Essays on Climate Shocks, Food Security, and Human Capital Formation in Ethiopia

Seyoum Teffera Mengesha

Doctoral School of Economics, Business, and Informatics

Supervisor:

Éva Berde, PhD

Corvinus University of Budapest

© Seyoum Teffera Mengesha

Corvinus University of Budapest

Doctoral School of Economics, Business, and Informatics

Essays on Climate Shocks, Food Security, and Human Capital Formation in Ethiopia

Doctoral Dissertation

Seyoum Teffera Mengesha

Budapest, Hungary, 2026

Table of Contents

ACKNOWLEDGEMENTS 4

INTRODUCTION 5

CHAPTER 1: HOUSEHOLD FOOD INSECURITY AND ADOLESCENTS' EDUCATIONAL ASPIRATIONS IN ETHIOPIA: A LONGITUDINAL ANALYSIS 9

Abstract 9

1. Introduction 11

2. Materials and Methods 15

3. Results 27

4. Discussion 38

5. Conclusions and Implications 45

CHAPTER 2: CLIMATE VULNERABILITY AND CHILD STUNTING IN ETHIOPIA: A SPATIO-TEMPORAL MULTILEVEL ANALYSIS USING DHS DATA, 2011–2019 59

Abstract 59

1. Introduction 61

2. Materials and Methods 66

3. Results 75

4. Discussion 86

5. Conclusion and Implications 95

CHAPTER 3: AGRICULTURAL EXTENSION AS CLIMATE INSURANCE: PANEL EVIDENCE FROM ETHIOPIAN SMALLHOLDERS 108

Abstract 108

1. Introduction 109

2. Materials and Methods 113

3. Results 122

4. Discussion 138

5. Conclusion and Implications 146

CONCLUSION AND POLICY SYNTHESIS 158

REFERENCES 163

List of Figures

Figure 1. Sample Construction and Attrition Flow 17
Figure 2. Distribution of educational aspirations by survey round 28
Figure 3. Comparison of Estimated Associations Between Food Insecurity and Educational Aspirations Across Models. 34
Figure 4. Sample Selection Flowchart 66
Figure 5. Regional climate vulnerability trajectories and exposure composition by survey year. 78
Figure 6. ORs for Climate Vulnerability and Individual-Level Covariates in the Primary Stunting Model 80
Figure 7. Sample Selection Process 114
Figure 8. Spatial Distribution of Rainfall Anomalies Across Survey Waves 126
Figure 9. Harvest Trajectories by Treatment Pattern, 2011/12–2015/16 129
Figure 10. Event Study: Harvest Coefficients Relative to First Extension Participation .. 131

List of Tables

Table 1. Sample Characteristics and Regional Distribution by Survey Round 27
Table 2. Food Insecurity Summary by Survey Round (HFIAS Score and HFIAP Category) 27
Table 3. Summary statistics of educational aspirations by survey round 29
Table 4. Within-Child Dynamics (Round 4 to Round 5) 29
Table 5. Variance Decomposition of Key Variables 30
Table 6. Cross-Sectional and Value-Added Associations 31
Table 7. Propensity Score Matching Estimates 32
Table 8. Within-Child Fixed-Effects Associations and First-Difference Verification 33
Table 9. Cognitive Pathway in the Food Insecurity–Aspiration Association 35
Table 10. Sample Characteristics by Survey Wave 76
Table 11. Regional Climate Vulnerability and Stunting Prevalence 77
Table 12. Temporal Interaction Model Results (Primary Model) 80
Table 13. Vulnerability Effects Across Livelihood Zones 82
Table 14. Association between standardized CVI and child stunting by FEWS NET livelihood-zone type, with region fixed effects. 82
Table 15. Descriptive Statistics and Extension Coverage 124
Table 16. Rainfall and Drought Incidence by Survey Wave and Location 126
Table 17. Outcome Trajectories and Pre-Trend Comparisons by Treatment Pattern 129

Table 18. Estimates of Lagged Extension Participation on Agricultural Outcomes..... 132
Table 19. Lagged Extension Participation Effects by Drought Conditions 132
Table 20. Propensity Score Matching: Sample Sizes and Covariate Balance 135
Table 21. Fixed Effects Estimates for Full and Propensity Score Matched Samples 136

ACKNOWLEDGEMENTS

First and foremost, I give thanks to Almighty *God*; without *His grace*, none of this would have been possible.

I am profoundly grateful to my supervisor, *Dr. Éva Berde*, whose *invaluable guidance and unwavering support* were instrumental throughout my doctoral journey. I also wish to thank *Dr. Anna Naszodi* for her *kind cooperation and insightful guidance*, which significantly contributed to the success of this dissertation.

I am thankful to the *faculty and staff* of the *Doctoral School of Economics, Business and Informatics (DSEBI)* at Corvinus University of Budapest for providing such an *intellectually stimulating environment*.

I extend my sincere gratitude to the *Hungarian Government*, the *Tempus Public Foundation* and the *Hungarian people* for awarding me the *Stipendium Hungaricum (SH)* Scholarship.

On a personal note, I express my deepest gratitude to *my family* for their constant *love* and *support* throughout this long journey. Last but not least, I thank my *fellow PhD students* for their *collegiality, shared experiences, and mutual support*.

INTRODUCTION

Human capital formation in developing economies represents an intertemporal investment process persistently disrupted by exogenous shocks. This concept, formalized by Schultz (1961) and Becker (1964) in the foundational human capital literature recognizes that investments in education, health, and nutrition are analogous to investments in physical capital, yielding returns over extended time horizons. In the Ethiopian context, where the structural transformation from a rain-fed agrarian society to a climate-resilient middle-income nation is the primary policy objective (FDRE, 2011a), this process reveals a fundamental tension between institutional expansion and environmental volatility.

Over the past two decades, Ethiopia has achieved a substantial expansion in educational access, exemplified by primary enrollment rates surging toward universality (Woldehanna & Araya, 2016) and a significant reduction in national stunting prevalence from 58% in 2000 to 37% by 2019 (EPHI & ICF, 2021). However, as this dissertation demonstrates, these aggregate successes often obscure a persistent pattern of household-level vulnerabilities in which climate and nutritional shocks threaten developmental trajectories, even as households maintain stable long-term aspirations.

This dissertation examines human capital formation across three complementary analytical dimensions: the psychological foundations of educational aspirations (Chapter 1), the biological vulnerabilities of child nutrition under climate stress (Chapter 2), and the institutional buffers protecting agricultural productivity during drought (Chapter 3). This research employs multiple theoretical frameworks, recognizing that human capital development operates through distinct mechanisms (psychological, physiological, and productive), each responding differently to material shocks and institutional interventions. While grounded primarily in the Beckerian human capital framework (Becker, 1964; Mincer, 1974), this dissertation also draws on Sen's capability approach (Sen, 1985, 1999) to understand how poverty and shocks affect not only resource constraints but also individuals' freedom to achieve valued functionings, that is, the states of being and doing they have reason to value, including educational attainment and health. The unifying methodological approach across all three essays is the rigorous application of panel data methods to distinguish causal effects from selection bias, leveraging the longitudinal depth of the Young Lives Ethiopia

study, the spatial breadth of the pooled cross-section Demographic and Health Surveys, and the detailed agricultural production longitudinal data from the Ethiopia Socioeconomic Survey.

The three essays advance a unified argument: while human capital aspirations in Ethiopia display remarkable psychological resilience to short-term shocks, the biological and productive foundations necessary to realize these aspirations remain acutely vulnerable to climate stress. Institutional interventions can partially bridge this gap, but only when designed as contingent buffers rather than unconditional growth enhancers. This reveals a fundamental tension in Ethiopia's human capital development: aspirations persist at stable levels even when the material capacity to fulfill them deteriorates under climate pressure.

The inquiry begins with the psychological dimension: the adolescent's capacity to aspire (Appadurai, 2004; Ray, 2006). A significant body of behavioral economics literature suggests that the cognitive burden of poverty should erode long-term educational goals. Banerjee and Duflo (2011) argue that poverty's immediate pressures overwhelm the cognitive resources required for long-term planning, a mechanism supported by neuroscience research linking stress to reduced prefrontal cortex function. This hypothesis was later empirically demonstrated in economic contexts by Mani et al. (2013) and Mullainathan and Shafir (2013), who demonstrated that the scarcity mindset significantly depletes cognitive bandwidth, leaving fewer mental resources for long-term human capital investments. However, by applying individual child fixed-effects models to the Younger cohort Young Lives panel in Chapter 1, I find evidence of aspirational resilience. The analysis suggests that the apparent relationship between food insecurity and educational aspirations is primarily driven by stable household characteristics rather than transitory shocks. Ethiopian adolescents maintain stable educational goals regardless of short-term resource fluctuations, a pattern consistent with Sen's (1985, 1999) distinction between capabilities (what people can actually achieve) and aspirations (what they wish to achieve), a gap that material and institutional constraints prevent from closing. This implies that the psychological demand for education is consistent; the bottleneck to educational mobility is therefore not a lack of ambition but the material and institutional constraints that prevent its realization.

However, while aspirations remain resilient to transitory shocks, the biological foundation of human capital (represented by physical growth and nutritional status) demonstrates acute

vulnerability to environmental stress. Chapter 2 examines this biological dimension, analyzing how multidimensional climate vulnerability affects child stunting across Ethiopia's diverse agroecological zones. Following the Intergovernmental Panel on Climate Change (IPCC, 2014b) framework for climate vulnerability assessment, I construct a Climate Vulnerability Index that integrates three components: exposure (rainfall and temperature anomalies), sensitivity (household water access, sanitation, rural residence, female household head and birth spacing), and adaptive capacity (maternal education and wealth). I link this index to high-resolution geolocated climate data and anthropometric records from three Ethiopia Demographic and Health Surveys (2011–2019). The analysis demonstrates that climate vulnerability significantly increases stunting risk, and this association is not attributable to a strengthening of climate exposure itself over time. The temporal interaction is instead concentrated in adaptive capacity and is not robust to design-based standard errors, even as average stunting rates have declined over the study period. This biological constraint represents a fundamental limitation on human capital returns (Becker, 1964; Schultz, 1981), as cognitive development and school performance are directly compromised by early-life nutritional deficits.

The contrast between aspirational stability and biological vulnerability raises a question: how can households translate their stable educational goals into actual attainment when the material and physical foundations are so fragile? Chapter 3 addresses this question by examining the role of state-led institutional support in buffering climate risk. Using propensity score matching and household fixed effects on the Ethiopia Socioeconomic Survey panel (2011–2016), the analysis demonstrates that agricultural extension functions as shock-contingent insurance: while showing no average productivity effect during normal years, it substantially offsets drought-induced harvest losses during severe rainfall deficits. No significant heterogeneity in this buffering effect was found by household size, education, or credit access. By stabilizing agricultural income during climate shocks, extension services may help protect the nutritional foundation identified as vulnerable in Chapter 2.

Taken together, these three essays reveal a multifaceted pattern in which psychological resilience coexists with biological vulnerability, and where institutional buffering can partially bridge the gap between stable aspirations and fragile material foundations.

These findings carry significant policy implications for Ethiopia's development strategy. The combination of aspirational resilience, biological vulnerability, and the shock-contingent effectiveness of institutional support suggests that human capital policy must operate simultaneously on multiple margins: expanding educational infrastructure to match existing demand, intensifying nutritional interventions in climate-vulnerable regions, and scaling state agricultural support systems that function as implicit social insurance. The concluding policy synthesis integrates these insights into actionable recommendations for Ethiopia's structural transformation agenda.

This dissertation proceeds as follows. Chapter 1 analyzes the relationship between food insecurity and educational aspirations among Ethiopian adolescents. Chapter 2 examines how climate vulnerability affects child stunting across Ethiopia's diverse agroecological zones. Chapter 3 examines agricultural extension as a form of climate insurance for smallholder farmers. Each chapter is structured as a self-contained research paper with its own abstract, introduction, materials and methods, results, discussion, and conclusion. A unified bibliography consolidating all references from the three chapters follows the main text.

CHAPTER 1: HOUSEHOLD FOOD INSECURITY AND ADOLESCENTS' EDUCATIONAL ASPIRATIONS IN ETHIOPIA: A LONGITUDINAL ANALYSIS

Abstract

Although Ethiopia has experienced dramatic educational expansion, food insecurity remains widespread, and its role in shaping adolescents' educational aspirations is poorly understood. Using longitudinal data from the Younger Cohort of the Young Lives Ethiopia study (Rounds 1–5, 2002–2016), this paper examines the association between household food insecurity and adolescents' educational aspirations. The analytical sample comprises 1,873 adolescents aged 12 and 15, providing up to 3,644 child-round observations. The study employs both cross-sectional analysis and panel fixed-effects models. Standard cross-sectional models show a significant negative association between food insecurity and educational aspirations ($\beta = -0.0593$, SE = 0.0172, $p = 0.002$), and value-added models (VAM) yield similarly significant but attenuated estimates. However, once time-invariant child and household characteristics are accounted for using child fixed-effects models, the association between changes in food insecurity and changes in educational aspirations is statistically indistinguishable from zero ($\beta = -0.0245$; SE = 0.0497; 95% CI [−0.122, 0.073]). This pattern is consistent across wealth strata. These findings suggest that adolescents' educational aspirations may be resilient to short-term changes in household food insecurity and that the observed cross-sectional gradient may largely reflect persistent differences in socioeconomic and household circumstances rather than a within-child response to changing food insecurity. One possible interpretation is that adolescents and their families view education as a pathway out of deprivation, sustaining high educational aspirations despite material hardship. The findings therefore suggest that policies seeking to support educational aspirations may need to address broader and more persistent conditions, including early childhood nutrition, household resources, and school quality, rather than relying primarily on adolescent feeding interventions to shift aspirations.

Keywords: Adolescence; Educational Aspirations; Ethiopia; Fixed Effects; Food Insecurity; Panel Data.

1. Introduction

Economic development remains a central objective for developing countries, as it is closely associated with higher productivity, increased employment, improved living standards, and greater economic opportunities. Education plays a fundamental role in achieving these outcomes by contributing to the accumulation of human capital. This view is consistent with the human capital framework, which views education as an investment that enhances individuals' skills and productivity, thereby increasing their potential earnings (Becker, 1964). Education is also widely regarded as an important driver of intergenerational social mobility, as educational attainment can influence the opportunities available to individuals relative to those of their parents. However, recent empirical evidence covering 153 countries shows that both absolute and relative intergenerational mobility in education are markedly lower in developing countries, where educational persistence across generations remains high (van der Weide et al., 2024). Within this broader context, Ethiopia provides an important setting for examining adolescents' educational aspirations and the factors that may influence their formation.

Between 2002 and 2016, Ethiopia experienced a dramatic expansion in education. Gross primary enrollment grew from 52% to over 100%[1]. Net enrollment, which counts only age-appropriate children, remained lower at approximately 84%. Simultaneously, the higher education sector expanded from two public universities in the early 1990s to 31 public universities and numerous private institutions by 2014 (Woldehanna & Araya, 2016; World Bank, 2018). This was coincided with a nominal GDP growth averaging 10% annually between 2004 and 2014 (World Bank Group, 2015). Accompanying this growth was a profound cultural shift, in which families increasingly viewed schooling not merely as an option, but as the primary pathway to economic survival and advancement (Tafere, 2014; Woldehanna & Araya, 2016).

However, the capacity of youth to translate expanded educational opportunities and shifting cultural expectations into economic success depends on demand-side factors, particularly educational aspirations that shape effort, investment, and persistence. Yet aspirations do not

[1] This figure can exceed 100% because it includes all enrolled students regardless of age while the denominator includes only the official primary school-age population.

form in a vacuum. They require a sufficiently nourished body and brain. Unfortunately, this biological foundation has been structurally compromised. Ethiopia has consistently ranked among the world's ten largest food crises by the number of people facing acute food insecurity, and is among the 36 countries that together account for roughly 90 percent of the world's stunted children (GNAFC, 2024; Sunil Rajkumar et al., 2012).

Understanding the relationship between food insecurity and children's educational aspirations is therefore particularly important for Ethiopia's human capital development and broader development strategy. Through the Seqota Declaration and National Nutrition Program II, the government has committed to ending child undernutrition by 2030, explicitly linking nutritional investments to long-term educational outcomes (FDRE Ministry of Health, 2015, 2024). The underlying assumption is that improving nutrition will enhance children's capacity to learn and aspire. However, whether nutritional improvements translate into higher educational aspirations remains inadequately tested during adolescence, a pivotal period when educational goals solidify and critical transitions from primary to secondary schooling take place.

Educational aspirations refer to the level of schooling that children hope to achieve (O'Donnell et al., 2025). In standard human capital theory, individuals invest in education by comparing expected returns against costs (Becker, 1964). Educational aspirations represent a key potential pathway through which poverty-related constraints may shape investment decisions. A child who perceives university as beyond their reach may reduce effort toward prerequisites, regardless of cognitive ability. Importantly, aspirations during early adolescence represent a leading indicator that precedes and shapes attainment. Studying aspirations at the formation stage (ages 12 to 15) rather than post-hoc attainment, allows examination of whether nutritional conditions are associated with educational goals before they translate into consequential investment decisions.

Empirical evidence suggests that aspirations are associated with subsequent educational attainment and labor-market outcomes, even after accounting for cognitive ability and family background (Genicot & Ray, 2017; Ray et al., 2006). Evidence from Ethiopia points in a similar direction, with forward-looking aspirations shown to predict agricultural investment and subsequent income growth (Bernard et al., 2014), while analyses of Young Lives data indicate that aspirations formed during adolescence are significant predictors of later

educational attainment (Favara, 2017; Serneels & Dercon, 2021). However, aspirations are strongly stratified by household resources, with 84% of children from the richest households aspiring to complete university, compared with only 67% of those from the poorest households (Favara, 2017). Moreover, substantial attrition occurs between aspiration and attainment, with only 61% of the cohort remaining enrolled in any education by age 19. These patterns raise the question of whether material deprivation, and particularly food insecurity, constrains adolescents' educational aspirations even amid unprecedented educational expansion.

Answering this question requires engaging two competing theoretical frameworks, each offering a distinct prediction about the potential effect of food insecurity on aspirations. First, the *scarcity hypothesis*, grounded in behavioral economics, posits that food insecurity directly constrains aspirations through cognitive bandwidth and agency mechanisms. Mullainathan and Shafir (2013) demonstrate that scarcity of any kind including food scarcity, which they use as a founding illustration of the theory through the *Minnesota Starvation Experiment*, captures cognitive attention and reduces the bandwidth available for long-term planning. Because aspiration formation requires this future-oriented cognitive capacity, food insecurity taxes the mental resources through which children construct and pursue educational goals, while chronic nutritional deficits can impair executive function and future-oriented thinking, which are essential for aspiration formation (Prado & Dewey, 2014). Furthermore, the psychological uncertainty associated with hunger can undermine self-efficacy, leading children to rationally lower their aspirations to align with constrained circumstances, a phenomenon known as *adaptive preferences* (Bandura, 1997; Sen, 1999). Under this framework, reducing food insecurity should increase educational aspirations, resulting in measurable within-child effects.

Second, the *structural selection hypothesis* posits that the association between food insecurity and aspirations is not causal, but rather reflects selection on stable structural characteristics. Under this framework, time-invariant factors such as parental human capital, institutionalized gender bias (Dercon & Singh, 2013), and the cultural norms that define a child's navigational map (Appadurai, 2004) jointly determine both household food security and educational goals. If this hypothesis is correct, aspirations are anchored in persistent structural contexts and should show limited within-child change when a child experiences

temporary nutritional shocks. This can generate a cross-sectional association that reflects stable differences between children and households and may attenuate or disappear under fixed-effects estimation. While Dercon and Singh (2013) and Appadurai (2004) establish that structural factors and gender norms set the baseline for aspirations, they do not address whether aspirations remain responsive to short-term material shocks like food insecurity within the same child.

These hypotheses yield testable predictions. The scarcity hypothesis predicts that within-child reductions in food insecurity should increase aspirations, consistent with the possibility that nutrition interventions could influence educational goals. The structural selection hypothesis predicts little or no within-child association in fixed-effects models despite strong cross-sectional associations. This suggests that if aspirations are anchored in early formed structural factors rather than transient nutritional shocks, improving food security may benefit health and cognitive development but will not shift aspirations without concurrent investments in school quality, economic opportunities, and pathways for social mobility. The present study examines which of these competing patterns is more consistent with the data.

It is important to note that, to our knowledge, no prior study has examined whether changes in food insecurity predict changes in educational aspirations among adolescents in Ethiopia or a comparable Sub-Saharan African context, using longitudinal methods that account for the panel structure of the data and the specific institutional and socioeconomic context. This study fills this gap by leveraging longitudinal data from the Young Lives Ethiopia younger cohort, which follows 1,873 children through the critical transition from primary to secondary schooling. The study focuses on the pivotal period between ages 12 and 15, when abstract hopes translate into concrete educational investments and family decisions about continued schooling.

The study addresses three specific research questions.

The first asks whether household food insecurity is associated with children's educational aspirations during adolescence (RQ1). This is examined using within-child fixed effects models that compare each child to themselves across survey rounds, thereby accounting for time-invariant differences between children and households. The second examines how structural inequalities, such as household wealth and rural or urban location, are associated with differences in educational aspirations and whether these differences persist

independently of within-child changes in food insecurity (RQ2). This is addressed through cross-sectional and stratified specifications that trace how the food insecurity coefficient changes as structural controls are introduced. The third investigates whether cognitive ability, measured by the Peabody Picture Vocabulary Test (PPVT), may account for part of the association between food insecurity and educational aspirations (RQ3). It further distinguishes whether this pattern reflects between-family differences or within-child responses to changes in food insecurity, using mediation analysis combined with variance decomposition and within-child cognitive models.

2. Materials and Methods

2.1 Data Source and Study Design

This study utilizes data from the Young Lives study, a longitudinal research project following 12,000 children across Ethiopia, India (Telangana and Andhra Pradesh), Peru, and Vietnam since 2002 (Outes-Leon & Sanchez, 2008). The project follows two cohorts in Ethiopia, comprising approximately 1,000 children in the older cohort and 2,000 in the younger cohort. The present analysis focuses exclusively on the younger Ethiopian cohort to leverage its larger sample size to enhance statistical precision. The younger cohort was first surveyed in Round 1 (2002), when the children were approximately one year old. They were subsequently followed at ages 5 (Round 2, 2006), 8 (Round 3, 2009), 12 (Round 4, 2013), and 15 (Round 5, 2016) (Pankhurst et al., 2018). In Rounds 1 and 2, data were collected exclusively from the primary caregiver, typically the mother, as the children were too young to be interviewed[2]. From Round 3 onwards, when the children were 8 years old, they were also interviewed directly, providing child-level data, including cognitive assessments.

The analysis focuses on the Younger Cohort's transition from early adolescence[3] (age 12) to mid-adolescence (age 15), a period that is both analytically and educationally significant in the Ethiopian context, as it captures the transition from primary to lower secondary education.

[2] Each survey round comprised multiple questionnaires administered at the household, caregiver, and community levels, covering food security, household wealth, household composition, and local service availability.

[3] The World Health Organization (WHO) defines adolescents as individuals aged 10–19 years and youth as individuals aged 15–24 years (WHO Regional Office for Africa, 2018) . We use the term “adolescence” rather than “youth” because the study period (ages 12–15 years) falls entirely within the adolescent age range.

Students ideally complete Grade 8 by age 14 and enter Grade 9 at age 15, though late enrollment and slow grade progression mean they often fall significantly behind expectations (Woldehanna et al., 2017). Consequently, aspirations measured during this window capture the shift from abstract childhood hopes to concrete, constraint-aware educational goals. Children's own educational aspirations were collected from Round 4 (age 12), making this interval well-suited to examining how aspirations evolve as children move from primary toward secondary school.

2.2 Sampling Strategy and Site Selection

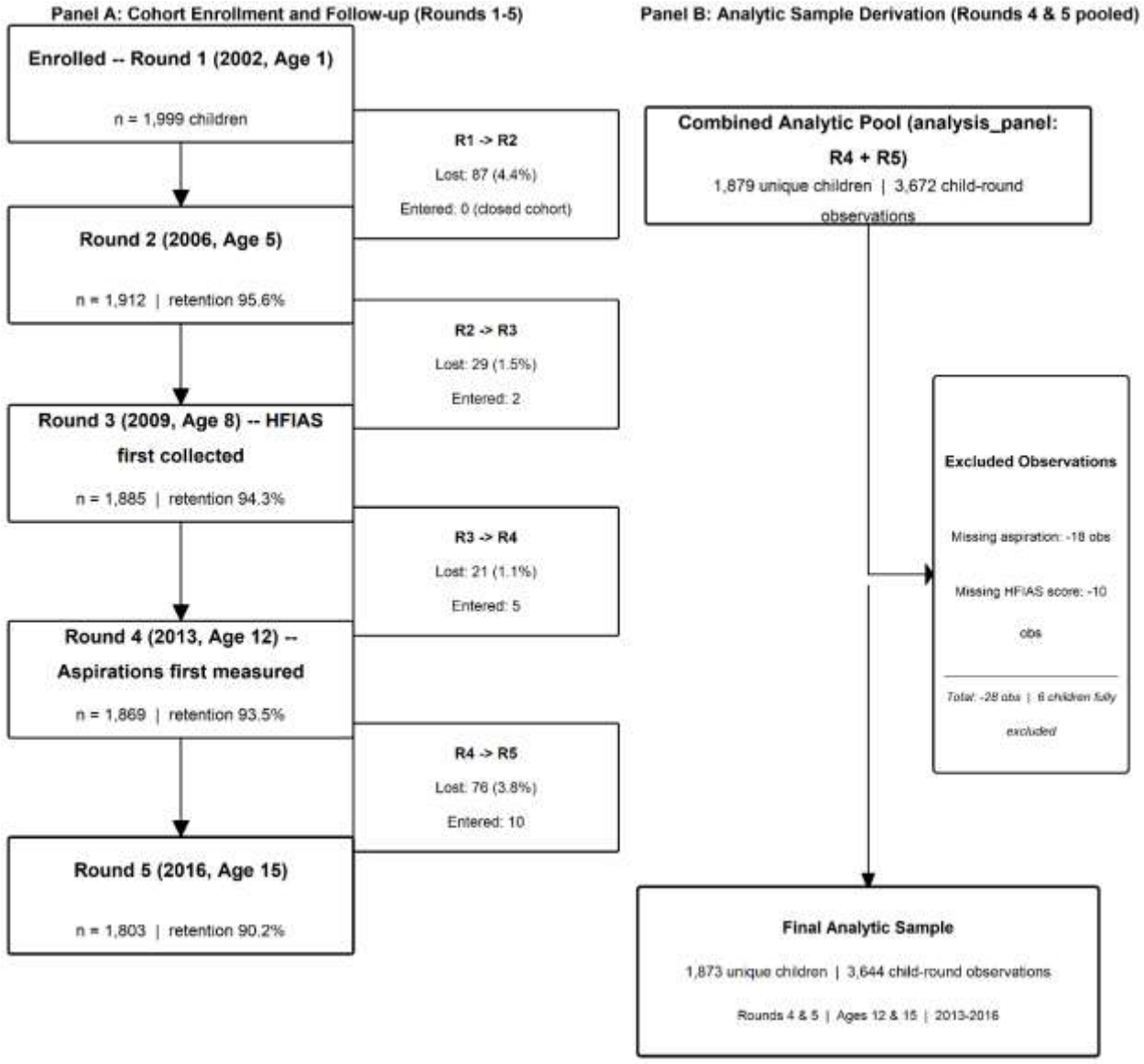
Young Lives employed a sentinel site sampling approach, selecting 20 sites across Ethiopia to represent diverse ecological zones, socioeconomic conditions, and demographic contexts (Outes-Leon & Sanchez, 2008; Pankhurst et al., 2018). The sites were selected through a systematic three-stage process based on national administrative structures, with regions selected first, followed by districts (woredas) and then local administrative areas (kebeles). To protect participant confidentiality, all site names are represented using Amharic food-related pseudonyms. The sites encompass both urban and rural areas across five administrative regions, including Addis Ababa (3 sites), Amhara (4 sites), Oromia (4 sites), SNNP (5 sites), and Tigray (4 sites). The analysis is restricted to these five main regions, excluding 13 observations with unresolved region assignments. Of these, 4 occur during the Round 3–4 transition and 9 during the Round 4–5 transition. *Five* children resided in regions outside the five main sentinel-site regions (Afar, Benishangul-Gumuz, and Somali), while the remaining *8* could not be assigned to a region because their household records could not be matched.

The 20 sentinel sites were selected with balanced representation of poverty levels, urban/rural areas, and food deficit status (Young Lives, 2018). At each site, approximately 100 households with age-appropriate children were randomly selected, yielding an initial sample of 2,000 children (Sanchez & Outes-Leon, 2008), though the exact sample size in the public dataset is 1,999.

Of the original 1,999 children enrolled at baseline (2002), 1,803 (90.2%) were followed through Round 5 (2016) (*see Table 1 for details*). Educational aspiration data were available for 1,859 children at age 12 (Round 4) and 1,795 at age 15 (Round 5). We restricted the

sample to observations with complete data on aspirations, food insecurity, and the key covariates mentioned below. The final analytic sample consisted of 3,644 child-round observations from 1,873 unique children. The sample construction process, including cohort enrollment, follow-up rounds, and observation-level exclusions due to missing data, is summarized in Figure 1.

Figure 1. Sample Construction and Attrition Flow



Although overall retention was high (90.2% over 15 years), high retention does not preclude selective attrition. We therefore assessed whether participants lost to follow-up differed systematically from completers at baseline (*Appendix Table A1*). Participants lost to follow-

up were more likely to reside in urban areas than completers (46.4% vs. 33.7%; $\chi^2 = 12.98$, $p < 0.001$), indicating that the unweighted analytic sample may underrepresent urban households. To account for this differential attrition, inverse probability weights were constructed using urbanicity, region, and household size and applied in all weighted analyses. Full details of the attrition analysis, baseline characteristics by completion status, and the logistic regression used to construct attrition weights are reported in supplementary Tables A1–A3.

2.3 Study Variables and Measurement

The study includes three categories of variables: the child's educational aspirations as the primary outcome, household food insecurity as the primary exposure, and a set of control variables comprising cognitive ability, household wealth, urban residence, stunting status, gender, and household size. Cognitive ability is included both as a control variable and as a potential mediator in the association between household food insecurity and educational aspirations. Across all analyses, the primary focus is on estimating the association between household food insecurity and educational aspirations. The operationalization and measurement of each variable are described below.

Educational Aspiration Measure

Educational aspirations are measured using children's responses to the question: "How many years of schooling do you hope [CHILD] will complete?" This self-report was collected directly from children in Rounds 4 and 5, when they were approximately 12 and 15 years old, respectively. We focus on aspirations rather than completed educational attainment for four reasons. First, aspirations develop earlier in childhood and capture motivational and cognitive processes that may be associated with educational decision-making under material constraints. They therefore provide a more direct measure for examining the scarcity hypothesis introduced in Section 1. Second, this age range captures a critical period in which school enrollment remains high while educational decisions become increasingly salient. Enrollment is 98% at age 12 and 93% at age 15, before the substantial decline observed by age 19, when only 61% of children remain enrolled (Favara, 2017). Third, aspirations are measured relatively close to the period in which household food insecurity is observed,

whereas completed educational attainment reflects experiences and circumstances accumulated over a longer period. This makes aspirations a more temporally proximate indicator of children's educational orientation. Fourth, measuring aspirations avoids right-censoring because most children in the sample have not yet completed their education at the time of measurement.[4]

Food Insecurity Measures

We measured food insecurity using the Household Food Insecurity Access Scale (HFIAS), a validated tool developed by USAID's Food and Nutrition Technical Assistance (FANTA) III project (Coates et al., 2007). The measure is restricted to Rounds 3, 4, and 5, as standardized HFIAS items were first included from Round 3 onwards. All nine standard HFIAS occurrence items and their corresponding frequency of occurrence follow-up questions were administered in each survey round. However, consistent with the annual Young Lives survey design, respondents reported food insecurity experiences over a *12-month recall period* rather than the four-week period specified in the original HFIAS methodology.

The Young Lives survey applied skip-routing rules that differ from the standard HFIAS protocol.[5] In Rounds 4 and 5, households reporting "No" to all four basic access items (Q1–Q4) were routed past Q5–Q7 entirely. Among households who reached Q5–Q7, items Q8 and Q9 were conditional on Q7=Yes. Missing values from both routing rules were imputed as

[4] A sixth round of the Young Lives Ethiopia survey was conducted by telephone in 2020/21, when cohort members were approximately 19–20 years old. Secondary school completion would generally be expected to be observable by this age, particularly given that 92.7% of participants remained enrolled at Round 5. Round 6 is nevertheless not incorporated here because its design and context differ substantially from the preceding five rounds. First, the shift from in-person to telephone interviewing may introduce measurement differences in self-reported educational outcomes, limiting comparability across rounds. Second, fieldwork coincided with COVID-19-related school closures and the Northern Ethiopia war, creating major contemporaneous disruptions to schooling that would complicate attribution of Round 6 educational outcomes to food insecurity experienced in earlier periods. Third, telephone follow-up may introduce selection because phone access is socially and economically patterned, potentially increasing the representation of urban and wealthier households among respondents. Finally, attrition may itself be selective: cohort members from households experiencing greater early food insecurity could have been more difficult to recontact because of migration, household changes, or marriage. Round 6 therefore warrants separate analysis that explicitly addresses these potential sources of measurement and selection bias.

[5] Under the standard HFIAS protocol, all nine occurrence questions (Q1–Q9) are administered to every household in sequence. Skipping applies only to the frequency sub-question paired with a "No" occurrence response (e.g., a "No" response to Q1 skips only Q1a, not Q2–Q9). The protocol does not provide for skipping blocks of items based on earlier responses, or for making later items conditional on a specific prior response.

zero after validation confirmed that the large majority of skip-routed households had zero scores on all preceding items (*see Appendix A.1*).

Frequency-of-occurrence follow-up responses (rated 1=rarely, 2=sometimes, 3=often) were collected for all nine items in all three rounds. We therefore compute the official HFIAS score. For each item, the frequency code is set to 0 if the occurrence response is "No", and to the recorded value (1=rarely, 2=sometimes, 3=often) if the occurrence response is "Yes". These nine codes are summed to produce a continuous score ranging from 0 to 27, with higher values indicating greater food insecurity severity. Households with a valid occurrence response of "Yes" on any item but a missing matching frequency response are coded as missing on the HFIAS score, which produces somewhat greater missingness than a simple occurrence-based count would (*see Appendix A.1*). The HFIAS score is the primary severity measure used in this study.

We additionally compute the official Household Food Insecurity Access Prevalence (HFIAP) classification, which assigns each household to one of four ordered categories (*Food Secure, Mildly, Moderately, or Severely Food Insecure*) based on a sequential set of rules applied to the occurrence and frequency responses, with the most severe matching condition determining the final category. Households reporting Q7 (no food of any kind), Q8 (sleeping hungry), or Q9 (a whole day without eating) at any frequency are classified as Severely Food Insecure, consistent with the original HFIAS specification that these three items indicate severe food insecurity regardless of how often they occur.

Within the Coates et al. (2007) quantity domain (Q5–Q9), we report one occurrence-based descriptive indicator, quantity reduction (Q5 or Q6 = Yes, regardless of frequency). This indicator captures the occurrence of quantity reduction but does not, by itself, measure its severity. Severity associated with Q5 and Q6 is instead captured by the HFIAP classification, which incorporates their frequency of occurrence, while severity associated with Q7–Q9 is captured by the HFIAP classification based on their occurrence alone.

To distinguish between transitory and chronic food insecurity, the study constructed a persistence ratio representing the proportion of analytic rounds (Rounds 4 and 5) in which a household was classified as moderately or severely food insecure according to the HFIAP classification. The resulting measure ranged from 0 to 1, with higher values indicating greater persistence of food insecurity over time. Round 3 data inform descriptive characterizations

of food insecurity trajectories but are not used to construct the persistence variable, consistent with the regression sample restriction described below. We additionally computed the mean HFIAS score across available rounds as an alternative continuous chronic-exposure measure.

Control Variables

To address RQ2, we include household wealth[6], urban residence, gender, household size, and stunting status. In the cross-sectional models discussed below, these variables capture the extent to which the raw food insecurity–aspiration association is explained by differences in household and individual characteristics. In the within-child fixed-effects models, time-invariant characteristics such as gender and urban residence are absorbed by the child fixed effect, while time-varying variables such as wealth, household size, and current stunting status are included to account for changes in household circumstances over time.

An important caveat applies to the wealth index specifically. The household wealth index may itself be partially determined by food insecurity. Households that are food insecure are likely also asset-poor. Including wealth alongside food insecurity in the same regression therefore risks a bad controls problem (Angrist & Pischke, 2009). The study therefore present specifications both with and without wealth controls. The model without wealth controls (Table 6, Model 1) is the preferred cross-sectional specification for interpreting the total association between food insecurity and aspirations. Models that add wealth are presented to show the degree of attenuation, but their food insecurity coefficients should be interpreted as partial effects conditional on wealth rather than total effects.

In the within-child fixed effects models, this concern is substantially mitigated because fixed effects difference out the stable component of wealth, and the residual within-child wealth variation captures short-term economic fluctuations rather than the structural poverty pathway. Full model specifications additionally adjust for household economic capacity (defined as the ability to raise emergency funds) and whether the household received financial support from social networks. Urban/rural residence was measured using the household-level 'typesite' variable, directly observed in Rounds 1 through 3. For Rounds 4–

[6] The household wealth index is a composite asset index constructed by the Young Lives study team from information on housing quality, ownership of consumer durables, and access to services (electricity, piped water, sanitation), standardized within each survey round to a 0–1 scale (Young Lives, 2018).

5, where this variable was unavailable in the public dataset, we assigned households their most recent observed classification (Round 3, or Round 2/1 if unavailable).

To capture early-life nutritional status and physiological capital, the study constructed height-for-age z-scores (HAZ) according to WHO growth standards (WHO, 2006). Following the standards, biologically implausible anthropometric values (HAZ < -6 or > 6 SD) were identified and excluded from the analysis. Children with a HAZ below -2 SD at age 1 were classified as "early stunted." Early stunting is treated as a time-invariant structural characteristic. It captures accumulated early-life nutritional disadvantages rather than current food insecurity. In cross-sectional models, it serves as an indicator of the depth of structural disadvantage that a child carries into adolescence.

Finally, to test whether cognitive ability mediates the food insecurity–aspiration relationship (RQ3), we incorporated standardized scores[7] from the Peabody Picture Vocabulary Test (PPVT) collected in Rounds 3, 4, and 5. These standardized scores were used to test whether cognitive ability mediated the relationship between food insecurity and aspirations (Table 9).

2.4 Analytical Approach

The analytical strategy proceeds in three stages, moving from descriptive associations to specifications that progressively account for observed and time-invariant differences between children. The first stage presents pooled OLS as a descriptive benchmark, estimating the cross-sectional association between household food insecurity and educational aspirations. The second stage employs value-added models (VAM) and propensity score matching (PSM) as intermediate specifications. VAM accounts for prior aspirations and observed covariates, while PSM compares children with similar observed characteristics across household food security groups. The third and primary stage employs within-child fixed-effects estimation, which uses variation within children over time to account for time-invariant child and household characteristics and examine whether changes in household food insecurity are associated with changes in adolescents' educational aspirations.

[7] Given differences in scoring systems across survey waves, PPVT scores were standardized within each round (mean = 0, SD = 1) prior to pooling. The PPVT has demonstrated strong cross-cultural validity within the Young Lives study context (Cueto & Muñoz, 2009).

Stage 1: Cross-Sectional Associations (Pooled OLS)

We estimated the following pooled OLS model:

$$Aspiration_{ir} = \beta_1 HFIAS_{\text{ir}} + X_{\text{ir}}\beta + \alpha_{\text{r}} + \gamma_j + \varepsilon_{\text{ir}}$$

where X_{ir} is a vector of child and household controls including gender, stunting status, urban residence, household size, household wealth, economic capacity, and social network support. α_{r} denotes round fixed effects and γ_j captures region fixed effects. We present specifications with and without wealth controls, given the potential bad-controls concern discussed above. This specification captures the cross-sectional association between household food insecurity and educational aspirations but does not account for unobserved, time-invariant characteristics that may be related to both food insecurity and aspirations.

Stage 2: Value-Added Model and Propensity Score Matching

To assess how the association between household food insecurity and educational aspirations changes when accounting for pre-existing differences, we estimate two benchmark specifications: a value-added model (VAM) and propensity score matching (PSM). These approaches address observed and pre-existing differences in different ways and therefore provide complementary benchmarks for the fixed-effects analysis.

2.1 Value-Added Model

The VAM accounts for pre-existing differences in educational aspirations by conditioning on the child's prior aspirations. The model is estimated using Round 5 observations only (age 15), conditioning on Round 4 aspirations (age 12), current household food insecurity, and the full set of household covariates. The specification is:

$$Aspiration_{i5} = \beta_0 + \beta_1 HFIAS_{i5} + \beta_2 Aspiration_{i4} + + X_{\text{ir}}\beta + \gamma_j + \varepsilon_i$$

where $Aspiration_{i4}$ is the lagged outcome and γ_j captures region fixed effects.

Unlike the within-child fixed-effects model, the VAM does not account for all time-invariant characteristics directly. Rather, it uses prior aspirations as a proxy for pre-existing differences that may be related to both household food insecurity and educational aspirations. The VAM is therefore interpreted as an associational specification rather than as a causal model. The analysis is restricted to children with valid Round 5 observations and non-missing Round 4 aspirations (N = 1,771 before covariate-induced exclusions).

2.2 Propensity Score Matching

PSM provides a second benchmark by comparing children with similar observed characteristics but different levels of household food insecurity. Because PSM relies on observed characteristics, it does not address differences arising from unobserved characteristics. The analysis uses 1:1 nearest-neighbor matching without replacement, with a caliper width of 0.2 standard deviations of the propensity score (Austin, 2011). The propensity score model conditions on gender, urban residence, stunting status, wealth index, household size, survey round, and region.

Matching quality is assessed using absolute standardized mean differences (|SMD|), with |SMD| < 0.1 indicating adequate balance. The analysis then estimates the association between household food insecurity and educational aspirations in the matched sample.

Stage 3: Primary Specification Using Within-Child Fixed Effects

The within-child fixed-effects model is the core specification of the paper, exploiting variation within the same child across rounds to account for time-invariant differences between children and their households. The model is specified as:

$$Aspiration_{ir} = \beta_1 HFIAS_{ir} + X_{ir}\beta + \alpha_i + \delta_r + \varepsilon_{ir}$$

where α_i represents child-specific fixed effects that absorb stable characteristics of the child and their household, including gender, urban residence, early stunting, parental education, and other time-invariant factors, while δ_r represents round fixed effects that account for common changes between ages 12 and 15.

The coefficient β_1 is identified from within-child changes in household food insecurity between ages 12 and 15, after accounting for common changes across rounds. Because time-invariant characteristics are absorbed by α_i, gender and urban residence are not included as explicit regressors in the fixed-effects specification. The vector X_{ir} therefore comprises time-varying covariates, including household size, wealth index, current stunting status, household economic capacity, and social network support.

If the scarcity hypothesis discussed in Section 1 is consistent with the observed data, children whose household food insecurity increases between ages 12 and 15 would be expected to exhibit lower aspirations relative to their earlier levels, while children whose household food insecurity decreases would be expected to exhibit higher aspirations. In contrast, if the structural selection hypothesis is consistent with the observed data, within-child changes in

household food insecurity would be expected to show little or no association with within-child changes in aspirations, because differences in aspirations may be largely related to stable structural characteristics.

We note that because the Young Lives Ethiopia study enrolled exactly one child per household, child fixed effects and household fixed effects are mathematically equivalent in this dataset. Each child fixed effect therefore absorbs all time-invariant household-level characteristics. Consequently, a separate household fixed-effects specification would not provide additional identifying variation.

Mediation Analysis

The study also examines whether cognitive ability may help explain the association between household food insecurity and educational aspirations, following the mediation framework described by Imai et al. (2010). The analysis estimates the association between household food insecurity and cognitive ability and the association between cognitive ability and educational aspirations conditional on household food insecurity. The indirect association is assessed using nonparametric bootstrapping with 1,000 simulations and percentile confidence intervals.

Two regression models are estimated: (1) a mediator model predicting standardized PPVT scores from HFIAS scores, and (2) an outcome model predicting educational aspirations from both PPVT scores and HFIAS scores. The baseline specification controls for gender, urban residence, stunting status, wealth index, survey round, and region fixed effects. A robustness specification additionally incorporates parental education. Parental education is incorporated using the missing-indicator method, which includes a binary indicator for missing values and assigns zero to the corresponding missing covariate value, allowing all observations to be retained while explicitly accounting for missingness.

To examine whether the associations in the proposed mediation sequence are primarily related to stable between-child or household differences or to within-child changes over time, we additionally estimate the mediator and outcome models using child fixed effects. These specifications provide a complementary assessment of whether the observed associations are evident after accounting for time-invariant child and household characteristics. In all regression specifications, standard errors are clustered at the primary sampling unit level (G

= 25) to account for spatial correlation and within-community dependence. The 25 clusters comprise the 20 original sentinel sites plus five region-based clusters created for the approximately 10.6% of observations that lacked a sentinel-site assignment, ensuring that all observations contribute to variance estimation.

All analyses were conducted using R software version 4.6.0 (R Core Team, 2026), using the $fixest$ package for fixed effects and clustered standard errors, $MatchIt$ for PSM, plm for panel data models, $mediate$ for ACME, and $MASS$ for ordered logit estimation. Unless otherwise noted, all tables report coefficients with standard errors clustered at the sentinel site level in parentheses, and statistical significance is denoted as: *† p<0.10, * p<0.05, ** p<0.01, *** p<0.001.*

3. Results

3.1 Descriptive Statistics

Descriptive statistics for the analytic sample (defined in Section 2) are presented in Table 1.

Table 1. Sample Characteristics and Regional Distribution by Survey Round

Variables	Rounds				
	Round 1	**Round 2**	**Round 3**	**Round 4**	**Round 5**
Year	2002	2006	2009	2013	2016
Age	1	5	8	12	15
N	1,999	1,912	1,885	1,869	1,803
Female	47.5%	47.2%	47.3%	47.2%	47.0%
Urban	35.0%	39.8%	39.7%	39.8%†	39.4%†
HFIAS (n)	--	--	1,870	1,863	1,799
Aspirations (n)	--	--	--	1,859	1,795
Regions					
Tigray	20.0% (400)	20.1% (385)	20.3% (383)	20.4% (381)	20.6% (371)
Amhara	20.0% (400)	20.0% (383)	20.2% (381)	19.6% (366)	19.7% (355)
Oromia	20.0% (399)	20.1% (385)	20.3% (383)	21.0% (392)	20.8% (375)
SNNPR	25.0% (500)	25.1% (479)	25.0% (472)	24.8% (463)	25.0% (451)
Addis Ababa	15.0% (300)	14.6% (280)	14.1% (266)	14.3% (267)	13.9% (251)

Notes: *† Urban status for Rounds 4–5 was imputed from the most recent prior measurement (see Section 2 for details). Five children in Addis Ababa with no prior urban record were assigned urban = 1, consistent with the Young Lives sentinel site classification of all Addis Ababa sites as urban. Standardized HFIAS data were first collected in Round 3 (age 8). Only Rounds 4–5 HFIAS data are included in the analytic sample as educational aspirations were first measured in Round 4 (age 12). Regional counts in parentheses.*

Food insecurity shifted during the adolescent transition (Table 2). Mean HFIAS scores declined from 5.58 in Round 4 to 5.27 in Round 5. Using the official HFIAP classification, Moderate-to-Severe insecurity (Categories 3–4 combined) declined from 63.8% at age 12 to 57.7% at age 15, while Mild insecurity (Category 2) increased from 15.9% to 26.5%. The proportion of Food Secure households (Category 1) declined from 20.0% to 15.7%. Consequently, Any Insecurity (Categories 2–4) increased from 79.7% to 84.1%, despite the decline in Moderate-to-Severe insecurity.

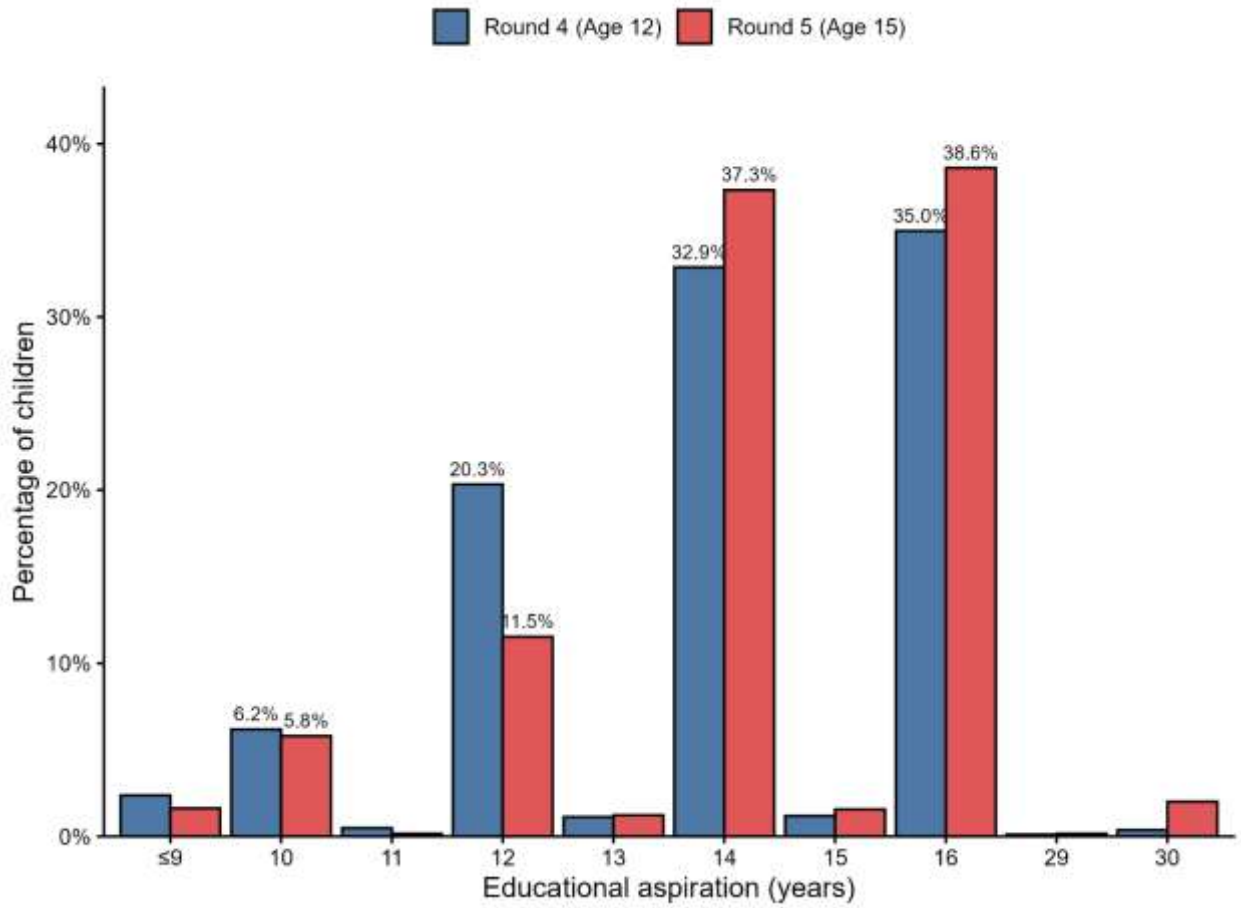
Table 2. Food Insecurity Summary by Survey Round (HFIAS Score and HFIAP Category)

Round	N	Median HFIAS	Mean HFIAS (SD)‡	Household Food Insecurity Access Prevalence (HFIAP)			
				Food Secure (%)	Mild (%)	Moderate to Severe (%)	Any Insecurity (%)
3‡	1,870	6	6.54 (4.41)	12.6	12.2	74.6	86.7
4	1,863	5	5.58 (4.38)	20.0	15.9	63.8	79.7
5	1,799	5	5.27 (4.07)	15.7	26.5	57.7	84.1
Change (R4→R5)	--	--	−0.31	−4.3	+10.6	−6.1	+4.4

***Notes**: ‡ Round 3 statistics are computed from the household-level HFIAS sample with valid scores (N = 1,870) and presented for contextual comparison only. Any Insecurity = HFIAP Categories 2–4 (i.e., not Food Secure). ⁑Standard deviations in parenthesis. Percentages in each row may not sum exactly to 100 due to rounding. N reflects households with non-missing HFIAS in the harmonized panel. See Methods for exclusion criteria applied to the child-level analytic sample.*

Educational aspirations exhibited a discrete distribution, with pronounced peaks at 10, 12, 14, and 16 years of schooling (Figure 2). At age 12 (Round 4), substantial concentrations were observed at 12 years (20.3%) and 16 years (35.0%). By age 15 (Round 5), the proportion aspiring to attain 12 years of education declined from 20.3% to 11.5%, while the proportions aspiring to 14 and 16 years increased. In particular, aspirations for 16 years of schooling increased from 35.0% to 38.6%. Mean educational aspirations increased from 13.9 years at age 12 to 14.5 years at age 15, while the median remained unchanged at 14 years (Table 3).

Figure 2. Distribution of educational aspirations by survey round



***Notes**: Bars represent the percentage of children reporting each desired level of completed schooling in Round 4 (age 12; N = 1,859) and Round 5 (age 15; N = 1,795). Percentages are calculated within each survey round.*

Table 3. Summary statistics of educational aspirations by survey round

Round	Age	N Aspiration	Mean years	SD	Median	N Enrollment	Enrollment Rate
Round 4	12	1,859	13.9	2.45	14	1869	94.6
Round 5	15	1,795	14.5	3.08	14	1769	92.7

Turning from the marginal distributions to individual-level dynamics, Table 4 reports within-child changes in food insecurity and educational aspirations between Rounds 4 and 5. Based on the continuous 0–27 HFIAS score, 45.9% of children experienced improved food security, 41.4% experienced worsening food security, and 12.8% remained stable, with a mean change of −0.35 points (SD = 4.81; N = 1,784). Among the 1,780 children with paired educational-aspiration observations, 1,732 had valid ordinal responses at both rounds. Of these, aspirations increased for 37.0%, decreased for 26.0%, and remained unchanged for 37.0%, with a mean change of +0.36 years (SD = 2.50). Overall, 87.3% of children experienced a change in HFIAS scores and 63.0% experienced a change in educational aspirations between the two rounds, indicating substantial within-child variation in both measures.

Table 4. Within-Child Dynamics (Round 4 to Round 5)

Direction of Change (Age 12 → 15)	Food Insecurity (HFIAS)	Educational Aspirations
Improved / Increased	45.9%	37.0%
Worsened / Decreased	41.4%	26.0%
Stable (No Change)	12.8%	37.0%
Mean Change	−0.35	+0.36 years
Standard Deviation of Change	4.81	2.50 years
N	1,784	1,732†

***Notes**: † Educational aspirations exclude children who reported a non-ordinal aspiration code at one or both rounds. Of the 1,780 children with paired aspiration observations, 48 (2.7%) were excluded because their reported aspiration corresponded to adult literacy, religious education, or another unspecified qualification at one or both rounds.*

Table 5 decomposes the variation in the principal time-varying variables into between-child and within-child components using the conventional child-means variance decomposition. Aspirations exhibit substantial within-child variation (Within SD = 1.75), while HFIAS also shows meaningful within-child variation (Within SD = 2.38). Although the child-means

decomposition attributes a larger share of total variance to between-child differences (61.2% for aspirations and 69.2% for HFIAS), these estimates can overstate between-child variance in two-wave panels because each child's mean is based on only two observations.

To address this limitation, Table 5 also reports intraclass correlation coefficients (ICCs) estimated using restricted maximum likelihood (REML). The REML estimates indicate that 19.3% of the variance in educational aspirations is attributable to stable between-child differences (ICC = 0.193), implying that 80.7% reflects within-child variation over time. Similarly, the REML ICC for HFIAS is 0.352, indicating that 64.8% of its total variance is attributable to within-child variation. These estimates indicate substantial within-child variation in both educational aspirations and household food insecurity, while illustrating the sensitivity of variance decomposition to the estimation method in short-panel data.

By contrast, the wealth index has a much higher REML ICC (0.761), indicating that most of its variation is between households (76.1%), with only 23.9% attributable to changes within households over time. Stunting status shows an intermediate level of stability (REML ICC = 0.487), with 48.7% of its variance attributable to between-child differences and 51.3% to within-child variation over time. Stunting is therefore less stable than wealth or PPVT, but more stable than aspirations or HFIAS. Overall, the variance decomposition indicates substantial within-child variation in educational aspirations and HFIAS, alongside greater stability in wealth, PPVT, and stunting status.

Table 5. Variance Decomposition of Key Variables

Variable	Overall Mean	Overall SD	Between SD	Within SD	Between variance (%)	Within variance (%)	REML ICC
Aspirations (Years)	14.2	2.79	2.18	1.75	61.2%	39.4%	0.193
HFIAS Score (0–27)	5.43	4.23	3.52	2.38	69.2%	31.6%	0.352
Wealth Index (0–1)	0.38	0.17	0.16	0.06	89.7%	11.7%	0.761
PPVT (standardized)	0.00	1.00	0.94	0.36	87.9%	12.6%	0.738
Stunted (0/1)	0.27	0.45	0.39	0.22	75.5%	25.0%	0.487

***Notes**: Between and within percentages are variance shares calculated from the child-means decomposition. ICC denotes the proportion of total variance attributable to between-child differences estimated using REML. The within-child share implied by the REML estimate equals 1−ICC.*

We also assessed whether unobserved individual effects are present in the panel data. Wooldridge's (2010) test rejected the null hypothesis of no unobserved effects ($z = 7.11$, $p < 0.001$), indicating substantial stable individual-level heterogeneity in the outcome process.

This finding provides a rationale for accounting for time-invariant differences between children and their households through a within-child fixed-effects specification.

3.2 Econometric Analysis

3.2.1 Cross-Sectional and Value-Added Associations

In the baseline pooled OLS specification (Table 6, Model 1), food insecurity was negatively associated with educational aspirations (β = −0.0593, SE = 0.0172, p = 0.002, N = 3,637). A one-unit increase in the HFIAS score was associated with 0.0593 fewer years of educational aspirations (p = 0.002), indicating a modest but statistically significant pooled association. Given the 0–27 range of the HFIAS, a 10-point difference in HFIAS corresponds to an estimated 0.59 year difference in educational aspirations.[8]

The association attenuated as household wealth was introduced. Adding household wealth (Model 2) reduced the coefficient's magnitude by 31.7%, to −0.0405 (SE = 0.0212, p = 0.068). Adding the full set of controls (Model 3) yielded a coefficient of −0.0414 (SE = 0.0222, p < 0.10), representing a 30.2% attenuation relative to Model 1. Both Models 2 and 3 therefore show weaker associations than the baseline specification and do not meet the conventional 5% significance threshold.

To situate the pooled estimates relative to a specification that conditions on prior educational aspirations, Table 6 (Model 4) presents estimates from the VAM. In this specification, the food insecurity coefficient remains statistically significant at the 5% level (β = −0.0544, SE = 0.0257, p = 0.045, N = 1,757). The lagged aspiration coefficient is positive and statistically significant (β = 0.145, SE = 0.0415, p = 0.002), indicating a positive association between prior and subsequent educational aspirations.

Table 6. Cross-Sectional and Value-Added Associations

Variable	Pooled OLS			(4) VAM
	(1) Baseline	(2) + Wealth	(3) Full Controls	
HFIAS Score	−0.0593**	−0.0405†	−0.0414†	−0.0544*
	(0.0172)	(0.0212)	(0.0222)	(0.0257)
Lagged Aspiration (age 12)	--	--	--	0.1450**
				(0.0415)
Female	0.0901 (0.1188)	0.1018	0.0990	0.1519
		(0.1140)	(0.1130)	(0.2013)

[8] For scale comparison, the one-unit coefficient (0.0593 years) represents approximately 0.42% of the overall mean educational aspiration of 14.2 years (*see Table 5*), calculated as (0.0593 / 14.2) × 100 = 0.42%.

Urban	1.403***	1.0724**	1.088**	1.091**
	(0.3195)	(0.3011)	(0.3033)	(0.3367)
Stunted	−0.3639***	−0.3161**	−0.3266**	−0.1302
	(0.0933)	(0.0924)	(0.0978)	(0.1672)
Household Size	0.0519*	0.0449†	0.0465†	0.0438
	(0.0223)	(0.0220)	(0.0235)	(0.0402)
Wealth Index (w_i)	--	1.8012*	1.833*	1.656†
		(0.6927)	(0.6591)	(0.9307)
Economic Capacity	--	--	0.0732	0.0701
			(0.0702)	(0.1035)
Received Support	--	--	−0.1663	−0.3294
			(0.1787)	(0.2346)
Social Network Support	--	--	0.0699	0.0668
			(0.0467)	(0.0526)
Observations	3,637	3,630	3,605	1,757
R^2	0.113	0.120	0.123	0.128
Within R^2	0.066	0.073	0.076	0.090

***Notes**: Time-invariant covariates (female, urban, stunted) are retained as explicit regressors in Models 1–4 because none of them include child fixed effects. Models 2–4 are subject to the bad controls concern discussed in Section 2 for wealth, economic capacity, received support, and social network support. Social network support and received support are only weakly correlated in both the pooled sample (r = 0.124, n = 3,641) and the Model 4 estimation sample (r = 0.092, n = 1,770), confirming they capture distinct constructs rather than duplicating one another.*

3.2.2 Propensity Score Matching Estimation

The PSM analysis in Table 7 shows that, among children with similar observed characteristics, those with high household food insecurity (HFIAS above the sample median of 5) had aspirations approximately 0.29 years lower than their matched counterparts (covariate-adjusted ATT = −0.287, SE = 0.099, p = 0.004; unadjusted ATT = −0.332, SE = 0.103, p = 0.001). This corresponds to approximately 2.0% of the mean aspiration level of 14.2 years.[9]

Table 7. Propensity Score Matching Estimates

Estimator	Coefficient (ATT)	S.E.	P-value	N
Unadjusted ATT	-0.332	0.103	0.001	2,728
Covariate-Adjusted ATT	-0.287	0.099	0.004	2,728

***Notes**: Treatment = high food insecurity (HFIAS above the sample median of 5). Of 1,671 treated observations, 1,364 were successfully matched, yielding 2,728 matched observations in total. All matching covariates met the prespecified balance criterion (|SMD| < 0.1; see Appendix Table A11 for full balance diagnostics). The covariate-adjusted ATT further adjusts for gender, urban residence, stunting status, wealth index, and household size in the outcome regression estimated on the matched sample.*

3.2.3 Fixed-effects Estimation

[9] The percentage is calculated as (0.287 / 14.2) × 100 = 2.02%, where 0.287 is the absolute value of the covariate-adjusted ATT from Table 7 and 14.2 is the overall mean educational aspiration reported in Table 5.

The child fixed-effects model examines whether within-child changes in household food insecurity between ages 12 and 15 are associated with changes in educational aspirations, while accounting for time-invariant observed and unobserved characteristics. The estimated association between HFIAS and educational aspirations was not statistically different from zero (β = −0.0245, SE = 0.0497, Table 8). Adding the full set of time-varying household controls left the estimate essentially unchanged (β = −0.0239, SE = 0.0485), indicating that the estimate was not materially affected by these controls (*Appendix Table A8*). The within R^2 indicates that the model explains only a small share of the variation in aspirations after accounting for child fixed effects.

As a robustness check, the fixed-effects model was re-estimated on the PSM-matched subsample (*Appendix Table A12*). The HFIAS coefficient remained statistically insignificant (β = −0.029, SE = 0.049, p = 0.558), indicating that the null FE association persists in the covariate-balanced sample.

Table 8. Within-Child Fixed-Effects Associations and First-Difference Verification

Variable	**Child FE (Bivariate)**	**Child FE (Full Controls)**	**First Difference (no baseline control)**
HFIAS Score	−0.0245 (0.0497)	−0.0239 (0.0485)	−0.0245 (0.0174)
Intercept	--	--	0.600*** (0.083)
Estimator	Within-child FE	Within-child FE	OLS on Δ scores
SE type	Clustered	Clustered	Conventional OLS
Child Fixed Effects	Yes	Yes	--
Round Fixed Effects	Yes	Yes	--
R^2 (Within)	0.00112	0.0106	0.00112
Observations	3,542	3,470	1,771

***Notes**: Standard errors in parentheses. The standard errors of FD and FE differ because the FE model clusters standard errors by sentinel site/region, while the FD model uses conventional standard errors.*

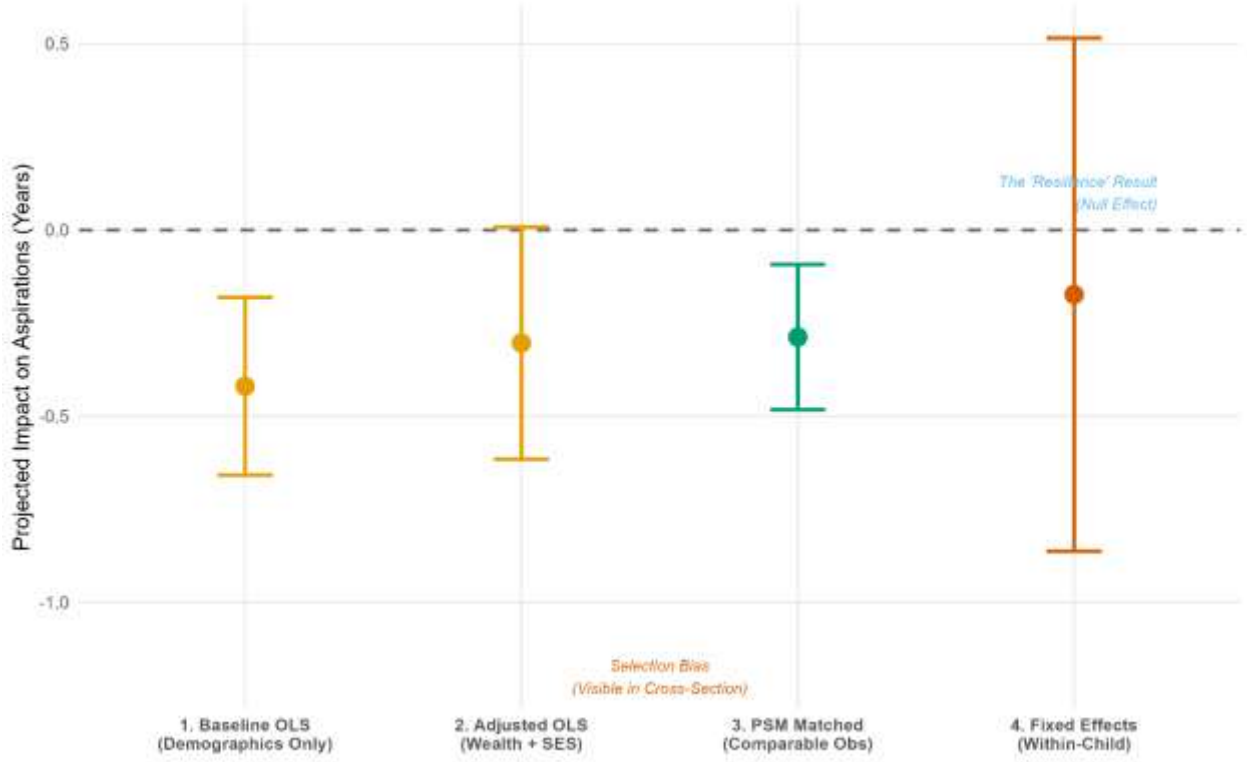
As a methodological verification, Table 8 also reports the first-difference (FD) specification, which produced an identical point estimate (β = −0.0245, SE = 0.0174), confirming the equivalence of the FE and FD point estimates in the two-period panel. In a two-period panel, FE and FD are algebraically equivalent and yield identical point estimates when estimated on the same observations with the same specification. Both specifications therefore yield the same substantive conclusion regarding the absence of a statistically detectable association. The standard errors differ because the two specifications use different variance estimators.

To assess whether measurement error in food insecurity could explain the null finding, we restricted the analysis to children with progressively larger within-child changes in HFIAS.

If attenuation from measurement error were driving the null result, the estimated coefficient would be expected to become increasingly negative as the minimum change threshold increased. However, as shown in *Appendix Table A13*, no such pattern emerges. The coefficient varies narrowly from −0.0249 to −0.0315 across thresholds of 1–4 HFIAS points and remains statistically insignificant across all thresholds (all $p > 0.5$). Thus, the estimated association does not strengthen among observations with larger within-child HFIAS changes, although this analysis cannot rule out measurement error as a source of attenuation.

Figure 3 shows progressive attenuation of the estimated association between household food insecurity and educational aspirations in specifications that account for within-child variation and time-invariant heterogeneity.

Figure 3. Comparison of Estimated Associations Between Food Insecurity and Educational Aspirations Across Models.



Note: *The PSM matched model reports the binary treatment contrast between high and low food insecurity, defined by the median HFIAS split. The baseline OLS, adjusted OLS, and fixed-effects estimates are rescaled from their per-unit HFIAS coefficients to the same high versus low contrast using the observed mean HFIAS difference between the two groups (approximately 7.07 points; Appendix Table A10). The original per-unit HFIAS coefficients are reported in Tables 6 and 8. The VAM is omitted because it is estimated on the restricted Round 5 sample (N = 1,757), whereas the other estimates use the pooled Round 4–5 sample, making it not directly comparable in this figure. The VAM estimate is reported in Table 6.*

3.2.4 Mediation Analysis

Having found no statistically detectable within-child association between changes in food insecurity and changes in educational aspirations, we next examine whether cognitive ability may help explain the observed food insecurity–aspiration association and whether this pattern is evident in cross-sectional and within-child specifications.

The cross-sectional mediation analysis revealed an association pattern consistent with cognitive ability being one potential component of the relationship between food insecurity and educational aspirations. A one-unit increase in the HFIAS score was associated with a reduction of 0.011 standard deviations in PPVT vocabulary scores ($\beta = -0.011$, SE = 0.003, $p < 0.01$), while higher PPVT scores were associated with higher educational aspirations ($\beta = 0.806$, SE = 0.062, $p < 0.001$). The estimated indirect association (ACME = −0.009, 95% CI [−0.015, −0.003], p = 0.002) is consistent with cognitive ability accounting for part of the cross-sectional association between food insecurity and aspirations. Adding parental education as a control attenuates the estimated indirect association modestly (ACME = −0.007, 95% CI [−0.013, −0.002], p = 0.008), but the association remains statistically detectable.

The key question is whether this pattern is primarily associated with within-child changes or with persistent differences between children and households. Variance decomposition (Table 5) shows that PPVT is largely stable over time, with a REML ICC of 0.738, indicating that approximately 74% of its variance is attributable to between-child differences, with the remainder reflecting within-child variation. The fixed-effects results provide a complementary assessment of whether the two components of the proposed pathway are also evident in within-child changes. Changes in food insecurity are not significantly associated with changes in PPVT scores (Path a: $\beta = -0.013$, SE = 0.008, p = 0.10, N = 3,114). In contrast, changes in PPVT scores remain positively associated with changes in aspirations after controlling for food insecurity (Path b: $\beta = 0.404$, SE = 0.133, $p < 0.01$, N = 3,100). The product of the two within-child coefficients is approximately −0.005 aspiration-years per unit increase in HFIAS.

Table 9. Cognitive Pathway in the Food Insecurity–Aspiration Association

Component	(1) Baseline OLS	(2) OLS with Parental Education	(3) Within-Child FE
Path a: Food Insecurity–Cognition	−0.011** (0.003)	−0.009** (0.003)	−0.013 (0.008)
Path b: Cognition–Aspirations	0.806*** (0.062)	0.806*** (0.062)	0.404** (0.133)
Indirect Association (ACME)	−0.009** (0.003)	−0.007** (0.003)	≈ −0.005 †
Direct Association (ADE)	−0.019 (0.012)	−0.019 (0.012)	--
Total Association	−0.027* (0.012)	−0.027* (0.012)	--
Observations	3,220	3,220	3,100–3,114

***Notes**: Standard errors in parentheses. Columns 1 and 2 use nonparametric bootstrap confidence intervals from 1,000 draws. Column 2 additionally controls for parental education years and a missing-data indicator. † The within-child indirect association is the product of the Path a and Path b fixed-effects coefficients and is reported for descriptive comparison only. No standard error or formal inference is available for this quantity. ACME (Average Causal Mediation Effect) and ADE (Average Direct Effect) are reported using the conventional mediation terminology. In this study, they are interpreted as statistical associations rather than causal effects.*

3.2.5 Heterogeneity Analysis

The overall finding of no statistically detectable within-child association raises the question of whether the association between food insecurity and educational aspirations differs across socioeconomic and contextual groups. We examine heterogeneity by socioeconomic status, gender, region, and early childhood nutritional status.

First, we examined gender heterogeneity and found no evidence of differential associations. In the primary fixed-effects specification, the coefficients for boys ($\beta = -0.024$) and girls ($\beta = -0.025$) were virtually identical, and a formal interaction test provided no evidence of a difference between the two coefficients (*Appendix Table A14*).

Stratification by wealth tercile revealed a cross-sectional gradient that was not supported by the within-child estimates (*Appendix Table A15*). The food insecurity–aspiration association was significant only among the poorest children in pooled OLS (T1: $\beta = -0.121$, SE = 0.035, $p = 0.002$), while the associations were not statistically different from zero in the middle (T2: $\beta = -0.012$, $p = 0.683$) and highest (T3: $\beta = -0.005$, $p = 0.842$) terciles. Within-child fixed-effects estimates were statistically indistinguishable from zero across all three strata (T1: $\beta = -0.062$, $p = 0.306$; T2: $\beta = -0.013$, $p = 0.855$; T3: $\beta = 0.010$, $p = 0.865$), and a formal joint test of the HFIAS-by-wealth-tercile interaction provided no evidence that the coefficients differed across terciles ($F(2,24) = 1.21$, $p = 0.316$). Thus, although the cross-sectional association was largest among the poorest children, the within-child estimates provided no evidence of differential associations across wealth groups.

Regional heterogeneity followed a similar pattern (*Appendix Table A16*). Among the five cross-sectional OLS models estimated separately by region, only Amhara showed a statistically significant association ($\beta = -0.133$, SE = 0.018, $p < 0.001$). Associations in SNNPR, Oromia, Tigray, and Addis Ababa were not statistically different from zero. Addis Ababa's near-zero association ($\beta = 0.001$, $p = 0.958$) should be interpreted cautiously given the distinct urban composition of this sentinel cluster. These regional estimates are based on pooled OLS and therefore carry the same limitations concerning unobserved differences between children and households as the aggregate cross-sectional estimates. Region-specific fixed-effects models were not estimable given the limited within-region sample sizes.

Finally, early childhood stunting at age 1 did not show evidence of modifying the association (*Appendix Table A17*). The interaction between HFIAS and stunting at age 1 was not statistically significant in the pooled OLS model ($\beta = -0.019$, SE = 0.021, $p = 0.356$). Concurrent stunting, measured contemporaneously in each round, is time-varying and is therefore included as a covariate in the full-controls fixed-effects specification (Table 8). Its within-child association with aspirations is small and statistically insignificant ($\beta = -0.14$, SE = 0.22), in contrast to the significant negative cross-sectional association reported in Table 6 ($\beta \approx -0.33$, $p < 0.01$). By contrast, early stunting measured at age 1 is a time-invariant characteristic and is absorbed by the child fixed effect, so its main effect cannot be separately estimated in this specification. Its potential moderating role is instead examined through the cross-sectional interaction reported in Appendix Table A17.

3.2.6 Robustness Checks

The core findings from both the cross-sectional and fixed-effects analyses are robust to a series of alternative specifications (see Appendix Tables A7 – A17). The OLS specifications generally retain a negative association between food insecurity and educational aspirations, although statistical significance varies across specifications, while the estimated within-child association remains statistically indistinguishable from zero throughout. Because multiple-testing adjustment would not change the substantive interpretation of these results, p-values are reported without adjustment throughout.

4. Discussion

Using the Younger Cohort of the Young Lives Ethiopia study (Rounds 1–5, 2002–2016), this study examines the relationship between food insecurity and educational aspirations during adolescence. The analytical sample comprises 1,873 adolescents observed at ages 12 (Round 4) and 15 (Round 5), providing 3,644 child-round observations. The study addresses three questions: whether food insecurity is associated with educational aspirations during adolescence (RQ1), whether the observed association is consistent with persistent structural differences rather than short-term changes in food insecurity (RQ2), and whether cognitive ability is associated with the food insecurity–aspiration relationship through within-child processes or stable between-child and household differences (RQ3).

The relatively stable composition of the analytic sample and regional distribution across survey rounds (Table 1) provides context for interpreting the changes in food insecurity and educational aspirations observed during adolescence. The divergence between HFIAP categories and the continuous HFIAS score (Table 2) suggests that the categorical shifts may partly reflect movement of households across severity thresholds rather than large changes in the underlying level of food insecurity. These findings highlight the value of considering both categorical and continuous measures when assessing changes in food insecurity over time.

The concentration of educational aspirations at 10, 12, 14, and 16 years of schooling in Figure 2 is broadly consistent with key educational milestones in the Ethiopian education system. The decline in aspirations for 12 years alongside increases at 14 and 16 years suggests a movement away from aspirations centered on secondary completion toward higher levels of educational attainment. At the same time, the unchanged median indicates that the shift was modest, with the central tendency of aspirations remaining at 14 years (Table 3). The individual-level dynamics further show that changes in food insecurity and educational aspirations were common between ages 12 and 15 (Table 4), with substantial proportions of children experiencing changes in both measures.

Consistent with these patterns, the variance decomposition indicates meaningful within-child variation in HFIAS and educational aspirations, although the extent of within-child variation differs across the study variables (Table 5). In contrast, wealth, PPVT, and stunting exhibit greater stability over time, with a larger share of their variation attributable to persistent

between-child or between-household differences. These patterns indicate that the longitudinal data contain substantial within-child variation in the key variables, while also revealing important differences in temporal stability across characteristics. The Wooldridge (2010) test for unobserved individual effects indicated substantial stable individual-level heterogeneity in the outcome process, providing a rationale for accounting for time-invariant differences between children and their households through a within-child fixed-effects specification.

The econometric findings show that food insecurity is negatively associated with educational aspirations in cross-sectional analyses and the VAM, but that this association is not statistically different from zero when the analysis uses within-child variation in the child fixed-effects models (Tables 6 and 8). The same pattern is evident across socioeconomic strata and in the cognitive mediation analysis (Table 9 and *Appendix Table A15*). Although VAM and fixed-effects models rely on different sources of variation and identifying assumptions, the attenuation of the HFIAS coefficient from the VAM to the child fixed-effects specification is consistent with the possibility that conditioning on prior aspirations does not account for the same time-invariant differences between children and households as the fixed-effects approach. Similarly, the smaller fixed-effects estimate in the PSM-matched subsample (*Appendix Table A12*) relative to the covariate-adjusted PSM estimate (Table 7) is consistent with the possibility that part of the cross-sectional association reflects time-invariant differences between children and households. However, differences in coefficient magnitude across *OLS, VAM, covariate-adjusted PSM, PSM-matched FE, and the primary FE specification* should not be interpreted as definitive evidence that unobserved characteristics account for the attenuation, because these estimators rely on different sources of variation and identifying assumptions.

These findings qualify the applicability of the scarcity mindset framework proposed by Mullainathan and Shafir (2013), introduced in Section 1, suggesting that scarcity-related cognitive mechanisms may not operate uniformly across developmental stages and outcomes. While the bandwidth-reduction mechanism may operate through chronic early-life nutritional deficits (Prado & Dewey, 2014), the present findings suggest that such mechanisms do not translate into short-term changes in educational aspirations during

adolescence. At this stage, aspirations appear to be more strongly embedded within persistent social and familial contexts than shaped by temporary fluctuations in material hardship.

By contrast, these findings are consistent with the *structural selection hypothesis*. Rather than establishing an effect of food insecurity, the observed cross-sectional association may partly reflect stable characteristics that influence both household food security and children's educational aspirations, producing an association that does not persist in the within-child specification. These characteristics may include child-specific factors such as cognitive ability and temperament; household-level factors such as parental preferences, valuation of education, and human capital; and broader community and institutional influences. The latter may include the cultural shift toward valuing education documented by Woldehanna and Araya (2016), and the expansion of schooling, which may have contributed to the normalization of university aspirations across socioeconomic groups.

Existing research using Young Lives India data has documented negative associations between food insecurity and educational outcomes, including years of schooling, vocabulary, and mathematics scores (Argaw et al., 2023; Aurino et al., 2019). Argaw et al. (2023) operationalize food insecurity as a categorical severity variable and a persistence ratio across three survey rounds, while Aurino et al. (2019) classify households into trajectory groups based on the timing and duration of exposure. Relative to this literature, the present study addresses three gaps. First, both prior studies examine achieved educational outcomes rather than aspirations. Aspirations operate earlier in the sequence of educational decision-making and provide a relatively more direct test of the behavioral mechanism posited by scarcity theory, which concerns how material hardship shapes forward-looking cognition and choice rather than downstream achievement. Second, both studies rely on VAMs without testing whether the observed associations hold under child fixed effects, which remove time-invariant child-level and household-level heterogeneity that baseline-score adjustment alone cannot address. Third, both studies draw on Indian data, whereas the present study uses Ethiopian data. Differences in findings may therefore reflect contextual factors specific to each setting rather than a general feature of the food insecurity–education relationship.

These design differences bear directly on how the present findings should be interpreted alongside the prior India studies. Notably, our own VAM estimate was also negative and statistically significant (*see Table 6*), consistent with the associations reported by Argaw et

al. (2023) and Aurino et al. (2019). It was only when we moved to a child fixed-effects specification that this association attenuated to statistical non-significance. This pattern suggests that VAMs, including those used in the prior India studies, may not fully distinguish within-child changes in food insecurity from persistent, time-invariant differences between children and households. We do not interpret this as a contradiction of their findings, however, because the studies differ in outcome (achievement versus aspirations), in the operationalization of food insecurity (categorical severity and persistence, or trajectory groups, versus a continuous HFIAS occurrence score), and in country context (India versus Ethiopia). We use a continuous occurrence-based HFIAS score at ages 12 and 15 rather than a categorical or trajectory-based classification because occurrence-based items describing severe deprivation are considered more cross-culturally comparable than moderate or subjective items such as reduced dietary quality, which may be more susceptible to context-specific interpretation Deitchler et al. (2011), as cited in Headey (2026). Given these differences, the discrepancy between our null within-child result and the significant associations reported in the India studies is better understood as reflecting differences in research questions, measurement, and identification strategy than as evidence against either set of findings.

The findings are supported by qualitative evidence from the same Young Lives Ethiopia Study. Tafere (2014), drawing on Rounds 2–3 survey data and longitudinal qualitative fieldwork from the Young Lives Ethiopia Older Cohort study (children born in 1994/95), finds that children and parents maintain high educational aspirations across different levels of poverty, measured by household consumption relative to a caloric-based poverty line rather than by food insecurity specifically. Where aspirations do change, these changes are more closely associated with children's progression through school grades than with poverty status. This suggests that educational aspirations are shaped less by household economic circumstances than by structural factors, including perceived returns to education, community expectations, and the rapid expansion of schooling. However, because Tafere's poverty measure is a consumption-based classification rather than a food-security indicator, the study provides only indirect support for the claim that food insecurity does not substantially affect educational aspirations. Consistent with this, Bernard et al. (2014) find that an aspiration-raising intervention in rural Ethiopia had no detectable effect among those

with the lowest baseline aspirations, suggesting that aspirations may be particularly resistant to change among those facing greater structural disadvantage. Together, these findings suggest that structural barriers to educational opportunity may operate through channels other than short-term changes in household food insecurity.

A related pattern appears in Mistry et al. (2009), who found that among Chinese American adolescents, the association between parent-reported family economic stress and educational engagement and attitudes was not consistently mediated by adolescents' own reports of financial constraints, despite using an individual fixed-effects design to address selection bias. This is broadly consistent with the present finding that an observed cross-sectional association does not necessarily translate into a corresponding within-child relationship. However, the comparison is indirect because Mistry et al. (2009) examined Chinese American adolescents and outcomes related to educational engagement and attitudes rather than Ethiopian adolescents' desired years of schooling, and used different measures of economic stress and financial constraints. These differences in population, exposure, mediator, and outcome limit the direct comparability of the two studies.

The results of this study also do not contradict evidence that early childhood malnutrition impairs cognitive development (Victora et al., 2021). Rather, they indicate that during adolescence, educational aspirations show no statistically detectable within-child association with changes in food insecurity, even though other outcomes may be affected. This pattern is consistent with the concept of aspirational resilience, understood as the persistence of educational goals despite changing material circumstances. The question is therefore not whether poor Ethiopian adolescents possess the capacity to aspire, but whether structural barriers prevent the realization of those aspirations. Food insecurity may affect attendance, cognitive performance, or grade progression while leaving educational goals relatively intact, particularly in a context where schooling has become a deeply entrenched cultural expectation (Tafere, 2014; Woldehanna & Araya, 2016). These possibilities, however, are not directly tested in the present study.

Regarding regional heterogeneity, the Amhara region is the only context in which cross-sectional OLS models show a significant negative association between food insecurity and aspirations (*Appendix Table A16*). However, this finding is also based on cross-sectional OLS and is therefore subject to the same confounding concerns as the aggregate cross-sectional

results. Its theoretical interest lies in the possibility that Amhara's historical exposure to severe drought episodes (Dercon, 2004) may make food insecurity a more salient signal of structural fragility in that specific context. Distinguishing this explanation from regional differences in reporting norms or unmeasured confounders requires further qualitative research.

This finding also highlights the importance of distinguishing between aspirations and attainment (*Appendix Table A18*), which are related but distinct constructs. In this sample, the correlation between aspirations and enrollment is modest (r = 0.18). Non-enrolled children report mean aspirations of 12.4 years compared to 14.4 years among enrolled children, a meaningful gap that confirms aspirations persist even when enrollment is disrupted. As Favara (2017) demonstrates, high aspirations in Ethiopia frequently fail to translate into attainment.

The findings should, however, be interpreted in light of the following limitations and their implications for interpretation. First, the analysis is observational, and the fixed-effects approach eliminates confounding from time-invariant unobserved characteristics but does not rule out residual confounding from unobserved time-varying factors that might jointly affect food security and aspirations between ages 12 and 15. The estimated coefficients should therefore be interpreted as conditional associations rather than causal effects.

Second, the null within-child association between food insecurity and educational aspirations should not be interpreted as evidence that changes in food insecurity are irrelevant to educational outcomes more broadly. Educational aspirations capture only one dimension of adolescents' educational trajectories, and changes in food insecurity may affect outcomes not examined here, including school attendance, grade progression, cognitive development, or school dropout. This distinction is important because educational aspirations may remain relatively stable even when other aspects of educational engagement respond to changes in household food insecurity.

Third, the study is based on two observation periods (ages 12 and 15), which limits the ability to examine how the association evolves across adolescence. A longer panel with additional observations would allow a richer examination of aspiration dynamics and would also provide a stronger basis for assessing whether aspirations at different ages represent

comparable repeated measures of the same underlying outcome. It would further allow investigation of whether associations emerge or differ at particular stages of adolescence.

Fourth, the null finding may partly reflect measurement error in HFIAS rather than the absence of an underlying association. Experience-based food insecurity measures may be affected by recall, translation, and social desirability biases (Headey, 2026). Under classical measurement error, such noise can attenuate estimated associations toward zero. This concern is particularly relevant for the fixed-effects analysis, which identifies associations from within-child changes in HFIAS and therefore relies on variation that may be smaller and more susceptible to measurement noise than cross-sectional differences between households. The issue is especially pertinent because the HFIAS in the study uses a 12-month recall period rather than the standard 4-week reference period.

Nevertheless, two features of the findings suggest that measurement error alone is unlikely to fully account for the null finding. First, the estimated HFIAS coefficient has a 95% confidence interval of −0.122 to 0.073. Second, the within-child association between HFIAS and PPVT scores is also statistically indistinguishable from zero (Path a), despite PPVT being measured through direct cognitive assessment rather than household recall. These results do not rule out HFIAS measurement error as a contributing factor, but suggest that the null finding is not specific to the aspiration outcome.

Finally, the fixed-effects estimates are less precisely estimated than the pooled OLS and VAM estimates because they rely only on within-child changes over time. The absence of statistical significance should therefore not be interpreted as precise evidence of no association. Rather, the estimates indicate that the study does not detect a statistically distinguishable within-child association given the available variation and precision of the data.

Future research could address these limitations by using food-insecurity measures with shorter recall periods and stronger evidence of reliability and cross-cultural validity, providing a more robust assessment of within-child associations between food insecurity and educational aspirations. Longer panel datasets extending from late adolescence into early adulthood would also help determine whether the observed patterns persist as young people experience important educational and labour-market transitions and whether aspirations remain stable or change at later stages of the life course. Where feasible, experimental or quasi-experimental designs could provide stronger evidence on whether changes in food

insecurity affect educational aspirations. Finally, examining whether similar patterns are observed in other Sub-Saharan African countries, and more broadly in countries across the Global South experiencing rapid educational expansion, would help establish the extent to which the findings from Ethiopia generalize to other contexts.

5. Conclusions and Implications

In a sample of 1,873 Ethiopian adolescents followed from ages 12 to 15, the study finds that household food insecurity is negatively associated with children's educational aspirations in cross-sectional analyses. This association is somewhat weaker in the VAM and becomes statistically insignificant in the within-child fixed-effects specification. The attenuation of the association across specifications suggests that persistent differences between children and households account for an important part of the observed cross-sectional gradient.

These findings suggest that the *scarcity mindset* framework (Mullainathan & Shafir, 2013) has limited explanatory power for explaining within-child changes in adolescents' longer-term educational aspirations associated with changes in household transitory food insecurity. This does not, however, rule out the possibility that scarcity may shape attention and decision-making under conditions of immediate constraint. Instead, the findings support the *structural selection* interpretation developed in Section 1, whereby the observed aspiration gradient appears to reflect persistent individual, household, and community differences. Barriers to educational mobility among Ethiopian adolescents therefore appear to be more closely related to persistent structural conditions than to short-term changes in household food insecurity.

The findings have several implications for research, practice, and policy. First, nutrition interventions should prioritize early childhood, when cognitive development is most responsive to nutritional inputs and the foundations of aspiration formation are established, rather than relying on adolescent feeding programs to influence short-run goal setting. Importantly, this does not imply that adolescent nutrition interventions lack value for health or educational outcomes other than aspirations. Rather, our findings provide no evidence that their effects on educational aspirations should be expected to operate through short-term changes in food insecurity. Second, interventions seeking to improve educational aspirations should therefore not rely on reducing short-term household food insecurity alone. Instead,

policies addressing the broader structural conditions that shape both food security and educational opportunity may be more relevant to narrowing inequalities in educational aspirations and, ultimately, attainment.

Third, while safety nets such as Ethiopia's Productive Safety Net Programme (PSNP)[10] can mitigate extreme deprivation, strengthening household assets and economic capacity remains essential for reducing the structural constraints associated with lower educational aspiration. Fourth, policy should focus not only on sustaining educational aspirations but also on reducing the structural constraints that prevent aspirations from translating into educational attainment. This would help narrow the gap between aspiration and attainment documented by Favara (2017). Finally, these findings carry direct implications for Ethiopia's integrated human capital strategy, particularly the Seqota Declaration's commitment to ending child undernutrition by 2030 (FDRE Ministry of Health, 2015). They also speak to the broader Sustainable Development Goals (SDGs), particularly SDG 1 (No Poverty), SDG 2 (Zero Hunger), and SDG 4 (Quality Education) (United Nations, 2015).

[10] The Productive Safety Net Programme (PSNP) is the Government of Ethiopia's flagship social protection programme, implemented by the Ministry of Agriculture with support from the World Bank and other development partners. Launched in 2005, it provides predictable cash and/or food transfers to chronically food-insecure and extremely poor rural households. The programme combines labor-intensive public works, including soil and water conservation, rural roads, and community infrastructure, for able-bodied beneficiaries with unconditional direct support for labor-constrained households. It also provides livelihoods support and shock-responsive assistance during droughts (FDRE Ministry of Agriculture, 2020). Following the completion of PSNP5 (2020–2025), the programme entered its sixth phase (PSNP6, 2026–2031), which places greater emphasis on livelihoods, job creation, human capital development, adaptive social protection, and climate resilience, while continuing to reach approximately eight million rural Ethiopians.

Appendices

Appendix A: Variable Definitions and Harmonization

A.1 Food Insecurity Measurement, based on Coates et al. (2007)

- **Instrument:** Household Food Insecurity Access Scale (HFIAS).
- **Rounds included:** Rounds 3 (age 8, 2009), 4 (age 12, 2013), and 5 (age 15, 2016). HFIAS items were first collected in Round 3. The regression analyses are restricted to Rounds 4 and 5, as these are the only waves containing both food insecurity and educational aspiration data.
- **Recall period:** Twelve months, consistent with the annual Young Lives survey design, rather than the standard four-week recall period used in the original HFIAS instrument.
- **Instrument completeness:** The Young Lives Ethiopia survey collected both binary occurrence responses (0 = No, 1 = Yes) and frequency-of-occurrence follow-up responses (1 = rarely, 2 = sometimes, 3 = often) for all nine HFIAS items across Rounds 3–5. This permits the computation of the standard 0–27 HFIAS score and the four-category Household Food Insecurity Access Prevalence (HFIAP) classification for each survey round.
- **Skip routing:** The standard HFIAS skip-routing protocol was modified in the Young Lives Ethiopia survey. In Rounds 4 and 5, households responding "No" to all four basic occurrence items (Q1–Q4) were routed past Q5–Q7. Items Q8 and Q9 were administered only if Q7 = Yes in Rounds 4 and 5 but were asked unconditionally in Round 3. Frequency-of-occurrence items followed the same skip structure as their corresponding occurrence items: when an occurrence response was "No," the corresponding frequency value was coded as 0. No frequency values were imputed for cases in which the occurrence response was "Yes", but the frequency response was missing.
- **Occurrence items missing value handling:** Skip-routed missing occurrence values were imputed to zero, with the validity of this imputation confirmed separately by round. In Round 4, all 144 households missing Q5–Q7 were confirmed to have Q1–Q4=0. In Round 3, only 2 of 4 households missing Q5–Q7 had Q1–Q4=0. Imputation to zero was therefore applied only to the confirmed cases, not unconditionally. Across

all rounds, 100% of cases missing Q8–Q9 had Q7=No, supporting unconditional imputation for those items. In Round 5, two households had a non-skip-routed missing value on at least one of Q1–Q4, preventing verification of the skip condition required to recode Q5–Q7 to zero. These cases represent only 2 of 1,821 households and have a negligible impact on the descriptive statistics.

- **HFIAS score missing value handling:** A household was coded as having a missing HFIAS score if any occurrence item was missing or if any occurrence response of "Yes" lacked the corresponding frequency-of-occurrence response. Observations with missing HFIAS scores were excluded from the regression analyses through listwise deletion.
- **Items included (all nine standard HFIAS occurrence questions, paired with a frequency-of-occurrence follow-up):**

1. In the past 12 months, did you worry that your household would run out of food before you could buy more or get food another way?
2. In the past 12 months, were you or any household member not able to eat the kinds of foods you wanted because of a lack of resources?
3. In the past 12 months, did you or any household member have to eat a limited variety of foods due to a lack of resources?
4. In the past 12 months, did you or any household member have to eat food you did not want to eat because of a lack of resources?
5. In the past 12 months, did you or any household member have to eat less (smaller portions) than you felt you needed because there was not enough food?
6. In the past 12 months, did you or any household member have to reduce the number of meals eaten per day because there was not enough food?
7. In the past 12 months, was there ever no food to eat in your household because of lack of money to get food?
8. In the past 12 months, did you or any household member go to sleep at night hungry because there was not enough food?
9. In the past 12 months, did you or any household member go a whole day and night without eating anything because there was not enough food?

Each occurrence item above is followed by a frequency-of-occurrence item ("How often did this happen? 1=Rarely [once or twice], 2=Sometimes [3–10 times], 3=Often [more than 10 times]"), asked only when the occurrence response is "Yes".

- **Scoring**
- **HFIAS score (primary measure):** For each of the nine items, the frequency code is 0 if occurrence=No, and the recorded value (1/2/3) if occurrence=Yes. The nine frequency codes are summed to produce a continuous score from 0 to 27. Higher scores indicate greater food insecurity severity.
- **HFIAP classification:** Households are assigned to one of four ordered categories using the official sequential decision rule, applied to the occurrence and frequency item responses, with categories assigned in order from least to most severe so that the most severe matching condition determines the final classification:
  - ✓ **Food Secure:** No food insecurity condition reported, or only rare worry about food (Q1).
  - ✓ **Mildly Food Insecure:** Worry occurring sometimes or often, and/or inability to eat preferred foods at any frequency, and/or eating a limited variety or undesired foods rarely but no reduction in meal size or number, and none of the three most severe conditions.
  - ✓ **Moderately Food Insecure:** Eating a limited variety or undesired foods sometimes or often, and/or reducing meal size or number rarely or sometimes but none of the three most severe conditions.
  - ✓ **Severely Food Insecure:** Reducing meal size or number often, or experiencing any of the three most severe conditions (no food of any kind; sleeping hungry; going a whole day without eating) at any frequency, even once.
  - ✓ Zero households with a valid HFIAS score were left unclassified by this rule in any round.
- **Persistence ratio and mean HFIAS score**: For households observed in both Round 4 and Round 5, the persistence ratio is the proportion of these two rounds in which the household was classified as Moderately or Severely Food Insecure (HFIAP category 3 or 4). It ranges from 0 to 1. The mean HFIAS score across available rounds is reported as an alternative continuous chronic-exposure measure.

- **Domain structure**:
  - ✓ Anxiety and uncertainty: Q1
  - ✓ Insufficient quality: Q2, Q3, Q4
  - ✓ Quantity reduction (occurrence-based descriptive; frequency-dependent severity is captured by the HFIAP classification rather than by this indicator): Q5, Q6.

A.2 Educational Aspirations

- **Question:** Children were asked directly: *"How many years of schooling do you hope [CHILD] will complete?"* Responses were provided by the child rather than the caregiver.
- **Measurement:** Educational aspirations were measured in Rounds 4 (age 12) and 5 (age 15).
- **Cleaning:** Implausible responses exceeding 30 years of schooling were recoded as missing.
- **Structure:** Educational aspirations are a discrete outcome with pronounced concentration at 12, 14, and 16 years of schooling. These peaks correspond closely to socially recognized educational milestones in the Ethiopian education system, including completion of secondary education, post-secondary qualifications, and university degree attainment. The observed clustering reflects children's tendency to express aspirations around salient educational pathways rather than measurement error or arbitrary response rounding. Ordered logit robustness checks indicate that modeling aspirations as an ordinal outcome does not materially change the main linear estimates.

A.3 Anthropometric Measures

- **Stunting:** Height-for-Age Z-score (HAZ) < -2 SD based on WHO Child Growth Standards (WHO, 2006).
- **Early Stunting:** Stunted at Age 1 (Round 1).
- **Concurrent Stunting:** Stunted at time of survey (Rounds 4–5).

A.4 Cognitive Assessment

- **Instrument:** Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4) (Dunn & Dunn, 2007).
- **Adaptation:** Translated and adapted for Ethiopia. Validated for local cultural context.
- **Standardization:** Scores were standardized within each survey round (Mean=0, SD=1) to account for differences in scoring methods (raw scores in R3 vs. percentiles in R4–5).

A.5 Demographic Controls

- **Urban/Rural:** measured contemporaneously in Rounds 1–3. For Rounds 4–5, the data were imputed using the most recent available measurement.
- **Wealth Index:** Composite asset index calculated by Young Lives, standardized within round.
- **Region:** Administrative regions (Addis Ababa, Amhara, Oromia, SNNPR, Tigray).
- **Economic Capacity:** Household ability to raise emergency funds (1,000 Birr within one month).
- **Received Support:** Binary indicator of whether the household received financial support from family/friends.

Appendix Tables

Attrition Diagnostics

Of the 1,999 children enrolled at baseline, 211 (10.6%) did not complete all five rounds and were classified as attriters. Attriters and completers differed significantly on three baseline characteristics. Attriters were more likely to live in urban areas (46.4% vs. 33.7%; $\chi^2 = 12.98$, $p < 0.001$) and came from smaller households (mean 5.36 vs. 5.76; $t = -2.49$, $p = 0.013$). Region composition also differed significantly ($\chi^2 = 16.77$, df = 4, $p = 0.002$), with Addis Ababa overrepresented among attriters (23.7% vs. 14.0% of completers). Gender did not differ significantly between groups (52.1% female among attriters vs. 47.0% among completers; $\chi^2 = 1.81$, $p = 0.179$).

Attriters differed from completers at baseline in region, household size, and urban residence (Tables A1–A2), motivating the use of inverse probability weights derived from a logistic model of dropout (*Appendix Table A3*). Applying these weights changes the cross-sectional coefficient by less than 1% (−0.0414 unweighted versus −0.0411 weighted; *Appendix Table*

A4) and leaves the fixed-effects estimate substantively unchanged (−0.0255, p = 0.614, weighted versus −0.0245, p = 0.627, unweighted; *Appendix Table A5*). Together, Tables A4 and A5 indicate that differential attrition does not materially alter either the cross-sectional or the within-child estimates reported in the main text.

Table A1. Attrition Bias Analysis (Completers vs. Non-Completers)

Variable	Test Statistic	P-value	Conclusion
Region	$\chi^2 = 16.77$, df = 4	0.002	Significant imbalance
Household size	t = −2.49	0.013	Significant imbalance
Urban residence	$\chi^2 = 12.98$, df = 1	< 0.001	Significant imbalance
Gender (female)	$\chi^2 = 1.81$, df = 1	0.179	Balanced

Notes: *Statistical significance set at p < 0.05. Significant predictors (region, household size, and urban residence) were subsequently used to construct inverse probability weights for the robustness checks reported in Tables A3–A4.*

Table A2. Baseline Characteristics by Attrition Status

Characteristic	Retained (n=1,788)	Attritors (n=211)	Difference	p-value
Region				0.002
- Addis Ababa	14.0%	23.7%	+9.7 pp	
- Tigray	20.7%	14.2%	−6.5 pp	
- Amhara	19.9%	20.9%	+1.0 pp	
- Oromia	20.2%	17.5%	−2.7 pp	
- SNNPR	25.2%	23.7%	−1.5 pp	
Household Size (Mean)	5.76	5.36	−0.40	0.013
Urban Residence	33.7%	46.4%	+12.7 pp	< 0.001
Female	47.0%	52.1%	+5.1 pp	0.179

Notes: *pp = percentage points, calculated as Attritor minus Retained. Completer definition requires presence in all five rounds (R1–R5). p-values from chi-square tests (categorical) and Welch t-tests (continuous). Attriters were significantly more likely to live in urban areas and to come from smaller households, and Addis Ababa was overrepresented among attriters. Gender did not differ significantly. Inverse probability weights were constructed from the baseline predictors in Appendix Table A3 to assess the sensitivity of the main estimates to differential attrition in Tables A4–A5.*

Table A3. Predictors of Study Dropout (Logistic Regression for IPW)

Predictor	Coefficient (SE)	Odds Ratio	p-value
Intercept	1.561 (0.317)	4.76	< 0.001
Region (ref = Addis Ababa)			
- Amhara	0.259 (0.257)	1.30	0.315
- Oromia	0.434 (0.265)	1.54	0.101
- SNNPR	0.278 (0.257)	1.32	0.280
- Tigray	0.631 (0.276)	1.88	0.022
Household Size	0.092 (0.036)	1.10	0.011
Female	−0.209 (0.147)	0.81	0.155
Urban	−0.371 (0.184)	0.69	0.044

Notes: *AIC = 1,335.5. Converged in 5 iterations. Tigray, household size, and urban residence were statistically significant predictors of dropout at the 5% level. Stabilized inverse probability weights were derived from this model. Mean = 1.000, SD = 0.043, trimmed range [0.939, 1.149], effective sample size = 1,785 of 1,788 completers (99.8%).*

The weighting scheme summarized in *Appendix Table A3* was then applied to the two central specifications reported in the main text (the pooled OLS full-controls model (Table 6, Model 3) and the bivariate child fixed-effects model (Table 8)) to assess sensitivity to differential attrition.

Table A4. IPW Cross-Sectional Attrition Correction

Model	Coefficient	SE	p-value	N
Unweighted OLS (full controls)	−0.0414	0.0222	0.075	3,605
IPW-Weighted OLS (full controls)	−0.0411	0.0221	0.0755	3,605

Notes: *IPW weighting changes the food insecurity coefficient by approximately 0.7%, indicating that the cross-sectional estimate is not sensitive to the attrition adjustment. The weighted model specification is identical to Table 6, Model 3 (full controls, including social network support), using trimmed observation-level IPW weights.*

Table A5. IPW Fixed Effects (Within-Child Attrition Correction)

Model	Coefficient	SE	p-value	N
FE Unweighted (bivariate)	−0.0245	0.0497	0.627	3,542
FE IPW-Weighted (bivariate)	−0.0255	0.0500	0.614	3,542

Notes: *Both specifications yield estimates that are not statistically different from zero ($p > 0.6$). The weighted and unweighted estimates are nearly identical, indicating that the within-child result is not sensitive to the attrition adjustment. The bivariate specification is used here (rather than the full-controls FE model) because it is the one directly comparable across both the weighted and unweighted estimation samples without additional covariate-driven missingness.*

Table A6. HFIAS Persistence Specification

Specification	Coefficient	SE	p-value	N	Description
Mean HFIAS	−0.0559	0.0283	0.060	3,547	Mean score across available rounds
Persistence Ratio	−0.338	0.319	0.299	3,547	Proportion of rounds with moderate/severe insecurity

Notes: *Pooled OLS with round and region fixed effects (full controls), N = 3,547 (children with valid HFIAS in at least two analytic rounds). Neither chronic-exposure measure reaches conventional statistical significance, although the mean HFIAS specification is close to the 0.10 threshold ($p = 0.060$). Because both measures are constructed by aggregating HFIAS across rounds, they do not provide the within-child variation required for the primary fixed-effects specification. Table 8 therefore remains the primary test of the within-child hypothesis.*

Robustness to Measurement, Functional Form, and Covariate Specification

The official HFIAP classification confirms the same distributional shift in food-insecurity severity described in Table 2. Chronic exposure, measured by mean HFIAS across rounds, shows a somewhat larger negative cross-sectional association (*Appendix Table A6*), although it does not reach conventional statistical significance. Because mean HFIAS is constructed by aggregating information across rounds, it does not provide the within-child variation required for the primary fixed-effects specification. Recoding HFIAS as a binary indicator

of any food insecurity or as a severe-food-insecurity indicator likewise produces statistically indistinguishable-from-zero fixed-effects estimates (*Appendix Table A7*). The HFIAS coefficient also remains similar as time-varying covariates are added sequentially, including social network support in the final specification (*Appendix Table A8*). An ordered logit specification yields a negative cross-sectional association in the same direction as the linear model (*Appendix Table A9*), but the coefficients and p-values from the two models are not directly comparable because they use different functional forms and scales. The substantive conclusion therefore rests on the within-child estimates rather than on equivalence between the cross-sectional functional forms.

Table A7. Robustness to Functional Form (Child Fixed Effects)

Specification	Coefficient	SE	p-value
Continuous HFIAS	−0.0245	0.0497	0.627
Binary (Any Insecurity > 0)	−0.2931	0.4529	0.524
Severe (HFIAS ≥ 6)	−0.1923	0.3370	0.574

Notes: *N = 3,542 throughout. The null result is robust across all three operationalizations of food insecurity exposure.*

Table A8. Stepwise FE Covariate Build-Up

Step	Specification	N	HFIAS coef	SE	p-value	Within R^2
1	Bivariate (HFIAS only)	3,542	−0.0245	0.0497	0.627	0.00112
2	+ Household size	3,542	−0.0261	0.0488	0.598	0.00689
3	+ Wealth index	3,528	−0.0290	0.0500	0.567	0.00780
4	+ Stunting status	3,516	−0.0289	0.0504	0.571	0.00809
5	+ Economic capacity, received support, social network support	3,470	−0.0239	0.0485	0.627	0.01059

Notes: *Time-invariant characteristics (gender, urban residence, region) are absorbed by child fixed effects throughout. The HFIAS coefficient is not statistically different from zero and stable across all five steps (range: −0.0239 to −0.0290, all p > 0.55).*

Table A9. Functional Form Robustness (Linear vs. Ordered Logit)

Specification	HFIAS Coefficient	SE	p-value
Linear OLS (pooled, M2)	−0.041	0.021	0.068
Ordered Logit (pooled)	−0.046	0.010	< 0.001

Notes: *The ordered logit coefficient represents the change in the latent propensity to be in a higher aspiration category per unit increase in HFIAS, whereas the OLS coefficient represents the change in years of schooling aspired to. The coefficients are therefore not directly comparable in magnitude. Both specifications yield negative cross-sectional associations, although statistical significance differs across the two functional forms. The within-child fixed-effects specification remains the relevant robustness test for the study's central identification strategy and yields a statistically indistinguishable-from-zero association.*

The PSM treatment threshold and the mean HFIAS difference used to rescale estimates in Figure 3 are reported in Appendix Table A10. The PSM achieves adequate balance on all matching covariates (all |SMD| < 0.1; *Appendix Table A11*), and restricting the fixed-effects model to this matched subsample yields a similarly null estimate (*Appendix Table A12*). As a sensitivity analysis for measurement-error attenuation, the fixed-effects model was also restricted to children with progressively larger observed within-child HFIAS changes. The resulting coefficients remain within a narrow range (approximately −0.025 to −0.032) and remain not statistically different from zero across all thresholds (*Appendix Table A13*).

Table A10. HFIAS Median Split and Figure 3 Rescaling Factor

Sample	N	Median HFIAS (split point)	Mean HFIAS (Low group)	Mean HFIAS (High group)	Mean Difference (scale factor)
Regression sample	3,644	5	2.17	9.24	7.07

***Notes**: HFIAS groups are defined using the median split at 5 (Low ≤ 5; High > 5). In the regression sample (N=3,644), the corresponding mean HFIAS values are 2.17 and 9.24, giving a mean difference of 7.07. This difference is the scale factor used to rescale the estimates in Figure 3. The same split in the PSM complete-case sample (N=3,630; Table A11) gives nearly identical values (2.17 and 9.23; difference = 7.06).*

Table A11. PSM Covariate Balance Diagnostics

Variable	SMD Before	SMD After	Balanced?
Wealth Index	−0.654	−0.072	Yes
Urban	−0.277	−0.041	Yes
Stunted	0.120	0.011	Yes
Household Size	−0.069	−0.024	Yes
Female	0.106	0.004	Yes

***Notes:** All covariates achieved balance after caliper matching (|SMD| < 0.1 for all variables). 1:1 nearest-neighbor matching, caliper = 0.2 SD of propensity score. Treatment = HFIAS above sample median (5). From 1,671 treated observations, 1,364 were matched; 307 fell outside the caliper. This balanced sample is the basis for the PSM-matched FE robustness check in Appendix Table A12.*

Table A12. FE on PSM-Matched Sample

Sample	N observations	N children	Coefficient	SE	p-value
Full sample	3,542	1,771	−0.0245	0.0497	0.627
PSM-matched subsample	3,220	1,686	−0.029	0.049	0.558

***Notes:** The within-child null effect persists in the covariate-balanced matched sample (Appendix Table A11), indicating that the null within-child estimate is not sensitive to restricting the sample to observations with comparable observed covariates.*

Table A13. FE Sensitivity by Minimum Within-Child HFIAS Change

Min. |ΔHFIAS|	N children (pre-singleton)	N observations (model)	Coefficient	SE	p-value	95% CI
1 point	1,546	3,092	−0.0249	0.0501	0.624	[−0.123, 0.073]
2 points	1,262	2,524	−0.0255	0.0487	0.605	[−0.121, 0.070]
3 points	952	1,904	−0.0278	0.0494	0.579	[−0.124, 0.069]
4 points	734	1,468	−0.0315	0.0502	0.536	[−0.130, 0.067]

Notes: *Each row restricts to children whose HFIAS score changed by at least the indicated amount between rounds. "Pre-singleton" N reflects children meeting the ΔHFIAS threshold before automatic removal of singletons (children observed in only one round after subsetting), which the fixed-effects estimator drops as uninformative; "N observations (model)" is the resulting estimation sample. The coefficient remains within a narrow range (−0.0249 to −0.0315) and remains statistically indistinguishable from zero across all thresholds (all p > 0.5). The absence of systematic movement toward larger negative estimates provides no evidence that classical measurement-error attenuation is driving the null result.*

Table A14. Food Insecurity Associations by Gender

Estimator	Boys	Girls
Child FE	−0.0237 (0.0566), p=0.679	−0.0254 (0.0614), p=0.683
Pooled OLS	−0.0420 (0.0220), p=0.068	−0.0381 (0.0261), p=0.158

Notes: *N for FE models: boys = 1,880, girls = 1,662 (after singleton removal). N for OLS models: boys = 1,922, girls = 1,708. Interaction terms are estimated on the full (non-stratified) sample (Pooled OLS N=3,630; Child FE N=3,542). Formal interaction tests likewise provide no evidence that the food-insecurity association differs by gender in either the pooled OLS or within-child fixed-effects specification.*

Table A15. SES Heterogeneity: Food Insecurity and Aspirations by Wealth Tercile

	T1: Low Wealth	T2: Mid Wealth	T3: High Wealth	T1 FE	T2 FE	T3 FE
HFIAS Score	−0.1214**	−0.0119	−0.0050	−0.0622	−0.0134	0.0095
	(0.0347)	(0.0288)	(0.0248)	(0.0596)	(0.0725)	(0.0557)
Observations	1,203	1,210	1,195	1,176	1,190	1,170

Notes: *Wealth terciles defined from Round 4 (baseline) household wealth index. OLS models include female, urban, stunted, wealth index, and round/region fixed effects. FE models include child and round fixed effects. A formal joint test of the hfias_score × wealth-tercile interaction in the fixed-effects specification found no significant difference across terciles (F(2,24) = 1.21, p = 0.316). The tercile-specific point estimates should not be interpreted as evidence of statistically different associations across socioeconomic strata. The significant OLS association in the lowest tercile (T1) is not observed in the corresponding within-child fixed-effects estimate.*

Table A16. Regional Heterogeneity (Pooled OLS Models)

Region	N	Coefficient	SE	p-value	Significant
Amhara	715	−0.133	0.0179	0.0007	***
Tigray	746	−0.0668	0.0713	0.402	
SNNPR	911	−0.0506	0.0573	0.411	
Oromia	734	−0.0008	0.0167	0.964	
Addis Ababa	524	0.0009	0.0163	0.958	

Notes: *Each row is a separate regression within that region. Full controls (female, urban, wealth index) plus round fixed effects. Region-specific fixed effects models were not estimable with sufficient power given within-region sample sizes. The Amhara association is a cross-sectional pattern and is subject to the same unobserved-confounding concerns as the aggregate cross-sectional estimate (Table 6); it should therefore not be interpreted as evidence of a causal effect.*

Table A17. Early Stunting × Food Insecurity Interaction

Variable	Coefficient	SE	p-value
HFIAS	−0.0315	0.0249	0.219
Stunted Age 1	−0.1106	0.1901	0.566
HFIAS × Stunted Age 1	−0.0195	0.0207	0.356
Female	0.1245	0.1195	0.308
Urban	1.0609	0.2986	0.0016
Wealth Index	1.9161	0.6604	0.0078

***Notes**: N = 3,500. Pooled OLS with region and round fixed effects. The interaction between HFIAS and stunting at age 1 is not statistically significant (p = 0.356), providing no evidence that early stunting moderates the cross-sectional association between concurrent food insecurity and aspirations.*

The robustness analysis also examines whether the pattern extends to current school enrollment. School enrollment is high (94.6% at age 12, 92.7% at age 15) and shows no relationship to food insecurity in either pooled OLS (β = 0.001, p > 0.10) or fixed effects (β = 0.0001, p > 0.10; *Appendix Table A18*). Aspirations and enrollment are only modestly correlated (r = 0.18) and are clearly distinct constructs; non-enrolled children still report substantial aspirations (12.4 versus 14.4 years for enrolled children). Because enrollment and aspirations are distinct outcomes, this analysis provides a complementary test: current enrollment is not detectably associated with HFIAS in either the pooled OLS or fixed-effects specification.

Table A18. Food Insecurity and School Enrollment (Robustness Check)

	Pooled OLS	Child Fixed Effects
HFIAS Score	0.0007 (0.0014)	0.0000916 (0.0023)
Female	Controlled	Absorbed by FE
Urban	Controlled	Absorbed by FE
Stunted	Controlled	Absorbed by FE
Wealth Index	Controlled	Controlled
Fixed Effects	Round + Region	Child + Round
Observations	3,600	3,490
R^2	0.0375	0.5631
Within R^2	0.0195	0.0000021

***Notes:** Dependent variable is current enrollment (1 = enrolled, 0 = not). Neither specification shows a statistically detectable association between HFIAS and current enrollment. The result provides complementary evidence that the null within-child association is not specific to the measurement of educational aspirations, although enrollment captures only one dimension of educational engagement.*

CHAPTER 2: CLIMATE VULNERABILITY AND CHILD STUNTING IN ETHIOPIA: A SPATIO-TEMPORAL MULTILEVEL ANALYSIS USING DHS DATA, 2011–2019

Abstract

In Ethiopia, dependence on rainfed agriculture makes food security and child nutrition vulnerable to climate variability. This vulnerability is particularly concerning given the persistently high burden of child stunting in the country, which declined only modestly from 44% in 2011 to 38% in 2016 and 37% in 2019. This study examined the association between household-level climate vulnerability, measured by the Climate Vulnerability Index (CVI), and childhood stunting in Ethiopia. We pooled data on 23,542 children from the 2011, 2016, and 2019 Ethiopia Demographic and Health Surveys (EDHS) and linked anthropometric and household data to satellite-derived climate indicators. We used multilevel logistic regression models incorporating random effects for administrative region and soil class and adjusted for child age, sex, breastfeeding status, household size, and survey year. A one-standard-deviation increase in household CVI was associated with 27% higher odds of stunting (aOR = 1.27; 95% CI, 1.20–1.34; $p < 0.001$), with the association modestly strengthening over time. Male sex and increasing child age were positively associated with stunting. Across livelihood systems, the association between CVI and stunting was positive but varied in magnitude. Decomposed analyses indicated that household sensitivity and adaptive capacity showed stronger associations with stunting than climate exposure alone, highlighting the greater relevance of the social and economic dimensions of vulnerability. Thus, addressing household sensitivity through improved water and sanitation and other social and demographic interventions, while strengthening adaptive capacity through maternal education, and household economic resources may support climate-resilient nutrition strategies tailored to Ethiopia's diverse livelihood systems.

Keywords: Climate vulnerability, child stunting, adaptive capacity, livelihood zones, Ethiopia, multilevel modeling, Demographic and Health Surveys

1. Introduction

The (economic) development trajectory of any country depends critically on the health and human capital of its children (Becker, 1964; Heckman, 2006). Yet this foundation is increasingly threatened by poor child nutritional outcomes, including stunting, particularly in low-income countries (World Bank, n.d.). The conditions underlying child nutritional outcomes are, in turn, being shaped by climate change.

Climate change poses a growing threat to global food security and nutrition, with significant implications for human health and development (IPCC, 2023). Low-income countries, particularly those in the Global South, are especially vulnerable to climate shocks due to their heavy reliance on rain-fed agriculture and limited adaptive capacity (Trisos et al., 2022). As a result, children under five in these settings face heightened risks of undernutrition (WHO, 2024).

Climate impacts can influence child nutrition through multiple pathways, including reduced agricultural productivity, increased food prices, compromised household food security, and heightened disease burden (Cooper et al., 2019; Kalhoff et al., 2023; Klapka et al., 2026; Lieber et al., 2022; Thiede & Strube, 2020). These pathways operate through three core dimensions of food and nutritional security (availability, access, and utilization), each plausibly disrupted by climatic variability (Thiede & Strube, 2020).[11] Importantly, Klapka et al. (2026) argue that climate change affects child nutrition primarily through indirect socioeconomic channels rather than through direct environmental exposure alone, highlighting the role of broader economic and social conditions in shaping nutritional outcomes.

Although children contribute minimally to the causes of climate change, they bear a disproportionate share of its health impacts (UNICEF, 2025). Projections estimate that by 2050, climate change will cause an additional 28 million children to be wasted and 40 million to be stunted globally (SADC & UNICEF, 2025), while nearly one billion children (approximately half of the world's total) are already exposed to high or extremely high water stress (Klapka et al., 2026). Evidence links extreme heat exposure to adverse birth outcomes

[11] The World Food Summit Plan of Action defines food security as follows: “Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life” (FAO, 1996).

such as preterm birth and low birth weight, as well as to long-term cognitive deficits associated with early malnutrition (Sheffield & Landrigan, 2011). Despite these risks, comprehensive empirical assessments of the effects of climate change on child nutrition outcomes remain scarce, particularly in sub-Saharan Africa (SSA), where integrated frameworks for assessing vulnerability are also limited (Klapka et al., 2026).

Child stunting, defined as height-for-age z-score (HAZ) below -2 standard deviations from the World Health Organization (WHO) Child Growth Standards median (WHO, 2006), serves as an indicator of chronic undernutrition reflecting long-term exposure to inadequate nutrition and recurrent infections, often driven by poor health environments (Black et al., 2013). In the context of climate vulnerability, stunting can also capture the long-term nutritional consequences of climate-related food insecurity and disease, particularly in settings where these stressors contribute to chronic nutritional deprivation (Save the Childern, 2023).

An estimated 149.2 million children under five were stunted globally in 2020 (Development Initiatives, 2021), a figure that remained largely stagnant at 148.1 million in 2022 (FAO et al., 2024), before rising to 150.2 million by 2024 with only 28 percent of countries on track to meet the goal of halving stunting by 2030 (UNICEF et al., 2025). These figures underscore the difficulty of achieving the WHO Global Targets 2025, which call for a 40 percent reduction in the number of stunted children under five from a baseline of 171 million in 2010 (WHO, 2025a).

Stunting is more than a public health concern. It constitutes a macroeconomic shock in the making, reflecting a loss of human capital that occurs before an individual even enters the labor market (Hoddinott et al., 2008). If climate change increases the incidence of stunting today, its economic consequences may materialize decades later through lower productivity and earnings, a narrower tax base, and higher public expenditure on health and social protection. By measuring stunting, we capture an early-life impairment with potentially persistent consequences for cognitive capacity, educational attainment, labor productivity and ultimately national economic development. In this sense, climate-induced stunting can become a mechanism through which environmental shocks are transmitted across generations, reinforcing the persistence of poverty and underdevelopment (Cooper et al.,

2019; Schmidt, 2019). Ethiopia provides a particularly important context for studying these relationships for several reasons, as discussed below.

Ethiopia, the second most populous country in Africa, with a population of over 132 million as of 2024, continues to face persistent poverty and food insecurity, creating conditions that heighten the risk of childhood stunting. Moreover, the country's predominantly rain-fed agricultural system, coupled with its dependence on primary product exports, makes it structurally vulnerable to climate variability and extreme weather events, which increasingly threaten livelihoods, infrastructure, and public health (UNICEF Ethiopia, 2025).

At the same time, despite remarkable economic growth averaging 9.1% annually between 2000 and 2017 (World Bank, 2020), child stunting in Ethiopia has remained persistently high, declining only modestly from 44% in 2011 (CSA & ICF, 2012) to 38% in 2016 (CSA & ICF, 2017) and 37% in 2019 (EPHI & ICF, 2021), before increasing again to 40% in 2024-25 (ESS & ICF, 2025). This coexistence of economic growth alongside persistent poverty, climate vulnerability, and high stunting rates underscores a critical research gap. Understanding how climate vulnerability interacts with household resilience to shape child nutrition outcomes is, therefore, essential for designing integrated, resilience-focused interventions.

Several factors moderate the relationship between climate change and undernutrition. For example, breastfeeding offers nutritional protection during periods of food insecurity (Tello et al., 2022), while unsafe water and sanitation increase the risk of diarrheal disease, which can contribute to nutrient loss and, when prolonged, severe malabsorption (WHO, 2011). These vulnerabilities are particularly pronounced in SSA, where 69% of the population lacked access to safely managed drinking water, 76% lacked safely managed sanitation, and 77% lacked basic hygiene services in 2022 (WHO & UNICEF, 2023). Household-level factors, particularly maternal education and access to resources, can significantly moderate these effects by enhancing health literacy, caregiving practices, and adaptive decision-making (Randell et al., 2021). Among these factors, caregiver education has been the most consistently investigated socioeconomic moderator in the climate–nutrition literature, with higher maternal education associated with reduced vulnerability to climate-related nutritional harm across diverse settings and climate hazards (Klapka et al., 2026).

The Intergovernmental Panel on Climate Change (IPCC) climate vulnerability framework provides a valuable lens for understanding climate-related risks by conceptualizing vulnerability across three dimensions: exposure (the presence of people, livelihoods, or assets in areas affected by climate hazards), sensitivity (the degree to which those systems are susceptible to harm), and adaptive capacity (the ability to adjust, respond, or cope with climate impacts) (IPCC, 2014a). Applied to child nutrition, this suggests that stunting results not solely from environmental shocks but from their interaction with household fragility and caregiver adaptive capacity.

Despite its conceptual appeal, few studies have empirically operationalized all three vulnerability dimensions simultaneously, particularly in SSA where both climate vulnerability and child undernutrition are widespread (Gasparrini et al., 2024; Schmidt, 2019). For instance, Randell et al. (2021) analyzed four rounds of Ethiopian Demographic and Health Survey (EDHS) data (2000, 2005, 2011 and 2016) for children aged 0-59 months linked with gridded rainfall and temperature data and found that children exposed to lower rainfall and higher maximum temperatures during their first year of life experienced significantly higher rates of stunting. Research by Davenport et al. (2017) examined children aged 12-35 months across 13 SSA countries using scenario-based projections and found that warming (1-2°C) combined with decreased seasonal rainfall was associated with increased child stunting.

Previous research has examined associations between climate variables and child nutrition outcomes in different settings (Cooper et al., 2019; Grace et al., 2012; Lieber et al., 2022), but four methodological gaps limit current understanding. First, existing studies predominantly examine single climate variables, typically rainfall or temperature in isolation, rather than composite indices capturing the full vulnerability pathway from hazard exposure through household sensitivity to adaptive capacity. This fragmentation has been documented for over a decade: an early systematic review of climate-undernutrition studies found that most examined single weather variables in isolation and relied on heterogeneous secondary data sources (Phalkey et al., 2015), and a more recent scoping review confirms the diversity of variables and lack of standardized analytical strategies still fundamentally limits comparability (Klapka et al., 2026). This fragmented approach cannot capture how

vulnerability emerges from the interaction of climate hazards with household fragility, consistent with IPCC frameworks.

Second, calculating vulnerability indices at the regional or district level, as done in virtually all prior work, systematically underestimates true vulnerability due to Jensen's inequality (Jensen, 1906). Specifically, because vulnerability is a nonlinear function of adaptive capacity, evaluating vulnerability at the mean level of adaptive capacity does not yield the mean vulnerability across households. Regional or district aggregation therefore obscures household-level heterogeneity and can bias vulnerability estimates downward. Third, while several studies document cross-sectional climate-nutrition associations, none have tested whether the magnitude of the climate-stunting relationship is itself changing over time, which is a critical question given that national stunting rates may decline even as climate sensitivity intensifies.

Fourth, spatial heterogeneity in climate-nutrition pathways across Ethiopia's diverse agroecological zones, including the role of soil-mediated agricultural productivity in these pathways, remains unexamined. Existing evidence suggests that this heterogeneity is substantively meaningful. Rainfall and temperature predict variation in child stunting and underweight differently across Ethiopia's highland, midland, and lowland zones, making a single model applied uniformly across agroecological contexts unlikely to be valid (Hagos et al., 2014). Similarly, the assumption that rural populations bear systematically greater climate-nutrition risk than urban populations has been challenged by multi-country evidence. Rural and urban children show similar vulnerability to heat-driven wasting under above-average temperature conditions, while urban children benefit significantly more than rural children when temperatures fall below average (Thiede & Strube, 2020). Both findings underscore the need for stratified analysis across livelihood zones, agroecological contexts, and settlement types.

This study addresses these gaps by employing a multilevel modeling approach to investigate the association between a composite climate vulnerability index (CVI) and child stunting in Ethiopia, while accounting for spatial and temporal heterogeneity. Specifically, this study addresses two key interconnected research questions:

1. What is the association between household-level climate vulnerability (and its underlying exposure, sensitivity, and adaptive capacity components) and child

stunting in Ethiopia, and how have these relationships changed between 2011 and 2019?

2. How do these associations vary across Ethiopia's regions, livelihood zones, and between rural and urban settings?

By doing so, the study strengthens the evidence base for understanding how climate vulnerability, livelihood conditions, and household adaptive capacity interact to shape child nutrition outcomes across Ethiopia's diverse agroecological settings. More broadly, this study contributes to Ethiopia's Climate-Resilient Green Economy (CRGE) Strategy (FDRE, 2011b) and advances several Sustainable Development Goals (SDGs) (United Nations General Assembly, 2015).

2. Materials and Methods

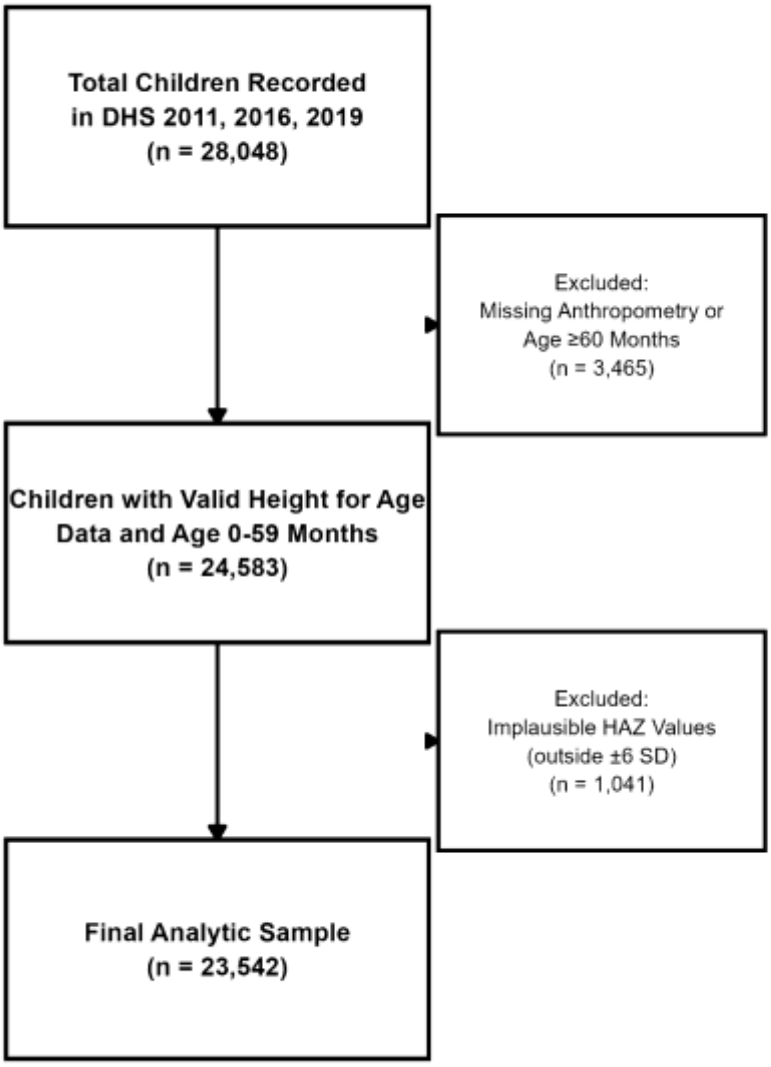
2.1 Data Sources

The study analyzed child stunting outcomes from three EDHS waves (2011, 2016, and 2019), spatially linked to CHIRPS rainfall estimates, ERA5-Land temperature reanalysis data, and cluster-level geocoordinates. Using the Children's Recode (KR) files[12], we identified 28,048 children born in the five years preceding each survey. After restricting the sample to children under 60 months of age with valid anthropometric measurements and applying WHO exclusion criteria for biologically implausible HAZ values (HAZ < −6 or > +6 SD) (WHO, 2006), 4,506 children (16.1%) were excluded, leaving a final analytical sample of 23,542 children (83.9% of the original sample) (see *Figure 4*).[13]

Figure 4. Sample Selection Flowchart

[12] Children's Recode" (KR file) is standard DHS Program terminology for the child-level dataset containing one record per eligible child under five years of age. The file includes child characteristics and selected variables pertaining to the child's mother and household (ICF, 2018).

[13] Spatial results are presented in tabular form (Tables 2 and A3) rather than as choropleth maps. Ethiopia's administrative boundaries have undergone substantial reorganization since the survey periods and the current study period, including the creation of several new regional states from the former Southern Nations, Nationalities and Peoples' Region. Moreover, some internal regional boundaries remain contested and have changed over time. Because a map would require selecting a particular boundary delineation that may not correspond to the administrative units used in the underlying survey sampling frames, we present regional comparisons numerically rather than cartographically. This approach avoids implying that the analysis endorses any particular current or historical administrative boundary delineation.



The EDHS provides nationally representative data through a two-stage stratified cluster sampling design, first selecting *primary sampling units (PSUs)* and then selecting *households*. The 2011 survey included 596 PSUs across 23 strata, the 2016 survey used 640 PSUs across 25 strata, and the 2019 Mini-DHS included 305 PSUs across 21 strata. *Appendix Table A19* summarizes survey characteristics across waves.

We linked the survey data with satellite-derived climate data to assess climate conditions during the periods preceding each survey. Climate data were obtained from two sources.

Precipitation data were obtained from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS v3.0), a 30+ year quasi-global rainfall dataset with a spatial resolution of 0.05º (approximately 5.5 km) (Funk et al., 2015). Temperature data were obtained from ERA5-Land monthly averaged data, a reanalysis dataset providing 2-m air temperature at a spatial resolution of 0.10º. Temperature data were extracted directly at each cluster's coordinates from the native ERA5-Land grid (Muñoz-Sabater et al., 2021).

These cluster-level precipitation and temperature anomalies were then averaged to the region-year level for use as the primary exposure measure, with the cluster-level version retained separately as a robustness check (Section 2.2). This region-level aggregation was prioritized for two reasons. First, DHS randomly displaces cluster coordinates to protect respondent confidentiality, by up to 2 km in urban areas and up to 5 km in rural areas, with a randomly selected 1% of rural clusters displaced by up to 10 km (Burgert et al., 2013). This displacement creates *positional uncertainty* in the spatial assignment of environmental exposures at the cluster level. Second, region-level exposure measures allowed us to retain the full analytical sample, as valid cluster GPS coordinates were unavailable for 650 children in the final analytical sample.

Additionally, soil data were obtained from the Harmonized World Soil Database (HWSD) v2.0 (FAO & IIASA, 2023), extracted at DHS cluster coordinates, and merged using the same procedure to account for agroecological heterogeneity. The HWSD mapping units characterize soil–landscape associations and provide information on soil properties relevant to local agroecological conditions. Soil class was retained at the cluster level because it was included as a control for agroecological heterogeneity rather than as a primary exposure.

2.2 Variables and Measurement

The primary outcome variable of the study was stunting, which reflects long-term nutritional deprivation (WHO, 2023). Following IPCC frameworks (IPCC, 2014a), we constructed a composite ***CVI*** as the primary treatment variable, incorporating three dimensions: exposure, sensitivity, and adaptive capacity, as detailed in *Appendix Table A20*.

The study primarily uses the multiplicative formulation of CVI, adapted from Mahapatra et al. (2021), who applied a structurally similar index to district-level agricultural vulnerability in India. The present study's methodological contribution in this regard lies in three

adaptations: (1) operationalizing all three vulnerability dimensions using DHS household-level microdata rather than district aggregates, thereby avoiding *Jensen's inequality bias*[14]; (2) constructing the exposure component from region-year climate anomalies matched to the DHS survey design, with cluster-specific satellite-derived anomalies incorporated as a robustness check; and (3) applying the index to child nutritional outcomes across Ethiopia's diverse agroecological zones.

Consistent with this framework, the CVI is calculated as:

$$CVI = \left(\frac{Climate_Exposure \times Sensitivity}{Adaptive_Capacity + 0.1}\right)$$

The constant (0.1) was selected to prevent division by zero in cases of extremely low adaptive capacity while maintaining meaningful scale differences.

Higher CVI values represent greater climate vulnerability. To reduce *aggregation bias*, we calculated the CVI at the household level before aggregating the resulting estimates within regions. This approach was important because adaptive capacity and sensitivity are correlated ($r = -0.54$), and adaptive capacity enters the index nonlinearly through the denominator. Calculating the CVI from regional mean values of the components rather than from household-level values resulted in an understatement of household-level vulnerability across regions, with a median of 51.6% and a range of 2.9% to 71.6% across the 33 region-survey year observations (*Appendix Table A21*). The understatement was smallest in Addis Ababa, ranging from 2.9% to 11.8% across survey rounds, and was largest in regions with greater within-region variation in sensitivity and adaptive capacity, consistent with the theoretical expectation that aggregation bias increases with household-level heterogeneity in the underlying components of the index. We therefore retained the household-level CVI estimates for subsequent regional aggregation.

The ***exposure dimension*** was calculated from precipitation and temperature anomalies computed at each DHS cluster over three-year periods preceding each survey, then averaged

[14] Jensen's inequality implies that applying a nonlinear function to an average does not generally equal the average of the function applied to individual observations (Jensen, 1906). Because adaptive capacity enters the CVI through the denominator, the index is a convex function of adaptive capacity, so region-level averages of the components can systematically understate household-level vulnerability relative to household-level construction. We therefore constructed the CVI at the household level and aggregated afterward. The exposure component, by contrast, is assigned at the region-year level and is therefore constant across households within a region-year, reflecting the spatial resolution of the climate data and the DHS cluster-displacement protocol rather than a limitation of the household-level measures.

to the region-year level for the primary specification's exposure term (2008–2010 for 2011; 2013–2015 for 2016; 2016–2018 for 2019). This three-year window was selected to capture cumulative nutritional impacts, as stunting reflects chronic undernutrition that develops over months to years (Black et al., 2013; Victora et al., 2021). The lag also ensures temporal precedence of climate exposure relative to outcome measurement. Although Ethiopia has distinct rainy seasons in some regions, we used annual rather than seasonal climate anomalies to maintain consistency across regions and to reflect the chronic, rather than short-term seasonal, processes underlying stunting.

Precipitation exposure was calculated as $\min(1, \frac{|P_i - \bar{P}|}{\bar{P}})$, where P_i represents precipitation in the period preceding survey wave i and $\bar{P}$ represents each cluster's mean precipitation over a fixed 16-year baseline period that is strictly prior to and immediately precedes that round's 3-year exposure period (1992–2007 for 2011; 1997–2012 for 2016; 2000–2015 for 2019). All three survey rounds had a full 16-year strictly prior baseline period.

We use deviations from long-term cluster-specific means (anomalies) rather than absolute precipitation levels. This approach captures departures from conditions that are typical for each cluster, thereby accounting for substantial spatial variation in baseline precipitation. It is also consistent with the idea that the effects of precipitation shocks on agricultural outcomes depend on deviations from locally expected conditions (Barrios et al., 2010). Anomalies also remove spatial confounding between regions, since absolute rainfall differs enormously between highland and lowland Ethiopia. Moreover, anomaly-based exposure metrics align with IPCC conceptualizations of climate hazard, which are defined relative to historical baselines (Ranasinghe et al., 2021). This anomaly-based approach nonetheless treats positive and negative deviations symmetrically, which may overstate vulnerability in hyper-arid regions such as Afar, where above-average rainfall relative to baseline can in some cases be beneficial rather than harmful. This limitation is revisited in the Discussion (Section 4).

Temperature exposure was calculated as $\min(1, \frac{|T_i - \bar{T}|}{2})$, where T_i represents the mean temperature in the period preceding survey wave i and $\bar{T}$ represents each cluster's mean temperature over the same fixed 16-year baseline period used for precipitation. This baseline period is strictly prior to and immediately precedes the corresponding 3-year exposure period.

We use 2°C as a scaling factor to put temperature anomalies on the same 0–1 scale as precipitation exposure. This choice is a methodological normalization rather than a claim that 2°C represents a biologically meaningful threshold.

The ***sensitivity dimension*** captures household characteristics that increase susceptibility to climate impacts. We constructed a composite score by averaging five component indicators (water risk, sanitation risk, birth interval, rural residence, and female headship), each scored on a common 0/0.5/1 scale. The *water risk* variable measured household water insecurity based on the WHO/UNICEF Joint Monitoring Programme (JMP) source-type classifications (ICF, 2018; WHO & UNICEF, 2021): 0 = improved sources (piped water, tube well/borehole, protected well, protected spring, rainwater, or delivered/packaged water); 0.5 = unimproved sources (unprotected well, unprotected spring, or other); and 1 = surface water (river, dam, lake, pond, stream, canal, or irrigation channel), representing the lowest rung of the JMP water ladder.

Sanitation risk was coded using an analogous JMP-based scale (ICF, 2018; WHO & UNICEF, 2021): 0 = improved facilities (flush/pour-flush to a sewer, septic tank, or pit latrine; ventilated improved pit latrine; pit latrine with slab; or composting toilet); 0.5 = unimproved facilities (flush to an unspecified location, pit latrine without a slab, bucket, hanging latrine, or other); and 1 = open defecation (no facility, bush, or field), representing the lowest rung of the JMP sanitation ladder. *Rural residence* was coded as a binary indicator (1 = rural, 0 = urban).

Biological vulnerability was measured using the preceding *birth interval*. Children born less than 24 months after a preceding sibling were classified as high risk (sensitivity score = 1.0), reflecting maternal nutritional depletion and reduced breastfeeding duration. Children born 24–35 months after a preceding sibling were assigned a moderate-risk score (0.5), whereas children born ≥36 months after a preceding sibling, as well as first-born children, were assigned no additional risk (score = 0.0). This classification aligns with WHO recommendations for optimal birth spacing (WHO, 2005) and with established associations between short birth intervals and child undernutrition (Conde-Agudelo et al., 2006; Rutstein, 2005).

Female headship was classified as an indicator of sensitivity rather than as a proxy for empowerment, while empowerment was conceptualized as an indicator of adaptive capacity.

In Ethiopia, female-headed households face distinct structural vulnerabilities, particularly among divorced and separated women, who experience higher levels of economic dependency and more limited access to productive assets than their male counterparts (Aguilar et al., 2015). This classification reflects the reality that female headship often results from marital dissolution or widowhood rather than autonomous choice, potentially reducing household access to adult labor and, consequently, resilience to climate shocks.

We assigned household wealth and maternal education to the ***adaptive capacity*** dimension. Although wealth may influence sensitivity, the IPCC (2014) framework conceptualizes it primarily as a system attribute that facilitates coping and recovery (Sharma & Ravindranath, 2019). Accordingly, to maintain theoretical consistency and avoid double-counting, wealth was treated exclusively as a determinant of adaptive capacity.

In addition to the CVI and its components, we included several covariates at the individual, household, and environmental levels to account for established predictors of child stunting. Child age in months and child sex (1 = Male, 0 = Female) were included as individual-level controls, given the established variation in stunting by age and sex. Stunting prevalence varies systematically with child age, as documented across 94 low- and middle-income countries (Karlsson et al., 2023), while boys have generally been found to experience higher stunting prevalence than girls in SSA (Wamani et al., 2007). We also incorporated breastfeeding status ($1 = yes, 0 = no$) as covariate, as it provides nutritional security during climate-induced food insecurity (Tello et al., 2022). Household size and soil class were also included to control for resource dilution and spatial agroecological variation, respectively.

2.3 Statistical Analysis

Multilevel modeling approaches are increasingly recognized as appropriate for analyzing climate-health relationships in nested data structures, where individuals are clustered within geographic units that share climate exposures (Leseba et al., 2025). Accordingly, the study employed multilevel logistic regression to account for the hierarchical structure of the DHS data, in which children are nested within clusters and clusters within regions. This approach accounts for the correlation among observations arising from their shared geographic contexts and allows the study to examine how factors at the individual and regional levels influence child stunting. For statistical modeling, the CVI was first square root transformed

to correct right skewness, while household size was log transformed. Both variables were then standardized using z score transformation.

Given the theoretical expectation of temporal changes in climate–vulnerability relationships outlined in Section 1, together with the superior model fit (i.e., lower AIC/BIC) observed in Section 3, we present the interaction model as our primary specification:

$$logit\left(P\left(stunting_{ijk}\right)\right) = \beta_0 + \beta_1 CVI_{std,\ ij} + \beta_2 Year_{c,i} + \beta_3\left(CVI_{std,ij} \times Year_{c,i}\right) + \beta_4 HouseholdSize_{std,ijk} + \beta_5 Breastfeeding_{ijk} + \beta_6 ChildAge_{ijk} + \beta_7 ChildSex_{ijk} + u_j + v_k$$

Where i indexes individual children, j indexes Ethiopia's eleven administrative regions, and k indexes soil class. CVI_{std} represents the standardized CVI, $HouseholdSize_{std}$ is the standardized household size, Breastfeeding is a binary indicator of current breastfeeding status, $ChildAge$ is age in months, and $ChildSex$ is a binary indicator coded as 1 for male and 0 for female. The term $Year_c$ denotes the survey year centered at 2015, which was used to aid model convergence and facilitate interpretation of the intercept. The terms u_j and v_k represent the random intercepts for region and soil class, respectively, accounting for spatial and agroecological heterogeneity. The soil class grouping comprised six HWSD texture classes and an explicit "missing" category for the 650 children in clusters without recorded GPS coordinates. This category was retained as a distinct level to preserve the sample size rather than exclude these observations from the model.

For robustness and diagnostic assessment, we conducted stratified analyses by livelihood zone (*Tables 13–14*) and tested alternative CVI formulations (additive and weighted) (*Appendix Tables A28–A29*). To control for multiple hypothesis testing across the four CVI coefficients from the livelihood zone stratified analysis, we applied the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) to maintain a false discovery rate of 0.05.

For stratified analyses examining whether the vulnerability–stunting relationship varied across broad production-system groupings, the study grouped Ethiopia's eleven regions into four livelihood categories: Highland (Tigray, Amhara, Benishangul-Gumuz), Pastoral (Afar, Somali, Gambela), Mixed (Oromia, Southern Nations, Nationalities and Peoples (SNNP)), and Urban (Addis Ababa, Dire Dawa, Harari). This classification was constructed for this robustness check based on commonly recognized agroecological and livelihood-system

distinctions and is not a validated livelihood-zone typology. It therefore simplifies varying degrees of sub-national heterogeneity. To assess this limitation, cluster-level livelihood-zone data from the Famine Early Warning Systems Network (FEWS NET) were examined to quantify within-region heterogeneity and to conduct a finer-grained stratified analysis using livelihood-zone type (LZTYPE). For LZTYPE categories represented across multiple regions, models additionally adjusted for region fixed effects. These analyses provide a finer-grained robustness check of whether the observed vulnerability–stunting association is preserved across different livelihood systems.

These comparisons across livelihood zones are reported as a supplementary robustness check and were not treated as a separate family of formal hypothesis tests. Accordingly, they were not subject to the Benjamini-Hochberg correction applied to the primary four-region stratification. Their purpose is to assess whether the direction and magnitude of the CVI–stunting association are consistent at a finer spatial resolution, rather than to establish independent statistical significance for each livelihood zone.

We also assessed the robustness of the CVI construction using a decomposed model in which exposure, sensitivity, and adaptive capacity were included as separate predictors. In addition, a categorical interaction between exposure and adaptive capacity tier (none, low, or higher capacity) was tested as an alternative to the continuous multiplicative form. A continuous exposure × adaptive capacity interaction was not fitted because exposure varies across only 33 region year units, providing insufficient variation to reliably estimate the interaction (*Appendix Table A27*). A continuous HAZ score outcome model was also estimated as an additional robustness check (*Appendix Table A24*). Multicollinearity was assessed using Variance Inflation Factors (VIF). All VIF values ranged from 1.00 to 1.16, indicating no meaningful collinearity. Selection bias due to missing anthropometric data, as defined in Section 2.1 and illustrated in Figure 4, was evaluated using logistic regression to compare included and excluded children (*Appendix Table A31*).

To estimate the potential impact of reducing climate vulnerability on stunting prevalence, we conducted counterfactual simulations using the interaction model. For each child in the primary model's fitted sample (n = 23,053), we generated predicted stunting probabilities, including region and soil class random effects, under two scenarios: (1) observed climate vulnerability and (2) a hypothetical 25% reduction in the raw CVI. The reduction was applied

to the raw CVI before its square-root transformation and z-score standardization, with all other covariates held constant. Thus, the counterfactual represents a 25% reduction in the underlying raw CVI rather than a 25% reduction in the standardized exposure. We selected this magnitude as a substantial but potentially achievable intervention scenario to assess the potential policy impact of reducing climate vulnerability, without prespecifying the mechanism through which the reduction might be achieved (e.g., reduced exposure, improved sanitation, or increased adaptive capacity). Predicted probabilities under each scenario were averaged using pooled sampling weights at the region-year and national levels, and the difference between scenarios was used to quantify the corresponding change in stunting prevalence.

All analyses incorporated DHS sampling weights. For the livelihood-zone stratified analyses and the alternative CVI formulation comparisons, weights were rescaled within survey round so that each round's total weight mass equaled its own sample size, ensuring equal representation across rounds within each subgroup. A sensitivity check compared the CVI coefficient under this rescaled weighting scheme against the primary specification's round-specific weights (*Appendix Table A28*). Unless otherwise stated, statistical significance was assessed as follows: ***p < 0.001, **p < 0.01, *p < 0.05, and †p < 0.1. Confidence intervals are presented at the 95% level.[15]

All models were implemented in R version 4.6.1 (R Core Team, 2026), using the *terra* and *raster* packages for satellite-derived climate data extraction; *lme4* for mixed-effects modeling; *survey*, *sandwich*, and *lmtest* for design-based and cluster-robust standard error estimation; *car* for multicollinearity diagnostics; *ggplot2* for data visualization; *sf* for spatial mapping; and *spdep* for spatial analysis.

3. Results

3.1 Descriptive Statistics

3.1.1 Sample Characteristics

[15] Post-hoc observed power was not calculated because it is determined by the observed test result and provides no additional information for interpreting the evidence beyond the p-value (Hoenig & Heisey, 2001).

The analytical sample included children from three EDHS waves (Table 10). The survey-weighted prevalence of stunting decreased from 44.3% in 2011 to 38.4% in 2016 and 36.9% in 2019. Mean HAZ showed a broadly similar, though non-monotonic, pattern, improving from −1.68 in 2011 to −1.44 in 2016 before declining slightly to −1.51 in 2019. Several sociodemographic and household-level indicators also improved over the study period. The proportion of mothers with no formal education decreased from 69.7% to 54.8%, rural residence declined from 83.9% to 76.9%, and the prevalence of unimproved sanitation decreased from 51.3% to 41.5%, although it remained common throughout the study period. Household size, breastfeeding prevalence, household wealth quintile, and mean standardized climate vulnerability (CVI) remained relatively stable across survey waves, with no clear monotonic trend. The prevalence of unimproved water sources followed a less consistent pattern, decreasing from 21.0% in 2011 to 13.4% in 2016 before increasing to 19.0% in 2019.

Table 10. Sample Characteristics by Survey Wave

Characteristic	2011	2016	2019
Total children	9,611	8,855	5,076
Stunted children (%, weighted)	44.3	38.4	36.9
Stunted children (%, unweighted)	42.3	36.4	35.8
Mean HAZ, weighted (SD)‡	-1.68 (1.76)	-1.44 (1.75)	-1.51 (1.55)
Mothers with no formal education (%)	69.7	63.7	54.8
Child age (months)[a]	29.4	28.5	28.8
Child sex (% male)	50.9	50.9	51.1
Household wealth quintile	2.7	2.6	2.7
Residence (% rural)	83.9	81.6	76.9
Household size (mean members)	6.2	6.1	6.2
Currently breastfeeding (%)	69.8	67.1	69.2
Unimproved water source (%)	21.0	13.4	19.0
Unimproved sanitation (%)	51.3	43.5	41.5
Standardized CVI, mean (SD)[b]	0.04 (0.98)	-0.06 (0.93)	0.03 (1.14)

***Notes**: [a]Mean age among children aged 0–59 months. [b]CVI standardized to mean = 0 and SD = 1 using the full pooled analytical sample (n = 23,542) to ensure comparability across survey waves. ‡HAZ mean is survey-weighted, SD is unweighted.*

3.1.2 Regional Patterns in Climate Vulnerability and Stunting

Climate vulnerability was calculated at the household level and then averaged within regions (Section 2.2), so regional CVI values represent mean household-level vulnerability. Regional patterns showed substantial variation in both mean CVI and stunting prevalence (*Table 11 and Appendix Figure A1*). Afar had the highest mean CVI (0.884), followed by Somali (0.440), and Amhara (0.370). Afar's mean CVI was approximately twice that of Somali.

Stunting prevalence was highest in Amhara (48.3%), followed by Tigray (47.0%) and Afar (46.6%).

The regional rankings for vulnerability and stunting were not consistently aligned. Afar ranked first in mean CVI but third in stunting prevalence, whereas Amhara ranked third in mean CVI but first in stunting prevalence. Somali also had relatively high mean CVI (0.440) but a lower stunting prevalence (29.8%). Tigray had a relatively low mean CVI (0.280) alongside a high stunting prevalence (47.0%). Addis Ababa had both the lowest mean CVI (0.023) and the lowest stunting prevalence (16.9%). These findings indicate substantial regional variation in both climate vulnerability and stunting prevalence, with the two measures not consistently aligned across regions.

Table 11. Regional Climate Vulnerability and Stunting Prevalence

Region	Mean Exposure	Mean Sensitivity	Mean Adaptive Capacity	Mean Vulnerability (CVI)	Mean Stunting	Total Children
Afar	0.172	0.591	0.106	0.884	0.466	2,290
Somali	0.102	0.539	0.154	0.440	0.298	2,447
Amhara	0.137	0.440	0.245	0.370	0.483	2,420
Dire Dawa	0.135	0.359	0.379	0.300	0.351	1,354
SNNP	0.120	0.456	0.293	0.294	0.408	3,066
Oromia	0.111	0.470	0.290	0.281	0.383	3,534
Tigray	0.111	0.435	0.291	0.280	0.470	2,444
Benishangul-Gumuz	0.101	0.448	0.233	0.268	0.455	2,021
Gambela	0.103	0.484	0.295	0.259	0.249	1,629
Harari	0.104	0.355	0.488	0.139	0.311	1,358
Addis Ababa	0.176	0.107	0.764	0.023	0.169	979

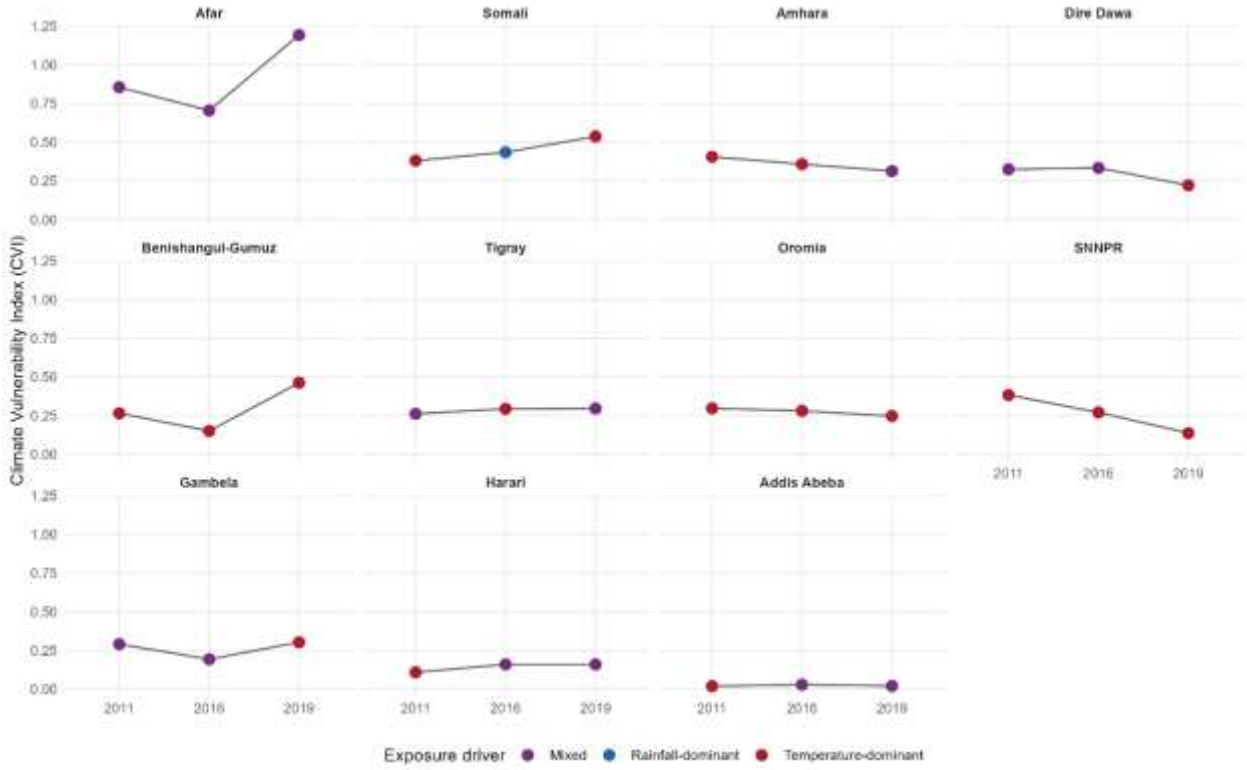
***Notes**: Regions are ordered by mean vulnerability (CVI), from highest to lowest.*

Figure 5 shows substantial regional variation in both the level and composition of CVI across the 2011, 2016, and 2019 survey rounds. Regions are ordered by mean CVI, with Afar, Somali, and Amhara identified as the three regions with the highest mean vulnerability. Afar consistently exhibited the highest CVI and showed a decline between 2011 and 2016 followed by a marked increase in 2019. Somali showed a progressive increase in CVI across the three rounds, while Amhara showed a gradual decline.

The composition of climate exposure also varied considerably across regions and over time. Afar remained relatively balanced between rainfall and temperature exposure, whereas Somali shifted sharply between temperature and rainfall dominance across survey rounds. Amhara was predominantly temperature-driven in the earlier rounds but became more

balanced by 2019. Across the remaining regions, exposure composition likewise varied, with some region-year observations strongly temperature dominant, others rainfall dominant, and others showing a mixed contribution.

Figure 5. Regional climate vulnerability trajectories and exposure composition by survey year.



***Notes**: Each panel represents one region and shows CVI across the 2011, 2016, and 2019 survey rounds. Point colors indicate the dominant exposure component. Blue = rainfall dominant (rainfall ≥70%), red = temperature dominant (rainfall ≤30%), and purple = mixed exposure. Numerical rainfall and temperature percentages underlying the classification are reported for each region-year observation.*

Descriptive region-survey year scatterplots (*Appendix Figure A2*) showed a positive association between sensitivity and stunting and an inverse association between adaptive capacity and stunting. The composite CVI was positively associated with regional stunting, whereas raw climate exposure alone showed a weak and non-monotonic pattern.

3.2 Econometric Analysis

3.2.1 Main Association Between Climate Vulnerability and Stunting

In the primary multilevel logistic regression model, higher climate vulnerability was associated with higher odds of child stunting (Table 12). Because the model included a CVI × survey-year interaction, the main CVI estimate represents the association at the reference survey year. Each standard-deviation increase in CVI was associated with 27% higher odds of stunting at the reference survey year (OR = 1.27, $p < 0.001$, 95% CI: 1.23–1.31).

A model excluding the CVI × survey-year interaction was estimated as a comparison (*Appendix Table A22*). The interaction model had slightly better fit than the no-interaction model (AIC = 29,074 vs. 29,084) and was therefore retained as the primary specification. The estimated CVI association was similar in both models (OR = 1.27 with the interaction vs. OR = 1.25 without the interaction). Thus, including the temporal interaction did not materially change the estimated main association between CVI and stunting.

Because the model-based standard errors do not account for the clustered and stratified DHS sampling design, additional sensitivity analyses were conducted using design-based and cluster-robust standard errors (*Appendix Table A23*). These analyses used models that matched the fixed-effect structure of the primary specification as closely as possible. The resulting CVI coefficients were similar across the three approaches, ranging from $\beta = 0.235$ to 0.255. The standard errors were larger under the design-based and cluster-robust approaches than under the model-based approach. Nevertheless, the association between CVI and stunting remained positive and statistically significant in all three specifications.

For the primary model coefficient, applying the design-based standard error yielded a 95% confidence interval corresponding to an OR of 1.20–1.34. This interval is reported in Table 12 as a design-based estimate. It should be distinguished from the confidence interval obtained directly from the design-based model in *Appendix Table A23*, which was OR = 1.22–1.36. The difference arises because the former combines the primary model coefficient with the design-based standard error, whereas the latter is based on the coefficient and standard error estimated jointly in the design-based model.

The CVI × survey-year interaction was statistically significant under model-based inference (OR = 1.02, 95% CI: 1.01, 1.03). However, the design-based confidence interval included the null boundary (OR = 1.02, 95% CI: 1.00–1.03). These results provide limited evidence that the magnitude of the CVI–stunting association varied across survey years, with weaker support under the more conservative design-based specification.

A decomposed extension of this model, in which exposure, sensitivity, and adaptive capacity were each interacted separately with survey year, indicated that this interaction was not attributable to climate exposure (β = 0.052, p = 0.851) but was concentrated in the adaptive-capacity component (β = −0.069, p < 0.001), with a directionally similar but non-significant pattern for sensitivity (β = −0.041, p = 0.151) (*Appendix Table A26*). These results indicate that the temporal interaction observed in the composite model reflects a strengthening protective association between adaptive capacity and stunting over time, rather than a strengthening relationship between climate exposure and stunting.

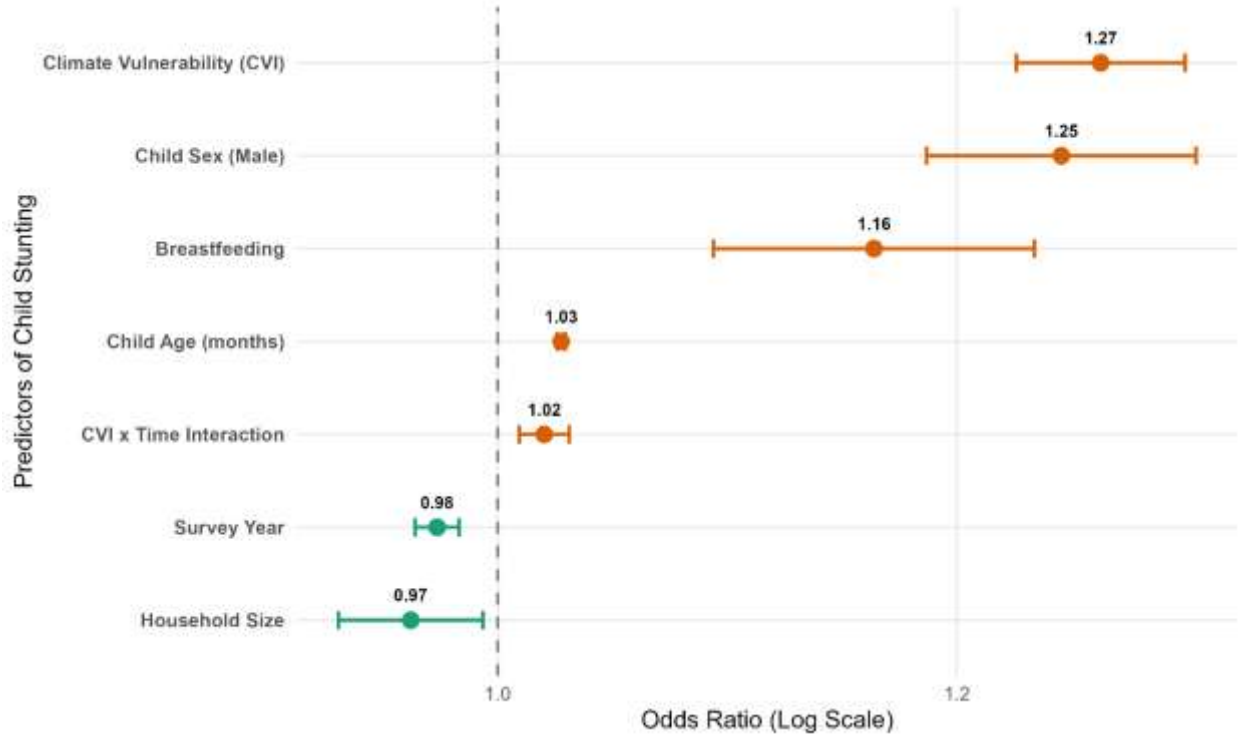
Among the covariates, child age was positively associated with stunting. Each additional month of age was associated with 3% higher odds of stunting (OR = 1.03, p < 0.001). Male children had higher odds of stunting than female children (OR = 1.25, p < 0.001). Breastfeeding was also positively associated with stunting after adjustment for child age (OR = 1.16, p < 0.001). Household size showed a small inverse association with stunting (OR = 0.97, p = 0.018). Figure 6 provides a visual summary of the estimated odds ratios (ORs) and corresponding 95% confidence intervals from the primary model.

Table 12. Temporal Interaction Model Results (Primary Model)

Variable	Coef. (β)	SE	z value	p-value	OR	95% CI (OR, model-based)	95% CI (OR, design-based)
Intercept	-1.582	0.146	-10.86	<0.001	0.21	[0.15, 0.27]	--
CVI (std.)	0.239	0.017	14.01	<0.001	1.27	[1.23, 1.31]	[1.20, 1.34]
Survey year	-0.024	0.004	-5.40	<0.001	0.98	[0.97, 0.98]	[0.96, 0.99]
Breastfeeding	0.149	0.033	4.59	<0.001	1.16	[1.09, 1.24]	[1.05, 1.28]
Household size	-0.035	0.015	-2.36	0.018	0.97	[0.94, 0.99]	[0.92, 1.01]
Child age (months)	0.025	0.001	29.36	<0.001	1.03	[1.02, 1.03]	[1.02, 1.03]
Child sex (male)	0.224	0.027	8.20	<0.001	1.25	[1.19, 1.32]	[1.15, 1.36]
CVI × Year	0.018	0.005	3.66	<0.001	1.02	[1.01, 1.03]	[1.00, 1.03]

***Notes**: AIC = 29,074; BIC = 29,155. Region variance = 0.084 (SD = 0.290); soil-class variance = 0.069 (SD = 0.263). The analysis included 23,053 children across 11 regions and 7 soil classes. The design-based confidence interval for CVI applies the design-based standard error to the CVI coefficient from the primary model.*

Figure 6. ORs for Climate Vulnerability and Individual-Level Covariates in the Primary Stunting Model



***Notes**: Points show ORs with 95% confidence intervals from the temporal interaction model. Risk factors are shown in red, protective factors in green, and non-significant effects in grey. Models control for region and soil class as random effects.*

3.2.2 CVI–Stunting Associations Across Broad Livelihood Zones

Having established a positive overall association between climate vulnerability and stunting, the analysis next examined whether this association was consistent across broad livelihood zones. These zones capture differences in livelihood and environmental contexts that may shape children's exposure and vulnerability to climate-related stressors (Section 2.3).

Stratified analyses showed a positive association between CVI and stunting in all four livelihood zones (Table 13). The largest point estimate was observed in the Urban zone (OR = 1.70, 95% CI: 1.22–2.38, raw p = 0.0017), followed by the Highland zone (OR = 1.30, 95% CI: 1.23–1.37, p < 0.001) and Mixed zone (OR = 1.29, 95% CI: 1.24–1.35, p < 0.001). The Pastoral zone had the smallest point estimate, although the association remained statistically significant (OR = 1.20, 95% CI: 1.05–1.36, p = 0.0067). Thus, the direction of the association was consistent across all four livelihood zones, although the magnitude of the point estimates varied.

After Benjamini–Hochberg adjustment for the four livelihood-zone comparisons, all four associations remained statistically significant. The largest adjusted p-value was 0.0067 for

the Pastoral zone, indicating that the findings remained statistically significant after accounting for multiple comparisons.

Table 13. Vulnerability Effects Across Livelihood Zones

Livelihood Zone	Sample Size	β (CVI)	SE	aOR	95% CI (OR)	Raw p-value	BH-adjusted p-value
Highland	6,885	0.261	0.029	1.30	1.23–1.37	<0.0001	<0.0001
Pastoral	6,366	0.179	0.066	1.20	1.05–1.36	0.0067	0.0067
Mixed	6,600	0.256	0.022	1.29	1.24–1.35	<0.0001	<0.0001
Urban	3,691	0.533	0.169	1.70	1.22–2.38	0.0017	0.0022

***Notes**: GLM models include region as a fixed effect and are weighted by round-normalized weights (see Section 2.3). Sample sizes reflect the total number of children in each livelihood zone (23,542). Models use complete cases and exclude 489 children with missing breastfeeding status, consistent with the primary model (Table 12).*

3.2.3 Variation by Livelihood-Zone Type

The broad livelihood-zone analysis showed that the positive CVI–stunting association was present across all four zones, although the magnitude of the point estimates varied. Because these broad categories encompass regions with differing agroecological and production characteristics, the analysis next used the finer-grained FEWS NET livelihood-zone type (LZTYPE) classification to examine the association across more specific livelihood systems. The LZTYPE classification does not map directly onto the four broad zones, and the Urban category was not represented in the analysis because it occurs only in Addis Ababa and does not meet the requirement for representation in at least two regions.

FEWS NET data showed considerable variation in livelihood-zone classification within regions. The number of distinct LZCODEs ranged from 1 in Addis Ababa to 44 in Oromia, with 41 in SNNP, 27 in Amhara, and 18 in Somali. The composition of LZTYPES also varied across regions. Oromia contained five LZTYPES, Somali four, and Afar, Gambela, and SNNP three each.

In models stratified by LZTYPE and adjusted for region, the CVI–stunting association was positive across the three cropping and agropastoral categories (Table 14). The association was strongest in Agropastoral zones, followed by Meher dominant and Belg dominant cropping zones. In Pastoral zones, the association was smaller and not statistically significant. These results indicate a consistent positive association across cropping and agropastoral livelihood types, whereas evidence of an association was weaker in pastoral zones.

Table 14. Association between standardized CVI and child stunting by FEWS NET livelihood-zone type, with region fixed effects.

LZTYPE	N	OR	95% CI	p-value
Cropping: Belg Dominant	1,668	1.27	1.15–1.41	<0.001
Cropping: Meher Dominant	14,978	1.30	1.25–1.35	<0.001
Pastoral	2,674	1.09	0.91–1.30	0.367
Agropastoral	2,061	1.34	1.10–1.62	0.004

***Notes**:*

- *Models were adjusted for breastfeeding, standardized household size, child age, and child sex. The analysis was restricted to LZTYPES represented in at least two regions and with at least 500 children. Belg refers to Ethiopia's short rainy season (approximately February–May) and the associated minor cropping cycle. Meher refers to the main rainy season (approximately June–September) and the associated major cropping cycle, which accounts for most of Ethiopia's annual cereal production.*
- *Sample sizes reflect complete cases for the covariates included in each model. The Urban livelihood-zone type is excluded because it is represented in only one region (Addis Ababa) and does not meet the eligibility criterion of representation in at least two regions. An additional 675 children lack a recorded LZTYPE due to missing cluster-level GPS coordinates. Models further exclude children with missing breastfeeding status via listwise deletion, consistent with the exclusions described for Table 12.*

3.2.4 Robustness and Alternative Specifications

The primary analysis used a binary stunting outcome and a composite climate vulnerability index. Several additional analyses were conducted to assess whether the main finding was sensitive to the choice of outcome, the representation of CVI components, and the formulation of the vulnerability index.

Using continuous HAZ as the outcome, the direction of the association was consistent with the primary analysis (*Appendix Table A24*). A one-standard-deviation increase in CVI was associated with a 0.186-standard-deviation decrease in HAZ ($\beta = -0.186$, SE = 0.013, $t = -14.20$, $p < 0.001$). The CVI × survey-year interaction was also statistically significant under model-based standard errors ($\beta = -0.0087$, $t = -2.28$, $p \approx 0.023$), providing some evidence that the CVI–HAZ association varied across survey years.

The analysis next examined the three components of CVI separately to assess whether the composite association was driven primarily by one dimension (*Appendix Table A25*). Sensitivity was positively associated with stunting (OR = 1.87, 95% CI: 1.55–2.24, $p < 0.001$), while adaptive capacity was inversely associated with stunting (OR = 0.53, 95% CI: 0.47–0.59, $p < 0.001$). Climate exposure was not independently associated with stunting (OR = 0.35, 95% CI: 0.06–2.15, p = 0.259). Adding exposure to a model containing sensitivity and adaptive capacity did not materially improve model fit (likelihood-ratio test: $\chi^2 = 0.82$, df = 1, p = 0.37). These results indicate that sensitivity and adaptive capacity contributed more strongly to the adjusted association than the exposure component in this specification.

As an additional robustness check, adaptive capacity was recoded into three tiers (none, low, and higher) and interacted categorically with climate exposure (*Appendix Table A27*). This specification was used as an alternative to the continuous exposure × adaptive-capacity interaction, which could not be estimated reliably because climate exposure varies across only 33 region-year units. The categorical model had an AIC of 28,997, which was 10 points higher than that of the decomposed specification (*Appendix Table A25*). Sensitivity remained a strong independent predictor of stunting (OR = 2.07, $p < 0.001$), consistent with the decomposed model. Among children in the higher-capacity reference category, climate exposure was not significantly associated with stunting (OR = 0.12, $p = 0.138$). Neither interaction term was statistically significant (exposure × low capacity: OR = 6.70, 95% CI: 0.20–226.32, $p = 0.289$; exposure × no capacity: OR = 5.64, 95% CI: 0.13–239.77, $p = 0.366$). Thus, there was no evidence that the association between climate exposure and stunting varied by adaptive-capacity tier. The wide confidence intervals for the interaction terms reflect the limited number of region-year exposure observations available to estimate these interactions.

To determine which dimension of vulnerability underlies the CVI × survey-year interaction reported in Table 12, exposure, sensitivity, and adaptive capacity were each interacted separately with survey year in an extension of the decomposed model (*Appendix Table A26*). The exposure × survey-year term was not statistically significant and the sensitivity × survey-year term was directionally consistent but did not reach significance. The adaptive-capacity × survey-year term was statistically significant (aOR = 0.93; CI: 0.90–0.97), indicating that the protective association between adaptive capacity and stunting strengthened across survey rounds. This pattern indicates that the composite CVI × survey-year interaction is driven primarily by a strengthening socioeconomic gradient in adaptive capacity rather than by a strengthening relationship between climate exposure and stunting.

Alternative formulations of CVI were then evaluated to determine whether the main association depended on the construction of the index (*Appendix Table A28*). The additive and weighted formulations both showed positive and statistically significant associations with stunting and had lower AIC values than the multiplicative formulation within this specification comparison. The multiplicative formulation was retained as the primary index because it was prespecified and reflects the conceptual structure of climate vulnerability as a

combination of exposure, sensitivity, and adaptive capacity. Because these alternative formulations were estimated using pooled sampling weights rescaled so that each survey round contributed equally to the analytic sample, rather than the round-specific survey weights used in the primary model, their AIC values are directly comparable with one another but not with the AIC of the primary model in Table 12.

The correlation structure of the alternative indices provided an additional basis for retaining the multiplicative formulation (*Appendix Table A29*). The multiplicative index showed moderate correlations with exposure (r = 0.317), sensitivity (r = 0.672), and adaptive capacity (r = −0.727). By contrast, the additive and weighted formulations were strongly correlated with sensitivity (r = 0.855 and 0.926, respectively) and adaptive capacity (r = −0.930 and −0.854), while showing near-zero correlations with exposure (r = 0.024 and 0.057). The multiplicative formulation therefore represented the three conceptual dimensions more evenly than the alternative formulations.

3.2.5 Spatial Clustering of Climate Vulnerability

The multilevel models presented above included region as a random intercept but did not explicitly model spatial dependence among regions. To assess whether regional climate vulnerability exhibited spatial clustering, global and local spatial autocorrelation in regional mean CVI was examined. Global spatial autocorrelation was assessed using both an analytical test of Moran's I and a Monte Carlo permutation test with 999 simulations. The analytical test indicated weak positive spatial autocorrelation that was not statistically significant (I = 0.098; expected I = −0.100; variance = 0.0281; standardized deviate = 1.179; p = 0.119). The Monte Carlo permutation test produced the same observed Moran's I (I = 0.098) and likewise indicated no statistically significant global spatial autocorrelation (observed rank = 909 of 1,000; p = 0.091). Thus, both approaches indicated weak positive but non-significant global spatial autocorrelation.

Local Indicators of Spatial Association (LISA) identified one nominally significant local association at the unadjusted α = 0.05 level. Tigray was classified as "Low-High" (local Moran's I = −0.302, p = 0.036), indicating a lower-vulnerability region surrounded by higher-vulnerability neighboring regions and therefore constituting a local spatial outlier rather than a High-High hotspot. The remaining ten regions were not statistically significant (*Appendix*

Table A30). With eleven separate regional tests conducted at an unadjusted 5% significance level, approximately one false-positive result would be expected by chance alone, so this local result should not be interpreted as strong evidence of a genuine spatial outlier. Tigray's "Low-High" classification likely reflects its sharp local contrast with neighboring Afar, which has the highest CVI in the dataset, rather than a broader pattern of spatial clustering. Given the small number of regions and the absence of significant global spatial autocorrelation, these findings do not warrant replacing the prespecified multilevel random-intercept specification with a more complex spatial model.

3.2.6 Counterfactual Simulation

Using the primary interaction model, a simulated 25% reduction in household-level CVI, with all other covariates held constant, was associated with a decrease in predicted national stunting prevalence from 40.5% to 39.4%, corresponding to an absolute reduction of 1.07 percentage points (pp) and a relative reduction of 2.81%.

The magnitude of the simulated reductions varied substantially across regions and survey years. The largest reductions were generally observed in regions with higher estimated climate vulnerability. Afar, Somali, and Amhara had the three highest regional mean CVI values and showed the largest reductions across survey rounds. In Afar, the reduction increased from 1.38 pp in 2011 to 1.97 pp in 2016 and 2.97 pp in 2019. The 2019 reduction was the largest observed in any region-year. Somali showed a similar increasing pattern, with reductions of 0.79, 1.39, and 1.87 pp. Amhara showed reductions of 0.94, 1.35, and 1.40 pp. Tigray had a comparatively moderate mean CVI but also showed a consistent increase across rounds, from 0.76 pp in 2011 to 1.22 pp in 2016 and 1.33 pp in 2019. Urban regions showed the smallest reductions. In Addis Ababa, reductions ranged from 0.11 to 0.20 pp across survey rounds. Full region-year results for all eleven regions are reported in *Appendix Table A33* and *Appendix Figure A3*.

4. Discussion

Using three waves of the EDHS (2011, 2016, and 2019) linked to cluster-specific climate data, this study examined the association between climate vulnerability, measured by the CVI, and child stunting in Ethiopia. It also assessed whether this association varied over time and across livelihood contexts.

Descriptively, mean raw CVI increased from 0.355 in 2011 to 0.384 in 2019, a period that encompassed the severe 2015–2016 and 2016–2017 droughts (Gebremeskel Haile et al., 2019; Philip et al., 2018). Over the same period, national stunting prevalence declined from 44.3% to 36.9%. This divergence is not necessarily paradoxical, as sustained economic growth and expanded social protection, including the Productive Safety Net Programme, may have helped buffer some of the potential nutritional consequences of increasing climate vulnerability (World Bank, 2020). However, such buffering may nonetheless be fragile, as social protection systems face growing pressure as climate-related poverty intensifies. Most importantly, these mechanisms were not directly examined in this study and therefore cannot be interpreted as causal explanations for the observed divergence.

Substantial regional variation was also observed in both climate vulnerability and stunting prevalence (*see Table 11 and Figure 5 for details*). However, regions with higher climate vulnerability did not consistently have higher levels of stunting. This suggests that climate vulnerability alone does not fully account for regional differences in child nutritional outcomes and that other socioeconomic, demographic, health, and environmental factors may also play important roles in explaining regional variation in stunting.

In the primary multilevel model, each standard-deviation increase in CVI was associated with 27% higher odds of stunting (Table 12; aOR = 1.27, design-based 95% CI: 1.20–1.34). The CVI × survey-year interaction was also positive and statistically significant when model-based standard errors were used (Table 12; aOR = 1.02, model-based 95% CI: 1.01–1.03). However, after accounting for the DHS's clustered and stratified sampling design, the point estimate remained above the null, although the confidence interval included the null (design-based 95% CI: 1.00–1.03), indicating that evidence for the interaction was sensitive to the method used to estimate uncertainty. Decomposing this interaction by CVI component clarified its source (see Appendix Table A26). The adaptive-capacity × survey-year term was significant ($\beta = -0.069$, $p < 0.001$), whereas the exposure × survey-year and sensitivity × survey-year terms were not. This pattern indicates that the temporal interaction reflects a strengthening protective association between adaptive capacity (maternal education and household wealth) and stunting across survey rounds, rather than a strengthening relationship between climate exposure and stunting.

This finding is consistent with IPCC vulnerability frameworks (IPCC, 2014a) and with Tirado et al. (2013), who describe how climate-related vulnerability can undermine food and nutrition security and increase the risk of undernutrition, particularly among vulnerable populations. It also aligns with prior Ethiopian evidence linking drought exposure to elevated stunting (Davenport et al., 2017; Randell et al., 2021). This is also consistent with a recent scoping review identifying stunting as the nutritional outcome most consistently affected across climate-hazard types in SSA and South Asia (Klapka et al., 2026). Comparable Ethiopian estimates include a 28% increase in the odds of severe stunting per 1°C rise in prenatal temperature (Randell et al., 2020), plausibly operating through heat-stress pathways such as placental damage and heat-induced preterm labor (Sheffield & Landrigan, 2011), and an approximately 8% increase in wasting risk associated with a substantial rise in temperature anomalies across SSA (Thiede & Strube, 2020).

This interpretation is consistent with Ethiopia's documented socioeconomic gains over the study period, including increases in maternal education and household wealth between 2011 and 2019, and with education's established role as a protective factor whose benefits may increase as returns to schooling and access to information expand (Striessnig et al., 2013). It does not support the hypothesis that the association between climate exposure and stunting intensified over time or that households' buffering capacity failed to keep pace with climate stress. If anything, the component-specific decomposition above suggests the opposite pattern. The significant CVI × survey-year interaction in the continuous HAZ model (Appendix Table A24) should likewise be interpreted cautiously, as it may reflect the same underlying adaptive-capacity trend rather than a climate-specific temporal dynamic. Longer and component-resolved exposure series would nevertheless be useful for assessing whether this pattern persists as additional survey rounds become available. Overall, the present evidence does not support a climate-specific interpretation of the temporal interaction.

Decomposing the CVI into its three components clarified which dimensions drive the composite association. Sensitivity and adaptive capacity were each independently associated with stunting, while climate exposure alone was not (OR = 0.35; p = 0.259; Appendix Table A25). The descriptive component plots (Appendix Figure A2) provide some context for this finding. High-exposure pastoralist regions such as Afar show lower stunting than several lower-exposure highland regions such as Amhara and Tigray. This pattern is consistent with,

although it does not establish, the possibility that adaptive capacity shapes how climatic exposure translates into nutritional risk. Exposure alone therefore appears insufficient to explain differences in nutritional outcomes. The same level of climatic exposure may have different nutritional consequences depending on households' sensitivity and capacity to adapt.

This pattern supports the choice of a multiplicative CVI formulation, which treats adaptive capacity as a modifier of the relationship between exposure and sensitivity rather than as an additive offset that can compensate for a low score on another component. Comparative work on IPCC-based composite indicators shows that additive formulations allow greater substitution across components, whereas multiplicative formulations ensure that a low value on one dimension constrains the overall score. This is consistent with the logic underlying the UNDP's move from an additive to a multiplicative structure in the Human Development Index (Choi, 2019).

The correlation structure provides empirical support for this choice (Appendix Table A29). The multiplicative CVI showed moderate correlations with all three components, including exposure ($r = 0.317$), sensitivity ($r = 0.672$), and adaptive capacity ($r = -0.727$). By contrast, the additive and weighted alternatives correlated much more strongly with sensitivity and adaptive capacity and showed near-zero correlations with exposure. This pattern suggests that these alternatives may function more as proxies for socioeconomic vulnerability than as genuinely multidimensional measures of climate vulnerability.

This balance should, however, be understood as a conceptual and measurement advantage rather than a statistical one. The decomposed specification had a lower AIC (28,987) than the composite multiplicative model (29,074), and, in a separate comparison using pooled survey weights, the additive and weighted formulations also showed lower AIC (27,513 and 27,528, respectively) than the multiplicative formulation (27,655). The multiplicative CVI was therefore retained because it preserves the theoretical structure of vulnerability as a joint function of exposure, sensitivity, and adaptive capacity.

The CVI–stunting association varied by livelihood context, indicating that the pathway from climate vulnerability to child nutrition depends on how households make a living. The largest association across all zones was observed in Urban areas (aOR = 1.70; 95% CI: 1.22–2.38; $p = 0.002$). This finding is particularly notable because Addis Ababa, Ethiopia's capital and

one of the country's two chartered city administrations, has the highest mean climate exposure among the eleven regions (Table 11). Two bodies of evidence suggest that the association is plausible. First, urban households may be particularly vulnerable to climate shocks transmitted through food-price volatility rather than through direct local production losses (Cohen & Garrett, 2010; Tacoli et al., 2013), indicating that climate-related nutritional effects need not operate primarily through crop failure. Second, multi-country evidence indicates that urban children may experience temperature-related malnutrition risks comparable to those of their rural counterparts (Thiede & Strube, 2020), suggesting that urban populations are not necessarily insulated from climate-related nutritional stress by their predominantly non-agrarian livelihoods. The relatively wider confidence interval for the Urban estimate is consistent with the smaller sample size in this zone and therefore reflects greater statistical uncertainty rather than necessarily a less meaningful association.

Among the three agrarian zones, Highland and Mixed livelihood systems showed the strongest associations (aOR = 1.30 and 1.29 respectively; both $p < 0.001$). Both patterns fit expectations from the agronomic literature: highland cereal systems are particularly temperature-sensitive, with warming above 2°C projected to reduce maize yields in Ethiopia's Central Rift Valley (Kassie et al., 2015), while mixed crop-livestock systems face a compounded exposure pathway, since a single climate shock can simultaneously depress crop yields and livestock productivity, amplifying income loss and narrowing dietary diversity (Thornton et al., 2014).

Pastoral zones showed the smallest, although still statistically significant, association (aOR = 1.20; 95% CI: 1.05–1.36; $p = 0.007$). This finding is notable because the zone includes Afar, one of the regions with the highest mean climate exposure in the sample. One plausible explanation is that pastoral mobility and livelihood diversification may provide households with a degree of buffering against climate shocks that is less available in other livelihood systems. This interpretation is consistent with evidence that wealthier Ethiopian pastoralists manage climate and livelihood risks partly through diversification into livestock and urban assets (Atalell et al., 2023; Coppock et al., 2018). If such strategies provide partial protection, conventional vulnerability proxies may overstate the extent to which exposure translates into nutritional risk, while understating the potential role of adaptive capacity (Klapka et al., 2026). However, because this study does not directly measure mobility, asset diversification,

or other buffering mechanisms, this explanation should be interpreted as a plausible account consistent with the observed pattern rather than as a mechanism demonstrated by the analysis. This pastoral finding did not replicate when we used the finer-grained, cluster-level FEWS NET livelihood typology (LZTYPE) rather than the region-based grouping (*Table 14*). Within pastoral clusters specifically, the association between CVI and stunting was smaller and no longer statistically significant. We therefore flag this discrepancy rather than attempt to resolve it. The two classifications differ in both their composition and spatial precision, and the available data do not allow us to determine which better captures the livelihood context most relevant to pastoral vulnerability. The discrepancy therefore underscores the sensitivity of the pastoral-zone finding to how livelihood context is defined and cautions against interpreting the region-based result as definitive evidence of a pastoral-specific association.

These findings from the livelihood-context and finer-grained analyses of the CVI–stunting association suggest that climate–nutrition programming should be differentiated by livelihood context. The particularly strong association observed in Urban areas is consistent with, though does not prove, a role for interventions that protect household food access from climate-related market disruptions, including food-price volatility and reduced affordability and availability of diverse foods. However, the strong Urban CVI–stunting association cannot be cleanly attributed to a market-mediated pathway, whereby food-price transmission affects households even in the absence of high direct climate exposure, rather than to direct climate sensitivity. These mechanisms are confounded within the Urban zone as constructed here and cannot be disentangled with the present study design (see Limitations).

In Highland and Mixed livelihood zones, where associations were also strong and statistically significant, geographically targeted climate–nutrition strategies should complement climate-resilient agricultural measures with actions that protect household income and dietary diversity from climate-related production losses. By contrast, the pastoral finding is less suitable as a basis for zone-specific policy targeting because the association was smaller and did not replicate under the finer-grained LZTYPE classification. This inconsistency does not imply that pastoral populations are protected from climate-related nutritional risk but highlights the need for further research to determine whether differences in adaptive capacity

and livelihood strategies help explain variation in the relationship between climate vulnerability and child nutritional outcomes.

The spatial analysis found no significant global clustering of climate vulnerability and no significant LISA hotspots or coldspots. Tigray was the only significant spatial outlier, representing a lower-vulnerability region surrounded by higher-vulnerability neighbors. With only one significant outlier among eleven regions, this does not indicate a broader spatial pattern. The lack of broad spatial dependence suggests that vulnerability may be shaped more strongly by local household, livelihood, and agroecological conditions than by proximity to other high-vulnerability areas. This interpretation is consistent with spatial and multilevel analyses of Ethiopian DHS data, which show that stunting clustering in Ethiopia is driven by local household and community-level factors, including wealth, maternal education, and anemia, rather than by spatial proximity or cross-boundary spillover effects (Hailu et al., 2020). However, the spatial pattern may also be influenced by measurement limitations. Evidence from 35 SSA countries shows that DHS anthropometric data quality degrades systematically with distance from settlements, suggesting that apparent spatial uniformity in remote high-vulnerability areas may partly reflect measurement limitations rather than true epidemiological homogeneity (Seidler et al., 2025).

The absence of broad spatial dependence does not imply uniform nutritional risk across settings. Global evidence shows that rural stunting (39.6%) exceeds urban prevalence (25.2%) and that climate effects differ across production systems (FAO et al., 2024). The regional heterogeneity observed here therefore supports geographically differentiated monitoring and programming under the WHO SDG 2.2.1 framework (WHO, 2025b), rather than uniform national targeting.

Among the additional control variables, Child age (OR = 1.03 per month; CI: 1.02, 1.03) and male sex (OR = 1.25; CI: 1.15, 1.36) were both significant predictors. The age finding is consistent with the cumulative nature of linear growth faltering: growth deficits accumulate progressively with age and do not plateau, meaning older children carry a higher burden of stunting reflecting prolonged exposure to nutritional insults (Benjamin-Chung et al., 2023; Leroy et al., 2014). This cumulative property is also why stunting, among nutritional outcomes, tends to be more sensitive to sustained deprivation than to isolated shocks (Klapka et al., 2026). The male disadvantage is equally well established: in SSA, boys are consistently

more likely to be stunted than girls, with a pooled adjusted OR of 1.18 across 16 DHS surveys (Wamani et al., 2007), a pattern confirmed specifically in Ethiopia where boys had significantly higher odds of stunting than girls across four rounds of nationally representative data (Sahiledengle et al., 2023).

The positive association between breastfeeding and stunting after adjusting for age (OR = 1.16; CI: 1.05, 1.28) likely reflects reverse causality. Caregivers often continue breastfeeding when a child shows signs of poor growth, so that after age adjustment, continued breastfeeding captures compensatory behavior rather than a causal nutritional effect. Thus, cross-sectional breastfeeding status may proxy prior growth faltering rather than represent a harmful exposure once age is held constant. This pattern is well documented in the breastfeeding literature: in developing-country settings, poorly growing infants are more likely to receive complementary foods earlier because of slower growth, while infants who continue exclusive breastfeeding tend to remain on acceptable growth trajectories, meaning that reverse causality operates in both directions and confounds observational estimates of breastfeeding effects on growth (Kramer & Kakuma, 2012).

These individual-level associations are consistent with established determinants of child growth but do not materially alter the interpretation of the primary climate-vulnerability finding. The protective role of adaptive capacity within the CVI is also evident. This component is proxied by maternal education and household wealth. Its effect is consistent with the concept of education as a "social vaccine" (Striessnig et al., 2013), and is reinforced by cross-study evidence identifying caregiver education as the single most consistently protective socioeconomic moderator of the climate-nutrition relationship across 37 studies spanning SSA and South Asia, with higher maternal education associated with reduced vulnerability across all major climate hazard types (Klapka et al., 2026). It also supports the emphasis on human development in Ethiopia's Climate Resilient Green Economy strategy (UNICEF Ethiopia, 2025). However, because maternal education and household wealth are components of the CVI itself, their individual associations should not be interpreted as independent causal effects of adaptive capacity.

The counterfactual simulation further illustrates the potential policy relevance of reducing climate vulnerability. A simulated 25% reduction in the CVI produced modest but potentially meaningful reductions in predicted stunting at the national level, with larger reductions in

Afar, Somali, Amhara, and Tigray regions. The relatively small national effect, coupled with substantial regional heterogeneity in the counterfactual estimates, supports geographically differentiated, climate-sensitive nutrition strategies. Interventions that address adaptive capacity and underlying sensitivity may have greater potential in regions where vulnerability is high and the modeled response to a reduction in vulnerability is largest. Nevertheless, these simulations should be interpreted as model-based scenarios rather than estimates of the causal effect of reducing climate vulnerability.

Several limitations warrant consideration. First, while political conflict and internal displacement are established drivers of food and nutrition insecurity, conflict event data were excluded from the final model because they were available only at a spatially aggregated level that was not directly comparable with the climate and vulnerability measures. In Ethiopia, areas that were historically more food secure experienced relatively high levels of conflict and displacement during the study period, whereas some of the most climate-vulnerable regions experienced lower levels of conflict. Including aggregate conflict measures could therefore introduce spatial aggregation bias and make it difficult to distinguish the association of climate vulnerability with stunting from that of conflict. Conversely, excluding conflict leaves open the possibility of unmeasured confounding. The direction and magnitude of any resulting bias cannot be determined from the available data.

Second, children with higher CVI were modestly more likely to have missing anthropometric data (unadjusted), an association that persisted, though attenuated, after adjusting for wealth, education, rural residence, and survey year (Appendix Table A31). This is consistent with broader DHS evidence that children without usable height data skew rural, poorer, and less educated (Finaret & Hutchinson, 2018), and that data quality itself declines with remoteness (Seidler et al., 2025), meaning the most remote, plausibly highest-CVI children may be under-represented in the analytic sample. If excluded high-CVI children are also more likely to be stunted, this would attenuate rather than inflate the observed association, but the true direction of bias cannot be confirmed from missingness patterns alone.

Third, the symmetric treatment of positive and negative rainfall anomalies in the precipitation exposure measure may overestimate vulnerability in hyper-arid regions such as Afar, where above-average rainfall can be beneficial rather than harmful. Future work should explore directional anomaly measures weighted by agroecological context. In addition, the linear

specification of temperature exposure does not capture potential nonlinear or threshold-based effects of temperature on child nutrition outcomes.

Fourth, Addis Ababa, Ethiopia's capital and one of the country's two chartered city administrations, also has the highest mean climate exposure of the eleven regions studied. This is driven predominantly by its temperature-anomaly component rather than rainfall (Figure 5), reflecting the tighter normalization cap applied to temperature (2°C) than to rainfall (100% of baseline) in the exposure index. Consequently, the strong Urban CVI–stunting association cannot be cleanly attributed to a market-mediated pathway, whereby food-price transmission affects households even in the absence of high direct climate exposure, rather than to direct climate sensitivity. These mechanisms are confounded within the Urban zone as constructed here and cannot be disentangled with the present study design.

Fifth, climate exposure was assigned at the region-year (or cluster) level, whereas the CVI's sensitivity and adaptive-capacity components were constructed at the household level. This spatial mismatch means the exposure term may not reflect the specific conditions experienced by an individual child or household, and the findings should be interpreted as evidence of an association between contextual climate vulnerability and stunting, rather than as an estimate of the effect of a child's own climate exposure.

Last but most important, the repeated cross-sectional survey design precludes definitive establishment of temporal precedence between CVI and child stunting. Residual confounding, measurement error in the vulnerability index, spatial aggregation of climate indicators, and selection into the analytic sample may also contribute to bias in the estimated associations. Accordingly, these associations should not be interpreted as causal effects.

5. Conclusion and Implications

This study found a consistent positive association between climate vulnerability and child stunting in Ethiopia. Children living in more climate-vulnerable settings had higher odds of stunting, but the association was not explained by climatic exposure alone. When the components of the CVI were examined separately, sensitivity and adaptive capacity were more strongly associated with stunting than raw climate exposure, indicating that the nutritional consequences of climate stress depend substantially on the social and household conditions through which that stress is experienced. The association also varied across

livelihood contexts. The strongest association was observed in urban settings, including Addis Ababa, which also had high direct climate exposure, while the association was smaller in pastoral settings. Together, these findings suggest that climate-related nutritional risk is not determined simply by where climate hazards are greatest. Rather, market dependence, livelihood systems, household resources, and capacity to adapt shape how climatic stress may be translated into child nutritional harm.

These findings point to three actionable intervention priorities. First, climate vulnerability assessments should be integrated into nutrition program targeting, particularly in high-vulnerability regions and livelihood contexts. Second, strengthening adaptive capacity through investments in maternal education and household wealth should be considered central to climate adaptation and nutrition strategies, alongside broader investments in livelihood resilience supported by the wider literature on women's empowerment and climate adaptation. Third, urban food systems require climate-sensitive interventions, including food-price monitoring and resilient urban–rural supply chains, because urban populations may experience substantial market-mediated climate risks in addition to direct climate exposure, which was in fact highest in Addis Ababa among the eleven regions studied. These priorities shift the policy focus from responding to climate shocks alone toward reducing the underlying conditions that make children and households nutritionally vulnerable to those shocks.

Future research should use longitudinal designs and higher-resolution spatial, market, and livelihood data to establish causal pathways and determine whether the relative importance of sensitivity and adaptive capacity observed here generalizes to other climate-vulnerable settings. These findings have implications for progress toward several Sustainable Development Goals (SDGs), including SDG 1 (No Poverty), SDG 2 (Zero Hunger), SDG 3 (Good Health and Well-being), SDG 4 (Quality Education), SDG 5 (Gender Equality), SDG 6 (Clean Water and Sanitation), and SDG 13 (Climate Action), particularly SDG target 2.2 on ending all forms of child malnutrition by 2030 (United Nations General Assembly, 2015). By linking climate vulnerability with child nutritional outcomes, the findings underscore the need to align climate adaptation, nutrition programming, and investments in human development.

Although grounded in Ethiopia, the findings support a broader shift in climate-sensitive nutrition policy: from asking where climate exposure is greatest to asking where climate stress is most likely to translate into nutritional harm. Integrating climate vulnerability, rather than climate exposure alone, into nutrition targeting may therefore help direct limited resources toward populations with the greatest combined climatic and structural vulnerability.

7. Appendices

Table A19. DHS Sample Characteristics Across Survey Waves (2011-2019)

Survey Year	Sample Size	Valid Cases*	PSUs	Strata	Stunting (%)	Prevalence	Design Effect	Effective Sample Size
2011	11,654	9,611	596	23	44.3		3.19	3,008
2016	10,641	8,855	640	25	38.4		3.51	2,524
2019	5,753	5,076	305	21	36.9		3.14	1,618
Total	28,048	23,542						7,150

***Notes**: *Excludes observations with implausible HAZ values (< -6 SD or > 6 SD). Final analytical sample fitted by the primary model: 23,053 (489 children excluded, all with missing breastfeeding status, concentrated entirely in the 2019 round among children aged 42–59 months). Design effects were computed as the ratio of design-based variance (accounting for clustering, stratification, and weighting via svydesign) to simple random sampling variance for the stunting indicator.*

The 2019 Mini-DHS included fewer children than the earlier surveys. Diagnostic checks (covariate distributions, stunting-prevalence trends, and design-effect-adjusted effective sample sizes) found no evidence that the smaller 2019 sample compromised comparability across waves.

Table A20. Components and Weighting of CVI

Component	Sub-component	Measurement	Coding
Exposure	Precipitation anomaly	$\min(1, \frac{\lvert P_i - \bar{P} \rvert}{\bar{P}})$	pmin(1, abs(rainfall_anomaly) / 100)†
	Temperature anomaly	$\min(1, \frac{\lvert T_i - \bar{T} \rvert}{2})$	pmin(1, abs(temp_anomaly) / 2)
Sensitivity	Water access	Improved/unimproved/surface water	v113: improved sources=0, unimproved sources=0.5, Surface water=1
	Sanitation quality	Improved/unimproved/open defecation	v116: improved facilities=0, unimproved facilities=0.5, Open defecation=1
	Rural residence	Urban/rural binary	v025: Urban=0, Rural=1
	Birth interval risk	Preceding birth interval	bord==1 (firstborn)=0; b11≥36mo=0; 24–35mo=0.5; <24mo=1
	Female household head	Male/female head	v151: Male=0, Female=1
Adaptive Capacity	Maternal education	None to higher education	v106: None=0, Primary=0.33, Secondary=0.67, Higher=1
	Wealth quintile	Poorest to richest	v190: Poorest/poorer=0, Middle=0.5, Richer/richest=1

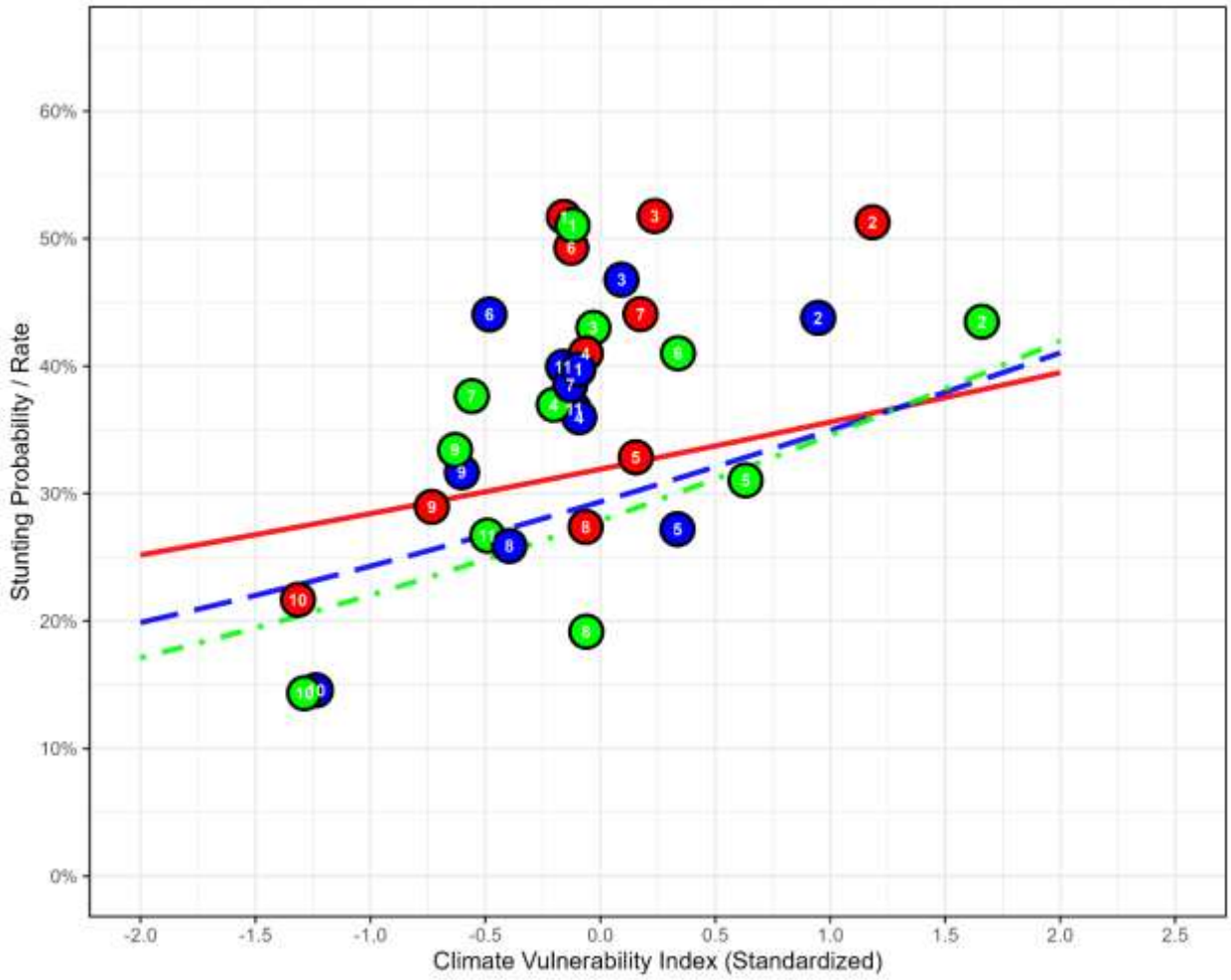
***Notes**: † rainfall_anomaly is the signed percentage deviation from each cluster's long-term mean rainfall, computed as 100 × $\frac{P_i - \bar{P}}{\bar{P}}$ (positive = wetter than baseline, negative = drought). Dividing by 100 converts this to proportion scale. The absolute value is applied separately, when computing the precipitation exposure sub-component. temp_anomaly follows the same underlying logic. It is a signed difference between the period mean and the baseline mean, measured in ºC. The absolute value is applied only at the temperature exposure sub-component step, rather than to the underlying anomaly itself. Unlike the percentage of baseline formulation, temp_anomaly is expressed on an additive ºC scale.*

Table A21. Household Level and Regional Mean Component CVI by Region-Survey Year (Aggregation Bias)

Region	Year	N	Mean Exposure	Mean Sensitivity	Mean Adaptive Capacity	CVI (Household Level)	CVI (Mean of Components)	Understatement (%)
Afar	2019	568	0.248	0.579	0.129	1.190	0.627	47.4
Afar	2011	893	0.160	0.606	0.112	0.855	0.457	46.5
Afar	2016	829	0.133	0.583	0.085	0.705	0.419	40.6
Somali	2019	538	0.105	0.602	0.093	0.537	0.328	38.9
Benishangul-Gumuz	2019	456	0.196	0.395	0.248	0.462	0.222	51.8
Somali	2016	1,154	0.098	0.533	0.143	0.434	0.215	50.4
Amhara	2011	1,068	0.139	0.453	0.214	0.405	0.201	50.4
SNNPR	2011	1,368	0.139	0.486	0.264	0.383	0.185	51.6
Somali	2011	755	0.107	0.501	0.213	0.381	0.171	55.1
Amhara	2016	887	0.138	0.439	0.266	0.358	0.165	54.0
Dire Dawa	2016	418	0.142	0.346	0.363	0.333	0.106	68.1
Dire Dawa	2011	584	0.143	0.396	0.343	0.324	0.128	60.6
Amhara	2019	465	0.132	0.411	0.275	0.313	0.145	53.8
Gambela	2019	386	0.150	0.475	0.365	0.302	0.153	49.2
Oromia	2011	1,515	0.115	0.489	0.299	0.297	0.140	52.7

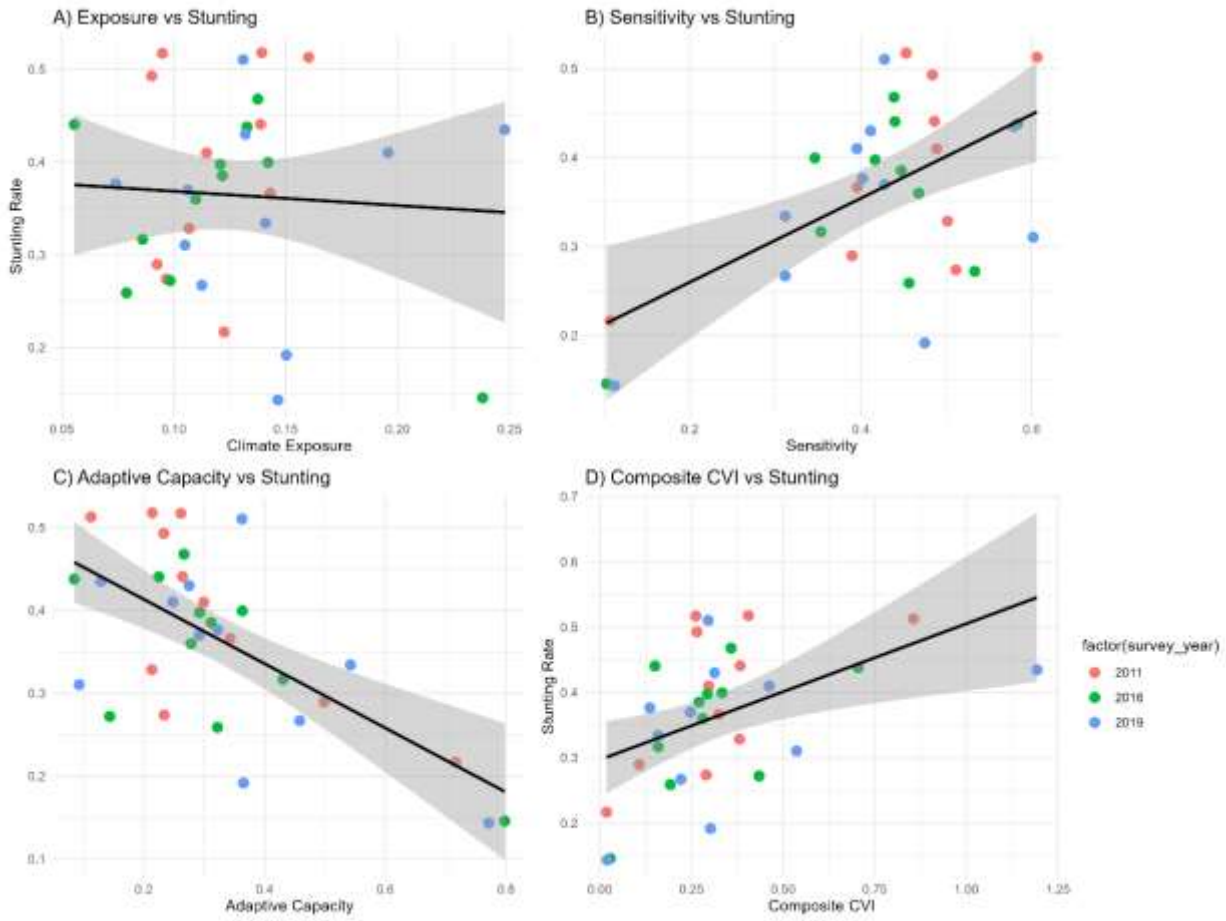
Tigray	2019	431	0.131	0.428	0.362	0.295	0.121	59.0
Tigray	2016	936	0.121	0.417	0.292	0.293	0.128	56.3
Gambela	2011	683	0.097	0.511	0.234	0.290	0.148	48.9
Oromia	2016	1,378	0.110	0.468	0.278	0.280	0.136	51.6
SNNPR	2016	1,103	0.122	0.447	0.312	0.270	0.132	51.1
Benishangul -Gumuz	2011	834	0.090	0.484	0.233	0.265	0.131	50.6
Tigray	2011	1,077	0.095	0.453	0.261	0.262	0.119	54.5
Oromia	2019	641	0.106	0.428	0.292	0.247	0.116	53.0
Dire Dawa	2019	352	0.113	0.312	0.458	0.221	0.063	71.6
Gambela	2016	560	0.079	0.456	0.322	0.192	0.085	55.6
Harari	2016	461	0.086	0.353	0.431	0.159	0.057	64.0
Harari	2019	386	0.141	0.312	0.542	0.159	0.068	56.9
Benishangul -Gumuz	2016	731	0.056	0.440	0.225	0.150	0.075	49.9
SNNPR	2019	595	0.074	0.402	0.323	0.136	0.070	48.3
Harari	2011	511	0.093	0.390	0.499	0.107	0.060	43.9
Addis Ababa	2016	398	0.238	0.103	0.798	0.028	0.027	2.9
Addis Ababa	2019	258	0.146	0.112	0.771	0.020	0.019	4.4
Addis Ababa	2011	323	0.122	0.107	0.716	0.018	0.016	11.8

Figure A1. Relationship Between Climate Vulnerability and Child Stunting Across Survey Years



***Notes**: Solid lines represent model predictions. Points represent observed regional averages. Survey Years: 2011 (Red solid line), 2016 (Blue long-dashed line), 2019 (Green dot-dashed line). Regional Numbers: 1. Tigray, 2. Afar, 3. Amhara, 4. Oromia, 5. Somali, 6. Benishangul-Gumuz, 7. SNNP, 8. Gambela, 9. Harari, 10. Addis Ababa, 11. Dire Dawa*

Figure A2. The Relationship between CVI components and Stunting Rate



***Notes**: Four-panel scatterplot showing regional-level associations between each CVI component and stunting rate across survey years (2011, 2016, 2019). Points represent regional means for each survey year. Lines show linear trend.*

Table A22. Model Results without Temporal Interaction (Supplementary)

Variable	Coef. (β)	SE	z value	p-value	OR	95% CI (β)	95% CI (OR)
Intercept	-1.579	0.142	-11.10	<0.001	0.21	[-1.858, -1.300]	[0.16, 0.27]
CVI (std.)	0.226	0.017	13.51	<0.001	1.25	[0.193, 0.259]	[1.21, 1.30]
Survey year	-0.026	0.004	-5.81	<0.001	0.97	[-0.035, -0.017]	[0.97, 0.98]
Breastfeeding	0.147	0.033	4.52	<0.001	1.16	[0.083, 0.211]	[1.09, 1.23]
Household size	-0.031	0.015	-2.10	0.035	0.97	[-0.059, -0.002]	[0.94, 1.00]
Child age (months)	0.025	0.001	29.42	<0.001	1.03	[0.024, 0.027]	[1.02, 1.03]
Child sex (male)	0.225	0.027	8.25	<0.001	1.25	[0.172, 0.279]	[1.19, 1.32]

***Notes**: AIC = 29,084; BIC = 29,157; region variance = 0.083 (SD = 0.287); soil-class variance = 0.063 (SD = 0.252); n = 23,053. This model excludes the CVI × survey-year interaction. The CVI effect (OR = 1.25) is close to the estimate from the primary interaction model (OR = 1.27), indicating that inclusion of the temporal interaction did not materially alter the estimated association between CVI and stunting. ΔAIC = 10 favors the interaction model as the primary specification (Table 12).*

Table A23. Comparison of Model Specifications and Standard Error Estimation Approaches for the CVI and Stunting Coefficient

Method	Model Specification	β (CVI)	SE	z	p-value	OR	95% CI (OR)
Model-based	Region and soil class as random intercepts.	0.239	0.017	14.01	<0.001	1.27	[1.23, 1.31]
Design-based	Region as fixed effect. PSU, strata, and weights modeled explicitly.	0.255	0.028	9.27	<0.001	1.29	[1.22, 1.36]
Cluster-robust	*R*egion and soil class as fixed effects. SEs clustered at enumeration-area level.	0.235	0.029	8.09	<0.001	1.26	[1.20, 1.34]

***Notes**: The model-based row corresponds to the primary specification reported in Table 12 and treats DHS sampling weights as case weights without accounting for clustering or stratification. The design based and cluster robust rows are proxy models that match the fixed effect structure of the primary specification as closely as each estimation method allows. Because the random intercept structure used in the primary model is not directly compatible with design based or cluster robust variance estimation, region and soil class enter these two models as fixed effects rather than random intercepts. Consequently, the coefficient estimates differ modestly across the three approaches, in addition to the differences in their standard errors. All three approaches produce a positive and statistically significant association between CVI and stunting. The 95% confidence interval of 1.20 to 1.34 reported for the CVI main effect in Table 12 is obtained by applying the design based standard error to the CVI coefficient from the primary model. It therefore differs from the 1.22 to 1.36 interval obtained from the design-based model's own coefficient and standard error.*

Table A24. Continuous HAZ Score Model (Robustness Check)

Variable	Coef. (β)	SE	t value	95% CI (β)
Intercept	-0.468	0.116	-4.04	[-0.694, -0.241]
CVI (std.)	-0.186	0.013	-14.20	[-0.211, -0.160]
Survey year	0.012	0.003	3.44	[0.005, 0.018]
Breastfeeding	-0.127	0.025	-5.07	[-0.176, -0.078]
Household size	0.063	0.011	5.64	[0.041, 0.084]
Child age (months)	-0.027	0.001	-41.93	[-0.028, -0.026]
Child sex (male)	-0.112	0.021	-5.39	[-0.153, -0.071]
CVI × Year	-0.0087	0.0038	-2.28	[-0.016, -0.001]

***Notes**: Outcome is HAZ (hw70/100). AIC = 107,483; BIC = 107,571. N = 23,053. Region and soil class were included as random effects.*

Table A25. Decomposed CVI Components Model

Variable	Coef. (β)	SE	z value	p-value	OR	95% CI (OR)
Intercept	-1.492	0.175	-8.50	<0.001	0.23	[0.16, 0.32]
Climate Exposure	-1.038	0.919	-1.13	0.259	0.35	[0.06, 2.15]
Sensitivity	0.623	0.094	6.66	<0.001	1.87	[1.55, 2.24]
Adaptive Capacity	-0.638	0.060	-10.62	<0.001	0.53	[0.47, 0.59]
Survey year	-0.027	0.005	-5.87	<0.001	0.97	[0.96, 0.98]
Breastfeeding	0.137	0.033	4.20	<0.001	1.15	[1.08, 1.22]
Household size	-0.040	0.015	-2.71	0.007	0.96	[0.93, 0.99]
Child age (months)	0.025	0.001	29.16	<0.001	1.03	[1.02, 1.03]
Child sex (male)	0.221	0.027	8.10	<0.001	1.25	[1.18, 1.32]

Notes: *AIC = 28,987; BIC = 29,076; n = 23,053. Region and soil class were included as random effects.*

Table A26. Component-Specific Time Interactions (Robustness Check for CVI × Survey-Year Interaction)

Term	B	SE	z	p	OR	CI
Exposure × Survey Year	0.052	0.277	0.19	0.851	1.05	[0.61, 1.81]
Sensitivity × Survey Year	−0.041	0.028	−1.44	0.151	0.96	[0.91, 1.01]
Adaptive Capacity × Survey Year	−0.069	0.019	−3.68	<0.001	0.93	[0.90, 0.97]

***Notes**: This table extends the decomposed model (Table A25) by interacting each of the three CVI components separately with survey year, rather than testing the composite CVI × survey-year interaction reported in Table 12. AIC = 28,980; BIC = 29,093; n = 23,053. Region variance = 0.0729 (SD = 0.270); soil-class variance = 0.0434 (SD = 0.208). Region and soil class were included as random effects. Models were weighted using round-specific survey weights.*

Table A27. Categorical Exposure × Adaptive-Capacity Tier Interaction (Robustness Check)

Variable	β	SE	z value	p-value	OR	95% CI (OR)
Intercept	−1.917	0.220	−8.72	<0.001	0.15	[0.10, 0.23]
Climate exposure	−2.160	1.455	−1.48	0.138	0.12	[0.01, 2.00]
Capacity tier: Low (ref: Higher)	0.165	0.218	0.76	0.448	1.18	[0.77, 1.81]
Capacity tier: None (ref: Higher)	0.261	0.233	1.12	0.262	1.30	[0.82, 2.05]
Sensitivity	0.730	0.091	8.01	<0.001	2.07	[1.74, 2.48]
Survey year	−0.025	0.005	−5.45	<0.001	0.98	[0.97, 0.98]
Breastfeeding	0.139	0.033	4.27	<0.001	1.15	[1.08, 1.23]
Household size	−0.045	0.015	−3.04	0.002	0.96	[0.93, 0.98]
Child age (months)	0.025	0.001	28.98	<0.001	1.03	[1.02, 1.03]
Child sex (male)	0.220	0.027	8.06	<0.001	1.25	[1.18, 1.31]
Exposure × Tier: Low	1.902	1.796	1.06	0.289	6.70	[0.20, 226.32]
Exposure × Tier: None	1.730	1.913	0.90	0.366	5.64	[0.13, 239.77]

***Notes**: AIC = 28,997; BIC = 29,109; region variance = 0.064 (SD = 0.254); soil-class variance = 0.037 (SD = 0.193); n = 23,053. Models were estimated using glmer with a binomial distribution, weighted by round-specific survey weights, with random intercepts for region and soil class. Capacity tier was defined as None (adaptive capacity index = 0), Low (>0 to ≤0.5), and Higher (>0.5; reference category). 95% CIs are Wald-based.*

Table A28. Alternative Vulnerability Index Formulations and Comparison

Index Formulation	Mathematical Form	β (CVI)	SE	p-value	OR	AIC
Multiplicative	(Exposure × Sensitivity) / (Adaptive Capacity + 0.1)	0.208	0.018	<0.001	1.23	27,655
Additive	Exposure + Sensitivity − Adaptive Capacity	0.290	0.017	<0.001	1.34	27,513
Weighted	0.4×Exposure + 0.4×Sensitivity − 0.2×Adaptive Capacity	0.289	0.018	<0.001	1.33	27,528

***Notes**: All three formulations were fitted on n = 23,053 observations, weighted by wt_pooled rather than wt_survey. AIC and standard errors are comparable within this table but should not be compared directly with those reported in Table 12. Refitting the winsorized multiplicative model using wt_survey gave β = 0.210, essentially unchanged from 0.208, indicating that the weighting choice did not materially affect the estimated association for the multiplicative formulation.*

Table A29. Correlation Between Vulnerability Indices and Their Components

Formulation	Correlation with Exposure	Correlation with Sensitivity	Correlation with Adaptive Capacity
Multiplicative	0.317	0.672	-0.727

Additive	0.024	0.855	-0.930
Weighted	0.057	0.926	-0.854

***Notes**: The multiplicative index showed a more balanced relationship with the three component dimensions than the additive and weighted alternatives. The latter showed near-zero correlations with exposure (0.024 and 0.057) but strong correlations with sensitivity (0.855 and 0.926), indicating that they were more strongly dominated by the sensitivity component. Thus, the multiplicative formulation better preserved variation across the three conceptual dimensions of vulnerability than the alternative formulations.*

Table A30. Local Moran's I (LISA) Results by Region

Region	Mean CVI	LISA Cluster	Local Moran's I	p-value
Afar	0.884	Not Significant	0.280	0.082
Somali	0.440	Not Significant	0.466	0.099
Amhara	0.370	Not Significant	0.121	0.206
Dire Dawa	0.300	Not Significant	-0.020	0.798
SNNP	0.294	Not Significant	0.033	0.706
Oromia	0.281	Not Significant	-0.009	0.831
Tigray	0.280	Low-High	-0.302	0.036
Benishangul-Gumuz	0.268	Not Significant	-0.005	0.992
Gambela	0.259	Not Significant	0.050	0.779
Harari	0.139	Not Significant	0.174	0.777
Addis Ababa	0.023	Not Significant	0.285	0.714

***Notse**: LISA classifications are based on local Moran's I at the unadjusted $\alpha = 0.05$ significance level. Global and local spatial statistics were calculated using regional CVI and the specified spatial adjacency weights.*

Table A31. Predictors of Missing Anthropometric Data (Selection Bias Analysis)

Predictor	B	SE	OR	95% CI (OR)	p-value
Intercept	-0.783	0.142	0.46	[0.35, 0.60]	<0.001
CVI	0.201	0.066	1.22	[1.07, 1.39]	0.002
Wealth Quintile	-0.045	0.018	0.96	[0.92, 0.99]	0.012
Education Level	-0.027	0.027	0.97	[0.92, 1.03]	0.306
Rural Residence	-0.389	0.054	0.68	[0.61, 0.75]	<0.001

***Notes**: The model was adjusted for survey year (factor). CVI remained significantly associated with missing anthropometric data after adjustment for wealth, education, rural residence, and survey year (OR = 1.22, 95% CI: 1.07–1.39, $p = 0.002$), indicating that children with higher climate vulnerability had modestly higher odds of missing anthropometric data, conditional on these covariates. The raw, unadjusted difference in mean CVI between excluded (0.369) and included (0.347) children (Table A32) was significant (Welch's t-test, $t = 4.0$, $df = 6,276$, $p < 0.001$), and the association persisted after adjustment, suggesting that the observed selection pattern was not fully explained by wealth or rural residence. Wealth ($p = 0.012$) and rural residence ($p < 0.001$) remained independently associated with missingness. Excluded children were, on average, poorer (mean wealth 2.63 vs. 2.68, $p = 0.026$) and more likely to reside in rural areas ($\chi^2 = 15.3$, $p < 0.001$). 23,542 of 28,048 children (83.9%) were retained in the analytic sample. Missingness was highest in Somali (22.8%) and lowest in Tigray (9.1%).*

Table A32. Comparison of Included and Excluded Observations

Group	N	Mean CVI	% Rural	Mean Wealth	% No Education	% Female-Headed
Included	23,542	0.347	81.5	2.68	64.2	19.3
Excluded	4,506	0.369	79.1	2.63	66.7	23.5

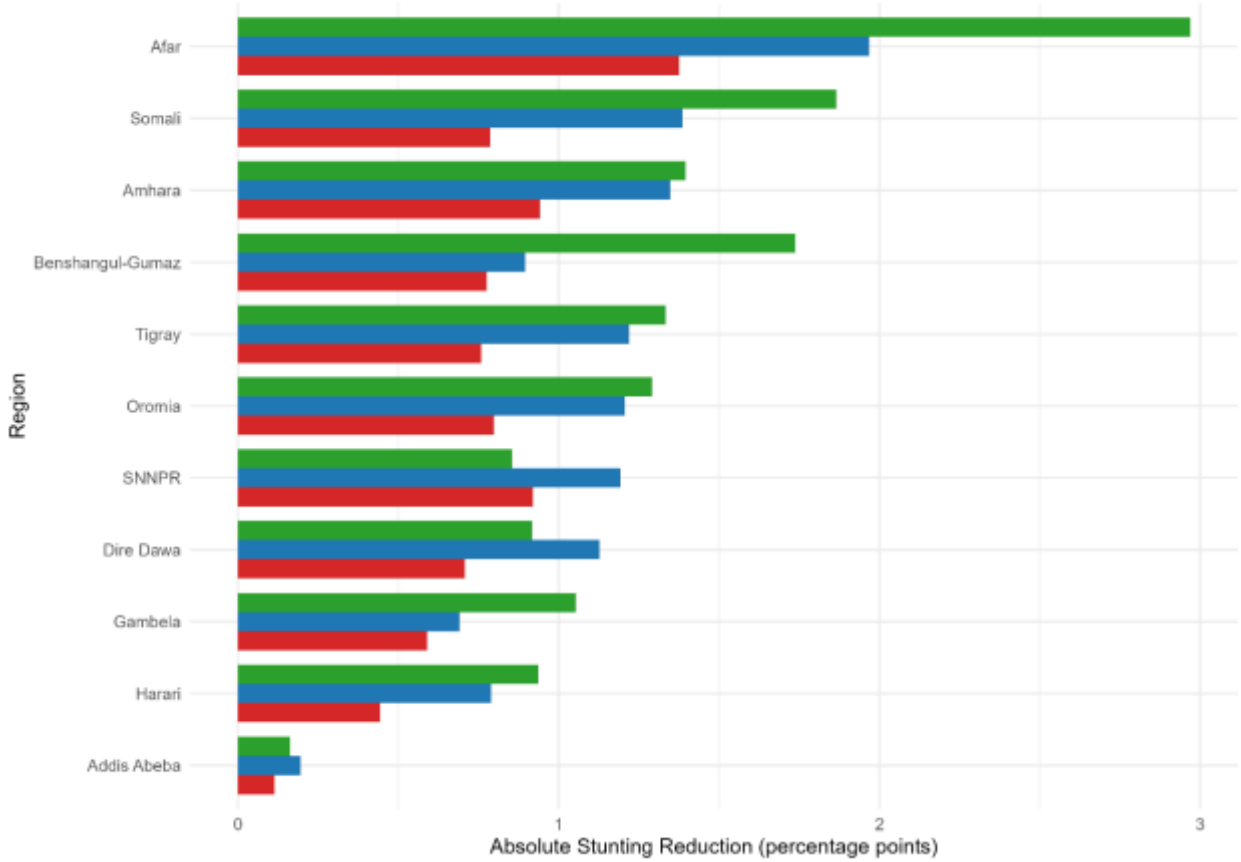
***Note**: Welch's t-test for the difference in mean CVI between included and excluded observations: $t = 4.0$, $df = 6,276$, $p < 0.001$.*

Table A33. Counterfactual Simulation Results

Region	Year	N	Baseline Stunting (%)	Counterfactual Stunting (%)	Absolute Reduction (pp)	Relative Reduction (%)
Addis Ababa	2011	323	21.3	21.2	0.11	0.54
Addis Ababa	2016	398	17.3	17.1	0.20	1.08
Addis Ababa	2019	224	14.5	14.3	0.16	1.09
Afar	2011	893	45.3	43.9	1.38	3.04
Afar	2016	829	42.7	40.7	1.97	4.54
Afar	2019	536	46.4	43.4	2.97	6.12
Amhara	2011	1,068	51.6	50.6	0.94	1.83
Amhara	2016	887	47.3	45.9	1.35	2.81
Amhara	2019	396	41.1	39.7	1.40	3.24
Benishangul-Gumuz	2011	834	46.1	45.3	0.78	1.70
Benishangul-Gumuz	2016	731	41.1	40.2	0.90	2.19
Benishangul-Gumuz	2019	409	41.8	40.1	1.74	3.95
Dire Dawa	2011	584	39.0	38.3	0.71	1.77
Dire Dawa	2016	418	35.6	34.5	1.13	2.83
Dire Dawa	2019	320	29.7	28.8	0.92	2.62
Gambela	2011	683	30.5	29.9	0.59	1.92
Gambela	2016	560	27.6	26.9	0.69	2.47
Gambela	2019	346	26.1	25.0	1.05	3.75
Harari	2011	511	35.7	35.2	0.44	1.22
Harari	2016	461	31.3	30.5	0.79	2.32
Harari	2019	348	28.8	27.9	0.94	2.96
Oromia	2011	1,515	41.0	40.2	0.80	1.96
Oromia	2016	1,378	37.2	36.0	1.21	3.22
Oromia	2019	588	34.2	32.9	1.29	3.62
SNNP	2011	1,368	44.3	43.3	0.92	2.08
SNNP	2016	1,103	39.7	38.5	1.19	2.97
SNNP	2019	534	33.0	32.1	0.85	2.54
Somali	2011	755	29.9	29.1	0.79	2.51
Somali	2016	1,154	29.0	27.6	1.39	4.71
Somali	2019	519	31.2	29.3	1.87	5.87
Tigray	2011	1,077	49.4	48.7	0.76	1.53
Tigray	2016	936	44.4	43.2	1.22	2.66
Tigray	2019	367	41.0	39.7	1.33	3.00

***Note**: National summary: baseline stunting prevalence = 40.5%; counterfactual prevalence = 39.4%; absolute reduction = 1.07 percentage points; relative reduction = 2.81%.*

Figure A3. Potential reduction in stunting prevalence from a 25% reduction in climate vulnerability, by region-survey year



***Note**: Survey year: 2011(Red bars), 2016 (Blue bars) and 2019 (Green bars).*

CHAPTER 3: AGRICULTURAL EXTENSION AS CLIMATE INSURANCE: PANEL EVIDENCE FROM ETHIOPIAN SMALLHOLDERS

Abstract

Agricultural extension services are increasingly recognized and promoted as a means of helping farmers increase productivity and adapt to climate change in sub-Saharan Africa. Yet evidence on whether these services raise average yields and input use, reduce climate-related production risk, or achieve these outcomes simultaneously remains limited. Using a balanced panel of 3,222 rural Ethiopian households observed in 2011/12, 2013/14, and 2015/16, we estimate household fixed-effects models of log-transformed harvest as a function of lagged extension participation. The results show that within households, harvest is 15.0% lower in waves following extension participation than in waves without it ($\beta = -0.163$, SE = 0.058, $p = 0.005$). To assess the sensitivity of this negative association to selection on observables, we use propensity score matching (PSM) to match households on covariates associated with extension participation. A probit selection model indicates that measured drought exposure does not significantly predict placement. The association remains similar in magnitude after matching (−17.4%, $p = 0.004$). In contrast, in a fixed-effects model that includes an interaction between lagged extension participation and drought exposure, extension participation significantly moderates drought-related harvest losses (interaction $\beta = 0.485$, SE = 0.144, $p < 0.001$), with a similar pattern after PSM. Because 2013/14 was anomalously wet, the drought-buffering estimate is identified primarily from the 2015/16 wave, and a cross-sectional specification using only 2015/16 data does not reproduce the interaction, indicating the result depends on the panel fixed-effects structure. These findings indicate that prior extension participation is associated with smaller harvest losses during droughts and lower harvest volatility across survey waves more generally, but not with higher average harvests. Thus, the contribution of extension appears to lie more in mitigating drought-related losses and harvest variability than in increasing average harvests. Evaluations of agricultural extension in Ethiopia and other drought-prone settings should therefore distinguish between

productivity effects and effects on harvest stability, particularly during adverse climatic conditions. For policy, this implies that the value of extension may be underestimated when performance is assessed solely by average yields.
Keywords: *Agricultural extension, climate adaptation, risk management, Ethiopia, panel data, fixed effects, propensity score matching*

1. Introduction

Economic development and the improvement of rural livelihoods remain critical priorities across the Global South, particularly in Sub-Saharan Africa (SSA). Yet progress towards these goals is closely tied to the performance and resilience of the agricultural sector. This interdependence is particularly important because, in many SSA economies, agriculture continues to serve as a fundamental pillar of national development, contributing significantly to gross domestic product, employment generation, food security, and export earnings. This dependence is reflected in the region's agricultural structure, with 55–62% of the workforce engaged in agriculture and 90–95% of cropland dependent on rainfed production. Such high levels of exposure and dependence are further compounded by limited adaptive capacity, driven by inadequate climate finance, weak institutions, and insufficient infrastructure (Trisos et al., 2022).

Ethiopia exemplifies this vulnerability, where roughly 80% of the population relies on rain-fed agriculture (World Bank Group, 2021), and recurrent droughts trigger food crises affecting millions. The 2011 Horn of Africa drought affected 13 million people across the East Africa region and required emergency assistance for 4.5 million Ethiopians from June to December 2011, a figure that declined to 3.2 million by early 2012 (OCHA, 2012; ReliefWeb, 2012). The 2015-2016 El Niño drought impacted over 10 million Ethiopians with livestock mortality exceeding 40% in pastoral areas (Government of Ethiopia, 2016), while the subsequent 2017 drought required emergency assistance for 8.5 million people (OCHA, 2017).

These severe and recurrent climate shocks, combined with Ethiopia's long-standing reliance on predominantly traditional farming systems, have contributed to persistent agricultural vulnerability (Sinore & Wang, 2024). Ethiopia has experienced a long history of drought-induced food crises. In the 1984 drought alone, approximately 8.7 million people in Ethiopia

were affected and about 1 million people died (FAO, 2000). In addition, the severe food crisis following the 2002 drought was only narrowly prevented from escalating into famine through large-scale emergency assistance (Sabates-Wheeler et al., 2022). These recurring challenges have underscored the need to strengthen institutional mechanisms that enhance resilience to climate variability and improve agricultural productivity.

In this context, Agricultural extension services are recognized and promoted as an important mechanism for helping farmers in sub-Saharan Africa increase agricultural productivity and adapt to climate change (Azzarri & Nico, 2022; Kalimba & Culas, 2020). Agricultural extension services encompass a range of organizations that support people working in agriculture by facilitating access to the knowledge, skills, and technologies needed to address agricultural challenges and improve their livelihoods (Davis et al., 2020).

Against this backdrop, successive Ethiopian governments have pursued reforms to expand and strengthen the agricultural extension system, aiming to improve farmers' productivity and resilience while supporting climate adaptation. Ethiopia's formal agricultural extension system emerged during the imperial period under *Emperor Haile Selassie I*, when the Imperial Ethiopian College of Agriculture and Mechanical Arts was established in 1953. The system subsequently underwent several institutional and programmatic changes before a major transformation in the early 1990s (Berhane et al., 2018). During this period, the Sasakawa Africa Association and Global 2000 (SG-2000) introduced a pilot program promoting improved agricultural practices through demonstrations, input packages, and close technical supervision (Sasakawa Africa Association, n.d.).

Following the reported success of the pilot, Ethiopia's transitional government adopted the approach nationally in 1995 through the Participatory Demonstration and Training Extension System (PADETES). PADETES provided the foundation for the subsequent expansion of Ethiopia's agricultural extension system. The system continued to evolve through subsequent reforms and regional adaptations, including the development of a new *Agricultural Extension Strategy* by the Government of Ethiopia in collaboration with the *Ethiopian Agricultural Transformation Agency* (Berhane et al., 2018).

Today, Ethiopia has one of Africa's largest agricultural extension systems, with tens of thousands of Development Agents (DAs) deployed across kebeles[16] to provide agronomic advice, facilitate access to agricultural inputs, and promote the adoption of improved farming practices (Berhane et al., 2018; Davis et al., 2010). This strategy positions extension as a central policy instrument for increasing productivity, supporting risk management, and building climate resilience (MoANR, 2017). The government's Climate-Resilient Green Economy (CRGE) Strategy reinforces this, positioning extension as central to emissions mitigation through improved crop and livestock management (FDRE, 2011a).

Despite the extensive reach of this extension network and its prominent role in national agricultural and climate policies, the core mechanism through which extension operates remains unclear. Moreover, official policy emphasizes production, as reflected in the National Agricultural Extension Strategy's focus on promoting improved technologies without adequate attention to value addition and marketing (MoANR, 2017, p. 22). Whether this production-oriented approach translates into sustained yield gains or whether extension instead operates through climate risk management channels remains an open empirical question. Addressing this issue requires first identifying who receives extension services and why, as selection into such programs fundamentally shapes what impact evaluations can reveal.

To date, rigorous evidence on whether extension services simultaneously improve average agricultural productivity and reduce farmers' exposure to climate-related production risks remains limited. Existing literature on agricultural extension in Ethiopia shows heterogeneous effects. Panel data from Ethiopia (1999–2009) suggest that extension agents initially increased the adoption of fertilizer and improved seeds, although these impacts weakened over time (Krishnan & Patnam, 2014). Similarly, evidence from another study shows that while extension services raised fertilizer use by 22 percentage points, they did not translate into measurable productivity gains (Berhane et al., 2018). Evidence from a cluster randomized controlled trial (RCT) indicates that extension support increased adoption of improved wheat varieties, but primarily among wealthier households (Yitayew et al., 2021).

[16] Ethiopia has a multilevel administrative structure organized hierarchically from the federal government to regional states and chartered cities, then to zones, woredas (districts), and finally ***kebeles*** (the lowest-level administrative units).

An instrumental variables study finds that extension services promoted chemical fertilizer adoption while reducing the use of organic fertilizer, with no detectable effects on pesticides, herbicides, or fungicides (Mesfin et al., 2022). Recent large-scale evidence on Ethiopia's fertilizer program revealed null effects on yields and household consumption despite successful input substitution (Assefa et al., 2025), suggesting that input adoption alone does not guarantee productivity gains. This disconnect may stem from allocative inefficiencies, such as the application of uniform fertilizer rates to non-responsive landscape positions (Desta et al., 2023), raising questions about whether extension programs deliver productivity gains or operate through alternative mechanisms such as risk management.

The theoretical relationship between extension services and agricultural outcomes remains ambiguous when viewed through the lens of production economics. The productivity pathway posits that extension drives modernization through technology adoption and input intensification to maximize yields (Anderson & Feder, 2004). However, this approach carries potential risks because high-input systems often exhibit risk-increasing properties (Just & Pope, 1979), making them more susceptible to steeper marginal losses during climate shocks when water scarcity renders fertilizer ineffective.

Conversely, the climate-insurance pathway suggests that extension may provide substantial utility by reducing downside exposure to shocks, even in the absence of consistent gains in average productivity, if it facilitates ex ante preparation through measures such as adoption of drought-tolerant varieties, investment in irrigation, and improved water management (Alam et al., 2024). This framing is consistent with evidence that input supply interventions can increase adoption without improving welfare outcomes (Assefa et al., 2025). Whether a negative average association with productivity reflects targeting, differences in risk-averse behavior among extension-using households, or other factors thus remains an empirical question.

Empirical evidence on whether agricultural extension primarily affects average productivity, production risk, or both in Ethiopia remains limited. Most studies in this literature examine input adoption or productivity rather than risk management, and few test whether extension benefits persist during climate shocks, a critical consideration for assessing their climate adaptation value. Understanding extension's role during climate stress is particularly important given Ethiopia's exposure to recurrent droughts and the government's explicit

policy goal of leveraging extension services to build climate resilience. Thus, this study addresses four inter-related research questions:

1. To what extent does drought exposure moderate the association between prior extension participation and agricultural outcomes and input use (harvest, fertilizer use, and irrigation), and is this pattern consistent with an insurance-like buffering function?
2. Does prior extension participation correlate with observable household and geographic characteristics at baseline, and do the estimated associations with total harvest and the drought-buffering interaction persist after accounting for observed baseline differences using propensity score matching?
3. Is prior extension participation associated with lower harvest volatility across survey waves overall, rather than only during specific drought episodes?
4. Do crop and livestock diversification help explain how prior extension participation buffers harvest losses during drought?

By addressing these questions, the study provides evidence on how agricultural extension is associated with both average productivity and drought-related production risk, treating these as distinct margins of impact rather than assuming that they move together. These findings are particularly relevant given Ethiopia's strategic goal of developing a climate-responsive, knowledge-driven extension system that enhances farmers' resilience and promotes environmental sustainability (MoANR, 2017). The study also contributes to several Sustainable Development Goals, particularly SDG 1 (No Poverty), SDG 2 (Zero Hunger), and SDG 13 (Climate Action) (United Nations General Assembly, 2015).

2. Materials and Methods

Data Sources and Sample Selection

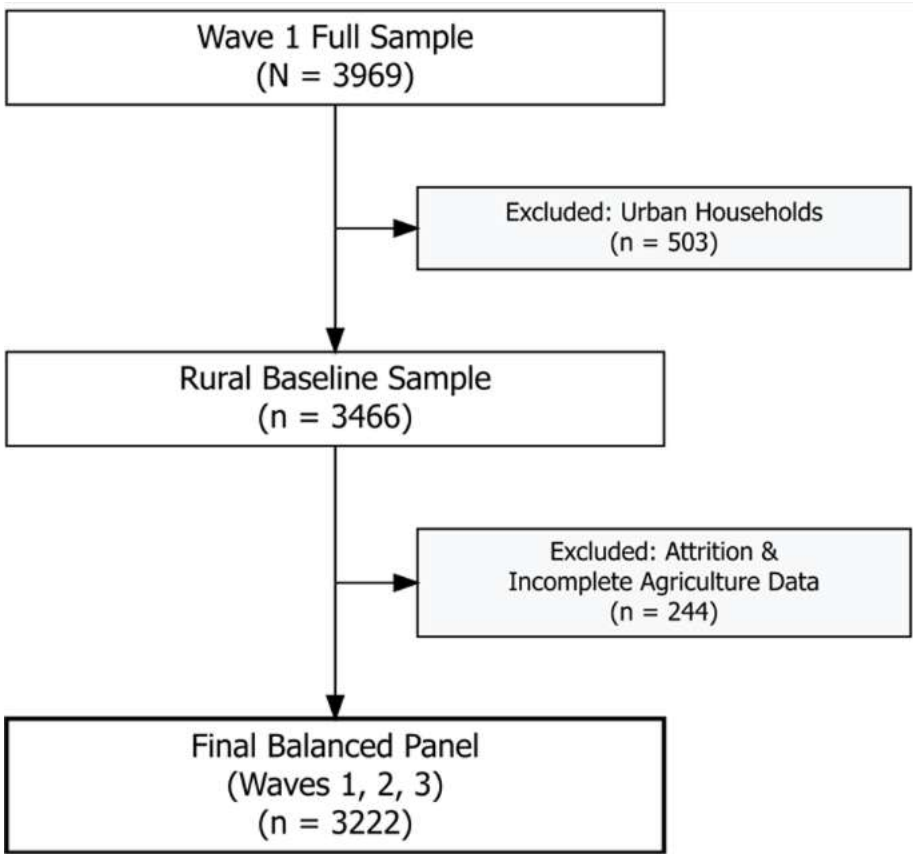
This study uses household-level longitudinal data from the Ethiopia Socioeconomic Survey (ESS), conducted by the Ethiopian Central Statistical Agency in partnership with the World Bank's Living Standards Measurement Study–Integrated Surveys on Agriculture (LSMS-ISA) program. The ESS is a nationally representative survey of rural and small-town households in Waves 1–3 (2011/12–2015/16), with the new Panel II introduced in Wave 4 (2018/19) and followed in Wave 5 (2021/22) (CSA, 2020b). The rural component follows

the two-panel design of the national ESS. There was no overlap in households between the two panels (*see Appendix Table A34, Panel A*).

This study focuses exclusively on Panel I because the two panels differ substantially in their sampling frames and panel structures, in addition to having no overlap in households, making direct pooling inappropriate for the present longitudinal analysis. In particular, Panel II differs from Panel I in both rural–urban composition and attrition patterns. Panel II had a substantially higher overall attrition rate than Panel I (35.6% versus 7.0%), driven largely by the exclusion of conflict-affected areas in Wave 5 (*see Appendix Table A34, Panel B*). In addition, the share of rural households declined from 87 percent in Wave 1 to 46 percent in Wave 4, and urban households in Wave 4 reported no agricultural participation (*Appendix Table A35*).

Having restricted the analysis to Panel I, we next construct the longitudinal estimation sample. Starting from 3,969 Wave 1 households, we restricted the sample to 3,466 rural households and retained households observed in all three waves, resulting in a balanced panel of 3,222 rural households and 9,666 household-wave observations (*Figure 7*). This represents a modest attrition of 244 rural households, corresponding to an attrition rate of 7.0 percent. The resulting balanced panel was used for the primary longitudinal analyses.

Figure 7. Sample Selection Process



Variable Construction

Extension participation was measured through two binary questions in the post-planting agricultural module. The first question (pp_s7q04, "***Do you participate in the extension program?***") asked whether the household participated in an extension program, capturing formal enrollment with DAs who deliver agronomic training, input facilitation, and technology demonstrations at the kebele level. The second question (pp_s7q08, "***Do you get advisory services?***") asked whether the household received agricultural advisory services, capturing services such as field visits and crop management consultations. A household was classified as an extension participant if it answered affirmatively to either question.

Following ESS survey protocol, the post-planting module was administered only to households with at least one member engaged in crop farming on owned or rented land, and the livestock module only to households engaged in livestock rearing (CSA, 2020a). Households absent from either module are coded as missing on the relevant variables rather

than as non-participants, and are excluded from the corresponding regression analyses. *Appendix Table A36* presents the cross-tabulation of these two questions across waves.

The composite two-question measure serves as the primary treatment definition. Two sensitivity analyses assess whether the findings are robust to alternative measurement choices. The first uses an alternative single-question definition based solely on pp_s7q04, while the second examines whether the results are sensitive to the choice of outcome transformation (*Appendix Table A46*).

We analyze four ***outcome variables*** in the main regression models: fertilizer adoption, total harvest, irrigation adoption, and production risk. ***Fertilizer adoption*** is a binary indicator of inorganic fertilizer use, derived from household reports in the extension module. A plot-level fertilizer measure from a separate survey module is used as a robustness check (see Sensitivity Analyses below). ***Total harvest*** aggregates crop production in kilograms across all plots and is transformed as *log(harvest + 1)* to address right-skew while retaining zero-harvest households. ***Irrigation adoption*** is a binary indicator of whether any household plot received irrigation during the main agricultural season. ***Production risk*** is measured as the coefficient of variation (CV) of household-level harvest across the three survey waves. The CV is calculated for households with at least two non-missing harvest observations.

Unlike the other three outcomes, production risk is a single household-level measure collapsed across waves and is therefore estimated in a cross-sectional rather than panel specification. For the production-risk analysis, the coefficient of variation (CV) is calculated for households with at least two non-missing harvest observations across the three survey waves, yielding 2,695 households eligible for CV construction. After further excluding households with missing regression covariates, 2,691 households remain in the baseline estimation sample. As a sensitivity analysis, we restrict the sample to the 1,442 households with non-missing harvest observations in all three waves; after excluding households with missing regression covariates, 1,440 households remain in the estimation sample.

Among the four outcome variables, total harvest is the only one carried through the full set of robustness and sensitivity analyses. Fertilizer and irrigation adoption are estimated in the main-effects and drought-interaction models but are not included in the full set of outcome-specific robustness analyses.

In addition to these outcomes, two variables are constructed as candidate mechanisms. ***Crop diversification*** is measured as the number of distinct crop codes harvested by each household. ***Livestock diversification*** is measured as the number of distinct livestock species owned, collapsed to the household level to avoid double-counting species reported by multiple agricultural holders. Households absent from the livestock module are assigned zero species owned. This coding treats households not administered the livestock module as having no livestock and therefore avoids treating structural absence from the module as conventional item nonresponse. The resulting measure should therefore be interpreted as a conservative measure of livestock diversity under the survey's module-selection rules.

Having defined the outcome variables, we next describe the time-varying rainfall shock measures and household-level controls used throughout the analysis.

Rainfall shocks and drought are constructed using household-specific long-run baselines based on historical rainfall conditions from a fixed pre-analysis reference period. This approach measures rainfall deviations relative to local climate norms rather than differences in average rainfall across locations. For each household's GPS location, we extract 30 years of monthly precipitation (1981–2010) from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS v2.0, 0.05° resolution) (Funk et al., 2015). For the small number of households whose GPS coordinates fall within more than one CHIRPS grid cell (56 of 3,222 households, spanning up to three cells), we average the relevant cell-level statistics rather than retaining only the first matched cell. This approach avoids assigning households to an arbitrary grid cell, particularly because cell-level long-run means differed by as much as 447 mm.

The standardized rainfall shock for household i in wave t is defined as:

$$Z_{it} = \left({(R_{it} - \mu_i)}/{\sigma_i} \right).$$

Here, R_{it} is total rainfall during household i 's growing season in wave t; μ_i is the household's long-run annual rainfall mean extracted from the ESS geo-covariate file separately for each wave; and σ_i is the household's long-run annual rainfall standard deviation calculated from the CHIRPS 1981–2010 series. Because the baseline statistics used here are defined annually, whereas R_{it} is a growing-season total, the numerator and baseline/standardization terms are defined over different temporal windows. We retain this construction because the ESS geo-covariate provides a household-specific long-run rainfall baseline based on a fixed reference

period and is used consistently throughout the analysis, including in the alternative drought definition described below. In the full analysis sample, the ESS geo-covariate baseline and corresponding CHIRPS annual mean are strongly correlated across waves (r = 0.87–0.88; n = 3,194–3,222), providing supporting evidence that the two products capture similar underlying rainfall conditions.

This baseline value nonetheless varies modestly across survey waves for the same household despite unchanged GPS coordinates, indicating differences in geo-covariate extraction across survey rounds rather than genuine changes in the underlying 1981–2010 reference-period climate. Among the 3,222 households with non-missing baseline values in all three waves, the median within-household range across waves is 64.0 mm, the mean range is 68.5 mm, and 269 households (8.3%) have a range exceeding 100 mm.

Accordingly, Z_{it} should be interpreted as growing-season rainfall relative to the household's long-run annual climate norm, rather than as a conventional growing-season standardized anomaly. We define ***drought*** as $\boldsymbol{Z_{it} < -1}$, corresponding to a growing-season rainfall total that is more than one household-specific standard deviation below the long-run annual rainfall mean. This definition ensures that drought classification in 2011 relies strictly on pre-2011 climate history. As a robustness check, we also apply an absolute threshold, classifying a season as a drought when growing-season rainfall falls more than 25% below the long-run annual rainfall average.

Time-varying controls include log-transformed household size (log(household size + 1)), years of education of the household head, and a binary indicator of credit access. Years of education is derived from the household head's reported highest completed grade or credential.

Empirical Estimation Strategy

The analysis proceeds in four stages. First, descriptive analyses characterize extension participation, agricultural outcomes, and rainfall exposure across waves. Descriptive statistics are calculated at the household level using the analysis sample relevant to each variable. For longitudinal variables, wave-specific statistics are reported separately, while baseline comparisons use Wave 1 characteristics. Prior to estimation, multicollinearity was assessed using pairwise Pearson correlations and variance inflation factors (VIFs). No

pairwise correlation among the regression covariates exceeded |r| = 0.35, and VIFs ranged from 1.01 to 1.02, indicating no evidence of problematic multicollinearity.

Second, the primary inferential analysis estimates the association between lagged extension participation and agricultural outcomes using household and wave fixed effects. The use of a lagged specification is motivated by the timing of the survey measures. Fertilizer adoption and extension participation are recorded in the same post-planting module, making a contemporaneous specification potentially problematic because the two measures may reflect the same period-specific conditions. We therefore use a one-period lag specification in which extension participation in wave t − 1 predicts agricultural outcomes in wave t. This lagged specification establishes temporal ordering and reduces concerns that contemporaneous agricultural conditions jointly determine extension participation and outcomes. For example, adverse agricultural conditions in wave t could induce a household to seek extension services during the same period, generating a spurious contemporaneous association. It also allows time for advisory recommendations to influence subsequent production decisions.

Third, the drought-buffering specification examines whether prior extension participation moderates the relationship between contemporaneous drought exposure and outcomes. Fourth, diagnostic and sensitivity analyses assess treatment timing, attrition, treatment-history dependence, selection on observables, outcome measurement, candidate mechanism channels, the stability of the rainfall baseline measure, and the stability of community-level DA exposure.

The baseline model is specified as follows:

$$Y_{it} = \beta Extension_{i,t-1} + \theta X_{it} + \phi_i + \eta_t + \varepsilon_{it}$$

where Y_{it} is the outcome for household i in wave t, $Extension_{i,t-1}$ indicates participation in the extension program in the previous survey wave, X_{it} is a vector of time-varying control variables, ϕ_i denotes household fixed effects and η_t represents wave fixed effects. The coefficient β captures the conditional association between prior extension participation and the current outcome after accounting for time-invariant household characteristics, observed time-varying covariates, and common wave-specific shocks. For all log-transformed harvest outcomes, coefficients are converted to percentage effects using $100 \times (exp^{\beta} - 1)$.

As a timing sensitivity analysis, the primary model is re-estimated using contemporaneous extension participation in place of the one-period lagged measure (*Appendix Table A43*).

To examine whether extension participation buffers the effects of climate shocks, we augment the baseline model with an ***interaction between*** lagged extension participation and contemporaneous drought:

$$Y_{it} = \beta_1 Extension_{i,t-1} + \beta_2 Drought_{it} + \beta_3 \left(Extension_{i,t-1} \times Drought_{it}\right) + \theta X_{it} + \phi_i + \eta_t + \varepsilon_{it}$$

The coefficient β_2 captures the average association of drought among households that did not participate in extension in the previous wave. The interaction coefficient β_3 captures the additional association of prior extension participation during drought years.

Because lagged extension status is undefined in Wave 1 (no t−1 observation is available) and no Wave 2 observations meet the drought criterion, the drought-related interaction in the main household fixed effect model is identified entirely by drought variation in Wave 3. As a complementary check, we therefore estimate a Wave 3-only cross-sectional specification with region fixed effects.

The Wave 3 specification also addresses potential contamination of the comparison group arising when households coded as untreated under the one-period lag retain benefits from earlier extension participation. We test this by adding a two-period lag of extension status as an additional control. Because the two-period lag of extension participation is non-missing only in Wave 3, it provides only one observation per household and therefore cannot be identified in a household fixed-effects model. The test is consequently conducted in the same Wave 3 cross-sectional specification.

To address the possibility that households with oscillating extension histories may retain benefits from earlier participation and thereby contaminate the comparison group defined by the one-period lag, we classify households according to their full three-wave extension histories as ***always-treated, never-treated, clean adoption, clean exit, or non-monotonic (oscillating)***. The specific treatment-history patterns are reported in *Table 15, Panel B*. We then re-estimate the drought-buffering interaction after excluding non-monotonic households as a sensitivity analysis.

To provide diagnostic evidence relevant to the identifying assumptions of the fixed-effects specification, we conduct an event-study comparison of households with a clean single-period adoption pattern against never-treated households, using relative time to first adoption (*see Figure 9*). Given the three-wave panel structure, this test is necessarily limited to two

pre-treatment periods and one post-treatment period. Pre-treatment coefficients are therefore treated as diagnostic evidence concerning pre-trends rather than as a definitive test of the identifying assumption.

As a complementary sensitivity analysis, we use ***PSM*** to assess the robustness of the primary results to selection on observable baseline characteristics. Extension participants are matched to comparable non-participants based on baseline drought exposure, rainfall z-scores, household size, education, credit access, distance to market, and the long-run rainfall baseline. We implement 1:1 nearest-neighbor matching with a 0.25 caliper using Wave 1 extension status and covariates. Balance diagnostics and the number of observations retained after matching are reported in the corresponding appendix results.

Two distinct data-completeness issues are addressed separately, namely ***missingness in the harvest outcome*** among households retained in the panel, and ***attrition of households*** from the panel itself.

Total harvest is missing for 2,622 of 9,666 household-wave observations in the balanced panel, even though these households remain in the sample. We assess the implications in three ways. First, we re-estimate the primary model treating missing harvest as zero. This analysis represents an alternative assumption about the meaning of missing harvest rather than a conventional robustness check. We separately re-estimate it after winsorizing extreme harvest values at the 1st and 99th percentiles. Second, we examine whether harvest missingness is related to drought and extension status. Households with no harvest data in any wave are distinguished from those missing data in only one or two waves, to assess whether missingness primarily reflects non-cultivation rather than systematic missingness related to the variables of interest. Third, we focus on the non-participant, drought-exposed cell ($Extension_{i,t-1} = 0$, $Drought_{it} = 1$), which has unusually high missingness that is not fully explained by non-cultivation. We therefore bound the interaction estimate by assigning the unexplained missing observations in this cell to the 1st and 99th percentiles of observed harvest within the same cell. This is a bounding exercise rather than a conventional robustness check. It shows how far the estimate could move under extreme assumptions about the missing observations, not what the estimate would be under a preferred alternative specification.

Separately, we assess whether the 244 households that dropped out of the panel (*Figure 7*) differ systematically from those retained. This is important because differential attrition could make the retained panel systematically different from the baseline sample and thereby affect the composition and external validity of the fixed-effects estimation sample. We first report unadjusted attrition rates by baseline extension status, household size, and region. We then test this using a logistic regression of panel dropout on baseline household size, drought exposure, extension participation, and region fixed effects. Several region indicators exhibit quasi-complete separation in this model, so we also re-estimate it using a bias-reduced (Firth-type) logistic regression and report both sets of estimates for comparison. Full results are presented in *Appendix Table A37*.

We conduct several additional analyses to assess the robustness and interpretation of the main drought-buffering result, including mechanism tests, climate-response and instrument diagnostics, alternative outcome and drought measures, sample restrictions, and heterogeneity and production-risk analyses.

All analyses were performed using R software version 4.6.1 (R Core Team, 2026). The *fixest* package was used for fixed-effects estimation, car for calculating variance inflation factors (VIFs) to assess multicollinearity, *MatchIt* for propensity score matching, *margins* for estimating average marginal effects, *brglm2* to assess the robustness of the attrition model to quasi-complete separation, *ggplot2* for figures, and *DiagrammeR* with *rsvg* for the sample-flow diagram (*Figure 7*). Unless otherwise noted, all panel models include household and wave fixed effects, with standard errors clustered at the household level. Statistical significance is denoted as † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

3. Results

Descriptive Statistics

The descriptive analysis is based on the balanced panel constructed as described in the previous section (*Figure 7*). Extension participation averaged 62.9 percent across waves and remained relatively stable over time (*Table 15, Panel A*). Among the 3,138 households with non-missing lagged extension status, 662 (21.1 percent) switched extension status between waves (*Table 15, Panel B*). Here, two distinct concepts of switching are relevant. ***FE-identifying switchers*** are households whose lagged extension status changes between Waves

2 and 3, thereby contributing within-household variation to the lagged fixed-effects estimates. ***Trajectory switchers*** are households whose participation histories are not constant across all three waves, including clean adoption, clean exit, and non-monotonic patterns. These households account for 37.9 percent of those with complete three-wave data and capture broader participation histories used in the descriptive and sensitivity analyses.

Extension participants exhibited higher fertilizer adoption and mean harvest than non-participants in the unadjusted comparisons reported in *Table 15, Panel C*, while irrigation use was similar across the two groups. Extension participants also reported a greater number of livestock types on average.

Harvest outcomes also show substantial temporal variation across survey waves (*Table 15, Panel D*). Mean reported harvest increased from 410 kg in 2011/12 to 1,362 kg in 2013/14 before declining slightly to 1,257 kg in 2015/16. These means should, however, be interpreted cautiously because the proportion of households with observed harvest data differs substantially across waves. Reported zero harvest was rare among households with observed harvest data in all three waves (1.7% in Wave 1 and less than 0.2% in Waves 2 and 3), indicating that the higher rate of missing harvest data in Wave 1 is not primarily attributable to widespread reported crop failure. Differences in mean harvest across waves may therefore reflect both differences in data completeness and variation in growing-season conditions.

Raw harvest is also highly right-skewed, with skewness ranging from 5.52 in Wave 1 to 25.90 in Wave 3 (*Table 15, Panel E*). To reduce the influence of extreme values, the regression analysis in the following econometric analysis subsection uses the transformation *log(total harvest + 1)*. Adding 1 preserves observations with zero harvest. The transformation substantially reduces the right tail, although the resulting distributions remain moderately left-skewed, with skewness ranging from −0.75 to −0.93 across waves. Accordingly, the log transformation should be viewed as a pragmatic approach to addressing the highly skewed harvest outcome, rather than as an assumption that the transformed outcome is normally distributed.

Table 15, Panel F summarizes the household and geographic covariates used as controls throughout the regression analysis. The average household in the sample had 5.1 members and was headed by an individual with 1.7 years of completed education, although this average is influenced by the large share of household heads with no formal schooling. About one-

quarter of households (25.0 percent) reported access to credit, and the average household was located 66.7 km from the nearest market. The long-run (1981–2010) rainfall baseline averaged 876 mm annually, with substantial variation across locations (SD = 288 mm), reflecting the geographic diversity of the sample.

Table 15. Descriptive Statistics and Extension Coverage

Panel A. *Extension Participation by Wave*

Year	Observations	Extension (%)	Extension Lag (%)
2011/12	3,222	61.9	--
2013/14	3,222	63.6	61.9
2015/16	3,222	63.1	63.6
Total	9,666	62.9	62.8

Panel B. *Extension Switcher Patterns*

Category	Households	Percent
Households with non-missing lagged extension status	3,138	100.0
➢ FE-identifying switchers	662	21.1
➢ Non-switchers	2,476	78.9
Households with complete three-wave data	2,744	100.0
➢ Trajectory switchers	1,039	37.9
▪ Clean adoption	444	16.2
▪ Clean exit	305	11.1
▪ Non-monotonic	290	10.6
➢ Always treated	1,234	45.0
➢ Never treated	471	17.2

Note*: Incomplete-data households (466) are excluded from the trajectory breakdown.*

Panel C. *Outcomes by Lagged Extension Status*

Outcome	No Extension (n = 2,193)	Extension (n = 3,708)	Difference (Extension − No Extension)
Fertilizer (%)	20.3	63.6	+43.3 pp
Irrigation (%)	11.4	11.0	−0.4 pp
Harvest (kg)	1,088	1,413	+325 kg
Livestock types	3.92	4.58	+0.66

Note*: Comparisons use lagged extension status across Waves 2 and 3 and include 5,901 observations with valid lagged extension status.*

Panel D. *Overall Statistics by Survey Wave*

Variable	Wave 1 (2011/12)	Wave 2 (2013/14)	Wave 3 (2015/16)
Fertilizer (%)	46.2	45.9	45.9
Irrigation (%)	11.9	11.1	10.9
Crop count	3.03	6.02	5.97
Harvest (kg)	410	1,362	1,257
Harvest data missing (%)	52.7	13.3	15.4
Observations with non-missing harvest (N)	1,524	2,795	2,725
Reported zero harvest (%)	1.7	0.0	0.1

Livestock types	2.61	2.96	5.25

Panel E. Harvest Distribution and Transformation

Statistic	Wave 1 (2011/12)	Wave 2 (2013/14)	Wave 3 (2015/16)
Raw harvest skewness	5.52	13.80	25.90
Log(harvest + 1) mean	4.97	6.56	6.36
Log(harvest + 1) median	5.30	6.69	6.54
Log(harvest + 1) skewness	−0.75	−0.87	−0.93

Panel F. Household and Geographic Covariates

Variable	N	Mean	SD
Household size	9,663	5.12	2.36
Household head education (years)	9,341	1.73	3.14
Credit access (%)	9,666	25.0	--
Distance to market (km)	9,638	66.7	50.6
Long-run rainfall baseline (mm)	9,638	876	288

***Notes**: Statistics are computed over the full balanced panel (9,666 household-wave observations). These variables enter as controls in all regression specifications reported in the Econometric Analysis section (Tables 18 and 19; full estimates in Appendix Tables A38–A39). Credit access is a binary indicator, and its SD is omitted because it is not an informative summary for a binary variable in this context (see Table A36 for the corresponding baseline comparison by extension status). N reflects the number of non-missing observations for each variable individually.*

As described in Section 2, the Extension variable is constructed from two questions combined into a composite measure. *Appendix Table A36* shows the holder-level response patterns underlying this indicator. Advisory-only participation declines across waves (31.1% → 26.0% → 22.1%), while reporting both program participation and advisory contact increases between Waves 1 and 2 (24.7% → 30.3%) and remains broadly stable in Wave 3. The share of holders reporting neither question affirmatively increases from Wave 1 to Wave 3. Program-only participation remains uncommon throughout, although it is somewhat higher in Wave 3. Because these are wave-specific holder-level distributions rather than a matched panel, the observed differences reflect both changes in reported participation and changes in holder composition across waves.

Turning to the climate conditions underlying the drought interaction, rainfall varied substantially across survey waves. Under the CHIRPS z-score definition ($z < -1$), 56.8 percent of households experienced drought in Wave 1 (2011/12), while Wave 2 (2013/14) was anomalously wet, with no households classified as experiencing drought (Table 16, Panel A). In Wave 3 (2015/16), 18.1 percent of households were classified as experiencing drought.

This variation provides substantial within-household variation in drought exposure across survey waves, which is important for estimating the extension-by-drought interaction.

This wave-to-wave variation also had a clear geographic dimension (Table 16, Panel B). In the waves with drought exposure, the central zones, roughly 8–12°N, had the highest incidence, while the northern zone had the lowest. In Wave 1, drought affected four in five households in the central-south zone. Wave 2 was uniformly wet, with no drought exposure in any zone. In Wave 3, the central-to-north gradient re-emerged, with the central-south zone again the most affected (29.3 percent) and the north unaffected. Figure 8 maps this spatial pattern across all three waves.

Table 16. Rainfall and Drought Incidence by Survey Wave and Location

Panel A. Rainfall and Drought Incidence by Wave

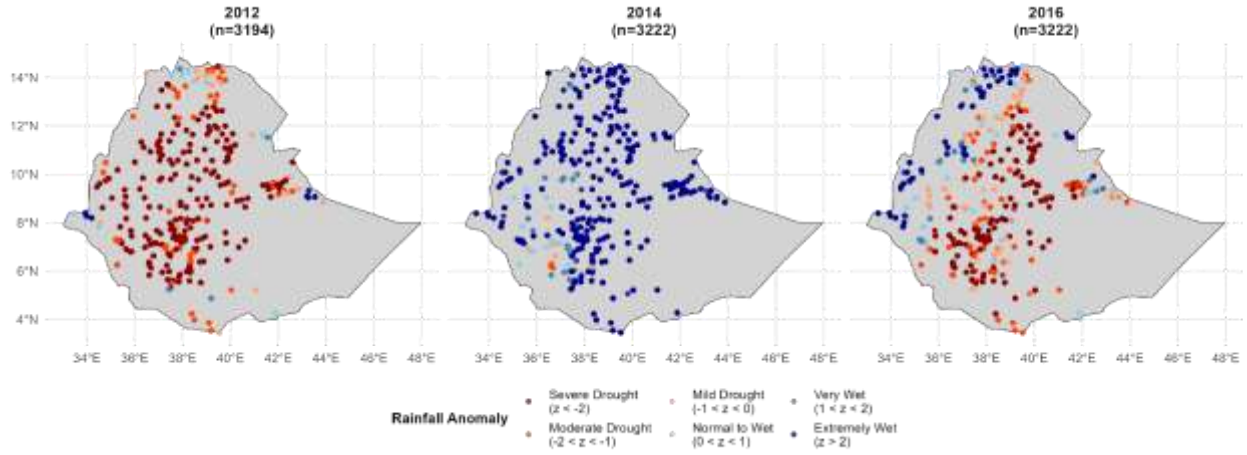
Year	Mean Rain (mm)	Baseline (mm)	Deviation (%)	Drought CHIRPS (%)	Drought 25% (%)
2011/12	719	847	−15.1	56.8	22.8
2013/14	1,586	902	+86.4	0.0	0.0
2015/16	946	880	+3.7	18.1	10.2

Panel B. Drought Exposure by Latitude Zone and Wave

Wave	Zone	N	Mean z-score	Drought (%)
1	South (3–8°N)	1,085	−0.87	45.4
1	Central-South (8–10°N)	992	−1.43	81.5
1	Central-North (10–12°N)	638	−1.48	72.1
1	North (12–18°N)	479	−0.15	11.3
2	South (3–8°N)	1,102	4.56	0.0
2	Central-South (8–10°N)	1,001	6.12	0.0
2	Central-North (10–12°N)	640	7.39	0.0
2	North (12–18°N)	479	5.68	0.0
3	South (3–8°N)	1,102	0.04	13.7
3	Central-South (8–10°N)	1,001	−0.02	29.3
3	Central-North (10–12°N)	641	0.88	21.5
3	North (12–18°N)	478	2.00	0.0

***Notes**: Zones are defined by latitude band (South: 3–8°N; Central-South: 8–10°N; Central-North: 10–12°N; North: 12–18°N). See Figure 8 for the spatial distribution of rainfall anomalies across the sample. Drought incidence is the share of households with a CHIRPS growing-season z-score below −1. N varies slightly across waves due to household attrition and re-listing between survey rounds. Panel A baseline is the household-specific 1981–2010 long-run annual mean from the geo-covariate file. Drought 25% is the alternative definition based on a 25 percent absolute deficit below the long-run annual mean.*

Figure 8. Spatial Distribution of Rainfall Anomalies Across Survey Waves



Notes: *Each point represents a sample household. Colors reflect location-specific CHIRPS rainfall z-scores (1981–2010 baseline). Twenty-eight Wave 1 households lack GPS coordinates and are excluded from the map but included in all regressions.*

Against this backdrop of geographically concentrated drought, extension coverage was slightly lower in drought zones in the raw data, at 60.5 percent compared with 64.3 percent in non-drought zones (*Appendix Table A38*). Baseline comparisons of extension participants and non-participants also reveal systematic differences in several characteristics (*Appendix Table A39*). Extension households were more likely to have access to credit, were slightly larger, and were located farther from markets, with higher mean rainfall z-scores and somewhat lower drought exposure. These raw differences indicate that extension participation was not randomly distributed across households. Whether this pattern reflects selection based specifically on drought exposure, after accounting for these other characteristics, is tested formally in the Econometric Analysis section below (*Appendix Table A40*).

The treatment trajectories raise a natural question: do extension outcomes evolve differently across households with different participation histories over time, and are pre-existing differences visible even before treatment status diverges? *Table 17* addresses these questions in three parts. Households are classified by their three-wave extension history, coded as a three-digit sequence indicating non-participation (0) or participation (1) in Waves 1, 2, and 3, respectively (e.g., “0-0-1” denotes a household that adopted extension only in Wave 3). Panel A presents mean log harvest and fertilizer adoption rates by treatment trajectory and wave. Panels B and C examine pre-existing differences between groups before their treatment

status diverges, addressing the concern that the patterns in Panel A could partly reflect pre-treatment differences rather than the effect of extension. Figure 9 plots the Panel A harvest trajectories with 95 percent confidence intervals.

In Panel A, always-treated households have the highest mean log harvest in every wave, while never-treated households have the lowest. Clean adopters show a sharp increase in harvest following first adoption, while fertilizer use continues to rise through Wave 3. Clean exiters show declining fertilizer use following program exit. The exit trajectories are particularly informative. Households with a 1-1-0 pattern show a statistically significant harvest decline from Wave 2 to Wave 3 of −0.358 log points (paired $t = -3.68$, $p < 0.001$), equivalent to a 30.1 percent reduction. Fertilizer use among this group falls from 42.9 percent in Wave 2 to 26.9 percent in Wave 3. A fixed-effects model comparing clean exit households with always-treated households estimates the exit effect at −0.160 log points (−14.8 percent, $p = 0.056$), marginally short of conventional significance, controlling for drought conditions. These results are directionally consistent with knowledge depreciation after program exit, although the estimate should be interpreted as suggestive rather than confirmatory.

These trajectory patterns should be interpreted with caution. Pre-existing differences between groups are already visible in Wave 1, before treatment status diverges. Panels B and C examine this concern directly. Panel B compares never-treated households with late adopters over the pre-adoption period, while Panel C compares two adoption-timing subgroups at the point where their treatment status first diverges. In Panel B, a harvest gap between never-treated households (0-0-0) and late adopters (0-0-1) is already visible in Wave 1 (0.487 log points, $p = 0.078$) and persists into Wave 2 (0.273 log points, $p = 0.021$), immediately before the late-adopting group receives extension in Wave 3. The persistence of this pre-treatment gap, together with its statistical significance in Wave 2, cautions against interpreting differences between these groups as evidence of a common pre-treatment trajectory.

In Panel C, early adopters (0-1-1) and late adopters (0-0-1), two subgroups that both eventually adopt extension but at different times, are statistically indistinguishable in Wave 1, before either group has adopted (5.00 vs. 4.91, $p = 0.685$). In Wave 2, when the early-adopting group first receives extension while the late-adopting group has not yet done so, a significant divergence emerges (6.74 vs. 6.21, $p < 0.001$), yielding a difference-in-differences

estimate of 0.439 log points (55.1 percent). In Wave 3, once both groups are treated, the gap narrows but persists (6.46 vs. 6.01, p = 0.002).

Table 17. Outcome Trajectories and Pre-Trend Comparisons by Treatment Pattern

Panel A. Outcomes by treatment trajectory and wave

Pattern	Wave	N	Mean log harvest	Fertilizer (%)	Irrigation (%)
Always treated	1	832	5.09	74.6	14.0
Always treated	2	1,224	6.93	75.8	11.0
Always treated	3	1,222	6.76	77.0	13.0
Clean adoption	1	240	4.97	37.5	10.8
Clean adoption	2	433	6.54	41.3	13.4
Clean adoption	3	427	6.28	47.8	11.7
Clean exit	1	164	4.91	45.1	12.2
Clean exit	2	290	6.36	36.2	10.0
Clean exit	3	268	6.05	23.9	7.1
Never treated	1	117	4.42	11.1	10.3
Never treated	2	324	5.94	6.8	15.0
Never treated	3	274	5.75	9.5	11.7

Panel B. Pre-Treatment Comparison: Never-Treated versus Late-Adopting Households

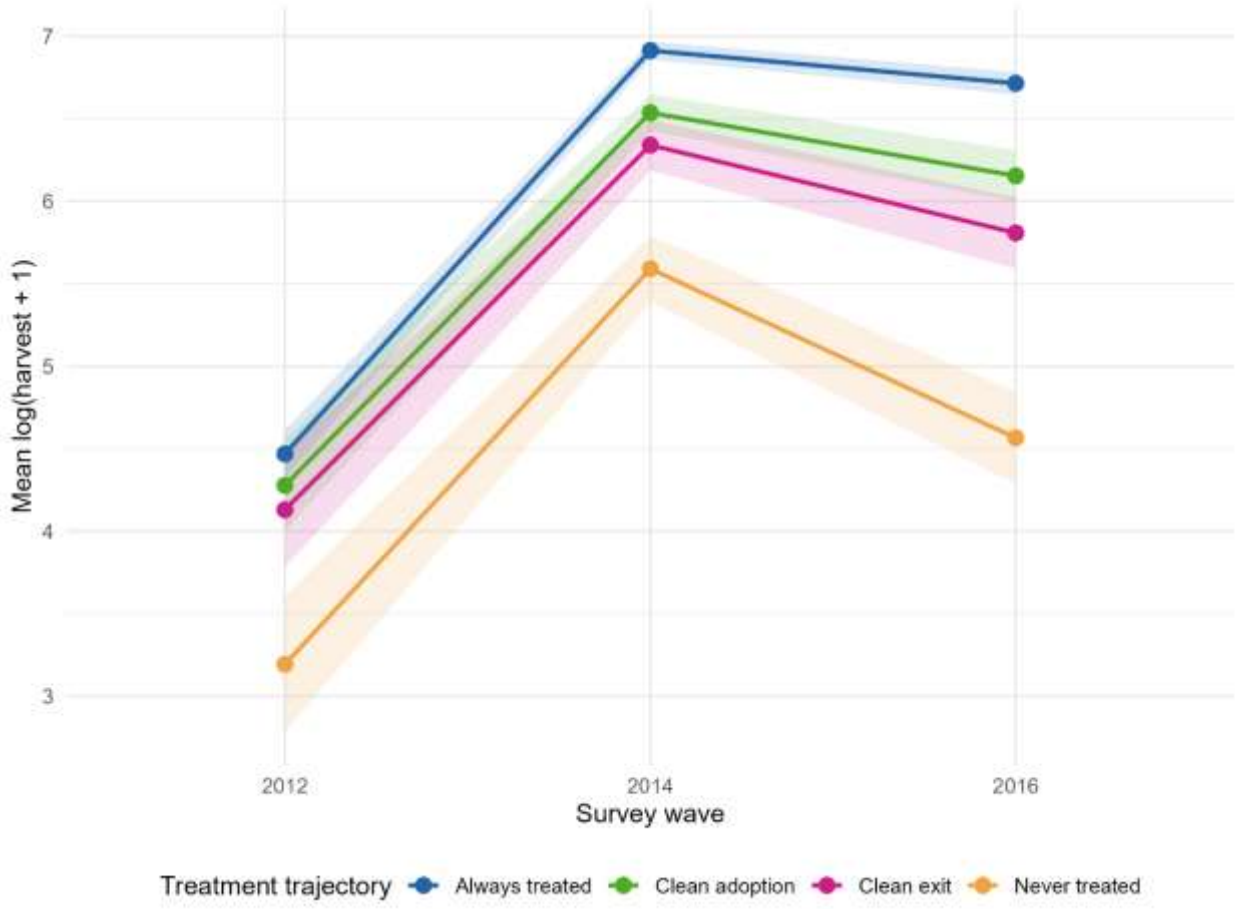
Wave	Never treated (0-0-0) mean log harvest	Late adopters (0-0-1) mean log harvest	Gap	p-value
1	4.42	4.91	0.487	0.078
2	5.94	6.21	0.273	0.021

Panel C. Adoption-Timing Comparison: Early versus Late Adopters

Wave	Early adopters (0-1-1) mean log harvest	Late adopters (0-0-1) mean log harvest	Gap	p-value
1	5.00	4.91	0.10	0.685
2	6.74	6.21	0.53	<0.001
3	6.46	6.01	0.45	0.002

***Notes**: Panel A shows observations with non-missing log harvest. Fertilizer and irrigation are reported as percentage of households. In Panel B, both groups are pre-treatment in both waves shown; the late-adopting group first receives extension in Wave 3. In Panel C, the early-adopting group receives extension beginning in Wave 2 and the late-adopting group beginning in Wave 3. Both groups are pre-treatment in Wave 1, and the Wave 2 gap forms the basis of the difference-in-differences estimate reported in the text. In all three panels, mean log harvest is log(total harvest kg + 1) and p-values in Panels B and C are from two-sample t-tests comparing mean log harvest between groups within each wave.*

Figure 9. Harvest Trajectories by Treatment Pattern, 2011/12–2015/16

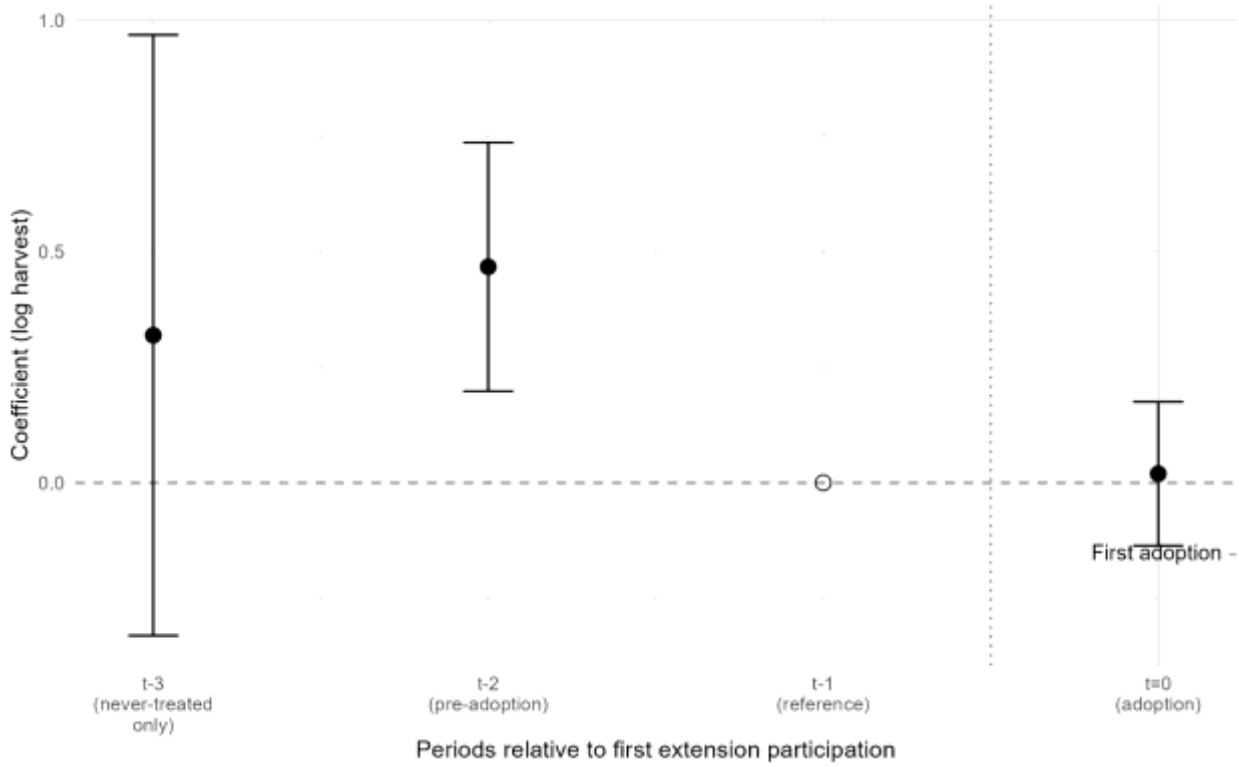


***Notes**: Mean log(total harvest kg + 1) by wave for always treated (n = 1,234), never treated (n = 471), clean adoption (n = 444), and clean exit (n = 305) households, corresponding to Table 15, Panel B. Shaded bands are 95 percent confidence intervals. The clean exit group's decline after Wave 2 (6.36 to 6.05) is consistent with the exit effect (−0.160 log points, p = 0.056) reported above.*

These comparisons, presented in *Table 17* and *Figure 9*, involve groups that differ systematically in participation history and show different outcome levels before their treatment status diverges. Figure 10 formalizes this diagnostic using a relative-time specification that indexes households by periods to first adoption, pooling early (0-1-1) and late (0-0-1) adopters with never-treated households and including household and year fixed effects. Because the panel contains only three waves, the reference period (t−1) is necessarily composite, combining never-treated households' Wave 3 observation, late adopters' Wave 2 observation, and early adopters' Wave 1 observation. Only one post-adoption period (t=0) is estimable; subsequent relative-time indicators are collinear with the year fixed effects. The t−2 coefficient (0.204, p=0.065) falls just short of conventional significance. This should be read alongside the raw Wave 2 comparison in Table 17, Panel B, which shows a statistically

significant pre-treatment gap (p = 0.021). Together, these do not support a clean parallel-trends assumption. We therefore interpret the pre-trend evidence as suggestive rather than conclusive.

Figure 10. Event Study: Harvest Coefficients Relative to First Extension Participation



***Notes**: Relative-time coefficients for log harvest around first extension adoption, pooling early (0-1-1) and late (0-0-1) adopters with never-treated households; household and year fixed effects; reference period t−1. The coefficients are 0.585 (p=0.021) for t−3, 0.204 (p=0.065) for t−2, and 0.054 (p=0.461) for t=0; t=1 is omitted due to collinearity with year fixed effects. N = 915 households and 1,657 observations.*

Econometric Analysis

Table 18 presents estimates from the baseline models. In the two-way fixed-effects specifications, lagged extension participation is associated with a 5.7 percentage point reduction in fertilizer adoption (p = 0.003) and a 15.0% reduction in total harvest (p = 0.005). Lagged extension participation is not significantly associated with irrigation adoption (p = 0.365). The final row reports a separate cross-sectional OLS model in which a one-unit increase in a household's mean lagged extension status, available for Waves 2–3, is associated with a 0.059 reduction in the within-household CV in harvest (p = 0.001).

Table 18. Estimates of Lagged Extension Participation on Agricultural Outcomes

Outcome	Model	Coeff.	SE	N	Households	Within R^2
Fertilizer adoption	Fixed effects	−0.057**	0.019	5,168	2,584	0.008
Log total harvest†	Fixed effects	−0.163**	0.058	4,580	2,290	0.007
Irrigation adoption	Fixed effects	−0.011	0.013	4,758	2,379	0.002
Harvest CV‡	Cross-sectional OLS	−0.059**	0.018	2,691	2,691	--

***Notes**: † Harvest percentage change computed as 100 × (e^{β} − 1); coefficients for the binary adoption outcomes are interpreted directly as percentage-point changes. ‡ Household's mean lagged extension status (Waves 2–3) regressed on the within-household CV in harvest. Cross-sectional OLS, since fixed effects do not apply to a time-invariant outcome. Full covariate estimates for all four models are reported in Appendix Table A41.*

Table 19 examines whether the association between extension participation and harvest differs under drought conditions. The drought main effect is −0.905 ($p < 0.001$), corresponding to a 59.5% reduction in harvest among non-extension households. The extension main effect is −0.209 ($p < 0.001$), while the extension × drought interaction is 0.485 ($p < 0.001$). Under drought conditions, the estimated association of lagged extension participation is 0.276 log points, corresponding to 31.8% higher harvest relative to non-extension households. The interaction estimate is sensitive to treatment timing. Under the contemporaneous specification, the interaction is −0.030 ($p = 0.778$), compared with 0.485 ($p < 0.001$) under the lagged specification (*Appendix Table A43*). Both specifications retain a significant negative drought main effect, but the interaction is significant only under lagged timing. Given the simultaneity concern associated with contemporaneous extension participation, this sensitivity limits the interpretation of the estimated buffering effect and does not provide evidence of a robust buffering effect.

For fertilizer adoption, neither the drought main effect nor the interaction is significant. Fertilizer use therefore does not appear to explain the harvest buffering estimate. For irrigation, the drought main effect is negative (−0.071, $p = 0.006$), while the interaction is positive but not significant (0.047, $p = 0.175$). Raw adoption rates (*Appendix Table A44*) show irrigation increasing during drought in both groups, with a larger increase among non-extension households. This pattern does not support an extension-driven irrigation response or differential irrigation adoption by extension status during drought.

Table 19. Lagged Extension Participation Effects by Drought Conditions

Outcome	**Extension**	**SE**	**Drought**	**SE**	**Extension × Drought**	**SE**	**N**	**Households**	**Within R^2**
Log total harvest†	−0.209***	0.058	−0.905***	0.117	0.485***	0.144	4,580	2,290	0.047

Fertilizer adoption	−0.055**	0.020	0.023	0.025	−0.019	0.033	5,168	2,584	0.008
Irrigation adoption	−0.016	0.013	−0.071**	0.026	0.047	0.035	4,758	2,379	0.006

***Notes**: † Harvest percentage changes computed as 100 × (e^{β} − 1). Coefficients for the binary adoption outcomes are interpreted directly as percentage-point changes. Full covariate estimates for all three models are reported in Appendix Table A42.*

Appendix Table A45 reports the extension × drought interaction under eight alternative specifications, in addition to the baseline. Seven of these alternatives retain household and wave fixed effects, and all seven yield a positive and statistically significant interaction. The estimates range from 0.389 under a harvest-per-crop outcome to 1.197 under the 25 percent rainfall-deficit drought definition. The interaction is also stable when missing harvest is imputed as zero (0.980), when harvest is winsorized at the 1st and 99th percentiles (0.478), and under an inverse hyperbolic sine harvest transformation (0.483, reported in *Appendix Table A46*). It remains stable after excluding the 290 households with oscillating extension histories (0.715), a specification in which the main effect becomes statistically insignificant. It also remains stable after excluding the 269 households whose long-run rainfall baseline swings by more than 100mm across waves (0.498), close to the 0.485 baseline estimate. Restricting the sample to households whose extension status was stable across Waves 2 and 3 yields a similar result (0.724), indicating that the buffering estimate is not driven specifically by switching households.

The eighth alternative is a Wave 3-only cross-sectional specification using region fixed effects rather than household fixed effects. This is the sole exception: the interaction is 0.016 and not statistically significant ($p = 0.919$). Wave 1 contributes no lagged-treatment observations, and Wave 2 had essentially no measured drought. As a result, this specification cannot draw on within-household variation in the way the primary model does. The buffering effect is therefore not independently confirmed under this alternative. We treat this as a limitation of the primary specification rather than as a robustness check that passes.

The drought-buffering association is not an artefact of how extension treatment is defined. Restricting treatment to the single formal-program question (pp_s7q04) produces an extension × drought interaction that remains large and highly significant ($\beta = 1.058$, SE = 0.134, $p < 0.001$), compared with $\beta = 0.485$ (SE = 0.144, $p < 0.001$) under the composite definition (*Appendix Table A46*, column 2 versus column 1). Both specifications also retain

a negative unconditional main effect of similar magnitude. These results indicate that the estimated drought-buffering effect is robust to a narrower definition of extension participation. The results are also robust to an alternative drought threshold, a harvest-per-crop outcome, and a plot-level fertilizer measure drawn from a separate survey module (Appendix, Table A47).

A separate but related concern is endogenous extension-seeking: if drought-exposed households are more likely to enroll in extension, the lagged extension measure could be correlated with drought exposure and complicate causal interpretation. We find no evidence of such a response. Among households not participating in extension, 37.4% of those exposed to drought in wave t entered extension in wave t+1, compared with 35.7% of those not exposed to drought ($\chi^2 = 0.47$, $p = 0.493$). Household fixed-effects linear probability models likewise show no significant association between prior drought and subsequent extension enrollment (AME = −2.5 pp, $p = 0.183$), including when restricting the sample to households not currently in extension (AME = +1.7 pp, $p = 0.638$) and when measuring drought as the household's rainfall deviation from other households in the same enumeration area ($\beta = -0.041$, $p = 0.648$). These results provide no evidence that drought systematically predicts subsequent extension enrollment and are therefore consistent with, but do not establish, an ex-ante interpretation of the buffering relationship (*Appendix Table A48*).

Candidate buffering mechanisms, crop diversification and livestock accumulation, are examined in the Appendix (Mechanism Analysis, Table A49). Neither mechanism exhibits a statistically significant extension × drought interaction. We also considered community-level proximity to extension agents as a potential instrument. However, neither candidate variable provides sufficient identifying variation for this purpose (Appendix, Instrument Feasibility, Table A50).

Baseline comparisons (*Appendix Table A39*) show that extension participation is associated with several observable household and geographic characteristics. To test whether this reflects selection specifically on drought exposure, we estimate a probit model of extension participation on baseline drought exposure and the same covariates (*Appendix Table A40*). Baseline drought exposure does not independently predict extension participation once these covariates are held constant (AME = −0.027, $p = 0.171$). Extension participation is instead associated with household size (AME = +0.014, $p = 0.001$), credit access (AME = +0.147, p

< 0.001), distance to market (AME = +0.002, p < 0.001), and the long-run rainfall baseline (AME = +0.0002, p < 0.001). Thus, although raw extension coverage differs between drought and non-drought locations (*Appendix Table A38*), this difference is not attributable to drought exposure after adjusting for observable characteristics. Because extension participation remains associated with these other observables, we use propensity-score matching alongside household fixed effects to improve comparability between extension and non-extension households, reported below.

The lagged specification addresses the direct form of within-wave simultaneity because participation measured in wave t − 1 precedes the outcome measured in wave t, but it does not eliminate other forms of time-varying selection or anticipatory behavior. Because the lagged specification reduces concerns about contemporaneous simultaneity but does not eliminate time-varying selection, we assess robustness to such selection using propensity-score matching (PSM). Nearest-neighbor matching (1:1), using Wave 1 extension status and baseline covariates and imposing a caliper of 0.25 standard deviations of the estimated propensity score, yields 988 matched pairs. Of the 1,703 treated households, 715 cannot be matched to a suitable comparison household within the specified caliper and are therefore excluded from the matched sample (*Table 20*).

Matching substantially improves covariate balance, reducing the standardized mean difference in rainfall z-score from 0.214 before matching to 0.072 after matching, with every covariate shown in Table 20 below 0.16 after matching. The propensity score itself improved from an SMD of 0.544 to 0.264. The unmatched treated households are less drought-exposed, farther from markets, and located in areas with higher baseline rainfall than the matched treated households. The PSM estimates therefore characterize the matched subpopulation rather than all extension participants.

Table 20. Propensity Score Matching: Sample Sizes and Covariate Balance

Panel A. Matching sample sizes

	Control	Treated
Before matching	1,034	1,703
After matching	988	988

Panel B. Standardized mean differences before and after matching

Covariate	Pre-match SMD	Post-match SMD
Drought (CHIRPS z-score, binary)	0.086	0.016

Rainfall z-score	0.214	0.072
Log(household size + 1)	0.140	0.101
Head education (years)	0.064	0.047
Credit access	0.273	0.156
Log(distance to market + 1)	0.205	0.087
Long-run rainfall baseline	0.266	0.120

Table 21 reports fixed-effects estimates for the baseline and drought-interaction specifications using the full and matched samples. In the specification without the drought interaction, the negative extension association is −0.191 (p = 0.004) in the matched sample, compared with −0.163 (p = 0.005) in the full sample. In the drought-interaction specification, the extension × drought association increases from 0.485 (p < 0.001) in the full sample to 0.584 (p < 0.001) in the matched sample. Thus, matching does not eliminate either the negative extension association or the drought-buffering association. However, because the estimates change after matching, the results do not establish that observable selection is irrelevant.

Table 21. Fixed Effects Estimates for Full and Propensity Score Matched Samples

Specification	Extension main	P	Extension × Drought	P	N
Full sample: no interaction	−0.163** (0.058)	0.005	--	--	4,580
PSM matched: no interaction	−0.191** (0.066)	0.004	--	--	3,244
Full sample: with interaction	−0.209*** (0.058)	<0.001	0.485*** (0.144)	<0.001	4,580
PSM matched: with interaction	−0.237*** (0.065)	<0.001	0.584*** (0.165)	<0.001	3,244

***Notes**: Standard errors in parentheses. The full and matched samples use the same model specifications. However, the matched sample is restricted to households retained by the PSM procedure.*

We also find no evidence of differential buffering by household size, credit access, or education. The three-way interactions are −0.168 (p = 0.556), −0.407 (p = 0.196), and 0.327 (p = 0.290), respectively (*Appendix Table A51*). The subgroup-specific two-way coefficients reported alongside these interactions should not be interpreted as evidence of statistically significant differences in the buffering effect between groups.

A potential concern is that the harvest estimates above are based on a selected sample of households that maintained cultivation. We therefore examine the nature and extent of harvest missingness before interpreting the magnitudes. Harvest data are missing for 2,622 of 9,666 observations (27.1 percent). Two structurally different groups account for this

missingness. The first comprises 315 households with no harvest data in any wave. These households exhibit very low extension coverage (10.3 percent), fertilizer use (5.2 percent), and irrigation (8.7 percent, on a small base of 46 valid observations), compared to 65.0 percent, 47.7 percent, and 11.9 percent respectively among households with at least some harvest data. Their permanent absence from the harvest module is consistent with structural non-participation in the cultivation/harvest module rather than repeated crop failure. The second group comprises 1,465 households missing harvest data in one or two waves. Among this group, 1,253 households are missing in one wave and 212 are missing in two waves.

Among households that do cultivate, harvest data availability differs sharply by extension status. During drought years, 42.9 percent of non-extension households lack harvest data compared to 4.2 percent of extension households. During non-drought years, the corresponding rates are 20.3 percent and 3.3 percent (*Appendix Table A52*). Zero or missing harvest is also significantly more common among non-participants than participants when tested separately by wave ($\chi^2 = 157.7$ for Wave 2, $\chi^2 = 420.9$ for Wave 3, both $p < 0.001$; *Appendix Table A52*, *Panel B*), consistent with this same composition effect rather than a floor effect unique to Wave 1. Logistic regression confirms that extension participation reduces the probability of missing harvest data substantially, with an odds ratio of 0.145 ($p < 0.001$; *Appendix Table A53*). When the always-missing non-farming households are excluded from this regression, the extension coefficient falls from −1.933 to −0.807, confirming that the large missingness association in the full sample is driven substantially by structural non-participation, although a substantial association remains among households with some harvest observations.

To assess whether panel attrition could threaten the validity of the estimates, a logistic regression of panel dropout was estimated using baseline household characteristics, including household size, baseline drought exposure, extension participation, and region fixed effects. Household size was the only household-level characteristic significantly associated with attrition. Each additional household member was associated with a 1.7 percentage-point reduction in the probability of dropout (AME = −0.017, $p < 0.001$). Baseline drought exposure (AME = −0.004, $p = 0.719$) and extension participation (AME = 0.003, $p = 0.793$) were not significantly associated with attrition after conditioning on the observed covariates and region fixed effects (*Appendix Table A37, Panel B*).

Unadjusted attrition rates were lower among extension participants than among non-participants, at 3.9 percent and 6.8 percent, respectively ($p < 0.001$). Attrition was also substantially higher among households with missing extension information, at 17.8 percent ($p < 0.001$), and varied across regions, ranging from 0.8 percent in Harari to 23.3 percent in Gambela (*Appendix Table A37, Panel A*). These differences were not observed for extension participation in the adjusted model. The high retention rate and the absence of a statistically significant adjusted association between extension participation and dropout reduce concerns about differential attrition bias in the fixed-effects estimates.

4. Discussion

This study examines whether prior agricultural extension participation in Ethiopia is associated with agricultural outcomes and input use, and whether these associations vary with drought exposure in ways consistent with an insurance-like buffering function (climate-contingent buffering relationship). It further examines whether these relationships are robust to observed differences between participants and non-participants, extend beyond specific drought episodes to encompass harvest volatility across survey waves, and may operate through crop and livestock diversification. The analysis uses the three-wave Panel I of the Ethiopia Socioeconomic Survey from 2011/12 to 2015/16.

The findings suggest that agricultural extension may serve more as a form of climate-risk buffering than as an unconditional productivity-enhancing intervention. In the baseline fixed-effects specifications, lagged extension participation is negatively and statistically significantly associated with harvest ($\beta = -0.163$, $p = 0.005$). At the same time, it is associated with lower within-household harvest volatility ($\beta = -0.059$, $p = 0.001$), and this association holds across survey waves generally rather than being tied to any single drought episode.

During drought, however, prior extension participation is associated with substantially smaller harvest losses. The extension × drought interaction itself is $\beta = 0.485$ ($p < 0.001$). Combined with the negative main effect, this implies a net 31.8 percent higher harvest for extension households relative to non-extension households specifically during drought years, even though the unconditional (non-drought) association between extension participation and harvest remains negative. For fertilizer adoption, by contrast, the extension × drought interaction is not statistically significant ($\beta = -0.019$, $p = 0.557$). Fertilizer use therefore does

not display the same drought-contingent pattern observed for harvest, a result broadly consistent with evidence that fertilizer allocation decisions in Ethiopia are often driven by plot-level heterogeneity and knowledge gaps rather than climate-risk considerations (Wossen et al., 2018).

The drought-buffering association is robust across seven of eight alternative specifications. It remains present when extension participation is defined using a narrower, single-question measure. A related but distinct analysis further shows that this association is not explained by drought-driven enrollment into the program (see below). The one specification that does not reproduce the interaction raises an important identification concern. This specification uses only Wave 3 data and region rather than household fixed effects. The panel-model interaction is also estimated almost entirely from Wave 3 variation. The results should therefore be interpreted as evidence of an association rather than as evidence of a causal buffering effect.

The findings remain similar after adjusting for observable baseline selection using propensity score matching (PSM). In the matched sample, the negative association between extension participation and harvest remains statistically significant ($\beta = -0.191$, $p = 0.004$). The interaction between prior extension participation and drought is also positive and statistically significant ($\beta = 0.584$, $p < 0.001$). Matching does not attenuate either association. The matched-sample estimate for the main effect is somewhat larger in magnitude than the full-sample estimate. The interaction estimate is also larger. These results suggest that observable baseline differences between extension and non-extension households do not explain either the negative unconditional association or the more favorable association during drought. However, because the matched estimates differ from the full-sample estimates, PSM does not establish that observable selection is irrelevant; rather, it shows that the qualitative pattern persists in the matched subpopulation.

Extension participation is associated with several baseline characteristics, including long-run rainfall conditions, market distance, household size, and credit access. However, baseline drought exposure does not independently predict participation after conditioning on these characteristics. This finding provides no evidence that households entered extension specifically in response to drought exposure. The observed drought-buffering association therefore does not appear to be driven simply by drought-induced enrollment.

Panel attrition is a related, though distinct, form of selection: households that dropped out of the panel do not differ significantly from surviving households on baseline drought exposure or extension participation once household size and region are accounted for (Appendix Table A37), suggesting that differential attrition is unlikely to be a major source of bias in the fixed-effects estimates.

The negative unconditional association between extension participation and harvest may instead reflect the characteristics and circumstances of households that participate in extension. It may also reflect differences in production environments that are not fully observed or other forms of time-varying selection that cannot be eliminated with the available data. These findings suggest that prior extension participation is associated with greater resilience to drought-related production losses, even though extension participation is negatively associated with unconditional harvest outcomes. However, the concentration of identifying variation in Wave 3 and the failure of the Wave 3-only specification to reproduce the interaction limit the strength of the causal interpretation. The evidence is therefore suggestive of a drought-buffering association but does not establish that agricultural extension causally improves resilience to drought.

The pattern of results is consistent with the climate-insurance mechanism proposed by Dercon and Christiaensen (2011), whereby interventions that reduce the consequences of adverse shocks can generate benefits that are not visible in unconditional average treatment effects. This finding illustrates why evaluations based only on unconditional average effects may overlook program benefits that emerge primarily under adverse climatic conditions. A related caution applies to adoption-focused metrics more broadly: in Ethiopia's fertilizer-blending facility program, substantial increases in input adoption did not translate into detectable gains in yields, crop revenue, or household consumption (Assefa et al., 2025), underscoring that adoption alone is an unreliable proxy for welfare impact.

The findings are more consistent with an ex-ante preparedness mechanism than with a contemporaneous-response mechanism. Several mechanisms could plausibly generate such a pattern, including variety choice, the timing of input decisions, soil and water management, or other changes in farm management. However, the mechanism analysis does not identify any of these channels directly, so the ex-ante preparedness interpretation remains suggestive rather than demonstrated.

The dynamic selection analysis provides additional evidence against a simple reactive-enrollment explanation. Drought exposure does not significantly predict subsequent extension participation either in the full-sample fixed-effects model (AME = −0.025, p = 0.183) or in the specification restricted to households not currently participating in extension (β = 0.017, p = 0.638). Within-community comparisons likewise provide no evidence of subsequent enrollment responses to rainfall deviations. These results are consistent with, but do not establish, an ex-ante interpretation of the buffering association.

The results are also consistent with a broader literature emphasizing that extension performance may depend on the conditions under which services are evaluated (Anderson & Feder, 2007). This distinction may help explain why evaluations emphasizing technology adoption or unconditional productivity effects can reach different conclusions about extension performance, including evidence that extension's influence on adoption is strong initially but fades relative to peer learning as coverage expands (Berhane et al., 2018; Krishnan & Patnam, 2014).

Differential missingness in harvest outcomes is an important qualification to the drought-buffering result. Harvest is substantially more likely to be observed among extension households than among non-extension households, particularly during drought. This pattern raises the possibility that part of the observed buffering association reflects differential selection into observed harvest outcomes rather than differences in realized production alone. The bounding exercise reinforces this concern: when unexplained missing observations in the key extension × drought cell are reassigned to the 99th percentile of observed harvest, the interaction changes sign and becomes statistically insignificant (β = −0.319, p = 0.064).

Treatment-history contamination provides a second qualification. Among Wave 3 households classified as untreated under the one-period lag, 28.9 percent had participated in extension in Wave 1. If knowledge or practices acquired through earlier participation persist beyond the period of reported participation, these households may retain some benefits of previous exposure while being classified as untreated in the main specification. The positive two-period lag association in the pooled regional model is consistent with this possibility, although it cannot be interpreted causally because a household fixed-effects specification is not estimable with only one Wave 3 observation per household. Importantly, adding the two-period lag to the Wave 3 cross-sectional model does not explain the failure of that cross-

section to reproduce the panel interaction: the interaction remains statistically insignificant both with and without the additional lag ($\beta = 0.016$, $p = 0.919$ and $\beta = 0.169$, $p = 0.304$, respectively).

The irrigation results provide limited support for interpreting the buffering association as evidence of capital-intensive adaptation. The extension × drought interaction for irrigation is positive but statistically insignificant ($\beta = 0.047$, $p = 0.175$), while the descriptive increase in irrigation during drought is larger among non-extension households. The data therefore do not identify irrigation as the principal channel underlying the estimated harvest buffering. One possible explanation is that irrigation responses depend on physical water infrastructure and access to capital, which advisory services alone may be unable to provide.

Crop diversification and livestock accumulation were also examined as candidate channels and show the same null pattern. The extension × drought interaction is not statistically significant for crop count ($\beta = -0.072$, $p = 0.725$) or livestock species owned ($\beta = -0.154$, $p = 0.557$). Neither diversification measure provides evidence of a buffering mechanism, and the irrigation results likewise do not support irrigation as the principal channel. The mechanism underlying the estimated harvest buffering therefore remains unidentified in this analysis.

Similarly, the heterogeneity analysis provides no statistical evidence that the estimated buffering association differs systematically by household size, credit access, or education. The three-way interaction coefficients are −0.168 ($p = 0.556$), −0.407 ($p = 0.196$), and 0.327 ($p = 0.290$), respectively. These null results should not be interpreted as evidence that the association is identical across households, particularly given the limited statistical power of three-way interaction models. Accordingly, the findings do not provide an empirical basis for targeting extension on the basis of household size, credit access, or education.

Several limitations constrain causal interpretation. First, the drought interaction is identified predominantly from Wave 3, as detailed above. A longer panel spanning multiple drought episodes would provide a stronger basis for distinguishing a general climate-buffering relationship from a wave-specific association. This concern is not merely hypothetical. In a separate long-running Ethiopian household panel covering 1999 to 2009, Krishnan and Patnam (2014) show that extension effects on technology adoption varied substantially across

survey rounds. An effect was detectable in 1999 but had largely disappeared by 2004 and 2009, even as extension coverage expanded considerably over the same period.

Second, the evidence on pre-trends supporting the parallel-trends assumption underlying the fixed-effects design is borderline rather than conclusive: the event-study coefficient at t−2 falls just short of conventional significance ($p = 0.065$), and the raw pre-treatment comparison between never-treated and late-adopting households shows a statistically significant gap in Wave 2 ($p = 0.021$). This does not invalidate the fixed-effects estimates but means that the identifying assumption cannot be confirmed with the available three-wave panel. More generally, the short three-wave panel provides only limited scope for diagnosing pre-treatment trends or sequential exogeneity.

Third, extension participation is measured as a binary indicator that does not capture the intensity, frequency, content, quality, or timing of advisory contact. Field assessments of the Ethiopian extension system describe it as predominantly top-down and supply-driven, with substantial variation across districts in development-agent capacity and turnover. They also document low capacity and morale among extension staff, high turnover, and shortages of operational budgets and facilities (Berhane et al., 2018; Gatdet, 2022; Gebremedhin et al., 2006), alongside gaps in the activity and resourcing of Farmer Training Centre management committees (MoANR, 2017). These differences suggest that a binary indicator may pool households receiving materially different levels and types of extension services under a single treatment label. This kind of proxy, which counts contact rather than measuring its content or quality, is a long-recognized limitation of extension evaluation. Because causal impacts are difficult to trace, programs and evaluators alike often rely on more easily observed input indicators such as contact frequency (Anderson & Feder, 2007). The composite construction of the indicator compounds this: because a household is coded as a participant if any agricultural holder answers affirmatively to either underlying question, the household-level measure may mask heterogeneity in extension exposure across agricultural holders within the same household.

Fourth, PSM is based on Wave 1 treatment status, whereas the principal regression uses time-varying lagged treatment. 13.8 percent of matched-sample observations have a lagged status that disagrees with the Wave-1 status used for matching, so the matched estimates should be read as sensitivity evidence rather than a definitive correction for selection.. The PSM

analysis therefore improves baseline comparability within the matched subpopulation but does not directly address time-varying selection in the fixed-effects models. Fifth, differential harvest missingness, discussed earlier, remains a threat beyond the composition effect already identified: reassigning unexplained missing observations in the extension × drought identifying cell to the 99th percentile of observed harvest reverses the sign of the interaction. Finally, residual treatment-history contamination, described above, may attenuate the estimated effect.

Under the sequential-exogeneity assumption, the coefficient on lagged extension participation is identified from within-household changes over time, conditional on household and wave fixed effects and the observed time-varying covariates, and can be interpreted as the effect of prior participation on the subsequent outcome. Household fixed effects remove time-invariant differences between households, and the lagged specification addresses direct within-wave simultaneity with the fertilizer outcome. Neither approach, however, rules out selection on unobserved time-varying factors, differences in treatment timing, or heterogeneous treatment effects. Because the analysis is observational, the estimated coefficients are best interpreted as conditional associations rather than established causal effects. Given the concentration of drought variation in Wave 3, the missing-harvest sensitivity, and the failure of the Wave 3-only specification to reproduce the interaction, this qualification is particularly important for the drought-buffering estimate.

The policy implications follow from treating climate-contingent performance as an important dimension of extension effectiveness. Evaluation frameworks could complement conventional productivity measures with indicators of drought-period harvest losses, production stability, and harvest-reporting completeness. The results also suggest that sustained engagement before climate shocks may be more relevant than reactive provision after shocks occur. This preparedness interpretation should, however, be treated with some caution.

Evidence from a separate Ethiopian panel shows that extension's marginal effect on technology adoption faded over time as neighbor-to-neighbor learning took hold (Krishnan & Patnam, 2014), raising the broader question of whether extension's influence on farm practice, preparedness-related or otherwise, can be sustained through continued investment in visits alone. The evidence does not justify targeting extension exclusively toward areas

with high drought frequency because the study does not estimate the welfare or cost-effectiveness of alternative targeting strategies. Combining extension with complementary investments in irrigation infrastructure, input finance, and risk-sharing mechanisms is therefore a plausible policy direction. This is consistent with Anderson and Feder (2007), who emphasize the importance of complementary policy conditions for the effectiveness of extension, although the relative effectiveness and cost-effectiveness of such combinations require separate evaluation.

Finally, the results point to the importance of service design and local adaptation. Rather than relying exclusively on blanket recommendations, advisory services could provide landscape-specific guidance that improves input efficiency under local soil and water conditions (Desta et al., 2023). This need is underscored by Ethiopia's own experience of shifting from decades of uniform nitrogen (N) and phosphorus (P) recommendations toward soil-mapped blended fertilizers. On-farm verification of the new blends has produced inconsistent results, with their benefits appearing highly dependent on how well recommendations were matched to local conditions (Erkossa et al., 2022). Evidence of within-household fertilizer misallocation across plots, with continued use on low-return parcels alongside non-use on high-return parcels, points to a similar conclusion. The authors attribute this pattern in part to farmers' limited knowledge of plot-specific returns and call for more targeted, site-specific advisory support (Wossen et al., 2018).

Complementarities between supply-side input access and farmer education have also been documented in Ethiopia. Exposure to demonstration plots strengthened the effect of expanded fertilizer access on adoption (Assefa et al., 2025), suggesting that extension-style engagement may be most effective when paired with a concrete improvement in access or technology, rather than delivered as a standalone service. This is consistent with the Ethiopian government's own extension strategy, which identifies client-orientation and demand-driven service delivery, rather than uniform, supply-driven provision, as a central unresolved gap in the current system (Berhane et al., 2018; MoANR, 2017). Evidence from other settings, including recent research on farmer preferences for advisory services, further suggests that service design, provider characteristics, and participation costs can influence engagement (Chiadmi & Fockaert, 2026).

5. Conclusion and Implications

This study finds that prior extension participation is associated with greater harvest resilience during drought, suggesting that the benefits of extension may depend on climatic conditions. However, the evidence is observational, and the drought-buffering association is not reproduced in the Wave 3 only specification. The findings therefore do not establish that extension causally increases drought resilience. Several priorities emerge from these findings. Future research should use longer household panels spanning multiple drought episodes to determine whether the association persists across shocks and to exploit repeated within-household variation in climate exposure. Stronger experimental or quasi-experimental designs could identify whether the content, timing, and intensity of extension shape drought preparedness and resilience. More detailed data on the timing and content of extension contacts would provide a further test of the preparedness mechanism suggested by the lagged results. Finally, research should examine extension alongside irrigation, credit, insurance, and input-market interventions, since advisory services alone may have limited capacity to protect farmers from severe climatic shocks.

Appendices

Table A34. ESS Panel Structure and Sample Retention

Panel A. Panel Composition and Household Retention

Panel	Wave	Year	Baseline HH	HH with Ag Data	Retention %
Panel I	1	2011/12	3,969	3,118	100.0
	2	2013/14	3,777	3,004	96.3
	3	2015/16	4,951	2,968	95.2
Panel II	4	2018/19	6,770	2,759	100.0
	5	2021/22	4,959	1,778	64.4

Notes: *Retention calculated relative to panel-specific Wave 1 agricultural baseline, restricted to households present in the corresponding wave's household roster (see Panel B for their geographic distribution where applicable). Panel I and Panel II have zero household overlap by design.*

Panel B. Zones and Woredas Fully Excluded from Wave 5

Region	Zone	Woreda	Households Lost
Tigray	Entire region	Entire region	676
Benishangul-Gumuz	Assosa	Entire zone	30
Afar	Fanti /Zone 4	Entire zone	59
Benishangul-Gumuz	Kamashi	Entire zone	10
Afar	Kilbati /Zone 2	Entire zone	148
Oromia	West Guji	Entire zone	16
Oromia	West Wellega	Entire zone	15
(+20 fully dropped woredas within otherwise-surviving zones: Amhara ×5, Afar ×3, Oromia ×5, Benishangul-Gumuz ×4, Somali ×3, SNNP ×1)			298
Total			1,**252**

Table A35. Agricultural Data Coverage by Rural/Urban Status

Wave	Location	Total HH	With Ag Data	Coverage %	Composition %
Wave 1 (2011/12)	Rural	3,466	2,944	84.9	87.3
	Urban	503	174	34.6	12.7
Wave 4 (2018/19)	Rural	3,115	2,759	88.6	46.0
	Urban	3,655	0	0.0	54.0

Notes*: Wave 1 "urban" households include peri-urban farmers, with 34.6% having agricultural data. Composition (%) indicates the share of the total sample in each location category within each wave. The Wave 4 urban sample represents purely urban areas, with no reported agricultural activity. Thus, the sample composition shifted from 87% rural in Wave 1 to 46% rural in Wave 4.*

Table A36. Extension Question Response Patterns by Wave (holder level)

Category	Wave 1 n	Wave 1 %	Wave 2 %	Wave 3 %
Neither question answered "Yes"	1,175	36.3	41.9	43.9
Advisory only	1,008	31.1	26.0	22.1
Both program and advisory	801	24.7	30.3	30.5
Program only	72	2.2	1.8	3.5
Missing on both questions	183	5.6	0	0

Notes*: One household may report multiple agricultural holders. This table is computed at the holder level, before the household-level aggregation used to construct the composite extension indicator (see Methods). Wave 1 percentages are computed among the 3,056 holders with a valid response to at least one question, excluding*

the 183 holders (5.6% of all Wave 1 holders) missing on both. Waves 2 and 3 have no missing responses, so their percentages use the full holder count as the denominator. Waves 2 and 3 have no missing responses. For the composite household-level indicator, holders missing on both questions are coded NA rather than 0. The table presents wave-specific holder-level response patterns and does not represent a matched panel of the same holders across waves.

Table A37. Attrition Analysis

Panel A. *Differential Attrition by Baseline Characteristics*

Characteristic	N	Attrited	Attrition Rate	Comparison	p-value
Household size					
▪ Attrited households	244	244	100.0%	Mean household size: 3.03	<0.001
▪ Surviving households	3,222	0	0.0%	Mean household size: 5.02	
Extension status					
▪ Missing extension data[1]	544	97	17.8%	vs. households with data	<0.001
▪ No extension	1,133	77	6.8%	vs. extension participants	<0.001
▪ Received extension	1,789	70	3.9%		
Region[2]					
▪ Gambela	120	28	23.3%		
▪ Somali	237	37	15.6%		
▪ Afar	120	18	15.0%		
▪ Benishangul Gumuz	120	9	7.5%		
▪ Amhara	728	50	6.9%		
▪ SNNP	885	51	5.8%		
▪ Tigray	360	20	5.6%		
▪ Oromia	656	28	4.3%		
▪ Dire Dawa	120	2	1.7%		
▪ Harari	120	1	0.8%		

Notes*: p-values are from two-sample proportion tests and t-tests. These are unadjusted comparisons and do not control for other covariates; the regression-adjusted association between extension participation and attrition is reported in Panel B.*

[1]Households without a recorded extension response (predominantly pastoral or mobile households not administered the module) are reported separately from "No extension," since they have substantially higher attrition (17.8%) and combining them with non-participants would overstate the extension-attrition relationship.

[2]Regional attrition rates are reported descriptively. Panel B tests region as a fixed effects control rather than reporting individual regional coefficients.

Panel B*. Logistic Regression Predictors of Attrition*

Variable	Standard logit AME	SE	Z	p-value	Firth-type AME	SE	z	p-value
Household size	−0.017	0.002	−7.30	<0.001	−0.017	0.002	−7.39	<0.001
Baseline drought	−0.004	0.010	−0.36	0.719	−0.003	0.011	−0.32	0.748
Extension participation	0.003	0.010	0.26	0.793	0.002	0.010	0.23	0.816

Notes: *Household-level coefficients (household size, drought, extension) are materially unchanged across the two models, indicating they are not affected by the separation problem. Extension participation is not significantly associated with attrition under either specification. Region-level coefficients are not reported in this panel, as region enters the model only as a fixed effect. Regional attrition patterns are shown descriptively in Panel A. The estimation sample contains 2,903 households with complete information on the relevant covariates. 563 households are excluded due to missing extension or other covariate information.*

Table A38. Extension Program Targeting by Drought Exposure (Wave 1)

Zone	Households	Extension Coverage (%)	Mean Rainfall (mm)	Mean Z-score
Drought zone ($z < -1$)	1,815	60.5	693	−1.72
Non-drought zone ($z \geq -1$)	1,379	64.3	753	−0.19
Difference	--	−3.8 pp	+60 mm	+1.53

***Notes**: Extension coverage is 3.8 percentage points lower in drought zones than in non-drought zones. The comparison is descriptive and does not imply that drought exposure itself determines extension participation.*

Table A39. Extension Program Targeting: Baseline Characteristics by Participation Status

Characteristic	No Extension	Extension	Difference	p-value
Drought exposure 2011/12 (%)	58.5	54.4	−4.0 pp	0.038
Rainfall z-score (mean)	−1.16	−0.98	+0.18	<0.001
Household size (mean)	5.1	5.4	+0.3	0.002
Head education (years)	1.4	1.5	+0.1	0.113
Credit access (%)	16.5	29.3	+12.8 pp	<0.001
Distance to market (km)	63.1	69.2	+6.1	0.002

***Notes**: N = 1,056 (No Extension) and N = 1,719 (Extension); 447 households with missing extension status are excluded. All characteristics are measured at Wave 1 baseline. Welch two-sample t-tests are used for continuous characteristics, while credit access is compared using a test of equality of proportions with continuity correction.*

Table A40. Probit Regression Predicting Extension Participation

Variable	Coefficient	SE	AME
Baseline drought	−0.074	0.054	−0.027
Log(household size + 1)	0.211**	0.065	+0.014**
Head education (years)	0.004	0.010	+0.002
Credit access	0.402***	0.060	+0.147***
Log(distance to market + 1)	0.132***	0.030	+0.002***
Long-run rainfall baseline	0.0005***	0.0001	+0.0002***

***Notes**: Probit regression using Wave 1 baseline data (N = 2,737). AMEs for household size and distance to market are expressed per unit of the underlying (untransformed) variable.*

Table A41. Full-Specification Estimates Underlying Table 18

Variables	(1) Fertilizer adoption	(2) Log total harvest	(3) Irrigation adoption	(4) Harvest CV
Extension (lagged)	−0.057** (0.019)	−0.163** (0.058)	−0.011 (0.013)	−0.059** (0.018)
log(household size + 1)	0.050 (0.034)	−0.041 (0.128)	−0.038 (0.029)	−0.033 (0.020)
Head education (years)	−0.011* (0.006)	−0.036. (0.020)	0.005 (0.004)	0.005* (0.002)
Credit access	0.026. (0.015)	−0.036 (0.046)	−0.016 (0.013)	0.035 (0.022)
Mean household harvest (kg)	--	--	--	1.60×10^{-5}*** (3.93×10^{-6})
Intercept	--	--	--	0.729*** (0.037)
Household FE	Yes	Yes	Yes	No
Wave FE	Yes	Yes	Yes	No
N	5,168	4,580	4,758	2,691
Households	2,584	2,290	2,379	2,691
Adj. R^2	0.650	0.513	0.454	0.013

***Notes**: In column 4, each household contributes one observation. Therefore, household-clustered standard errors are equivalent to heteroskedasticity-robust standard errors. Column 4 uses the household's mean lagged*

extension status across Waves 2–3, regressed cross-sectionally on the within-household harvest CV. Household and year fixed effects do not apply to this time-invariant specification. Extension coefficient in column (2) corresponds to a 15.0% change in harvest via 100 × (e^{β} − 1). Coefficients in columns (1), (3), and (4) are interpreted directly (percentage points for columns 1 and 3 and CV units for column 4). As a sensitivity analysis, the production-risk estimate is robust to restricting the analysis to households with complete three-wave harvest histories. The coefficient is −0.076 (p = 0.001) in the restricted sample, compared with −0.059 (p = 0.001) in the primary specification. The restricted analysis includes 1,440 households after excluding observations with missing regression covariates.

Table A42. Full-Specification Estimates Underlying Table 19

Variables	(1) Log total harvest	(2) Fertilizer adoption	(3) Irrigation adoption
Extension (lagged)	−0.209*** (0.058)	−0.055** (0.020)	−0.016 (0.013)
Drought (CHIRPS z-score)	−0.905*** (0.117)	0.023 (0.025)	−0.071** (0.026)
Extension × Drought	0.485*** (0.144)	−0.019 (0.033)	0.047 (0.035)
log(household size + 1)	−0.066 (0.125)	0.050 (0.034)	−0.040 (0.029)
Head education (years)	−0.039† (0.020)	−0.011* (0.006)	0.005 (0.004)
Credit access	0.004 (0.045)	0.024† (0.015)	−0.013 (0.013)
N	4,580	5,168	4,758
Households	2,290	2,584	2,379
Adj. R^2	0.532	0.650	0.456
Within R^2	0.047	0.008	0.006

***Notes**: Extension × Drought coefficient in column (1) corresponds to a 62.5% relative harvest change via 100 × (e^{β} − 1). Coefficients in columns (2) and (3) are interpreted directly as percentage points.*

Table A43. Contemporaneous versus lagged extension specification on harvest outcome

Specification	B	SE	% effect (transform)	N
Contemporaneous ➢ main effect only	0.160**	0.052	+17.3%	6,522
Lagged ➢ main effect only	−0.163**	0.058	−15.0%	4,580
Contemporaneous × drought ➢ extension main effect	0.162**	0.053	+17.6%	6,519
Contemporaneous × drought ➢ drought main effect	−0.312**	0.096	−26.8%	6,519
Contemporaneous × drought ➢ interaction	−0.030	0.106	−2.9%	6,519
Lagged × drought ➢ extension main effect	−0.209***	0.058	−18.8%	4,580
Lagged × drought ➢ drought main effect	−0.905***	0.117	−59.5%	4,580
Lagged × drought ➢ interaction	0.485***	0.144	+62.5%	4,580

***Notes**: Dependent variable is log(total harvest kg + 1) in all rows. All models include household and year fixed effects. Percentage effects use 100 × (e^{β} − 1).*

Table A44. Irrigation Adoption by Drought Exposure and Lagged Extension Status

Drought	Extension (lagged)	N	Irrigation (%)
No	No	1,897	10.7
No	Yes	3,447	10.9
Yes	No	296	16.7
Yes	Yes	261	12.6

***Notes**: Raw (unadjusted) adoption rates, not regression estimates. Sample restricted to observations with non-missing drought status and lagged extension status.*

Table A45. Robustness Checks: Extension × Drought Interaction

#	Specification	Coefficient	SE	N
1	Baseline (CHIRPS z-score)	0.485***	0.144	4,580
2	Missing harvest imputed as zero	0.980***	0.182	5,200
3	Winsorized at P1/P99	0.478***	0.141	4,580
4	Harvest per crop	0.389**	0.128	4,580
5	Alternative drought (25% deficit)	1.197***	0.207	4,580
6	Excluding oscillating switchers	0.715***	0.153	4,092
7	Excluding geo-covariate-unstable households	0.498***	0.144	4,174
8	Restricted to stable extension-status subsample (Waves 2–3)	0.724***	0.160	3,462
9	Wave 3 only (region FE)	0.016	0.160	2,582

***Notes**: All coefficients are the extension × drought interaction from the primary log-harvest specification unless noted. Row 5 uses a binary drought indicator. Row 7 excludes 269 households whose long-run rainfall baseline swings by more than 100mm across waves (see Variable Construction). Row 8 restricts to the 76.0 percent of Wave 2–3 comparison households whose lagged extension status did not change between waves (23.96 percent switched). The extension main effect is not separately identified in this subsample as it is collinear with household fixed effects.*

Table A46. Sensitivity of the Drought-Buffering Interaction to Extension Definition and Harvest Transformation

Variables	(1) Composite def., log(harvest+1) [baseline]	(2) Single-question def., log(harvest+1)	(3) Composite def., IHS(harvest)
Extension (lagged)	−0.209*** (0.058)	−0.176*** (0.046)	−0.215*** (0.059)
Drought (CHIRPS z-score)	−0.905*** (0.117)	−0.964*** (0.092)	−0.908*** (0.119)
Extension × Drought	0.485*** (0.144)	1.058*** (0.134)	0.483*** (0.146)
➢ As % change in harvest ratio	62.5%	188.1%	n/a[1]
Household FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes
Observations	4,580	4,548	4,580
Households	2,290	2,274	2,290
Within R^2	0.047	0.064	0.047

***Notes**: Column (1) reproduces the primary specification (Table 19) for reference. Column (2) restricts the extension definition to the single formal-program question (pp_s7q04), excluding advisory-services contact. Column (3) replaces log(harvest+1) with the inverse hyperbolic sine of harvest, which accommodates zero-harvest observations without an additive offset. [1]The percentage-change transformation $100\times(\exp^{\beta}-1)$ is exact only for the log specification and is not reported for column (3). Coefficients are similar in sign and significance across all three columns. The buffering result does not depend on the extension-definition or harvest-transformation choice.*

Table A47. Alternative Outcome and Drought Measures

	(1) Alternative drought threshold (25% deficit)	(2) Harvest per crop	(3) Plot-level fertilizer
Extension (lagged)	−0.217*** (0.058)	−0.160** (0.050)	−0.053** (0.020)
Drought	−1.196*** (0.173)	−0.617*** (0.105)	0.046 (0.041)
Extension × Drought	1.197*** (0.207)	0.389** (0.128)	0.059 (0.046)
Household FE	Yes	Yes	Yes

Wave FE	Yes	Yes	Yes
Observations	4,580	4,580	4,760

***Notes**: Column (1) replaces the CHIRPS z-score drought definition with a binary indicator equal to 1 when growing-season rainfall falls more than 25 percent below the household's long-run annual baseline; the percentage-change transformation does not apply to this binary specification. Column (2) uses harvest divided by the number of valid harvest rows per household as the outcome, in place of total harvest. Column (3) uses a plot-level fertilizer measure from a separate survey module (post-planting plot section), rather than the extension-module fertilizer question.*

Table A48. Does Drought Predict Subsequent Extension Enrollment?

Specification	Estimate	SE	p-value	N
Transition rate into extension, prior drought	37.4%	--	0.493[1]	609
Transition rate into extension, no prior drought	35.7%	--	--	1,547
FE-LPM: drought → extension (all households)	−0.025	0.019	0.183	5,772
FE-LPM: drought → extension (non-participants only)	0.017	0.037	0.638	1,200
FE-LPM: within-EA drought deviation → extension	−0.041	0.090	0.648	5,772

***Notes**: Household and year fixed effects included in all regression models. [1]Two-sample proportion test.*

Table A49. Mechanism Analysis: Crop and Livestock Diversification

Panel A. Extension × drought effects on candidate mechanism variables

Variables	Crop count	Livestock species count
Extension (lagged)	−0.120 (0.105)	−0.149 (0.126)
Drought (CHIRPS z-score)	−0.259† (0.153)	0.202 (0.214)
Extension × Drought	−0.072 (0.206)	−0.154 (0.262)
Household FE	Yes	Yes
Wave FE	Yes	Yes
Observations	4,754	5,200

Panel B. Crop diversification as a mediator of the harvest interaction

Variables	Harvest model (no mediator)	Harvest model (with crop count)
Extension × Drought	0.485*** (0.144)	0.352* (0.140)
Crop count	--	0.086*** (0.010)
Crop count × Drought	--	0.102*** (0.018)
Observations	4,580	4,580

***Note**: Both models include household and year fixed effects and the standard control set.*

Instrument feasibility

We investigated whether community-level proximity to extension agents could serve as an instrument for household extension participation. The ESS community questionnaire provides two candidate variables for this purpose, namely whether a development agent resides in the community (cs6q08) and the distance to the nearest agent office (cs6q09_1). The distance variable is recorded only when no agent is resident, following the survey's structural skip pattern, leaving insufficient coverage for an instrument. The binary agent-presence variable has near-complete coverage but provides little within-household variation

over time, which is required for identification in the household fixed-effects specification; we assess this temporal persistence directly by comparing agent-presence status across Waves 1–3 among households observed in all three waves. Agent presence proves highly stable across waves, further limiting its usefulness as a source of within-household identifying variation. This temporal-persistence check is diagnostic only. Because neither measure provides sufficient identifying variation for an instrument (*Appendix Table A50*).

Table A50. Instrument Feasibility Diagnostics

Panel A. Household-level coverage of candidate instruments (Wave 1)

Variable	Valid N	Coverage
Agent-presence indicator (cs6q08)	3,191	99.0%
Distance to nearest agent (cs6q09_1)	223	6.9%

Panel B. Temporal stability of agent-presence status across waves

	N	Share
Households with valid readings in all three waves	3,305	--
Status unchanged across all three waves	2,756	83.4%
Status changed at least once	549	16.6%

***Notes**: Distance coverage is computed among 32 enumeration areas with a non-resident agent. Agent-presence coverage and stability are computed among all Wave 1 households and communities, respectively.*

Heterogeneity

Appendix Table A51 reports the full three-way interaction models underlying the heterogeneity results discussed in the main text. For each dimension, the table shows the extension × drought coefficient evaluated at the reference category, alongside the triple-interaction coefficient testing whether this association differs for the comparison category. As noted in the main text, none of the three triple interactions reaches conventional significance.

Table A51. Heterogeneity of the Drought-Buffering Interaction by Household Size, Education, and Credit Access

Subgroup dimension	Interaction, reference group	Reference group	Triple interaction: comparison group	Comparison group	Significant difference?
Household size	0.527** (0.181)	Small households	−0.168 (0.285)	Large households	No
Education	0.360* (0.175)	Uneducated head	0.327 (0.309)	Educated head	No
Credit access	0.536** (0.167)	No credit access	−0.407 (0.315)	Has credit access	No

***Notes**: SEs in parenthesis. "Interaction, reference group" is the extension × drought coefficient from a three-way interaction model (extension_lag × drought_zscore_chirps × subgroup), evaluated at the reference (omitted) category. "Triple interaction" is the coefficient on the three-way term, testing whether the buffering interaction differs significantly for the comparison group relative to the reference group. All models include household and year fixed effects. None of the three triple interactions is statistically significant at conventional levels. No differential-buffering claim by household size, education, or credit access is supported by this analysis.*

Table A52. Harvest Data Missingness by Drought Exposure, Extension Status, and Wave

Panel A. Harvest availability by drought and lagged extension status

	No Drought	Drought	Difference
No extension	79.7%	57.1%	−22.6 pp
Extension	96.7%	95.8%	−0.9 pp
Difference	+17.0 pp	+38.7 pp	--

***Notes**: The Difference column reports Drought minus No Drought within each extension-status row; the Difference row reports Extension minus No Extension within each drought-status column.*

Panel B. Zero/missing harvest observations by wave and lagged extension status

Wave	Ext. lag = 0, zero/missing %	Ext. lag = 1, zero/missing %	χ^2	p-value
2	16.2%	2.9%	157.7	<0.001
3	30.1%	4.0%	420.9	<0.001

***Notes**: Panel A percentages are the share of household-wave observations with non-missing harvest data, among observations with valid lagged extension status and non-missing CHIRPS z-score; drought defined as z-score below −1. Panel B chi-square tests are computed separately by wave (extension_lag is undefined in Wave 1, so only Waves 2–3 appear). Both panels indicate harvest-data completeness is substantially lower among non-participant households, consistent with the composition effect discussed in the text (Appendix Table A53).*

Table A53. Logistic Regression Predicting Missing Harvest Data

Variable	Coefficient	SE	Odds Ratio
Drought	1.202***	0.141	3.328
Extension (lagged)	−1.933***	0.124	0.145
Drought × Extension	−0.993**	0.359	0.370
Log(household size+1)	−0.812***	0.132	0.444
Head education	0.003	0.019	1.003
Credit access	−0.054	0.113	0.947

***Notes**: N = 5,645. Model estimated using feglm with binomial family. The large extension (lagged) coefficient reflects a composition effect: extension_lag = 0 contains a disproportionate share of always-missing non-farming households. When these 315 households are excluded, the extension coefficient falls to −0.807 ($p < 0.001$), confirming the composition interpretation.*

Robustness Assessment of the Drought-Buffering Interaction

The extension × drought interaction (β_3) is the paper's central estimate of whether prior extension participation buffers households against drought-related harvest losses. The baseline estimate is $\beta_3 = 0.485$ ($p < 0.001$), reported in *Table 19*. The five checks below assess whether this estimate is robust to alternative samples, timing assumptions, missing-

data treatments, and pre-trend tests. Four of the five checks raise concerns about the robustness of β_3. The fifth, subgroup consistency, does not raise such concerns, but it does not resolve the concerns raised by the other four checks either. Accordingly, β_3 should be presented as an association observed under the primary specification rather than as a causal effect (*Appendix Table A54*).

Table A54. Robustness Assessment of the Drought-Buffering Interaction (β_3)

No	Check	What was tested	Result	What it means
1	Source of identifying variation	Whether the interaction holds when estimated on Wave 3 alone (with region fixed effects, since household fixed effects cannot be estimated from a single wave)	Interaction is statistically insignificant (β = 0.016, p = 0.92)	The baseline result is driven almost entirely by Wave 3 data, not by within-household variation in drought exposure across survey waves. Wave 1 supplies no lagged-treatment observations and Wave 2 had essentially no measured drought, so Wave 3 is the only wave contributing drought cases to the interaction.
2	Timing of the treatment variable	Whether the buffering result holds using contemporaneous (same-wave) extension status instead of one-wave lagged status	Not significant under the contemporaneous specification (β = −0.030, p = 0.78), versus significant under the lagged specification (β = 0.485, p < 0.001)	The result is specific to the lagged-treatment specification and does not replicate when treatment and outcome are measured in the same wave.
3	Sensitivity to missing harvest data	Whether the result is stable if the harvest values missing specifically in the "no extension, drought" comparison group are set to the 1st or 99th percentile of observed harvest	Sign of the interaction reverses under the 99th-percentile assumption (β = −0.319, p = 0.064). The 1st-percentile assumption instead strengthens the estimate (β = 0.961, p < 0.001), illustrating the width of the bound	The estimate depends on 54 households whose harvest was never recorded, in the subgroup with the least other data available. We do not know why their harvest is missing, and the data cannot tell us. Assuming a low harvest for these households strengthens the buffering result, while assuming a high harvest reverses it. The sign of the effect in this subgroup cannot therefore be established.
4	Pre-treatment trends	Whether outcomes for late adopters were already trending differently from non-adopters before either group received extension	It is borderline (p = 0.065). It is not significant at 5%, but it is not a clean pass either	A key assumption behind the fixed-effects design, that adopters and non-adopters were on comparable trajectories before treatment, is not clearly confirmed by the data.
5	Consistency across subgroups	Whether the buffering effect differs by household size, education, or credit access	No subgroup shows a statistically significant difference in the buffering effect	The result does not appear to be concentrated in, or driven by, any particular type of household.

CONCLUSION AND POLICY SYNTHESIS

This dissertation examined how climate shocks and resource constraints affect human capital formation in Ethiopia through three distinct but interrelated mechanisms: psychological aspirations among adolescents, biological development in early childhood, and productive stability in smallholder agriculture. Drawing on panel data methods applied to three complementary Ethiopian datasets (the Young Lives Study, Demographic and Health Surveys, and Ethiopia Socioeconomic Survey) the analysis reveals a paradoxical pattern: aspirations remain resilient to shocks, biological development shows acute vulnerability, and institutional support can partially bridge this gap.

In Chapter 1, while simple cross-sectional models suggest that food insecurity significantly reduces educational aspirations ($\beta = -0.0593$, $p = 0.002$), individual child fixed-effects models show that this negative correlation attenuates to a statistically null estimate once persistent, unobserved household characteristics are accounted for. Once time-invariant factors (such as parental agency, community norms, and family culture) are accounted for through fixed-effects estimation, the coefficient becomes statistically insignificant and substantively null ($\beta = -0.0245$, SE = 0.0497, 95% CI: −0.122, 0.073). This pattern reveals aspirational resilience: Ethiopian adolescents anchor their educational goals in enduring structural factors rather than adjusting them in response to transitory nutritional shocks. This finding resonates with Sen's (1985, 1999) capability approach, which emphasizes that individuals' valued functionings and sense of agency can remain consistent even when material circumstances fluctuate. Moreover, it underscores Sen's critical distinction between aspirations (valued ends) and capabilities (actual freedoms to achieve those ends). Ethiopian youth possess high educational aspirations but lack the material capabilities and institutional support to realize them. The bottleneck to educational mobility is not a deficit of ambition but a deficit of resources and opportunities (Tafere, 2014).

Chapter 2 reveals a fundamentally different empirical pattern. In contrast to the null association observed for aspirations, climate vulnerability is strongly and positively associated with child stunting. Across pooled survey waves, and after adjusting for a set of observable child, maternal, household, and community characteristics using multilevel models and accounting for missing data, higher climate vulnerability is consistently linked

to substantially greater odds of stunting (aOR = 1.27, 95% CI: 1.20–1.34; Table 12). Moreover, a survey-year interaction (β = 0.018, aOR = 1.02) was significant only under model-based standard errors (95% CI: 1.01–1.03); once the survey design was accounted for, the interval included the null (95% CI: 1.00–1.03), and the interaction was driven by a strengthening protective effect of adaptive capacity rather than intensifying climate vulnerability, even as average national stunting rates have declined by 7.4 percentage points between 2011 and 2019. This pattern is consistent with well-established biological pathways through which early-life nutritional deprivation is associated with lasting cognitive deficits and diminished educational returns (Victora et al., 2021).

Chapter 3 demonstrates yet another pattern. Extension coverage is lower in drought zones than in non-drought zones, and drought exposure does not independently predict extension participation once other covariates are held constant; selection instead reflects household size, credit access, market distance, and rainfall baseline. After correcting for this selection through propensity score matching, extension shows a statistically significant negative association with harvest overall ($\beta 1 = -0.191$, $p = 0.004$ after PSM). However, the evidence suggests extension may generate value when climate fails, though this should be interpreted with caution given data limitations noted above. During severe droughts, extension participation is associated with substantially smaller harvest losses (interaction coefficient, PSM-matched, $\beta 3 = 0.584$, $p < 0.001$), consistent with, though not proof of, a shock-contingent insurance function. This association is not reproduced in the Wave 3 only specification. Extension-participating households achieve harvests roughly 31.8 percent higher than non-participating households in the same drought year in the full (unmatched) sample, which is a similar but not numerically identical result to the PSM-matched estimate above.

However, neither irrigation nor fertilizer adoption exhibits a statistically significant extension × drought interaction. Similarly, mechanism analyses examining crop diversification and livestock accumulation find no significant interaction effects. No significant heterogeneity in the drought-buffering effect was found by household size, education, or credit access; the point estimate for uneducated household heads was in fact the smallest of the three groups tested. This shock-contingent value aligns with Schultz's (1975) theory of human capital in agriculture.

These findings provide guidance for Ethiopia's integrated human capital strategy, particularly as outlined in the Seqota Declaration (aiming to end child undernutrition by 2030) and the National Agricultural Extension Strategy.

Remove structural barriers to converting educational demand into attainment. Chapter 1 shows that Ethiopian adolescents' educational aspirations are already high and stable, remaining largely unaffected by transitory food insecurity. The policy challenge is therefore not to generate demand for schooling, but to remove the structural constraints that prevent existing aspirations from translating into educational attainment. These constraints include limited school access and affordability, weak household economic capacity, and inadequate early-life nutritional foundations. Policy should focus on strengthening the enabling conditions that allow aspirations to become sustained investments in schooling. This includes complementing safety nets such as the Productive Safety Net Programme (PSNP) with asset-building interventions that strengthen household resilience to economic shocks. Such measures can help protect household resources so that existing aspirations are not undermined by the resource constraints and institutional gaps identified across this dissertation as barriers to attainment.

Integrate nutritional interventions with climate adaptation infrastructure. Conventional food aid or supplementation programs cannot address the strong, persistent relationship between climate vulnerability and child stunting. Policy must prioritize structural determinants of adaptive capacity: water security, maternal education, and sanitation infrastructure that protect children's biological development during the critical first 1,000 days. Given that stunting directly reduces returns to educational investments (Becker, 1964; Victora et al., 2021), these interventions constitute essential infrastructure for human capital formation, not merely social welfare programs.

Reframe agricultural extension as climate insurance rather than technology diffusion. Extension delivers greatest value during drought years, when it prevents harvest collapse that would otherwise trigger household disinvestment in children's nutrition and education. Extension curricula should explicitly incorporate drought-responsive advisory services (drought-triggered irrigation guidance, crop switching recommendations, emergency livestock management) rather than offering uniform agronomic advice. Geographic deployment should prioritize drought-prone zones, and program evaluation should measure

downside risk protection (harvest stability during shocks) rather than average productivity gains.

Couple extension with complementary physical infrastructure. No significant differences in the drought-buffering effect were found across education, household-size, or credit-access subgroups. Extension investments, therefore, offer efficiency gains (protecting aggregate productivity); this dissertation does not provide evidence of differential equity benefits to disadvantaged households. Extension must be coupled with complementary investments in water infrastructure, credit access, and input supply chains to enable farmers to act on knowledge they receive. This supports Schultz's (1961, 1975) argument that human capital in agriculture requires not only cognitive skills but also access to complementary physical capital and market institutions.

Methodologically, this dissertation highlights the importance of longitudinal data and within-unit identification in development microeconometrics. The systematic discrepancies between cross-sectional and propensity score estimates, as well as fixed-effects specifications, across Chapters 1 and 3 demonstrate that much existing policy evidence in Sub-Saharan Africa, which relies predominantly on cross-sectional surveys, likely misidentifies the true causal impact of both shocks and interventions. While propensity score matching can address selection on observables, it cannot eliminate bias from time-invariant unobservables such as parental motivation, community culture, or historical agricultural knowledge. Only panel methods that exploit within-household variation over time can credibly isolate causal effects in such settings (Wooldridge, 2010).

In conclusion, the adolescents of Ethiopia do not lack ambition; their educational aspirations remain stable and high regardless of transitory food insecurity. However, they inhabit an environment in which climate volatility periodically undermines both the biological and material foundations required to transform aspiration into achievement. In Sen's (1999) terminology, they possess aspirations but lack capabilities; in Becker's (1964) framework, they face binding constraints on the production inputs necessary for human capital formation. The physical growth necessary for cognitive development is increasingly vulnerable to climate-driven nutritional stress, and the household livelihoods that finance school continuation face repeated drought-induced collapse.

Yet this dissertation also demonstrates that well-designed institutional interventions can partially bridge this gap. Agricultural extension, when reframed as climate insurance and coupled with complementary infrastructure, can stabilize the material foundation upon which aspirations depend. The policy challenge for Ethiopia is therefore not to generate demand for education, as that demand already exists, but rather to build the climate-resilient infrastructure and institutional support systems that enable households to act on their existing aspirations. Future research should examine how these protective mechanisms operate over longer time horizons and whether early-life institutional support generates persistent improvements in educational attainment and economic mobility.

REFERENCES

Aguilar, A., Carranza, E., Goldstein, M., Kilic, T., & Oseni, G. (2015). Decomposition of gender differentials in agricultural productivity in Ethiopia. *Agricultural Economics*, *46*(3), 311–334. https://doi.org/10.1111/agec.12167

Alam, M. J., Sarma, P. K., Begum, I. A., Connor, J., Crase, L., Sayem, S. M., & McKenzie, A. M. (2024). Agricultural extension service, technology adoption, and production risk nexus: Evidence from Bangladesh. *Heliyon*, *10*(14). https://doi.org/10.1016/j.heliyon.2024.e34226

Anderson, J., & Feder, G. (2007). Agricultural Extension. In R. Evenson & P. Pingali (Eds.), *Handbook of Agricultural Economics* (Vol. 3, pp. 2343–2378). Elsevier.

Anderson, J. R., & Feder, G. (2004). Agricultural Extension: Good Intentions and Hard Realities. *The World Bank Research Observer*, *19*(1), 41–60. https://doi.org/10.1093/wbro/lkh013

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Appadurai, A. (2004). The Capacity to Aspire: Culture and the Terms of Recognition. In *Culture and public action* (pp. 59–84). Stanford University Press.

Argaw, T. L., Fledderjohann, J., Aurino, E., & Vellakkal, S. (2023). Children's Educational Outcomes and Persistence and Severity of Household Food Insecurity in India: Longitudinal Evidence from Young Lives. *The Journal of Nutrition*, *153*(4), 1101–1110. https://doi.org/10.1016/j.tjnut.2023.02.008

Assefa, T., McCullough, E., & Berhane, G. (2025). Evaluating large-scale government investments in fertilizer adoption: The Ethiopian experience. *American Journal of Agricultural Economics*. https://doi.org/10.1111/ajae.70007

Atalell, K. A., Techane, M. A., Terefe, B., & Tamir, T. T. (2023). Mapping stunted children in Ethiopia using two decades of data between 2000 and 2019. A geospatial analysis through the Bayesian approach. *Journal of Health, Population and Nutrition*, *42*(1), 113. https://doi.org/10.1186/s41043-023-00412-3

Aurino, E., Fledderjohann, J., & Vellakkal, S. (2019). Inequalities in adolescent learning: Does the timing and persistence of food insecurity at home matter? *Economics of Education Review*, *70*, 94–108. https://doi.org/10.1016/j.econedurev.2019.03.003

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Azzarri, C., & Nico, G. (2022). Sex-disaggregated agricultural extension and weather variability in Africa south of the Sahara. *World Development*, *155*, 105897. https://doi.org/10.1016/j.worlddev.2022.105897

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/ Henry Holt & Co.

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

Barrios, S., Bertinelli, L., & Strobl, E. (2010). Trends in Rainfall and Economic Growth in Africa: a Neglected Cause of the African Growth Tragedy. *The Review of Economics and Statistics*, *92*(2), 350–366.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press.

Benjamin-Chung, J., Mertens, A., Colford, J. M., Hubbard, A. E., van der Laan, M. J., Coyle, J., Sofrygin, O., Cai, W., Nguyen, A., Pokpongkiat, N. N., Djajadi, S., Seth, A., Jilek, W., Jung, E., Chung, E. O., Rosete, S., Hejazi, N., Malenica, I., Li, H.,…The Ki Child Growth, C. (2023). Early-childhood linear growth faltering in low- and middle-income countries. *Nature*, *621*(7979), 550–557. https://doi.org/10.1038/s41586-023-06418-5

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berhane, G., Ragasa, C., Abate, G. T., & Assefa, T. W. (2018). *The state of agricultural extension services in Ethiopia and their contribution to agricultural productivity*.

Bernard, T., Dercon, S., Orkin, K., & Taffesse, A. S. (2014). The Future in Mind: Aspirations and Forward-Looking Behaviour in Rural Ethiopia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2514590

Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., Martorell, R., & Uauy, R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, *382*(9890), 427–451. https://doi.org/10.1016/S0140-6736(13)60937-X

Burgert, C. R., Colston, J., Roy, T., & Zachary, B. (2013). *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys* (DHS Spatial Analysis Reports No. 7, Issue. http://dhsprogram.com/pubs/pdf/SAR7/SAR7.pdf

Chiadmi, I., & Fockaert, L. (2026). Designing information policies for soil and water quality: Farmers' preferences and prior participation effects for advisory services in Flanders. *Journal of Rural Studies*, *127*, 104346. https://doi.org/10.1016/j.jrurstud.2026.104346

Choi, H. I. (2019). Assessment of Aggregation Frameworks for Composite Indicators in Measuring Flood Vulnerability to Climate Change. *Scientific Reports*, *9*(1), 19371. https://doi.org/10.1038/s41598-019-55994-y

Coates, J., Swindale, A., & Bilinsky, P. (2007). *Household Food Insecurity Access Scale (HFIAS) for Measurement of Food Access: Indicator Guide*. https://www.fantaproject.org/monitoring-and-evaluation/household-food-insecurity-access-scale-hfias

Cohen, M. J., & Garrett, J. L. (2010). The food price crisis and urban food (in)security. *Environment and Urbanization*, *22*(2), 467–482. https://doi.org/10.1177/0956247810380375

Conde-Agudelo, A., Rosas-Bermúdez, A., & Kafury-Goeta, A. C. (2006). Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. *Jama*, *295*(15), 1809–1823. https://doi.org/10.1001/jama.295.15.1809

Cooper, M. W., Brown, M. E., Hochrainer-Stigler, S., Pflug, G., McCallum, I., Fritz, S., Silva, J., & Zvoleff, A. (2019). Mapping the effects of drought on child stunting. *Proceedings of the National Academy of Sciences*, *116*(35), 17219–17224. https://doi.org/10.1073/pnas.1905228116

Coppock, D. L., Bailey, D., Ibrahim, M., & Tezera, S. (2018). Diversified Investments of Wealthy Ethiopian Pastoralists Include Livestock and Urban Assets That Better Manage Risk. *Rangeland Ecology & Management*, *71*(1), 138–148. https://doi.org/10.1016/j.rama.2017.05.004

CSA. (2020a). *Ethiopia Socioeconomic Survey 2013-2014*.

CSA. (2020b). *Ethiopia Socioeconomic Survey 2015-2016, Wave 3*.

CSA, & ICF. (2012). *Ethiopia Demographic and Health Survey 2011*. https://dhsprogram.com/publications/publication-fr255-dhs-final-reports.cfm

CSA, & ICF. (2017). *Ethiopia Demographic and Health Survey 2016*. http://dhsprogram.com/pubs/pdf/FR328/FR328.pdf

Cueto, S., & Muñoz, I. (2009). *Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives* (Technical Notes, Issue. https://www.younglives.org.uk/publications/psychometric-characteristics-cognitive-development-and-achievement-instruments-round-2

Davenport, F., Grace, K., Funk, C., & Shukla, S. (2017). Child health outcomes in sub-Saharan Africa: A comparison of changes in climate and socio-economic factors. *Global Environmental Change*, *46*, 72–87. https://doi.org/10.1016/j.gloenvcha.2017.04.009

Davis, K., Swanson, B., Amudavi, D., Mekonnen, D. A., Flohrs, A., Riese, J., Lamb, C., & Zerfu, E. (2010). *In-depth assessment of the public agricultural extension system of Ethiopia and recommendations for improvement*. International Food Policy Research Institute (IFPRI). https://www.ifpri.org/publication/depth-assessment-public-agricultural-extension-system-ethiopia-and-recommendations

Davis, K. E., Babu, S. C., & Ragasa, C. (2020). *Agricultural extension: Global status and performance in selected countries*. Intl Food Policy Res Inst. https://doi.org/10.2499/9780896293755

Deitchler, M., Ballard, T., Swindale, A., & Coates, J. (2010). *Validation of a measure of household hunger for cross-cultural use*. F. a. A. O. o. t. U. Nations. https://www.fao.org/agrifood-economics/publications/detail/en/c/122329/

Deitchler, M., Ballard, T., Swindale, A., & Coates, J. (2011). *Household Hunger Scale: A Simple Measure for Cross-Cultural Use*. https://coilink.org/20.500.12592/rzckbw

Dercon, S. (2004). Growth and shocks: evidence from rural Ethiopia. In *Macroeconomic policies and poverty* (pp. 308–329). Routledge.

Dercon, S., & Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics*, *96*(2), 159–173. https://doi.org/10.1016/j.jdeveco.2010.08.003

Dercon, S., & Singh, A. (2013). From Nutrition to Aspirations and Self-Efficacy: Gender Bias over Time among Children in Four Countries. *World Development*, *45*, 31–50. https://doi.org/10.1016/j.worlddev.2012.12.001

Desta, G., Legesse, G., Agegnehu, G., Tigabie, A., Nagaraji, S., Gashaw, T., Degefu, T., Ayalew, B., Addis, A., Getachew, T., Managido, D., Bazie, Z., Abathun, T., Abera, A., Dache, A., Adissie, S., Sebnie, W., Feyisa, T., Yakob, G.,…Harawa, R. (2023). Landscape-based nutrient application in wheat and teff mixed farming systems of Ethiopia: farmer and extension agent demand driven approach. *Frontiers in Sustainable Food Systems*, *7*. https://doi.org/10.3389/fsufs.2023.1241850

Development Initiatives. (2021). *2021 Global Nutrition Report: The state of global nutrition*. D. I. P. R. Ltd. https://www.un.org/nutrition/sites/www.un.org.nutrition/files/global_nutrition_report_2021.pdf

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test--Fourth Edition (PPVT-4)*.

EPHI, & ICF. (2021). *Ethiopia Mini Demographic and Health Survey 2019: Final Report*. https://dhsprogram.com/pubs/pdf/FR363/FR363.pdf

Erkossa, T., Laekemariam, F., Abera, W., & Tamene, L. (2022). Evolution of soil fertility research and development in Ethiopia: From reconnaissance to data-mining approaches. *Experimental Agriculture*, *58*, e4, Article e4. https://doi.org/10.1017/S0014479721000235

ESS, & ICF. (2025). *Ethiopia Demographic and Health Survey 2024–25 Key Indicators Report (English)*. https://dhsprogram.com/publications/publication-PR163-Preliminary-Reports-Key-Indicators-Reports.cfm

FAO. (1996). *Rome Declaration on World Food Security and World Food Summit Plan of Action*. Rome: Food and Agriculture Organization of the United Nations Retrieved from https://www.fao.org/4/w3613e/w3613e00.htm

FAO. (2000). *The elimination of food insecurity in the Horn of Africa*. https://www.fao.org/4/x8406e/x8406e00.htm

FAO, IFAD, UNICEF, WFP, & WHO. (2024). *The State of Food Security and Nutrition in the World 2024*. https://openknowledge.fao.org/server/api/core/bitstreams/39dbc6d1-58eb-4aac-bd8a-47a8a2c07c67/content/state-food-security-and-nutrition-2024/global-nutrition-targets-trends.html

FAO, & IIASA. (2023). *Harmonized World Soil Database Version 2.0* (F. I. I. f. A. S. A. (IIASA), Ed.). https://doi.org/10.4060/cc3823en

Favara, M. (2017). Do Dreams Come True? Aspirations and Educational Attainments of Ethiopian Boys and Girls. *Journal of African Economies*, *26*(5), 561–583. https://doi.org/10.1093/jae/ejx018

FDRE. (2011a). *Climate-resilient green economy strategy*. https://www.preventionweb.net/publication/ethiopia-climate-resilient-green-economy-strategy

FDRE. (2011b). *Ethiopia's Climate-Resilient Green Economy Strategy*. Retrieved from https://faolex.fao.org/docs/pdf/eth186137.pdf

FDRE Ministry of Agriculture. (2020). *Productive Safety Net Programme: General Programme Implementation Manual*. Addis Ababa, Ethiopia: Ministry of Agriculture, Federal Democratic Republic of Ethiopia, Retrieved from https://capacity4dev.europa.eu/media/111774/download/7a71b028-2624-4f28-aaa6-f4806681318a_en

FDRE Ministry of Health. (2015). *Seqota Declaration*. Retrieved from https://www.moh.gov.et/en/initiatives-4-col/Seqota_Declaration?language_content_entity=en

FDRE Ministry of Health. (2024). *Resource Mobilization Plan for the Seqota Declaration Expansion and Scale-up Phases*. https://r4d.org/resources/seqota-declaration-resource-mobilization-plan/

Finaret, A. B., & Hutchinson, M. (2018). Missingness of Height Data from the Demographic and Health Surveys in Africa between 1991 and 2016 Was Not Random but Is Unlikely to Have Major Implications for Biases in Estimating Stunting Prevalence or the Determinants of Child Height. *The Journal of Nutrition*, *148*(5), 781–789. https://doi.org/10.1093/jn/nxy037

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066. https://doi.org/10.1038/sdata.2015.66

Gasparrini, A., Vicedo-Cabrera, A. M., & Tobias, A. (2024). The Multi-Country Multi-City Collaborative Research Network: An international research consortium investigating environment, climate, and health. *Environmental Epidemiology*, *8*(5), e339. https://doi.org/10.1097/ee9.0000000000000339

Gatdet, C. (2022). The Ethiopian agricultural extension services: A mixed perspective. *Cogent Food & Agriculture*, *8*(1), 2132848. https://doi.org/10.1080/23311932.2022.2132848

Gebremedhin, B., Hoekstra, D., & Tegegne, A. (2006). *Commercialization of Ethiopian agriculture: Extension service from input supplier to knowledge broker and facilitator*. International Livestock Research Institute.

Gebremeskel Haile, G., Tang, Q., Sun, S., Huang, Z., Zhang, X., & Liu, X. (2019). Droughts in East Africa: Causes, impacts and resilience. *Earth-Science Reviews*, *193*, 146–161. https://doi.org/10.1016/j.earscirev.2019.04.015

Genicot, G., & Ray, D. (2017). Aspirations and Inequality. *Econometrica*, *85*(2), 489–519. https://doi.org/10.3982/ecta13865

GNAFC. (2024). *Food Crisis Trends: Global Report on Food Crises 2024*. Global Network Against Food Crises. https://www.fightfoodcrises.net/country/ethiopia

Government of Ethiopia. (2016). Ethiopia Humanitarian Requirements Document 2016. In. Addis Ababa: National Disaster Risk Management Commission.

Grace, K., Davenport, F., Funk, C., & Lerner, A. M. (2012). Child malnutrition and climate in Sub-Saharan Africa: An analysis of recent trends in Kenya. *Applied Geography*, *35*(1), 405–413. https://doi.org/10.1016/j.apgeog.2012.06.017

Hagos, S., Lunde, T., Mariam, D. H., Woldehanna, T., & Lindtjørn, B. (2014). Climate change, crop production and child under nutrition in Ethiopia; a longitudinal panel study. *BMC Public Health*, *14*(1), 884. https://doi.org/10.1186/1471-2458-14-884

Hailu, B. A., Bogale, G. G., & Beyene, J. (2020). Spatial heterogeneity and factors influencing stunting and severe stunting among under-5 children in Ethiopia: spatial and multilevel analysis. *Scientific Reports*, *10*(1), 16427. https://doi.org/10.1038/s41598-020-73572-5

Headey, D. (2026). The state of global food insecurity measurement: A mix of methods, and a mix of messages. *Food Policy*, *142*, 103121. https://doi.org/10.1016/j.foodpol.2026.103121

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, *312*(5782), 1900–1902. https://doi.org/10.1126/science.1128898

Hoddinott, J., Maluccio, J. A., Behrman, J. R., Flores, R., & Martorell, R. (2008). Effect of a nutrition intervention during early childhood on economic productivity in

Guatemalan adults. *The Lancet*, *371*(9610), 411–416. https://doi.org/10.1016/S0140-6736(08)60205-6

Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. *The American Statistician*, *55*(1), 19–24. https://doi.org/10.1198/000313001300339897

ICF. (2018). *Demographic and Health Surveys Standard Recode Manual for DHS7*. T. D. a. H. S. Program. https://dhsprogram.com/publications/publication-dhsg4-dhs-questionnaires-and-manuals.cfm

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. https://doi.org/10.1037/a0020761

IPCC. (2014a). *Climate Change 2014 Impacts, Adaptation, and Vulnerability*. Cambridge University Press. https://doi.org/10.1017/cbo9781107415379

IPCC. (2014b). *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. C. U. Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-FrontMatterA_FINAL.pdf

IPCC. (2023). *Climate Change 2022 – Impacts, Adaptation and Vulnerability* (C. U. Press, Ed.). Intergovernmental Panel on Climate Change. https://doi.org/10.1017/9781009325844

Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, *30*(1), 175–193. https://doi.org/10.1007/BF02418571

Just, R. E., & Pope, R. D. (1979). Production Function Estimation and Related Risk Considerations. *American Journal of Agricultural Economics*, *61*(2), 276–284. https://doi.org/10.2307/1239732

Kalhoff, H., Sinningen, K., Belgardt, A., Kersting, M., & Luecke, T. (2023). Climate change and fluid status in children: early education as one response to an emerging public health problem. *Public Health Nutrition*, *26*(12), 2891–2894. https://doi.org/10.1017/s1368980023002562

Kalimba, U. B., & Culas, R. J. (2020). Climate change and farmers' adaptation: Extension and capacity building of smallholder farmers in Sub-Saharan Africa. *Global climate change and environmental policy: Agriculture perspectives*, 379–410.

Karlsson, O., Kim, R., Moloney, G. M., Hasman, A., & Subramanian, S. V. (2023). Patterns in child stunting by age: A cross-sectional study of 94 low- and middle-income countries. *Maternal & Child Nutrition*, *19*(4), e13537. https://doi.org/10.1111/mcn.13537

Kassie, B. T., Asseng, S., Rotter, R. P., Hengsdijk, H., Ruane, A. C., & Van Ittersum, M. K. (2015). Exploring climate change impacts and adaptation options for maize production in the Central Rift Valley of Ethiopia using different climate change scenarios and crop models. *Climatic Change*, *129*(1-2), 145–158. https://doi.org/10.1007/s10584-014-1322-x

Klapka, C. S., Barbosa, B. B., Magalhães, A. R., Pereira, E. B., Schettino, J. P., Domingos, A. L., Carioca, A. A. F., Lourenço, B. H., Garcia, L. M. T., Jacob, M. C. M., López-Olmedo, N., Gomes, S. M., Albuquerque, U. P., & de Carvalho, A. M. (2026). Exploring the Effects of Climate Change on Child Malnutrition: A Scoping Review.

Journal of Human Nutrition and Dietetics, *39*(2), e70220. https://doi.org/10.1111/jhn.70220

Kramer, M. S., & Kakuma, R. (2012). Optimal duration of exclusive breastfeeding. *Cochrane Database of Systematic Reviews*(8). https://doi.org/10.1002/14651858.CD003517.pub2

Krishnan, P., & Patnam, M. (2014). Neighbors and extension agents in Ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, *96*(1), 308–327. https://doi.org/10.1111/ajae.70007

Leroy, J. L., Ruel, M., Habicht, J.-P., & Frongillo, E. A. (2014). Linear Growth Deficit Continues to Accumulate beyond the First 1000 Days in Low- and Middle-Income Countries: Global Evidence from 51 National Surveys. *The Journal of Nutrition*, *144*(9), 1460–1466. https://doi.org/10.3945/jn.114.191981

Leseba, N., Vermaak, K., Makatjane, T., & Lebuso, M. (2025). A multilevel analysis of factors associated with stunting among children under five years in Lesotho: a study of the lesotho multiple cluster indicator survey 2018. *Journal of Health, Population and Nutrition*, *44*(1). https://doi.org/10.1186/s41043-025-00901-7

Lieber, M., Chin-Hong, P., Kelly, K., Dandu, M., & Weiser, S. D. (2022). A systematic review and meta-analysis assessing the impact of droughts, flooding, and climate variability on malnutrition. *Global Public Health*, *17*(1), 68–82. https://doi.org/10.1080/17441692.2020.1860247

Mahapatra, B., Walia, M., Rao, C. A. R., Raju, B. M. K., & Saggurti, N. (2021). Vulnerability of agriculture to climate change increases the risk of child malnutrition: Evidence from a large-scale observational study in India. *PLoS One*, *16*(6), e0253637. https://doi.org/10.1371/journal.pone.0253637

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, *341*(6149), 976–980. https://doi.org/10.1126/science.1238041

Mesfin, H., Tessema, Y. M., Tirivayi, N., & Nillesen, E. (2022). The impact of agricultural extension service on the uptake of various agricultural technologies in Ethiopia. *Africa Development/Afrique et Développement*, *47*(4), 77–106.

Mincer, J. (1974). Schooling, experience, and earnings. Human behavior & social institutions no. 2.

Mistry, R. S., Benner, A. D., Tan, C. S., & Kim, S. Y. (2009). Family economic stress and academic well-being among Chinese-American youth: The influence of adolescents' perceptions of economic strain. *Journal of family psychology*, *23*(3), 279. https://doi.org/10.1037/a0015403

MoANR. (2017). *Agricultural Extension Strategy of Ethiopia*. FDRE Ministry of Agriculture and Natural Resources

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why Having Too Little Means So Much*. Times Books.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

O'Donnell, A. W., Redmond, G., Zhou, H., Skattebol, J., Wang, J. J. J., Reynolds, K. J., & MacDougall, C. (2025). Adolescents' educational aspirations and expectations: the interaction between school experiences, region, and financial disadvantage. *Social Psychology of Education*, *28*(1), 180. https://doi.org/10.1007/s11218-025-10099-5

OCHA. (2012). *Humanitarian Requirements 2012: Joint Government and Humanitarian Partners' Document*. https://reliefweb.int/report/ethiopia/humanitarian-requirements-2012-joint-government-and-humanitarian-partners-document

OCHA. (2017). Ethiopia Weekly Humanitarian Bulletin, 31 July 2017. In. New York: UN Office for the Coordination of Humanitarian Affairs.

Outes-Leon, I., & Sanchez, A. (2008). *An assessment of the Young Lives sampling approach in Ethiopia*. https://ora.ox.ac.uk/objects/uuid:1cc6b2df-3bf4-4aa9-8abd-f65d9cc1665d

Pankhurst, A., Woldehanna, T., Araya, M., Tafere, Y., Rossiter, J., Tiumelissan, A., & Birhanu, K. (2018). *Young Lives Ethiopia: Lessons from longitudinal research with the children of the millennium*.

Phalkey, R. K., Aranda-Jan, C., Marx, S., Höfle, B., & Sauerborn, R. (2015). Systematic review of current efforts to quantify the impacts of climate change on undernutrition. *Proceedings of the National Academy of Sciences*, *112*(33), E4522–E4529. https://doi.org/10.1073/pnas.1409769112

Philip, S., Kew, S. F., Jan Van Oldenborgh, G., Otto, F., O'Keefe, S., Haustein, K., King, A., Zegeye, A., Eshetu, Z., Hailemariam, K., Singh, R., Jjemba, E., Funk, C., & Cullen, H. (2018). Attribution Analysis of the Ethiopian Drought of 2015. *Journal of Climate*, *31*(6), 2465–2486. https://doi.org/10.1175/jcli-d-17-0274.1

Prado, E. L., & Dewey, K. G. (2014). Nutrition and brain development in early life. *Nutrition Reviews*, *72*(4), 267–284. https://doi.org/10.1111/nure.12102

R Core Team. (2026). *R: A Language and Environment for Statistical Computing*. In (Version 4.6.1) R Foundation for Statistical Computing. https://www.r-project.org/

Ranasinghe, R., Ruane, A. C., Vautard, R., Arnell, N., Coppola, E., Cruz, F. A., Dessai, S., Islam, A. S., Rahimi, M., Ruiz Carrascal, D., Sillmann, J., Sylla, M. B., Tebaldi, C., Wang, W., & Zaaboul, R. (2021). Climate Change Information for Regional Impact and for Risk Assessment. In V. Masson-Delmotte (Ed.), *Climate Change 2021: The Physical Science Basis* (pp. 1767–1926). Cambridge University Press. https://doi.org/10.1017/9781009157896.014

Randell, H., Grace, K., & Bakhtsiyarava, M. (2021). Climatic conditions and infant care: implications for child nutrition in rural Ethiopia. *Population and Environment*, *42*(4), 524–552. https://doi.org/10.1007/s11111-020-00373-3

Randell, H., Gray, C., & Grace, K. (2020). Stunted from the start: Early life weather conditions and child undernutrition in Ethiopia. *Social Science & Medicine*, *261*, 113234. https://doi.org/10.1016/j.socscimed.2020.113234

Ray, D. (2006). Aspirations, poverty, and economic change. *Understanding poverty*, *1*(2006), 409–421.

Ray, D., Banerjee, A. V., Bénabou, R., & Mookherjee, D. (2006). Aspirations, Poverty, and Economic Change. In *Understanding poverty* (pp. 0). Oxford University Press. https://doi.org/10.1093/0195305191.003.0028

ReliefWeb. (2012). *Horn of Africa Crisis: 2011-2012*. https://reliefweb.int/disaster/dr-2011-000029-ken

Rutstein, S. O. (2005). Effects of preceding birth intervals on neonatal, infant and under-five years mortality and nutritional status in developing countries: evidence from the demographic and health surveys. *International Journal of Gynecology & Obstetrics*, *89*(S1), S7–S24. https://doi.org/10.1016/j.ijgo.2004.11.012

Sabates-Wheeler, R., Hirvonen, K., Lind, J., & Hoddinott, J. (2022). Expanding Social Protection Coverage with Humanitarian Aid: Lessons on Targeting and Transfer Values from Ethiopia. *The Journal of Development Studies*, *58*(10), 1981–2000. https://doi.org/10.1080/00220388.2022.2096443

SADC, & UNICEF. (2025). *Climate change and child malnutrition: A double threat to resilience in Southern Africa*. https://www.sadc.int/document/climate-change-and-child-malnutrition-double-threat-resilience-southern-africa

Sahiledengle, B., Mwanri, L., Blumenberg, C., & Agho, K. E. (2023). Gender-specific disaggregated analysis of childhood undernutrition in Ethiopia: evidence from 2000–2016 nationwide survey. *BMC Public Health*, *23*(1), 2040. https://doi.org/10.1186/s12889-023-16907-x

Sanchez, A., & Outes-Leon, I. (2008). *An Assessment of the Young Lives Sampling Approach in Ethiopia* (Technical Notes, Issue.

Sasakawa Africa Association. (n.d.). *Ethiopia*. https://saa-safe.org/www/ethiopia.html

Save the Childern. (2023). *Twice Threatened: Nearly Half of All Stunted Children Live in Countries Most Vulnerable to Climate Change*. https://www.savethechildren.net/news/twice-threatened-nearly-half-all-stunted-children-live-countries-most-vulnerable-climate-change

Schmidt, C. W. (2019). The Future of Stunting: Potential Scenarios under Climate Change. *Environmental Health Perspectives*, *127*(4), 044002. https://doi.org/10.1289/ehp5049

Schultz, T. W. (1961). Investment in human capital. *The American economic review*, *51*(1), 1–17.

Schultz, T. W. (1975). The value of the ability to deal with disequilibria. *Journal of Economic Literature*, *13*(3), 827–846.

Schultz, T. W. (1981). *Investing in people: The economics of population quality*. Univ of California Press.

Seidler, V., Utazi, E. C., Finaret, A. B., Luckeneder, S., Zens, G., Bodarenko, M., Smith, A. W., Bradley, S. E. K., Tatem, A. J., & Webb, P. (2025). Subnational variations in the quality of household survey data in sub-Saharan Africa. *Nature Communications*, *16*(1), 3771. https://doi.org/10.1038/s41467-025-58776-5

Sen, A. (1985). Commodities and capabilities. *North-Holland*.

Sen, A. (1999). Development as freedom. *Oxford University Press*.

Serneels, P., & Dercon, S. (2021). Aspirations, Poverty, and Education. Evidence from India. *The Journal of Development Studies*, *57*(1), 163–183. https://doi.org/10.1080/00220388.2020.1806242

Sharma, J., & Ravindranath, N. H. (2019). Applying IPCC 2014 framework for hazard-specific vulnerability assessment under climate change. *Environmental Research Communications*, *1*(5), 051004. https://doi.org/10.1088/2515-7620/ab24ed

Sheffield, P. E., & Landrigan, P. J. (2011). Global Climate Change and Children's Health: Threats and Strategies for Prevention. *Environmental Health Perspectives*, *119*(3), 291–298. https://doi.org/10.1289/ehp.1002233

Sinore, T., & Wang, F. (2024). Impact of climate change on agriculture and adaptation strategies in Ethiopia: A meta-analysis. *Heliyon*, *10*(4). https://doi.org/10.1016/j.heliyon.2024.e26103

Striessnig, E., Lutz, W., & Patt, A. G. (2013). Effects of Educational Attainment on Climate Risk Vulnerability. *Ecology and Society*, *18*(1).

Sunil Rajkumar, A., Gaukler, C., & Tilahun, J. (2012). *Combating Malnutrition in Ethiopia - An Evidence Based Approach for Sustained results*. https://documents1.worldbank.org/curated/en/456711468024616593/pdf/660480PUB0EPI10mbating0malnutrition.pdf

Tacoli, C., Bukhari, B., & Fisher, S. (2013). *Urban poverty, food security and climate change* (9781843699149). https://www.iied.org/10623iied

Tafere, Y. (2014). *Education aspirations and barriers to achievement for young people in Ethiopia*. Y. Lives. https://www.younglives-ethiopia.org/sites/default/files/YL-WP120_Tafere_Education%2520Aspirations%2520and%2520Barriers%2520to%2520Achievement.pdf

Tello, B., Rivadeneira, M. F., Moncayo, A. L., Buitrón, J., Astudillo, F., Estrella, A., & Torres, A. L. (2022). Breastfeeding, feeding practices and stunting in indigenous Ecuadorians under 2 years of age. *International Breastfeeding Journal*, *17*(1). https://doi.org/10.1186/s13006-022-00461-0

Thiede, B. C., & Strube, J. (2020). Climate variability and child nutrition: Findings from sub-Saharan Africa. *Global Environmental Change*, *65*, 102192. https://doi.org/10.1016/j.gloenvcha.2020.102192

Thornton, P. K., Ericksen, P. J., Herrero, M., & Challinor, A. J. (2014). Climate variability and vulnerability to climate change: a review. *Global Change Biology*, *20*(11), 3313–3328. https://doi.org/10.1111/gcb.12581

Tirado, M. C., Crahay, P., Mahy, L., Zanev, C., Neira, M., Msangi, S., Brown, R., Scaramella, C., Coitinho, D. C., & Müller, A. (2013). Climate Change and Nutrition: Creating a Climate for Nutrition Security. *Food and Nutrition Bulletin*, *34*(4), 533–547. https://doi.org/10.1177/156482651303400415

Trisos, C. H., Adelekan, I. O., Totin, E., Ayanlade, A., Efitre, J., Gemeda, A., Kalaba, K., Lennard, C., Masao, C., Mgaya, Y., Ngaruiya, G., Olago, D., Simpson, N. P., & Zakieldeen, S. (2022). Africa. In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1285–1455). Cambridge University Press. https://doi.org/10.1017/9781009325844.011

UNICEF. (2025). *Global Outlook 2025: Prospects for Children – Building Resilient Systems for Children's Futures*. https://www.unicef.org/innocenti/media/10341/file/UNICEF-Innocenti-Prospects-for-Children-Global-Outlook-2025.pdf

UNICEF, WHO, & WB. (2025). *Levels and trends in child malnutrition: UNICEF / WHO / World Bank Group Joint Child Malnutrition Estimates. Key findings of the 2025 edition*. https://data.unicef.org/resources/jme/

UNICEF Ethiopia. (2025). *A Climate Landscape Analysis for Children in Ethiopia*. https://knowledge.unicef.org/child-rights/resource/climate-landscape-analysis-children-ethiopia

United Nations. (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. United Nations Retrieved from https://www.un.org/sustainabledevelopment/%20development-agenda/

United Nations General Assembly. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development* (A/RES/70/1). https://sdgs.un.org/2030agenda

van der Weide, R., Lakner, C., Mahler, D. G., Narayan, A., & Gupta, R. (2024). Intergenerational mobility around the world: A new database. *Journal of Development Economics*, *166*, 103167. https://doi.org/10.1016/j.jdeveco.2023.103167

Victora, C. G., Christian, P., Vidaletti, L. P., Gatica-Domínguez, G., Menon, P., & Black, R. E. (2021). Revisiting maternal and child undernutrition in low-income and middle-income countries: variable progress towards an unfinished agenda. *The Lancet*, *397*(10282), 1388–1399. https://doi.org/10.1016/s0140-6736(21)00394-9

Wamani, H., Åstrøm, A. N., Peterson, S., Tumwine, J. K., & Tylleskär, T. (2007). Boys are more stunted than girls in Sub-Saharan Africa: a meta-analysis of 16 demographic and health surveys. *BMC Pediatrics*, *7*(1), 17. https://doi.org/10.1186/1471-2431-7-17

WHO. (2005). *Report of a WHO technical consultation on birth spacing: Geneva, Switzerland 13–15 June 2005*. W. H. Organization.

WHO. (2006). *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. https://www.who.int/publications/i/item/924154693X

WHO. (2011). *Water, sanitation and hygiene interventions and the prevention of diarrhoea*. https://www.who.int/tools/elena/bbc/wsh-diarrhoea

WHO. (2023). *Child malnutrition: Stunting among children under 5 years of age*. https://www.who.int/data/gho/indicator-metadata-registry/imr-details/72

WHO. (2024). *Climate change and child health: Protecting the most vulnerable*. https://www.who.int/publications/i/item/9789240085350

WHO. (2025a). *Global Targets 2025: To improve maternal, infant and young child nutrition*. https://www.who.int/teams/nutrition-and-food-safety/global-targets-2025

WHO. (2025b). *Prevalence of stunting in children under 5 (%)*. World Health Organization. Retrieved June 24, 2025 from https://data.who.int/indicators/i/A5A7413/5F8A486

WHO, & UNICEF. (2021). *Progress on household drinking water, sanitation and hygiene 2000–2020: Five years into the SDGs*. https://www.who.int/publications/i/item/9789240030848

WHO, & UNICEF. (2023). *Progress on household drinking-water, sanitation and hygiene 2000-2022: Special focus on gender*. https://www.who.int/publications/m/item/progress-on-household-drinking-water--sanitation-and-hygiene-2000-2022---special-focus-on-gender

WHO Regional Office for Africa. (2018). *Regional Atlas on Adolescent and Youth Health: Monitor the Health Status and Trend of Adolescent and Youth in Afric*. https://iris.who.int/handle/10665/272365

Woldehanna, T., & Araya, M. (2016). *Educational inequalities among children and young people in Ethiopia*. https://ora.ox.ac.uk/objects/uuid:12bce8a1-6be8-49eb-a104-82756b3530e9

Woldehanna, T., Araya, M., & Pankhurst, A. (2017). *Education and Learning: Preliminary Findings from the 2016 Young Lives Survey (Round 5): Ethiopia*. https://www.younglives.org.uk/sites/default/files/ETHIOPIA-Education&Learning-Factsheet-Oct17_1.pdf

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

World Bank. (2018). *World Development Indicators: Ethiopia*. https://databank.worldbank.org/reports.aspx?source=2&country=ETH#

World Bank. (2020). *Ethiopia Poverty Assessment: Harnessing Continued Growth for Accelerated Poverty Reduction*. https://openknowledge.worldbank.org/handle/10986/33544

World Bank. (n.d.). *Nutrition*. World Bank. https://www.worldbank.org/ext/en/topic/health/nutrition

World Bank Group. (2015). *Ethiopia's Great Run: The Growth Acceleration and How to Pace It*. https://doi.org/10.1596/23333

World Bank Group. (2021). Climate Change Knowledge Portal for Policymakers and Practitioners: Ethiopia. In: World Bank.

Wossen, T., Gatiso, T. T., & Kassie, M. (2018). Estimating returns to fertilizer adoption with unobserved heterogeneity: Evidence from Ethiopia. *Food and Energy Security*, *8*(2), e00156. https://doi.org/10.1002/fes3.156

Yitayew, A., Abdulai, A., Yigezu, Y. A., Deneke, T. T., & Kassie, G. T. (2021). Impact of agricultural extension services on the adoption of improved wheat variety in Ethiopia: A cluster randomized controlled trial. *World Development*, *146*, 105605. https://doi.org/10.1016/j.worlddev.2021.105605

Young Lives. (2018). *Young Lives survey design and sampling (Round 5): Ethiopia*. https://ora.ox.ac.uk/objects/uuid:de9a6e42-6d28-4494-97bc-7372ee413f8e