	Obs.≥8
[7, 8]	Prop. female students	0.251	2103	2265
[6, 9]	Age difference	0.308	4286	4573
[5, 10]	Juntos district	0.000	23503	9099
[4, 11]	Juntos district	0.000	24408	11529
[3, 12]	District Juntos eligible	0.000	25221	13715

Notes: Variable refers to the predetermined covariate with the lowest difference in means p-value across the cutoff within that window. List excludes two variables that are not balanced as shown in Table 4.4. Significance level for rejecting the null, $H_0 : Z_i(1) = Z_i(0) \forall i$, is set at $\alpha \geq 0.15$, as recommended by Cattaneo et al. (2024).

Table D.11: Effects of JEC on Younger Sibling Outcomes: Alternative Window

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Sections≥8	0.095 [0.000]***	0.017 [0.078]*	0.037 [0.000]***	0.106 [0.000]***	0.036 [0.000]***	0.046 [0.000]***
Control mean	-0.405	0.286	0.776	-0.235	0.227	0.620
Window	[6,9]	[6,9]	[6,9]	[6,9]	[6,9]	[6,9]
N	8859	8859	8859	8859	8859	8859

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for all windows.

Table D.12: Effects of JEC on Younger Sibling Outcomes: Including 2018 Survey Wave

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Sections≥8	0.088 [0.014]**	-0.003 [0.864]	-0.020 [0.246]	0.095 [0.011]**	0.015 [0.364]	0.046 [0.014]**
Control mean	-0.387	0.281	0.748	-0.226	0.239	0.625
N	5539	5539	5539	5539	5539	5539

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table D.13: Effects of JEC on Younger Sibling Outcomes, by Survey Wave and Grade

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Panel A: Year 2015 2nd grade						
Sections≥8	0.183	0.060	0.013	0.099	0.009	0.069
	[0.018]**	[0.113]	[0.575]	[0.190]	[0.791]	[0.068]*
Control mean	-0.381	0.328	0.902	-0.269	0.241	0.627
N	1805	1805	1805	1805	1805	1805
Panel B: Year 2016 2nd grade						
Sections≥8	0.032	0.001	0.028	0.041	0.043	-0.019
	[0.631]	[0.979]	[0.175]	[0.599]	[0.234]	[0.588]
Control mean	-0.355	0.346	0.914	-0.049	0.293	0.716
N	1321	1321	1321	1321	1321	1321
Panel C: Year 2016 4th grade						
Sections≥8	0.156	-0.020	0.057	0.222	0.033	0.109
	[0.024]**	[0.481]	[0.136]	[0.003]***	[0.267]	[0.004]***
Control mean	-0.454	0.177	0.484	-0.333	0.176	0.532
N	1242	1242	1242	1242	1242	1242

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

References

Abramowitz, M., & Stegun, I. A. (1968). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Vol. 55). U.S. Government Printing Office.

Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects. *American Political Science Review*, *110*(3), 512–529.

Agüero, J. (2016). *Evaluación de impacto de la jornada escolar completa*. GRADE and FORGE.

Agüero, J., Favara, M., Porter, C., & Sanchez, A. (2021). *Do More School Resources Increase Learning Outcomes? Evidence from an Extended School-Day Reform*.

Aguirre, J., & Matta, J. (2021). Walking in your footsteps: Sibling spillovers in higher education choices. *Economics of Education Review*, *80*, 102062.

Aiken, C. E., & Ozanne, S. E. (2014). Transgenerational developmental programming. *Human Reproduction Update*, *20*(1), 63–75.

Akresh, R., Halim, D., & Kleemans, M. (2023). Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia. *The Economic Journal*, *133*(650), 582–612.

Alcázar, L. (2016). *Evaluación del diseño y proceso de implementación del modelo de jornada escolar completa (jec) para educación secundaria a nivel nacional. informe final*. GRADE and FORGE.

Almeida, R., Bresolin, A., Borges, B., Mendes, K., & Menezes-Filho, N. (2016). *Assessing the Impacts of Mais Educação on Educational Outcomes: Evidence between 2007 and 2011* (No. 7644). The World Bank.

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality Psychology and Economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (pp. 1–181, Vol. 4). Elsevier.

Almond, D., & Currie, J. (2011). Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, *25*(3), 153–172.

Almond, D., Currie, J., & Duque, V. (2018). Childhood Circumstances and Adult Outcomes: Act II. *Journal of Economic Literature*, *56*(4), 1360–1446.

Almond, D., Currie, J., & Herrmann, M. (2012). From infant to mother: Early disease environment and future maternal health. *Labour Economics*, *19*(4), 475–483.

Almond, D., Edlund, L., Li, H., & Zhang, J. (2007). *Long-Term Effects Of The 1959-1961 China Famine: Mainland China and Hong Kong*.

Almond, D., & Mazumder, B. (2013). Fetal Origins and Parental Responses. *Annual Review of Economics*, *5*(1), 37–56.

Altmejd, A., Barrios-Fernández, A., Drlje, M., Goodman, J., Hurwitz, M., Kovac, D., Mulhern, C., Neilson, C., & Smith, J. (2021). O Brother, Where Start Thou?

Sibling Spillovers on College and Major Choice in Four Countries. *The Quarterly Journal of Economics*, *136*(3), 1831–1886.

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2020). Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia. *American Economic Review*, *110*(1), 48–85.

Auffhammer, M., Hsiang, S. M., Schlenker, W., & Sobel, A. (2013). Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy*, *7*(2), 181–198.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). School Quality and the Gender Gap in Educational Achievement. *American Economic Review*, *106*(5), 289–295.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2019). Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes. *American Economic Journal: Applied Economics*, *11*(3), 338–381.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2023). Males at the Tails: How Socioeconomic Status Shapes the Gender Gap. *The Economic Journal*, *133*(656), 3136–3152.

Bale, T. L. (2015). Epigenetic and transgenerational reprogramming of brain development. *Nature Reviews: Neuroscience*, *16*(6), 332–344.

Barker, D. J. (1990). The fetal and infant origins of adult disease. *British Medical Journal*, *301*(6761), 1111–1111.

Barrios-Fernández, A., & Bovini, G. (2021). It's time to learn: School institutions and returns to instruction time. *Economics of Education Review*, *80*, 102068.

Beard, A. S., & Blaser, M. J. (2002). The ecology of height: The effect of microbial transmission on human height. *Perspectives in Biology and Medicine*, *45*(4), 475–498.

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, *4*(3), S1–S39.

Bedard, K., & Dhuey, E. (2006). The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. *The Quarterly Journal of Economics*, *121*(4), 1437–1472.

Begueria, S., Vicente-Serrano, S. M., & Angulo-Martinez, M. (2010). A multiscalar global drought dataset: The SPEIbase: A new gridded product for the analysis of drought variability and impacts. *Bulletin of the American Meteorological Society*, *91*(10), 1351–1354.

Begueria, S., Vicente-Serrano, S. M., Fergus, R.-G., & Borja, L. G. (2023). *SPEIbase v.2.9* (Version 2.9).

Behrman, J. R., & Rosenzweig, M. R. (2002). Does Increasing Women's Schooling Raise the Schooling of the Next Generation? *American Economic Review*, *92*(1), 323–334.

Behrman, J. R., Schott, W., Mani, S., Crookston, B. T., Dearden, K., Duc, L. T., Fernald, L. C. H., & Stein, A. D. (2017). Intergenerational Transmission of Poverty and Inequality: Parental Resources and Schooling Attainment and Children's Hu-

man Capital in Ethiopia, India, Peru, and Vietnam. *Economic Development and Cultural Change*, *65*(4), 657–697.

Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review*, *28*(5), 629–640.

Bellemare, M. F., Lee, Y. N., & Novak, L. (2021). Contract farming as partial insurance. *World Development*, *140*, 105274.

Benin, M. H., & Johnson, D. R. (1984). Sibling Similarities in Educational Attainment: A Comparison of Like-Sex and Cross-Sex Sibling Pairs. *Sociology of Education*, *57*(1), 11–21.

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

Berry, J., Dizon-Ross, R., & Jagnani, M. (2020). *Not Playing Favorites: An Experiment on Parental Fairness Preferences*.

Beusenberg, M., Orley, J. H., & World Health Organization Division of Mental Health. (1994). *A User's guide to the self reporting questionnaire (SRQ)*. World Health Organization.

Bevis, L. E. M., & Barrett, C. B. (2015). Decomposing Intergenerational Income Elasticity: The Gender-differentiated Contribution of Capital Transmission in Rural Philippines. *World Development*, *74*, 233–252.

Bevis, L. E. M., & Villa, K. (2022). Intergenerational Transmission of Maternal Health: Evidence from Cebu, the Philippines. *Journal of Human Resources*, *57*(5), 1425–1465.

Bhalotra, S., & Rawlings, S. (2013). Gradients of the Intergenerational Transmission of Health in Developing Countries. *The Review of Economics and Statistics*, *95*(2), 660–672.

Bhalotra, S., & Rawlings, S. B. (2011). Intergenerational persistence in health in developing countries: The penalty of gender inequality? *Journal of Public Economics*, *95*(3), 286–299.

Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., Roth, J., & Simonsen, M. (2021). Sibling Spillovers. *The Economic Journal*, *131*(633), 101–128.

Black, S. E., & Devereux, P. J. (2011). Recent Developments in Intergenerational Mobility. In D. Card & O. Ashenfelter (Eds.), *Handbook of Labor Economics* (pp. 1487–1541, Vol. 4). Elsevier.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. *American Economic Review*, *95*(1).

Bound, J., Brown, C., & Mathiowetz, N. (2001, January 1). Measurement Error in Survey Data. In J. J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics* (pp. 3705–3843, Vol. 5). Elsevier.

Boyden, J., Duc, L. T., Galeb, S., Penny, M., Sanchez, A., & Woldehanna, T. (2018). *Young Lives: An International Study of Childhood Poverty: Round 5, 2016 [data collection]*. SN: 8357, UK Data Service.

Brando, J. F., & Santos, R. J. (2015). *La Niña y los niños: Effects of an unexpected winter on early life human capital and family responses*.

Breining, S. N. (2014). The presence of ADHD: Spillovers between siblings. *Economics Letters*, *124*(3), 469–473.

Briones, K. (2017). 'How Many Rooms Are There in Your House?': Constructing the Young Lives Wealth Index. *Young Lives Technical Note*, *43*.

Briones, K. (2018). A Guide to Young Lives Rounds 1 to 5 Constructed Files. *Young Lives Technical Note*.

Cabrera-Hernández, F. (2020). Does Lengthening the School Day Increase School Value-Added? Evidence from a Mid-Income Country. *The Journal of Development Studies*, *56*(2), 314–335.

Caliendo, M., Cobb-Clark, D. A., & Uhlendorff, A. (2015). Locus of Control and Job Search Strategies. *The Review of Economics and Statistics*, *97*(1), 88–103.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, *90*(3), 414–427.

Cantril, H. (1965). The pattern of human concerns. *Rutgers University Press*.

Carrillo, B. (2020). Early Rainfall Shocks and Later-Life Outcomes: Evidence from Colombia. *The World Bank Economic Review*, *34*(1), 179–209.

Carroll, J. B. (1963). A Model of School Learning. *Teachers college record*, *64*(8), 1–9.

Caruso, G. D. (2015). *Intergenerational Transmission of Shocks in Early Life: Evidence from the Tanzania Great Flood of 1993*.

Case, A., & Paxson, C. (2008). Height, Health, and Cognitive Function at Older Ages. *American Economic Review*, *98*(2), 463–467.

Cattaneo, M. D., Frandsen, B. R., & Titiunik, R. (2015). Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. *Journal of Causal Inference*, *3*(1), 1–24.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2024). *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Cambridge University Press.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press.

Cattaneo, M. D., Keele, L., & Titiunik, R. (2023). Covariate Adjustment in Regression Discontinuity Designs. In J. R. Zubizarreta, E. A. Stuart, D. S. Small, & P. R. Rosenbaum (Eds.), *Handbook of Matching and Weighting in Causal Inference* (pp. 153–168). Chapman and Hall.

Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2017). Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality. *Journal of Policy Analysis and Management*, *36*(3), 643–681.

Cerdan-Infantes, P., & Vermeersch, C. (2007). *More Time Is Better : An Evaluation Of The Full Time School Program In Uruguay* (No. 4167). The World Bank.

Chang, C. H., Ferris, D. L., Johnson, R. E., Rosen, C. C., & Tan, J. A. (2012). Core Self-Evaluations: A Review and Evaluation of the Literature. *Journal of Management*, *38*(1), 81–128.

Chang, G., Favara, M., & Novella, R. (2022). The origins of cognitive skills and non-cognitive skills: The long-term effect of in-utero rainfall shocks in India. *Economics & Human Biology*, *44*, 101089.

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

Costa, P., & McCrae, R. (2008). The revised neo personality inventory (neo-pi-r). In *The SAGE handbook of personality theory and assessment* (2nd ed., pp. 179–198, Vol. 2).

Costello, A. B., & Osborne, J. W. (2005). Exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(7), 1–9.

Couttenier, M., & Soubeyran, R. (2014). Drought and Civil War in Sub-Saharan Africa. *The Economic Journal*, *124*(575), 201–244.

Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31–47.

Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, *43*(4), 738–782.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, *78*(3), 883–931.

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

Currie, J., & Moretti, E. (2007). Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight. *Journal of Labor Economics*, *25*(2), 231–263.

Currie, J., & Vogl, T. (2013). Early-Life Health and Adult Circumstance in Developing Countries. *Annual Review of Economics*, *5*, 1–36.

Dabelea, D., Hanson, R. L., Lindsay, R. S., Pettitt, D. J., Imperatore, G., Gabir, M. M., Roumain, J., Bennett, P. H., & Knowler, W. C. (2000). Intrauterine exposure to diabetes conveys risks for type 2 diabetes and obesity: A study of discordant sibships. *Diabetes*, *49*(12), 2208–2211.

Dahl, G. B., Rooth, D.-O., & Stenberg, A. (2023). *Intergenerational and Sibling Spillovers in High School Majors*.

Danysh, H. E., Gilman, R. H., Wells, J. C., Pan, W. K., Zaitchik, B., Gonzálvez, G., Alvarez, M., & Checkley, W. (2014). El Niño adversely affected childhood stature and lean mass in northern Peru. *Climate Change Responses*, *1*(1), 7.

De Onis, M. (2017). Child Growth and Development. In S. De Pee, D. Taren, & M. W. Bloem (Eds.), *Nutrition and Health in a Developing World* (pp. 119–141). Springer International Publishing.

De Onis, M., & Habicht, J. (1996). Anthropometric reference data for international use: Recommendations from a World Health Organization Expert Committee. *The American Journal of Clinical Nutrition*, *64*(4), 650–658.

De Onis, M., Onyango, A. W., Borghi, E., Siyam, A., Nishida, C., & Siekmann, J. (2007). Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization*, *85*(9), 660–667.

Del Bono, E., Francesconi, M., Kelly, Y., & Sacker, A. (2016). Early Maternal Time Investment and Early Child Outcomes. *The Economic Journal*, *126*(596), F96–F135.

Dell, M., Jones, B. F., & Olken, B. A. (2014). What Do We Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature*, *52*(3), 740–798.

Dercon, S., & Sanchez, A. (2013). Height in mid childhood and psychosocial competencies in late childhood: Evidence from four developing countries. *Economics and Human Biology*, *11*(4), 426–432.

Dinno, A. (2009). Exploring the Sensitivity of Horn's Parallel Analysis to the Distributional Form of Random Data. *Multivariate Behavioral Research*, *44*(3), 362–388.

Doyle, O., & Jernström, L. (2023). Second-generational impacts: The legacy of prenatal shocks on human capital. In K. F. Zimmermann (Ed.), *Handbook of labor, human resources and population economics* (pp. 1–25). Springer International Publishing.

Drake, A. J., & Walker, B. R. (2004). The intergenerational effects of fetal programming: Non-genomic mechanisms for the inheritance of low birth weight and cardiovascular risk. *Journal of Endocrinology*, *180*(1), 1–16.

Drake, A. J., & Liu, L. (2010). Intergenerational transmission of programmed effects: Public health consequences. *Trends in Endocrinology & Metabolism*, *21*(4), 206–213.

Duc, L. T., & Tam, T. N. M. (2015). Growth in Middle Childhood and Early Adolescence, and its Association with Cognitive and Non-cognitive Skills at the Age of 15 Years: Evidence from Vietnam. *Young Lives Working Paper*, *138*.

Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). *Test de Vocabulario en Imágenes Peabody: Adaptación hispanoamericana*. American Guidance Service Circle Pines, MN.

Duque, V. (2017). Early-life conditions and child development: Evidence from a violent conflict. *SSM - Population Health*, *3*, 121–131.

Duque, V., Rosales-Rueda, M., & Sanchez, F. (2019). *How Do Early-Life Shocks Interact with Subsequent Human Capital Investments? Evidence from Administrative Data*.

Dustan, A. (2018). Family networks and school choice. *Journal of Development Economics*, *134*, 372–391.

East, C. N., & Page, M. E. (2020). How do early-life health experiences affect future generations' equality of opportunity? In L. Tach, R. Dunifon, & D. L. Miller (Eds.), *Confronting inequality: How policies and practices shape children's opportunities.* (pp. 67–96). American Psychological Association.

Edo, M., & Nistal, M. (2022). *More time less time? The effect of lengthening the school day on learning trajectories*.

Emanuel, I., Filakti, H., Alberman, E., & Evans, S. J. (1992). Intergenerational studies of human birthweight from the 1958 birth cohort. 1. Evidence for a multigenerational effect. *British Journal of Obstetrics and Gynaecology*, *99*(1), 67–74.

Ersoy, E., & Forshaw, R. (2023). *The Effects of Longer Secondary School Days on Parents' Labour Market Outcomes: Evidence from the Jornada Escolar Completa Reform in Peru*.

Escobal, J., & Flores, E. (2008). An assessment of the Young Lives sampling approach in Peru. *Young Lives Technical Note*, *3*.

Escobal, J., & Flores, E. (2009). Maternal Migration and Child Well-being in Peru. *Young Lives Working Paper*, *56*.

Escobar, X., & Sanchez Castro, A. L. (2021). *Evaluación de diseño e implementación de la Jornada Escolar Completa: Lecciones aprendidas del modelo de jornada escolar ampliado en escuelas públicas*. Ministerio de Educación.

Espinoza Revollo, P., & Scott, D. (2022). Cognitive and Achievement Tests in the Young Lives Study. *Young Lives Technical Note*, *53*.

Fan, W., & Porter, C. (2020). Reinforcement or compensation? Parental responses to children's revealed human capital levels. *Journal of Population Economics*, *33*(1), 233–270.

Favara, M. (2018). Maternal group participation and child nutritional status in Peru. *Review of Development Economics*, *22*(2), 459–483.

Favara, M., Crivello, G., Penny, M., Porter, C., Revathi, E., Sánchez, A., Scott, D., Duc, L. T., Woldehanna, T., & McKay, A. (2022). Cohort Profile Update: The Young Lives study. *International Journal of Epidemiology*, *50*(6), 1784–1785e.

Figlio, D., Holden, K. L., & Ozek, U. (2018). Do students benefit from longer school days? Regression discontinuity evidence from Florida's additional hour of literacy instruction. *Economics of Education Review*, *67*, 171–183.

Figlio, D. N., Karbownik, K., & Özek, U. (2023). *Sibling Spillovers May Enhance the Efficacy of Targeted School Policies*.

Fletcher, J. M. (2013). The effects of personality traits on adult labor market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization*, *89*, 122–135.

Frandsen, B. R. (2017). Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete. In *Regression discontinuity designs: Theory and applications.* (pp. 281–315). Emerald Publishing Limited.

Fung, W., & Ha, W. (2010). Intergenerational Effects of the 1959–61 China Famine. In R. Fuentes-Nieva & P. A. Seck (Eds.), *Risk, Shocks, and Human Development* (pp. 222–254). Palgrave Macmillan UK.

Garganta, S., Pinto, M. F., & Zentner, J. (2022). *Extended School Day and Teenage Fertility in Dominican Republic*. Inter-American Development Bank.

Gelman, A., & Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, *37*(3), 447–456.

Gloor, M., Brienen, R. J. W., Galbraith, D., Feldpausch, T. R., Schöngart, J., Guyot, J.-L., Espinoza, J. C., Lloyd, J., & Phillips, O. L. (2013). Intensification of the Amazon hydrological cycle over the last two decades. *Geophysical Research Letters*, *40*(9), 1729–1733.

Gluckman, P., & Hanson, M. (2004). *The fetal matrix: Evolution, development and disease*. Cambridge University Press.

Godfrey, K. M., & Barker, D. J. (2000). Fetal nutrition and adult disease. *The American Journal of Clinical Nutrition*, *71*, 1344S–52S.

Goldman, N., Moreno, L., & Westoff, C. F. (1989). Collection of Survey Data on Contraception: An Evaluation of an Experiment in Peru. *Studies in Family Planning*, *20*(3), 147–157.

Goodman, J., Hurwitz, M., Smith, J., & Fox, J. (2015). The relationship between siblings' college choices: Evidence from one million SAT-taking families. *Economics of Education Review*, *48*, 75–85.

Gopnik, A., Meltzoff, A., & Kuhl, P. (1999). *The scientist in the crib: What early learning tells us about the mind,* Harper Collins.

Grätz, M., & Torche, F. (2016). Compensation or Reinforcement? The Stratification of Parental Responses to Children's Early Ability. *Demography*, *53*(6), 1883–1904.

Gurantz, O., Hurwitz, M., & Smith, J. (2020). Sibling effects on high school exam taking and performance. *Journal of Economic Behavior & Organization*, *178*, 534–549.

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, *19*(2), 149–161.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209.

Harari, M., & Ferrara, E. L. (2018). Conflict, Climate, and Cells: A Disaggregated Analysis. *The Review of Economics and Statistics*, *100*(4), 594–608.

Hayes, M. J., Svoboda, M. D., Wiihite, D. A., & Vanyarkho, O. V. (1999). Monitoring the 1996 Drought Using the Standardized Precipitation Index. *Bulletin of the American Meteorological Society*, *80*, 429–438.

Haylock, M. R., Peterson, T. C., Alves, L. M., Ambrizzi, T., Anunciação, Y. M. T., Baez, J., Barros, V. R., Berlato, M. A., Bidegain, M., Coronel, G., Corradi, V., Garcia, V. J., Grimm, A. M., Karoly, D., Marengo, J. A., Marino, M. B., Moncunill, D. F., Nechet, D., Quintana, J., ... Vincent, L. A. (2006). Trends in Total and Extreme South American Rainfall in 1960–2000 and Links with Sea Surface Temperature. *Journal of Climate*, *19*, 1490–1512.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, *103*(6), 2052–2086.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, *104*(33), 13250–13255.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464.

Hidalgo Arestegui, A. (2021). *Long term effects of an extended school-day programme: Evidence from a longitudinal study in Peru*.

Hilger, A., Nordman, C. J., & Sarr, L. R. (2022). Which Skills Matter for What Type of Worker? Cognitive Skills, Personality Traits, Hiring Channels and Wages in Bangladesh. *Indian Journal of Human Development*, *16*(2), 219–247.

Hoffmann, A. B. (2014). *Adverse in utero experiences and distortions of the sex ratio: Culling or scarring effects in relation to prenatal fasting?*

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185.

Huber, M. (2019). *A review of causal mediation analysis for assessing direct and indirect treatment effects* (No. 500). University of Fribourg.

Hyland, M., & Russ, J. (2019). Water as destiny – The long-term impacts of drought in sub-Saharan Africa. *World Development*, *115*, 30–45.

Ibáñez, L., Potau, N., Enriquez, G., & de Zegher, F. (2000). Reduced uterine and ovarian size in adolescent girls born small for gestational age. *Pediatric Research*, *47*(5), 575–577.

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635.

Joensen, J. S., & Nielsen, H. S. (2018). Spillovers in education choice. *Journal of Public Economics*, *157*, 158–183.

Joffe, M. M., & Greene, T. (2009). Related Causal Frameworks for Surrogate Outcomes. *Biometrics*, *65*(2), 530–538.

John, O. P. (1990). The "big five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research*. Guilford Press.

Judge, T. A., & Hurst, C. (2007). The benefits and possible costs of positive core self-evaluations: A review and agenda for future research. *Positive organizational behavior*, 159–174.

Judge, T. A., Locke, E. A., Durham, C. C., & Kluger, A. N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology*, *83*(1), 17–34.

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, *20*(1), 141–151.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401–415.

Karbownik, K., & Özek, U. (2023). Setting a Good Example?: Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design. *Journal of Human Resources*, *58*(5), 1567–1607.

Kerwin, J. T., & Thornton, R. L. (2021). Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures. *The Review of Economics and Statistics*, *103*(2), 251–264.

Khan, M. (2021). *Intergenerational consequences of in-utero and early-life conditions*.

Kim, S., Deng, Q., Fleisher, B. M., & Li, S. (2014). The Lasting Impact of Parental Early Life Malnutrition on Their Offspring: Evidence from the China Great Leap Forward Famine. *World Development*, *54*, 232–242.

Klebanoff, M. A., Secher, N. J., Mednick, B. R., & Schulsinger, C. (1999). Maternal size at birth and the development of hypertension during pregnancy: A test of the Barker hypothesis. *Archives of Internal Medicine*, *159*(14), 1607–1612.

Kochar, A. (1999). Smoothing Consumption by Smoothing Income: Hours-of-Work Responses to Idiosyncratic Agricultural Shocks in Rural India. *The Review of Economics and Statistics*, *81*(1), 50–61.

Kolesár, M., & Rothe, C. (2018). Inference in Regression Discontinuity Designs with a Discrete Running Variable. *American Economic Review*, *108*(8), 2277–2304.

Krutikova, S., & Lilleør, H. (2015). Fetal Origins of Personality: Effects of early life circumstances on adult personality traits.

Laajaj, R., & Macours, K. (2021). Measuring Skills in Developing Countries. *Journal of Human Resources*, *56*(4), 1254–1295.

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, *142*(2), 675–697.

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355.

Leight, J., Glewwe, P., & Park, A. (2015). *The impact of early childhood rainfall shocks on the evolution of cognitive and non-cognitive skills*.

Leon, J. (2020). Equating Cognitive Scores Across Rounds and Cohorts for Young Lives in Ethiopia, India, Peru and Vietnam. *Young Lives Technical Note*, *51*.

Leon, J., Guerrero, G., Cueto, S., & Glewwe, P. (2021). *What difference do schools make? A mixed methods study in secondary schools in Peru* (No. 114). GRADE. Lima.

Li, Q., & An, L. (2015). Intergenerational health consequences of the 1959–1961 Great Famine on children in rural China. *Economics & Human Biology*, *18*, 27–40.

Li, W., & Yang, D. T. (2005). The great leap forward: Anatomy of a central planning disaster. *Journal of Political Economy*, *113*(4), 840–877.

Llach, J., Adrogué, C., Gigaglia, M., & Orgales, C. R. (2009). Do Longer School Days Have Enduring Educational, Occupational, or Income Effects? A Natural Experiment in Buenos Aires, Argentina. *Economía*, *10*(1), 1–43.

Lloyd-Hughes, B., & Saunders, M. A. (2002). A drought climatology for Europe. *International Journal of Climatology*, *22*(13), 1571–1592.

Lumey, L. H. (1992). Decreased birthweights in infants after maternal in utero exposure to the Dutch famine of 1944–1945. *Paediatric and Perinatal Epidemiology*, *6*(2), 240–253.

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

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, *341*(6149), 976–980.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, *60*(3), 531–542.

Marsh, H. W. (1990). Self Description Questionnaire-I.

Marshall, W. A., & Tanner, J. M. (1969). Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood*, *44*(235), 291.

Marshall, W. A., & Tanner, J. M. (1970). Variations in the Pattern of Pubertal Changes in Boys. *Archives of Disease in Childhood*, *45*(239), 13–23.

Martorell, R., & Habicht, J.-P. (1986). Growth in early childhood in developing countries. *Human growth: a comprehensive treatise*, *3*, 241–262.

Matsuura, K., & Willmott. (2018). *Terrestrial air temperature and precipitation: 1900–2017 gridded monthly time series* (Version 5.01).

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., & Rajani, R. (2019). Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania. *The Quarterly Journal of Economics*, *134*(3), 1627–1673.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714.

McEwan, P. J., & Shapiro, J. S. (2008). The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates. *Journal of Human Resources*, *43*(1), 1–29.

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, *17*(22), 179–183.

McQuade, G., & Favara, M. (2024). Identifying climate shocks in Young Lives communities: Estimating weather conditions using a global gridded time series. *Young Lives Technical Note*, *57*.

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1), 202–216.

Moorthy, A. (2021). The effect of rainfall shocks on early childhood development in Uganda. *Economics Letters*, *200*, 109764.

Mueller, G., & Plug, E. (2006). Estimating the Effect of Personality on Male and Female Earnings. *ILR Review*, *60*(1), 3–22.

Nguyen, G., & Nguyen, T. T. (2020). Exposure to weather shocks: A comparison between self-reported record and extreme weather data. *Economic Analysis and Policy*, *65*, 117–138.

Nicoletti, C., & Rabe, B. (2019). Sibling spillover effects in school achievement. *Journal of Applied Econometrics*, *34*(4), 482–501.

Noack, C., Olma, T., & Rothe, C. (2023). *Flexible Covariate Adjustments in Regression Discontinuity Designs*.

Nordman, C. J., Sarr, L. R., & Sharma, S. (2019). Skills, personality traits, and gender wage gaps: Evidence from Bangladesh. *Oxford Economic Papers*, *71*(3), 687–708.

Nordman, C. J., Sharma, S., & Sunder, N. (2022). Here Comes the Rain Again: Productivity Shocks, Educational Investments, and Child Work. *Economic Development and Cultural Change*, *70*(3), 1041–1063.

Nübler, L., Austrian, K., Maluccio, J. A., & Pinchoff, J. (2021). Rainfall shocks, cognitive development and educational attainment among adolescents in a drought-prone region in Kenya. *Environment and Development Economics*, *26*(5–6), 466–487.

OECD. (2011). When Students Repeat Grades or Are Transferred Out of School.

Oettinger, G. S. (2000). Sibling Similarity in High School Graduation Outcomes: Causal Interdependency or Unobserved Heterogeneity? *Southern Economic Journal*, *66*(3), 631–648.

Ortega, J. (2018). Dentro del colegio y lejos del embarazo: El efecto de la Jornada Escolar Completa (JEC) sobre el embarazo adolescente en el Perú. *Investigaciones*.

Osborne, J. W., Costello, A. B., & Kellow, J. T. (2014). *Best practices in exploratory factor analysis*. CreateSpace Independent Publishing.

Osborne, J. W. (2015). What is rotating in exploratory factor analysis? *Practical Assessment, Research, and Evaluation*, *20*(1).

Pablo, A., Peter, H., & Evans, D. (2015, June 16). *Extending the School Day in Latin America and the Caribbean*.

Painter, R. C., Westendorp, R. G., de Rooij, S. R., Osmond, C., Barker, D. J., & Roseboom, T. J. (2008). Increased reproductive success of women after prenatal undernutrition. *Human Reproduction*, *23*(11), 2591–2595.

Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the School Day or School Year: A Systematic Review of Research (1985-2009). *Review of Educational Research*, *80*(3), 401–436.

Pathania, V. (2007). *The long run impact of drought at birth on height of women in rural India*.

Perez-Escamilla, R., Bermudez, O., Buccini, G. S., Kumanyika, S., Lutter, C. K., Monsivais, P., & Victora, C. (2018). Nutrition disparities and the global burden of malnutrition. *BMJ*, *361*, k2252.

Persson, P., Qiu, X., & Rossin-Slater, M. (2021). *Family Spillover Effects of Marginal Diagnoses: The Case of ADHD*.

Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature*, *465*(7299), 721–727.

Pitt, M. M., Rosenzweig, M. R., & Hassan, M. N. (1990). Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries. *The American Economic Review*, *80*(5), 1139–1156.

Qureshi, J. A. (2018a). Siblings, Teachers, and Spillovers on Academic Achievement. *Journal of Human Resources*, *53*(1), 272–297.

Qureshi, J. A. (2018b). Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital. *The Economic Journal*, *128*(616), 3285–3319.

Ramírez, I. J., & Briones, F. (2017). Understanding the El Niño Costero of 2017: The Definition Problem and Challenges of Climate Forecasting and Disaster Responses. *International Journal of Disaster Risk Science*, *8*(4), 489–492.

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*, *7*(9–12), 1393–1512.

Rocha, R., & Soares, R. R. (2015). Water scarcity and birth outcomes in the Brazilian semiarid. *Journal of Development Economics*, *112*, 72–91.

Rodrigues, M., & Campos Flores, A. (2021). *Evaluación de impacto de la Jornada Escolar Completa: Ampliación de la jornada escolar en escuelas públicas*. Ministerio de Educación.

Romano, J. P., & Wolf, M. (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica*, *73*(4), 1237–1282.

Rosales-Rueda, M. (2018). The impact of early life shocks on human capital formation: Evidence from El Niño floods in Ecuador. *Journal of Health Economics*, *62*, 13–44.

Rosenbaum, P. R. (1984). The Consequences of Adjustment for a Concomitant Variable that Has Been Affected by the Treatment. *Royal Statistical Society Journal Series A: General*, *147*(5), 656–666.

Rosenzweig, M. R., & Stark, O. (1989). Consumption Smoothing, Migration, and Marriage: Evidence from Rural India. *Journal of Political Economy*, *97*(4), 905–926.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*(1), 1–28.

Rougeaux, E., Miranda, J. J., Fewtrell, M., & Wells, J. C. K. (2022). Maternal internal migration and child growth and nutritional health in Peru: An analysis of the demographic and health surveys from 1991 to 2017. *BMC Public Health*, *22*(1), 37.

Royer, H., & Witman, A. (2014). Intergenerational Effects on Health – In Utero and Early Life. In *Encyclopedia of Health Economics* (pp. 83–90). Elsevier.

Royer, H. (2009). Separated at Girth: US Twin Estimates of the Effects of Birth Weight. *American Economic Journal: Applied Economics*, *1*(1), 49–85.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701.

Saavedra, J., & Gutierrez, M. (2020). Peru: A Wholesale Reform Fueled by an Obsession with Learning and Equity. In F. M. Reimers (Ed.), *Audacious Education Purposes: How Governments Transform the Goals of Education Systems* (pp. 153–180). Springer International Publishing.

Sacerdote, B. (2014). Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics*, *6*(1), 253–272.

Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010). Crop planting dates: An analysis of global patterns. *Global Ecology and Biogeography*, *19*(5), 607–620.

Sanchez, A., & Favara, M. (2019). Consequences of Teenage Childbearing in Peru. *Young Lives Working Paper*, *185*.

Santos, M. P., Turner, B., & Chaparro, M. P. (2021). The double burden of malnutrition in Peru: An update with a focus on social inequities. *The American Journal of Clinical Nutrition*, *113*(4), 865–873.

Schore, A. N. (2001). Effects of a secure attachment relationship on right brain development, affect regulation, and infant mental health. *Infant Mental Health Journal: Infancy and Early Childhood*, *22*(1–2), 7–66.

Schwarzer, R., & Jerusalem, M. (1995). Optimistic Self-Beliefs as a Resource Factor in Coping with Stress. In S. E. Hobfoll & M. W. de Vries (Eds.), *Extreme Stress and Communities: Impact and Intervention* (pp. 159–177). Springer Netherlands.

Sekhon, J. S., & Titiunik, R. (2017). On interpreting the regression discontinuity design as a local experiment. *Advances in Econometrics*, *38*, 1–28.

Sevim, D., Baranov, V., Bhalotra, S. R., Maselko, J., & Biroli, P. (2023). *Socioemotional skills in early childhood: Evidence from a maternal psychosocial intervention*. IZA Discussion Papers.

Shah, M., & Steinberg, B. M. (2017). Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital. *Journal of Political Economy*, *125*(2), 527–561.

Shoji, M. (2023). Gendered effects of early childhood weather shocks on locus of control: Evidence from 28 agricultural countries. *Journal of Population Economics*, *36*(3), 1363–1393.

Silventoinen, K. (2003). Determinants of variation in adult body height. *Journal of Biosocial Science*, *35*(2), 263–285.

Skinner, M. K. (2008). What is an epigenetic transgenerational phenotype? F3 or F2. *Reproductive Toxicology*, *25*(1), 2–6.

Skinner, M. K. (2014). Environmental stress and epigenetic transgenerational inheritance. *BMC Medicine*, *12*(1), 153.

Steelman, L. C., Powell, B., Werum, R., & Carter, S. (2002). Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges. *Annual Review of Sociology*, *28*, 243–269.

Stein, A. D., & Lumey, L. (2000). The Relationship between Maternal and Offspring Birth Weights after Maternal Prenatal Famine Exposure: The Dutch Famine Birth Cohort Study. *Human Biology*, *72*(4), 641–654.

Tafere, K. (2017). *Inter-generational Effects of Early Childhood Shocks on Human Capital: Evidence from Ethiopia*.

Thai, T. Q., & Falaris, E. M. (2014). Child Schooling, Child Health, and Rainfall Shocks: Evidence from Rural Vietnam. *The Journal of Development Studies*, *50*(7), 1025–1037.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*(6), 309–317.

Thompson, O. (2014). Genetic mechanisms in the intergenerational transmission of health. *Journal of Health Economics*, *35*, 132–146.

Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, *113*(485), F3–F33.

Todd, P. E., & Wolpin, K. I. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, *1*(1), 1–46.

Trinh, T.-A., Feeny, S., & Posso, A. (2021). Rainfall shocks and child health: The role of parental mental health. *Climate and Development*, *13*(1), 34–48.

Tuan, T., Harpham, T., & Huong, N. (2004). Validity and Reliability of the Self-reporting Questionnaire 20 Items in Vietnam. *Hong Kong Journal of Psychiatry*, *14*(3), 15–18.

USAID. (2011). *Peru climate change vulnerability and adaptation desktop study*.

Van Den Berg, G. J., & Pinger, P. R. (2016). Transgenerational effects of childhood conditions on third generation health and education outcomes. *Economics & Human Biology*, *23*, 103–120.

VanderWeele, T. J. (2009). Mediation and mechanism. *European Journal of Epidemiology*, *24*(5), 217–224.

VanderWeele, T. J. (2011). Controlled direct and mediated effects: Definition, identification and bounds. *Scandinavian journal of statistics, theory and applications*, *38*(3), 551–563.

VanderWeele, T. J. (2012). Mediation analysis with multiple versions of the mediator. *Epidemiology*, *23*(3), 454–463.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327.

Venkataramani, A. S. (2011). The intergenerational transmission of height: Evidence from rural Vietnam. *Health Economics*, *20*(12), 1448–1467.

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (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., Beguería, S., Lorenzo-Lacruz, J., Camarero, J. J., López-Moreno, J. I., Azorin-Molina, C., Revuelto, J., Morán-Tejeda, E., & Sanchez-Lorenzo, A. (2012). Performance of Drought Indices for Ecological, Agricultural, and Hydrological Applications. *Earth Interactions*, *16*(10), 1–27.

Walker, S. P., Chang, S. M., Wright, A. S., Pinto, R., Heckman, J. J., & Grantham-McGregor, S. M. (2022). Cognitive, psychosocial, and behaviour gains at age 31 years from the Jamaica early childhood stimulation trial. *Journal of Child Psychology and Psychiatry*, *63*(6), 626–635.

Webb, D. (2024). Critical Periods in Cognitive and Socioemotional Development: Evidence from Weather Shocks in Indonesia. *The Economic Journal*, *134*(660), 1637–1665.

WHO. (1995). *Physical status: The use and interpretation of anthropometry. Report of a WHO Expert Committee.* (No. 854).

WHO. (2006). Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. *Acta Paediatrica*, *95*, 38–46.

WHO. (2017). *The double burden of malnutrition*. World Health Organization. Geneva.

Yi, J., Heckman, J. J., Zhang, J., & Conti, G. (2015). Early Health Shocks, Intra-Household Resource Allocation and Child Outcomes. *The Economic Journal*, *125*(588), F347–F371.

Yorke, L., & Ogando Portela, M. J. (2018). Psychosocial scales in the Young Lives Round 4 Survey: Selection, adaptation and validation. *Young Lives Technical Note*, *45*.

Youngson, N. A., & Whitelaw, E. (2008). Transgenerational Epigenetic Effects. *Annual Review of Genomics and Human Genetics*, *9*, 233–257.

Zang, E., Tan, P. L., & Cook, P. J. (2023). Sibling Spillovers: Having an Academically Successful Older Sibling May Be More Important for Children in Disadvantaged Families. *American Journal of Sociology*, *128*(5), 1529–1571.

Zimmermann, L. (2020). Remember when it rained – Schooling responses to shocks in India. *World Development*, *126*, 104705.