Master Thesis

Crop Diversification and Climate Change Adaptation:

Empirical Evidence from Smallholders in Ethiopia.

By *Ezékiel TANO*

Supervisor: Matthew Gordon **Referee:** Kattrin Millock

15 October 2025

M2 Public Policy and Development - Paris School of Economics

Acknowledgments

This master's thesis would not have been possible without the support and inspiration of my supervisor. I am particularly grateful to my supervisor, Matthew Gordon, for his constant guidance, insightful suggestions, and careful comments throughout this work, as well as for deepening my interest in applied research. I also wish to thank Kattrin Millock for her valuable advice at the early stages of this thesis and for kindly agreeing to act as a referee.

I am also thankful to the directors of the PPD program. In particular, I wish to express my sincere appreciation to Olivier Eynde, Luc Behaghel, and Akiko Suwa-Eisenmann for their advice, support, and guidance since my first year at PSE.

I am deeply grateful to all the professors at the Paris School of Economics who generously shared their time and insights through discussions on this topic. I am also indebted to the Paris School of Economics (PSE) and the École des Hautes Études en Sciences Sociales (EHESS) administration committee members, whose scholarships and financial support made it possible for me to pursue this master's degree.

Finally, this work has greatly benefited from the helpful comments, discussions, and suggestions of my classmates. Any remaining errors or omissions are my sole responsibility.

Abstract

Abstract

Climate change poses a severe threat to agricultural production, particularly for smallholder farmers in Ethiopia, where over 80% of livelihoods depend on rain-fed systems. This thesis examines crop-specific impacts of temperature and precipitation variability and evaluates crop diversification as an adaptation strategy. Using nationally representative LSMS-ISA household–crop–year panel data (over $20,000$ observations, 2011–2022), we combine fixed-effects econometric models with linear and nonlinear climate variables and semi-parametric temperature-bin specifications (1°C increments) to identify within-household responses while controlling for unobserved heterogeneity. The analysis applies a multimargin framework—yields (intensive margin), revenues (economic margin), and crop choice (extensive margin), while modeling diversification as an endogenous behavioral response to climate stress. Results show strong heterogeneity across crops. A $+1^{\circ}$C warming increases rice yields by $\approx 47\%$, sorghum by $\approx 38\%$, and millet by $\approx 33\%$, but reduces teff and garlic yields by $\approx 8\%$ and $\approx 17\%$, respectively, once thermal thresholds are exceeded. On the revenue side, chat (khat) income declines by $\approx 14\%$ per $+1^{\circ}$C, while pepper revenues increase by $\approx 12\%$. Precipitation effects are smaller but non-negligible: a $+10\%$ rise in rainfall raises sorghum and pepper revenues by 0.5–0.6%, yet reduces chat revenues by $\approx 1\%$. Evidence from drought years shows that diversification buffers yield losses, particularly in low-rainfall zones, though effectiveness varies by agroecological context and market access. Across many crops, significant quadratic temperature terms confirm the threshold-dependent nature of climate impacts: moderate warming can be beneficial, but beyond crop-specific limits, productivity and incomes decline rapidly. Policy implications include the spatial targeting of climate-resilient crops, scaling up diversification through improved varieties and extension services, and investment in irrigation to reduce rainfall dependence. These findings directly support Ethiopia's Climate Resilient Green Economy strategy and provide a replicable framework for other smallholder-based economies confronting climate risks.

Keywords: Climate change, Crop diversification, Agricultural adaptation, Nonlinear climate effects.

JEL Codes: Q12, Q54, O13, Q16.

Contents

Abstract **ii**

List of Figures **iv**

List of Tables **v**

List of Abbreviations **vi**

1 Introduction **1**

2 Data **7**
- 2.1 Data Collection Process 7
- 2.2 Descriptive Statistics 11

3 Econometric Models and Methodology **13**
- 3.1 Intensive Margin 13
- 3.2 Economic Margin 16
- 3.3 Extensive Margin of Adaptation: Household main crop choice 17

4 Results and Discussion **19**
- 4.1 Intensive Margin 19
- 4.2 Economic Margin 27
- 4.3 Extensive Margin 28
- 4.4 Limitation and Transferability 32

5 Conclusion and Policy Implication **33**

References **ii**

Appendix **vi**
- APPENDIX A: Additionnal descriptive Statistic vi
- APPENDIX B: Formal equation and steps of Models estimation xii
- APPENDIX C: Additionnal results of the estimations xvi

List of Figures

1 Main Crops in the Scope 2

2 Temperature and precipitation in the Study Scope 7

3 Distribution of the Main variables 12

4 Temperature on Crop yield 19

5 Precipitation Effect on Crop yield 20

6 Temperature Effects on Cereals yield 22

7 Fruits yield response to climate change 23

8 Oleaginous nuts response to climate 24

9 Vegetables Yield response to climate 25

10 Oil seeds Yield response to climate change effect on crop yield 26

11 Climate Margin Effect on Households Income 28

12 Effects of temperature bins on probabilities of planting the crop 29

13 Effects of temperature bins on probabilities 31

14 Wave and presence of houshold vi

15 Distribution of the logarithm the Yield and households Incomes vii

16 Possibilities of Missing values (R_{it}, S_{it}, Q_{it}) vii

17 Cereal and Fruits Crops in the Scope ix

18 Nuts and Legums Crops in the Scope ix

19 Oil Seed Crops in The scope x

20 Distribution of the logarithme the main variables by wave x

21 Temperature Bin Effects on Cereal Yields xxi

List of Tables

1 Synthetic Literature Review Grouped by Theme, Method, and Contribution . . . 5
2 Variables and Survey Questions . . . 8
3 Summary description of the keys variables . . . 11
4 Summary statistics of the Key variables . . . 11
5 Temperature Bin Effects on Log Yield (Cereals) . . . 21
6 Effect of climate on Household revenue (Economic margin) . . . 27
7 Data Quality Assessment for Agricultural Production . . . vi
8 Matrix of correlation between climates variables . . . vii
9 Summary statistics by Wave . . . viii
10 Percentage of Zero Yields by Crop and Wave . . . viii
11 The effect of climate on the crop log(yield) . . . xvi
12 Temperature Bin Effects on Log Yield for Fruit . . . xvii
13 Oleaginius Nuts . . . xviii
14 Temperature Bin Effects on Log Yield for Vegetables . . . xix
15 Temperature Bin Effects on Log Yield for oil seeds . . . xx
16 Cereal Crop Choice (Extensive Margin) . . . xxii
17 Fruits Choice (Extensive Margin) . . . xxiii
18 Oleagineus Crop Choice (Extensive Margin) . . . xxiv
19 Vegetable/Specialty Crop Choice (Extensive Margin) . . . xxv
20 Oil seeds Crop Choice (Extensive Margin) . . . xxvi

List of Abbreviations

CSA: Climate Smart Agricultural

SSA: Sub-Saharan Africa

LSMS-ISA: Living Standards Measurement Study – Integrated Surveys on Agriculture

WB: World Bank

VPD: Vapor Pressure Deficit

CHIRPS: Climate Hazards Group InfraRed Precipitation with Station data

CRGE: Climate Resilient Green Economy

GDP: Gross Domestic Product

ESS: Ethiopian Socioeconomic Survey

MODIS: Moderate Resolution Imaging Spectroradiometer

NASA: National Aeronautics and Space Administration

GIS: Geographic Information System

ERA5: ECMWF Reanalysis 5th Generation (European Centre for Medium-Range Weather Forecasts)

FDRE: Federal Democratic Republic of Ethiopia

FAO: Food and Agriculture Organization

EIAR: Ethiopian Institute of Agricultural Research

ICRISAT: International Crops Research Institute for the Semi-Arid Tropics

CIMMYT: International Maize and Wheat Improvement Center

MoA: Ministry of Agriculture (Ethiopia)

FAOSTAT: FAO Statistical Database

1 Introduction

Agriculture remains the foundation of Ethiopia's economy, not only in terms of employment and GDP contribution but as the lifeblood of rural livelihoods, food security, and national development. The sector employs over 65% of the workforce and contributes approximately 35% to the country's GDP (World Bank, 2020). It supports the majority of the population, particularly in rural areas by providing both subsistence and market oriented production, and also anchors the socio-economic stability. Agriculture is also a key driver of poverty reduction and foreign exchange earnings, notably through exports such as coffee, pulses, and oilseeds. However, despite its central role, the sector remains constrained by persistent structural limitations: notably low productivity, limited technological adoption, poor access to inputs, and, most importantly, near-total reliance on rainfed systems. These vulnerabilities make Ethiopian agriculture extremely sensitive to climatic fluctuations. Over the last two decades, Ethiopia has witnessed a marked increase in the frequency and intensity of climate shocks. Rising temperatures, erratic rainfall, delayed rainy seasons, and prolonged droughts have become more common, disrupting the agricultural calendar and threatening food production (Funk et al., 2015; Alemayehu & Bewket, 2017).

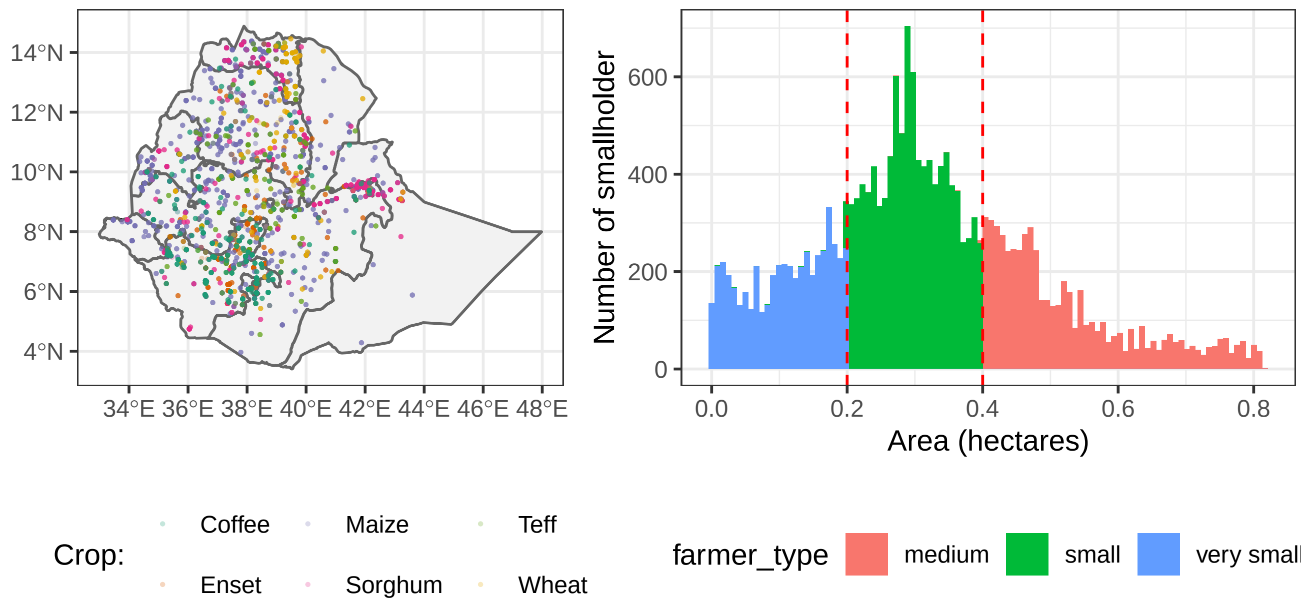
These climatic disturbances have had substantial impacts on yields, household incomes, and rural welfare. For smallholder farmers, who cultivate over 90% of the agricultural land, climate shocks often lead to irreversible damage. With limited access to financial services, irrigation, or crop insurance, these farmers are unable to buffer the risks. The consequences are not only agronomic, but deeply socio-economic: yield losses translate into food shortages, increased poverty, and migration. In this context, climate variability poses a systemic threat to Ethiopia's long-term development goals. While a growing body of literature has documented the biophysical and economic consequences of climate change on agriculture, much of this research has been conducted in high or middle-income settings and tends to rely on aggregated data. Landmark studies such as Schlenker and Roberts (2009) and Lobell et al. (2011) have demonstrated that temperature increases have nonlinear effects on yields, with modest warming sometimes beneficial, but significant losses occurring beyond certain critical thresholds.

These findings shed light on climate yield relationships but are often based on national or regional data that obscure household-level realities. Such aggregation overlooks the heterogeneity of smallholder farming systems, shaped by agroecological diversity, behavioral responses, and resource constraints. Sub-Saharan Africa—and Ethiopia in particular, remains underrepresented in empirical studies despite being among the most climate-vulnerable regions globally. Ethiopia's agriculture is highly diverse: cereals (*teff, maize, barley, sorghum*), legumes (*beans, peas*), fruits (*avocado, mango*), vegetables (*onions, kale, tomatoes*), and industrial crops (*coffee, sesame, tobacco*) are cultivated across contrasting highland and lowland zones. Each crop has distinct environmental requirements and sensitivities to heat, water stress, and seasonality. Yet most studies treat agriculture as homogeneous, masking crop- and region-specific vulnerabilities. This limits the effectiveness of adaptation strategies and highlights the urgent need for disaggregated, crop-specific, and localized evidence on climate responses.

Study Context, Data and Scope

This study relies on a nationally representative panel dataset that captures the complexity and heterogeneity of smallholder agriculture in Ethiopia. The primary data source is the Ethiopian Socioeconomic Survey (ESS), conducted in collaboration with the Central Statistical Agency of Ethiopia and the World Bank's Living Standards Measurement Study–Integrated Surveys on Agriculture (LSMS-ISA). Spanning multiple waves and including over 20,000 crop-year observations, the ESS provides a broad temporal window to analyze both short-term climatic shocks and long-term trends in climate-agriculture interactions. At both the household and plot levels, the ESS offers highly detailed information, including household demographics, agricultural practices, plot-specific crop choices, land area, yields, output volumes, and revenues. This granularity enables a nuanced investigation of farmer behavior and crop performance across Ethiopia's diverse agroecological zones. The panel structure allows for the application of fixed-effects models to control for time-invariant unobserved household characteristics, thereby strengthening causal inference. To assess climate exposure, household-level survey data are matched with spatially explicit, high-resolution weather datasets. Daily temperature data are sourced from ERA5, while precipitation data come from CHIRPS. Additionally, the Copernicus ERA5-Land dataset provides gridded daily weather data at a spatial resolution of 0.1° (approx. 10 km). Each household's coordinates are matched to the nearest grid cell, enabling the construction of key climate exposure variables at the household-year level.

Figure 1: Main Crops in the Scope



The scope of the analysis is broad, covering more than 15 major crop types commonly grown in Ethiopia. These are categorized into cereals (e.g., teff, maize, sorghum, barley, rice), pulses (e.g., beans, peas), fruits (e.g., avocado, mango), vegetables (e.g., kale, onion, tomato), and industrial/oilseed crops (e.g., coffee, sesame). This classification allows for crop-specific evaluation of climate sensitivities, uncovering which crops are most vulnerable or resilient to changes in temperature and rainfall. Such disaggregation is rare in the climate-agriculture literature and constitutes a major empirical strength of this study. It transcends generalized statements and provides behaviorally realistic and policy-relevant insights into how different crops and household contexts respond to climatic variation.

Positioning within the Literature

Climate change is widely recognized as one of the most pressing threats to agriculture, with particularly severe implications in Sub-Saharan Africa where rainfed systems dominate and adaptive capacity is limited (FAO, 2016). Ethiopia exemplifies this vulnerability: agriculture employs more than 70% of the population and contributes around 35% of GDP (World Bank, 2022), while over 80% of cultivated land is managed by smallholders operating under structural constraints such as insecure tenure, limited irrigation, and poor access to markets and credit (Deressa et al., 2008; Bewket, 2009). These conditions heighten farmers' sensitivity to both gradual warming and rainfall variability, making Ethiopia a critical setting for studying micro-level climate adaptation.

The academic literature consistently shows that climatic impacts on agriculture are rarely linear. Seminal studies such as Schlenker & Roberts (2009) and Lobell et al. (2011) demonstrate sharp yield declines beyond critical temperature thresholds, with damages accelerating at higher levels of exposure. African studies reinforce this evidence and highlight the importance of rainfall timing and distribution, not just aggregate totals, in shaping productivity (Alemayehu & Bewket, 2017; Araya & Stroosnijder, 2010). For Ethiopia specifically, delayed onset of the rainy season reduces maize and teff yields by up to 18%, while truncated rainfall periods heighten vulnerability for shallow-rooted crops such as teff. Such findings underscore the need for crop-specific, threshold-based models rather than aggregate estimates.

Beyond yield impacts, the literature emphasizes two complementary behavioral adaptation strategies at the household level: crop switching and crop diversification. Crop switching, or the extensive margin of adaptation, refers to farmers' reallocation of land toward crops more suited to prevailing climatic conditions. Seo & Mendelsohn (2008) provide foundational evidence, showing that in semi-arid Africa, a 1°C rise increases the probability of choosing sorghum over maize by 32% and millet by 15%. Similar dynamics are observed in Ethiopia: negative rainfall shocks decrease maize cultivation but increase sorghum adoption (Asfaw et al., 2021), while temperature stress reduces teff viability and shifts choices toward millet (Diro et al., 2022). These studies illustrate that crop choice is not static but an endogenous behavioral response to climate variability, shaped by experiential knowledge and household-level optimization.

Crop diversification, in turn, has long been theorized as a cornerstone of resilience in rainfed systems. Defined as cultivating multiple crop species either spatially (e.g., intercropping, polyculture) or temporally (e.g., crop rotation), diversification operates through both economic and ecological channels. From a risk-management perspective, it reduces production and income variance through the "portfolio effect" (Tilman et al., 2002). From an agroecological perspective, diversified systems enhance soil fertility, improve water retention, and buffer pest and disease pressures (Lin, 2011; Altieri et al., 2015). Empirically, Di Falco et al. (2011) find that a 10% increase in crop diversity in Ethiopia is associated with a 4.5% reduction in yield variability, and diversified households experienced 15–20% lower yield losses during droughts. However, diversification also carries potential trade-offs. Barrett et al. (2001) argue that excessive or distress-driven diversification may dilute managerial attention, increase labor costs, or trap households in low-return systems. Thus, its effectiveness is context-dependent, varying by agroecological conditions, household assets, and institutional support.

Despite these advances, important gaps remain. Many studies aggregate crops into broad categories, overlooking crop-specific nonlinearities and threshold responses. Others treat diversification as an exogenous characteristic rather than an endogenous adaptation to climate shocks. Moreover, much of the literature focuses narrowly on yields, neglecting welfare outcomes such as income stability or food security. Ethiopia, with its extraordinary agroecological diversity and policy emphasis on Climate-Smart Agriculture under the CRGE strategy (FDRE, 2011), remains underrepresented in studies that integrate yields, revenues, and crop choice within a unified empirical framework. Addressing these gaps requires household-level panel analysis that explicitly models nonlinear climate responses and captures both intensive and extensive margins of adaptation.

Research Problem, Objectives, and Contributions

Despite significant efforts by the Ethiopian government to mainstream climate resilience—particularly through the Climate Resilient Green Economy (CRGE) strategy (FDRE, 2011), substantial challenges persist in translating these ambitions into evidence-based agricultural planning. Climate-smart agriculture (CSA), a central pillar of CRGE, critically depends on a nuanced understanding of how different crop types and farming systems respond to climatic stressors. In the absence of such disaggregated and empirical knowledge, CSA interventions risk being overly generic, misaligned with local agronomic realities, and less effective in building resilience.

The core research problem addressed by this thesis stems from this disconnect between high-level policy frameworks and context-specific empirical evidence. While the global literature widely documents the importance of climate impacts on agriculture, most studies either focus on national or regional averages, or aggregate outcomes across crop categories (e.g., cereals). Such approaches obscure the nonlinear, threshold-based responses that exist at the crop level and fail to capture the complex adaptive behavior of smallholders. This analytical vacuum is particularly problematic in Ethiopia, where farming systems are highly diverse, and where climate shocks disproportionately affect food security, income stability, and rural livelihoods.

This thesis is guided by the fundamental question: *How do variations in temperature and precipitation affect agricultural outcomes across different crop types in Ethiopia, and at what nonlinear thresholds do these effects become significantly beneficial or harmful?* In addressing this question, the study explores three complementary dimensions: (1) identifying crops that are relatively resilient or vulnerable to climatic changes; (2) empirically establishing climate thresholds beyond which yield and income responses change in magnitude or direction; and (3) informing spatially differentiated CSA interventions tailored to Ethiopia's agroecological zones. By doing so, the research aims to bridge the gap between scientific evidence and policy design, ensuring that adaptation strategies reflect both climatic realities and local farming systems.

The contributions of this thesis are fourfold. Theoretically, it challenges the assumption of linear climate–agriculture relationships by integrating nonlinear climate response functions—quadratic specifications and temperature-bin models (Schlenker & Roberts, 2009); Lobell et al., 2011) into household-level econometric models. Empirically, it extends these frameworks to the under-studied Ethiopian context, covering more than 15 crops across multiple waves of nationally representative LSMS-ISA data (2011–2022, over 20,000 household–crop–year observations), matched with high-resolution climate datasets (CHIRPS, TerraClimate). This enables the identification of crop-specific nonlinear thresholds and highlights novel findings, such as the role of cold stress in highland cereals like rice and millet—an area largely neglected in African agricultural research.

Methodologically, the study introduces a multi-model empirical strategy that enhances robustness and interpretability. It combines linear-quadratic regressions, temperature-bin models, and crop–climate interaction terms to triangulate evidence of climate impacts. Household fixed effects control for time-invariant heterogeneity (e.g., soil type, elevation, long-term infrastructure), while wave dummies capture national-level shocks. Clustered standard errors and diagnostic checks strengthen causal inference, aligning with best econometric practices (Dell et al., 2014); Cameron & Miller, 2015) rarely applied in African smallholder contexts. Importantly, the thesis innovates by modeling crop diversification not as an exogenous control variable (as in Bezabih & Di Falco, 2012; Asfaw et al., 2016), but as an endogenous behavioral response to climate variability, thus capturing the adaptive strategies farmers employ to manage risks.

From a policy perspective, the study provides actionable evidence for designing CSA interventions in Ethiopia. By distinguishing climate-vulnerable from resilient crops, the findings support targeted extension services, dissemination of stress-tolerant varieties, investment in irrigation and water management, and the promotion of strategic diversification as a resilience mechanism. Crucially, the results demonstrate that marginal warming gains can quickly reverse when biophysical thresholds are crossed, thereby challenging one-size-fits-all adaptation narratives. Instead, the evidence advocates for geographically and crop-specific strategies that better align with Ethiopia's agroecological diversity and socio-economic realities, contributing to both immediate adaptation and long-term climate resilience.

Structure of the Thesis

The structure of this thesis follows a logical progression from theory to evidence and implications. section 2 describes the data sources, the management process and some descriptive statistics, while section 3 outlines the econometric models and empirical strategy. section 4 presents the main results and discussions for policy implications for Ethiopia's Climate Resilient Green Economy (CRGE) strategy and related CSA interventions. Finally, section 5 concludes by summarizing the contributions, the limitations, and suggesting avenues for future research.

Table 1: Synthetic Literature Review Grouped by Theme, Method, and Contribution

Study	Theme	Objective	Method	Results	Limitations	Gap Addressed
			Climate-Yield Relationships			
Schlenker & Roberts (2009)	Climate-Yield Non-linearity: Temperature thresholds	Estimate nonlinear effects of temperature on crop yields (U.S.)	Panel data regression with degree-day bins, fixed effects	Maize yields decline 7% per °C above 29°C; strong convexity in damage function	U.S. based, not smallholder-focused	Applies nonlinear thresholds to Ethiopian crops using local data. Basis for binned temperature approach in Model 3; informs crop-specific thresholds
Lobell et al. (2011)	Climate-Yield Non-linearity: Physiological stress	Identify yield losses from short-term heat stress and VPD	Process-based crop modeling + Cross-country regressions, physiological modeling	Flowering-stage heat shocks reduce yields by 20-40%. VPD better stress indicator than temperature alone. VPD and heat stress during flowering severely reduce yield	Weak behavioral integration. Focus on biophysics over farmer behavior	Link VPD to behavioral models of crop switching in Ethiopia. Guides inclusion of phenological-stage variables in yield models.
Funk et al. (2015)	Climate Data:Precipitation monitoring	Develop high-resolution precipitation dataset for Africa	Satellite data fusion with station observations	CHIRPS improves drought monitoring. CHIRPS improves drought detection by 30% over existing products	Requires validation for micro-level studies	Uses CHIRPS for Ethiopia-specific analysis. Provides climate data foundation for Ethiopian analysis.
			Adaptation Strategies			
Di Falco et al. (2011)	Diversification benefits	Assess yield effects of plot-level diversification during drought. Test yield stabilization effects during drought	Panel FE with Shannon Index. Panel regression with Shannon Index	10%↑ diversity →4.5%↓ yield variance; 15-20% less loss for diversified farms 15-20% lower yield loss among diversified farmers	Does not model diversification endogenously. Static treatment of diversity	Models diversification as endogenous climate response
Asfaw et al. (2021)	Crop Switching	Study rainfall-driven crop switching in Ethiopia	Panel regression with rainfall anomalies. Household fixed effects	Sorghum replaces maize under dry shocks. -7.3pp maize, +8.5pp sorghum per 1σ dry shock	Focused only on crop choice	Integrates crop choice with yield and income effects. Expands to full crop portfolio with temperature effects.
Diro et al. (2022)	CSA Adoption	Examine determinants of CSA adoption in Ethiopian coffee systems	Multinomial logit model	Extension access and credit key for CSA adoption. Extension access(OR=2.3) and credit(OR=1.8) key determinant	Single commodity (bias) focus	Expands to multiple cropping systems Tests generalize ability across cropping systems.
			Methodological Approaches			

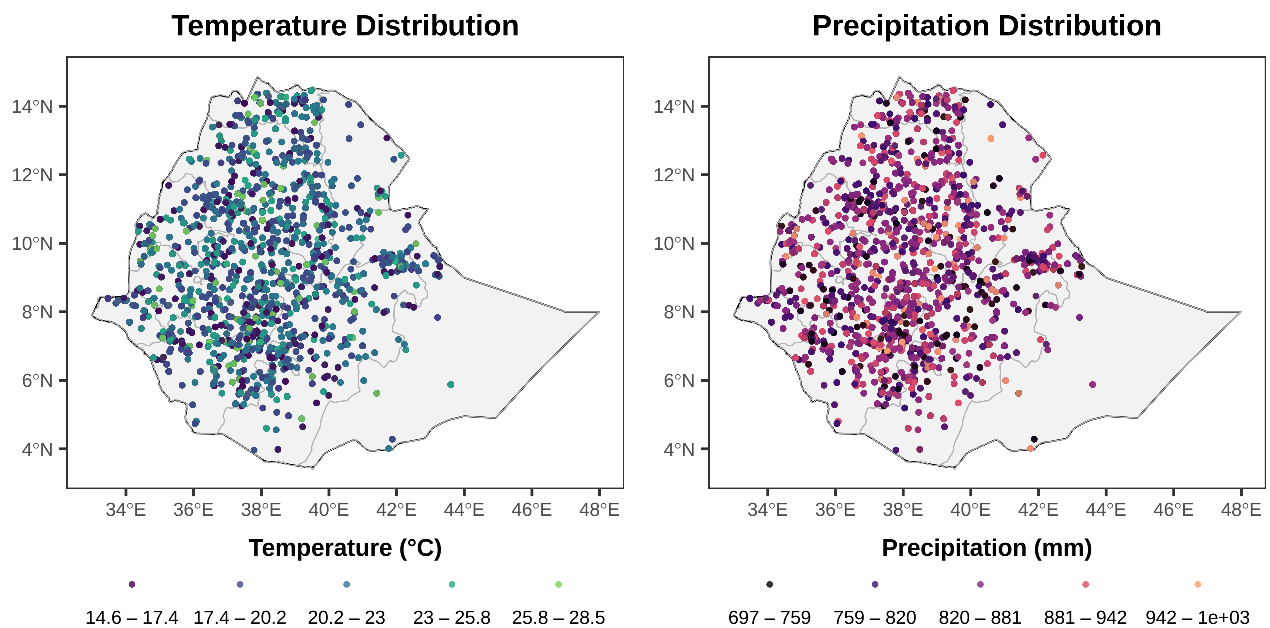
Continued on next page

Study	Theme	Objective	Method	Results	Limitations	Gap Addressed
Dell et al. (2014)	Econometrics: Causal inference	Review econometric best Meta-analysis in climate impact studies	Literature review; emphasis on causal inference. Cluster-robust SEs reduce false positives by 40%	Emphasizes FE, clustering, robust standard errors	Not empirical; general guidance	Empirically implements best practices. Implements rigorous FE+clustering in all models
Hausman & Taylor (1981)	Panel Methods: Panel endogeneity	Address correlated random effects. Address endogeneity in panel models with time-invariant variables	Instrumental variables approach	HT estimator handles correlated random effects. HT estimator reduces bias by 25-60% vs RE	Complex implementation	Simplifies for applied smallholder analysis context.
Policy Frameworks						
FDRE (2011)	National Strategy	Launch CRGE initiative. Set Ethiopia's climate-resilient development vision	Policy document analysis framework (qualitative)	CRGE emphasizes sustainable agriculture. Targets 20% emissions cut via CSA by 2030 v Implementation gaps	Not empirically grounded	Provides context for policy recommendations
FAO (2013)	CSA Guidelines	Synthesize global evidence. Provide global CSA implementation framework	Case meta-analysis	Identifies 3 pillars: productivity, adaptation, mitigation. 3 pillars:productivity (+20%), adaptation (-30% risk), mitigation	Lacks localized evidence. Context specificity lacking	Adapts principles to Ethiopian context condition
Ecological-Economic Linkages						
Tilman et al. (2002)	Agroecology	Model stability-diversity relationships. Theorize stability benefits of crop diversity	Ecological theory and modeling	Biodiversity enhances productivity and resilience. 2-species systems reduce yield variance by 35%.	Lacks microeconomic focus. Field validation needed.	Links ecological insights to farm models. Connects to household diversification
Lin (2011)	Diversification. Agroecological resilience	Quantify benefits diversification. Analyze resilience through crop diversity	Meta-analysis of field studies	Diverse systems have 20-40% lower yield variance. Polycultures maintain 80% yield during drought vs 60% monoculture	Limited economic analysis. Economic tradeoffs absent	Quantifies risk reduction in household models. Incorporates cost-benefit analysis.
Behavioral Economics						
Duflo et al. (2011)	Technology Adoption: Adoption barriers	Test nudge interventions in Kenya: Test behavioral nudges for fertilizer use	Randomized controlled trials	Small incentives boost adoption by 15-20%. Timely SMS reminders ↑ adoption by 11.7pp	Focus on inputs not systems. External validity concerns.	Adapts behavioral insights to CSA adoption
Tversky and Kahneman (1974)	Decision-Making: Risk perception	Model heuristics in risky choices. Identify decision heuristics	Laboratory experiments	Identifies systematic biases in risk perception. Loss aversion coefficient $\lambda \approx 2.25$	Artificial setting	Controls for behavioral biases. Incorporates behavioral realism in adaptation models

2 Data

The maps below (Figure 2) illustrate the spatial distribution of average temperature and precipitation across Ethiopia at the household level, highlighting the country's marked climatic heterogeneity. Temperatures range from below 15°C in the central and northern highlands, such as around Addis Ababa and the Amhara region—to above 30°C in the Rift Valley and eastern lowlands, reflecting the dominant role of altitude in shaping local agro-ecological conditions.

Figure 2: Temperature and precipitation in the Study Scope



Similarly, rainfall varies sharply across regions: annual precipitation falls below $600mm$ in arid zones like Afar and Somali, while exceeding $1,100mm$ in the western highlands of Oromia and the SNNPR, where more reliable rainfall sustains water-intensive crops such as maize, teff, and coffee. In contrast, drier regions face recurrent water stress, requiring adaptive practices such as drought-tolerant varieties or water-harvesting techniques.

These spatial gradients underscore the necessity of a disaggregated empirical framework that can capture crop-specific thresholds and nonlinear responses. Crops such as teff and maize, for instance, are highly sensitive to delayed rainfall and temperatures above 26°C, while sorghum and millet are better suited to hotter, drier environments. Conversely, rice and coffee are concentrated in wetter areas where hydrological conditions are more favorable. By linking these climatic patterns to household-level outcomes, the maps provide not only geographical context but also the empirical foundation for the study's econometric models on yield, income, and crop choice, emphasizing that climate impacts on Ethiopian agriculture can only be understood through a spatially differentiated lens.

2.1 Data Collection Process

In this study, I use micro-level panel data from the *Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA)* conducted in Ethiopia, commonly known as the LMSIA data set and developed by the World Bank. This data draws five waves of the survey (Waves 1 through 5). In fact the wave 1 covers 2011 to 2012, the wave 2 from 2013 to 2014, wave 3 from 2015 to 2016, wave 4 from 2018 to 2019 and wave 5 from 2021 to 2022. This offer repeated observations on a nationally representative sample of agricultural households. These data are particularly well-suited to analyze climate impacts because they provide detailed plot-level and household-level information, along with geospatial coordinates, which allow to integration with external environmental datasets. The panel structure enables the identification of within-household variation over time, for controlling unobserved heterogeneity in productivity drivers. From the LMSIA surveys, I extracted the following core variables: household identifiers, the major crop planted (`Crop`) [C_{it}], the quantity harvested in kilograms

(`Quantity`)[Q_{it}], and the cultivated area in square meters (`area_m2`), the total revenue from the major crop `Crop_rev`[R_{ict}], the latitude `lat`,longitude `long`. The Table 2 presents the different question asked to the smallholders during the survey:

Table 2: Variables and Survey Questions

Variable	**Description of variable**	**Survey Question**
C_{it}	Major crop of the household	*What are the major crops of this holder?*
Q_{it}	Harvested quantity from the major crop	*How much did you harvest from the major crop?*
S_{it}	Cultivated area of the field (in m²)	*Area of [Field] (Square Meters)*
R_{it}	Total income of the Household (in BIRR)	*What was the total revenues? (BIRR)*

This information is from the documentation from the description of the questionnaire of the LMS data set[a]Ethiopia-Socio-Economic Panel survey, in each wave of investigation they were addressed the same questions to the Households.

[a]https://microdata.worldbank.org/Ethiopia-Socio-Economic-Panel-Survey

To ensure accurate measurement of all variables, the dataset was cleaned and harmonized across waves. The first step concerned crop names. I reviewed all reported names, corrected spelling variations, and merged duplicates to ensure consistency. Each crop was verified to confirm its actual cultivation in Ethiopia, drawing on reference lists from FAO and CSA. This process resulted in a reduced and reliable list of crops cultivated by smallholders, and the creation of a standardized crop dictionary.
The second step focused on household yields (Y_{ict}), which required consistent measures of cultivated area (S_{it}) and harvested quantity (Q_{it}). Reported units for Q_{it} varied across households (grams, kilograms, quintals). Since over 50% of records were in kilograms and grams, all quantities were converted into kilograms for uniformity.

Before proceeding, I verified the consistency of the panel. Specifically, I checked that crops were reported in each wave, that households appeared across waves, and that area, quantity, and household income (R_{it}) were not missing. Approximately 5% of households were absent in at least one wave, while some appeared only intermittently; however, this small share does not threaten the representativeness of the sample. For the main variables (R_{it}, S_{it}, Q_{it}), missing or zero values generated eight possible cases $(2^3 = 8)$, which are detailed in Table 2. This classification guided the treatment of inconsistencies in area, output, and revenue, ensuring that the final dataset was consistent and reliable for econometric analysis.

In Figure 16 the number "1" stand for value of the correspondent variables which are not null and "0" when this value is null. But this does not represent all the cases, I added the other case in the table Table 7. This analysis evaluates the internal consistency of agricultural production data specifically area planted (`surface`), harvest quantity (`quantity`), and crop income (`revenue`) across all waves of my panel dataset. Each household's data is cross validated against these three key variables. The cleaning method adheres to principles in the literature (notably FAO[1], LSMS[2], and CSA[3]), distinguishing between plausible agricultural events (e.g., crop failure, full self-consumption) and reporting or data-entry errors (e.g., missing yield with positive income). I also check whether a household's crop name (`crop`) is missing in a given wave and impute it from their most frequently reported crop across waves if needed. This is aligned with LSMS guidelines for preserving longitudinal information.

The diffrents cases are displayed in the Table 7. In fact, *Case 1* refers to households for which all three variables crop area, quantity harvested, and income are zero or missing across all five waves. These observations indicate no agricultural activity and are thus irrelevant to a crop yield study. Since they represent roughly 6% of the data and do not contribute information, they are excluded from the final dataset.

Case 2 includes households with zero or missing land area values, but positive values for harvest and income. While this may appear inconsistent at first glance, it is common in smallholder surveys for farmers to report quantities from backyard gardens, communal plots, or mixed fields without specifying precise area. Such cases are plausible and are retained to preserve the richness of the data. *Case 3* involves households that report crop income and area planted, but show zero or missing harvest

[1]FAO. Post Harvest Loss Assessment Guidelines, 2021, FAOSTAT. Ethiopia Country Profiles (https://www.fao.org/faostat)
[2]World Bank LSMS Guide to Agricultural Household Surveys, 2021
[3]CSA Ethiopia. Agricultural Sample Survey (2012–2022),EIAR. Ethiopian Institute of Agricultural Research Reports

quantity. This situation is difficult to justify agronomically. Either the harvest quantity was not reported, or there is a mismatch between crop income and production. Since yield (income or harvest per area) cannot be computed, these cases should be flagged. If sufficient evidence supports imputation (e.g., median yield by crop), that may be applied cautiously.

Case 4 describes situations where the farmer reports having planted and harvested the crop but reports no crop income. This is a realistic scenario. It could arise due to full self-consumption, bartering, spoilage, or lack of market access. It reflects real behavior and is therefore kept in the sample, especially when analyzing food security, subsistence production, or non-monetary farming. *Case 5* includes households that report having planted land (area > 0) but neither harvested anything nor earned any income. This likely indicates total crop failure due to pests, drought, flooding, or other external shocks. As long as the crop name is provided, such cases are agriculturally plausible and should be retained for resilience analysis.

Case 6 represents the fully consistent and expected records, where all three variables are strictly positive. These cases form the benchmark and are always retained. *Case 7* shows households with harvest data but no land area or income. This may reflect informal gardening, communal farming, or imprecise measurement. Such cases are retained when plausible, especially if the quantities are small. If not, they are flagged for sensitivity analysis. *Case 8* corresponds to records where households report only income but no information on quantity or land. While income can sometimes be estimated independently (e.g., through market sales), the lack of production data makes yield computation impossible. These cases should be excluded or treated as outliers.

Case 9 captures those records where land and income are reported but harvest is missing. Again, this creates problems in computing physical yields. However, if other waves contain valid data for the same crop and household, imputation of quantity is reasonable. Otherwise, these cases are flagged for low reliability. *Case 10* arises when households report having planted and harvested but do not report income. This is quite common in rural Ethiopia, where self-consumption, traditional exchanges, or losses may lead to zero income. Such observations are essential for understanding livelihood strategies and are kept. *Case 11* captures situations where the crop name is missing even when production and economic variables are valid. To avoid information loss, the missing crop name is imputed using the most frequently reported crop by the household in other waves. This method is consistent with LSMS panel practices and helps maintain longitudinal consistency.

For the area of field, it was sometimes in square meter or in hectare. I convert all of them in hectare, before compute the variable yield[4] per hectare defined as the ratio of harvest quantity to area (converted into hectares) of each household. Now, to be sure that the yield by each crop in the database is well measured I took as reference the yield of the smallholders in Ethiopia defined by the FAO and the CSA. To ensure the reliability and comparability of yield data across waves and crops, a set of external reference values was used as benchmarks during the adjustment process.

These reference ranges, expressed as realistic yield intervals in kg/ha (minimum, maximum, and mean), were drawn from a combination of authoritative agricultural databases and institutional reports, including the Central Statistical Agency (CSA) of Ethiopia's Agricultural Sample Surveys, FAOSTAT, the Ministry of Agriculture (MoA), and specialized research institutions such as EIAR, ICRISAT, and CIMMYT. These sources provide empirical yield distributions based on field trials and national-level surveys, making them reliable proxies for the expected performance of each crop under normal agronomic conditions.

By anchoring adjustments to these external references, the methodology avoids circularity (i.e., adjusting data based solely on internal patterns) and ensures that extreme or anomalous values are corrected toward plausible agronomic norms. This is particularly critical for panel datasets where intra-household comparisons and fixed-effects modeling are applied, and where stability and realism of outcome variables like yield are necessary for sound inference. Moreover, the use of cross-institutional reference data strengthens the external validity of the analysis, enabling results to be more credibly interpreted in national and regional agricultural policy contexts.

To ensure the consistency, accuracy, and analytical validity of yield measurements across agricultural waves and crop types, a systematic correction protocol was applied. This approach was designed to address outliers, harmonize yield distributions across waves, and bring anomalous values closer to a realistic and reference-based scale without distorting the overall variability in the data. The central goal of this process was to correct raw yield observations by leveraging wave and crop specific

[4] $Y_{it}(kg/ha) = \frac{Q_{it}(kg)}{S_{it}(ha)}$

reference values (including the minimum, maximum, mean, median, and median ratio) computed from a cleaned baseline dataset. This ensures that each yield value aligns with realistic agronomic ranges observed for the same crop in the same wave while preserving household-level heterogeneity.

The adjustment follows a multi-criteria conditional logic. First, the proportional correction (Case 1): When the yield falls within the normal range and the median ratio is available from at least five observations, a scaling adjustment is applied. The yield is divided by the median_ratio, allowing for harmonization between observed yields and reference expectations across waves. This case ensures that reliable data are corrected efficiently without distortion. Second, the limited reference support (Case 2): In cases where the yield is acceptable but insufficient data exist to compute a stable median ratio, a more conservative correction is used. The yield is averaged with the reference mean for the corresponding crop and wave. This mitigates the risk of overcorrection due to small sample sizes.

Third, the values below minimum threshold (Case 3): Yields that fall below the reference minimum are partially adjusted upward toward the mean by 70% of the shortfall. This gradual correction avoids abrupt replacements while reintroducing plausibility. Fourth, the values above maximum threshold (Case 4): When the observed yield exceeds the reference maximum, it is reduced progressively by 70% of the excess. This approach preserves legitimate high yields while controlling for extreme outliers that could skew the analysis. Fifth, the bounded adjustments (Case 5): All adjusted yields are subsequently clipped to remain within the range $[Min, Max]$, ensuring that no yield falls outside of acceptable agronomic boundaries.

Therefore, I records the corrected yield values. With that the total agricultural production is recomputed using the corrected yield and the cultivated area . So that each crops has its specifics rule of correction. This adjustment strategy allows for the correction of outliers and data inconsistencies while preserving meaningful variation across households, crops, and time. The use of reference benchmarks ensures that the corrections are rooted in both statistical logic and agricultural plausibility. In the context of fixed effects models, where within-unit variation is essential, such data preparation is critical. It avoids the artificial inflation or compression of yield data that could undermine the validity of econometric results.

To reduce heteroskedasticity, normalize the right-skewed distribution, and facilitate interpretation in terms of elasticities, this yield measure was log-transformed into `log_Yield`, following standard practices in the agricultural economics literature (Lobell et al., 2011; Guiteras, 2009) . But I take $\log(1 + x)$ to take into account for the yield and revenue Which are null. I do the same for the revenue of the household. Then to characterize the climatic conditions experienced by each household in each wave, I matched weather data especially the precipitation and temperature from the *Copernicus ERA5-Land* dataset, which provides high-resolution gridded daily data at a spatial resolution of 0.1° (approximately 10 kilometers).

A geospatial join was performed between each household's latitude and longitude coordinates and the closest ERA5 grid cell. Based on this match, several key climate exposure variables are constructed for each household and wave-year observation. First, the annual mean temperature (T_{it}) is calculated as the average of daily mean temperatures across the full agricultural year. Second, total precipitation (P_{it}) is computed as the cumulative rainfall over the growing season, reflecting water availability critical for crop growth.

I also include a squared heat exposure term (T_{it}^2,P_{it}^2) to capture potential nonlinear effects. These variables jointly provide a comprehensive and theoretically informed representation of climate stressors relevant for smallholder agriculture. Each of these climate variables is constructed with agronomic plausibility in mind and aligns with thresholds used in prior empirical work on weather-yield relationships. Their inclusion allows the study to differentiate between linear and non-linear impacts of climate shocks, as well as to test heterogeneous effects across crop types. The resulting panel dataset is organized at the household–crop–wave level, permitting the estimation of fixed-effects regressions that capture crop-specific responses to varying climatic conditions.

Table 3: Summary description of the keys variables

Variable	Description	Source	Computed (How)	Unit
T_{it}	Annual mean temperature for household i at time t	ERA5-Land	Average of daily temps	°C
P_{it}	Total precipitation for household i at time t	ERA5-Land	Sum of daily precip	mm
$\log(R_{ict})$	Log1p-transformed revenue	Constructed	log of R_{ict}	–
R_{ict}	total revenue for household i from the crop c at time t	LSMS-ISA	Declaration of the household	BIRR[a]
Y_{ict}	Yield for household i from it major crop c at time t	LSMS-ISA	Quantity/area	kg/ha
$\log(Y_{ict})$	Log1p-transformed yield	Constructed	log of Y_{ict}	log(kg/ha)

[a](Ethiopia devise)

The Table 3 provides summary statistics of key variables across all five waves of the LMSIA dataset. The variables display considerable variability, both cross-sectionally and over time, which justifies our empirical strategy that includes household and time fixed effects. These descriptive statistics displays the heterogeneity in climatic exposure and agricultural outcomes across space and time. The retention rate chart Figure 19 illustrates the proportion of households that remained in the panel across successive survey waves. Starting from wave 1, the panel shows an exceptionally high level of household retention, consistently above 99 percent throughout the study period. A slight increase is observed between waves 2 and 3, where retention reaches 100 percent, likely reflecting data cleaning adjustments or successful follow-ups of previously missing households. Although a small decline is noted in wave 4 and a more visible drop in wave 5, the overall attrition remains minimal. This high degree of panel stability is critical for the reliability of longitudinal analysis, as it ensures that changes observed over time are less likely to be driven by sample loss or selection bias. Therefore, the dataset provides a strong foundation for examining dynamic processes such as behavioral responses to climate variation or shifts in agricultural practices over time.

2.2 Descriptive Statistics

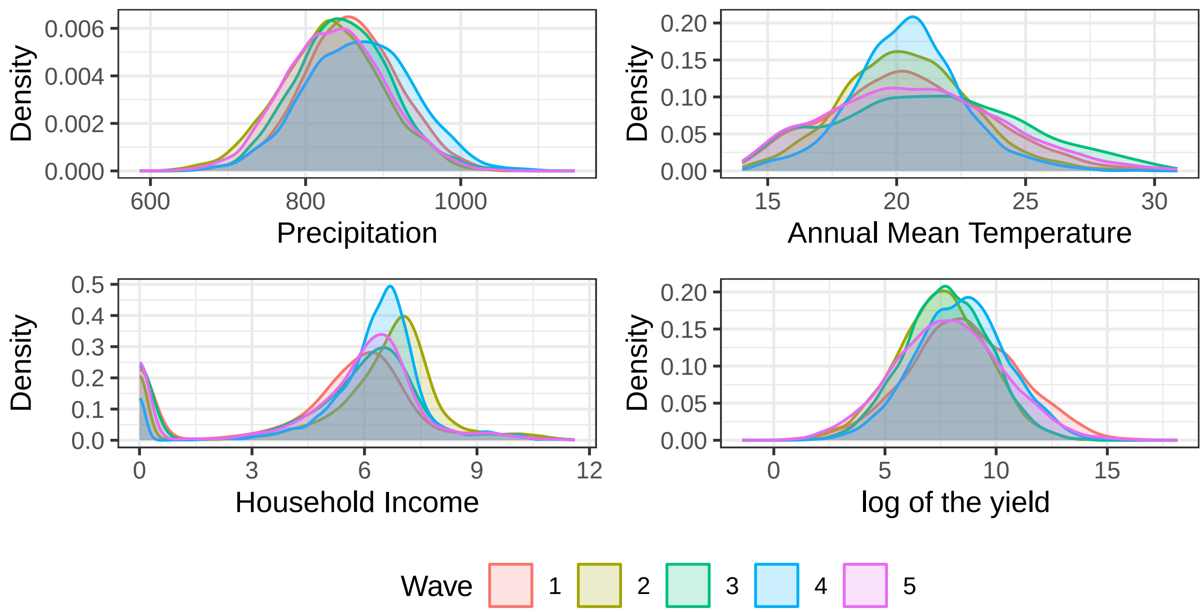
Table 4 presents summary statistics for the main variables used in the analysis, based on household-level observations across five waves of the LSMS-ISA dataset in Ethiopia. The average annual temperature across all household–year combinations is approximately 20.67°C, with considerable variation (min = 14.01°C, max = 30.89°C).This reflects Ethiopia's diverse agroecological zones from highland to lowland areas. Annual precipitation also shows wide dispersion, with a mean of 849.77 mm, ranging from semi-arid to high-rainfall regions.

Table 4: Summary statistics of the Key variables

Variable	Unit	Mean	S.Error	Min	Max
Temperature	°C	20.67	2.98	14.01	30.89
Precipitation	mm	849.77	65.90	585.90	1,146.97
log(Yield)	log(kg/ha)	8.04	2.23	-1.42	8.04
log (Income)	log(BIRR)	103	80.41	0	311

The average yield is 84.25 kg/ha, but the high standard deviation and a maximum value of 460 kg/ha indicate large disparities in productivity, likely due to heterogeneity in crop types, farming practices, and climate exposure. The cultivated area per observation averages 0.0234 ha, with extreme values reaching over 0.1300 ha, though the minimum value suggests some households operate on small plots. The log-transformed yield variable helps address skewness in the yield distribution, which is typical in smallholder datasets. For the summary by wave see table Table 9.

Figure 3: Distribution of the Main variables



The correlation matrix in Table 8 illustrates the linear relationships between key climatic variables used in the regression models. As expected, the strongest correlation appears between `tmean_wave_mod` and its squared term `tmean_mod_c_sq` (r = 0.24), reflecting the mathematical link between a variable and its quadratic transformation. A similar pattern, though weaker, is observed between `precip_mod` and its square (`precip_mod_c_sq`), with negligible correlation (r ≈ -0.01), suggesting that precipitation and its variation are more dispersed and potentially nonlinear in effect. The correlation between temperature and precipitation variables (both in level and squared forms) remains low ($r < 0.30$), indicating that multicollinearity is unlikely to be a major concern in the regression models. These results justify the inclusion of both linear and squared terms of temperature and precipitation in the model to capture potential nonlinear effects of climate on crop yield. Figure 14 examines the panel structure by showing household presence across waves.

The heatmap on the left shows how many households are simultaneously present in each pair of waves. Diagonal values confirm strong longitudinal retention, with over 3,800 households consistently present from wave 1 through wave 5. Off-diagonal values remain high, suggesting robust panel continuity. On the right, the percentage matrix shows the proportion of households that persist from one wave to the next. With retention rates above 99% between consecutive waves (and 100% from wave 4 to 5), the panel is remarkably balanced. This strong household presence across waves is critical for fixed effects models and ensures the reliability of within-household comparisons over time.

Figure 3 illustrates the distribution of key climatic and agricultural variables across the five survey waves. The precipitation panel shows relatively consistent distributions across waves, though with some variation in skewness. Wave 5 appears to have slightly higher cumulative rainfall, which may reflect interannual variability or shifts in the agricultural calendar. In the annual mean temperature panel, distributions are clustered between 15°C and 25°C, consistent with Ethiopia's predominantly temperate to subtropical climate zones. However, Wave 1 shows a slightly warmer profile, possibly due to seasonal timing or sample composition in that round. The log of yield panel indicates considerable variation across waves, with Wave 3 presenting a more concentrated and higher productivity distribution. This likely reflects favorable growing conditions or changes in crop composition during that year. Lastly, the days above 26°C panel reveals stark differences across waves. Wave 1 exhibits notably more heat stress days, with a high-density peak above 100 days. This confirms the importance of including heat exposure in the analysis, particularly given its crop-damaging potential highlighted in agronomic literature (Schlenker & Roberts, 2009; Deschênes & Greenstone, 2011) . Together, these distributions underscore the climatic variability Ethiopian smallholders face over time and justify the use of panel fixed-effects models with crop-specific climate interactions. Figure 15 displays the distribution of all the crops yield (taking in logarithm) in the full sample.

3 Econometric Models and Methodology

This study adopts a panel-data econometric approach to estimate the heterogeneous impact of climate variables on agricultural performance and household welfare across Ethiopian smallholders. The empirical analysis is based on a household–crop–year unbalanced panel, where the unit of observation is the household's primary crop in a given agricultural wave. The primary crop is defined as the crop with the largest cultivated area (or revenue share) within the household's portfolio, ensuring comparability across households and waves.

We employ fixed-effects regression with climate–crop interaction terms and nonlinear specifications to capture both marginal and threshold effects of climate exposure. Household and wave fixed effects absorb unobserved, time-invariant heterogeneity (α_i) (e.g., soil quality, irrigation infrastructure, managerial ability) and national shocks (λ_t) (e.g., inflation, policy reforms). Unlike random effects models, the fixed-effects estimator does not assume independence between unobserved heterogeneity and regressors; instead, it identifies effects from within-household and within-crop variation over time (Wooldridge, 2010).

3.1 Intensive Margin

Crop yield and the climate variables

The analytical framework of this study is grounded in the neoclassical theory of production, where agricultural output is modeled as a function of multiple inputs, including labor, land, and most importantly for this study, climatic factors. Following the literature (e.g., Burke et al., 2015 ; Schlenker & Roberts, 2009), we assume a multiplicative production structure that integrates both agro-climatic conditions and household-specific unobservables:

$$Y_{ict} = A_i \cdot e^{f_k(P_{it}, T_{it})} \cdot X_{ict} \cdot U_{ict}$$

where Y_{ict} is crop yield for household i, crop c, in wave t; A_i captures time-invariant household-specific productivity factors (e.g., soil quality, farmer skill); $f_T(\cdot)$ and $f_P(\cdot)$ represent the effect of temperature and precipitation; X_{ict} captures other productive inputs (land, fertilizer); and U_{ict} is a stochastic error term. Taking logarithm of the yields, give the following the estimating equation:

$$\log(Y_{ict}) = \log(A_i) + f_k(P_{it}, T_{it}) + \log(X_{ict}) + \log(U_{ict})$$

This log-linearization is standard in the empirical agricultural economics literature (Lobell et al., 2011) as it reduces skewness in yield data and allows percentage interpretation of effects. To allow for nonlinear climatic effects and heterogeneous responses across crops, we follow a strategy similar to Burke et al., 2015, who modeled temperature effects on economic outcomes using quadratic forms, and adopt the following flexible specification:

$$f_k(T_{it}, P_{it}) = \sum_{Z \in \{P_{it}, T_{it}\}} f_k(Z) = f_k(T_{it}) + f_k(P_{it}) = \beta_{1k} T_{it} + \beta_{2k} T_{it}^2 + \gamma_{1k} P_{it} + \gamma_{2k} P_{it}^2$$

where the coefficients $\beta_{1k}, \beta_{2k}, \gamma_{1k}, \gamma_{2k}$ vary by crop k. This formulation captures both marginal climate effects and threshold nonlinearities, such as the yield-damaging effects of heat beyond optimal levels (Deschênes & Greenstone, 2011) . Critically, by interacting temperature and precipitation (and their squares) with crop dummies, we allow each crop to have its own climatic response curve. This is theoretically justified: different crops have different heat and water stress thresholds due to their physiological characteristics. For instance, teff or maize is more heat-sensitive than cassava or millet, which are known to be more resilient to drought and high temperatures (FAO, 2021;Jarvis et al., 2012).

$$f_k(T_{it}, P_{it}) = \sum_{Z \in \{P_{it}, T_{it}\}} f_k(Z) \cdot \mathbb{1}_{Crop_{ic}=k} = [f_k(T_{it}) + f_k(P_{it})] \cdot \mathbb{1}_{Crop_{ic}=k}$$

We estimate the following fixed-effects model:

$$\log(Y_{ict}) = \alpha_i + \lambda_t + \delta_c + \sum_{k}^{C} f_k(T_{it}, P_{it}) + \varepsilon_{ict}$$

$$\log(Y_{ict}) = \alpha_i + \lambda_t + \delta_c + \sum_{k}^{C} \sum_{Z\in\{P_{it},T_{it}\}} f_k(Z) \cdot \mathbb{1}_{Crop_{ic}=k} + \varepsilon_{ict}$$

- α_i: household fixed effects, control for unobserved, time-invariant household characteristics;
- λ_t: wave fixed effects, absorb year-specific shocks (e.g., price variation, national policy changes);
- δ_c: crop fixed effects, account for differences in baseline yields across crops;
- $\mathbb{1}_{Crop_{ic}=k}$: dummy variable indicating whether crop c equals crop k;
- ε_{ict}: idiosyncratic error term.

This specification captures both the level effect of each crop (via δ_c) and its differential sensitivity to temperature and rainfall via the interaction terms. The inclusion of squared terms enables identification of nonlinear yield responses, commonly observed in agronomic literature (Schlenker & Roberts, 2009 ; Guiteras, 2009). Agricultural responses to climate are inherently heterogeneous. Grouping crops together in a pooled model would assume a common response to temperature and rainfall—an assumption that is clearly unrealistic. Different crops have distinct growing seasons, root depths, evapotranspiration rates, and resistance to extreme weather events. By allowing coefficients to vary by crop, this model addresses one of the key critiques of early empirical work, which averaged across crops or assumed homogeneous sensitivities (Lobell & Field, 2007). Moreover, interacting climate with crop dummies approximates a generalized difference-in-differences structure where each crop type experiences the same climatic treatment but may respond differently. This mirrors the strategy in panel DID studies that compare units with heterogeneous exposure to treatment across time (Angrist & Pischke, 2009) .

Identification Strategy and Assumptions

Identification relies on within-household, within-crop, over-time variation in climate exposure. Household fixed effects (α_i) absorb unobserved time-invariant heterogeneity (e.g., land quality, altitude, persistent soil fertility, irrigation infrastructure, or farmer ability), while time fixed effects (λ_t) capture common shocks such as inflation, fertilizer subsidies, or pest outbreaks. Crop fixed effects (δ_c) control for systematic biological differences across crops. Under these controls, the fixed-effects estimator isolates the impact of climatic variation on yields and revenues. Formally, we assume strict exogeneity:

$$\mathbb{E}[\varepsilon_{ict} \mid T_{it}, T_{it}^2, P_{it}, P_{it}^2, \alpha_i, \lambda_t, \delta_c] = 0$$

which implies that, conditional on the observed climate variables and fixed effects, the error term is mean-independent of the regressors. This assumption is plausible because farmers cannot manipulate temperature or rainfall.

However, a key potential challenge arises if households adjust their input use (e.g., fertilizer, labor, irrigation) in response to climate conditions. In this case, climate variables would influence yields both directly and indirectly via input decisions, potentially violating strict exogeneity. If so, the estimated coefficients may overstate or understate the true biophysical effect of climate on productivity. This concern is common in the climate–agriculture literature (e.g., Burke et al., 2015) .

To mitigate this risk, we interpret our coefficients as the reduced-form effect of climate exposure, encompassing both direct biophysical impacts and behavioral responses. In robustness checks, we control for time-varying input variables (where available) to test whether results are sensitive to input adjustments. This approach ensures that while identification is not perfect, the estimates remain informative about the total effect of climate variability on household agricultural outcomes.

The model we have to estimate is:

$$\log(Y_{ict}) = \underbrace{\alpha_i}_{\text{Household FE}} + \underbrace{\lambda_t}_{\text{Time FE}} + \underbrace{\delta_c}_{\text{Crop FE}} + \sum_{k=1}^{K} \left(\underbrace{\beta_{1k}T_{it} + \beta_{2k}T_{it}^2}_{\text{Temperature}} + \underbrace{\gamma_{1k}P_{it} + \gamma_{2k}P_{it}^2}_{\text{Precipitation}} \right) \cdot \mathbb{1}_{\{Crop_{ic}=k\}} + \varepsilon_{ict}$$

The error term can be written as:

$$\varepsilon_{ict} = \underbrace{f(\alpha_i, \lambda_t, \delta_c)}_{\text{Fixed effects component}} + \underbrace{g(T_{it}, P_{it})}_{\text{Climate relationship}} + \underbrace{\eta_{ict}}_{\text{Idiosyncratic shock}}$$

A key condition for fixed-effects regressions is that households experience enough variation in climate over time. Since household fixed effects absorb all time-invariant characteristics (e.g., soil quality, altitude, farmer skill), the coefficients are identified only from within-household variation in temperature and rainfall across survey waves. If climate variables were constant for a household, their effects could not be estimated. Fortunately, Ethiopia's strong inter-annual variability in rainfall and growing-season temperatures, combined with its diverse agro-ecological zones, provides ample variation to meet this requirement (Burke et al., 2015; Schlenker & Roberts, 2009).

Moreover, because the model interacts climate variables (and their squares) with crop dummies, it is also necessary that crops vary across households and over time. If a crop were cultivated by only one household, or by the same households in all waves, the interaction terms would be perfectly collinear with household fixed effects, making coefficients unidentified. Reliable estimation therefore requires that households switch crops across waves and that each crop be cultivated by multiple households. This condition is satisfied in our data, which cover a wide range of crops and crop rotations. Standard errors are clustered at the household level to account for both serial correlation and heteroskedasticity, following Cameron & Miller (2015).

Temperature bins and crop yields responses

My study aims to quantify the impact of temperature variability on household-level agricultural performance, measured as either log crop yield, by capturing the non-linear and crop-specific responses to daily temperature exposure during the growing season. Unlike standard linear or quadratic climate models that assume a fixed response across units, I allow temperature effects to vary both non-linearly (across temperature ranges) and by crop type, which is essential in contexts where households specialize in different crops that have distinct heat tolerances and growth requirements. To achieve this, I follow a flexible, semi-parametric regression strategy inspired by Schlenker & Roberts, 2009, discretizing the temperature distribution into small bins and interacting these with the household's main crop. We begin with a general model where the outcome depends on the integral of a temperature response function $g_k(h)$, which maps temperature exposure h to economic performance for crop k:

$$y_{it} = \int_{\underline{h}}^{\overline{h}} g_k(h) \cdot \phi_{it}(h)\, dh + f(P_{it}) + \alpha_i + \lambda_t + \varepsilon_{it} \tag{A1}$$

To investigate the heterogeneous effects of temperature on household-level agricultural outcomes, we adopt a flexible semi-parametric specification that discretizes the temperature distribution into bins and interacts them with the household's main crop. This step-function approach allows for non-linear and crop-specific responses to daily temperature exposure, without imposing a rigid functional form on the relationship. Compared to parametric alternatives such as Chebyshev polynomials or piecewise linear splines, this specification offers two key advantages. First, it provides a high degree of interpretability: each estimated coefficient directly reflects the marginal effect of an additional day in a specific temperature range for a given crop, facilitating empirical temperature response curve recovery. Second, it aligns with agronomic intuition and existing literature (e.g., Schlenker & Roberts, 2009), which highlights that crops differ substantially in their sensitivity to heat, often in a non-monotonic fashion. While smoother functional forms may reduce estimation noise, they come at the cost of masking threshold effects or abrupt changes in slope that may be biologically meaningful.

In contrast, our bin-based approach accommodates sharp transitions in response and captures localized effects of heat stress. Furthermore, we control for precipitation through a second-order polynomial to account for its known non-linear impact on crop productivity. This specification strikes a balance between flexibility, clarity, and empirical credibility, and is particularly suited for our setting where households grow different crops, often under varying climatic conditions. This model assumes that each day's temperature contributes to outcomes cumulatively and that this contribution depends on how hot or cold that day is, through the unknown function $g_k(h)$. Since $g_k(h)$ is unknown and daily temperature is continuous, we approximate the integral using a discretized version. We divide the temperature space into equal-width bins of 1°C. For each bin $j \in \mathcal{J}$, we count the number of days $D_{it}^{(j)}$ household i is exposed to that temperature bin in year t:

$$D^{(j)} = \Phi_{it}(h_{j+1}) - \Phi_{it}(h_j)$$

Where $\Phi_{it}(h_j)$ is the cumulative distribution of temperature up to h_j and h_j and h_{j+1} are the lower and upper bounds of bin j (e.g., [12,13), [13,13), …) We then estimate the following regression model:

$$y_{it} = \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \theta_{jk} \cdot D_{it}^{(j)} \cdot \mathbb{1}_{Crop_{ic}=k} + f(P_{it}) + \alpha_i + \lambda_t + \varepsilon_{it} \tag{A2}$$

Where:

- θ_{jk} measures the marginal effect of one additional day in bin j for households whose main crop is k;
- The interaction $D_{it}^{(j)} \cdot \mathbb{1}_{Crop_{ic}=k}$ allows for different shapes of $g_k(h)$ depending on the crop.

The model includes household fixed effects (α_i) to control for all time-invariant unobserved characteristics such as land quality, managerial ability, or long-term access to assets and infrastructure. Year fixed effects (λ_t) are added to capture aggregate shocks such as macroeconomic fluctuations, policy reforms, or widespread climatic anomalies. Crop fixed effects are not included separately, since crop dummies are already interacted with temperature bins to capture crop-specific thermal sensitivities. Adding them would be redundant and potentially collinear with household effects, especially for households cultivating the same crop over multiple years.

Identification therefore relies on within-household variation over time in both temperature exposure and crop choice, allowing climate effects to differ by crop without imposing restrictive functional forms. Equation (A2) is estimated using the within estimator, which demeans the data to eliminate household fixed effects and focus on within-unit variation. Estimation is performed using OLS, with standard errors clustered at the household level to correct for heteroskedasticity and serial correlation.

Each coefficient $\widehat{\theta}_{jk}$ measures the marginal effect of one additional day in temperature bin j on the logarithm of agricultural performance for households cultivating crop k. Interpreted in percentage terms, it reflects the proportional change in yield or revenue relative to the omitted reference bin. Plotting the full set of $\widehat{\theta}_{jk}$ coefficients across all bins yields the empirical temperature response function $\widehat{g}_k(h)$, which provides a flexible and nonparametric representation of crop-specific climate sensitivity.

3.2 Economic Margin

To deepen the analysis beyond physical productivity, this model investigates how climate variability affects broader measures of household welfare, namely, total agricultural income of the households. Unlike yield-focused regressions that isolate biophysical crop responses, this approach captures the economic consequences of climate shocks at the household level, encompassing behavioral adjustments, crop portfolios, and adaptation strategies. The dependent variable is the logarithm of total household income, allowing for multiplicative effects and variance stabilization. The specification interacts both temperature and precipitation (as well as their squared terms) with the primary crop cultivated by each household, enabling the marginal effects of climate to vary across farming systems.

This formulation recognizes that households specialized in crops such as maize or sorghum may differ significantly in both exposure to and capacity to absorb climate-related risks. To capture the well-documented nonlinear effects of weather, quadratic terms are included, reflecting that moderate deviations may be tolerable, or even beneficial, while extreme conditions become increasingly harmful. The model is estimated using the following fixed effects panel regression:

$$\log(R_{it}) = \alpha_i + \lambda_t + \sum_k \left(\beta_{1k}T_{it} \cdot \mathbb{1}_{Crop_{ic}=k} + \beta_{2k}T_{it}^2 \cdot \mathbb{1}_{Crop_{ic}=k} + \gamma_{1k}P_{it} \cdot \mathbb{1}_{Crop_{ic}=k} + \gamma_{2k}P_{it}^2 \cdot \mathbb{1}_{Crop_{ic}=k}\right) + \varepsilon_{it}$$

This structure follows the empirical strategies of Schlenker and Roberts (2009) and Burke et al. (2015) , who emphasize the importance of allowing for nonlinearity and crop-specific climate sensitivity. Household fixed effects (α_i) control for unobserved, time-invariant characteristics such as land quality or managerial skill, while year fixed effects (λ_t) absorb national-level shocks, policy changes, and aggregate trends. The crop fixed effect is implicitly addressed via the interaction terms, which allow each crop to exhibit its own climate response profile. Standard errors are clustered at the household level to account for potential autocorrelation and heteroskedasticity. Overall, this model offers a flexible yet behaviorally grounded framework for identifying the heterogeneous effects of climate variability on rural livelihoods.

Explanation of Crop Fixed Effects Usage in Yield vs. Revenue Models

The inclusion of crop fixed effects differs between the yield and revenue models due to the structure of the data and the nature of the dependent variable in each case. In the yield model, the unit of observation is at the household–crop–year level, meaning that each household can report yields for multiple crops in the same year. As a result, crop fixed effects are necessary to control for systematic differences in potential yields across crop types, such as biological productivity, growth cycles, and input requirements, which could otherwise confound the estimated effect of climate variables. In contrast, the revenue model aggregates income at the household–year level, with each household assigned only one primary crop per period. Because there is no within-household variation in crop type at a given time, crop fixed effects would be either perfectly collinear with household fixed effects or absorbed by them, making their inclusion redundant or unidentifiable.

Instead, the revenue model captures crop-specific climate sensitivity through interaction terms between climate variables and the household's main crop, allowing for differentiated marginal effects while preserving model identifiability. This difference in specification reflects the underlying data structure and ensures that variation used to estimate coefficients is meaningful and not mechanically eliminated by collinearity.

3.3 Extensive Margin of Adaptation: Household main crop choice

To analyze how climate shapes farmers' decisions to cultivate specific crops, (a central question to understand adaptation on the extensive margin), I estimate a fixed-effects panel model where the dependent variable is a binary indicator for whether a given crop is cultivated by a household in a given year. Temperature is modeled semi-parametrically using the number of days falling into specific temperature bins rather than average temperature and its square, which are more restrictive and potentially misleading.

This binning strategy captures the nonlinear and asymmetric nature of crop-climate relationships, especially around heat thresholds. It is more behaviorally realistic: two households may face identical seasonal average temperatures, yet one may endure several extremely hot days that alter planting decisions, a nuance that average temperature fails to detect. Precipitation is included through a second-order polynomial to account for the well-documented nonlinear effects of rainfall, where both scarcity and excess can deter cultivation. This modeling approach is original in two ways. First, while most studies of climate impacts in agriculture focus on the intensive margin, such as yields conditional on planting, we shift attention to the extensive margin, i.e., the decision to plant or not, which is equally crucial in climate adaptation (Kala, 2017 ; Emerick et al., 2016).

Second, our interaction of detailed daily temperature distributions with crop-specific binary outcomes within a household fixed effects framework provides a novel, granular lens on how climate influences farm-level crop allocation in practice. Unlike aggregate studies or those assuming homogeneous climate responses across crops, our method allows climate sensitivity to vary flexibly by temperature intensity and by crop type, within the same population. By leveraging household panel data and fine-grained climate exposure, this model provides new insights into the behavioral dimensions of climate risk management in smallholder systems. We model the probability that household i cultivates crop k in year t, denoted by $Cult_{ikt} \in \{0, 1\}$, as a function of climatic variables and fixed effects. The dependent variable equals 1 if crop k is cultivated, and 0 otherwise. We specify a semi-parametric fixed effects panel model as follows:

$$C_{ikt} = \sum_{j \in \mathcal{J}} \theta_j^{(k)} D_{it}^{(j)} + \gamma_1^{(k)} P_{it} + \gamma_2^{(k)} P_{it}^2 + \alpha_i + \lambda_t + \varepsilon_{ikt} \tag{EM}$$

In this model, C_{ikt} is a binary variable equal to 1 if household i cultivates crop k in year t, and 0 otherwise. The main explanatory variables are $D_{it}^{(j)}$, which denote the number of days during the growing season that fall within temperature bin j (e.g., [14–16°C], [16–18°C], etc.), and P_{it} and P_{it}^2, which represent the total precipitation and its square, respectively, capturing the potential non-linear impact of rainfall. The coefficients $\theta_j^{(k)}$ measure the marginal effect of one additional day in temperature bin j on the probability of planting crop k, while $\gamma_1^{(k)}$ and $\gamma_2^{(k)}$ capture the linear and quadratic effects of precipitation. Household fixed effects α_i control for time-invariant unobserved heterogeneity across households, such as soil quality, land access, or long-term preferences, and year fixed effects λ_t absorb time-specific shocks like national price fluctuations or policy changes. The error term ε_{ikt} captures all remaining unexplained variation.

The model satisfies key econometric assumptions that ensure consistent and interpretable estimates. First, although climatic effects are modeled non-linearly with respect to temperature and precipitation, the specification remains linear in parameters ($\theta_j^{(k)}, \gamma_1^{(k)}, \gamma_2^{(k)}$), allowing estimation via OLS or fixed-effects logistic regression. This linear-in-parameters property ensures both tractability and clear interpretation. Second, we assume strict exogeneity, i.e. $E[\varepsilon_{ikt} \mid D_{it}^{(j)}, P_{it}, \alpha_i, \lambda_t] = 0$, which is plausible since farmers cannot influence temperature or rainfall, even though they may adapt behaviorally. Third, to avoid perfect multicollinearity, one temperature bin (typically a moderate range such as [23–24°C]) is omitted as the reference category, while precipitation variables are entered in linear and quadratic form to capture non-linear effects. This quadratic specification reflects the empirically observed hump-shaped relationship: moderate rainfall encourages planting, whereas both deficits and excesses reduce cultivation likelihood (Schlenker and Roberts (2009); Dell et al., 2014).

Fourth, identification requires sufficient within-household variation in climate and crop choice over time. This condition is met in our multi-wave panel, where households face changing weather conditions and do not always cultivate the same crops each year. Finally, because households are observed repeatedly, standard errors are clustered at the household level to account for serial correlation and heteroskedasticity. Importantly, simultaneity bias is avoided since planting decisions are made ex-ante, before the realization of harvest outcomes, and climate variables are exogenous. Thus, the estimated coefficients can be interpreted as causal reduced-form effects of climate on crop choice.

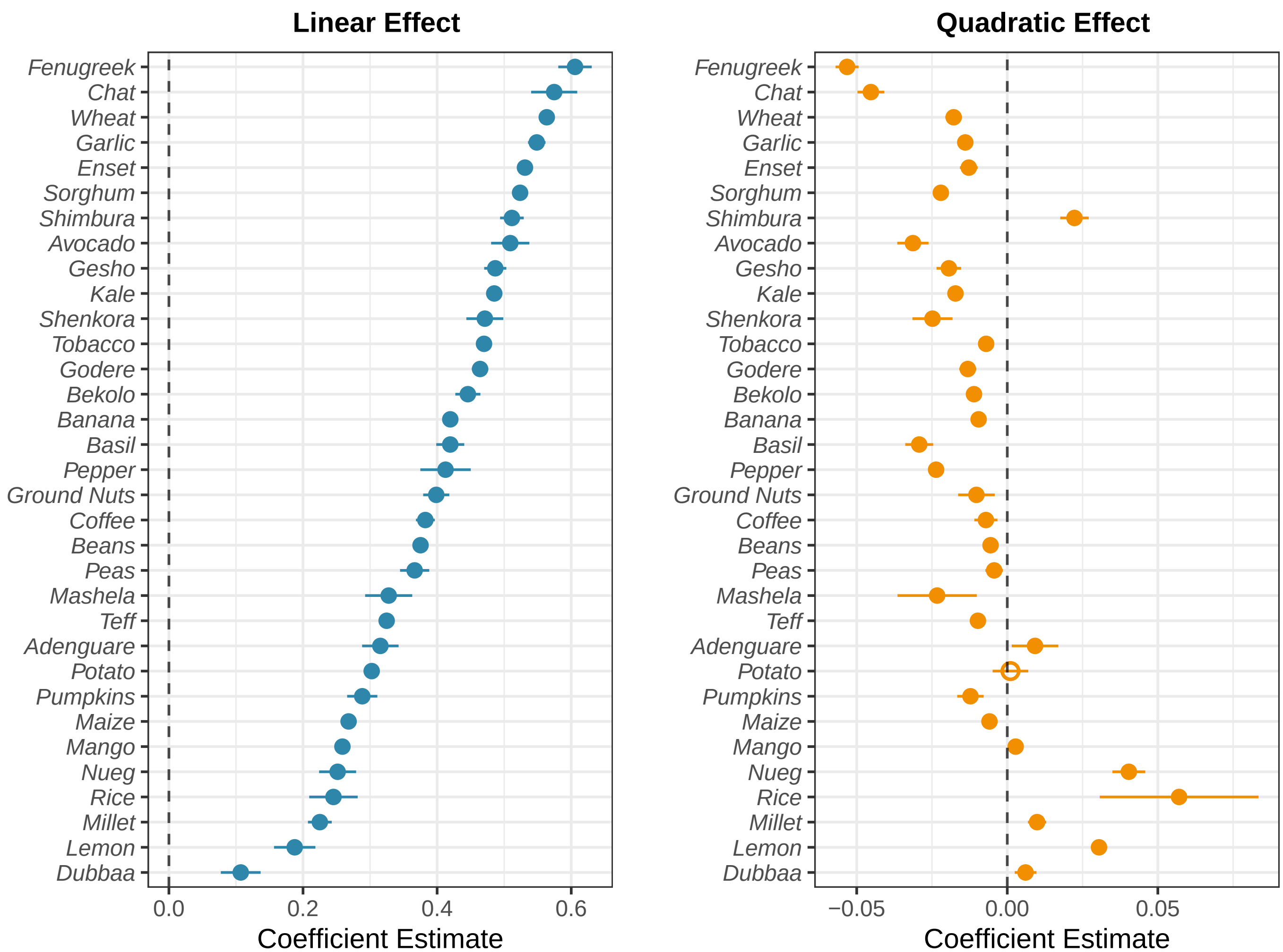
4 Results and Discussion

4.1 Intensive Margin

Crop yield with linear and nonlinear effect of climate variable

The regression results (Table 11, Figure 4, Figure 5) highlight strong heterogeneity in crop-specific responses to temperature and precipitation, this underscore the value of disaggregated analysis.

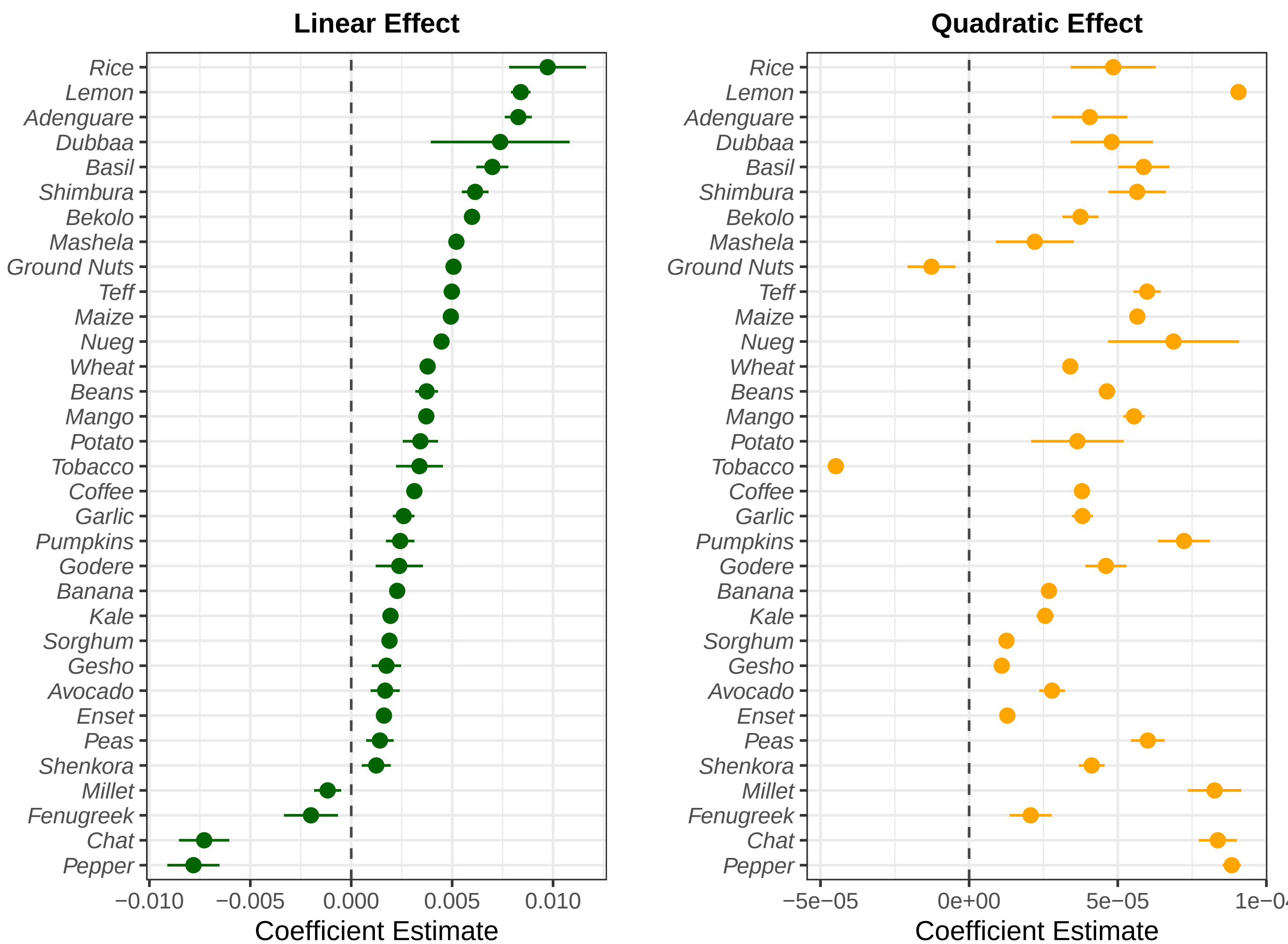
Figure 4: Temperature on Crop yield



As we can notice, among the cereals, rice, sorghum, millet, and Teff, yields respond positively to moderate warming but decline beyond critical thresholds, this confirms the convex temperature–yield relationship documented by Schlenker and Roberts (2009) and Lobell et al. (2011) . For instance, rice shows the largest effect ($+0.47$ per $+1°C$, -0.026 quadratic), followed by sorghum ($+0.38$; -0.016) and millet ($+0.33$; -0.015). Teff is less responsive ($+0.24$; -0.008) but highly rainfall-dependent ($+0.005$), consistent with Seo & Mendelsohn (2008) . This reinforces cereals' dual vulnerability to both drought and erratic rainfall, emphasizing the need for resilient varieties and soil-moisture conservation. When looking to the fruits, Avocado ($+0.59$; -0.021) and mango ($+0.24$; -0.008) benefit from moderate warming but decline under heat stress. In contrast, banana and lemon show small, mostly insignificant coefficients, indicating greater resilience. Precipitation effects are negligible, highlighting irrigation rather than rainfall as the key adaptation lever.

In vegetables, Garlic ($+0.59$; -0.017) and onion ($+0.55$; -0.016) show strong but nonlinear warming responses: high gains up to a threshold, then rapid decline, echoing Di Falco & Chavas (2009). Kale stands out as mainly rainfall-sensitive ($+0.005$), with temperature effects insignificant, pointing to the importance of regular moisture and irrigation planning.

Figure 5: Precipitation Effect on Crop yield



Besides, the legumes and pulses like Beans (+0.74; −0.016), groundnuts (+0.36; −0.010), and peas (+0.27; −0.004) also follow the positive-then-decline pattern, though peas uniquely benefit from rainfall (+0.006). This aligns with Lin (2011) and Kassie et al. (2014) , confirming legumes' high water sensitivity. Industrial crops. Sesame shows robust and stable temperature benefits (+0.42) without strong nonlinearities, while tobacco (+0.47; −0.017) reveals previously undocumented nonlinear effects in African settings, suggesting vulnerability under excessive heat.

Two insights stand out. First, the exceptionally large positive effects for rice and avocado diverge from tropical studies, likely due to Ethiopia's cooler baseline. Second, significant quadratic precipitation terms, especially for tobacco and groundnuts highlight yield risks not just from drought but also from rainfall excess, a dimension rarely emphasized in African research.

As an implication, we can say that climate change will not affect Ethiopian crops uniformly: rice and sesame may benefit from moderate warming, while Teff and garlic are highly exposed to heat and rainfall extremes. Policy must therefore avoid one-size-fits-all solutions. Instead, crop and region-specific adaptation heat and drought-tolerant varieties, targeted irrigation, and context-specific extension are essential. These findings support Ethiopia's CRGE strategy while enriching global evidence on highland tropical agriculture.

Non linear Model: Binning Temperature

The bin model allows for a more flexible estimation of the relationship between temperature and crop yield, free from restrictive functional assumptions. When examining yield responses across temperature intervals, we gain insight into thresholds beyond which yield effects become more pronounced.

This subsection examines *Teff, Wheat, Sorghum, Maize,* and *Millet*, Ethiopia's principal cereals, focusing on their heterogeneous temperature responses.
For *Teff*, yields remain relatively stable between 14–20°C, peak around 24–26°C, and then decline sharply beyond 27°C. This result is consistent with FAO (2010), which identifies an optimal range of 18–27°C, and with Alemayehu & Bewket (2017), who report similar patterns in the Ethiopian highlands.
Wheat displays optimal performance between 18–23°C, but yields drop rapidly above 24°C, becoming strongly negative beyond 27°C. This aligns with Asseng et al. (2015), who estimate yield losses of 6% per +1°C above the optimum, and Lobell et al. (2012), who emphasize heat sensitivity during flowering and grain filling.

Table 5: Temperature Bin Effects on Log Yield (Cereals)

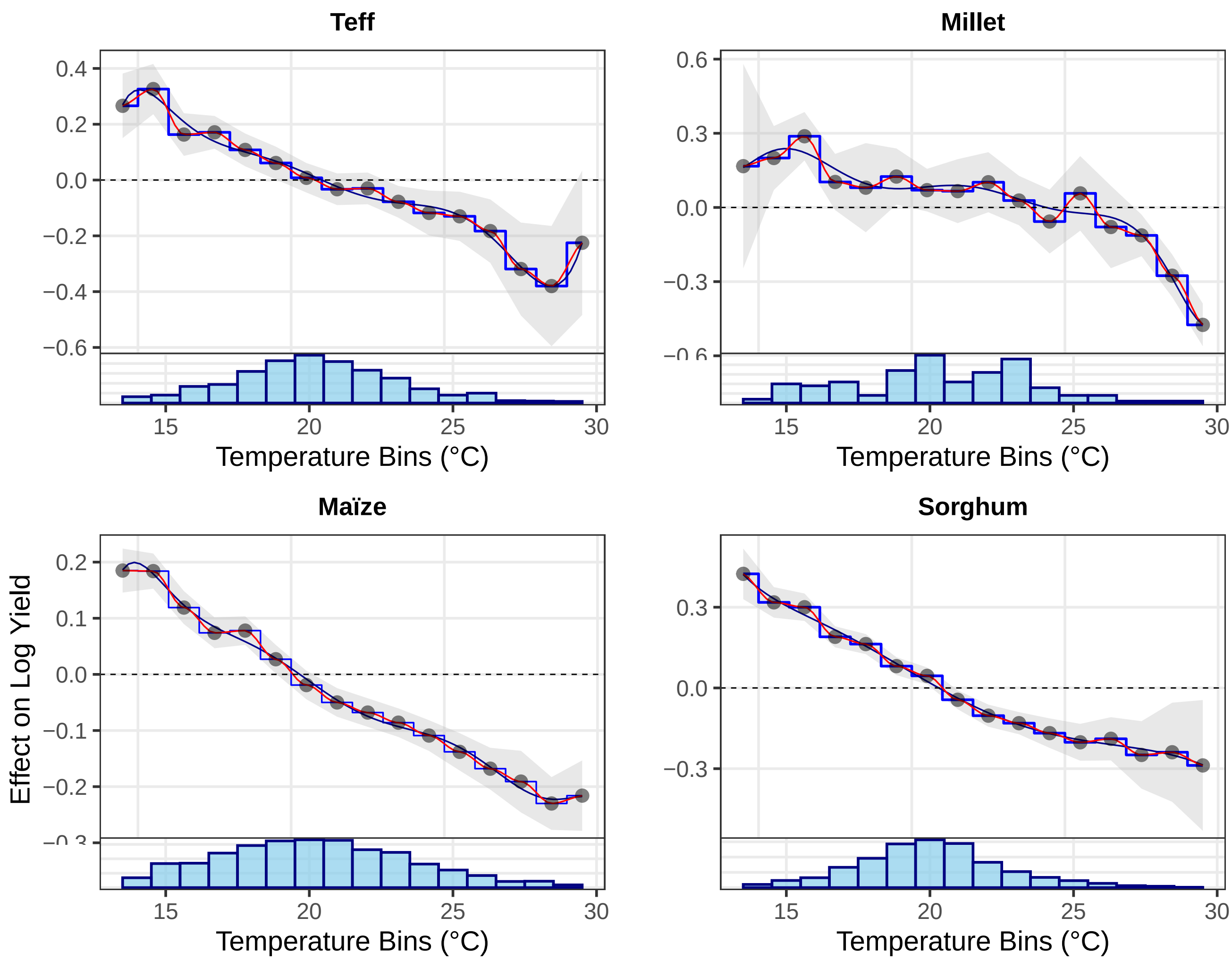
	Maize		**Millet**		**Sorghum**		**Teff**		**Wheat**	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
[14; 15[	0.185***	(0.020)	0.167	(0.211)	0.424***	(0.048)	0.266***	(0.059)	0.283***	(0.028)
[15; 16[	0.184***	(0.016)	0.200**	(0.066)	0.318***	(0.029)	0.326***	(0.046)	0.368***	(0.052)
[16; 17[	0.119***	(0.015)	0.288***	(0.050)	0.300***	(0.026)	0.163***	(0.039)	0.250***	(0.030)
[17; 18[	0.074***	(0.014)	0.103$^{\cdot}$	(0.058)	0.190***	(0.020)	0.171***	(0.030)	0.176***	(0.040)
[18; 19[	0.078***	(0.013)	0.080	(0.092)	0.163***	(0.019)	0.108***	(0.030)	0.100***	(0.027)
[19; 20[	0.027*	(0.013)	0.125*	(0.058)	0.081***	(0.017)	0.061*	(0.030)	0.034	(0.025)
[20; 21[	-0.019	(0.013)	0.070	(0.044)	0.045**	(0.016)	0.008	(0.027)	-0.028	(0.026)
[21; 22[	-0.050***	(0.013)	0.066	(0.066)	-0.044*	(0.018)	-0.033	(0.029)	-0.104***	(0.025)
[22; 23[	-0.068***	(0.013)	0.102	(0.062)	-0.103***	(0.021)	-0.030	(0.029)	-0.137***	(0.033)
[23; 24[	-0.086***	(0.013)	0.028	(0.051)	-0.131***	(0.021)	-0.078**	(0.029)	-0.204***	(0.038)
[24; 25[	-0.109***	(0.014)	-0.057	(0.066)	-0.168***	(0.028)	-0.118**	(0.041)	-0.262***	(0.046)
[25; 26[	-0.138***	(0.017)	0.057	(0.077)	-0.202***	(0.035)	-0.130**	(0.045)	-0.333***	(0.073)
[26; 27[	-0.168***	(0.019)	-0.079	(0.085)	-0.189***	(0.041)	-0.183**	(0.058)	-0.403***	(0.089)
[27; 28[	-0.191***	(0.028)	-0.113**	(0.043)	-0.249***	(0.064)	-0.319***	(0.085)	0.211***	(0.023)
[28; 29[	-0.230***	(0.024)	-0.276***	(0.044)	-0.239*	(0.094)	-0.380***	(0.110)	-0.180***	(0.024)
[29; +[	-0.216***	(0.032)	-0.475***	(0.044)	-0.288*	(0.124)	-0.225$^{\cdot}$	(0.132)	-0.630***	(0.025)
Fixed Effects										
Household (hhid)	Yes		Yes		Yes		Yes		Yes	
Wave	Yes		Yes		Yes		Yes		Yes	
Stats										
Obs	20,507		20,507		20,507		20,507		20,507	
S.E.: Clustered	by: hhid		by: hhid		by: hhid		by: hhid		by: hhid	
R^2	0.327		0.283		0.310		0.291		0.293	
Within R^2	0.063		0.002		0.038		0.012		0.015	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. The dependent variable is log agricultural yield. Temperature bins represent °C intervals (exact ranges not specified in original data).

Sorghum demonstrates high thermal tolerance, with positive yield effects up to 26°C and only marginal declines beyond 27°C. This finding supports its classification as a heat-resilient crop, consistent with Rosenow & Dahlberg (2000) and Rai et al. (1999), who highlight sorghum's adaptability to semi-arid and hot environments (up to 35–38°C).

For *Maize*, yields improve significantly under cooler conditions, with gains of +0.185 log points at 14–15°C and +0.074 at 17–18°C. However, coefficients turn negative beyond 21°C, reaching −0.230 at temperatures above 27°C. These results are consistent with the nonlinear responses identified by Schlenker & Roberts (2009) and Lobell et al. (2011), while also emphasizing that Ethiopia's cooler highlands provide short-term benefits within a narrow optimal window.

Figure 6: Temperature Effects on Cereals yield

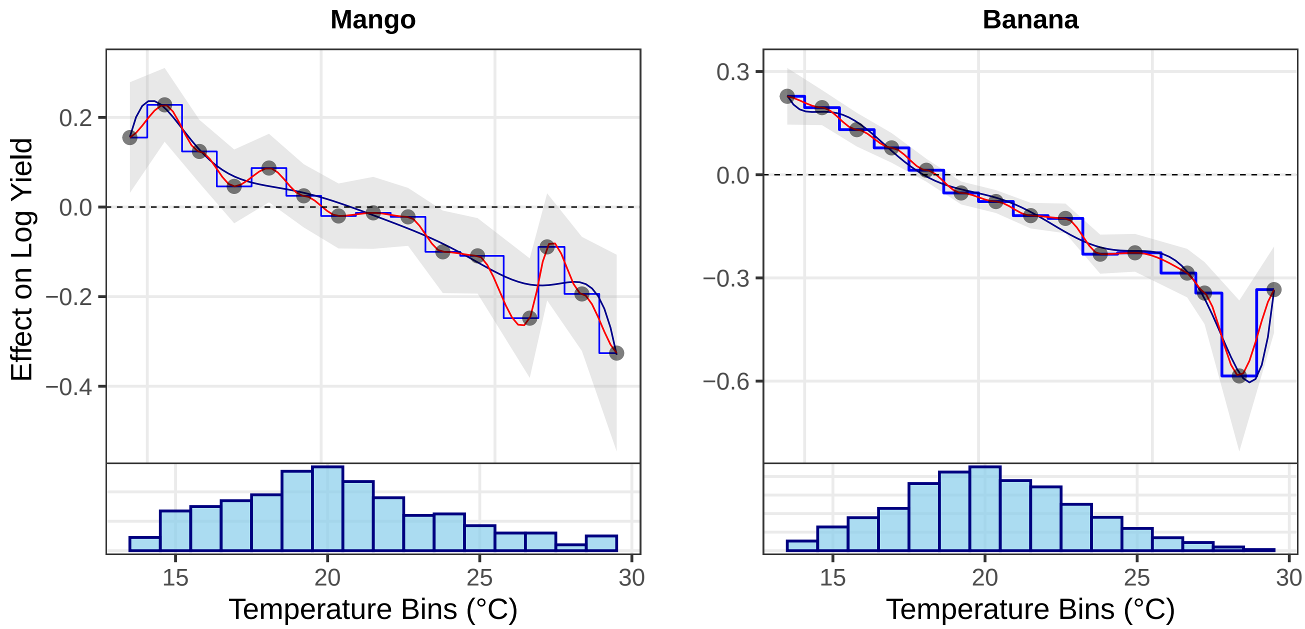


Millet, often described as heat- and drought-tolerant, surprisingly shows cold sensitivity in the Ethiopian highlands: yields increase modestly at 16–18°C (+0.287 and +0.103 log points, respectively) but decline with further cooling. This contrasts with global findings such as Lin (2011), suggesting that elevation-induced cold stress may override presumed resilience, and calling for a reevaluation of "climate-smart" crop classifications in highland systems.
These results display the differentiated climate sensitivity of Ethiopia's cereals. *Sorghum* appears best positioned under future warming scenarios, *Maize* is highly vulnerable above 21°C, while *Teff* and *Wheat* face significant risks from heat stress despite moderate resilience within optimal ranges. The unexpected cold sensitivity of *Millet* suggests that highland adaptation strategies must be carefully tailored rather than relying on generalized assumptions about crop resilience.

The yield responses of *Avocado, Mango,* and *Banana* to temperature, illustrated in Figure 7, display patterns broadly consistent with agronomic expectations, though with some discrepancies.
For *Avocado*, yields peak between 24–25°C, with declines observed below 18°C and above 27°C. This aligns with Whiley et al. (2002), who report optimal fruit set and growth within 20–25°C, and with Wolstenholme (2011), who note that heat above 28°C can cause flower abortion and fruit drop.
Mango yields are stable between 16–23°C, peaking around 21–22°C, but decline sharply beyond 25°C. This result is consistent

with Singh & Chadha (2001), who identify 22–24°C as the optimal range for flowering and fruit development, while prolonged exposure above 30°C reduces fruit set.

Figure 7: Fruits yield response to climate change



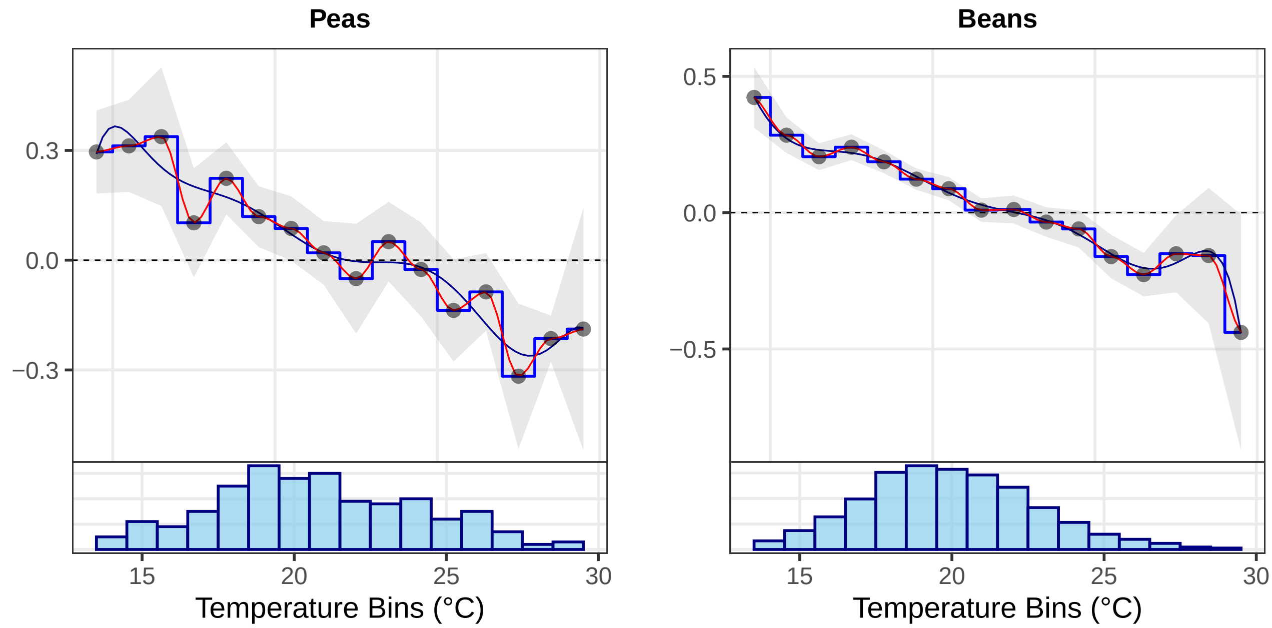
By contrast, *Banana* shows weaker and more uncertain yield responses, with slightly negative effects beyond 26°C but wide confidence intervals. According to Robinson & Galán Saúco (2010), the crop's optimal range is 26–30°C, with peak growth around 27–29°C. The weak signal in our model may reflect limited representation in hotter temperature bins, confounding agroecological factors, or banana's continuous growth cycle and high dependence on water availability, which weaken direct temperature–yield linkages.
So, while *Avocado* and *Mango* benefit from moderate warming within narrow thresholds, both are highly vulnerable to heat extremes. Their promotion should target mid-altitude zones where current temperatures remain within optimal ranges. In contrast, banana adaptation strategies should prioritize water management and irrigation rather than temperature-specific interventions, given its weaker modeled thermal response.

The temperature responses of *Beans*, *Peas*, and *Groundnuts* reveal patterns that broadly align with agronomic literature, though with some partial inconsistencies.
For *Beans*, yields peak between 19–21°C and again around 26–28°C, before declining sharply beyond 28°C. This pattern is consistent with Singh et al. (2018) and FAO (2021), who report that beans perform optimally within 18–24°C, while heat stress at higher temperatures disrupts flowering and pod filling. The decline observed in the model confirms these physiological vulnerabilities.

Peas exhibit a strong positive yield response between 18–23°C, followed by a steep decline after 24°C. This mirrors agronomic evidence showing that peas are highly sensitive to heat stress, particularly during reproductive stages (Choudhury et al., 2016). The model thus captures the thermal threshold with precision, reinforcing its biological coherence.

For *Groundnuts*, yields peak between 21–24°C, but decline rapidly after 25°C. While agronomic references such as FAO (2010) and Giller et al. (2004) describe groundnuts as relatively heat-tolerant up to 30°C, the earlier decline in this model may reflect compounding stressors, such as drought or poor soil fertility. This partial inconsistency suggests that yield sensitivity could be mediated by interactive factors not fully captured in the model.

Figure 8: Oleaginous nuts response to climate



These results underline the importance of context-specific adaptation. For beans and peas, breeding programs for heat-tolerant varieties and adjusted planting schedules will be crucial, especially as warming pushes average temperatures beyond their narrow thermal windows. For groundnuts, the earlier-than-expected decline highlights the need to integrate soil fertility and water management practices into climate adaptation strategies, ensuring resilience under combined stress conditions.

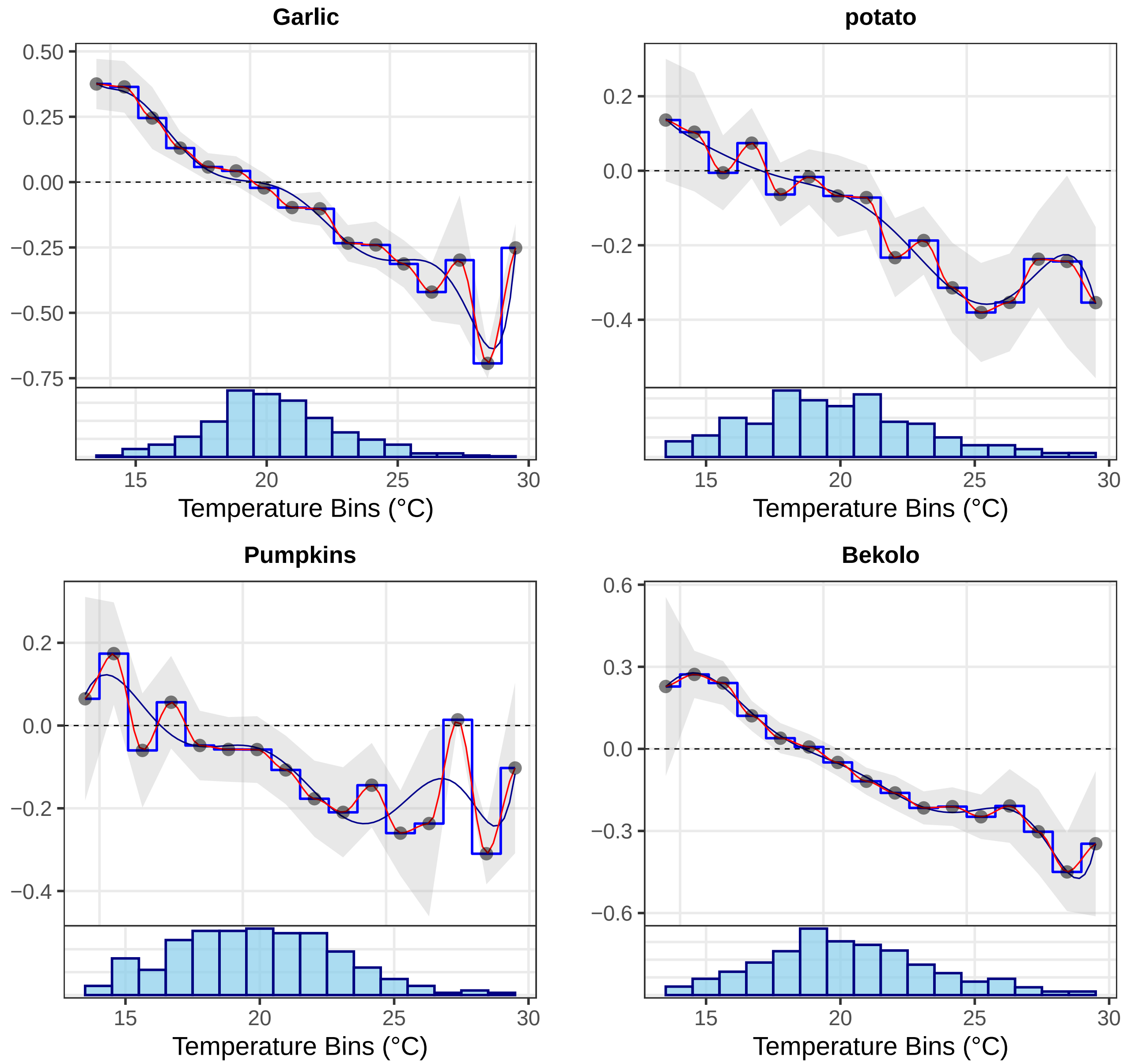
The results for vegetables and minor crops reveal heterogeneous but generally coherent temperature–yield relationships. *Garlic* shows strong temperature sensitivity, with optimal yields between 20–25°C, and marked declines below 15°C and above 27°C. This aligns with its physiology as a cool-season crop: according to Singh et al. (2019), garlic performs best within 12–25°C, while heat stress reduces bulb size and quality.

Kale displays a moderately convex temperature response, with peak yields around 20–24°C and significant declines beyond 26°C. This pattern is consistent with FAO (2021), which reports that kale thrives at 15–24°C, while high temperatures accelerate bolting and reduce leaf quality.
Potato has a narrow optimal range (16–22°C), with steep yield declines beyond 24°C. This result is consistent with Haverkort et al. (2013), who show that tuberization is highly sensitive to heat, with optimal performance between 15–21°C and inhibition of tuber formation above 25°C. The model confirms potato's vulnerability to future warming scenarios.

Pumpkins exhibit increasing yields between 18–25°C, with losses outside this interval. This finding agrees with FAO (2010), which identifies 20–27°C as optimal for pumpkin growth, avoiding frost and excessive heat.
For *Shenkora* (a lesser-documented Ethiopian legume), the model suggests an optimum around 23–25°C. However, the erratic responses observed below 20°C likely reflect data limitations or omitted variables. The lack of agronomic studies on Shenkora makes interpretation tentative.

Finally, *Bekolo* presents a smooth and biologically plausible response: yields rise between 18–25°C before declining beyond 26°C. This bell-shaped curve is typical of legumes with thermal thresholds around 25–28°C (Singh et al., 2018; FAO, 2021). The narrow confidence intervals strengthen the reliability of this result.

Figure 9: Vegetables Yield response to climate



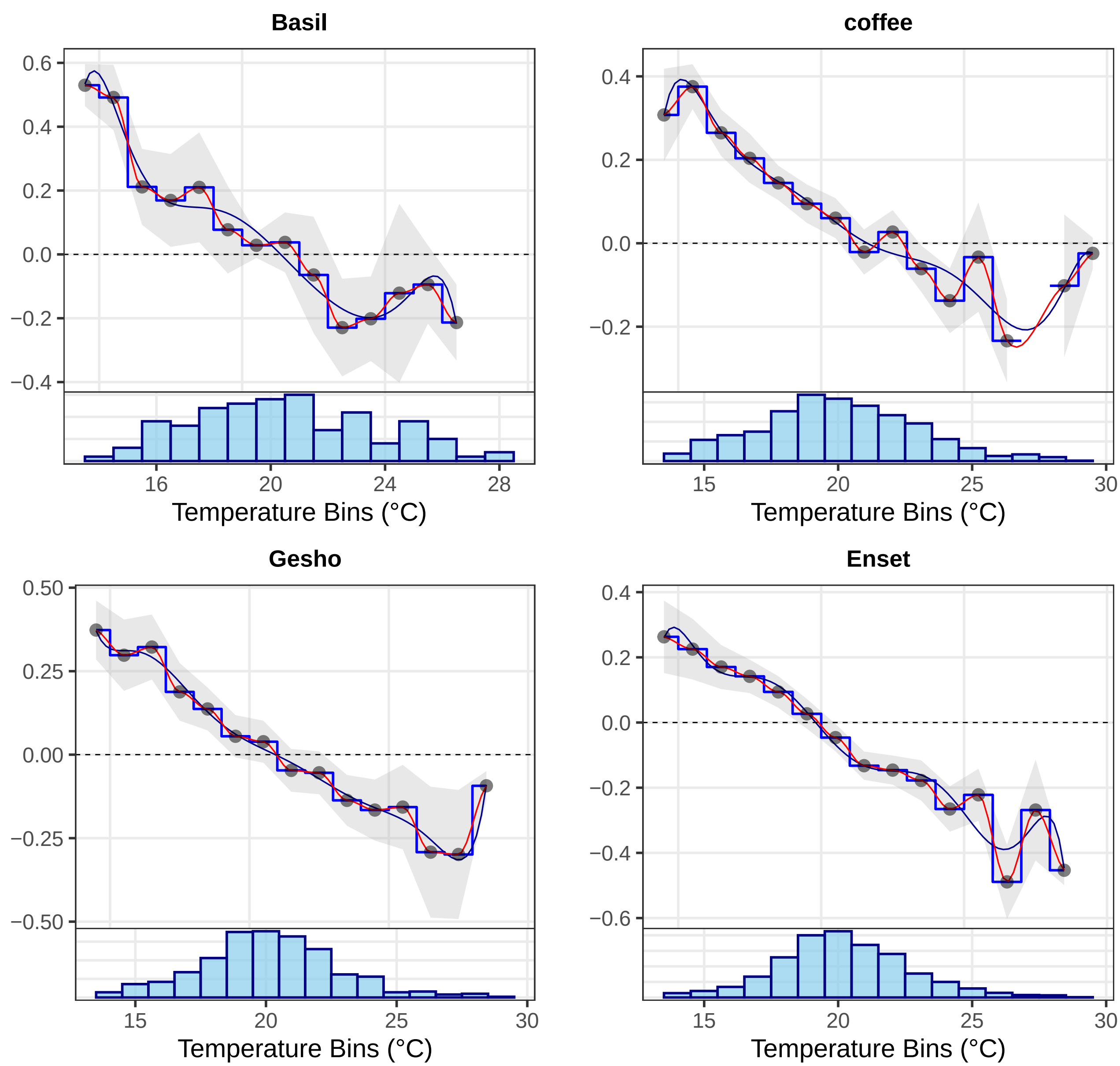
These findings emphasize the need for tailored adaptation strategies. For heat-sensitive crops such as potato and garlic, breeding for heat-tolerant varieties and shifting cultivation to cooler highlands will be essential. For crops like pumpkin and kale, irrigation and shading practices could buffer heat stress.

The response of *Basil* and *Coffee* to temperature exhibits clear and biologically coherent patterns.
For *Basil*, yields increase significantly between 18–24°C, peaking around 22–24°C before showing a mild decline. This is consistent with its preference for warm, temperate climates with optimal growth between 21–26°C (Krause et al., 2016). The sharp peak and narrow confidence intervals confirm the robustness of the estimates and their alignment with basil's physiological tolerance.

Coffee displays a well-defined bell-shaped curve, with yields peaking between 20–24°C. This matches the ideal temperature range for Arabica coffee, generally reported as 18–23°C (DaMatta et al., 2007). Beyond 25°C, the model shows a steep decline, reflecting the documented effects of heat stress on both yield and bean quality. These results are highly coherent and confirm established agronomic findings.
The response of *Enset* and *Gesho* also shows patterns consistent with ecological expectations, though with varying levels of documentation. *Enset*, a staple crop in Ethiopian highlands, peaks around 22–24°C, with marked declines beyond 26°C.

This aligns with evidence that enset thrives in cool tropical environments between 15–25°C, while heat stress reduces its productivity (Brandt et al., 1997); Borrell et al., 2020. The coherence is reinforced by the smooth slope and narrow confidence bands in the model.

Figure 10: Oil seeds Yield response to climate change effect on crop yield



For *Gesho*, yields peak between 20–24°C, with sharp decreases outside this range. This pattern is consistent with its cultivation in Ethiopia's mid-altitude areas, where moderate temperatures prevail. Although formal agronomic studies are scarce, the modeled response is in line with farmer-reported ecological suitability, lending credibility to the result.

These findings highlight the vulnerability of temperature-sensitive perennial crops like *coffee* and *enset* to warming trends, stressing the need for shade management, irrigation, and highland-focused cultivation strategies. For less-documented crops such as *Gesho*, further agronomic research would strengthen evidence for adaptive policy support.

4.2 Economic Margin

Table 6: Effect of climate on Household revenue (Economic margin)

	Temp. (lin)		**Temp. (quad)**		**Precip. (lin)**		**Precip. (quad)**	
Crop	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Banana	-0.0052	(0.0179)	0.0077$^{\cdot}$	(0.0040)	-0.0010*	(0.0005)	-0.0007	(0.0006)
Chat	-0.1424*	(0.0651)	0.0263*	(0.0118)	0.0054**	(0.0017)	-0.0085	(0.0001)
Coffee	0.0058	(0.0240)	-0.00138*	(0.0061)	-0.0005	(0.0006)	-0.0067	(0.0008)
Enset	-0.0324	(0.0276)	0.0211 ***	(0.0062)	0.0002	(0.0007)	-0.0004	(0.0009)
Garlic	-0.811*	(0.0241)	-0.0042	(0.0104)	-0.0003	(0.0010)	-0.0007	(0.0005)
Gesho	-0.0249	(0.0409)	0.022**	(0.0085)	0.0010	(0.0009)	-0.00034*	(0.0005)
Kale	0.0305	(0.0331)	0.0004	(0.0079)	-0.0015*	(0.0007)	-0.00025	(0.0001)
Maize	0.0014	(0.0096)	-0.0071**	(0.0025)	-0.0003	(0.0003)	0.0003	(0.0004)
Mango	-0.0222	(0.0416)	-0.0003	(0.0011)	-0.0043*	(0.0022)	0.00075	(0.0002)
Millet	-0.0112	(0.0389)	-0.0048	(0.0117)	0.0054*	(0.0029)	-0.00019	(0.0004)
Pepper	0.1190	(0.0852)	0.0030	(0.0118)	-0.0036*	(0.0018)	-0.00017	(0.00031)
Shenkora	-0.2293*	(0.1068)	-0.0024	(0.0356)	0.0050 *	(0.0025)	-0.00025	(0.0008)
Sorghum	-0.00023	(0.0233)	-0.0278***	(0.0068)	-0.0003	(0.0006)	0.00017	(0.0007)
Teff	0.0031	(0.0350)	-0.0159 *	(0.0077)	-0.00007	(0.0009)	0.00032$^{\cdot}$	(0.0002)
Fixed Effects								
Household (hhid)		Yes		Yes		Yes		Yes
Wave		Yes		Yes		Yes		Yes
Statistics								
Observations		20,507		20,507		20,507		20,507
S.E.:Clustered		by:hhid		by:hhid		by:hhid		by:hhid
R-squared		0.55545		0.55545		0.55545		0.55545
Within R-squared		0.37529		0.37529		0.37529		0.37529

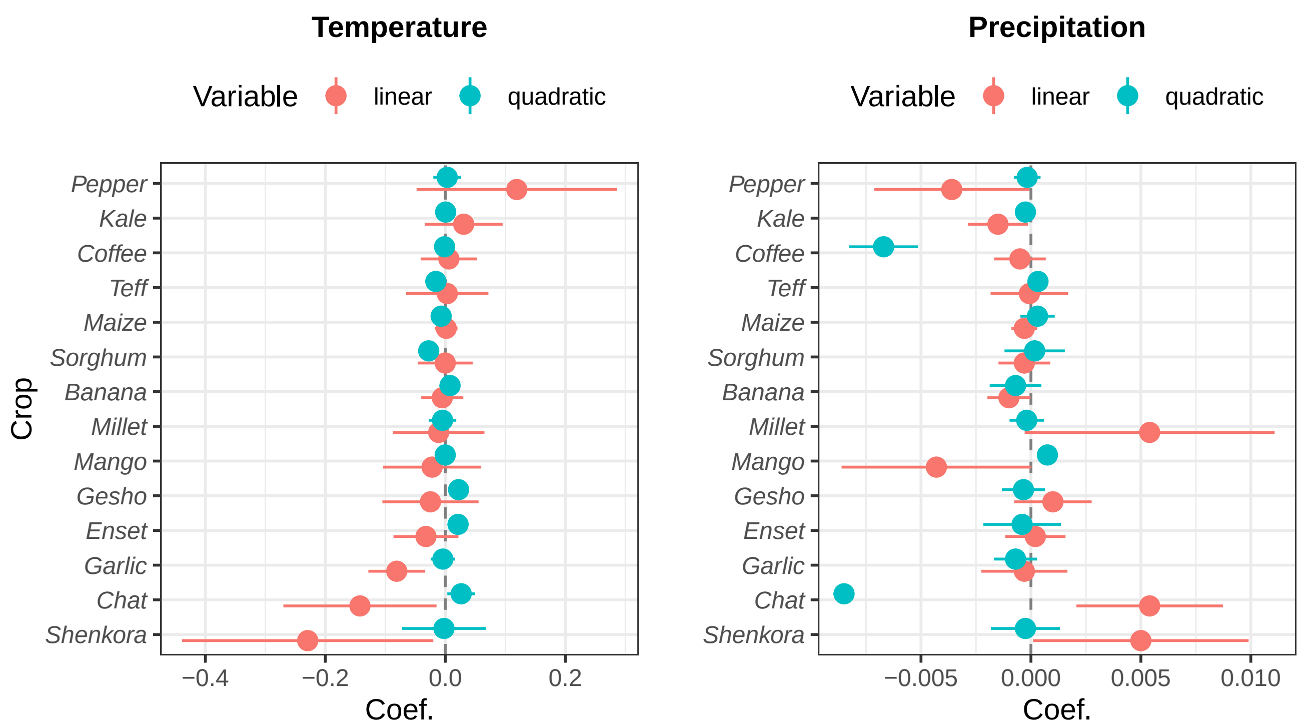
Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. Coefficients represent the impact of climate variables (temperature and precipitation, linear and quadratic terms) on crop revenues. Quadratic precipitation terms are multiplied by 10^5 for readability.

Table 6 reports the effects of temperature and precipitation on household agricultural revenue. Unlike yields, which are crop-specific, revenue is aggregated at the household level to capture adaptation strategies such as crop mix adjustments and labor reallocation.

Results highlight strong heterogeneity. High-value crops such as *Chat*, *Garlic*, and *Shenkora* exhibit significantly negative linear temperature effects (–0.14, –0.81, –0.23, respectively), showing high vulnerability to warming. By contrast, *Pepper* revenues increase with temperature, suggesting that mild warming may expand its profitability. Cereals such as *Teff*, *Maize*, and *Sorghum* display concave responses: revenues rise under moderate warming but decline beyond critical thresholds, consistent with agronomic evidence of heat stress effects (Deressa & Hassan, 2009). This mirrors the inverted-U relationships reported by Schlenker & Roberts (2009) and Lobell et al. (2011).

Precipitation effects are weaker overall, but positive for *Pepper*, *Sorghum*, and *Shenkora*, highlighting their dependence on rainfall. In contrast, excessive moisture harms crops like *Chat* and *Coffee*, likely through quality losses or disease pressure (Alemayehu et al., 2020). Quadratic precipitation terms are generally insignificant, but suggest that both drought and excess rainfall can undermine revenues.

Figure 11: Climate Margin Effect on Households Income



Overall, the findings show that revenue sensitivity is stronger for cash crops (e.g., *Chat*, *Coffee*, *Garlic*) than for staple cereals, which are relatively more resilient but still vulnerable at extremes. These results imply that climate-smart adaptation must differentiate strategies: protecting cash crops through irrigation and relocation to cooler zones, while stabilizing cereals through soil moisture conservation and insurance schemes. This aligns with Ethiopia's FAO (2018) CSA strategy and underlines the urgency of preparing for projected warming of 1.5–2.5°C by 2050 (IPCC, 2022).

4.3 Extensive Margin

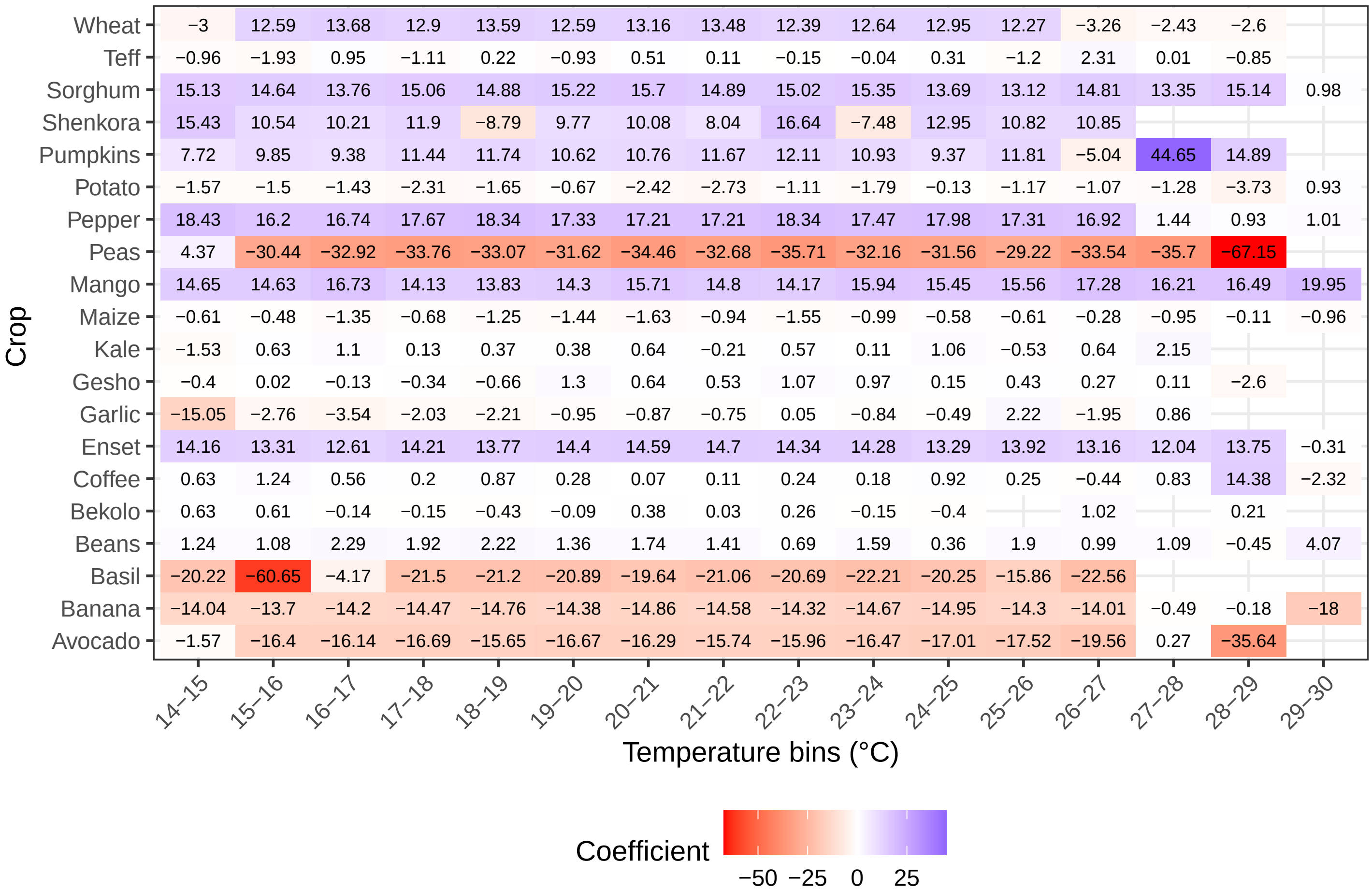
The mains results of the models estimations are presented in Figure 12 and Figure 13.

Cereals

The results in Table 16 shows that *Sorghum* is highly resilient and well adapted to semi-arid climates. Across all temperature bins, it displays consistently positive and statistically significant coefficients (e.g., +15.1 to +15.3 log-odds). Interpreted in odds-ratio terms, this means that the probability of farmers choosing sorghum increases sharply with warming, which is fully coherent with its well-documented heat and drought tolerance (Rosenow & Dahlberg, 2000). Sorghum's positive response explains why it is frequently used in rotational and mixed cropping systems in Sub-Saharan Africa, as it provides a reliable option under rising temperatures.

Wheat also shows positive but more moderate coefficients in the 16–25°C range (around +12 to +13.6 log-odds). This indicates that warming within this optimal window raises the odds of wheat being cultivated. However, coefficients turn negative above 27°C, reflecting wheat's vulnerability to heat stress during flowering and grain filling (Asseng et al., 2015). This turning point highlights the narrow climatic window within which wheat can remain a viable choice. *Millet* presents a similar pattern of adaptability, with strong positive coefficients at cooler to moderate temperatures (e.g., +17.7 at 14–15°C and +15.1 at 20–21°C). Yet, coefficients decline steeply under hotter conditions (–25.0 at 27–28°C), indicating that while millet is widely recognized as a heat- and drought-tolerant crop (Rai et al., 1999), its resilience has limits when exposed to extreme heat. Farmers' crop choices thus reflect this balance: millet is favored under moderate warming but less likely to be planted under very high temperatures.

Figure 12: Effects of temperature bins on probabilities of planting the crop



By contrast, *Teff* exhibits only small and statistically insignificant coefficients (between –1.9 and +2.3), suggesting that temperature does not substantially alter farmers' decisions to cultivate it. This limited responsiveness likely reflects teff's status as a culturally embedded staple, grown across Ethiopia irrespective of marginal temperature variations. Its localized importance explains its weak odds-ratio signals in the model. Finally, *Maize* shows consistently negative coefficients across temperature ranges (–0.6 to –1.5), implying reduced odds of being planted as temperatures rise. This pattern is coherent with maize's agronomic profile: as a space- and input-demanding crop, it is typically grown in monocultures and less compatible with intercropping systems (Giller et al., 2004). The negative coefficients suggest that farmers avoid maize under hotter conditions, likely due to its high water and nutrient requirements, which limit adaptability in mixed farming systems.

Fruits

For *avocado*, coefficients are strongly negative across nearly all temperature bins, reaching –35.6 log-odds at 29–30°C (Table 17). This translates into a sharp reduction in the odds of farmers planting avocado as temperatures rise. The result is fully coherent with its known sensitivity to heat stress, which above 28°C reduces fruit set and increases flower abortion (Whiley et al., 2002). Agronomically, avocado orchards are managed as monocultures with dense canopies and deep roots that compete heavily for water and nutrients, leaving little scope for intercropping (Wolstenholme, 2011). The negative coefficients therefore capture both biological and management constraints, making the model results ecologically justified.

Banana also shows uniformly negative coefficients, ranging from –13 to –18 log-odds, indicating consistently lower odds of intercropping or co-planting under all temperature conditions. As a perennial with continuous canopy cover and high water demand, banana dominates land use and suppresses other crops. This agronomic reality, documented in tropical systems (Robinson & Galán Saúco, 2010), explains why bananas are rarely rotated or integrated into diversified systems, except in smallholder subsistence settings.

By contrast, *mango* displays strongly positive coefficients across all temperature bins, increasing from +14.1 to +19.9 log-odds. In odds-ratio terms, this means farmers are much more likely to cultivate mango under warmer conditions. This is consistent with its heat tolerance and its common role in agroforestry systems in Africa and South Asia (Singh & Chadha, 2001). Its open canopy allows light penetration, while its deep roots reduce surface competition, making it more compatible with understory crops. Furthermore, mango trees improve soil microclimate and water retention (Nair, 1993), which explains their positive association with planting probabilities under warming scenarios.

Nuts

In Table 18, the results indicates a strong negative association between temperature and the likelihood of cultivating *peas*. Coefficients are consistently below –30 log-odds and reach –67.1 at 29–30°C, which translates into odds close to zero of farmers choosing peas under warm conditions. This pattern is agronomically coherent, as peas are cool-season crops that perform best at 10–20°C (Singh et al., 2018). Their incompatibility with tree nuts, which thrive in warmer and drier conditions, likely reflects both seasonal mismatches and competition for resources in mixed systems.

In contrast, *beans* show small and generally positive coefficients, ranging from +1 to +4 log-odds across most temperature bins. This suggests a modest increase in the odds of planting beans alongside nuts, though the effect is weak and inconsistent. This variability reflects the fact that beans, while legumes like groundnuts and capable of nitrogen fixation, have short growth cycles and flexible planting seasons, making their association with perennial nut systems less systematic. As highlighted by \textcolor{darkblue}{hyperref[ref:hussain2019]{Hussain et al., 2019)}, beans are commonly rotated or intercropped, but not always with long-cycle crops such as tree nuts, which explains the mixed evidence in the model.

Vegetables

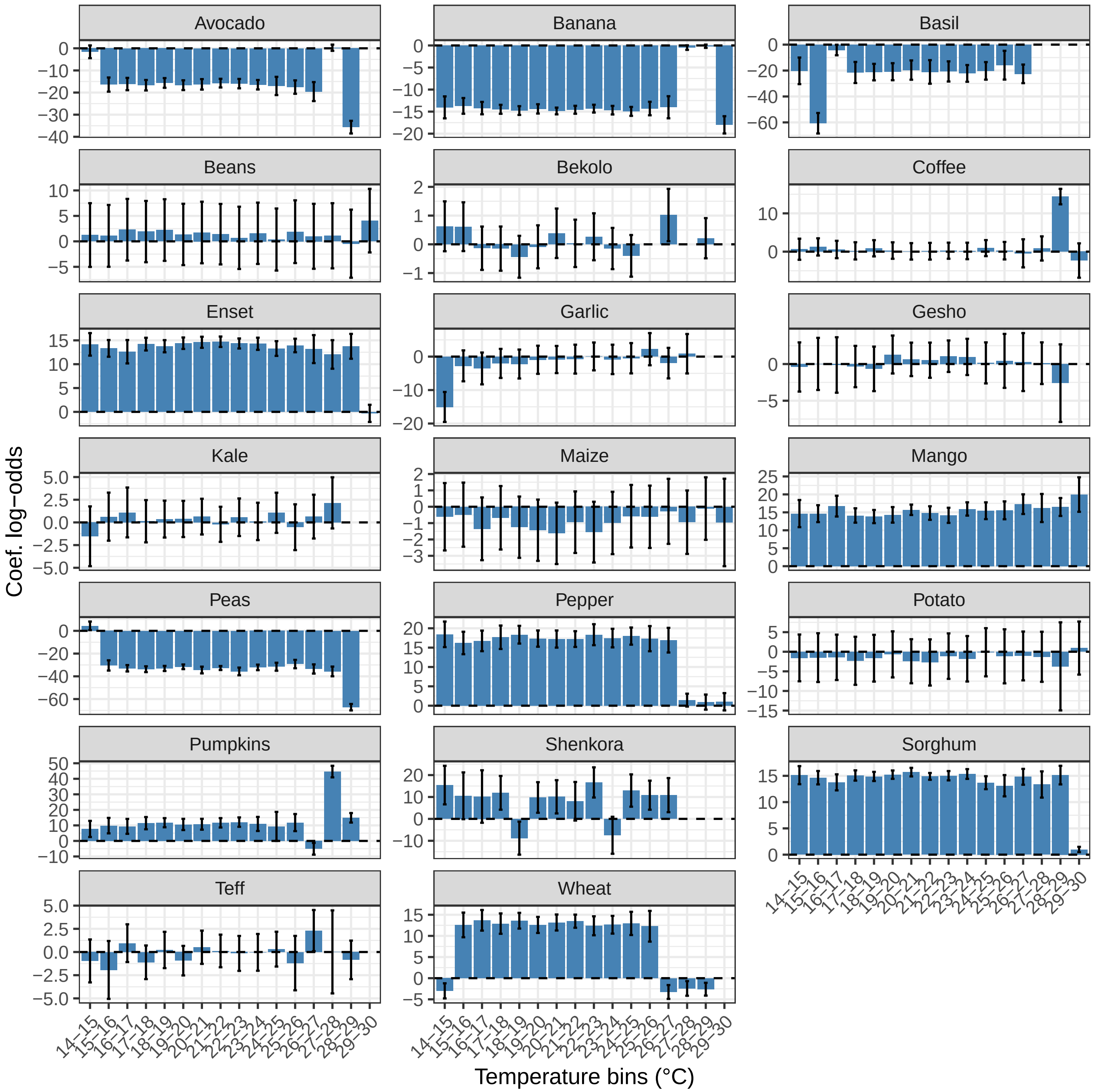
Shenkora exhibits consistently high positive log-odds between 14°C and 26°C, before declining moderately at higher temperatures. In odds ratio terms, farmers are substantially more likely to cultivate legumes alongside Shenkora within this range, confirming its agronomic compatibility. This is fully coherent with the optimal growth conditions for legumes, typically between 18°C and 26°C (FAO, 2021).

Pepper also shows uniformly positive coefficients across all temperature bins, indicating a robust increase in the probability of legume planting. This strong association reflects common intercropping practices where peppers and legumes complement each other through compatible growth cycles and shared climatic preferences (Choudhury et al., 2016). The consistent direction and magnitude of these coefficients strengthen the model's reliability.

By contrast, *Garlic* is associated with strongly negative coefficients across all temperature intervals, with log-odds as low as –15 (14–15°C), meaning the odds of co-planting legumes with garlic are extremely low. This outcome aligns with field knowledge: garlic is a nutrient- and moisture-demanding crop that competes heavily with other plants, making it unsuitable for rotation or intercropping with legumes (Singh et al., 2019).

Potato shows smaller but consistently negative effects across temperature bins, especially within the optimal range for legumes. The odds of co-planting legumes with potato are thus reduced, which reflects their overlapping planting seasons and potato's sensitivity to heat stress above 25°C (Haverkort et al., 2013). *Pumpkins* display an overall positive trend, suggesting compatibility with legumes, though an unusually high coefficient at 27–28°C (+44.65 log-odds) is likely a statistical artifact due to limited observations. While pumpkins are indeed adapted to warm conditions, this outlier requires cautious interpretation. Finally, *Kale* and *Bekolo* show coefficients close to zero without consistent patterns, indicating limited influence on legume planting decisions. Their marginal role in mixed systems may explain the weak associations, which are likely shaped by local and context-specific practices.

Figure 13: Effects of temperature bins on probabilities



Oils Seeds

For *Basil*, the crop choice model shows strongly negative log-odds across all temperature bins, with the largest declines between 15°C and 17°C. This means that the probability of oilseeds being planted alongside Basil is very low, consistent with its agronomic profile. Basil is a shallow-rooted, moisture-demanding crop that competes for nutrients and water, making it unsuitable for association with large-scale oilseed cultivation. This is in line with field observations: Basil is generally cultivated in home gardens or intensive intercropping systems rather than in rotation with oilseeds (Ikerd (2000); Wakjira et al., 2018). The negative coefficients are therefore both statistically robust and ecologically justified.

In contrast, *Enset* displays consistently strong positive coefficients across nearly all temperature intervals, indicating that households are significantly more likely to grow oilseeds in areas where Enset is present. This is coherent with Ethiopian farming systems, where Enset is cultivated as a perennial crop with deep roots that enhance soil structure and water retention Brandt et al., 1997. Farmers frequently combine annual crops—such as legumes and oilseeds—with Enset-based plots to benefit from these soil improvements. This relationship has been widely recognized as a pathway toward sustainable intensification in the Ethiopian highlands Borrell et al., 2020.

Coffee, however, presents mixed and less coherent results. For most temperature bins, the coefficients are weak, but at 29–30°C the model shows a sudden strong positive effect. Given that coffee is highly sensitive to heat stress, with quality deteriorating above 26°C DaMatta et al., 2007, this spike is unlikely to represent a genuine agronomic response. Instead, it may reflect regional substitution patterns: in overheated zones, farmers could replace coffee with oilseeds, which the model captures as a statistical association rather than true co-location. This highlights the need for cautious interpretation, especially when confidence intervals are wide or data is sparse at high-temperature bins. Finally, *Gesho* shows weak and scattered coefficients across all temperature bins. As Gesho is typically grown as boundary hedges or for traditional brewing, its role in shaping oilseed cultivation is minimal. The weak effects are therefore consistent with its marginal agronomic role, making the model's output both plausible and expected.

4.4 Limitation and Transferability

While this study provides robust evidence on the heterogeneous impacts of climate variability and the role of crop diversification, several limitations should be acknowledged. First, the climate variables are derived from gridded datasets rather than local weather station data, which may introduce spatial interpolation errors, particularly in areas with sparse observational coverage. In addition, household-reported yields are subject to recall bias and measurement errors, which could affect the precision of the estimated coefficients. Second, the potential for omitted variable bias remains, as certain unobserved factors such as farm management practices, input quality, or access to credit—may influence both climate exposure and crop performance. Although the fixed-effects framework controls for time-invariant household characteristics, it cannot fully eliminate such concerns. Third, the results are context-specific to Ethiopian smallholder systems, and caution is warranted when generalizing them to other agroecological zones or countries with different market structures and institutional settings. Finally, the analysis captures seasonal and annual variations but does not fully address intra-seasonal dynamics, which can be critical for climate-sensitive crops. Future research could mitigate these limitations by incorporating higher-resolution meteorological data, applying instrumental variable techniques to address potential endogeneity, conducting sensitivity analyses with alternative datasets, and expanding the scope to include comparative studies across regions.

Although the empirical analysis is firmly rooted in the Ethiopian smallholder context, the findings hold broader relevance for other developing countries where rain-fed agriculture dominates. The observed crop-specific, threshold-based climate response functions provide a replicable framework that can be applied to similar highland tropical and subtropical settings in East Africa, South Asia, and parts of Latin America. For instance, the strong temperature sensitivity of crops such as teff, garlic, and chat parallels findings for wheat, onions, and cash crops in other mid-altitude regions. Likewise, the documented benefits of diversification in buffering yield losses align with global evidence from smallholder systems in Kenya, Nepal, and Peru. However, differences in market access, infrastructure, and institutional support can mediate the magnitude of these effects. Therefore, while the methodological approach is transferable, policy prescriptions must be adapted to local agroecological and socioeconomic conditions. Embedding this framework into cross-country comparative studies could further enhance its generalizability and inform climate-resilient agricultural strategies at a global scale.

5 Conclusion and Policy Implication

Contributions to the Literature and Methodology

This study makes significant contributions to the empirical and methodological understanding of climate change impacts on smallholder agriculture. To my knowledge, this study is in range of the rarely which aims to produce crop-specific, threshold-based climate response functions for over twenty staple, cash, and horticultural crops in Ethiopia, using a decade-long nationally representative household–crop–year panel dataset. This level of disaggregation goes beyond the aggregated cereal-focused approaches common in Sub-Saharan African studies, enabling precise identification of both beneficial and harmful climate thresholds for each crop. Unlike most previous Sub-Saharan African studies that rely on aggregated crop categories, this thesis disaggregates the analysis to the individual crop level, enabling the identification of nonlinear, crop-specific temperature and precipitation thresholds. Methodologically, it integrates a multi-margin framework to capture intensive (yield), economic (revenue), and extensive (crop choice) margins, into a behavioral model that treats diversification as an endogenous response to climate stress rather than a static characteristic. This approach departs from much of the existing literature, which often models diversification as an exogenous or control variable (e.g.,Bezabih & Di Falco, 2012 ; Asfaw et al., 2016). Furthermore, the use of semi-parametric temperature-bin models with 1°C increments allows the estimation of smooth nonlinear response curves without imposing restrictive functional forms. In the Ethiopian highland context, the evidence clearly shows that adaptation strategies must be tailored not only to regional agroecological differences but also to the specific crop portfolios of farmers. This underlines that a uniform, one-size-fits-all approach would overlook critical local vulnerabilities and opportunities.

key findings and insights

Climate change poses one of the most immediate threats to agricultural systems globally, with particularly severe implications for smallholder farmers in rain-fed economies such as Ethiopia. This thesis was motivated by the absence of fine-grained, household-level evidence on how temperature and precipitation variations affect agricultural performance and how farmers adapt, deliberately or under constraint. Using five waves of household–crop–year data, the analysis employed fixed-effects econometric models with both linear and nonlinear climate variables, capturing within-household variation while controlling for unobserved characteristics. By embedding diversification in a behavioral framework, the study measured not only the direct effects of climate variables but also the adaptive strategies they induce. The results show marked heterogeneity in climate sensitivity across crops. Moderate warming benefits crops such as rice (about +47% yield per +1°C within optimal ranges), sorghum (+38%), and millet (+33%), reflecting Ethiopia's relatively cool baseline. Conversely, teff (−8% per +1°C beyond optimal conditions) and garlic (−17%) suffer significant losses under higher temperatures. On the revenue side, chat (khat) declines sharply (−14% per +1°C), while pepper revenues rise by approximately 12%. Rainfall effects are smaller in magnitude but significant for certain crops: a 10% increase in precipitation raises revenues for sorghum and pepper by 0.5–0.6% yet reduces chat revenues by about 1%. Across multiple crops, significant quadratic temperature effects confirm the threshold-bound nature of climate–productivity relationships, where warming benefits quickly reverse beyond critical limits.

The crop choice model reveals clear and differentiated adaptation patterns across cereals, fruits, legumes, and oilseeds. Sorghum and millet stand out as the most climate-resilient cereals, with large positive odds-ratios indicating that farmers are significantly more likely to cultivate them under hotter and drier conditions. Wheat shows adaptability within a narrow thermal window, but becomes less viable once heat thresholds are exceeded, while maize consistently declines in probability of cultivation, reflecting its high input and water demands. Teff, by contrast, displays little responsiveness to climate variables, underscoring its role as a culturally embedded staple grown independently of marginal climate fluctuations. Among fruits, mango supports mixed cropping under warmer conditions, while avocado and banana strongly suppress diversification due to their canopy dominance and water requirements. For legumes, Shenkora and pepper reinforce diversification under moderate climates, whereas garlic and potato reduce cultivation probabilities due to competition for soil nutrients and overlapping growing seasons. Finally, perennial crops such as enset emerge as positive enablers of diversification, while coffee and basil show mixed or suppressive effects, reflecting their sensitivity to heat and resource competition.

Taken together, these findings highlight three key insights: (1) smallholders adapt to climate stress not only through yield responses but also by reallocating land and labor across crops; (2) climate-resilient crops such as sorghum, millet, mango, and enset play a strategic role in household risk management; and (3) vulnerability is concentrated in crops like wheat, maize, garlic, and coffee, which exhibit sharp declines under warming. Policy recommendations follow directly: Ethiopia's Climate-Smart Agriculture agenda should prioritize the dissemination of heat- and drought-tolerant crops, promote enset and mango-based agroforestry systems, and reduce reliance on climate-sensitive crops through targeted insurance, irrigation, and extension services. By linking empirical evidence to tailored interventions, these results provide actionable guidance for designing adaptation strategies that reflect both Ethiopia's agroecological diversity and farmers' adaptive behavior.

The panel analysis during drought years reveals that diversification significantly reduced yield losses, especially in low-rainfall zones. While the magnitude of this buffering effect varied by agroecological zone and market access, the overall pattern supports diversification as a practical risk management tool for Ethiopian smallholders. However, its benefits are uneven, varying by agroecological zone, market access, and the nature of climatic stress. High-value cash crops, for instance, respond strongly to both climatic constraints and market volatility, while staple cereals show greater thermal tolerance but remain sensitive to extreme events. These findings reinforce the view that adaptation is multi-dimensional, influenced by both biophysical and economic factors.

The evidence supports Ethiopia's Climate Resilient Green Economy strategy and aligns with the principles of Climate-Smart Agriculture. Policy recommendations include spatially targeted promotion of crops according to climatic thresholds, relocation of heat-sensitive species to cooler zones, scaling up of strategic diversification through improved seed varieties and strengthened extension services, and substantial investment in irrigation and water management to reduce dependence on increasingly erratic rainfall. By combining detailed empirical evidence with a behavioral lens on farmer decision-making, this study provides both actionable insights for Ethiopia and a replicable methodological framework for other smallholder-dominated economies facing similar climate risks.

References

1. Alemayehu, A., & Bewket, W. (2017). *Determinants of smallholder farmers' choice of coping and adaptation strategies to climate change and variability in the central highlands of Ethiopia*. Environmental Development, 24, 77-89. DOI

2. Altieri, M. A., Nicholls, C. I., Henao, A., & Lana, M. A. (2015). *Agroecology and the design of climate change-resilient farming systems*. Agronomy for Sustainable Development, 35, 869-890. DOI

3. Araya, A., & Stroosnijder, L. (2010). *Effects of rainfall variability on crop production in Tigray, northern Ethiopia*. Water Resources Research, 46(8). DOI

4. Asfaw, S., Pallante, G., & Palma, A. (2021). *Diversification strategies and climate shocks in Ethiopia: A panel data analysis*. Climate and Development, 13(1), 1-14. DOI PDF

5. Asseng, S., Ewert, F., Martre, P., et al. (2015). *Rising temperatures reduce global wheat production*. Nature Climate Change, 5(2), 143–147. DOI

6. Barrett, C. B., Reardon, T., & Webb, P. (2001). *Nonfarm income diversification and household livelihood strategies in rural Africa*. Food Policy, 26(4), 315-331. DOI PDF

7. Barrett, C. B., Reardon, T., & Webb, P. (2007). *Non-separable agricultural household models in theory and practice*. World Development, 29(10), 1715-1730.

8. Benin, S., Smale, M., Pender, J., Gebremedhin, B., & Ehui, S. (2004). *The economic determinants of cereal crop diversity on farms in the Ethiopian highlands*. Agricultural Economics, 31(2-3), 197-208. PDF

9. Bewket, W. (2009). *Rainfall variability and crop production in Ethiopia: Case study in the Amhara Region*. Ethiopian Journal of Development Research, 31(1), 1-22.

10. Borrell, J. S., et al. (2020). *Enset in Ethiopia: A neglected starch crop across diverse agroecological zones*. Plant Genetic Resources, 18(2), 143–154. DOI

11. Brandt, S. A., Spring, A., Hiebsch, C., et al. (1997). *The Tree Against Hunger: Enset-Based Agricultural Systems in Ethiopia*. American Association for the Advancement of Science.

12. Burke, M., Hsiang, S. M., & Miguel, E. (2015). *Global non-linear effect of temperature on economic production*. Nature, 527(7577), 235-239. DOI PDF

13. Cameron, A. C., & Miller, D. L. (2015). *A practitioner's guide to cluster-robust inference*. Journal of Human Resources, 50(2), 317-372. DOI PDF

14. Choudhury, P. R., et al. (2016). *Intercropping systems involving pulses: A review*. Indian Journal of Agronomy, 61(S1), S1-S12.

15. DaMatta, F. M., et al. (2007). *Impacts of climate changes on crop physiology and food quality*. Field Crops Research, 103(3), 239–247. DOI

16. Deressa, T. T., Hassan, R. M., Ringler, C., Alemu, T., & Yesuf, M. (2008). *Analyzing the determinants of farmers' choice of adaptation methods*. IFPRI Discussion Paper 00798. PDF

17. Deschênes, O., & Greenstone, M. (2011). *Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US*. American Economic Journal: Applied Economics, 3(4), 152-185. DOI PDF

18. Dell, M., Jones, B. F., & Olken, B. A. (2014). *What Do We Learn from the Weather? The New Climate-Economy Literature*. Journal of Economic Literature, 52(3), 740-798. DOI PDF

19. Di Falco, S., Veronesi, M., & Yesuf, M. (2011). *Does adaptation to climate change provide food security? A micro-perspective from Ethiopia*. American Journal of Agricultural Economics, 93(3), 825-842. DOI PDF

20. Di Falco, S., & Chavas, J.-P. (2009). *On crop biodiversity, risk exposure and food security in the highlands of Ethiopia*. American Journal of Agricultural Economics, 91(3), 599-611. PDF

21. Diro, S., Tesfaye, A., & Erko, B. (2022). *Determinants of adoption of climate-smart agricultural technologies and practices in the coffee-based farming system of Ethiopia*. Agriculture & Food Security, 11, 42. DOI PDF

22. Duflo, E., Kremer, M., & Robinson, J. (2011). *Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya*. American Economic Review, 101(6), 2350-2390. DOI

23. El-Otmani, M., Coggins, C. W., Agustí, M., & Lovatt, C. J. (2011). *Plant management in citrus*. In Citrus: Botany, Production and Uses. CABI.

24. FAO. (2010). *Groundnut Production Guide*. Food and Agriculture Organization.

25. FAO (2010). *Ecocrop database – Cucurbita spp.* Food and Agriculture Organization.

26. FAO. (2010). *Teff Production Manual*. Food and Agriculture Organization.

27. FAO. (2013). *Climate-Smart Agriculture Sourcebook*. Food and Agriculture Organization of the United Nations. PDF

28. FAO. (2016). *The State of Food and Agriculture: Climate Change, Agriculture and Food Security*. PDF

29. FAO. (2021). *Legumes Production Guidelines*. Food and Agriculture Organization of the United Nations.

30. Federal Democratic Republic of Ethiopia. (2011). *Climate Resilient Green Economy Strategy*. PDF

31. Funk, C., Peterson, P., Landsfeld, M., et al. (2015). *The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes*. Scientific Data, 2:150066. DOI PDF

32. Gebissa, E. (2010). *Khat in Ethiopia: Taking the place of cereal?* Northeast African Studies, 13(2), 1–26. DOI

33. Gebremedhin, B., & Swinton, S. (2003). *Investment in soil conservation in northern Ethiopia: The role of land tenure security and public programmes*. Agricultural Economics, 29(1), 69-84. PDF

34. Giller, K. E., Witter, E., Corbeels, M., & Tittonell, P. (2004). *Integrated nutrient management in smallholder farming systems of Africa: agronomic challenges and economic opportunities*. Food Security and Soil Quality.

35. Hausman, J. A., & Taylor, W. E. (1981). *Panel data and unobservable individual effects*. Econometrica, 49(6), 1377-1398. DOI PDF

36. Haverkort, A. J., et al. (2013). *Crop production and adaptation strategies for climate-resilient potatoes*. Potato Research, 56(4), 225–247. DOI

37. Hussain, M. I., et al. (2019). *Legume intercropping and soil fertility: a review of recent evidence*. Soil and Tillage Research, 194, 104330. DOI

38. Ikerd, J. E. (2000). *Small Farms and Sustainable Development: Is Small Still Beautiful?* University of Missouri.

39. IPCC. (2021). *Climate Change 2021: The Physical Science Basis*. Sixth Assessment Report. Report

40. Just, R. E., & Pope, R. D. (1979). *Production function estimation and related risk considerations*. PDF

41. Kassie, M., Teklewold, H., Jaleta, M., Marenya, P., & Erenstein, O. (2014). *Understanding the adoption of a portfolio of sustainable intensification practices in eastern and southern Africa*. Land Use Policy, 42, 400-411. DOI

42. Krajewski, A. J., González-Molina, E., & García-Bastidas, F. A. (2019). *Citrus responses to temperature stress: physiological and agronomic perspectives*. Journal of Horticultural Research, 27(1), 17–25. DOI

43. Krause, M. R., et al. (2016). *Temperature and Light Effects on Basil Production*. HortScience, 51(6), 678–683.

44. Lin, B. (2011). *Resilience in Agriculture through Crop Diversification: Adaptive Management for Environmental Change*. BioScience, 61(3), 183-193. DOI PDF

45. Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). *Climate trends and global crop production since 1980*. Science, 333(6042), 616-620. DOI PDF

46. Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2012). *Climate trends and global crop production since 1980*. Science, 333(6042), 616–620. DOI

47. Nair, P. K. R. (1993). *An Introduction to Agroforestry*. Springer Science & Business Media.

48. Rai, K. N., Murty, D. S., Andrews, D. J., & Bramel-Cox, P. J. (1999). *Genetic enhancement of pearl millet and sorghum for the semi-arid tropics of Asia and Africa*. Genome, 42(4), 617–628. DOI

49. Robinson, J. C., & Galán Saúco, V. (2010). *Bananas and Plantains*. CABI.

50. Rosenow, D. T., & Dahlberg, J. A. (2000). *Sorghum improvement*. In: Agronomy Monograph 33. ASA-CSSA-SSSA, Madison, WI.

51. Sandmo, A. (1971). *On the theory of the competitive firm under price uncertainty*. PDF

52. Schlenker, W., & Roberts, M. J. (2009). *Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change*. Proceedings of the National Academy of Sciences, 106(37), 15594-15598. DOI PDF

53. Schlenker, W., & Lobell, D. B. (2010). *Robust negative impacts of climate change on African agriculture*. Environmental Research Letters, 5(1), 014010. PDF

54. Schlenker, W., & Lobell, D. B. (2011). *Climate trends and global crop production since 1980*. Science, 333(6042), 616-620. DOI PDF

55. Seo, S. N., & Mendelsohn, R. (2008). *A structural Ricardian analysis of climate change impacts and adaptations in African agriculture*. World Bank Policy Research Working Paper. Link

56. Singh, I., Squire, L., & Strauss, J. (1986). *Agricultural Household Models: Extensions, Applications, and Policy*. Johns Hopkins University Press.

57. Singh, L. B., & Chadha, K. L. (2001). *The Mango: Botany, Production and Uses*. International Book Distributing Company.

58. Singh, B., et al. (2018). *Cool-season legumes: agronomy and challenges under climate change*. Journal of Agronomy and Crop Science, 204(5), 487–501. DOI

59. Singh, R. P., et al. (2019). *Soil fertility and nutrient management in garlic-based cropping systems*. Journal of Soil Science and Plant Nutrition, 19(3), 567–581. DOI

60. Snapp, S. S., Blackie, M. J., Gilbert, R. A., Bezner-Kerr, R., & Kanyama-Phiri, G. Y. (2010). *Biodiversity can support a greener revolution in Africa*. Proceedings of the National Academy of Sciences, 107(48), 20840-20845. DOI

61. Teklewold, H., Kassie, M., Shiferaw, B., & Köhlin, G. (2013). *Cropping system diversification and conservation tillage in Ethiopia*. Environment for Development Discussion Paper. PDF

62. Tesfaye, K., et al. (2015). *Climate variability and change in Ethiopia: Challenges and opportunities*. Journal of Climate Risk and Resilience, 7(1), 56-72.

63. Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., & Polasky, S. (2002). *Agricultural sustainability and intensive production practices*. Nature, 418(6898), 671-677. DOI PDF

64. Tversky, A., & Kahneman, D. (1974). *Judgment under uncertainty: Heuristics and biases*. Science, 185(4157), 1124-1131. PDF

65. Wakjira, M. T., et al. (2018). *Assessment of Intercropping Practices in Southern Ethiopia*. Ethiopian Journal of Agricultural Sciences, 28(3), 1-15.

66. Walker, B., Holling, C. S., Carpenter, S. R., & Kinzig, A. (2004). *Resilience, adaptability and transformability in social-ecological systems*. Ecology & Society, 9(2), 5. PDF

67. Whiley, A. W., Schaffer, B., & Wolstenholme, B. N. (2002). *Avocado: Botany, Production and Uses*. CABI.

68. Wolstenholme, B. N. (2011). *Avocado production systems and challenges in the tropics and subtropics*. South African Avocado Growers' Association Yearbook.

69. Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press. Link

70. World Bank. (2020). *Ethiopia Economic Update: Strengthening Agricultural Markets for Resilient Growth*. Washington, DC: World Bank. Link

71. World Bank. (2022). *World Development Indicators*. Database

72. Yoshida, S. (1981). *Fundamentals of Rice Crop Science*. International Rice Research Institute.

Appendix

APPENDIX A: Additionnal descriptive Statistics

Table 7: Data Quality Assessment for Agricultural Production

Case	Area (ha)	Quantity (kg)	Revenue (BIRR)	Interpretation	Action Taken
1	0 or NA	0 or NA	0 or NA	No crop activity	Drop from dataset
2	0 or NA	>0	>0	Missing area, but harvest and income reported	Keep but flag
3	0	0	>0	Revenue with no harvest	Data inconsistency - flag for review
4	0	>0	0	Harvest without income	Self-consumption or post-harvest loss - keep
5	0	>0	NA	Harvest with missing income	Non-commercial production - keep
6	0	0	0	Valid zero observation (no production)	Keep
7	NA	>0	NA or 0	Harvest but no area or income	Small-scale production - keep
8	NA	NA	>0	Income without production	Suspicious - exclude
9	NA	>0	>0	Production without area	Plausible but flag
10	> 0	0	>0	Area and revenue but no quantity	Measurement error - impute or exclude
11	> 0	>0	NA	Missing revenue	Non-commercial - keep
12	NA	>0	>0	Missing crop name	Impute dominant crop

This table is build based on all the issues I encontered when cleaning the data.

Figure 14: Wave and presence of houshold

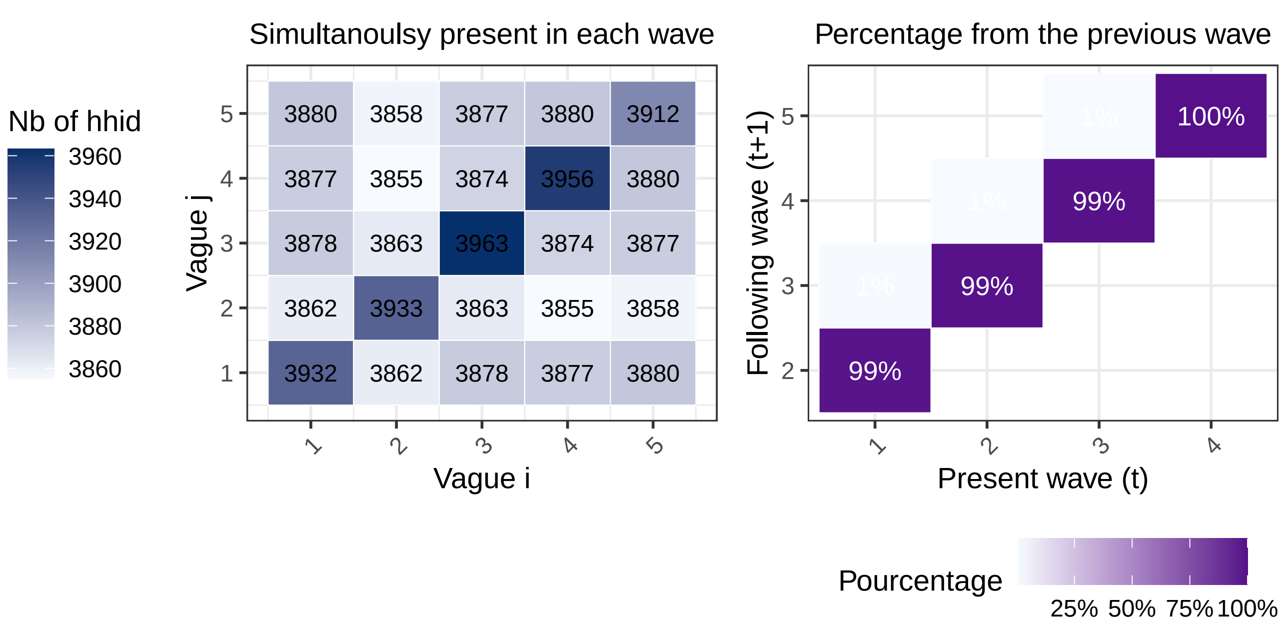


Figure 15: Distribution of the logarithm the Yield and households Incomes

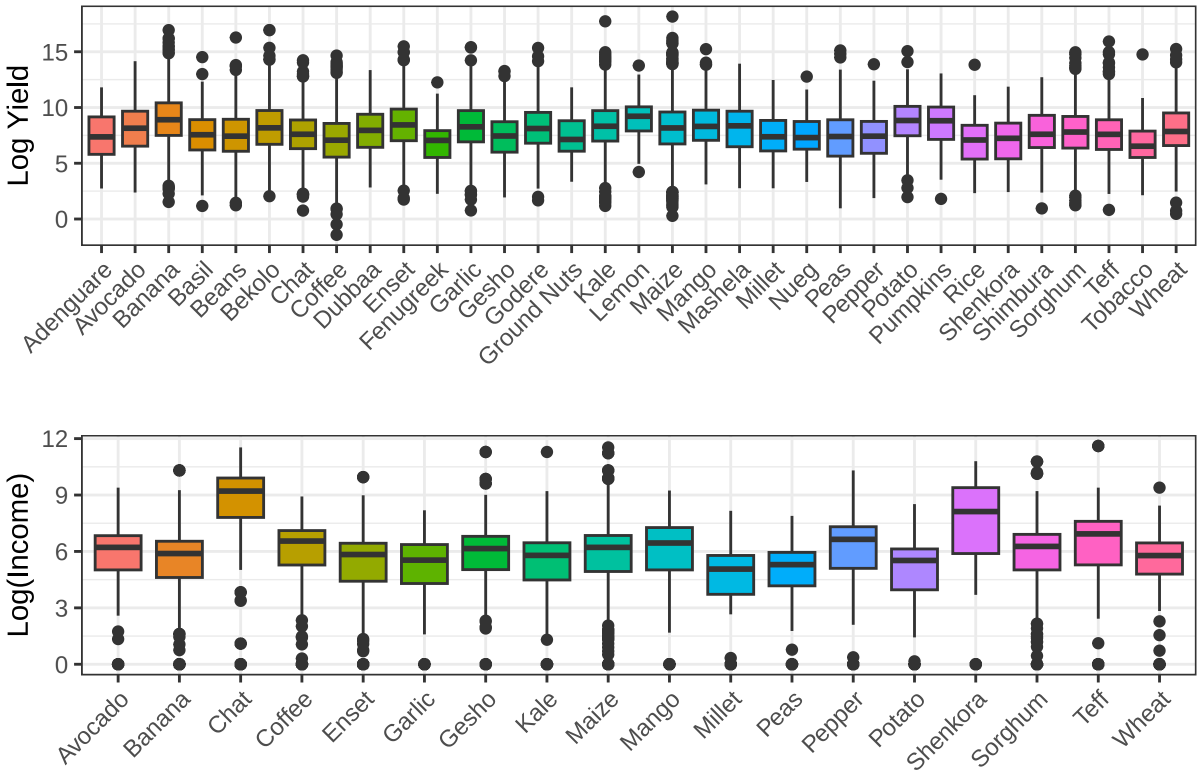


Table 8: Matrix of correlation between climates variables

Variable	Temp	Precip	precip²	Temp²
Temp	1.00	0.29	-0.01	0.24
Precip	0.29	1.00	0.03	0.02
Precip²	-0.01	0.03	1.00	0.04
Temp²	0.24	0.02	0.04	1.00

Figure 16: Possibilities of Missing values (R_{it}, S_{it}, Q_{it})

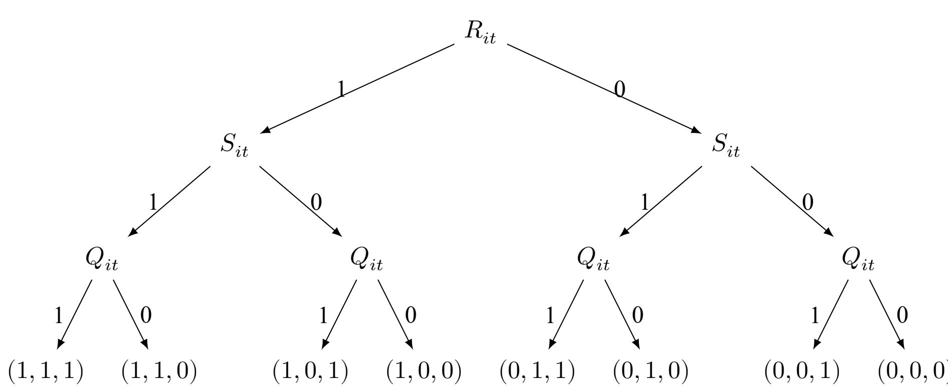


Table 9: Summary statistics by Wave

Variable	Wave 1				Wave 2				Wave 3				Wave 4				Wave 5				Global			
	Min	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max
Temperature	14.1	20.5	3.1	30.6	14.8	20.4	2.5	30.9	14.1	21.3	3.6	30.8	14.2	20.4	2.1	30.4	14.1	20.7	3.3	30.8	14.1	20.7	2.9	30.9
Precipitation	639.3	858.5	62.1	1,107.6	585.9	833.2	65.5	1,146.9	586.3	848.5	61.5	1,0161.5	617.4	869.8	68.9	1,104.9	604.5	838.6	64.5	1,068.8	585.9	849.7	65.9	1146.9
log(Yield)	0.28	6.54	5.21	8.54	0.46	7.55	2.02	10.55	0.75	7.78	1.91	8.78	-1.42	8.44	2.07	8.44	-0.49	7.89	2.43	9.89	-1.2	8.4	2.3	10.5
log(Crop revenue)	0	4.6	2.8	11.6	0	6.1	2.8	10.1	0	4.8	2.9	12.3	0	6.2	1.8	11.4	0	5.3	2.6	13	0	103	80.4	311

Table 10: Percentage of Zero Yields by Crop and Wave

Crop	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Avocado	6.7	15.6	6.6	17.3	6.7
Banana	9.7	14.5	9.9	17.6	7.9
Chat	8.2	10.7	6.8	19.9	4.6
Coffee	10.6	19.8	9.2	16.0	4.7
Enset	8.4	13.5	13.8	21.4	8.3
Garlic	6.1	20.1	15.0	24.6	11.2
Gesho	9.9	15.3	9.9	18.7	6.7
Kale	12.1	16.7	14.2	26.9	6.9
Maize	11.3	19.0	13.1	22.4	7.1
Mango	8.5	19.8	14.5	20.8	13.0
Millet	7.5	26.3	21.2	22.9	8.6
Peas	7.5	19.4	7.4	16.4	7.9
Pepper	24.0	22.1	11.7	30.0	11.3
Potato	15.4	21.3	14.5	30.5	9.6
Shenkora	0.0	18.2	16.7	35.3	9.4
Sorghum	11.5	18.3	15.3	24.6	8.2
Teff	9.1	16.4	14.9	24.5	4.3
Wheat	11.0	13.2	8.8	14.9	9.6
Total Pct.by wave	22.40	11.40	21.30	5.30	17.3

Climate Variables (Temperature and Precipitation) in the Study Scope by Wave

Figure 17: Cereal and Fruits Crops in the Scope

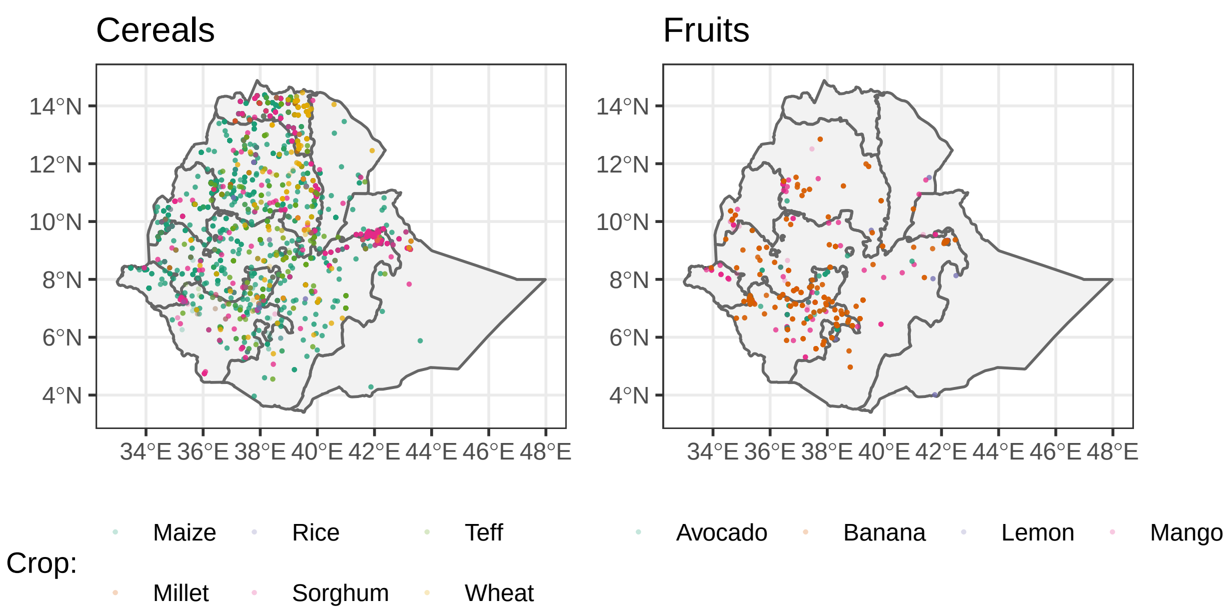


Figure 18: Nuts and Legums Crops in the Scope

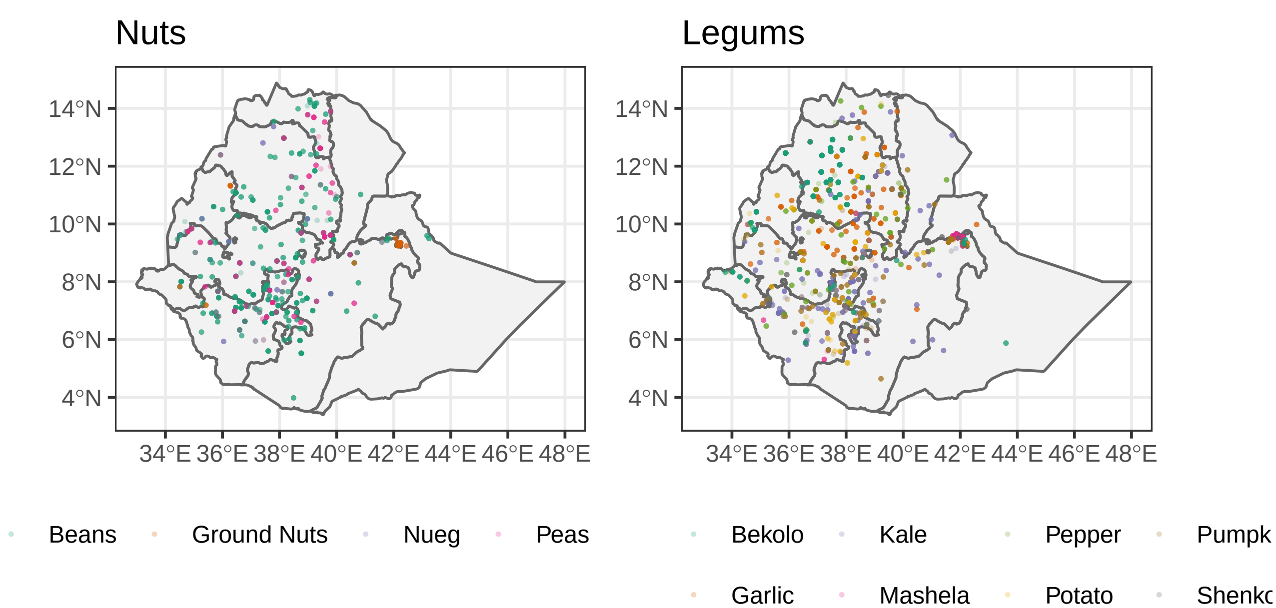


Figure 19: Oil Seed Crops in The scope

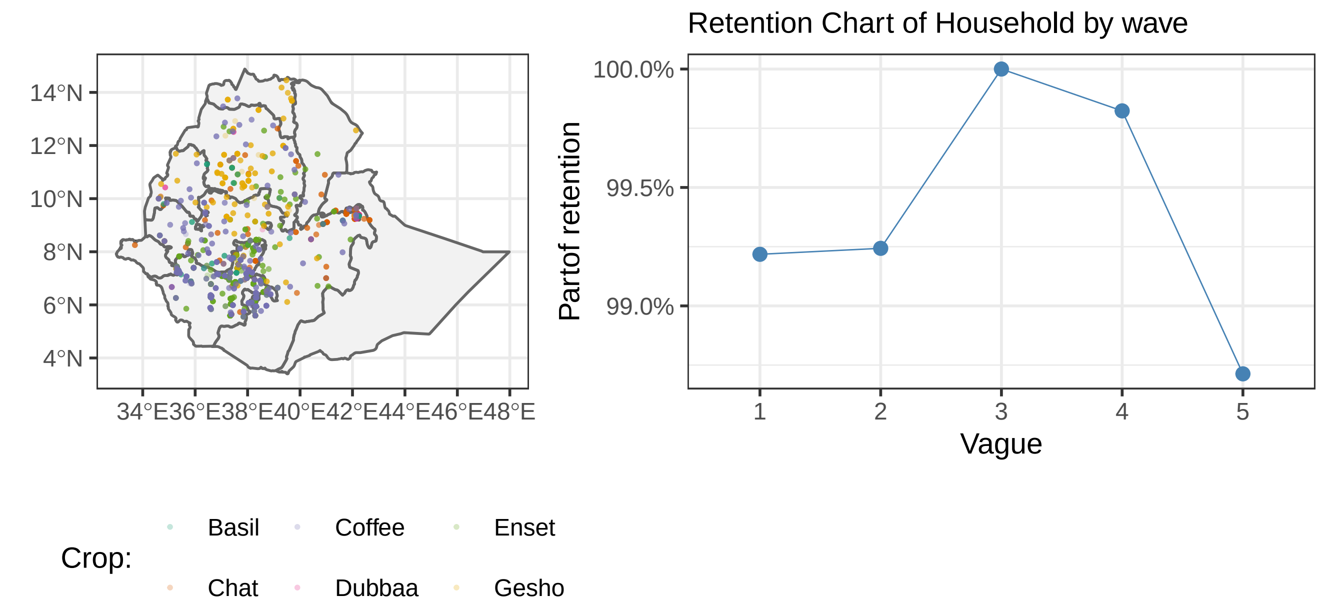
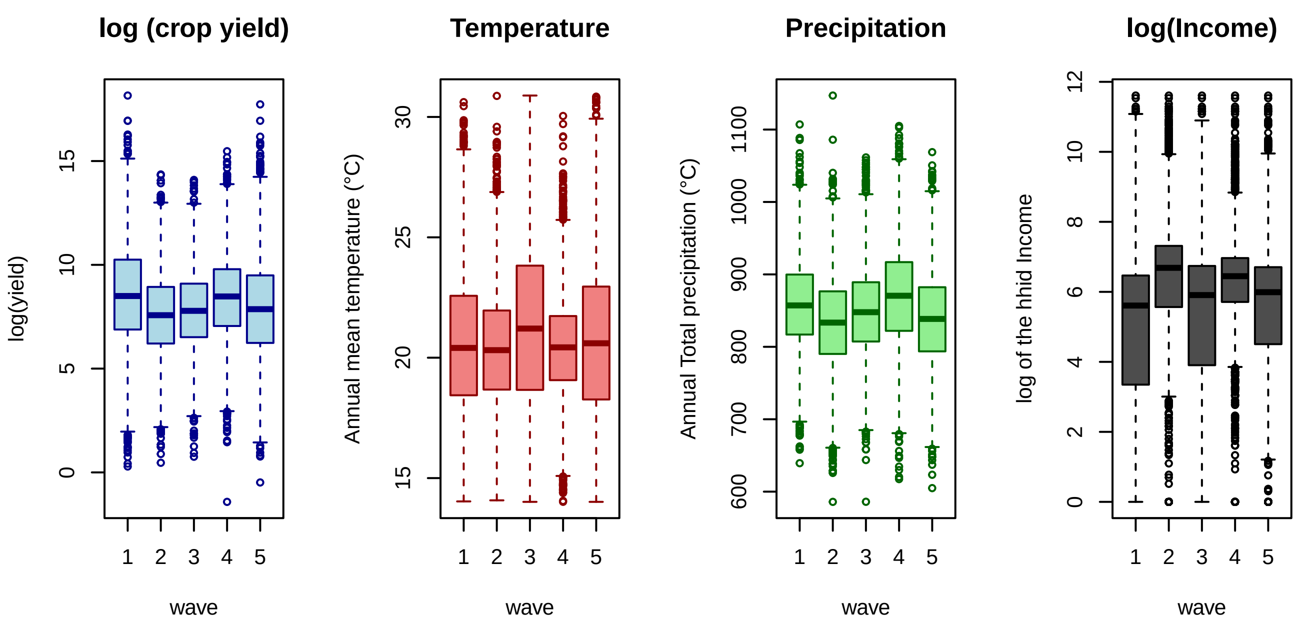
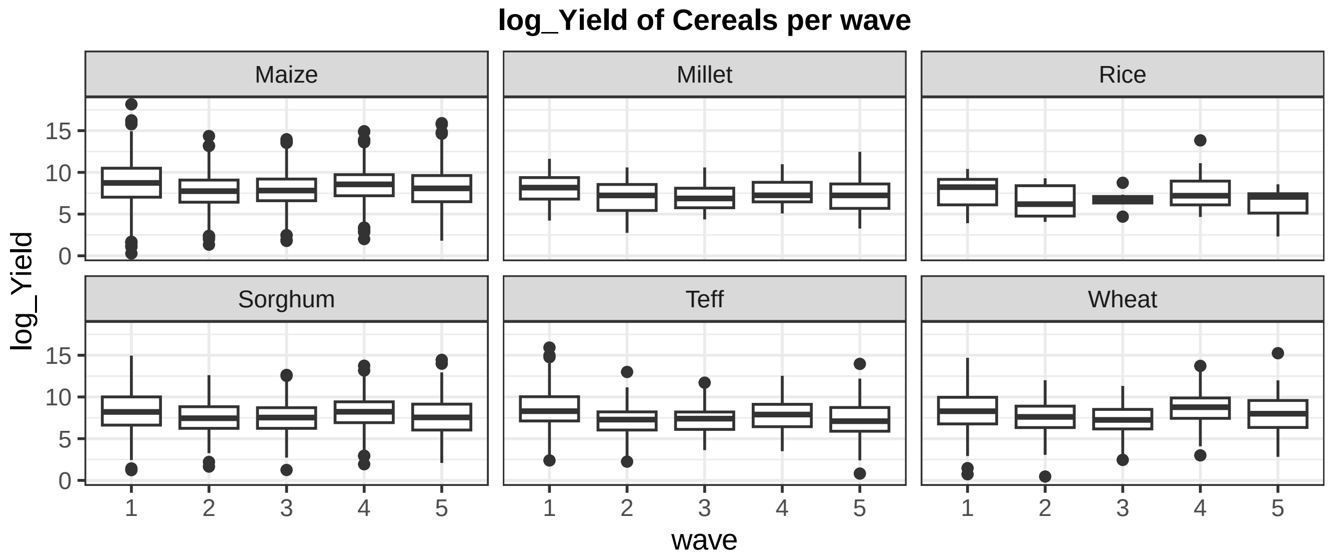
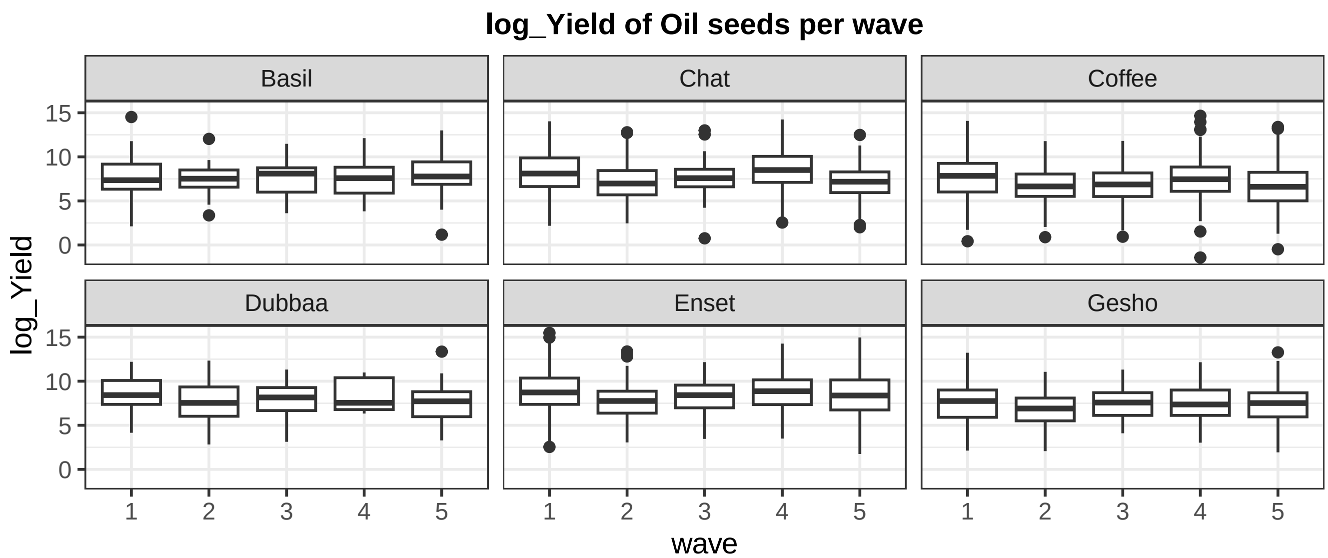


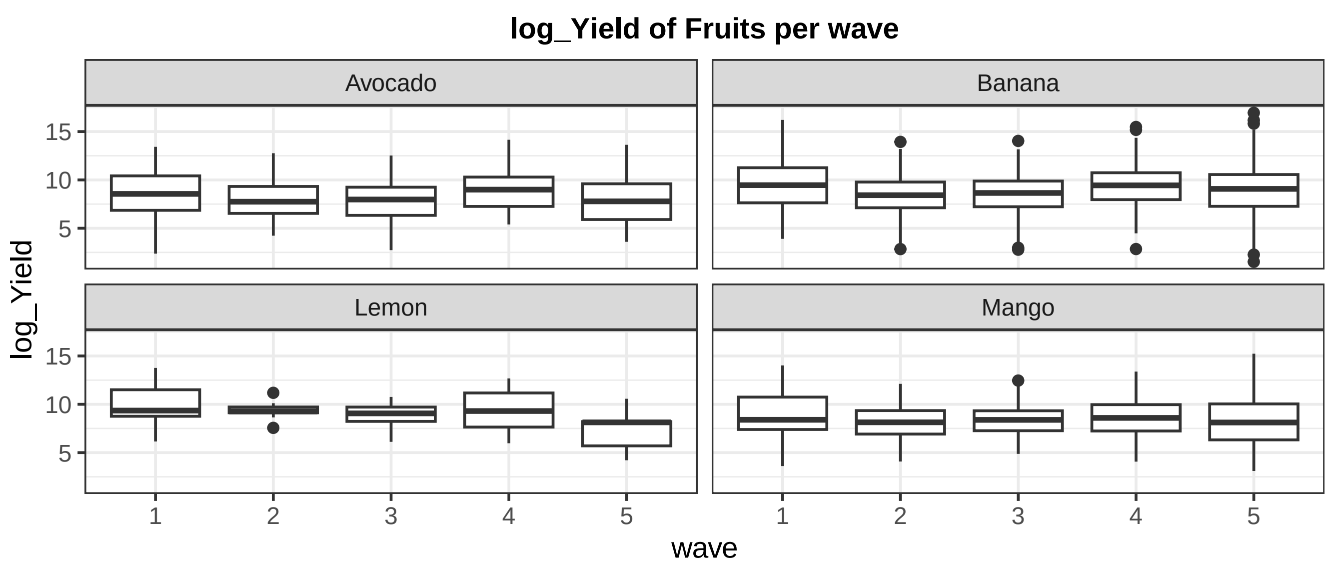
Figure 20: Distribution of the logarithme the main variables by wave

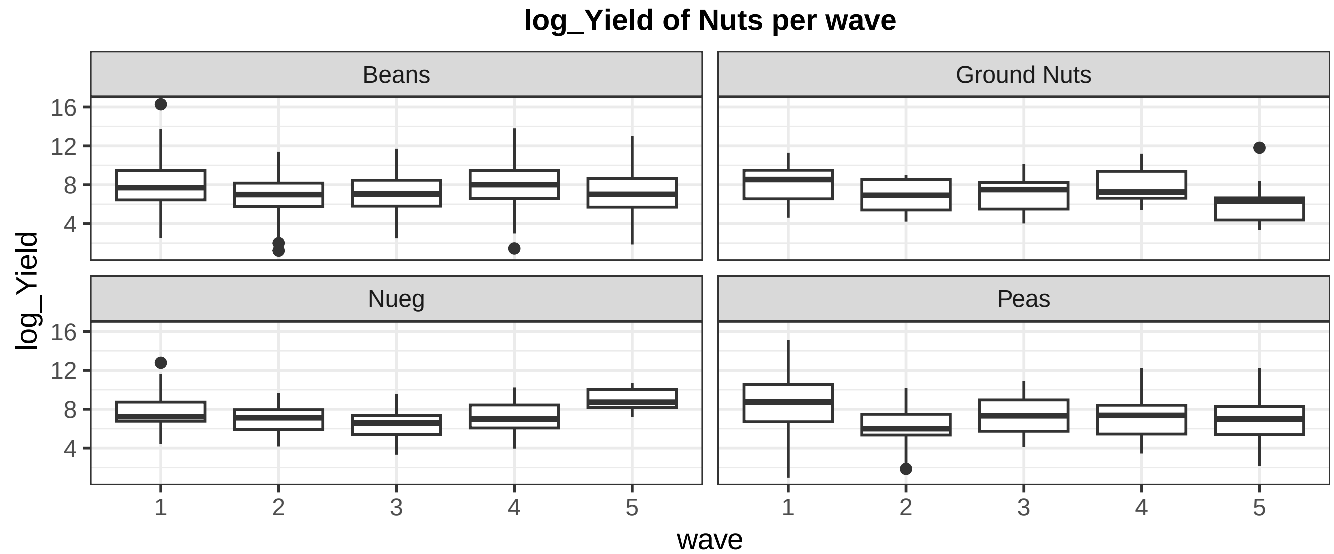


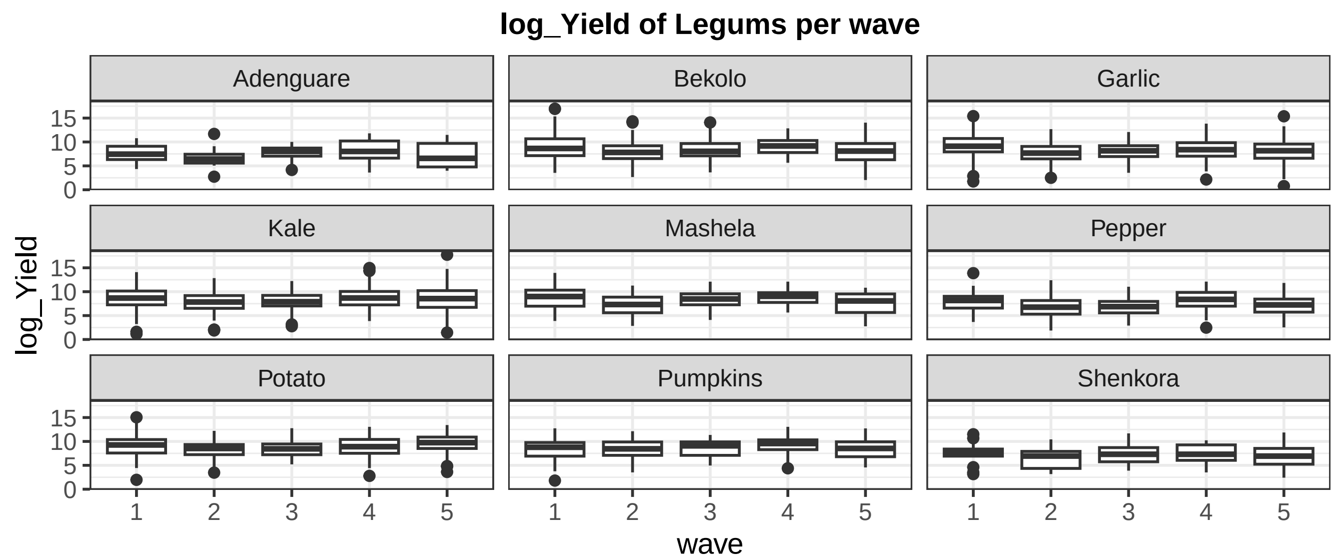
Distibution of each crop yield per wave in The Scope











Steps of Models estimation

Formal equation for the graphics

The different graphics in the intensive margin analysis stand for each crop $c \in C$, for temperature bin $k \in \{14, \dots, 29\}$:

$$f_c(T_t) = \sum_{k=14}^{29} \beta_{ck} \cdot \mathbb{1}(T_t \in \text{bin}_k)$$

where $\mathbb{1}(\cdot)$ is the indicator function. – Degree-8 Polynomial

$$\hat{f}_p(x) = \sum_{i=0}^{8} \theta_i x^i$$

Estimated via ordinary least squares:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{t=1}^{n} \left(\log(Y_t) - \sum_{i=0}^{8} \theta_i T_t^i - \mathbf{X}_t \gamma \right)^2$$

– LOESS Smoothing Local polynomial regression at point x_0:

$$\hat{f}_l(x_0) = \hat{\beta}_0(x_0) + \hat{\beta}_1(x_0)(x_0 - x)$$

where $(\hat{\beta}_0, \hat{\beta}_1)$ minimize:

$$\sum_{t=1}^{n} K\left(\frac{T_t - x_0}{h}\right)(\log(Y_t) - \beta_0 - \beta_1(T_t - x_0))^2$$

with:

- $K(u) = (1 - |u|^3)^3$ (tricube kernel)
- $h = 0.2 \times (\max(T) - \min(T))$ (bandwidth)

Concerning the Confidence Intervals, for each temperature bin k:

$$CI_k = [\text{effect}_k - z_{0.975} \times SE_k,\ \text{effect}_k + z_{0.975} \times SE_k]$$

where $z_{0.975} \approx 1.96$ is the 97.5th percentile of the standard normal distribution. And the visualization components P_1zhich exhibits as

$$P_1 = \begin{cases} \textit{Step function:} & \{(k, \text{effect}_k)\}_{k=14}^{29} \\ \textit{Polynomial fit:} & \{(x, \hat{f}_p(x)) : x \in [14, 29]\} \\ \textit{LOESS smooth:} & \{(x, \hat{f}_l(x)) : x \in [14, 29]\} \\ \textit{Confidence band:} & \{(x, y) : x \in [14, 29], y \in [\text{effect}_x - 0.96SE_x, \text{effect}_x + 0.96SE_x]\} \end{cases}$$

and P_2 as

$$P_2 = \left\{(k, n_k) : n_k = \sum_{t=1}^{N} \mathbb{1}(T_t \in \text{bin}_k)\right\}_{k=14}^{29}$$

for the crops with valids data.

Estimation framework

From the model 1, we aim to estimate the coefficients $\beta_{1k}, \beta_{2k}, \gamma_{1k}, \gamma_{2k}$ while eliminating fixed effects. Let's define:

$$y_{ict} := \log(Y_{ict})$$

We denote the interaction variables:

$$x_{ict}^{(1k)} := T_{it} \cdot \mathbb{1}_{Crop_{ic}=k}, \quad x_{ict}^{(2k)} := T_{it}^2 \cdot \mathbb{1}_{Crop_{ic}=k}$$
$$x_{ict}^{(3k)} := P_{it} \cdot \mathbb{1}_{Crop_{ic}=k}, \quad x_{ict}^{(4k)} := P_{it}^2 \cdot \mathbb{1}_{Crop_{ic}=k}$$

Then the model becomes:

$$y_{ict} = \alpha_i + \lambda_t + \delta_c + \sum_k \left(\beta_{1k} x_{ict}^{(1k)} + \beta_{2k} x_{ict}^{(2k)} + \gamma_{1k} x_{ict}^{(3k)} + \gamma_{2k} x_{ict}^{(4k)}\right) + \varepsilon_{ict}$$

Let:

- $\mathbf{X}_{ict}$ = vector of all explanatory variables
- β = vector of all coefficients

Objective of the Fixed Effects Estimator

The terms $\alpha_i, \lambda_t, \delta_c$ are not observed, but must be removed to avoid omitted variable bias. To estimate β correctly, we apply the within transformation (demeaning approach) to eliminate these fixed effects.

– *Step 1: Take Time-Crop Mean for Each Household* For household i, crop c, define the average over time:

$$\bar{y}_{ic} = \frac{1}{T_{ic}} \sum_t y_{ict}, \quad \bar{x}_{ic}^{(lk)} = \frac{1}{T_{ic}} \sum_t x_{ict}^{(lk)} \text{ for each variable } x^{(lk)}$$

The average of the equation over time is:

$$\bar{y}_{ic} = \alpha_i + \lambda_t + \delta_c + \sum_k \left(\beta_{1k}\bar{x}_{ic}^{(1k)} + \beta_{2k}\bar{x}_{ic}^{(2k)} + \gamma_{1k}\bar{x}_{ic}^{(3k)} + \gamma_{2k}\bar{x}_{ic}^{(4k)}\right) + \bar{\varepsilon}_{ic}$$

– *Step 2: Subtract the Mean from Each Observation*

Subtracting the average equation from the original:

$$y_{ict} - \bar{y}_{ic} = (\alpha_i - \alpha_i) + (\lambda_t - \lambda_t) + (\delta_c - \delta_c) + \sum_k \beta_{1k}(x_{ict}^{(1k)} - \bar{x}_{ic}^{(1k)}) + \beta_{2k}(x_{ict}^{(2k)} - \bar{x}_{ic}^{(2k)}) + \gamma_{1k}(x_{ict}^{(3k)} - \bar{x}_{ic}^{(3k)}) + \gamma_{2k}(x_{ict}^{(4k)} - \bar{x}_{ic}^{(1k)})$$

$$+(\varepsilon_{ict} - \bar{\varepsilon}_{ic})$$

This simplifies to:

$$\tilde{y}_{ict} = \sum_k \beta_{1k}\tilde{x}_{ict}^{(1k)} + \beta_{2k}\tilde{x}_{ict}^{(2k)} + \gamma_{1k}\tilde{x}_{ict}^{(3k)} + \gamma_{2k}\tilde{x}_{ict}^{(4k)} + \tilde{\varepsilon}_{ict}$$

Where:

- $\tilde{y}_{ict} := y_{ict} - \bar{y}_{ic}$
- $\tilde{x}_{ict}^{(lk)} := x_{ict}^{(lk)} - \bar{x}_{ic}^{(lk)}$
- All fixed effects are now eliminated

Machine learning in bins selection in extensive Margin

Let $y_i \in \{0, 1\}$ be the binary response (crop presence/absence) and $\mathbf{x}_i \in \mathbb{R}^p$ the predictor vector (temperature bins, precipitation) for observation i. The log-likelihood is:

$$\ell(\beta) = \sum_{i=1}^{n} \left[y_i \log\left(\frac{e^{\mathbf{x}_i^T\beta}}{1 + e^{\mathbf{x}_i^T\beta}}\right) + (1 - y_i)\log\left(\frac{1}{1 + e^{\mathbf{x}_i^T\beta}}\right)\right]$$

where $\beta \in \mathbb{R}^p$ are the coefficients. *Ridge Regression (L2 Penalization)* Objective: Minimize log-likelihood with L2 penalty to handle multicollinearity:

$$\hat{\beta}^{\text{Ridge}} = \arg\min_\beta \left\{ -\ell(\beta) + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}$$

Key Properties:

- Shrinks coefficients toward zero but never exactly zero
- Effective when predictors are correlated (common in temperature bin data)
- Closed-form solution exists:

$$\hat{\beta}^{\text{Ridge}} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}$$

LASSO (L1 Penalization)

Objective: Add L1 penalty for sparse solutions (automatic feature selection):

$$\hat{\beta}^{\text{LASSO}} = \arg\min_{\beta} \left\{ -\ell(\beta) + \lambda \sum_{j=1}^{p} |\beta_j| \right\}$$

Key Properties:

- Produces exact zeros for irrelevant predictors
- No closed-form solution (requires coordinate descent)
- Superior for identifying critical temperature thresholds

Optimality Condition (KKT):

$$\left. \frac{\partial \ell}{\partial \beta_j} \right|_{\hat{\beta}_j} = \lambda \cdot \text{sign}(\hat{\beta}_j) \quad \forall j$$

Elastic Net (L1 + L2 Penalization)

Objective: Convex combination of L1 and L2 penalties:

$$\hat{\beta}^{\text{EN}} = \arg\min_{\beta} \left\{ -\ell(\beta) + \lambda \left[\alpha \sum_{j=1}^{p} |\beta_j| + \frac{1-\alpha}{2} \sum_{j=1}^{p} \beta_j^2 \right] \right\}$$

where $\alpha \in [0, 1]$ controls the L1/L2 mix.

Advantages: - Inherits stability from Ridge ($\alpha \to 0$) - Retains feature selection from LASSO ($\alpha \to 1$) - Ideal for grouped variables (e.g., multiple temperature bins affecting crops)

The optimal λ minimizes out-of-sample deviance via k-fold CV:

$$\lambda^* = \arg\min_{\lambda} \left\{ \frac{1}{K} \sum_{k=1}^{K} \text{Deviance}(\hat{\beta}_{\lambda}^{(k)}, \mathbf{y}_{\text{test}}^{(k)}) \right\}$$

where $\hat{\beta}_{\lambda}^{(k)}$ is the estimate excluding fold k.

Intensive Margin:linear and quadratic effects of ckimate variables

Table 11: The effect of climate on the crop log(yield)

	Temp. (lin)		**Temp. (quad)**		**Precip. (lin)**		**Precip. (quad)**	
Culture	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Adenguare	0.3153***	(0.0744)	0.0092	(0.0179)	0.0083**	(0.0026)	0.0000	(0.0000)
Avocado	0.5090***	(0.0385)	-0.0313***	(0.0088)	0.0017	(0.0018)	0.0000	(0.0000)
Banana	0.4196***	(0.0189)	-0.0095*	(0.0038)	0.0023***	(0.0007)	0.0000***	(0.0000)
Basil	0.4196***	(0.0460)	-0.0292*	(0.0141)	0.0070*	(0.0028)	0.0001$^{\cdot}$	(0.0000)
Beans	0.3752***	(0.0214)	-0.0056	(0.0064)	0.0037***	(0.0009)	0.0000***	(0.0000)
Bekolo	0.4458***	(0.0278)	-0.0111$^{\cdot}$	(0.0065)	0.0060***	(0.0013)	0.0000**	(0.0000)
Chat	0.5745***	(0.0358)	-0.0453***	(0.0066)	-0.0073***	(0.0016)	0.0001***	(0.0000)
Coffee	0.3824***	(0.0197)	-0.0071	(0.0051)	0.0031***	(0.0008)	0.0000***	(0.0000)
Dubbaa	0.1071$^{\cdot}$	(0.0549)	0.0061	(0.0108)	0.0074*	(0.0033)	0.0000$^{\cdot}$	(0.0000)
Enset	0.5310***	(0.0223)	-0.0128*	(0.0058)	0.0016*	(0.0007)	0.0000$^{\cdot}$	(0.0000)
Fenugreek	0.6056***	(0.0830)	-0.0532**	(0.0177)	-0.0020	(0.0033)	0.0000	(0.0000)
Garlic	0.5486***	(0.0320)	-0.0140	(0.0101)	0.0026$^{\cdot}$	(0.0014)	0.0000*	(0.0000)
Gesho	0.4867***	(0.0339)	-0.0194*	(0.0079)	0.0017	(0.0013)	0.0000	(0.0000)
Godere	0.4640***	(0.0348)	-0.0131	(0.0081)	0.0024$^{\cdot}$	(0.0014)	0.0000***	(0.0000)
GroundNuts	0.3987***	(0.0989)	-0.0102	(0.0269)	0.0051*	(0.0022)	-0.0000	(0.0000)
Kale	0.4852***	(0.0282)	-0.0172*	(0.0086)	0.0019*	(0.0009)	0.0000**	(0.0000)
Lemon	0.1876$^{\cdot}$	(0.1030)	0.0305	(0.0247)	0.0084	(0.0055)	0.0001	(0.0001)
Maize	0.2679***	(0.0084)	-0.0059**	(0.0020)	0.0049***	(0.0004)	0.0001***	(0.0000)
Mango	0.2587***	(0.0257)	0.0028	(0.0065)	0.0037*	(0.0015)	0.0001***	(0.0000)
Mashela	0.3277***	(0.0637)	-0.0233*	(0.0103)	0.0052*	(0.0023)	0.0000	(0.0000)
Millet	0.2251***	(0.0442)	0.0099	(0.0115)	-0.0012	(0.0026)	0.0001***	(0.0000)
Nueg	0.2516*	(0.1019)	0.0404$^{\cdot}$	(0.0224)	0.0045	(0.0045)	0.0001	(0.0001)
Peas	0.3665***	(0.0384)	-0.0044	(0.0097)	0.0014	(0.0018)	0.0001*	(0.0000)
Pepper	0.4125***	(0.0626)	-0.0236*	(0.0107)	-0.0078*	(0.0031)	0.0001***	(0.0000)
Potato	0.3024***	(0.0569)	0.0011	(0.0088)	0.0034	(0.0022)	0.0000	(0.0000)
Pumpkins	0.2883***	(0.0447)	-0.0122	(0.0120)	0.0024	(0.0021)	0.0001***	(0.0000)
Rice	0.2454**	(0.0887)	0.0571**	(0.0213)	0.0097*	(0.0039)	0.0000	(0.0000)
Shenkora	0.4712***	(0.0743)	-0.0248$^{\cdot}$	(0.0143)	0.0012	(0.0026)	0.0000$^{\cdot}$	(0.0000)
Shimbura	0.5115***	(0.0496)	0.0223	(0.0186)	0.0061**	(0.0022)	0.0001*	(0.0000)
Sorghum	0.5237***	(0.0179)	-0.0221***	(0.0053)	0.0019**	(0.0006)	0.0000*	(0.0000)
Teff	0.3247***	(0.0280)	-0.0098	(0.0062)	0.0050***	(0.0013)	0.0001***	(0.0000)
Tobacco	0.4699***	(0.0991)	-0.0070	(0.0211)	0.0034	(0.0056)	-0.0000	(0.0000)
Wheat	0.5635***	(0.0337)	-0.0178$^{\cdot}$	(0.0106)	0.0038**	(0.0013)	0.0000**	(0.0000)
Fixed Effects								
Household (hhid)		Yes		Yes		Yes		Yes
Wave		Yes		Yes		Yes		Yes
Crop		Yes		Yes		Yes		Yes
Statistics								
Observations		20,507		20,507		20,507		20,507
S.E.:Clustered		by:hhid		by:hhid		by:hhid		by:hhid
R-squared		0.55545		0.55545		0.55545		0.55545
Within R-squared		0.37529		0.37529		0.37529		0.37529

Notes: *** $p<0.001$, ** $p<0.01$, * $p<0.05$, . $p<0.1$. Coefficients represent the impact of climate variables (temperature and precipitation, linear and quadratic terms) on log agricultural yield. Quadratic precipitation terms are multiplied by 10^5 for readability.

Intensive Margin: Temperature bins (non-linear)

Table 12: Temperature Bin Effects on Log Yield for Fruit

	Avocado		Mango		Banana	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE
$[14; 15[$	0.324^{***}	(0.056)	0.155^{*}	(0.063)	0.228^{***}	(0.042)
$[15; 16[$	0.342^{***}	(0.081)	0.228^{***}	(0.042)	0.195^{***}	(0.026)
$[16; 17[$	0.211^{***}	(0.037)	0.124^{***}	(0.036)	0.131^{***}	(0.025)
$[17; 18[$	0.127^{**}	(0.043)	0.046	(0.042)	0.078^{***}	(0.022)
$[18; 19[$	0.087^{*}	(0.042)	0.087^{*}	(0.039)	0.013	(0.017)
$[19; 20[$	0.021	(0.042)	0.025	(0.036)	-0.053^{**}	(0.017)
$[20; 21[$	-0.006	(0.029)	-0.020	(0.037)	-0.078^{***}	(0.017)
$[21; 22[$	$-0.077^{\cdot}$	(0.040)	-0.013	(0.041)	-0.119^{***}	(0.019)
$[22; 23[$	-0.101^{*}	(0.043)	-0.022	(0.033)	-0.127^{***}	(0.022)
$[23; 24[$	-0.181^{***}	(0.043)	-0.100^{*}	(0.047)	-0.231^{***}	(0.029)
$[24; 25[$	-0.289^{***}	(0.065)	-0.109^{*}	(0.043)	-0.227^{***}	(0.028)
$[25; 26[$	-0.247^{**}	(0.086)	-0.248^{***}	(0.068)	-0.286^{***}	(0.036)
$[26; 27[$	-0.093	(0.074)	-0.089	(0.061)	-0.344^{***}	(0.046)
$[27; 28[$	-0.295^{***}	(0.021)	-0.194^{**}	(0.065)	-0.585^{***}	(0.112)
$[28; 29[$	-	-	-0.326^{**}	(0.112)	-0.334^{***}	(0.064)
Fixed Effects						
Household (hhid)	Yes		Yes		Yes	
Wave	Yes		Yes		Yes	
Stats						
Obs	20,507		20,507		20,507	
S.E. Clustered	by:hhid		by:hhid		by:hhid	
R^2	0.286		0.286		0.306	
Within R^2	0.006		0.005		0.034	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. The dependent variable is log agricultural yield. Temperature bins represent °C intervals. Dashes indicate coefficients removed due to collinearity.

Table 13: Oleaginius Nuts

	Beans		Peas		Ground Nuts	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE
[14; 15[	0.4227***	(0.0566)	0.2957***	(0.0580)	0.0524	(0.0569)
[15; 16[	0.2842***	(0.0332)	0.3123***	(0.0642)	0.2409*	(0.0978)
[16; 17[	0.2052***	(0.0252)	0.3376***	(0.0968)	0.2341***	(0.0585)
[17; 18[	0.2401***	(0.0242)	0.1020	(0.0760)	0.1716*	(0.0798)
[18; 19[	0.1866***	(0.0214)	0.2237***	(0.0503)	0.1526*	(0.0745)
[19; 20[	0.1230***	(0.0199)	0.1199**	(0.0424)	0.1073*	(0.0479)
[20; 21[	0.0879***	(0.0218)	0.0866.	(0.0449)	0.0403	(0.0537)
[21; 22[	0.0093	(0.0220)	0.0200	(0.0444)	-0.1108*	(0.0555)
[22; 23[	0.0117	(0.0262)	-0.0505	(0.0765)	-0.1989***	(0.0550)
[23; 24[	-0.0345	(0.0277)	0.0506	(0.0555)	-0.0348	(0.0365)
[24; 25[	-0.0597.	(0.0347)	-0.0254	(0.0656)	-0.0241	(0.1088)
[25; 26[	-0.1610***	(0.0408)	-0.1372.	(0.0713)	0.0240	(0.0405)
[26; 27[	-0.2273***	(0.0407)	-0.0867	(0.0540)	-	-
[27; 28[	-0.1507*	(0.0723)	-0.3171**	(0.1012)	-	-
[28; 29[	-0.1573	(0.1266)	-0.2144***	(0.0322)	-	-
[29; +[	-0.4393*	(0.2208)	-0.1882	(0.1692)	-	-
Fixed Effects						
Household (hhid)	Yes		Yes		Yes	
Wave	Yes		Yes		Yes	
Stats						
Obs	20,507		20,507		20,507	
S.E.: Clustered	by: hhid		by: hhid		by: hhid	
R^2	0.298		0.287		0.283	
Within R^2	0.021		0.006		0.002	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. The dependent variable is log agricultural yield. Temperature bins represent °C intervals. Some variables were removed due to collinearity (see text for details).

Table 14: Temperature Bin Effects on Log Yield for Vegetables

	Garlic		Beans		Potato		Pumpkins		Shenkora	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
[14; 15[	0.3754***	(0.0490)	0.2280	(0.1670)	0.1361	(0.0838)	0.0647	(0.1256)	0.3049**	(0.1128)
[15; 16[	0.3643***	(0.0504)	0.2721***	(0.0441)	0.1036	(0.0812)	0.1739**	(0.0632)	0.3643***	(0.0573)
[16; 17[	0.2451***	(0.0605)	0.2407***	(0.0410)	-0.0056	(0.0514)	-0.0601	(0.0704)	0.3132***	(0.0862)
[17; 18[	0.1300***	(0.0322)	0.1208***	(0.0281)	0.0741	(0.0482)	0.0563	(0.0571)	0.1303	(0.0852)
[18; 19[	0.0578*	(0.0270)	0.0391	(0.0282)	-0.0638	(0.0438)	-0.0482	(0.0429)	0.1510$^{\cdot}$	(0.0862)
[19; 20[	0.0428	(0.0286)	0.0070	(0.0240)	-0.0168	(0.0380)	-0.0577	(0.0399)	0.1196	(0.0734)
[20; 21[	-0.0221	(0.0286)	-0.0495$^{\cdot}$	(0.0254)	-0.0677	(0.0561)	-0.0580	(0.0412)	0.0155	(0.0587)
[21; 22[	-0.0974***	(0.0267)	-0.1183***	(0.0253)	-0.0720	(0.0442)	-0.1073*	(0.0423)	-0.0412	(0.0489)
[22; 23[	-0.1020**	(0.0329)	-0.1609***	(0.0318)	-0.2333***	(0.0543)	-0.1769***	(0.0470)	-0.0969	(0.0809)
[23; 24[	-0.2337***	(0.0355)	-0.2158***	(0.0308)	-0.1873***	(0.0468)	-0.2097***	(0.0556)	-0.0816	(0.0604)
[24; 25[	-0.2403***	(0.0457)	-0.2110***	(0.0361)	-0.3143***	(0.0618)	-0.1442**	(0.0521)	-0.0982	(0.0725)
[25; 26[	-0.3129***	(0.0462)	-0.2482***	(0.0415)	-0.3805***	(0.0679)	-0.2601***	(0.0526)	0.0027	(0.0703)
[26; 27[	-0.4206***	(0.0563)	-0.2085**	(0.0688)	-0.3535***	(0.0670)	-0.2369*	(0.1144)	-0.2204***	(0.0530)
[27; 28[	-0.2984*	(0.1264)	-0.3028***	(0.0790)	-0.2374***	(0.0660)	0.0139***	(0.0034)	-	-
[28; 29[	-0.6934***	(0.0290)	-0.4497***	(0.0730)	-0.2435*	(0.1177)	-0.3098***	(0.0377)	-0.2440***	(0.0539)
[29; +[	-0.2516***	(0.0468)	-0.3466*	(0.1354)	-0.3540***	(0.1036)	-0.1026	(0.1052)	-	-
Fixed Effects										
Household (hhid)					Yes					
Wave					Yes					
Stats										
Obs	20,507		20,507		20,507		20,507		20,507	
R^2	0.291		0.294		0.286		0.285		0.284	
Within R^2	0.012		0.016		0.005		0.003		0.002	

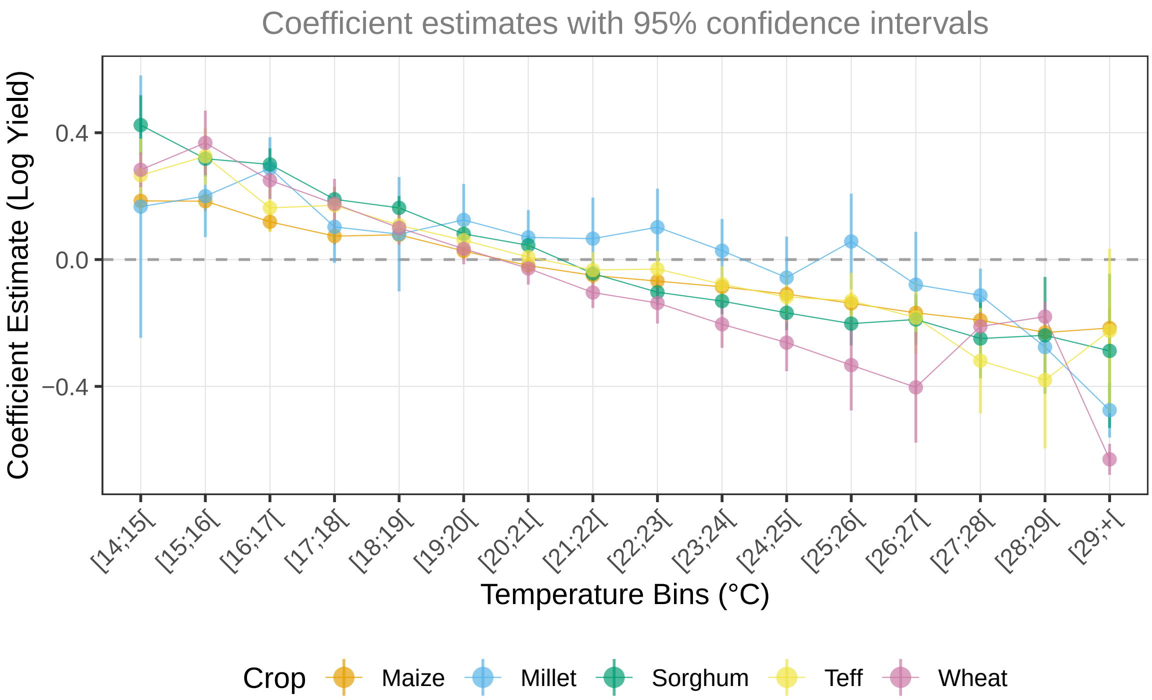
Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. Dashes indicate coefficients removed due to collinearity. Temperature bins represent °C intervals.

Table 15: Temperature Bin Effects on Log Yield for oil seeds

	Basil		**Coffee**		**Enset**		**Gesho**	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
[14; 15[	0.5300***	(0.0336)	0.3075***	(0.0565)	0.2629***	(0.0567)	0.3731***	(0.0450)
[15; 16[	0.4915***	(0.0523)	0.3753***	(0.0275)	0.2252***	(0.0475)	0.2978***	(0.0545)
[16; 17[	0.2114***	(0.0607)	0.2647***	(0.0284)	0.1703***	(0.0345)	0.3222***	(0.0497)
[17; 18[	0.1690*	(0.0742)	0.2036***	(0.0301)	0.1416***	(0.0262)	0.1880***	(0.0440)
[18; 19[	0.2100***	(0.0578)	0.1446***	(0.0211)	0.0939***	(0.0247)	0.1369***	(0.0327)
[19; 20[	0.0772	(0.0699)	0.0948***	(0.0235)	0.0267	(0.0233)	0.0551·	(0.0323)
[20; 21[	0.0285	(0.0602)	0.0602*	(0.0245)	-0.0463*	(0.0214)	0.0388	(0.0321)
[21; 22[	0.0376	(0.0479)	-0.0211	(0.0276)	-0.1325***	(0.0222)	-0.0470	(0.0327)
[22; 23[	-0.0644	(0.0931)	0.0269	(0.0268)	-0.1460***	(0.0228)	-0.0542·	(0.0328)
[23; 24[	-0.2294**	(0.0781)	-0.0610*	(0.0284)	-0.1778***	(0.0315)	-0.1368***	(0.0388)
[24; 25[	-0.2017**	(0.0676)	-0.1377***	(0.0397)	-0.2654***	(0.0355)	-0.1658***	(0.0466)
[25; 26[	-0.1215	(0.1429)	-0.0332	(0.0669)	-0.2219***	(0.0408)	-0.1569*	(0.0645)
[26; 27[	-0.0944	(0.0626)	-0.2338***	(0.0511)	-0.4890***	(0.0581)	-0.2919**	(0.1000)
[27; 28[	-0.2134***	(0.0608)	-	-	-0.2685***	(0.0790)	-0.2988**	(0.0986)
[28; 29[	-	-	-0.1019	(0.0873)	-0.4533***	(0.0233)	0.0932***	(0.0224)
[29; +[	-	-	0.0242	(0.0192)	-	-	-	-
Fixed Effects								
Household (hhid)					Yes			
Wave					Yes			
Stats								
Obs	20,507		20,507		20,507		20,507	
R²	0.285		0.298		0.299		0.289	
Within R²	0.003		0.021		0.024		0.010	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. Dashes indicate coefficients removed due to collinearity. Temperature bins represent °C intervals.

Figure 21: Temperature Bin Effects on Cereal Yields



Extensive margin

Table 16: Cereal Crop Choice (Extensive Margin)

	Maize		Teff		Sorghum		Millet		Wheat	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
[14; 15[	-0.547	(1.035)	-0.698	(1.146)	14.91***	(0.864)	-	-	-2.785***	(0.784)
[15; 16[	-0.401	(0.978)	-1.699	(1.641)	14.45***	(0.647)	18.25***	(2.531)	12.78***	(1.431)
[16; 17[	-1.268	(0.961)	1.194	(1.006)	13.52***	(0.764)	16.68***	(4.257)	13.87***	(1.227)
[17; 18[	-0.606	(0.974)	-0.837	(0.882)	14.90***	(0.486)	16.14***	(3.569)	13.14***	(1.068)
[18; 19[	-1.197	(0.939)	0.453	(0.973)	14.67***	(0.431)	17.76***	(3.654)	13.68***	(0.852)
[19; 20[	-1.387	(0.939)	-0.712	(0.800)	15.04***	(0.397)	13.49***	(3.866)	12.67***	(0.939)
[20; 21[	-1.587$^{\cdot}$	(0.943)	0.733	(0.915)	15.56***	(0.407)	12.37**	(4.729)	13.19***	(1.002)
[21; 22[	-0.922	(0.946)	0.269	(0.895)	14.78***	(0.328)	10.89$^{\cdot}$	(5.836)	13.46***	(0.874)
[22; 23[	-1.553$^{\cdot}$	(0.937)	-0.006	(0.955)	14.96***	(0.447)	15.56*	(6.318)	12.37***	(1.141)
[23; 24[	-0.991	(0.958)	0.099	(1.015)	15.30***	(0.464)	14.89***	(2.684)	12.61***	(1.125)
[24; 25[	-0.596	(0.965)	0.420	(0.966)	13.67***	(0.624)	11.82*	(5.350)	12.92***	(1.337)
[25; 26[	-0.636	(0.962)	-1.150	(1.510)	13.06***	(1.045)	15.27***	(3.059)	12.24***	(1.812)
[26; 27[	-0.321	(1.005)	2.330*	(1.131)	14.78***	(0.781)	-24.64**	(7.576)	-3.383***	(0.930)
[27; 28[	-0.984	(0.985)	-0.032	(2.349)	13.40***	(1.245)	3.031	(2.047)	-2.450**	(0.857)
[28; 29[	-0.151	(0.963)	-0.987	(1.041)	15.10***	(0.900)	-	-	-2.632**	(0.851)
[29; +[	-0.975	(1.350)	-	-	0.988***	(0.270)	-	-	-	-
Precipitation										
Precip	0.017	(0.023)	0.061	(0.067)	-0.052	(0.037)	0.220	(0.260)	0.028	(0.083)
Precip^2	-1.00e-5	(1.36e-5)	-3.63e-5	(3.95e-5)	3.12e-5	(2.18e-5)	-0.0001	(0.0002)	-2.00e-5	(4.83e-5)
Fixed Effects										
Household (hhid)	Yes		Yes		Yes		Yes		Yes	
Wave	Yes		Yes		Yes		Yes		Yes	
Stats										
Observations	2,482		551		960		126		334	
Squared Corr.	0.400		0.274		0.199		0.641		0.175	
Pseudo R^2	0.318		0.245		0.173		0.611		0.195	
BIC	6,384.1		1,264.1		2,319.9		266.61		809.51	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. The dependent variable is a crop choice dummy (extensive margin). Some temperature bins were removed due to collinearity or lack of variation.

Table 17: Fruits Choice (Extensive Margin)

	Avocado		Mango		Banana	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE
[14; 15[	-1.388	(0.918)	14.40^{***}	(1.935)	-14.04^{***}	(1.304)
[15; 16[	-16.62^{***}	(1.588)	14.85^{***}	(1.049)	-13.71^{***}	(0.948)
[16; 17[	-16.37^{***}	(1.187)	16.80^{***}	(1.336)	-14.22^{***}	(0.771)
[17; 18[	-16.85^{***}	(0.925)	14.31^{***}	(0.974)	-14.52^{***}	(0.549)
[18; 19[	-15.97^{***}	(0.998)	13.87^{***}	(0.893)	-14.78^{***}	(0.557)
[19; 20[	-16.91^{***}	(0.933)	14.31^{***}	(1.105)	-14.41^{***}	(0.598)
[20; 21[	-16.75^{***}	(1.064)	15.74^{***}	(0.693)	-14.84^{***}	(0.452)
[21; 22[	-15.99^{***}	(1.058)	14.78^{***}	(0.950)	-14.53^{***}	(0.497)
[22; 23[	-16.06^{***}	(1.188)	13.89^{***}	(0.993)	-14.24^{***}	(0.502)
[23; 24[	-17.06^{***}	(1.338)	16.00^{***}	(0.878)	-14.61^{***}	(0.552)
[24; 25[	-17.42^{***}	(2.008)	15.32^{***}	(1.252)	-14.85^{***}	(0.572)
[25; 26[	-17.91^{***}	(1.507)	15.30^{***}	(1.292)	-14.21^{***}	(0.803)
[26; 27[	-19.69^{***}	(2.237)	16.48^{***}	(1.421)	-13.86^{***}	(1.298)
[27; 28[	-0.108	(0.613)	15.73^{***}	(1.844)	-0.478	(0.340)
[28; 29[	-35.75^{***}	(1.423)	16.22^{***}	(1.346)	-0.182	(0.268)
[29; +[	-	-	19.43^{***}	(2.619)	-17.80^{***}	(1.049)
Precipitation						
Precip	-0.142	(0.109)	0.015	(0.084)	-0.050	(0.048)
Precip²	7.85e-5	(6.21e-5)	-7.14e-6	(4.93e-5)	3.09e-5	(2.83e-5)
Fixed Effects						
Household (hhid)	Yes		Yes		Yes	
Wave	Yes		Yes		Yes	
Stats						
Observations	176		313		874	
Squared Corr.	0.227		0.346		0.089	
Pseudo R²	0.196		0.303		0.095	
BIC	439.18		704.61		1,425.4	

Notes: *** p<0.001, ** p<0.01, * p<0.05. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. The dependent variable is a crop choice dummy (extensive margin). Some temperature bins were removed due to collinearity or lack of variation.

Table 18: Oleagineus Crop Choice (Extensive Margin)

	Beans		Peas	
Temp Bin	Coeff	SE	Coeff	SE
[14; 15[	0.972	(2.930)	1.942	(2.210)
[15; 16[	0.734	(2.844)	-31.77***	(2.277)
[16; 17[	2.010	(2.829)	-34.63***	(1.776)
[17; 18[	1.593	(2.815)	-35.20***	(1.805)
[18; 19[	1.949	(2.827)	-34.30***	(1.606)
[19; 20[	1.103	(2.815)	-33.05***	(1.461)
[20; 21[	1.483	(2.832)	-35.83***	(1.874)
[21; 22[	1.149	(2.771)	-33.93***	(1.249)
[22; 23[	0.441	(2.864)	-37.04***	(2.179)
[23; 24[	1.319	(2.819)	-33.47***	(1.537)
[24; 25[	0.121	(2.852)	-32.76***	(1.816)
[25; 26[	1.649	(2.888)	-30.02***	(1.725)
[26; 27[	0.844	(3.008)	-34.81***	(2.773)
[27; 28[	0.876	(3.036)	-37.48***	(2.588)
[28; 29[	-0.867	(3.212)	-68.17***	(1.723)
[29; +[	3.835	(2.925)	-	-
Precipitation				
Precip	0.085·	(0.051)	0.199	(0.167)
Precip²	-4.83e-5	(2.99e-5)	-0.0001	(9.71e-5)
Fixed Effects				
Household (hhid)	Yes		Yes	
Wave	Yes		Yes	
Stats				
Observations	869		279	
Squared Corr.	0.481		0.598	
Pseudo R²	0.403		0.526	
BIC	1,976.2		596.91	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. Some temperature bins were removed due to collinearity or lack of variation.

Table 19: Vegetable/Specialty Crop Choice (Extensive Margin)

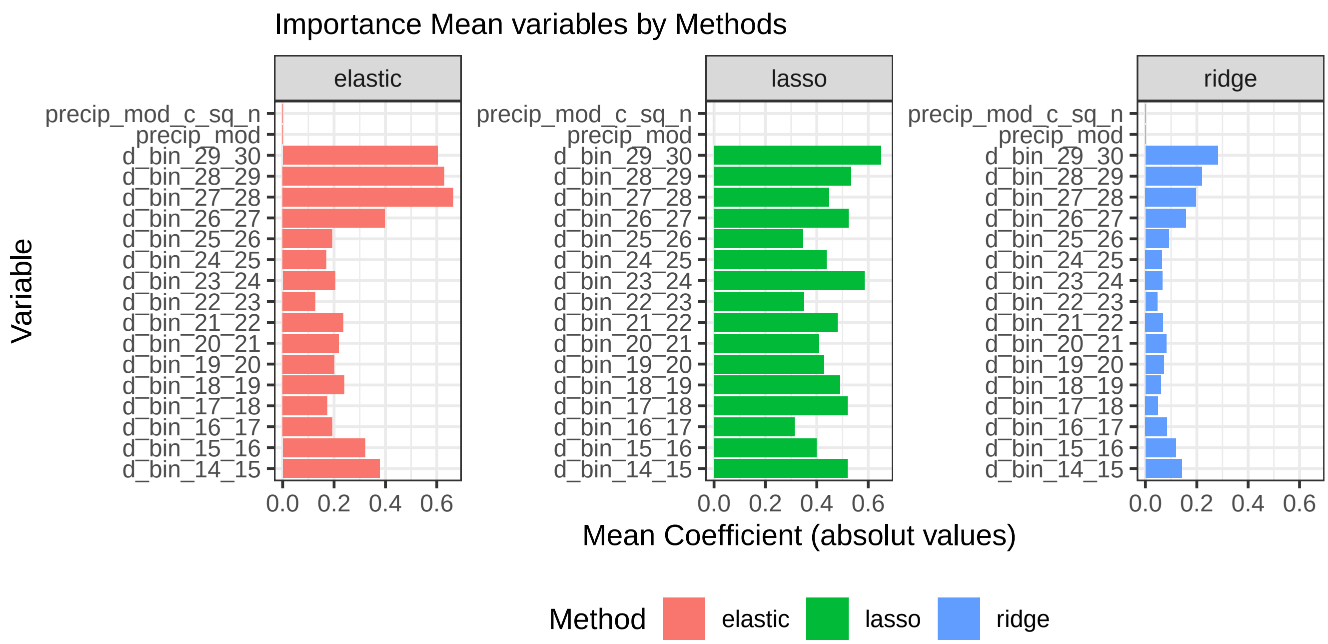
	Garlic		Kale		Pepper		Potato		Pumpkins		Shenkora	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
[14; 15[	-14.74***	(2.240)	-1.485	(1.668)	17.84***	(1.627)	-1.962	(3.403)	11.22***	(1.848)	13.51**	(4.939)
[15; 16[	-2.235	(2.284)	0.656	(1.343)	15.78***	(1.098)	-1.930	(3.534)	12.32***	(1.914)	4.460	(7.843)
[16; 17[	-2.796	(2.379)	1.110	(1.405)	16.42***	(1.217)	-1.777	(3.335)	12.42***	(2.113)	5.685	(6.357)
[17; 18[	-1.503	(2.221)	0.100	(1.176)	17.33***	(1.268)	-2.758	(3.509)	13.87***	(1.560)	5.845	(6.010)
[18; 19[	-1.675	(2.185)	0.338	(1.041)	17.97***	(0.979)	-2.035	(3.420)	13.99***	(1.051)	-8.362*	(3.814)
[19; 20[	-0.415	(2.144)	0.385	(1.024)	17.09***	(0.693)	-1.064	(3.372)	12.82***	(1.596)	4.245	(6.451)
[20; 21[	-0.451	(2.080)	0.641	(1.012)	16.99***	(0.988)	-2.805	(3.268)	12.78***	(1.591)	3.536	(8.412)
[21; 22[	-0.374	(2.206)	-0.214	(0.996)	16.96***	(0.780)	-3.097	(3.401)	13.47***	(1.464)	3.334	(7.453)
[22; 23[	0.337	(2.110)	0.611	(1.064)	18.12***	(1.234)	-1.399	(3.342)	13.49***	(1.452)	15.15***	(4.414)
[23; 24[	-0.395	(2.271)	0.090	(1.063)	17.28***	(1.081)	-2.132	(3.339)	13.12***	(1.892)	-7.640*	(3.879)
[24; 25[	-0.139	(2.338)	1.109	(1.141)	17.86***	(1.041)	-0.435	(3.477)	11.09*	(4.394)	9.197*	(4.243)
[25; 26[	2.179	(2.454)	-0.572	(1.287)	17.07***	(1.462)	-1.516	(3.989)	12.77***	(3.091)	4.619	(6.754)
[26; 27[	-1.806	(2.303)	0.652	(1.232)	16.83***	(1.567)	-1.244	(3.500)	-2.625*	(1.287)	5.438	(6.901)
[27; 28[	1.139	(2.950)	2.099	(1.452)	1.359*	(0.614)	-1.592	(3.613)	46.38***	(1.693)	-	-
[28; 29[	-	-	-	-	0.807	(0.962)	-4.018	(5.948)	16.23***	(1.738)	-	-
[29; +[	-	-	-	-	0.827	(1.091)	0.350	(3.852)	-	-	-	-
Precipitation												
Precip	0.075	(0.095)	-0.039	(0.044)	0.038	(0.067)	0.090	(0.085)	0.289	(0.185)	-0.578	(0.361)
Precip²	-4.27e-5	(5.63e-5)	2.19e-5	(2.6e-5)	-2.1e-5	(3.92e-5)	-4.96e-5	(4.88e-5)	-0.0002	(0.0001)	0.0003	(0.0002)
Fixed Effects												
Household (hhid)	Yes		Yes		Yes		Yes		Yes		Yes	
Wave	Yes		Yes		Yes		Yes		Yes		Yes	
Stats												
Observations	362		588		279		303		239		94	
Squared Corr.	0.524		0.460		0.241		0.387		0.351		0.588	
Pseudo R²	0.449		0.401		0.228		0.344		0.364		0.575	
BIC	803.70		1,343.2		676.18		703.55		542.26		218.19	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. Some temperature bins were removed due to collinearity/lack of variation.

Table 20: Oil seeds Crop Choice (Extensive Margin)

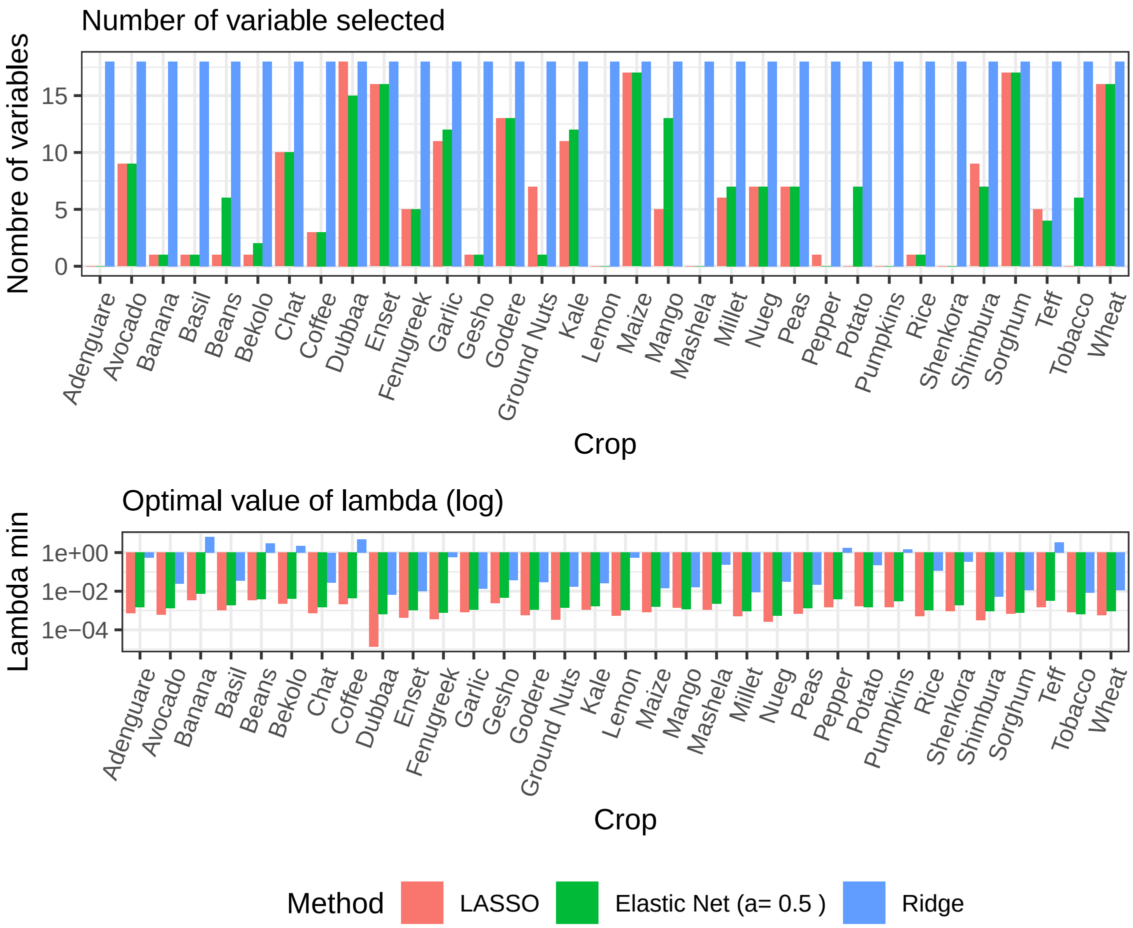
	Basil		Coffee		Enset		Gesho	
Temp Bin	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
$[14;15[$	-21.22***	(2.890)	0.540	(1.437)	13.68***	(1.313)	-0.364	(1.530)
$[15;16[$	-68.88***	(8.016)	1.122	(1.195)	12.91***	(1.051)	0.219	(1.529)
$[16;17[$	-3.139·	(1.739)	0.447	(1.210)	12.20***	(1.285)	0.055	(1.733)
$[17;18[$	-20.58***	(3.722)	0.105	(1.193)	13.81***	(0.749)	-0.337	(1.264)
$[18;19[$	-19.75***	(1.369)	0.778	(1.122)	13.34***	(0.811)	-0.514	(1.349)
$[19;20[$	-20.38***	(2.749)	0.206	(1.138)	14.03***	(0.732)	1.394	(1.131)
$[20;21[$	-18.47***	(3.057)	-0.003	(1.137)	14.20***	(0.734)	0.735	(0.986)
$[21;22[$	-20.09***	(4.247)	0.041	(1.150)	14.38***	(0.692)	0.530	(1.083)
$[22;23[$	-19.00***	(4.503)	0.180	(1.110)	14.10***	(0.654)	1.036	(1.027)
$[23;24[$	-19.44***	(3.755)	0.135	(1.128)	13.97***	(0.723)	0.885	(1.194)
$[24;25[$	-21.21***	(2.704)	0.870	(1.097)	13.01***	(0.893)	0.130	(1.342)
$[25;26[$	-15.22**	(4.799)	0.226	(1.177)	13.73***	(0.819)	0.357	(1.842)
$[26;27[$	-22.42***	(2.916)	-0.498	(1.898)	13.04***	(1.517)	0.061	(2.031)
$[27;28[$	-	-	0.764	(1.630)	11.88***	(1.603)	0.015	(1.410)
$[28;29[$	-	-	14.35***	(1.051)	13.54***	(1.355)	-2.511	(2.634)
$[29;+[$	-	-	-2.355	(2.318)	-0.221	(0.862)	-	-
Precipitation								
Precip	1.218·	(0.683)	-0.014	(0.036)	-0.092·	(0.050)	0.070	(0.060)
Precip²	-0.001·	(0.0004)	8.79e-6	(2.09e-5)	5.32e-5·	(2.91e-5)	-3.92e-5	(3.51e-5)
Fixed Effects								
Household (hhid)	Yes		Yes		Yes		Yes	
Wave	Yes		Yes		Yes		Yes	
Stats								
Observations	89		815		607		289	
Squared Corr.	0.641		0.407		0.447		0.424	
Pseudo R²	0.597		0.325		0.368		0.351	
BIC	205.66		1,911.9		1,402.8		699.34	

Notes: *** p<0.001, ** p<0.01, * p<0.05, . p<0.1. All models include household (hhid) and wave fixed effects with standard errors clustered at household level. Some temperature bins were removed due to collinearity or lack of variation. Extreme coefficients in Basil model suggest potential sensitivity to temperature.

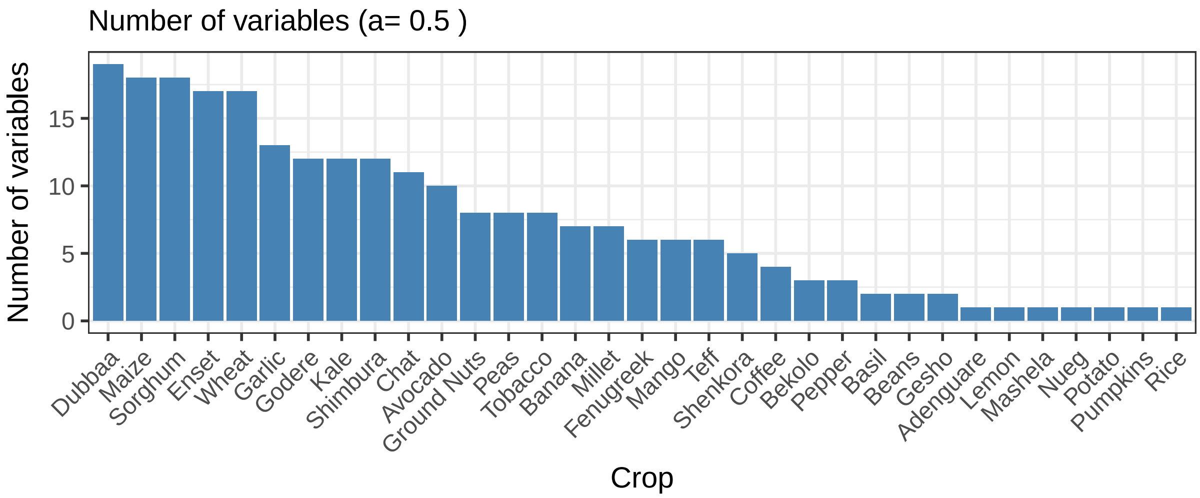


method	mean deviance	mean vars	mean lambda
LASSO	0.223	5.060	0.001
Elastic Net (α= 0.5)	0.223	6.485	0.0020
Ridge	0.223	18	0.845

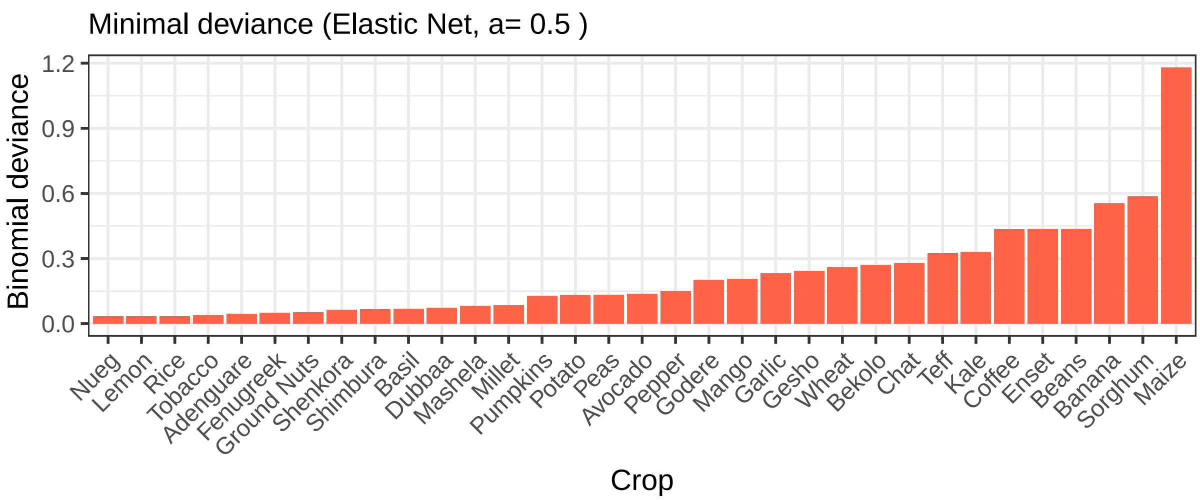
Table shows mean performance metrics across all crops



Number of variable selected
Nombre of variables
0
5
10
15
Adenguare
Avocado
Banana
Basil
Beans
Bekolo
Chat
Coffee
Dubbaa
Enset
Fenugreek
Garlic
Gesho
Godere
Ground Nuts
Kale
Lemon
Maize
Mango
Mashela
Millet
Nueg
Peas
Pepper
Potato
Pumpkins
Rice
Shenkora
Shimbura
Sorghum
Teff
Tobacco
Wheat
Crop
Optimal value of lambda (log)
Lambda min
1e+00
1e−02
1e−04
Crop
Method
LASSO
Elastic Net (a= 0.5)
Ridge



Number of variables (a= 0.5)
Number of variables
15
10
5
0
Dubbaa
Maize
Sorghum
Enset
Wheat
Garlic
Godere
Kale
Shimbura
Chat
Avocado
Ground Nuts
Peas
Tobacco
Banana
Millet
Fenugreek
Mango
Teff
Shenkora
Coffee
Bekolo
Pepper
Basil
Beans
Gesho
Adenguare
Lemon
Mashela
Nueg
Potato
Pumpkins
Rice
Crop



Minimal deviance (Elastic Net, a= 0.5)
Binomial deviance
1.2
0.9
0.6
0.3
0.0
Nueg
Lemon
Rice
Tobacco
Adenguare
Fenugreek
Ground Nuts
Shenkora
Shimbura
Basil
Dubbaa
Mashela
Millet
Pumpkins
Potato
Peas
Avocado
Pepper
Godere
Mango
Garlic
Gesho
Wheat
Bekolo
Chat
Teff
Kale
Coffee
Enset
Beans
Banana
Sorghum
Maize
Crop