Shifting Seasons: Agricultural Adaptation and Resilience in Africa

Iván Kim Taveras *

JOB MARKET PAPER

Link to most recent version

March 26, 2026

Abstract

Climate risk threatens Sub-Saharan Africa's rainfed agriculture. Using microdata from six countries, I show that a one-week delay in the onset of the rainy season reduces yields by 2% and consumption by 1%. Damages disproportionately harm the most vulnerable, with the effects being most pronounced on female-managed plots, while education and wealth build resilience. Farmers adapt by delaying planting, but this is insufficient due to informational frictions. False onsets—misleading early rains followed by a dry spell—trigger premature planting, increasing damages. The negative impacts are concentrated in locations experiencing long-term climatic shifts, indicating a persistent failure to adapt. Projecting these damages forward reveals a substantial threat: cumulative discounted losses from 2025 to 2050 could reach up to 10% of 2024 real GDP. These findings establish shifts in seasonal timing as a first-order economic threat and highlight the value of short-range forecasts in mitigating this risk.

Keywords: Rainy Season Onset, Agricultural Productivity, Climate Change Adaptation, Africa, LSMS
JEL Codes: Q12, Q54, O13

*__Kim Taveras__: Bocconi University (e-mail: ivan.kimtaveras@phd.unibocconi.it). I am grateful to my advisors, Jérôme Adda, Alex Armand, and Joseph-Simon Görlach. I also thank Erika Deserranno, Sarah Eichmeyer, Jared Gars, Katarina Kuske, Thomas Le Barbanchon, Karen Macours, Cristóbal Ruiz-Tagle, Enrico Stivella, Julian Streyczek, and Guido Tabellini, as well as seminar participants at Bocconi University, NEUDC 2025, the II Spanish Workshop in Development Economics, LEAP, the Milan PhD Workshop, and the Oxford Development Economics Workshop for valuable comments and suggestions. All remaining errors are my own.

1 Introduction

Sub-Saharan Africa's reliance on rainfed agriculture makes its population exceptionally vulnerable to weather risk. In a region where approximately 95% of cropland is rainfed and 55-62% of the workforce is employed in agriculture (IPCC, 2022), shifts in weather patterns pose a substantial threat to livelihoods. The *timing* of the rainy season is a particularly important feature of weather that is often overlooked. The onset of the rainy season has a direct biophysical impact on agricultural productivity by shifting crops' exposure to weather conditions throughout critical growth stages, and an indirect effect through an adaptation channel. Through this channel, weather is not merely an input but also an uncertain signal that informs farmers' decisions across several dimensions but most importantly their planting schedule (Kala, 2019). As climate change evolves, understanding farmers' adaptation to variation in timing is essential. This requires both identifying the strategies farmers employ and the frictions that limit the effectiveness of their responses or turn their heuristics into costly mistakes. First, it can inform the design of interventions to address the growing impacts of climate change; the strategies farmers already employ are, by revealed preference, those they deem most cost-effective among their feasible options (Hultgren et al., 2025). Second, it helps identify the most promising areas for intervention, especially where the adoption of agricultural technologies is low (Suri et al., 2024).

This paper addresses a central question for climate resilience: How do weather signals shape the adaptation to, and economic impacts of, shifting seasons? To answer this, the analysis proceeds in three parts. First, I quantify the causal impacts of short-run variation in the onset of the rainy season on agricultural yields and household welfare. Second, I examine the adaptation strategies farmers employ and the informational frictions that limit their response. Finally, I link these short-run semi-elasticities to long-term climatic trends to project the future economic damages. The analysis focuses on six Sub-Saharan African countries: Ethiopia, Malawi, Mali, Niger, Nigeria, and Tanzania, which collectively represent 35% of the region's population and 40% of its economy (The World Bank, 2025). Because the start of the rainy season cannot be observed directly, I use daily weather and soil data to identify the first date when conditions are suitable for planting.

Before the main empirical analysis, I establish the climatic context by documenting a significant long-term shift in the timing of the rainy season. The data reveal a stark trend for the rainy season onset: across all six countries, it has shifted significantly later between 1979 and 2024. The magnitude of this delay ranges from approximately 6 days in Tanzania to over 25 days in Nigeria. The trend for cessation dates is more varied; in some countries, such as Niger and Nigeria, the season's end has also become later, while in others like Mali and Tanzania, it now arrives earlier. The net result for most countries is a discernible shortening of the rainy season. For example, the growing period in Mali has shortened by approximately 26 days, while in

Niger, where the later onset was offset by a later cessation, the length of the rainy season has remained broadly unchanged.

My identification strategy is designed to first quantify the net economic impacts of onset timing, and then to disentangle the role of the adaptation channel in driving these impacts. The strategy exploits substantial year-to-year variability in onset timing around long-term trends, focusing on short-run, plausibly exogenous shocks. To isolate these shocks, I compare variation within locations over time, controlling for local trends and regional conditions using a multi-way fixed effects structure that includes high-resolution grid-cell fixed effects and climatic-area-by-country-by-year fixed effects. To control for the direct biophysical channel—such as changes in precipitation patterns resulting from a delayed rainy season—the model includes a rich set of weather controls. By accounting for these aggregate weather impacts, the strategy estimates the net effect of a pure timing shock, distinct from overall seasonal weather.

I find that a one-week delay lowers crop yields by approximately 2%. This effect is asymmetric, driven entirely by later-than-usual onsets, and is also short-lived, with no spillover to future agricultural seasons. These findings are robust across several specification checks, and the statistical inference holds when accounting for spatial correlation in onset dates.

A delayed onset also harms household welfare, lowering per capita consumption by 1-1.2% per week of delay and worsening women's nutritional status. It is therefore critical to understand if and how farmers respond. I find that they primarily adjust along two low-cost margins: they plant later and use more seeds, with limited evidence of other short-run changes in fertiliser use, crop choice, or labour allocation. Focusing on the most intuitive strategy—shifting the planting schedule—, the analysis reveals that the adjustment is insufficient, as a one-week delay in the onset leads to a planting delay of just over half a day. To confirm that this muted adaptation response is a real phenomenon and not merely a statistical artifact driven by the coarse, monthly nature of the survey data, I design a simulation exercise to correct for this measurement error and find that the true daily-level planting adjustment, while larger than the initial estimate, remains modest. The key puzzle, therefore, is why farmers do not delay planting more substantially when the rainy season arrives late.

Part of the answer lies in exploring who is most affected and how they adapt. The damages from a delayed onset are not evenly distributed; the negative impacts are borne disproportionately by the most vulnerable, with plots managed by women suffering significantly larger yield losses. Conversely, resilience is systematically linked to several farmer characteristics which induce heterogeneous incentives for adaptation. Education and household wealth make farmers less constrained, which should, in theory, improve their ability to execute optimal adaptive strategies like adjusting planting dates. In contrast, technologies such as irrigation make them less vulnerable to the timing of the onset, weakening their incentive to adapt on this margin.

This heterogeneity raises an important question: do these resilient farmers achieve better out-

comes because they are better at adapting their planting schedules? The evidence suggests this is largely not the case. An analysis of planting decisions reveals no statistically significant difference in timing adjustments for farmers with more education or more assets. The findings for technology are mixed and align with their differing incentives: access to irrigation is associated with a tendency to delay planting *less*, likely because a reliable water source reduces farmers' dependence on the rainy season for optimal growing conditions. In contrast, farmers using improved seeds are the only group to show a significantly more responsive timing adjustment. The fact that most resilient groups do not exhibit superior adaptation on this key margin suggests their resilience stems from an ability to buffer shocks rather than from superior timing decisions.

The answer to both puzzles—why adaptation is insufficient overall, and why superior adaptation on timing is not observed even among less constrained farmers—points to an informational barrier. First, farmer awareness of the underlying shock appears low and non-linear. Households are not more likely to report having suffered a negative shock to their crops in response to a typical late onset; awareness is triggered only by extreme events when the onset is exceptionally late. Even then, farmers appear to misattribute the cause, reporting the shock as a drought rather than linking the damage to the shift in seasonal timing.

Second, farmers can be misled by incorrect signals. I introduce the concept of *false onsets*: early rains followed by a dry spell. Farmers often adapt by following a common heuristic: they wait for a significant rainy event, often verifying the soil moisture, and then plant immediately afterwards (see, for instance, Marteau et al. 2011). A false onset is so damaging because it turns this reasonable strategy into a costly mistake. During the dry spell following a false onset, crops are likely to fail. Indeed, I find that farmers exposed to a false onset during a later-than-usual season plant earlier, which then requires replanting when the original seeds fail to grow. Ultimately, this combination of a false start and a late true onset more than doubles the negative impact on productivity.

The findings in this paper, which estimate the cost of short-run weather variability, have direct implications for understanding the long-term economic consequences of climate change. The analysis shows that the negative impact of a delayed onset on productivity is driven entirely by locations already experiencing a secular trend towards later rainy seasons. This evidence of a persistent failure to adapt to gradual climatic shifts provides the justification for using the estimated short-run damages as a guide for future losses. To quantify this threat, I project the Net Present Value (NPV) of damages to the year 2050. The results are stark: under a *Business as Usual* scenario, cumulative discounted losses could reach approximately 10% of real GDP in 2024 for the most affected country, Mali. While this projection is grounded in observed behavior, assuming a complete absence of future adaptation may be unrealistic. I therefore supplement the main analysis by considering two optimistic counterfactuals: one where the yield impact of onset delays linearly decreases to zero by 2050, and another that accounts for

the mitigating effect of future irrigation expansion. Even under these more optimistic scenarios, the remaining economic costs are substantial. Furthermore, the analysis reveals a significant *Sustainability Benefit*: a less severe climate path could avert over $20 billion in damages in Nigeria alone. These findings establish that subtle shifts in seasonal timing, a less-studied feature of climate change, represent a first-order economic threat to livelihoods in the region.

This paper contributes to three main strands of literature. First, it makes a novel contribution to the literature on adaptation to climate change in low-income settings. A large body of work, comprehensively reviewed by Carleton et al. (2024), documents how farmers in low-income settings adapt to weather shocks along various margins. Examples from this literature are geographically diverse, including studies on adjusting agricultural inputs in Kenya (Jagnani et al., 2020), investing in irrigation in India (Taraz, 2017), adopting soil and water conservation practices in Peru (Tambet and Stopnitzky, 2021), and changing land use in both India and Peru (Aragón et al., 2021). However, this literature has largely focused on responses to shocks in temperature and total precipitation, with less attention paid to the *timing* of the rainy season. This paper argues that this is an important omission because shocks to timing introduce unique informational frictions, adding a layer of complexity to a setting where learning is already complex and costly (Laajaj and Macours, 2024). My analysis shows these frictions have real consequences: farmers' adaptation is ultimately insufficient. By documenting their spontaneous ex-post responses, I find that while the most intuitive margins are adjusting planting dates and seed use, these adjustments are often too small or poorly timed to prevent losses, presenting a puzzle as to what limits a more effective response.

Second, this paper contributes to the extensive literature on information and knowledge constraints in agriculture. A large body of work documents how farmers may forgo profitable strategies, particularly technology adoption, because of constraints (for a review, see Suri et al. 2024). I show that the insufficiency of adaptation is driven by two key informational frictions. The first is a failure of attribution. Because year-to-year onset variation is not always a discrete, extreme event, its impact has low salience. My analysis shows that farmers struggle to distinguish it from normal productivity fluctuations; only when the delay is large enough do they report an adverse shock, and even then, they tend to misattribute it to drought rather than the underlying timing shift. This finding is consistent with Patel (2025), which, in the context of soil salinity and agricultural productivity in Bangladesh, provides a framework where such attribution failures arise naturally from an identification problem: farmers interpret ambiguous signals through the lens of their priors, leading to persistent misperceptions, especially when environmental shifts are subtle rather than salient. This paper provides evidence on this phenomenon in the context of seasonal timing in SSA. I introduce and test the importance of the *false onset*—early rains followed by a dry spell. As climate change systematically delays the onset, farmers' search for planting signals increasingly occurs at the wrong time of year, turning a traditionally reliable heuristic into a costly mistake that more than doubles the neg-

ative impact on productivity. These insights are useful for policy design, as addressing these informational constraints through tools like reliable short-range forecasts could enhance the strategies farmers already find intuitive, a potentially more effective approach than promoting costly and complex technologies that have historically suffered from low take-up (Suri et al., 2024). Together, these findings motivate the expansion of reliable weather forecasts to serve as an early-warning system, contributing to recent work on the value of both short-range (Camacho and Conover, 2019; Yegbemey et al., 2023; Rudder and Viviano, 2024) and long-range forecasts (Giné et al., 2018; Rosenzweig and Udry, 2019; Burlig et al., 2024).

Third, this paper contributes to the literature quantifying the economic impacts of weather shocks by examining a critical, yet understudied, dimension of weather risk. An extensive literature provides robust evidence on the impacts of shocks to temperature and rainfall quantity on agricultural and economic outcomes (Deschênes et al., 2009; Deschênes and Moretti, 2009; Barreca et al., 2016; Aragón et al., 2021). This paper complements this body of work by analysing the economic consequences of a different feature of weather: the timing of the rainy season. In the rainfed agricultural systems that dominate Sub-Saharan Africa, a delay in the onset can shift critical crop growth stages into periods of greater heat stress or misalign harvest with labor cycles, even when total seasonal rainfall is normal. Despite its agronomic importance, the economic literature on onset timing is sparse, with the few existing studies focused on India (Amale et al., 2023; Burlig et al., 2024). I complement existing evidence by providing multi-country evidence from Sub-Saharan Africa, the region most reliant on rainfed agriculture. To do so, I introduce a robust, high-resolution agronomic measure of onset, trace its effects beyond agricultural production to welfare indicators, and uncover how these damages are borne disproportionately by the most vulnerable farmers (Udry, 1996). By first estimating the impact of onset timing and then identifying its mechanisms, my findings show that these damages stem largely from responses to uncertain signals, not just from inevitable biophysical constraints.

The rest of the paper is structured as follows. Section 2 discusses the causal pathways linking the rainy season onset to crop yields. Section 3 describes the data sources and construction of key variables. Section 3.3 presents the evidence on long-term trends in the rainy season calendar. Section 4 outlines the empirical strategy in detail. Section 5 presents the main results, and Section 6 discusses the findings and concludes.

2 Linking the onset of the rainy season to crop yields

The timing of the rainy season onset affects agricultural productivity through two distinct pathways, as illustrated in Figure 1: a direct **biophysical channel** that impacts crop growth, and an indirect **adaptation channel** that operates through farmer decision-making.

The biophysical channel: Direct impacts on crop growth. Shifts in the actual onset date of the rainy season directly change the biophysical conditions crops experience throughout their development. The timing of the onset largely dictates when the entire growing period begins. This shift means all subsequent crop life cycle stages occur at different points in the calendar year, potentially exposing them to less favourable weather and soil conditions regarding temperature, sunlight, and moisture, even if the total length of the season were unchanged. For instance, critical phases like germination, flowering, or grain and tuber filling might be pushed into periods of excessive heat or suboptimal solar radiation, thereby negatively impacting yields irrespective of season length (Jägermeyr and Frieler, 2018; Yang et al., 2024). Of course, a common direct consequence of a delayed onset is also a shortening of the effective growing season, a trend observed in the study regions (see Section 3.3). Such a reduction in available time can prevent crops from accumulating sufficient warmth or completing their development, leading to lower yields. Lastly, these onset-driven climatic shifts can indirectly lower yields by changing the exposure to pests and diseases, whose life cycles also respond to environmental changes (Yang et al., 2024).

The adaptation channel: Weather as a signal for farmer decisions. Beyond its direct effects, year-to-year variation in the onset is an important source of uncertainty and a critical **signal** for agricultural producers. In anticipation of, and in response to, this signal, farmers must make crucial management decisions that significantly influence eventual crop yields. These decisions span a range of practices, including the choice of crops to plant, the selection of specific crop varieties, the quantity and type of inputs (such as seeds, fertilisers, and pesticides), the allocation of labour, and, crucially, the timing of planting activities (Jägermeyr and Frieler, 2018; Marteau et al., 2011; Minoli et al., 2022). To interpret this uncertain signal, farmers often rely on local knowledge and heuristics, such as observing rainfall patterns and soil moisture, to guide their planting decisions and mitigate risks like early-season dry spells (Wolf et al., 2015; Marteau et al., 2011). The effectiveness and feasibility of these adjustments in the face of onset variability are central to understanding agricultural outcomes and represent the core of the adaptation channel. Persistent changes in seasonal timing can also make traditionally grown crop varieties less suitable over time, requiring farmers to adapt their crop choices in the long run.

The role of other timing features. While both the length and cessation of the rainy season are important, the relationship with productivity is not always straightforward. For example, an extended season is not inherently superior if those additional days expose the crop to end-of-season stresses like drought or increased pest pressure. Indeed, adaptive farming practices might involve selecting cultivars or adjusting planting schedules to achieve a strategically shorter but better-timed growing period to avoid such terminal stresses (Minoli et al., 2022). As the empirical evidence detailed in Section 5 will show, it is the year-to-year change in *onset timing* that is the primary aspect of seasonal timing that causally affects productivity in the

sampled countries, with cessation timing and overall length having less significant independent effects once the onset is controlled for.

3 Data

I combine detailed household and plot-level survey data with high-resolution climate information to analyse the impact of rainfall onset timing. Appendix A.1 provides further details on variable construction and specific data sources.

3.1 Household, individual, and plot-level data.

The core household, individual, and plot-level information is sourced from a recently released harmonised panel dataset compiling the Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) waves for seven Sub-Saharan African countries (World Bank, 2024). I use data for six of these countries: Ethiopia, Malawi, Mali, Niger, Nigeria, and Tanzania. While some central and northern regions of Ethiopia experience a secondary, sporadic rainy season (the *Belg*), the analysis focuses on the main agricultural cycle tied to the primary *Kiremt* rainy season. Uganda is excluded from the analysis because its predominantly bimodal rainfall pattern is unsuitable for my current methodology for identifying a single main rainy season onset, which relies on defining a consistent agroclimatic year. The included survey waves span the period from 2008 to 2022. Although the harmonised dataset is meant to be a panel, I treat the data as repeated cross-sections as agricultural plots cannot be reliably tracked across waves. Appendix A.1 provides a detailed list of the specific survey waves included for each country (see Table A2).

The LSMS-ISA surveys are designed to be representative of the household and smallholder agriculture sectors within the sampled countries, though it is important to note that specific sampling and coverage might vary slightly by country and wave.[1] Figure A1 shows the broad geographical coverage of the sampled LSMS communities. Locations are spread widely across the geography of each country. Naturally, the sample excludes hyper-arid areas where agriculture is not practiced, which is most evident in the desert regions of northern Mali and Niger. The surveys collect comprehensive data relevant to this study. The household questionnaire records demographics, education, labour participation, household nonfarm activities, shocks experienced, and crucially provides an annualised per capita measure of consumption. The agricultural questionnaires gather detailed plot-level information, including inputs, crop choices, yields, and farming practices such as planting and harvest timing. For all included surveys, I focus on the main agricultural growing season. Importantly, the harmonised dataset provides

[1] Refer to official LSMS-ISA documentation and World Bank (2024) for details.

plot area measurements based primarily on Global Positioning System (GPS) readings, addressing known inaccuracies in self-reported areas; imputation methods based on self-reports and administrative data are used when GPS measures are missing (see World Bank, 2024, for details). The dataset also provides standardised groupings for crops, which I use to analyse crop choice decisions.[2] The harmonised dataset defines the main crop on a plot as the crop within these categories having the highest reported monetary value.

To construct the final estimating samples, I applied several restrictions. Survey waves without geocoordinates were excluded. I also excluded the 2019–2021 Tanzania Wave 5, as its multi-year survey period and lack of planting-year information made it impossible to reliably match plots to a specific agricultural season. Only plots reporting a complete harvest are included, as incorporating yields from incomplete harvests could introduce bias into the productivity estimates. This decision is supported by formal tests showing that the timing of the rainy season onset does not significantly predict the probability of observing a fully harvested plot; Table B1, Column (1), presents this result. Lastly, I retain only plots where planting occurred within the same calendar year as the survey interview, a restriction that drops approximately 5% of plots and primarily excludes perennial crops. These criteria define the final plot-level sample used for productivity analyses.

After applying these criteria, the final sample for plot-level analysis comprises approximately 130,000 plot-wave observations. For the household-level analyses of consumption, I restrict the sample to the approximately 29,000 household-wave observations from the agricultural module that report a complete harvest for all plots. This represents 75% of all household-wave observations. The individual-level sample includes all working-age individuals from these selected households.

This restriction on complete harvests for the consumption analysis is necessary because the surveys provide an annualised consumption measure without specifying its timing relative to the harvest. Including households yet to complete their harvest would mean comparing post-harvest households with those still in their lean season, the period between planting and harvesting, introducing significant unobserved heterogeneity. While this raises a valid concern about selection bias–for instance, a late onset could delay a harvest, making it more likely the survey occurs before completion–I explicitly test for this. Column (2) of Table B1 confirms that the onset timing does not significantly predict the probability of observing a complete harvest, suggesting the restriction does not systematically bias the results. Furthermore, as shown in Table 2, the main findings on consumption are robust to including all households, regardless of their main occupation sector. Appendix A.1 details the construction of key outcome and control variables.

[2]The categories are barley, wheat, rice, sorghum, maize, millet, perennials such as fruit and tree crops, legumes, root crops, nuts, and others.

3.2 Nutritional outcomes from Demographic and Health Surveys.

To examine whether the impacts of rainfall onset variability extend to nutritional outcomes, I incorporate data from all 25 available DHS waves for the sampled countries that provide geocoded cluster information. A detailed list of the specific DHS survey years I included for each country is available in Table A2. DHS surveys are nationally representative, collecting comprehensive data on health and population with a particular focus on maternal and child health. Women aged 15–49 are the primary respondents, providing information on household characteristics, birth histories, wealth, and human capital, alongside objective anthropometric measurements (weight and height) for themselves and their children under five.

For this analysis, I primarily use samples of adult female respondents with available objective anthropometric data: a main sample from agricultural households, and a full sample of all women to test for spillovers. The key nutritional outcomes I examine for women are weight-for-height z-scores and indicators for whether she is classified as wasted or underweight. Analyses using data for children under five (weight-for-height z-scores, wasting, underweight) are presented in the appendix. Appendix A.1 details the construction of all nutritional variables.

3.3 Onset of the rainy season and weather variables.

I rely on daily climate information sourced from a dataset of agrometeorological indicators derived from ERA5 reanalysis, provided by the Copernicus Climate Change Service (C3S) Climate Data Store (Boogaard et al., 2020). This dataset covers the period January 1979 to December 2024 at a $0.1° \times 0.1°$ spatial resolution, corresponding to approximately 11×11 kilometres at the equator. Throughout the text, I will use the terms $0.1° \times 0.1°$ grid cell and *location* interchangeably. I extract daily 2-meter minimum and maximum temperature, 2-meter dewpoint temperature, total precipitation, surface solar radiation, and 10-meter wind speed magnitude. These data provide the basis for calculating the rainy season onset and for constructing the weather variables used as controls in my empirical analysis (as detailed in Section 4).

To define the rainy season calendar, I calculate its onset, cessation, and overall length using a dynamic agronomic model that simulates the daily soil water balance (Takele and Dell'Acqua, 2023). The model requires a defined search window, comprising an earliest possible onset, a latest possible onset, and a latest possible cessation date. To set these parameters on a location-specific basis and account for regional differences, I first establish the usual start and end dates of the wet season by analysing long-term daily rainfall patterns, following Dunning et al. (2016).[3] Using these location-specific dates, I define the search window as 90 days be-

[3] Appendix A.1 provides details on how these typical start and end dates for each location are derived from historical rainfall data.

fore and after the usual start for the onset, and 90 days after the usual end for the cessation. Within this window, the model simulates day-to-day changes in soil moisture by combining daily weather data with information on local elevation (Danielson and Gesch, 2011) and the soil's water-holding capacity (Leenaars et al., 2018).

The *onset* date is then identified based on criteria ensuring sustained water availability suitable for crop growth: the ratio of actual-to-potential evapotranspiration (ETa/PET) must exceed 0.5 for seven consecutive days, followed by a 20-day period where simulated plant available water remains above the wilting point. The *cessation* date is defined as the first day when the ETa/PET ratio falls below 0.5 for seven consecutive days, followed by 12 consecutive non-growing days during which plant-available water remains below the wilting point over the root zone. The *length* of the rainy season is then calculated as the total number of days from the identified onset date to cessation date. Finally, the calculated onset and cessation dates for each location and year are converted into the corresponding week number of the year.

This agronomic definition, grounded in soil water availability, is employed for two primary reasons. First, by incorporating potential evapotranspiration (atmospheric water demand) and soil water holding capacity, it moves beyond rainfall thresholds to identify when sufficient moisture is likely available in the root zone to support germination and early plant growth. This provides a robust indicator of the start of the viable planting window, distinguishing true onsets from potential false starts caused by initial rains followed by dry, high-evaporation conditions. While the optimal planting date within this window varies depending on the maturation period of specific crops, this calculated onset marks the earliest point at which planting becomes reliably feasible for rainfed agriculture.

Second, this focus on soil moisture aligns closely with the heuristics farmers themselves report using when making planting decisions. Rather than relying solely on rainfall patterns, farmers commonly assess soil moisture directly to judge planting suitability. Evidence supporting the importance of soil moisture cues in farmer decision-making comes from diverse contexts, including Ghana (Antwi-Agyei et al., 2022), semi-arid India (Giné et al., 2018), and Niger (Marteau et al., 2011). Therefore, an onset definition based on simulated soil water balance provides a measure that is not only agronomically relevant but also resonates with practical farming strategies.

LSMS-ISA and DHS communities are matched to their corresponding $0.1° \times 0.1°$ ERA5 grid cell using their GPS coordinates to link survey data with the calculated onset week and the grid-level weather controls derived from the same ERA5-derived dataset. Finally, to account for broad agro-ecological characteristics, locations are classified into climatic areas based on an updated global Köppen-Geiger climate classification (Metzger et al., 2023). The detailed classes are aggregated into Tropical, Arid, and Temperate zones to ensure sufficient observations for fixed effect estimation. Appendix A.1 provides further details on data processing.

The rainy season calendar over time. My analysis primarily focuses on the onset of the rainy season for two main reasons. First, while the entire rainy season calendar exhibits year-to-year variation, the timing of the onset is particularly critical. It is around this period that farmers make crucial planting decisions which have substantial implications for plot productivity (see Section 2). Second, as this section will show, onset timing across the sampled countries reveals a clear, consistent pattern of change over recent decades, a characteristic less evident in cessation dates.

To estimate the trends, I constructed a panel dataset at the $0.1° \times 0.1°$ grid cell level, covering the period 1979–2024. For each cell and year, I use the day of year (DOY) for the onset, cessation, and length of the rainy season, as detailed in this section. This sample is restricted to only those grid cells that spatially match the locations of households in the LSMS-ISA surveys. I then estimate country-specific trends using a fixed-effects model, with standard errors clustered at the $0.5° \times 0.5°$ grid-cell level (approximately 55×55 kilometres).

Figure 2 presents the country-specific trend analyses for the onset variable. The results consistently show that the onset of the rainy season is occurring later across all sampled countries. The total delay over the historical period is substantial, ranging from 5.8 days in Tanzania to 25.1 days in Nigeria.

Beyond the shift in average onset timing, I also explored whether the year-to-year variability is changing.[4] To do this, I de-trended the onset data using the residuals from the superior-fitting quadratic models. Figure B2 plots the interquartile range (IQR) of these residuals within 5-year periods, revealing differing trends across the sample. In Ethiopia, Malawi, and Tanzania, there is an increasing trend in the IQR. This growing unpredictability has direct consequences for farmers: it makes traditional knowledge less reliable and increases the risk of costly planting errors, such as being misled by a false onset. Conversely, Mali, Niger, and Nigeria show a tendency towards decreasing IQR over time. For farmers in these regions, the rainy season, while still shifting, is becoming more predictable year-to-year, which can lower uncertainty and facilitate adaptation.

Regarding the cessation of the rainy season (Figure B3), trends are more heterogeneous. While Ethiopia, Niger and Nigeria exhibit a tendency towards later cessation dates, Malawi, Mali, and Tanzania show the season ending earlier. Despite these varied patterns, the net effect on the length of the rainy season (Figure B4) is a discernible decrease over time across most of the study area, with the exception of Niger, where the season length remains roughly unchanged. This reduction in the growing period is primarily driven by the consistent and significant delay I observe in the season's onset, reinforcing the importance of onset timing as the principal determinant of changes in the effective length of the agricultural season.

[4] Testing for a time trend in the interquartile range of the onset residuals is conceptually equivalent to testing for a form of heteroskedasticity where the error variance changes over time, e.g., $\sigma^2_{lt} = f(t)$. This speaks to whether the *predictability* of the onset is changing.

4 Empirical strategy

To estimate the causal effect of rainy season onset timing on agricultural, economic, and nutritional outcomes, I leverage temporal variation in onset dates within specific, fine-resolution geographic locations. Identification relies on comparing outcomes within the same $0.1° \times 0.1°$ grid cell across years, exploiting variations in onset timing relative to the cell's long-run average conditions and net of local trends. This approach matches survey units via their geocoordinates to grid cell ℓ, within climatic area a and country c, in year t. The analysis occurs at both the plot and household level for the LSMS-ISA data, and at the woman-level for the DHS data, as detailed in Section 3.

Given this identification strategy, the benchmark specification I estimate is:

$$Y_{i\ell t} = \beta W_{\ell t} + \mathbf{X}'_{i\ell t}\gamma + \mu_\ell + \delta_{cat} + \epsilon_{i\ell t} \tag{1}$$

Here, $Y_{i\ell t}$ represents the outcome of interest for unit i (a plot, household, or individual). Key outcomes from the LSMS-ISA include agricultural productivity at the plot-level, household per capita consumption at the household-level, and farmer adaptation strategies. From the DHS, I examine women's nutritional outcomes at the individual-level. The key independent variable, $W_{\ell t}$, is the calendar week of the rainy season onset. The coefficient of interest, β, thus captures the change in the outcome associated with a one-week delay. As context for the magnitude of these year-to-year shocks, the pooled trend across all sampled cells reveals a significant long-term delay in onset dates of approximately 18 days between 1979 and 2024 (see Section 3.3). Conditional on the model's fixed effects, a one standard deviation shift in onset timing corresponds to approximately two weeks.

Identification of the causal effect β relies on the multi-way fixed effects structure, which isolates plausibly exogenous short-run deviations in $W_{\ell t}$. The model includes two key sets of fixed effects. First, μ_ℓ represents high-resolution grid-cell fixed effects ($0.1° \times 0.1°$). These *location fixed effects* account for all time-invariant characteristics specific to location ℓ, such as its average climate, soil properties, and market access. However, detecting the potentially subtle impact on nutritional outcomes is challenging, particularly with the more limited sample of agricultural households in the DHS of approximately 59,000 woman-year observations. To ensure the model retains sufficient statistical power for this specific analysis, the resolution for these fixed effects is therefore coarsened to the $0.5° \times 0.5°$ grid-cell level. Second, δ_{cat} represents a full set of fixed effects for each unique combination of climatic area, country, and year. These *spatially specific time effects* are crucial for identification for two primary reasons: they flexibly control for any underlying secular trends in outcomes, and they absorb any unobserved shocks common to a climatic region within a given country in a particular year. As detailed in Appendix A.2, a grid cell ℓ is considerably smaller than any climatic area within a country (Fig-

ure A1), and each climatic area-by-country group contains a large number of unique grid cells with sampled households (Table A3). This nested variation allows the model to distinguish between broad, regional shocks and time-invariant local characteristics.

The vector $\mathbf{X}_{i\ell t}$ incorporates available time-varying controls to account for remaining observable differences. First, and most importantly for the conceptual framework outlined in Section 2, a comprehensive set of weather controls is included to separate the effect of onset timing from aggregate seasonal weather. By accounting for total rainfall and temperature patterns, these controls allow the model to estimate the net impact of a shift in seasonal timing, holding aggregate conditions constant. The analysis then proceeds to test the hypothesis that adaptation failures are a key driver of these net negative impacts. This distinction is crucial for policy. As the subsequent analysis will show, a large portion of these damages stems from identifiable adaptation failures—rather than inevitable biophysical consequences—which implies that these losses can be mitigated with targeted interventions, such as improved weather forecasts. These weather controls, measured at the grid-cell level over the calendar year, include: the number of days with precipitation below the 25th and above the 75th percentiles of the historical distribution, daily maximum relative humidity, average daily temperature, total annual precipitation, and harmful degree days (days with maximum temperature exceeding the 90th percentile). Appendix B.3 presents estimates using these same controls aggregated over different temporal windows.

Second, specific controls are also included depending on the level of analysis. For plot-level outcomes, controls include household size, total farm area, and plot manager characteristics (age, sex, marital status, education). For household-level outcomes, controls include indicators for any household member's education, non-farm enterprise ownership, number of plots, and an urban indicator. For individual-level nutritional outcomes from the DHS, controls include the woman's years of education, her age and age squared, household head's age, household size, and an urban indicator.

Finally, $\epsilon_{i\ell t}$ represents the idiosyncratic error term. Because outcomes and the onset of the rainy season exhibit spatial correlation, I cluster standard errors at a larger spatial unit: the $0.5° \times 0.5°$ grid-cell level. All descriptive statistics and regression analyses use the survey weights provided by LSMS-ISA and DHS. These weights are adjusted for the pooling of multiple survey waves per country to ensure representativeness. For plot-level analyses, household weights are further rescaled by the number of plots per household. Appendix A.3 provides full details.

The demanding fixed-effects structure may raise concerns about whether sufficient variation remains for identification. For transparency, all regression tables report the number of observations used for identification and those dropped due to singleton groups. Furthermore, I provide two pieces of evidence in Appendix B.2 to address these concerns directly. A variance de-

composition (Figure B5) confirms that substantial within-location variation remains in the key variables after accounting for fixed effects. A plot of the residualised onset of the rainy season (Figure B6) shows a symmetric, zero-centred distribution consistent with an idiosyncratic shock, providing supportive evidence for the exogeneity assumption.

5 Results

This section presents the main empirical findings. I begin by showing that a delayed rainy season onset reduces agricultural productivity and, in turn, harms household welfare, as measured by consumption and nutritional outcomes. I then investigate how farmers adapt to this shock, focusing on adjusting their planting schedules to offset losses. Finally, I explore a key challenge to this strategy: how a false onset, a misleading signal that prompts early planting, can lead to costly mistakes and worsen outcomes.

5.1 The impact of rainy season onset on agricultural productivity

The analysis begins by assessing the impact of rainy season onset timing on agricultural productivity. My findings indicate that a delayed start to the rainy season causally reduces agricultural yields.

Table 1 presents the estimated impact of a one-week delay in the rainy season onset on agricultural productivity, measured as the logarithm of yield in constant 2020 US dollars per hectare. Across all specifications, a later onset is associated with a statistically significant reduction in yields. In my preferred specification (Column 3), which includes a comprehensive set of location fixed effects ($0.1° \times 0.1°$ grid cell), climatic area by country by year fixed effects, weather controls, and plot-level demographic and farm characteristics, a one-week delay in the rainy season onset causes an approximate 2% decrease in agricultural yields. This effect is statistically significant at conventional levels.

The magnitude of this impact is economically important. To put this into perspective, a one standard deviation delay in the onset of the rainy season–which corresponds to one and a half weeks, conditional on the included fixed effects–leads to a reduction in agricultural productivity of about 3%. This estimate is remarkably stable: the impact of a one-week delay is a consistent 2% reduction in yields across all specifications, irrespective of the inclusion of weather or plot-level controls (Table 1, Columns 1-3).

To further explore the nature of this relationship, I check whether the impact of onset timing is uniform or asymmetric. For this, I did not use the continuous onset week variable. Instead, I binned each location-year's onset timing based on its historical distribution within its $0.5° \times 0.5°$ grid cell. Specifically, I created indicator variables for whether the onset fell into the

bottom, middle, or top tercile. I then re-estimated equation (1) using indicator variables for the earliest and latest terciles, with the middle tercile serving as the omitted category. The results, presented in Figure 3, reveal a stark asymmetry. An onset occurring in the earliest third of the historical distribution has an estimated effect on log yields of -0.3%, which is statistically indistinguishable from zero when compared to an onset in the middle tercile. In contrast, an onset occurring in the latest third of the historical distribution leads to a statistically significant reduction in log yields by approximately 12%, compared to an onset in the middle tercile. This confirms that the adverse productivity consequences are predominantly driven by significantly later-than-usual rainy seasons. This finding holds when controlling for grid-cell-specific linear trends (Figure B7), which I implement at the $0.5° \times 0.5°$ binning level. This check motivated by recent concerns in the literature about potential biases in binned climate analyses (Jones et al., 2025). See Appendix B.2 for further details.

My benchmark empirical strategy, detailed in Section 4, imposes a stringent requirement on the data by leveraging only within-cell variation in onset timing, net of controls and local trends. Consequently, any remaining threat to identification would need to stem from an unobserved, cell-level time-varying factor that is correlated with the year-to-year deviations in onset timing and directly affects agricultural productivity through a channel other than the onset timing itself. Such a specific confounder is arguably less likely given that the timing of the rainy season, and its interannual variability, is substantially influenced by large-scale climatic patterns, such as the El Niño Southern Oscillation, the Indian Ocean Dipole, and Tropical Atlantic Variability, which affect precipitation regimes and climate extremes across Sub-Saharan Africa (Zita et al., 2025).

Nevertheless, I verify the robustness of this core finding through an extensive series of checks, detailed in Appendix B.3. First, the negative and significant effect of onset timing remains stable even when employing different definitions of local trends by varying the geographic level at which such trends are defined; this is noteworthy as altering the local trend specification effectively changes the definition of the onset shock by modifying the variation used for identification (Table B3). Second, the impact is confirmed to be contemporaneous: a placebo test using future onset (t+1) shows a coefficient close to zero and statistically insignificant, while lagged onset (t-1) also has no significant effect on current yields, indicating the impact is not persistent across years (Figure B8). Third, when examining other features of the rainy season, I find that cessation timing or rainy season duration do not significantly affect productivity once onset is controlled for, reinforcing the primary role of onset (Table B4). Fourth, the onset effect is not merely capturing other weather-related shocks, as it remains robust to the inclusion of several weather controls defined at finer temporal frequencies (quarterly and monthly) that capture variations in temperature, precipitation patterns including droughts and floods, and relative humidity (Table B5). Fifth, the finding is robust to various alternative transformations of the yield variable (Table B6) and alternative measures of output (Table B7). Notably, the

magnitude of the yield reduction is larger when measured in kilograms per hectare than in US dollars, suggesting a potential local price increase in response to the negative supply shock. To address this potential attenuation bias, I re-estimate the model valuing output at constant national median prices per crop (see Section B.3 for details); this yields a larger negative impact of approximately 2.7%, confirming that local price responses likely dampen the benchmark estimate (Table B8). Additionally, to ensure results are not driven by composition effects from aggregating diverse crop types, I restrict the sample to cereal crops only; the negative impact remains consistent in magnitude and significance (Table B9). The shock also affects the extensive margin, as a one standard deviation later onset increases the probability of crop failure by approximately 5 percent. Lastly, statistical inference is robust, holding under different clustering assumptions for standard errors (Table B10) and confirmed by permutation-based inference (Figure B9).

5.2 Impact on household welfare: Consumption and nutrition

The reduction in agricultural productivity documented above has direct consequences for household welfare, with economic impacts that affect the broader local economy.

First, the shock to agricultural output translates into lower household spending. For agricultural households, the group most directly affected, a later onset of the rainy season leads to a statistically significant decrease in per capita consumption. As shown in Table 2, Columns (1)–(3) indicate that a one-week delay reduces per capita consumption by approximately 1.0% to 1.2%. To put this magnitude in perspective, a one standard deviation shift in onset timing reduces consumption by approximately 1.5%.

Importantly, these negative economic consequences are not confined to agricultural households. As shown in Columns (4)-(6) of Table 2, the effect on per capita consumption remains statistically significant for the full sample. This indicates the presence of local spillover effects, as the initial shock to agricultural income propagates to the non-farm economy. A primary channel for this spillover is a contraction in local demand: as farmers' incomes fall, they reduce their spending on local goods and services, depressing economic activity for the entire community. This is consistent with the finding in Table B16 that a late onset is associated with a significant reduction in wage employment opportunities. Beyond this direct spillover, a broader general equilibrium effect could also be at play, operating through prices. A smaller harvest may increase local food prices, which would reduce the real purchasing power of all households and transmit the shock throughout the entire local economy.

Table 3 shows that a delay in the onset of the rainy season leads to a statistically significant deterioration in the short-term nutritional status of women in agricultural households. Specifically, a one-week delay is associated with a 0.007 standard deviation decrease in the weight-for-height z-score. While this effect is statistically significant, its magnitude is modest. The standardised

effect indicates that for a one-standard-deviation delay in the onset, the z-score decreases by approximately 2% of the sample mean.

Importantly, this nutritional decline is not severe enough to increase the likelihood of women being classified as clinically wasted or underweight, as shown by the statistically insignificant results in columns (2) and (3). The size of the impact on weight-for-height is consistent with the analysis of productivity and consumption decreases. Table 3 shows that a delay in the onset of the rainy season leads to a statistically significant deterioration in the short-term nutritional status of women in agricultural households.

This negative shock, however, does not appear to be transmitted to young children. Using the anthropometric data for children, I find no statistically significant impact of a delayed onset on weight-for-height z-scores or the probabilities of being wasted or considered underweight (Table B19). This divergence—a nutritional cost for mothers but not for children—could reflect incomplete risk sharing within households, where adults, particularly mothers, shield younger members from shocks, potentially at their own expense (Dercon and Krishnan, 2000).

The main analysis employs location fixed effects at the $0.5° \times 0.5°$ grid-cell level to maintain statistical power. Table B20 re-estimates the model with the finer $0.1° \times 0.1°$ fixed effects used in the agricultural analysis. In this more demanding specification, the point estimate remains negative and similar in magnitude, but the larger standard errors render it statistically insignificant.

Furthermore, the negative nutritional impact of a late onset extends beyond agricultural households. When the sample is expanded to include all women, a delayed onset still causes a statistically significant decrease in weight-for-height z-scores (Table B21), suggesting the presence of general equilibrium spillovers that affect the wider community.

5.3 Farmer adaptation strategies in response to onset variation

Given these significant adverse impacts, I next analyse how farmers respond to year-to-year variability in the start of the rainy season. I examine adjustments in key agricultural practices, focusing on four primary margins: the timing of planting and other on-farm activities, the intensity and type of input use, choices regarding crop types, and the allocation of household labour to off-farm sectors.

The most direct way farmers can respond to a varying onset is by adjusting their planting schedule. Table 4 (Column 1) shows that a later start of the rainy season leads to a delay in farmers' planting. Specifically, a one-week delay in the agronomically-defined onset week causes an average delay in the approximated first planting day of the year by 0.56 days (p-value < 0.05).

The estimated responsiveness from the main text—an average planting delay of just half a day

for a full week of onset delay—is difficult to interpret directly. The coarse, monthly nature of the planting data introduces two potential and opposing sources of measurement error, making the direction of the net bias theoretically unclear.

The first and more intuitive source is attenuation bias. As shown in Figure B13, the median planting *month* is remarkably stable over time, with a Pearson correlation of approximately 0.90 across survey waves. This stability, however, can mask significant daily-level adjustments. For the 90% of plots where planting is completed within a single calendar month, any shift in planting dates that does not cross a month-end boundary (e.g., a 10-day delay from June 5th to June 15th) is recorded as a zero-day change. This systematically understates the true responsiveness of farmers' planting decisions.

Conversely, the data may also suffer from an opposing boundary-crossing bias that exaggerates responsiveness. If a small, multi-day delay in the onset causes a farmer's planting date to shift from the end of one month to the beginning of the next (e.g., a 5-day delay from May 30th to June 4th), the monthly approximation would record this as a 30-day shift. In such cases, the regression would overstate the true sensitivity to the onset shock.

To resolve this ambiguity and estimate the net effect of these two opposing biases, I implement a simulation exercise to recover the unbiased daily-level coefficient. The full methodology is detailed in Appendix B.6, but in brief, I use the observed monthly planting data to fit location-specific daily probability distributions, which I then use in an iterative search algorithm to find the true daily coefficient that would produce my observed monthly coefficient. I conduct this exercise using two different distributional assumptions: a parametric skew-normal and a non-parametric kernel density estimation (KDE).

The simulation results confirm that while both biases exist, the attenuation from unobserved within-month adjustments is the dominant effect. As shown in Figure B12, the corrected daily coefficient is larger in magnitude than the original estimate. However, the magnitude of this correction is modest. Using the KDE-based draws, the corrected coefficient is only 29% larger than the biased estimate. This finding is significant. While the simulation corrects for a statistical bias, the relatively small size of the correction implies that the true daily-level response of farmers is itself limited. In other words, the measurement error is not masking a large underlying behavioral response. This suggests that farmers are not adjusting their planting schedules to the onset of the rains as flexibly as might be expected, pointing towards the presence of other frictions or decision-making heuristics that Section 5.5 explores.

Beyond timing, a later onset week causes a statistically significant increase in seed quantity. As shown in Table 4 (Column 4), the coefficient of 0.032 indicates that a one-week delay in onset leads to an approximate 3.2% increase in seed quantity used per hectare. This could reflect a strategy to ensure adequate plant density under potentially less favourable or more uncertain early growing conditions. Alternatively, an increase in seed quantity might also indicate

re-planting efforts following early crop failure, a channel for which I will present suggestive evidence when discussing the impacts of false onsets in Section 5.5. This increase in quantity, however, does not translate into higher expenditure on seeds. Table B13 (Column 1) in Appendix B.5 shows no significant change in seed value per hectare, suggesting that farmers likely use more of the same type of seeds they customarily use, such as saved seeds or common local varieties, rather than switching to more expensive, potentially improved, varieties as an immediate response to onset variation.

In contrast to these adjustments in planting timing and seed quantity, I find limited evidence that onset timing causes short-run changes in other major input applications. Table 4 (Column 5) shows no statistically significant change in the probability of using fertiliser, with the small coefficient indicating a negligible and insignificant effect on the likelihood of fertiliser application. Further analysis in Table B13 (Columns 2 and 3) indicates no significant causal effect of onset week on the quantity of inorganic fertiliser applied (kg/ha), nor on the value of inorganic fertiliser (USD/ha), where the point estimate for value suggests a decrease but is not statistically distinguishable from zero. Similarly, the probability of using pesticides (Table 4, Column 6) does not respond significantly to onset timing. I also find no significant adjustments in on-farm labour intensity. Table B13 (Columns 4 and 5) shows that onset week does not significantly change either total family labour days per hectare or hired labour days per hectare; the point estimate for hired labour days indicates a decrease, but this effect is not statistically significant.

When examining crop choice, I find no significant shifts in the main crop grown in response to the onset week of a particular year. Table 4 (Columns 7-9) shows no statistically significant change in the probability that the main crop on the plot is a cereal, a tuber, or a legume. However, a more detailed analysis of which crops are present on a plot reveals a specific adjustment margin. While a later onset does not affect the probability of a plot containing cereals or legumes, it does lead to a small but statistically significant decrease in the probability of it containing tubers (Table B14). This aligns with the finding on the composition of plot value, where a later onset also causes a minor decrease in the share of value attributed to tubers. Overall, these findings suggest that while major decisions regarding crop portfolios are largely inflexible in the face of year-to-year onset variability, farmers may make minor adjustments away from certain crops like tubers when the season starts late.

Furthermore, I explored whether onset timing influences broader land use decisions at the household level. As detailed in Table B15, I find no statistically significant impact of onset week on extensive margin decisions such as the number of fallow plots or the number of plots cultivated. Similarly, there is no significant effect on the overall scale of land operation, as measured by the logarithm of total cultivated area or the logarithm of total farm size. This suggests that households do not significantly alter these decisions in immediate response to year-to-year variations in onset timing.

Finally, I investigate whether households reallocate labour to off-farm sectors as an adaptive response. Contrary to a simple story of labour substitution, the evidence suggests that a delayed onset reduces local wage-employment opportunities. Using individual-level data from the LSMS-ISA, I find that a one-week delay in the onset of the rainy season causes a small but statistically significant decrease in the probability of individuals engaging in any wage work. Specifically, Panel A of Table B16 shows a 0.1 percentage point decrease for wage work over the past 12 months, representing a 2% reduction relative to the sample mean, and a 0.2 percentage point decrease for work in the past 7 days, a 3.7% reduction. Panel B provides evidence that this finding holds for the general working-age population and is not driven by other members of the household, as it remains significant when the sample is restricted to only household heads. Furthermore, there is no significant effect on working specifically in agriculture or services, nor on engagement in household business activities.

Taken together, these results suggest that rather than households readily shifting their labour to off-farm work, opportunities for local wage employment diminish when the agricultural season starts late, consistent with a negative shock to the local economy. An alternative explanation involving an increase of on-farm labour demand seems unlikely, as direct measures of on-farm labour days (both family and hired) do not show a significant increase; if anything, the point estimates for hired labour days, while not statistically significant, suggest a decrease (Table B13).

5.4 Heterogeneity of impacts: Vulnerability, resilience, and adaptation

While the average effect of a delayed onset on productivity is negative, the impact is not uniform. To identify the characteristics associated with vulnerability and resilience, I explore how the effect varies across key farmer, household, and technological dimensions. The results, presented in Table 5, reveal a significant and economically meaningful degree of heterogeneity.

Socioeconomic and demographic factors play a role in shaping a household's resilience. Column (1) shows that the negative impact of a delayed onset is exacerbated on plots managed by women. The interaction term is negative and significant (-0.006), indicating that the yield loss for female-managed plots is approximately a third larger than for male-managed plots in the face of the same one-week delay. Conversely, human capital appears to be a key mitigating factor. As shown in column (2), the interaction with a manager's formal education is positive and significant (0.006), implying that for educated managers, the yield penalty is reduced by over a quarter. Economic resources provide a similar buffering effect. Column (3) shows that households with above-median assets suffer a significantly smaller yield loss, suggesting that wealthier households are better able to deploy capital to smooth the shock. Finally, the analysis highlights the critical role of technology. For irrigated plots, the positive and significant interaction term (0.019) in Column (4) almost perfectly cancels out the baseline negative effect.

The use of improved seeds, shown in Column (5), also provides a statistically significant, albeit smaller, mitigating effect.

I explore additional dimensions of heterogeneity in Appendix Table B24. Notably, I find that market integration—measured by whether the household sold any produce (Column 5)—is associated with significantly greater resilience, consistent with the hypothesis that market access facilitates consumption smoothing or is correlated with higher adaptive capacity. Furthermore, while both female land management and poverty (lower assets) predict greater vulnerability, it is worth noting that coming from a poor household and having at least one female-managed plot are not strongly correlated in this sample; thus, while both factors drive effects in the expected direction, they likely capture distinct sources of disadvantage.

A crucial question that arises from these findings is whether these more resilient groups achieve better outcomes because they adapt their practices more effectively. I test this by examining if they adjust their planting schedules more responsively to a delayed onset. The results, shown in Table B25, suggest this is generally not the case. There is no statistically significant evidence that farmers with more education or more assets adjust their planting dates any differently than their less-resourced counterparts. The findings for technology are more nuanced. For irrigated plots, there is weak evidence that farmers delay planting *less*, perhaps because access to water reduces the urgency of optimal timing. In contrast, farmers using improved seeds are the only group to show a significantly more responsive adjustment, delaying their planting by an additional 0.34 days for every week of onset delay.

Taken together, these findings paint a clear picture. While all rainfed agriculture is exposed to shifting seasonal timing, economic damages are disproportionately borne by marginalised producers. Resilience is systematically linked to education, wealth, and access to adaptive technologies. However, the source of this resilience is not fully explained by a comprehensively superior management response to the shock. For what is arguably the most critical margin—planting time—the analysis reveals that most resilient farmers do not adapt their schedules more effectively. The response on other input use is more nuanced. While wealthier and more educated farmers show a statistically significant differential increase in the use of inputs like pesticides, the magnitude of this adjustment is economically small (see Table B26). The most meaningful adaptive response observed is in seed use, where more educated farmers significantly increase the quantity planted after a late onset. This selective and partial pattern of adaptation—absent on the crucial timing margin, marginal for some inputs, and robust only for seeds—suggests that the resilience of better-off farmers stems primarily from their greater capacity to buffer shocks, rather than from consistently superior real-time management. This implies that significant informational barriers to optimal adaptation remain, leaving room for improvement across the board. Furthermore, an examination of how these heterogeneous impacts have changed over the sample period reveals a concerning trend: the vulnerability gap for female-managed plots has significantly widened in more recent years, while the protec-

tive effect of formal education has strengthened (Table B12). This suggests that underlying inequalities may be amplifying the impacts of shifting seasonal timing over time.

5.5 Limits to adaptation: False onsets and farmer awareness

While strategically delaying planting can significantly mitigate productivity losses—a finding supported by an instrumental variable analysis presented in Appendix B.7—the fact that not all farmers perfectly adjust their timing suggests that adaptation is not free of risk. In the sample, farmers appear to plant too early, and therefore miss the agronomically correct window, on approximately one-third of all plots. Specifically, 34% of plots are planted before the modelled onset date—recall that this is the earliest point at which soil moisture is sufficient to reliably sustain a crop. Planting before this window opens is a high-risk strategy, as it exposes seeds to early-season dry spells that can lead to crop failure. While one might hypothesize that this early planting is driven by off-farm labour commitments, this is unlikely to be the primary explanation, as only 3% of working-age individuals in the sample report having any wage employment. A more plausible explanation is that farmers' decisions are constrained by misleading weather signals and imperfect awareness of onset delays. In this section, I provide suggestive evidence for both of these channels.

A primary source of confusing signals for farmers is a *false onset* of the rains. I develop a precise, data-driven definition to capture this specific weather pattern. The process first defines a relevant *search window* for each location, symmetrically set at 60 days around its long-term average onset date. This local baseline is calculated as the mean onset for each $0.5° \times 0.5°$ grid cell using annual data since 2000. This window helps to avoid identifying spurious weather events far outside the planting season. Within this plausible window, a false onset is identified by a two-part sequence. First, there must be a *wet spell* that could plausibly trigger planting: two consecutive days with at least 20mm of cumulative rainfall (Marteau et al., 2011). Second, to be considered *false*, this must be followed by a damaging *dry spell*: seven consecutive days with less than 0.1mm of rain each, which must begin within 20 days of the wet spell. For a location-year to be classified as a false onset, this entire wet-then-dry sequence must conclude 21 days before the date of the true agronomical onset, to avoid any overlap. This is a relatively rare event, affecting 4.5% of plot-years in the sample, and on average, it occurs 26 days before the true onset. To test its impact, I extend my benchmark regression to include a dummy variable for whether a false onset occurred and its interaction with the onset week, using a fully flexible specification where all fixed effects are also interacted with the false onset indicator.

A false onset should only harm productivity if it tricks farmers into planting at the wrong time. The results in Table 6 show this is precisely what happens. Column (1) reveals that when a false onset occurs, farmers react by planting significantly earlier than they otherwise would. The interaction term shows that for each week the true onset is delayed, the presence of a

false signal causes planting to occur approximately one day earlier. The consequences of this mistake are evident in Column (2). For each week of onset delay, a false onset is associated with an additional 10% increase in seed quantity, a substantial rise that strongly suggests a narrative of initial crop failure followed by costly replanting. Ultimately, this disruption has a dramatic impact on productivity. As shown in Column (3), the negative effect of a one-week delay in the rainy season is amplified more than three-fold when it is preceded by a false onset, with the total yield loss increasing from 2.3% to 8.9%. While the evidence on planting dates and seed quantities points to timing errors as the primary damage channel, other agronomic or economic pathways could also be at play. I explicitly test for these alternatives in Table B23. The results show no significant change in pesticide use, casting doubt on the hypothesis that false onsets primarily harm yields by increasing pest prevalence. Similarly, I find no effect on the use of hired labour, ruling out a labour market response as an alternative damage channel. The only significant effect is a sharp decrease in fertilizer use. This finding reinforces the main hypothesis: after a false onset induces a planting error and subsequent crop failure, farmers reduce their investment on the compromised plot. This core finding is robust to the specific parameters of the false onset definition (Table B22). Results are qualitatively unchanged when using alternative definitions. For instance, I test for a wider search window (70 days) and require a longer, more damaging dry spell (10 days) to ensure the signal is truly relevant. In these alternative specifications, a false onset continues to significantly worsen the impact of a delayed season on yields, with the total negative effect remaining large and statistically significant.

Beyond being misled by false onsets, a second informational barrier appears to be the low salience of the underlying shock itself. To explore whether farmers perceive and report the negative shock from a delayed rainy season onset—a potential barrier to adaptation—I use self-reported, plot-level data from the LSMS-ISA surveys asking if a plot experienced a negative shock. I examine two binary outcomes: suffering (1) any crop shock, and (2) a drought shock specifically. I estimate equation (1), also replacing the continuous onset variable with indicators for the top and bottom terciles of the historical distribution to test for non-linearities in perception. The results, presented in Table 7, indicate that farmers' awareness is indeed non-linear. A typical delay in the rainy season does not significantly increase the probability that a farmer reports a negative shock, suggesting the effect may be difficult to distinguish from normal year-to-year variation. Awareness is triggered only by extreme events; farmers are significantly more likely to report a shock only when the onset is exceptionally late (in the top tercile). Even then, they tend to misattribute the shock primarily to drought rather than the shift in timing. These patterns of non-linear awareness and misattribution hold when controlling for grid-cell-specific linear trends. This robustness check, detailed in Appendix B.2 (Table B2), addresses potential biases from unobserved local trends highlighted by Jones et al. (2025).

Given these apparent informational constraints, access to timely information should mitigate

losses by helping farmers better anticipate the true onset or react more swiftly (Burlig et al., 2024). To explore this, I examine whether household ownership of mobile phones, radios, and televisions—assets commonly associated with information access—moderates the impact of a delayed onset on productivity. Following the same approach as in Section 5.4, I find that owning any of these assets is associated with significantly smaller yield losses following a delayed onset. These plots experience yield reductions that are approximately 20% smaller than those in households without such assets (Table B27). To mitigate concerns about omitted variables, these specifications control for a household's level of wealth (proxied by an indicator for being below the median asset index), which itself is not strongly correlated with ownership of any single information asset. While this evidence is correlational, it provides suggestive support that informational barriers are indeed a key constraint, and that improving farmers' access to information could enhance resilience.

5.6 Projecting future damages

A key question is whether the estimated short-run effects of weather shocks can inform our understanding of the long-term consequences of climate change. A common approach to testing for long-run adaptation is to compare locations with different initial climatic conditions; for instance, locations with higher average temperatures might be better adapted to a heatwave (Dell et al., 2014). For the timing of the rainy season, however, the absolute day of the year is less important than its deviation from local expectations. I therefore adapt this approach to compare locations based on their long-term trends in onset timing. If farmers adapt to persistent environmental change, one might expect those in locations where the onset is secularly arriving later to be more resilient to a late-onset shock than farmers in areas with a more stable climate.

To test this, I classify locations as either *trending* or *non-trending* based on the long-run linear trend in onset timing over the last four decades, estimated at the $0.5° \times 0.5°$ grid-cell level. A location is defined as trending if it exhibits a statistically significant positive trend. I then expand equation (1) by interacting the onset week of the rainy season with an indicator for being in a trending location.

The results, presented in Table 8, provide strong evidence of a failure to adapt. The negative impact of a delayed onset on agricultural yields is driven entirely by locations already experiencing a long-term climate trend toward later rainy seasons. In non-trending locations, a later-than-usual onset has no statistically significant effect on productivity. In stark contrast, in trending locations, a one-week delay causes a significant yield loss. These vulnerable, trending locations constitute the vast majority of the sample, representing approximately 90% of plot-wave observations.

This lack of long-run adaptation is further corroborated by two pieces of evidence: a failure to learn even from recent, year-to-year experience, and an increase in the effect of a delayed onset

on productivity in recent years. First, if farmers were updating their strategies based on past weather, a delayed onset in the previous year ($t-1$) should lead them to adjust their behavior in the current year (t). I test for this inter-annual learning across all major adaptation margins and find no evidence of a systematic response. As shown in Table B17, the previous year's onset has no statistically significant effect on planting dates, input use, or harvest timing. Second, I test whether the average impact of onset delays has diminished within the sample period itself, which could indicate adaptation over the medium term. Interacting the onset week with an indicator for later survey waves in countries with sufficient panel length (Ethiopia, Malawi, Nigeria) reveals that the negative impact has significantly increased in more recent years (Table B11).

Taken together, these three pieces of evidence suggest a persistent failure to adapt, which provides the justification for using short-run estimates to project future damages. To quantify this threat, I project the Net Present Value (NPV) of damages from 2025 to 2050, presented as a percentage of 2024 real GDP. The calculation combines my benchmark estimate of a 2% yield loss per week with country-specific climate shocks. Given the evidence that the negative impact of onset delays has intensified in recent years (Table B11), employing the average effect across the entire sample period likely yields a conservative estimate of future damages. These shocks are scenario-dependent: under the Business as Usual scenario, I use the full projected onset delay from Figure B1, while the Toward Sustainability scenario assumes a 50% less severe delay, reflecting climate mitigation efforts. The annual loss is calculated in absolute monetary terms by combining these shocks with FAO scenarios for future real agricultural shares and real GDP growth. The model also accounts for a data-driven expansion of irrigated land, which is forecast based on historical trends (see Figure C1); however, these projections suggest that only Ethiopia, Mali, and Tanzania will see meaningful increases in irrigation by 2050. Finally, I use a time-varying discount rate from the DICE-2023 model. The full methodology is detailed in Appendix C.

Figure 4 presents the main finding. Cumulative damages will be high under either climate path, but a more sustainable trajectory can offset a sizeable portion of the economic losses. Under the Business as Usual scenario (solid red line), cumulative discounted damages by 2050 are substantial, reaching approximately 10% of 2024 GDP in Ethiopia and Mali. The shaded area quantifies the Sustainability Benefit of a less severe climate shock combined with a more resilient economy, averting over $4.3 billion in damages for Ethiopia and $20.2 billion for Nigeria.

Of course, this projection is pessimistic as it does not allow for long-run adaptation beyond the expansion of irrigation. Figure 5 addresses this by comparing the baseline Business as Usual damages with an optimistic counterfactual where farmers adapt successfully, causing the impact of onset delays to decrease linearly to zero by 2050. While this Adaptation Benefit is significant, the remaining costs are still substantial.

In my model, I already include irrigation as a key adaptation technology. However, I found its projected effect is limited. To test this further, I simulated an aggressive (and costly) scenario that expanded irrigated farmland to 20% of the total by 2050. Comparing the results to the baseline reveals that the benefits of this expansion are modest for most countries (see Figure C2).

Finally, it is crucial to acknowledge the large uncertainty surrounding these estimates. The projections for GDP, agricultural shares, climate shocks, and adaptation capacity are all subject to uncertainty. While it is not possible to model this fully, Figure C3 and Table C1 present the Business as Usual projections using the 95% confidence interval of the benchmark yield loss coefficient. They reveal that even under the most optimistic lower-bound estimate, the cumulative economic damages remain large and economically significant for all countries.

6 Conclusion

This paper investigates the consequences of shifting seasonal timing for agricultural communities in Sub-Saharan Africa. Using decades of high-resolution climate data, I document a clear trend: the onset of the rainy season is occurring progressively later across six major African countries, leading to shorter growing seasons. The analysis shows that these delays have significant negative consequences. A one-week delay in the onset causally reduces agricultural yields by approximately 2%. When the onset is exceptionally late (in the latest tercile of the historical distribution), these losses deepen to roughly 12%. To put this magnitude in perspective, this impact is remarkably comparable to the damages caused by other major climate shocks, such as the 12% yield reduction from severe drought found by Corno et al. (2020) and the 11% decline from heat stress estimated by Aragón et al. (2021). This highlights that shifts in seasonal timing are a first-order threat comparable to traditional weather shocks.

These damages are not evenly distributed; they are borne disproportionately by the most vulnerable, particularly on plots managed by women, while education, wealth, and technology build resilience. In response, farmers adapt primarily along low-cost margins—delaying planting and increasing seed quantities—but these adjustments are insufficient to prevent losses to productivity and welfare.

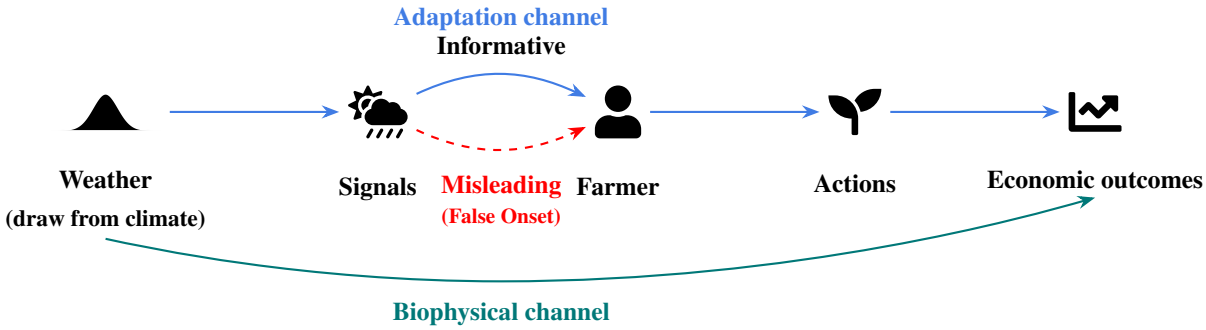
A central puzzle emerging from the analysis is why farmers do not fully leverage planting date adjustments, a strategy this paper shows is both powerful and inexpensive. The evidence points strongly toward informational constraints as the primary barrier. This conclusion is reinforced by the finding that even farmers who are more resilient to the initial shock do not systematically adapt their planting timing more effectively. The problem is not merely a lack of resources but a widespread informational challenge, best illustrated by the *false onset*. This paper demonstrates that these misleading early rains trick farmers into planting prematurely,

a mistake that ultimately more than doubles the negative impact of a late onset of the rainy season.

The scale of these projected damages lends urgency to several clear policy implications. First, the demonstrated vulnerability of farmers to false onsets highlights a high potential return for investments in accessible and reliable short-range weather forecasts. Such services could create enormous value not by replacing farmer knowledge, but by helping to differentiate a true onset from a false one. Second, the heterogeneity of the impacts underscores that interventions must be targeted. Policy should not only deliver weather information but also work to reduce the underlying vulnerabilities that amplify damages, for instance, by improving women's access to resources and promoting education. Third, the limited adaptation along the crop choice margin suggests a need for continued investment in agricultural R&D to develop and disseminate seed varieties with more flexible growing periods. Finally, the evidence of negative spillovers to the entire local economy points to a need for social protection systems that are responsive enough to respond to shocks in seasonal timing, not just catastrophic events like major droughts.

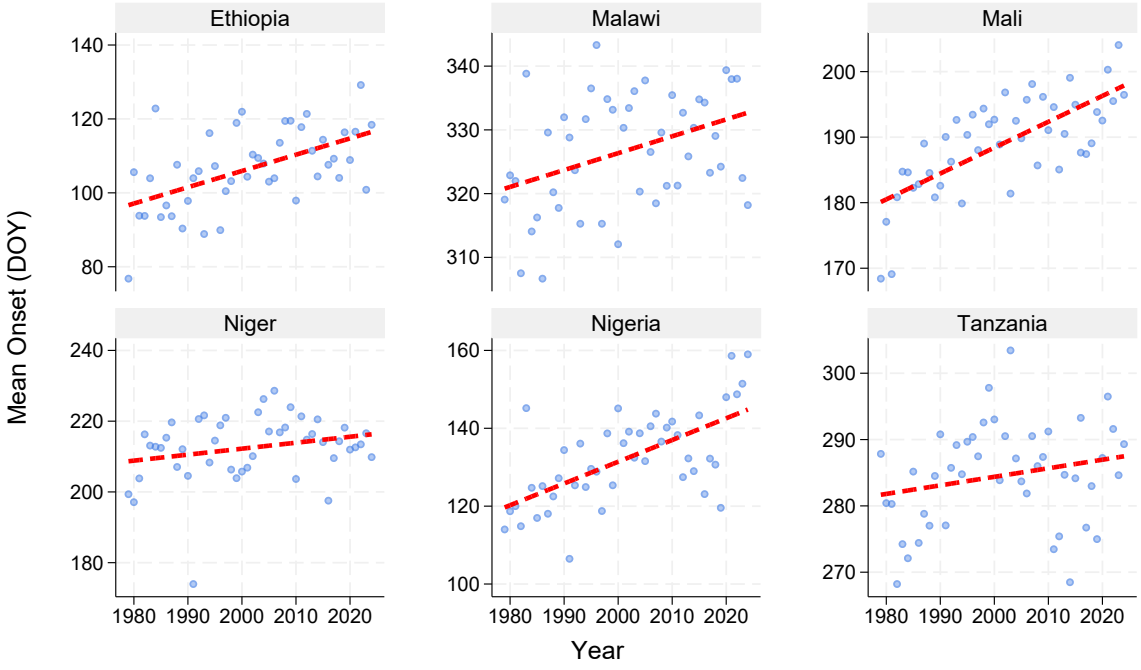
The findings in this paper point to several promising avenues for future research. Building on this analysis, future work using true panel data could offer deeper insights into the dynamic, household-level adaptation strategies that unfold over multiple seasons. Further research could also explore a broader set of adaptation margins not explicitly covered here, including migration, the role of social networks, and investments in irrigation. Finally, linking these climate shocks to local market data would be a valuable extension, illuminating how market access and price signals mediate both the damages and the capacity for adaptation, providing a richer understanding of this critical climate vulnerability.

Figure 1: Channels of impact from shifting seasonal timing



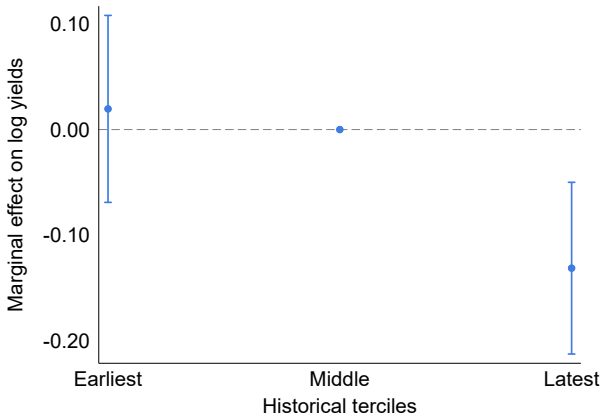
Note. Two primary channels through which the timing of the onset of the rainy season affects economic outcomes. The *Biophysical channel* represents the direct impact of weather conditions (e.g., temperature, moisture) on crop growth and productivity, independent of farmer actions. The *Adaptation channel* represents how farmers' responses to weather signals affect productivity. The empirical strategy first estimates the net impact of onset timing (controlling for the biophysical channel) and then uses subsequent analysis to identify the adaptation channel as the key mechanism, as detailed in Section 4.

Figure 2: Country-specific trends in rainy season onset



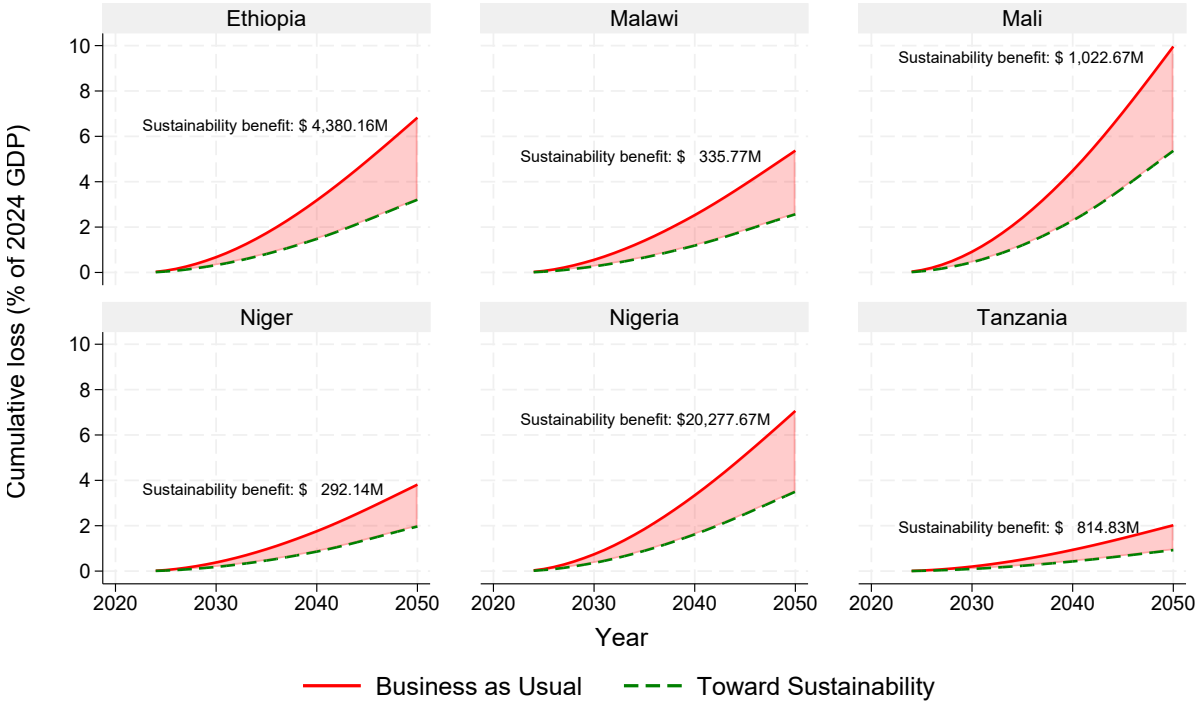
Note. Each panel in the figure displays the annual mean onset day of the year (DOY) for $0.1° \times 0.1°$ grid-cells matched to LSMS-ISA survey locations within the respective country from 1979–2024. The *Mean onset (DOY)* is the average across space for each year of the day of the year when the rainy season started conditions specified in Section 3 are met. The dashed red line represents the fitted linear trend, estimated using country-specific fixed-effects models. The total predicted change in the mean onset day of the year from 1979–2024, as implied by the model, is: Ethiopia (19.8 days), Malawi (17.8 days), Mali (11.9 days), Niger (7.6 days), Nigeria (25.1 days), and Tanzania (5.8 days).

Figure 3: Onset of the rainy season – asymmetric effects



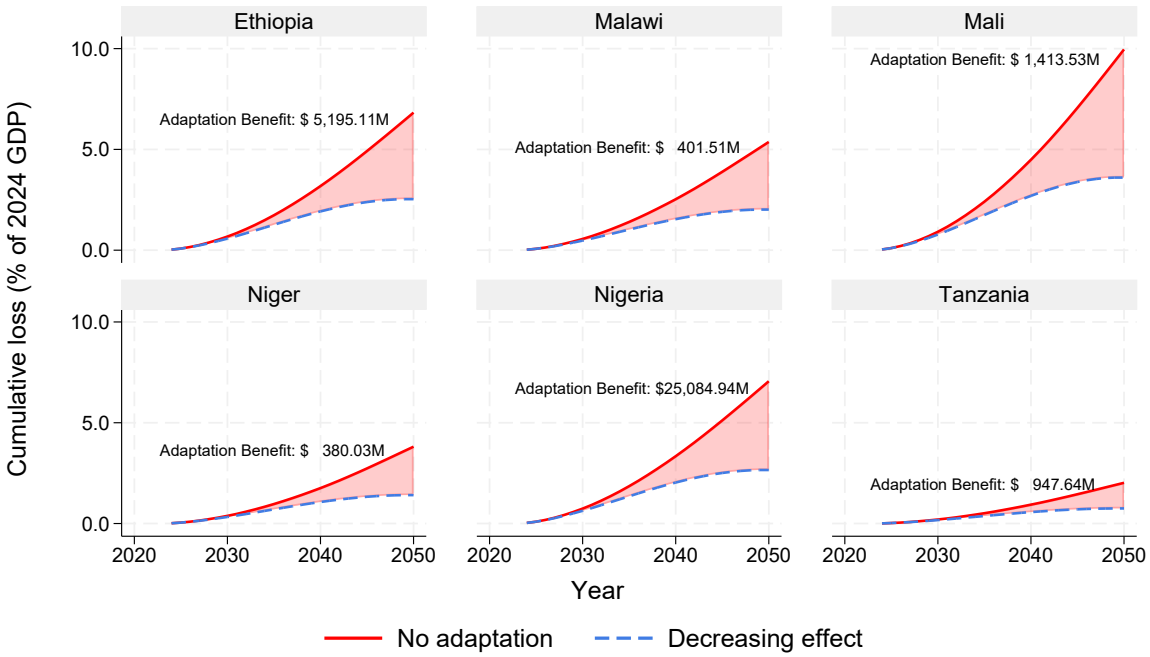
Note. The figure displays OLS regression coefficients for indicator variables representing the bottom tercile and top tercile of the historical rainy season onset distribution within $0.5° \times 0.5°$ grid cells. The middle tercile is the omitted reference category. The dependent variable is log agricultural yields (2020 USD per hectare). The regression includes the benchmark set of fixed effects ($0.1° \times 0.1°$ grid cell and climatic area by country by year), weather controls, and plot-level controls as specified in Equation 1. Bars represent 95% confidence intervals based on standard errors clustered at the $0.5° \times 0.5°$ grid-cell level.

Figure 4: Projected damages by 2050



Note. Cumulative discounted damages from 2025 to 2050, expressed as a percentage of 2024 real GDP. Damages are calculated using the benchmark estimate of a 2% yield loss per week of onset delay. The solid line shows damages under the FAO's Business as Usual economic scenario and the full projected onset delay. The dashed line shows damages under the Toward Sustainability scenario, which assumes a 50% less severe onset delay. The shaded area represents the economic benefit of the sustainable path. All projections account for the mitigating effect of future irrigation expansion and use a time-varying discount rate from the DICE-2023 model. See Appendix C for full details.

Figure 5: Accounting for future adaptation



Note. Cumulative discounted damages from 2025 to 2050 under the Business as Usual scenario with an optimistic counterfactual, expressed as a percentage of 2024 real GDP. The solid line (No adaptation) shows the baseline projection using a constant 2% yield loss per week of onset delay. The dashed line (Decreasing effect) shows a scenario where this yield impact linearly declines to zero between 2025 and 2050. The shaded area represents the total economic benefit of this optimistic, catch-all adaptation scenario, quantified in millions of US dollars.

Table 1: Onset of the rainy season and agricultural productivity

Dependent variable:	**Log yields (2020 USD per hectare)**		
	(1)	(2)	(3)
Onset of the rainy season	-0.022	-0.021	-0.021
	(0.007)	(0.008)	(0.007)
	[0.002]	[0.005]	[0.002]
Standardised effect	-0.036	-0.033	-0.033
Mean (dep. var.)	5.744	5.744	5.744
Identifying observations	130,556	130,556	129,919
Singleton observations	182	182	180
Countries	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022
Weather controls	No	Yes	Yes
Demographic controls	No	No	Yes

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. A full list of controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Table 2: Onset of the rainy season and per capita consumption

Dependent variable:	**Log per capita consumption**					
	Agricultural households			**All households**		
	(1)	(2)	(3)	(4)	(5)	(6)
Onset of the rainy season	-0.012	-0.010	-0.010	-0.008	-0.007	-0.007
	(0.004)	(0.005)	(0.005)	(0.002)	(0.003)	(0.003)
	[0.006]	[0.041]	[0.047]	[0.001]	[0.006]	[0.010]
Standardised effect	-0.020	-0.016	-0.015	-0.016	-0.014	-0.013
Mean (dep. var.)	5.970	5.970	5.970	6.227	6.227	6.225
Identifying observations	29,335	29,335	29,269	70,748	70,748	70,298
Singleton observations	338	338	337	457	457	457
Countries	6	6	6	6	6	6
Survey year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022
Weather controls	No	Yes	Yes	No	Yes	Yes
Demographic controls	No	No	Yes	No	No	Yes

Note. OLS regression estimates of Equation 1. The dependent variable is the log of per capita consumption, measured in 2020 US dollars. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. In Columns (1)–(3), the sample is restricted to households engaged in the agricultural sector and reporting a complete harvest. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Table 3: Onset of the rainy season and women's nutrition

Dependent variable:	**Weight-for-height**	**Wasted**	**Underweight**
	(1)	(2)	(3)
Onset of the rainy season	-0.007	0.001	0.001
	(0.004)	(0.001)	(0.001)
	[0.040]	[0.261]	[0.378]
Standardised effect	-0.016	0.001	0.001
Mean (dep. var.)	-0.850	0.110	0.142
Identifying observations	59,014	59,014	51,733
Singleton observations	16	16	15
Countries	6	6	6
Interview year range	1995–2018	1995–2018	1992–2022

Note. OLS regression estimates of Equation 1. The data are from the DHS (Croft et al., 2018) for women in agricultural households. The dependent variables are defined as follows. *Weight-for-height* is a z-score, representing the standard deviation from the median of the DHS reference population. *Wasted* is an indicator for a weight-for-height z-score below -2. *Underweight* is an indicator for a BMI below 18.5. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include weather controls, individual-level controls, climatic area by country by year fixed effects, and location fixed effects at the $0.5° \times 0.5°$ grid-cell level. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides further details.

Table 4: Onset of the rainy season and crop management

Dependent variable:	Timing			Inputs			Main crop is		
	Planting (DOY)	**Planting duration**	**Harvest (DOY)**	**Seeds**	**Fertiliser**	**Pesticide**	**Cereals**	**Tubers**	**Legumes**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Onset of the rainy season	0.557	-0.008	0.620	0.032	-0.001	-0.000	0.000	-0.001	0.000
	(0.222)	(0.003)	(0.942)	(0.011)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
	[0.012]	[0.016]	[0.510]	[0.006]	[0.664]	[0.771]	[0.817]	[0.410]	[0.910]
Standardised effect	0.842	-0.012	0.936	0.046	-0.001	-0.001	0.001	-0.001	0.000
Mean (dep. var.)	188.934	1.105	250.370	3.453	0.516	0.066	0.572	0.069	0.122
Identifying observations	125,609	125,609	106,970	120,669	135,081	135,306	136,855	136,855	136,855
Singleton observations	74	74	117	197	173	173	175	175	175
Countries	5	5	5	6	6	6	6	6	6
Planting year range	2009–2022	2009–2022	2009–2022	2009–2022	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variables are defined as follows. *Planting (DOY)* is the approximated first planting day of the year; this approximation assumes planting occurred on the 15th of the reported month. *Planting duration* is the number of distinct months planting occurred on the plot. *Harvest (DOY)* is the approximated first harvest day of the year, assuming harvest occurred on the 15th of the reported month. *Seed* is the log of seed quantity used on the plot, measured in kilograms per hectare. *Fertiliser* is an indicator variable taking the value 1 if any fertiliser, whether organic or inorganic, was used, and 0 otherwise. *Pesticide* is an indicator variable that equals 1 if any pesticide was used and 0 otherwise. *Cereals*, *Tubers*, and *Legumes* are indicator variables taking the value 1 if the main crop by production value on the plot belongs to the respective category, and 0 otherwise. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors, which are clustered at the $0.5° \times 0.5°$ grid-cell level, are reported in parentheses. *p*-values are reported in brackets. Appendix A.1 provides further details on variable construction, selected surveys, and weighting procedures.

Table 5: Heterogeneous effects on agricultural productivity

Dependent variable:	**Log yields (2020 USD per hectare)**				
	(1)	(2)	(3)	(4)	(5)
Onset of the rainy season	-0.020	-0.024	-0.015	-0.021	-0.017
	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)
	[0.004]	[0.001]	[0.022]	[0.002]	[0.098]
× Manager is female	-0.006				
	(0.001)				
	[0.000]				
× Manager has formal education		0.006			
		(0.002)			
		[0.001]			
× Below median assets			-0.006		
			(0.002)		
			[0.001]		
× Irrigated plot				0.019	
				(0.007)	
				[0.009]	
× Improved seeds					0.005
					(0.002)
					[0.033]
Mean (dep. var.)	5.743	5.743	5.769	5.748	5.705
Identifying observations	129,805	129,805	118,366	128,373	105,557
Singleton observations	180	180	179	177	215
Countries	6	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022	2011–2022

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Table 6: The role of false onsets

Dependent variable:	**Planting (DOY)**	**Seeds**	**Log yields (2020 USD per hectare)**
	(1)	(2)	(3)
Onset of the rainy season	0.446	0.033	-0.024
	(0.226)	(0.012)	(0.007)
	[0.049]	[0.009]	[0.001]
× False onset	-0.982	0.099	-0.062
	(0.422)	(0.029)	(0.031)
	[0.020]	[0.001]	[0.043]
Mean (dep. var.)	188.916	3.453	5.744
Identifying observations	125,588	120,650	129,899
Singleton observations	95	216	200
Countries	5	6	6
Planting year range	2009–2022	2009–2022	2008–2022

Note. OLS regression estimates based on Equation (1). *Planting (DOY)* is the approximated first planting day of the year; this approximation assumes planting occurred on the 15th of the reported month. *Seeds* is the log of seed quantity used on the plot, measured in kilograms per hectare. *Log yields (2020 USD per hectare)* is the log of yields, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. The analysis examines the interaction between the rainy season onset week and a dummy variable indicating a *False onset*. Benchmark fixed effects are interacted with the false onset dummy. All specifications include a full set of weather and plot-level controls. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides further details.

Table 7: Farmers' awareness

Dependent variable:	**Crops shock**		**Drought shock**	
	(1)	(2)	(3)	(4)
Onset of the rainy season	0.003		0.001	
	(0.002)		(0.003)	
	[0.301]		[0.821]	
Earliest onset tercile		-0.005		0.006
		(0.017)		(0.016)
		[0.769]		[0.709]
Latest onset tercile		0.031		0.034
		(0.012)		(0.013)
		[0.007]		[0.008]
Standardised effect	0.004		0.001	
Mean (dep. var.)	0.411	0.411	0.202	0.202
Identifying observations	136,420	136,420	132,044	132,044
Singleton observations	173	173	180	180
Countries	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variable in each column is a binary indicator for whether a farmer reported a negative shock on a given plot. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. *Earliest onset tercile* and *Latest onset tercile* are indicators for whether the onset of the rainy season for a location falls into the earliest or latest tercile of its historical distribution, with the middle tercile serving as the omitted category. These terciles are defined separately for each $0.5° \times 0.5°$ grid cell based on its long-term (1979–2020) onset history. For comparison, Panel A presents results using the continuous onset week variable. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects, climatic area by country by year fixed effects, and a full set of weather and plot-level controls. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides detailed information on all variables.

Table 8: Heterogeneity of onset impacts by long-run climate trend

Dependent variable:	**Log yields (2020 USD per hectare)**		
	(1)	(2)	(3)
Onset of the rainy season	0.005	0.004	-0.000
	(0.009)	(0.009)	(0.008)
	[0.595]	[0.656]	[0.981]
× Positive and significant trend	-0.033	-0.032	-0.026
	(0.011)	(0.012)	(0.011)
	[0.003]	[0.009]	[0.017]
Mean (dep. var.)	5.744	5.744	5.744
Identifying observations	130,556	130,556	129,919
Singleton observations	182	182	180
Countries	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022
Weather controls	No	Yes	Yes
Demographic controls	No	No	Yes

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. A location is designated as *trending* if its $0.5° \times 0.5°$ grid cell has a statistically significant positive long-run trend in onset. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. A full list of controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

References

Amale, H. S., P. S. Birthal, and D. S. Negi (2023). Delayed Monsoon, Irrigation and Crop Yields. *Agricultural Economics 54*(1), 77–94.

Antwi-Agyei, P., A. J. Dougill, R. Doku, and R. C. Abaidoo (2022). The Role of Soil Moisture Information in Developing Robust Climate Services for Smallholder Farmers: Evidence from Ghana. *Agronomy 12*(2), 541.

Aragón, F. M., F. Oteiza, and J. P. Rud (2021, February). Climate Change and Agriculture: Subsistence Farmers' Response to Extreme Heat. *American Economic Journal: Economic Policy 13*(1), 1–35.

Barreca, A., K. Clay, O. Deschênes, M. Greenstone, and J. S. Shapiro (2016). Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century. *Journal of Political Economy 124*(1), 105–159.

Boogaard, H., J. Schubert, A. De Wit, J. Lazebnik, R. Hutjes, and G. Van der Grijn (2020). Agrometeorological Indicators from 1979 to Present Derived from Reanalysis. Accessed on 01-03-2025.

Burlig, F., A. Jina, E. M. Kelley, G. V. Lane, and H. Sahai (2024, 2). Long-Range Forecasts as Climate Adaptation: Experimental Evidence from Developing-Country Agriculture. Working Paper 32173, National Bureau of Economic Research.

Camacho, A. and E. Conover (2019). The Impact of Receiving SMS Price and Weather Information on Small Scale Farmers in Colombia. *World Development 123*, 104596.

Carleton, T., E. Duflo, B. K. Jack, and G. Zappalà (2024). Chapter 4 - Adaptation to climate change. In L. Barrage and S. Hsiang (Eds.), *Handbook of the Economics of Climate Change*, Volume 1, pp. 143–248. North-Holland.

Corno, L., N. Hildebrandt, and A. Voena (2020). Age of Marriage, Weather Shocks, and the Direction of Marriage Payments. *Econometrica 88*(3), 879–915.

Croft, T. N., A. M. J. Marshall, and C. K. Allen (2018). Guide to DHS Statistics. Demographic and Health Surveys Program.

Danielson, J. J. and D. B. Gesch (2011). Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010). Open-File Report 2011-1073, U.S. Geological Survey.

Dell, M., B. F. Jones, and B. A. Olken (2014). What Do We Learn from the Weather? the New Climate-Economy Literature. *Journal of Economic Literature 52*(3), 740–98.

Dercon, S. and P. Krishnan (2000). In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia. *Journal of Political Economy 108*(4), 688–727.

Deschênes, O., M. Greenstone, and J. Guryan (2009). Climate Change and Birth Weight. *American Economic Review 99*(2), 211–17.

Deschênes, O. and E. Moretti (2009). Extreme Weather Events, Mortality, and Migration. *The Review of Economics and Statistics 91*(4), 659–681.

Dunning, C. M., E. C. L. Black, and R. P. Allan (2016). The Onset and Cessation of Seasonal Rainfall over Africa. *Journal of Geophysical Research: Atmospheres 121*(19), 11405–11424.

FAO (2018). The future of food and agriculture – Alternative pathways to 2050. Technical report, FAO, Rome.

Giné, X., R. M. Townsend, and J. I. Vickery (2018). Forecasting When It Matters: Evidence from Semi-Arid India. *Journal of Development Economics 131*, 150–167.

Hultgren, A., T. Carleton, M. Delgado, D. R. Gergel, M. Greenstone, T. Houser, S. Hsiang, A. Jina, R. E. Kopp, S. B. Malevich, K. E. McCusker, T. Mayer, I. Nath, J. Rising, A. Rode, and J. Yuan (2025, jun). Impacts of Climate Change on Global Agriculture Accounting for Adaptation. *Nature 642*(8068), 644–652.

IPCC (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. In press.

IPCC (2022). Africa. In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, and B. Rama (Eds.), *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 1285–1455. Cambridge, UK and New York, NY, USA: Cambridge University Press.

Jägermeyr, J. and K. Frieler (2018). Spatial Variations in Crop Growing Seasons Pivotal to Reproduce Global Fluctuations in Maize and Wheat Yields. *Science Advances 4*(11), eaat4517.

Jagnani, M., C. B. Barrett, Y. Liu, and L. You (2020, 05). Within-Season Producer Response to Warmer Temperatures: Defensive Investments by Kenyan Farmers. *The Economic Journal 131*(633), 392–419.

Jones, B. F., J. Moscona, B. A. Olken, and C. von Dessauer (2025, September). With or Without U? Binning Bias and the Causal Effects of Temperature Shocks.

Kala, N. (2019). Learning, adaptation, and climate uncertainty: evidence from Indian agriculture. Technical report, MIT Sloan School of Management.

Klein Goldewijk, K., A. Beusen, J. Doelman, and E. Stehfest (2017). Anthropogenic land use estimates for the Holocene – HYDE 3.2. *Earth System Science Data 9*(2), 927–953.

Laajaj, R. and K. Macours (2024, April). The Complexity of Multidimensional Learning in Agriculture.

Leenaars, J. G. B., L. Claessens, G. B. M. Heuvelink, T. Hengl, M. Ruiperez González, L. G. J. van Bussel, N. Guilpart, H. Yang, and K. G. Cassman (2018). Mapping Rootable Depth and Root Zone Plant-Available Water Holding Capacity of the Soil of Sub-Saharan Africa. *Geoderma 324*, 18–36.

Marteau, R., B. Sultan, V. Moron, A. Alhassane, C. Baron, and S. B. Traoré (2011). The Onset of the Rainy Season and Farmers' Sowing Strategy for Pearl Millet Cultivation in Southwest Niger. *Agricultural and Forest Meteorology 151*(10), 1356–1369.

Metzger, M. J., J. Rombouts, L. Reino, M. Campos, A. Lomba, C. A. Guerra, I. Sattler, S. Ceauşu, D. Singing, S. Herrando, L. Brotons, F. Sayol, J. Santana, A. Delgado, M. J. Santos, J. P. Honrado, S. Schindler, D. Kübler, L. Mo, D. Poursanidis, A. Didem, P. Crawford, L. Halada, N. Abrantes, T. Domingos, and H. M. Pereira (2023). Integrative Climate Classification for the Global Land Surface. *Scientific Data 10*(1), 708.

Minoli, S., J. Jägermeyr, S. Asseng, A. Urfels, and C. Müller (2022). Global Crop Yields Can Be Lifted by Timely Adaptation of Growing Periods to Climate Change. *Nature Communications 13*(1), 7079.

Patel, D. (2025, March). Environmental Beliefs and Adaptation to Climate Change. Working Paper.

Rosenzweig, M. R. and C. R. Udry (2019, 5). Assessing the Benefits of Long-Run Weather Forecasting for the Rural Poor: Farmer Investments and Worker Migration in a Dynamic Equilibrium Model. Working Paper 25894, National Bureau of Economic Research.

Rudder, J. and D. Viviano (2024). Learning from Weather Forecasts and Short-Run Adaptation: Evidence from an At-Scale Experiment. Working Paper.

Suri, T., C. Udry, J. C. Aker, C. B. Barrett, L. F. Bergquist, M. Carter, L. Casaburi, R. D. Osei, D. Gollin, V. Hoffmann, T. Jayne, N. Karachiwalla, H. Kazianga, J. Magruder, H. Michelson, M. Startz, and E. Tjernstrom (2024, 3). Agricultural Technology in Africa. *VoxDevLit 5*(2).

Takele, R. and M. Dell'Acqua (2023). *AquaBEHER: Estimation and Prediction of Wet Season Calendar and Soil Water Balance for Agriculture*. R package version 1.4.0.

Tambet, H. and Y. Stopnitzky (2021). Climate Adaptation and Conservation Agriculture among Peruvian Farmers. *American Journal of Agricultural Economics 103*(3), 900–922.

Taraz, V. (2017). Adaptation to climate change: historical evidence from the Indian monsoon. *Environment and Development Economics 22*(5), 517–545.

The World Bank (2025). World Development Indicators. Accessed on June 20, 2025.

Udry, C. (1996). Gender, Agricultural Production, and the Theory of the Household. *Journal of Political Economy 104*(5), 1010–1046.

Wolf, J., K. Ouattara, and I. Supit (2015). Sowing Rules for Estimating Rainfed Yield Potential of Sorghum and Maize in Burkina Faso. *Agricultural and Forest Meteorology 214-215*, 208–218.

Wollburg, P., T. Bentze, Y. Lu, C. Udry, and D. Gollin (2024). Crop yields fail to rise in smallholder farming systems in sub-Saharan Africa. *Proceedings of the National Academy of Sciences 121*(21).

World Bank (2024). A Longitudinal Cross-Country Dataset on Agricultural Productivity and Welfare in Sub-Saharan Africa. Policy Research Working Paper, World Bank Group, Washington, D.C.

Yang, Y., D. Tilman, Z. Jin, P. Smith, C. B. Barrett, Y.-G. Zhu, J. Burney, P. D'Odorico, P. Fantke, J. Fargione, J. C. Finlay, M. C. Rulli, L. Sloat, K. J. van Groenigen, P. C. West, L. Ziska, A. M. Michalak, the Clim-Ag Team, D. B. Lobell, M. Clark, J. Colquhoun, T. Garg, K. A. Garrett, C. Geels, R. R. Hernandez, M. Herrero, W. D. Hutchison, M. Jain, J. M. Jungers, B. Liu, N. D. Mueller, A. Ortiz-Bobea, J. Schewe, J. Song, J. Verheyen, P. Vitousek, Y. Wada, L. Xia, X. Zhang, and M. Zhuang (2024). Climate Change Exacerbates the Environmental Impacts of Agriculture. *Science 385*(6713).

Yegbemey, R. N., G. Bensch, and C. Vance (2023). Weather information and agricultural outcomes: Evidence from a pilot field experiment in Benin. *World Development 167*(C).

Zita, L. E., F. Justino, C. Gurjão, J. Adamu, and M. Talacuece (2025). Spatio-Temporal Characteristics of Climate Extremes in Sub-Saharan Africa and Potential Impact of Oceanic Teleconnections. *Atmosphere 16*(1), 86.

APPENDIX

Supplementary material to *Shifting Seasons: Agricultural Adaptation and Resilience in Africa*

Iván Kim Taveras

A Data and methodological procedures **2**
A.1 Variables and data sources . . . 2
A.2 Descriptive statistics . . . 3
A.3 Weight Construction . . . 6

B Supplementary results **7**
B.1 Trends in rainy season onset, cessation, and length . . . 7
B.2 Issues related to identification . . . 10
B.3 Robustness of impact on productivity . . . 13
B.4 Stability of impacts on productivity over time . . . 19
B.5 Supplementary results on adaptation strategies . . . 20
B.6 Simulation methodology to correct for aggregation bias . . . 24
B.7 The role of adjusting planting dates . . . 28
B.8 Additional results for nutritional outcomes . . . 29
B.9 Stability of planting months . . . 30
B.10 False onsets: Robustness and alternative agronomical pathways . . . 31
B.11 Additional results on heterogeneity . . . 33

C Methodology for Projecting Economic Damages **36**

A Data and methodological procedures

A.1 Variables and data sources

Variable	Description
Agricultural productivity	Two primary measures of agricultural yields are used: output in constant 2020 US dollars per hectare and output in kilograms per hectare. These are calculated at the plot level.
Climatic areas	Household GPS coordinates from the LSMS-ISA surveys are spatially matched to define unique climatic areas based on the Köppen-Geiger climate classification (Metzger et al., 2023). Due to the requirements of fixed effect estimation needing sufficient observations within each group, these detailed classifications are aggregated into three broader zones: Tropical, Temperate, and Arid.
Crop types and main crop	Standardised crop groupings are utilised as defined in the LSMS-ISA harmonised dataset (World Bank, 2024). There are ten categories created to ensure comparability across countries: barley, wheat, rice, sorghum, maize, millet, perennials (e.g., fruit and tree crops), legumes, root crops, and nuts. Crops that do not fall into any of these categories are classified as *others*. The *main crop* on a given plot is identified as the crop within these categories that has the highest production value, following the methodology of World Bank (2024).
False onset	An area is considered exposed to a false onset if, within a 30-day window around its usual onset date (see *Usual rainy season dates*), it experiences a 2-day wet spell with at least 20mm of rainfall, which is then followed by a 7-day dry spell (where a day counts as dry if it receives less than 0.1 mm of rain) within the subsequent 20 days.
Inputs	Plot-level information is collected on the use of various agricultural inputs. *Seeds* refers to the quantity of seeds used, measured in kilograms per hectare for a given plot. *Seed value* is the total value of seeds used per hectare on a plot, expressed in constant 2020 US dollars (see *Unit values*). *Improved seeds* is an indicator variable taking the value 1 if the seeds planted are not traditional varieties (i.e., are improved or high-yielding), and 0 otherwise. *Fertiliser* is an indicator variable that equals 1 if either organic or inorganic fertilisers were used on the plot, and 0 otherwise. Similarly, *Pesticides* is an indicator variable taking the value 1 if pesticides were applied to the plot, and 0 otherwise.
Irrigated	An indicator variable at the plot level that takes the value 1 if any form of irrigation was used on the plot during the agricultural season, and 0 otherwise.
Labour outcomes	Individual-level data from LSMS-ISA are used to construct indicator variables for individuals of working age regarding their employment activities. *Any wage work (Last 12 months)* indicates if an individual reports working for a wage in the past twelve months. *Agriculture (Last 12 months)* indicates if an individual reports wage work specifically in the agricultural sector in the past twelve months. *Services (Last 12 months)* indicates if an individual reports wage work in the service sector in the past twelve months. For a shorter recall period, *Any wage work (Last 7 days)* indicates if an individual reports working for a wage in the past seven days. *Household business (Last 7 days)* indicates if they report working for a household business in the past seven days. Each of these is an indicator variable taking the value 1 if the condition is met, and 0 otherwise.
Nutrition	The DHS records objective measurements performed by its data collection team, utilizing the CDC Standard Deviation-derived Growth Reference Curves for standardized distributions (Croft et al., 2018). The indicators used are as follows: *w/h (weight-for-height)* represents the z-score from the reference curve, with *wasted* being an indicator variable equal to 1 if the *w/h* z-score is less than -2, and 0 otherwise. Similarly, *h/a (height-for-age)* is the z-score from the reference curve, where *stunted* is an indicator variable equal to 1 if the *h/a* z-score is less than -2, and 0 otherwise. *BMI* is defined as the ratio of weight in kilograms to the square of height in meters, excluding pregnant women. Finally, *Underweight* is an indicator variable set to 1 if a woman's BMI is below 18.5.
Per capita consumption	Household consumption is aggregated per capita and expressed in constant 2020 US dollars. This aggregate generally excludes expenditure on rent and durables. It is important to note that Tanzania Wave 5, which includes durables, is excluded from the main analyses as detailed in Section 3. Furthermore, comparable consumption aggregates are not available for Malawi Waves 3 and 4, and Mali Wave 2.
Plot area	The area of each agricultural plot is measured in hectares. This information is primarily sourced from GPS device measurements as provided in the LSMS-ISA harmonised dataset (World Bank, 2024). In instances where GPS measures are unavailable, imputation methods based on self-reports and administrative data are employed, as detailed by World Bank (2024).
Proxies for access to information	Individual indicators are created for specific assets: *Owns mobile phone* (equal to 1 if the household owns any mobile telephone, including basic, feature, regular, or smart phones where distinguished), *Owns Television* (equal to 1 for television set ownership), and *Owns Radio* (equal to 1 for radio ownership, including radio cassette or radio with flash drive). Additionally, for Nigeria only, an indicator *Internet Access* is included, sourced from a different survey module. Note that specific item labels used in the construction vary slightly across countries and waves. The construction aims for maximum consistency across surveys.

(continued on next page)

Variable	Description
Rainy season calendar	Variables summarising the rainy season calendar are computed for each $0.1° \times 0.1°$ grid cell and year using the AquaBEHER R package (Takele and Dell'Acqua, 2023). This package employs an agronomic definition based on soil water balance. The package requires the user to specify an earliest possible onset date, a latest possible onset date, and a latest possible cessation date for each location. These are determined flexibly: the earliest onset is set 90 days before the location's usual onset date, the latest onset is 90 days after the usual onset date, and the latest cessation is 90 days after the usual cessation date (see *Usual rainy season dates*). The *Onset* is then defined as the first day when the actual-to-potential evapotranspiration ratio is greater than 0.5 for seven consecutive days, followed by a 20-day period in which plant-available water remains above the wilting point over the root zone of the soil layer. The *Cessation* is the first day when the actual-to-potential evapotranspiration ratio falls below 0.5 for seven consecutive days, followed by 12 consecutive non-growing days during which plant-available water remains below the wilting point over the root zone. The *Length* of the rainy season is the total number of days from the onset to the cessation of the season.
Unit values	To enable cross-country analysis, key agricultural input and output variables are valued in constant 2020 US dollars, following the methodology of the LSMS-ISA harmonised dataset (World Bank, 2024). This involves calculating median prices for specific categories of inputs and outputs at the enumeration area (EA) level. If fewer than 10 price observations exist at the EA level, medians are calculated using progressively larger administrative units up to the national level. These derived prices are then multiplied by standardised quantities. The resulting local currency unit values are converted to 2020 US dollars using annual exchange rates and CPI data from the World Bank Open Data Initiative.
Usual rainy season dates	The usual onset and cessation dates of the rainy season for each $0.1° \times 0.1°$ grid cell are determined following the methodology of Dunning et al. (2016). This method first computes the climatological mean rainfall for each day of the calendar year and the overall climatological daily mean rainfall for the location. From these, the climatological cumulative daily rainfall anomaly is calculated for each day of the year. The day of the minimum in this cumulative anomaly series marks the beginning of the climatological water season (usual onset), and the day of the maximum in the series marks its end (usual cessation). These usual dates provide the baseline for defining the search window in the AquaBEHER package (see *Rainy season calendar*).
Weather controls	Daily climate variables are sourced from the ERA5 reanalysis dataset ($0.1° \times 0.1°$ resolution), provided by the Copernicus Climate Change Service (C3S) Climate Data Store (Boogaard et al., 2020). For use as controls in regression analyses, these daily data are aggregated. Unless otherwise specified, variables are aggregated over the calendar year. As a robustness check, alternative aggregation periods such as quarterly and monthly aggregations are used. *Temperature* is the average of the daily minimum and maximum temperatures, expressed in $°C$. *Total precipitation* is the sum of daily precipitation over the aggregation period, measured in mm. *Relative humidity* is a measure of the amount of water vapour in the air compared to the maximum amount of water vapour the air can hold at that temperature, expressed as a percentage. *Harmful degree days* quantifies the number of days where the maximum temperature exceeds the 90th percentile of a location's historical temperature distribution. Additionally, the number of days with precipitation above the 50th, 75th, and 90th percentiles of a location's historical distribution, and below the 5th, 15th, and 25th percentiles, are computed to capture variations in rainfall intensity and dry spells. For these percentile-based measures, the historical distributions are calculated at the $0.5° \times 0.5°$ grid-cell level to ensure robust local benchmarks.

Note. For time-varying variables, missing values are linearly interpolated.

Table A2: Sampled surveys

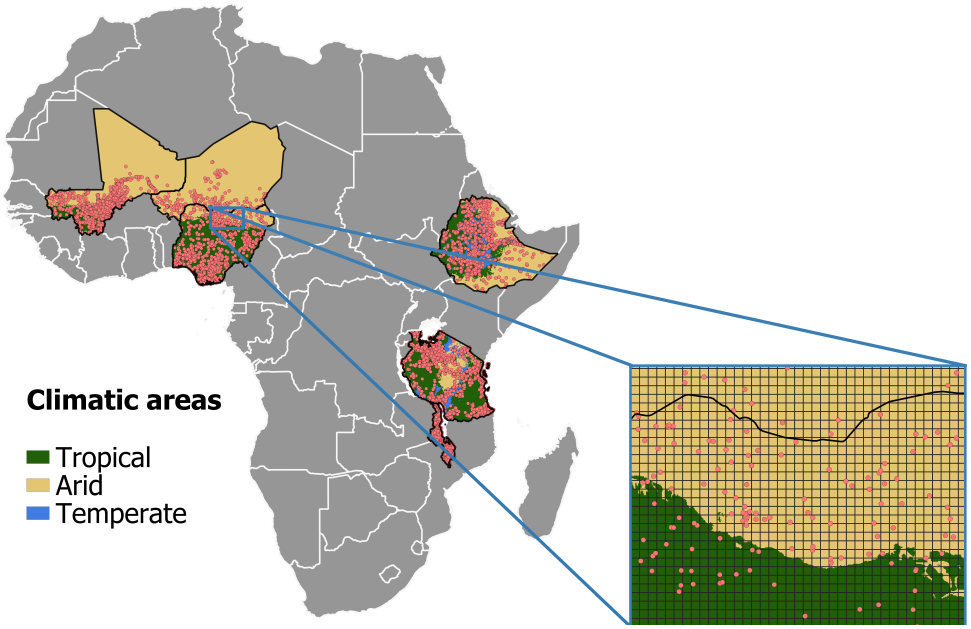
	LSMS-ISA		DHS
Country	**Survey years**	**Waves**	**Survey years**
Ethiopia	2011, 2013, 2015, 2018, 2022	1–5	2000, 2005, 2011, 2016, 2019
Malawi	2010, 2013, 2016, 2019	1–4	2000, 2004, 2010, 2015
Mali	2014, 2017	1–2	1995, 2001, 2006, 2012, 2018
Niger	2011, 2014	1–2	1992, 1998, 2012
Nigeria	2010, 2012, 2015, 2018	1–4	1990, 2003, 2008, 2013, 2018
Tanzania	2008, 2010, 2012	1–3	1999, 2010, 2015, 2022

Note. Year reported corresponds to the start year for surveys spanning multiple calendar years.

A.2 Descriptive statistics

This subsection presents descriptive statistics for LSMS coordinates on sampled countries along with the geographic coverage of these coordinates within country and climatic area.

Figure A1: Area covered by the study



Note. The map displays the locations of sampled LSMS communities (pink dots) across the three primary climatic zones within the six study countries—Tropical, Arid, and Temperate—based on Metzger et al. (2023). These large colored regions correspond to the climatic area-by-country groups that define the spatially-specific time effects (δ_{cat}). The inset, showing the border between Nigeria and Niger, illustrates the resolution of the $0.1° \times 0.1°$ grid cells where location fixed effects (μ_ℓ) are applied.

Table A3: Geographical distribution of sampled $0.1° \times 0.1°$ grid cells within climatic area-by-country groups

Country	**Climatic Area**	**No. of unique sampled grid cells** ($0.1° \times 0.1°$)
Ethiopia	Arid	372
	Temperate	903
	Tropical	890
Malawi	Arid	1
	Temperate	339
	Tropical	171
Mali	Arid	876
	Tropical	654
Niger	Arid	246
Nigeria	Arid	467
	Tropical	2100
Tanzania	Arid	30
	Temperate	343
	Tropical	1461
Total		**8,853**

Note: Number of unique $0.1° \times 0.1°$ grid cells containing LSMS communities that are nested within each of the broader climatic area-by-country groups. These counts correspond to the number of location fixed effects (μ_ℓ) being absorbed by each spatially-specific time fixed effect (δ_{cat}) in the benchmark regression model.

Table A4: Descriptive statistics

	Mean (1)	Std. dev. (2)	Min (3)	Max (4)	Observations (5)
A. Plot					
Yield (2020 USD/ha)	1,574.65	3,263.98	0.00	19,227.22	135,726
Yield (kg/ha)	4,706.72	9,465.56	0.00	53,049.94	136,348
Crop failure (zero yield)	0.03	0.18	0.00	1.00	136,348
Plot suffered crop shock	0.35	0.48	0.00	1.00	137,507
False onset	0.05	0.22	0.00	1.00	137,948
Onset of the rainy season (week of year)	24.74	12.46	1.00	50.00	137,948
Planting month	5.90	2.27	1.00	12.00	126,440
Number of planting months	1.22	0.72	1.00	12.00	126,440
Seed quantity (kg/ha)	501.62	1,612.81	0.00	8,532.42	123,314
Seed value (2020 USD/ha)	235.61	688.89	0.00	3,908.35	123,514
Used fertiliser	0.51	0.50	0.00	1.00	136,099
Used pesticides	0.10	0.30	0.00	1.00	136,363
Main crop is a cereal	0.54	0.50	0.00	1.00	137,948
Main crop is a tuber	0.15	0.35	0.00	1.00	137,948
Main crop is a legume	0.11	0.32	0.00	1.00	137,948
Number of seasonal crops	1.66	0.97	0.00	15.00	137,820
Female plot manager	0.20	0.40	0.00	1.00	137,828
Plot manager age	47.05	15.01	0.00	100.00	137,646
Plot manager has formal education	0.50	0.50	0.00	1.00	137,605
Plot manager has wage work (past 12m)	0.10	0.30	0.00	1.00	135,852
Plot manager has wage work (past 7d)	0.10	0.30	0.00	1.00	135,226
B. Households					
Per capita consumption (2020 USD)	544.25	428.44	88.12	3,669.61	29,683
Household size	5.91	3.49	1.00	84.00	36,024
HH head age	47.63	15.11	8.00	100.00	35,985
Female HH head	0.17	0.37	0.00	1.00	35,999
HH head has any wage work (past 12m)	0.11	0.31	0.00	1.00	35,962
HH head has non-farm wage work (past 12m)	0.08	0.27	0.00	1.00	35,962
Has primary-age children	0.76	0.43	0.00	1.00	31,756
Farm size (ha)	1.80	3.09	0.02	29.78	35,486
Number of fallow plots	0.14	0.49	0.00	12.00	35,709
Number of cultivated plots	4.73	5.28	0.00	51.00	36,091
Owns non-farm enterprise	0.34	0.47	0.00	1.00	36,064
Household asset index	0.06	0.92	-1.22	10.00	36,065
Urban household	0.10	0.30	0.00	1.00	36,091
Onset of the rainy season (week of year)	24.17	12.19	1.00	50.00	36,081
C. Individuals					
Has any wage work (past 12m)	0.03	0.18	0.00	1.00	208,521
Has any wage work (past 7d)	0.03	0.18	0.00	1.00	205,742
Female	0.50	0.50	0.00	1.00	211,194
Age	22.33	18.91	0.00	100.00	210,547
Married	0.35	0.48	0.00	9.00	204,155
Has formal education	0.54	0.50	0.00	1.00	203,053
Onset of the rainy season (week of year)	24.75	11.27	1.00	50.00	213,312
D. Adult women (DHS)					
Has any wage work (past 12m)	0.88	0.32	0.00	1.00	92,737
Age	18.96	3.40	10.00	46.00	79,901
Married	0.82	0.39	0.00	1.00	92,737
Years of education	3.46	3.77	0.00	20.00	92,730
HH head age	42.43	13.43	13.00	97.00	92,642
HH head is female	0.18	0.39	0.00	1.00	92,737
Household size	4.46	0.92	1.00	5.00	92,737
BMI	21.88	3.54	12.02	59.72	61,558
Underweight (BMI < 18.5)	0.14	0.34	0.00	1.00	51,796
Wasted	0.10	0.31	0.00	1.00	59,073
Weight-for-height z-score	-0.83	1.00	-3.99	5.96	59,073
Stunted	0.19	0.39	0.00	1.00	59,712
Height-for-age z-score	-1.09	1.06	-5.99	5.70	59,712
Onset of the rainy season (week of year)	33.28	14.49	1.00	50.00	92,737
E. ERA5 cells					
Onset of the rainy season (week of year)	26.04	12.75	1.00	50.00	255,072

Continued on next page

Table A4 – continued from previous page

	Mean (1)	Std. dev. (2)	Min (3)	Max (4)	Observations (5)
Cessation of the rainy season (week of year)	34.37	11.31	1.00	52.00	170,676
Length of the rainy season	121.18	55.74	25.00	337.00	170,676
False onset	0.02	0.14	0.00	1.00	255,507
Days with precipitation below 25th percentile	31.13	45.03	0.00	269.00	255,438
Days with precipitation above 75th percentile	93.56	22.82	0.00	226.00	255,438
Maximum relative humidity	74.17	15.52	32.93	99.93	255,438
Average daily temperature (Celsius)	24.36	4.24	8.48	31.30	255,438
Total precipitation (mm)	1,261.35	826.63	75.92	11,772.02	255,438
Harmful degree days	34.45	50.61	0.00	1,275.74	255,438

Note. To account for outliers, the following continuous variables are winsorised at the 1st and 99th percentiles: Yield (2020 USD/ha), Yield (kg/ha), Seed quantity (kg/ha), Seed value (2020 USD/ha), Per capita consumption (2020 USD), and Farm size (ha). Appendix A.1 provides detailed information on all variables, selected surveys, and weighting procedures.

A.3 Weight Construction

I start from the LSMS household sampling weight w_i. To pool survey waves within a country while preserving representativeness, I apply a country–wave factor $f_{ct} = W_{ct}/W_c$, where $W_{ct} = \sum_{i \in (c,t)} w_i$ is the sum of LSMS weights in country c, wave t, and $W_c = \sum_{t'} W_{ct'}$ sums over all waves used for that country. The final household weight is $w_i^* = w_i f_{ct}$. For plot-level analyses, if household i reports P_i plots, each plot p receives $w_{ip}^* = w_i^*/P_i$, ensuring households contribute proportionally regardless of plot count. I use w_i^* or w_{ip}^* for all descriptive statistics and regressions, according to the unit of analysis.

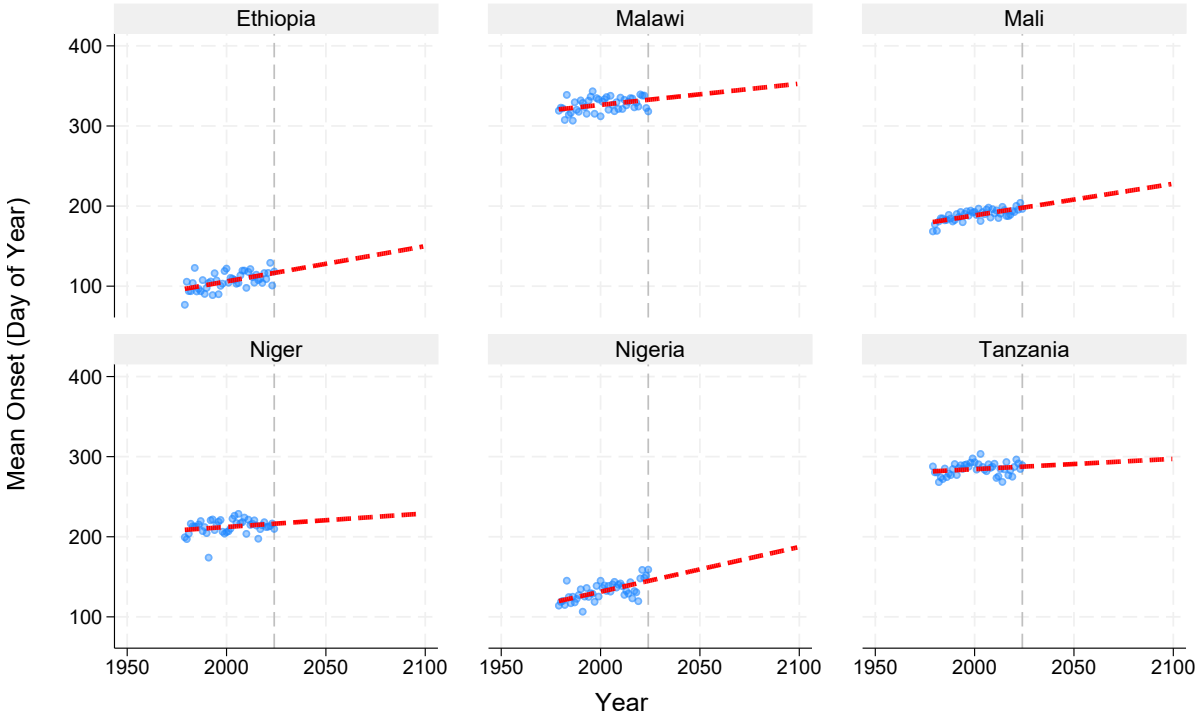
B Supplementary results

B.1 Trends in rainy season onset, cessation, and length

This appendix provides supplementary figures regarding the long-term trends in rainy season onset, cessation, and length, complementing the analysis presented in Section 3.3. The trend estimation methodology, including the fixed-effects model and data construction, is detailed in the main text. All trends cover the period 1979–2024 for $0.1° \times 0.1°$ grid cells matched to LSMS-ISA survey locations, with standard errors clustered at the $0.5° \times 0.5°$ level.

Forecast to 2100. Figure B1 plots the forecast to 2100. By 2100, the linear trend model predicts an average delay of at least ten days in all countries, with extreme cases like Ethiopia and Nigeria of more than a month.

Figure B1: Long-range forecast of the onset of the rainy season to 2100



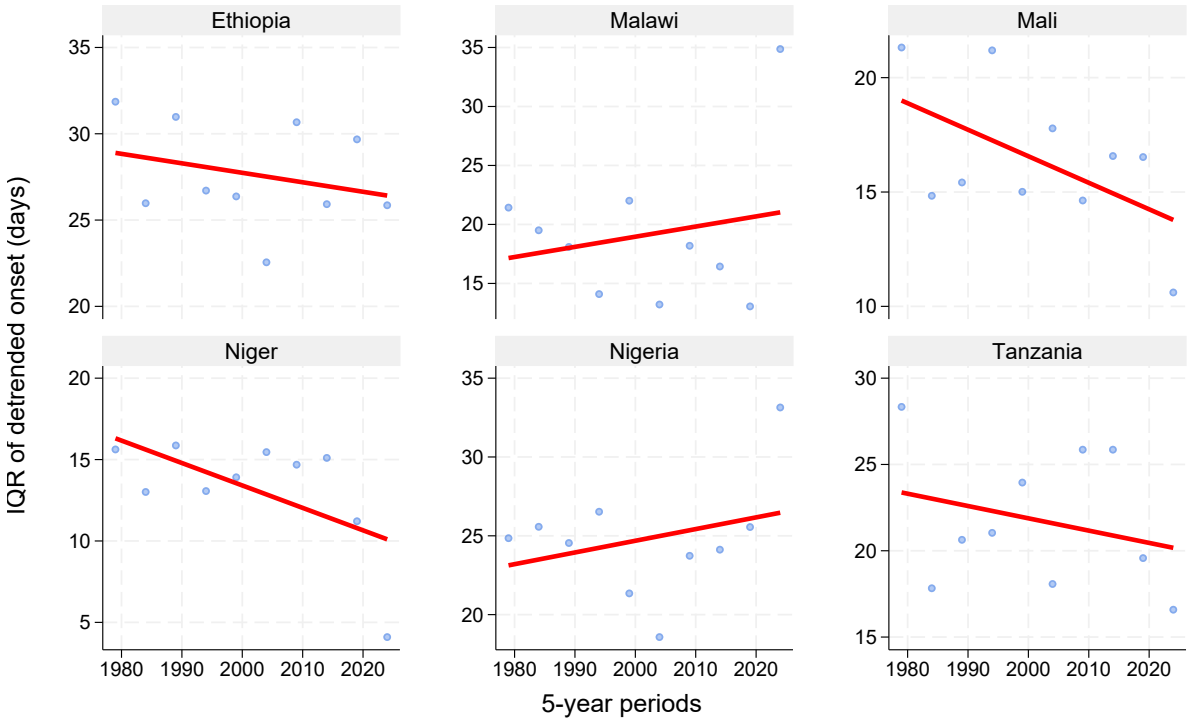
	Forecasted change in mean onset	
Country	By 2050	By 2100
Ethiopia	11.5	33.5
Malawi	6.9	20.1
Mali	10.3	30.0
Niger	4.4	12.8
Nigeria	14.5	42.4
Tanzania	3.4	9.8

Note. Comparison of the long-range forecast of rainy season onset to 2100 using a linear versus. Each panel shows the annual mean onset (in DOY) for a country, with blue dots representing the historical data (1979–2024). The dashed red line show the out-of-sample forecasts from the linear model. The model is estimated using only historical data, and the trends are then extrapolated. The vertical dashed line marks the year 2024, the beginning of the forecast period. The accompanying table summarizes the total change in onset days predicted by the linear model for 2050 and 2100.

Trends in onset variability. Beyond shifts in the average onset timing, I also examined

changes in its year-to-year variability. To do this, I first de-trended the onset data for each country by calculating the residuals from the country-specific fixed-effects models. Then, for each country, I computed the interquartile range (IQR) of these de-trended onset residuals within 5-year rolling windows. Figure B2 plots the evolution of this IQR over time for each country.

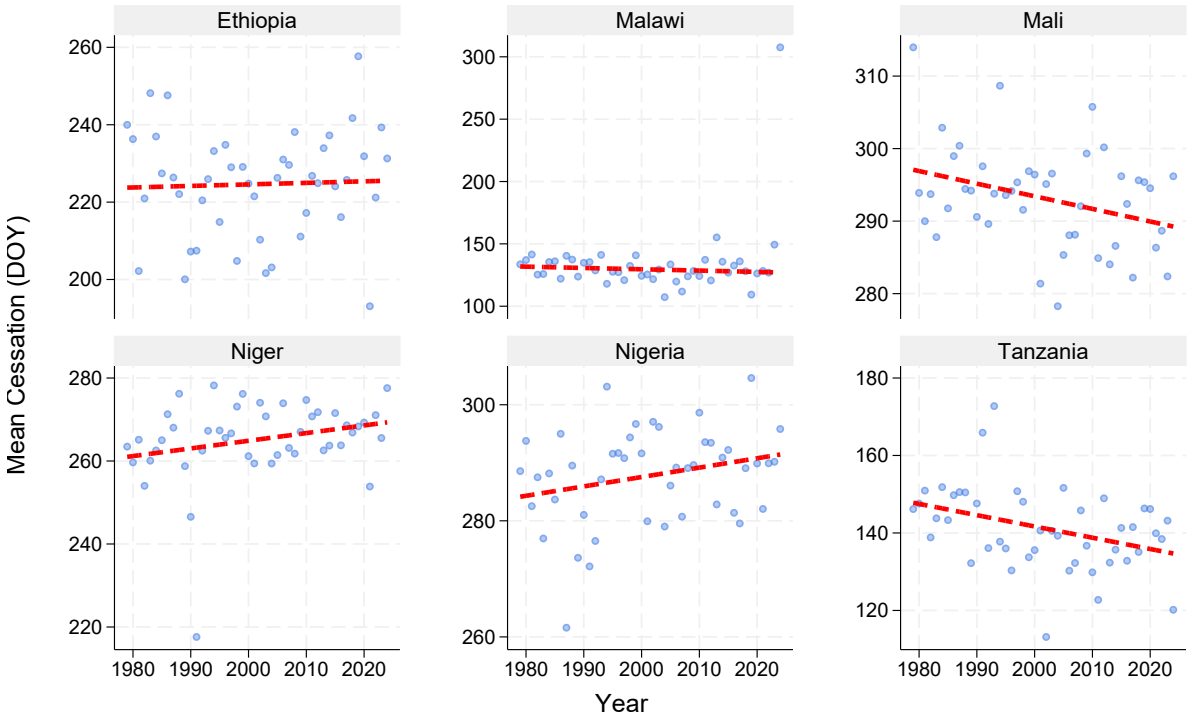
Figure B2: Country-specific trends in rainy season onset variability



Note: Each panel shows the 5-year rolling interquartile range (IQR) of de-trended onset Day of Year (DOY) residuals for LSMS-ISA matched grid cells within the respective country. De-trending is based on country-specific fixed-effects models. The red line represents the linear trend fitted to these 5-year IQR values. This figure illustrates changes in the year-to-year predictability of onset timing. A negative trend (e.g., Mali, Niger) indicates a decrease in variability, meaning onset timing has become more predictable. A positive trend (e.g., Malawi, Nigeria) indicates an increase in variability, suggesting onset timing has become less predictable.

Cessation trends. As noted in Section 3.3, trends in the cessation of the rainy season are more varied than onset trends. Figure B3 presents the country-specific trend analyses for the cessation DOY, highlighting this heterogeneity. For instance, while Niger and Nigeria show a tendency towards later cessation, Malawi, Mali, and Tanzania exhibit trends towards earlier cessation. A pooled analysis for cessation (not shown) consequently reveals a relatively flat and statistically insignificant overall trend.

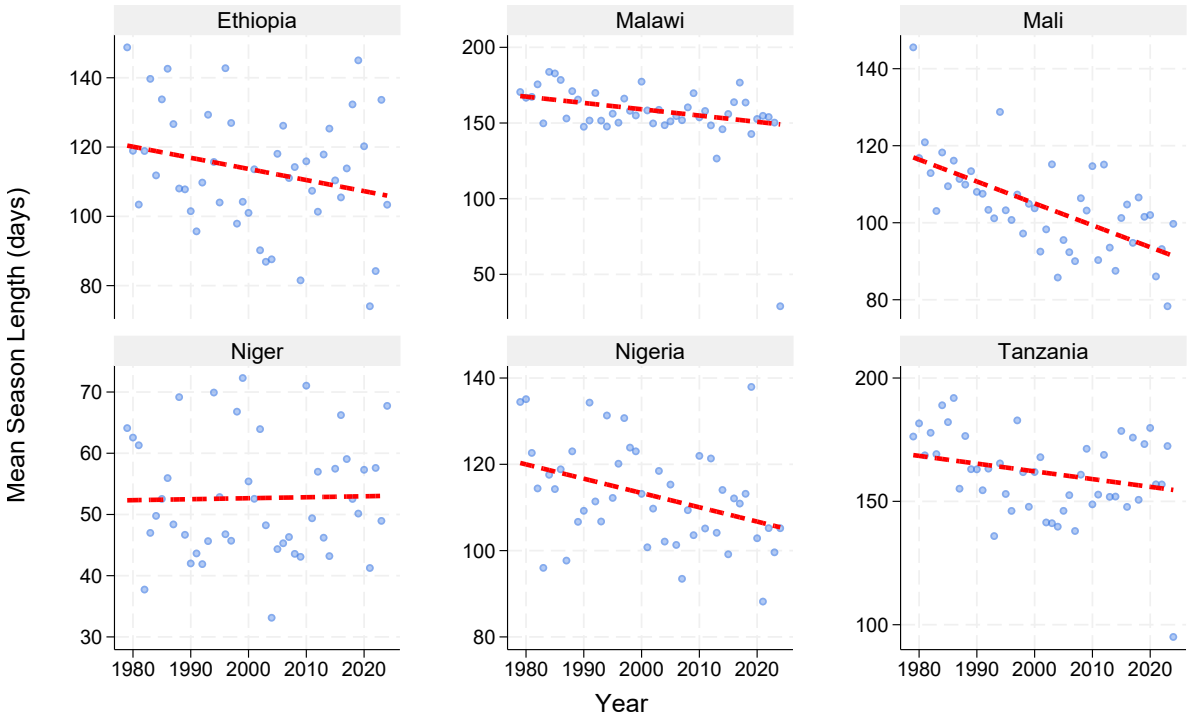
Figure B3: Country-specific trends in rainy season cessation day of year



Note. Each panel displays the annual mean cessation day of the year (DOY) for 0.1°× 0.1° grid cells matched to LSMS-ISA survey locations within the respective country from 1979–2024. The *Cessation of the rainy season* is the day of the year when conditions specified in Section 3 are met. The dashed red line represents the fitted linear trend, estimated using country-specific fixed-effects models. The total predicted change in the mean cessation day of the year from 1979–2024, as implied by the model, is: Ethiopia (1.8 days), Malawi (-7.8 days), Mali (-4.7 days), Niger (8.3 days), Nigeria (7.3 days), and Tanzania (-13.1 days).

Season length trends. The interplay of later onsets and diverse cessation patterns generally leads to a decrease in the length of the rainy season across the study area. The consistent and significant delay in the season's onset is the primary driver of this reduction. Figure B4 illustrates these country-specific trends in the length of the rainy season.

Figure B4: Country-specific trends in rainy season length



Note. Each panel displays the annual mean of the length of the rainy season for 0.1°× 0.1° grid cells matched to LSMS-ISA survey locations within the respective country from 1979–2024. The dashed red line represents the fitted linear trend, estimated using country-specific fixed-effects models. The total predicted change in the mean length of the rainy season in days over the full sample period, as implied by the model, is: Ethiopia (-14.4 days), Malawi (-25.6 days), Mali (-18.7 days), Niger (0.7 days), Nigeria (-14.9 days), and Tanzania (-14.1 days).

B.2 Issues related to identification

Selection bias. A primary concern for identification is the potential for selection bias stemming from the sample restrictions used in the main analyses. The plot-level productivity analysis is restricted to plots that report a complete harvest, while the household-level consumption analysis uses an even stricter criterion, including only households for which *all* reported plots have a complete harvest. This could bias the estimates if the timing of the rainy season onset systematically affects the probability of an observation being included in these samples. For instance, a delayed onset could push the harvest date past the survey's fieldwork period, leading to an underrepresentation of late-onset plots and households.

To formally test for this at both levels, I estimate linear probability models where the dependent variable is an indicator for meeting the respective sample selection criterion. As shown in Table B1, the results provide confidence that selection bias is not a significant concern. At the plot level, the coefficient on the rainy season onset week is small and not statistically significant, indicating that a late onset does not predict whether an individual plot reports a complete harvest. This result holds at the household level as well; onset timing does not significantly predict the probability that a household reports a complete harvest for all of its plots.

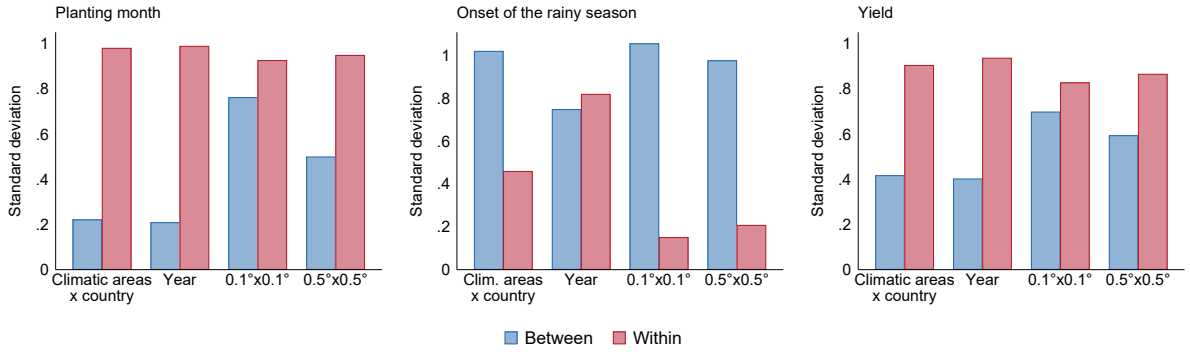
Table B1: Onset of the rainy season and selection into analysis sample

Dependent variable:	**Complete harvest (plot)**	**Complete harvest (household)**
	(1)	(2)
Onset of the rainy season	-0.002	-0.002
	(0.001)	(0.003)
	[0.151]	[0.543]
Standardised effect	-0.003	-0.003
Mean (dep. var.)	0.904	0.732
Identifying observations	148,854	43,330
Singleton observations	177	370
Countries	6	6
Survey year range	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. The main independent variable, *Onset of the rainy season*, is the calendar week when the agronomic conditions specified in Section 3 are met. All specifications include high-resolution ($0.1°\times 0.1°$) grid-cell fixed effects, a full set of climatic area-by-country-by-year fixed effects, and a comprehensive set of plot-level and weather controls detailed in Section 3. Standard errors, clustered at the $0.5°\times 0.5°$ grid-cell level, are in parentheses; *p*-values are in brackets. Appendix A.1 provides further details on all variable definitions, survey data, and weighting procedures.

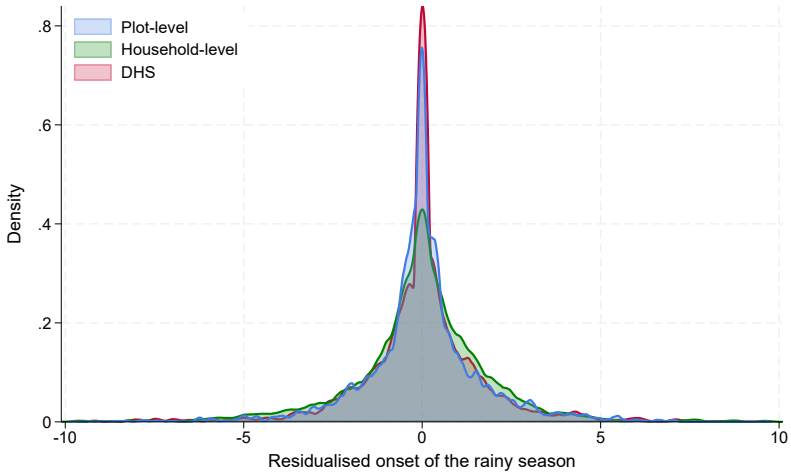
Variation and identification. A common concern with demanding fixed-effects models is whether sufficient variation remains in the data to precisely identify the parameters of interest. This section presents two figures to address this concern. Figure B5 presents a variance decomposition for the main outcome and explanatory variables to show the primary sources of variation in the data. Figure B6 then plots the kernel density of the residualised onset of the rainy season. This residual represents the plausibly exogenous variation used for identification, and the figure is used to visually inspect its distribution for any signs of systematic bias.

Figure B5: Between and within variation decomposition of onset and yields



Note. Between and within standard deviations for three standardised variables: planting month, onset of the rainy season, and log yield. The decomposition is shown for three different panel data structures: climatic area by country, year, and $0.1° \times 0.1°$ grid cell. The sample used for this decomposition is the plot-level LSMS-ISA sample as described in Section 3.

Figure B6: Residualised onset of the rainy season



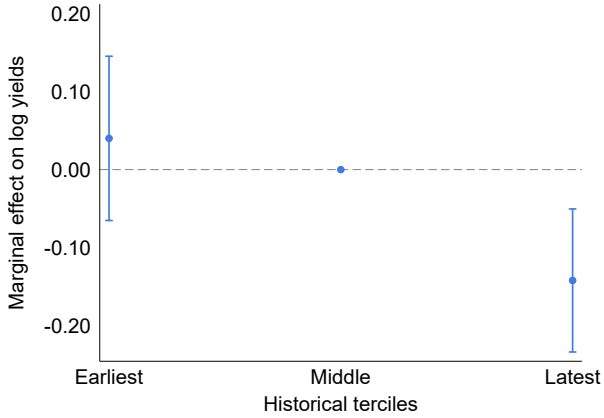
Note. The figure plots the kernel density of the residualised onset of the rainy season for the three main samples used in the analysis: the plot-level agricultural sample, the household-level consumption sample, and the women's nutrition sample (DHS). For each sample, the residuals are obtained by regressing the onset week on the full set of fixed effects used in its corresponding benchmark specification. For visual clarity, the distributions are trimmed at -10 and 10. The onset of the rainy season is defined as the week of the year when the agronomic conditions specified in Section 3 are met.

Robustness to place-specific trends. Recent literature highlights that panel analyses of climate impacts can produce biased estimates if unobserved, place-specific trends correlate with weather variables, a concern particularly relevant when using binned regressors (Jones et al., 2025). While my main analysis relies on a continuous measure of rainy season onset, this issue is pertinent to the tercile indicators I use in two secondary analyses: testing for asymmetric effects on yields (Figure 3) and examining farmer awareness of shocks (Table 7).

To address this potential issue, I augment my benchmark specification by controlling for grid-cell-specific linear trends, which Jones et al. (2025) identify as the most suitable strategy for a short panel. Given that I bin the onset of the rainy season at the $0.5° \times 0.5°$ level, I implement these trends by interacting fixed effects for each grid cell with a linear year trend. Although a linear trend assumption warrants caution, yields have largely stagnated in my study countries (Wollburg et al., 2024), making it a reasonable first-order approximation for smooth, unobserved spatial changes. Jones et al. (2025) provide further support, finding that crop yield estimates are less susceptible to this form of bias.

As I show in Figure B7 and Table B2, my findings are robust to this more demanding specification. The inclusion of these trends does not alter the key results: the negative impact on agricultural yields remains driven entirely by late onsets, and farmers' awareness of a shock is still only triggered by exceptionally late onsets.

Figure B7: Asymmetric effects — Correcting for binning bias



Note. The figure displays OLS regression coefficients for indicator variables representing the bottom tercile and top tercile of the historical rainy season onset distribution within $0.5° \times 0.5°$ grid cells. The middle tercile is the omitted reference category. The dependent variable is log agricultural yields (2020 USD per hectare). The regression augments equation (1) by including interactions between the $0.5° \times 0.5°$ grid-cell fixed effects and a linear time trend. Bars represent 95% confidence intervals based on standard errors clustered at the $0.5° \times 0.5°$ grid-cell level.

Table B2: Farmers' awareness — Correcting for binning bias

Dependent variable:	**Crops shock**	**Drought shock**
	(1)	(2)
Earliest onset tercile	-0.011	0.019
	(0.018)	(0.018)
	[0.530]	[0.284]
Latest onset tercile	0.029	0.039
	(0.014)	(0.014)
	[0.044]	[0.005]
Standardised effect		
Mean (dep. var.)	0.411	0.202
Identifying observations	136,420	132,044
Singleton observations	173	180
Countries	6	6
Planting year range	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variable in each column is a binary indicator for whether a farmer reported a negative shock on a given plot. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. *Earliest onset tercile* and *Latest onset tercile* are indicators for whether the onset of the rainy season for a location falls into the earliest or latest tercile of its historical distribution, with the middle tercile serving as the omitted category. These terciles are defined separately for each $0.5° \times 0.5°$ grid cell based on its long-term (1979–2020) onset history. For comparison, Panel A presents results using the continuous onset week variable. All specifications augment equation (1) by including interactions between the $0.5° \times 0.5°$ grid-cell fixed effects and a linear time trend. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides detailed information on all variables.

B.3 Robustness of impact on productivity

This appendix presents a series of robustness checks for the main findings on the impact of rainy season onset on agricultural productivity, as discussed in Section 5.1 and presented in Table 1.

Alternative local trends. The stability of the estimated onset coefficient was tested against alternative specifications that incorporate different ways of controlling for local time trends. Introducing local time trends effectively changes the definition of the onset shock, as it alters the variation from which the effect is identified by removing different local temporal patterns.

Table B3 shows that the negative impact of a delayed onset on agricultural productivity remains consistent even when employing these more demanding specifications for local trends, which vary the geographic level at which such trends are defined.

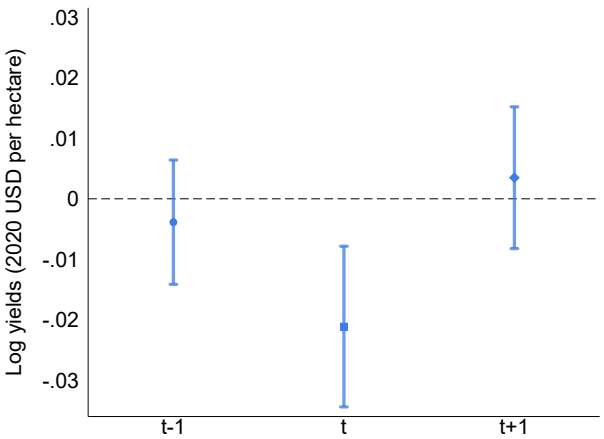
Table B3: Alternative local trends

Dependent variable:	**Log yields (2020 USD per hectare)**			
	(1)	(2)	(3)	(4)
Onset of the rainy season	-0.017	-0.019	-0.016	-0.020
	(0.007)	(0.007)	(0.007)	(0.008)
	[0.012]	[0.004]	[0.034]	[0.012]
Standardised effect	-0.027	-0.030	-0.022	-0.022
Mean (dep. var.)	5.744	5.744	5.744	5.744
Identifying observations	129,919	129,919	129,919	129,913
Singleton observations	180	180	180	186
Countries	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022
Level of local trend	Country	Climatic area	Admin level 1	Admin level 2

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects. The columns differ by the inclusion of location by planting year trends. A full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures. The standardised effect shows the impact of a one standard deviation change in the onset week.

Past and future onsets. To assess whether the estimated impact is genuinely contemporaneous, I introduced past (one-year lag) and future (one-year lead, a placebo) season onset dates into the benchmark specification (equation 1). As shown in Figure B8, neither past nor future onset dates significantly affect current season productivity.

Figure B8: Timing



Note. Marginal effects of rainy season onset on log yields measured in 2020 USD per hectare. Estimates are based on equation 1. On the x-axis, *t* refers to the onset week of the actual planting year, *t-1* refers to the onset week from the previous year, and *t+1* refers to the onset week from the subsequent year, serving as a placebo test. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

The roles of rainy season cessation and length. To better understand the distinct role of rainy season onset relative to other timing features, I examined specifications that also incorporate

cessation timing and total season length. Importantly, the agronomic conditions defining cessation are not met in all location-years, leading to missing values for cessation and, consequently, length for some observations. Therefore, in Table B4, I restrict the sample to observations with non-missing cessation and length data.

Table B4: Cessation and length of the rainy season

Dependent variable:	**Log yields (2020 USD per hectare)**				
	(1)	(2)	(3)	(4)	(5)
Cessation of the rainy season	-0.004	-0.003			
	(0.004)	(0.004)			
	[0.254]	[0.450]			
Onset of the rainy season		-0.016		-0.018	-0.017
		(0.008)		(0.008)	(0.008)
		[0.045]		[0.025]	[0.029]
Length in weeks			-0.003	-0.003	
			(0.004)	(0.004)	
			[0.514]	[0.426]	
Mean (dep. var.)	5.616	5.616	5.616	5.616	5.616
Identifying observations	106,511	106,511	106,511	106,511	106,511
Singleton observations	97	97	97	97	0
Countries	6	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare.*Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met; *Cessation of the rainy season* is the week of the year when conditions specified in Section 3 are met; and *Length* is the total number of weeks the rainy season lasted, calculated from the onset and cessation dates. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects, climatic area by country by year fixed effects, weather controls, and plot-level controls (a full list of these controls is presented in Section 3). The sample for these regressions is restricted to observations with non-missing data for cessation and length. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are reported in parentheses; p-values are reported in brackets. Appendix A.1 provides further details on variable construction.

Alternative specifications for weather controls. Table B5 presents estimates of equation 1 using weather variables aggregated at different frequencies.

Table B5: Alternative weather controls

Dependent variable:	**Log yields (2020 USD per hectare)**		
	(1)	(2)	(3)
Onset of the rainy season	-0.021	-0.023	-0.019
	(0.007)	(0.007)	(0.007)
	[0.002]	[0.002]	[0.009]
Standardised effect	-0.033	-0.033	-0.027
Mean (dep. var.)	5.744	5.744	5.744
Identifying observations	129,919	129,919	129,919
Singleton observations	180	180	180
Countries	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022
Time aggregation	Yearly	Quarterly	Monthly

Note: OLS regression estimates of Equation 1. The dependent variables is the log of yield (measured in 2020 US dollars per hectare). *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects, climatic area by country by year fixed effects, and plot-level controls (for yield regressions) or household-level controls (for consumption regressions) as listed in Section 3. The columns differ by the temporal aggregation of the weather controls: Columns (1) and (3) use weather controls aggregated at the quarterly level, while Columns (2) and (4) use weather controls aggregated at the monthly level. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are reported in parentheses; p-values are reported in brackets. Appendix A.1 provides further details on variable construction.

Alternative transformations of yield. The negative impact of a delayed onset on agricultural productivity persists when using alternative transformations of the yield variable. Table B6

shows results using log(1+yield), the inverse hyperbolic sine (IHS) transformation of yield, and yield levels (after trimming the 1st and 99th percentiles to mitigate outlier influence).

Table B6: Alternative transformations of yield

Dependent variable transformation	**Yield (trimmed)** (1)	**Log(1+Yield)** (2)	**IHS(Yield)** (3)
Onset of the rainy season	-23.513	-0.028	-0.030
	(10.898)	(0.009)	(0.009)
	[0.031]	[0.001]	[0.001]
Standardised effect	-36.476	-0.044	-0.046
Mean (dep. var.)	801.653	5.547	6.203
Identifying observations	133,530	134,880	134,880
Singleton observations	182	181	181
Countries	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022

Note: OLS regression estimates of Equation 1. Each column uses a different transformation of yield (measured in 2020 USD per hectare) as the dependent variable. Column (1) uses yield levels (measured in 2020 US dollars per hectare), with the 1st and 99th percentiles trimmed. Column (2) uses the log of (1 + yield). Column (3) uses the inverse hyperbolic sine (IHS) transformation of yield. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects, climatic area by country by year fixed effects, weather controls, and plot-level controls (a full list of these controls is presented in Section 3). Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are reported in parentheses; *p*-values are reported in brackets. Appendix A.1 provides further details on variable construction.

Alternative measures of output. The negative impact of a delayed onset extends to other measures of farm output. Table B7 presents results for the impact of onset timing on: output in kilograms, output in constant 2020 US dollars, log of yields in kilograms per hectare, and an indicator for crop failure, defined as zero reported yield.

Table B7: Alternative output measures

Dependent Variable:	**Log output (2020 USD)** (1)	**Log output (kg)** (2)	**Log yield (kg per hectare)** (3)	**Crop failure** (4)
Onset of the rainy season	-0.019	-0.028	-0.031	0.002
	(0.008)	(0.008)	(0.006)	(0.001)
	[0.015]	[0.000]	[0.000]	[0.047]
Standardised effect	-0.030	-0.044	-0.049	0.003
Mean (dep. var.)	4.331	5.407	6.824	0.037
Identifying observations	131,185	131,756	130,475	135,504
Singleton observations	179	178	179	178
Countries	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are reported in parentheses; *p*-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Constant national prices. The benchmark yield measure is valued using median prices at the enumeration area level. If a delayed onset causes a local negative supply shock, local prices may rise, attenuating the estimated negative impact on revenue. To address this, I construct an alternative yield measure where output is valued using constant national median prices. I use a hierarchical approach: prices are first assigned based on the national median for the specific *crop*; if unavailable, I use the national median for the *crop type*; finally, I index any remaining

gaps using a country-year weighted price index. Table B8 reestimates Table 1.

Table B8: Onset of the rainy season and agricultural productivity: Constant national prices

Dependent variable:	**Log yields (2020 USD per hectare)**		
	(1)	(2)	(3)
Onset of the rainy season	-0.028***	-0.026***	-0.027***
	(0.008)	(0.009)	(0.009)
Standardised effect	-0.045	-0.041	-0.042
Mean (dep. var.)	5.751	5.751	5.751
Adjusted R^2	0.454	0.454	0.491
Observations	130,777	130,777	130,142
Planting year range	2008–2022	2008–2022	2008–2022
Weather controls	No	Yes	Yes
Demographic controls	No	No	Yes

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. Local prices per crop are calculated at constant national levels using a hierarchical procedure detailed in Section B.3. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. A full list of controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Cereal crops subsample. To ensure that the results are not driven by composition effects from aggregating diverse crop types with different physical properties (for instance, heavier tubers vs. lighter cereals), I compute plot-level yields for cereals. Table B9 reestimates Table 1.

Table B9: Onset of the rainy season and cereals productivity

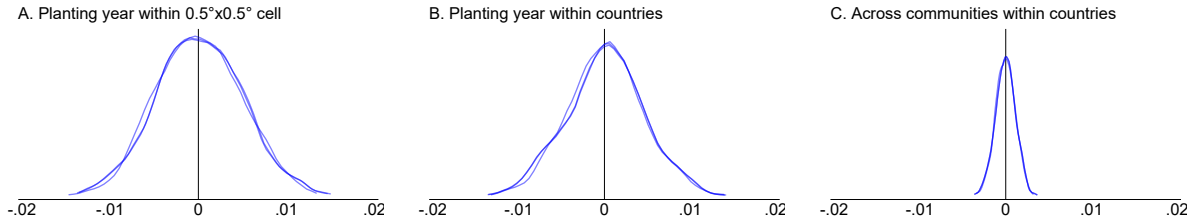
Dependent variable:	**Log yields (2020 USD per hectare)**		
	(1)	(2)	(3)
Onset of the rainy season	-0.023	-0.023	-0.019
	(0.011)	(0.012)	(0.010)
	[0.037]	[0.051]	[0.064]
Standardised effect	-0.034	-0.032	-0.027
Mean (dep. var.)	6.762	6.762	6.761
Identifying observations	73,946	73,946	73,614
Singleton observations	113	113	113
Countries	5	5	5
Planting year range	2009–2022	2009–2022	2009–2022
Weather controls	No	Yes	Yes
Demographic controls	No	No	Yes

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield of cereals grown in a plot, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. A full list of controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Permutation-based inference. To further evaluate the statistical robustness of the estimated effect of rainy season onset on agricultural productivity, I implemented a series of permutation tests. These tests involve re-estimating the impact of rainy season onset (equation 1, using specifications from Table 1) across 500 iterations for each of three distinct randomisation scenarios. Scenario A permutes planting years within local $0.5° \times 0.5°$ grid cells, effectively reassigning the timing of the onset shock experienced by a household to that of another household in a different wave but within the same local area. Scenario B applies a similar permutation of planting years but within the broader scope of each country. Scenario C directly tests the importance of

an observation's specific geographic location by randomising its geographic identifiers within its country. This process reassigns the experienced weather and onset data to the household while holding its original planting year constant.

Figure B9: The effect on yields: permutation-based inference



		Empirical p-values		
	Table 1 coefficients	A	B	C
Benchmark				
specification (1)	-0.020	0.000	0.000	0.000
specification (2)	-0.020	0.000	0.000	0.000
specification (3)	-0.019	0.000	0.000	0.000

Note. Empirical distributions of estimated coefficients for *Onset of the rainy season* from 500 permutations for three randomisation scenarios. Panel A: Planting year permuted within $0.5° \times 0.5°$ grid cells. Panel B: Planting year permuted within countries. Panel C: Geographic location (and thus onset/weather) permuted across communities within countries, holding planting year constant. Each distribution is generated by re-estimating the three benchmark specifications from Table 1. The sub-table reports the empirical p-values, indicating the proportion of permuted coefficients that are as extreme or more extreme than the actual estimated coefficients. All models include the full set of fixed effects and controls as per Equation 1. Appendix A.1 provides detailed information on variables.

Alternative clustering of standard errors. The benchmark analysis clusters standard errors at the $0.5° \times 0.5°$ grid-cell level to account for potential spatial correlation in errors. To assess the sensitivity of statistic inference to this choice, Table B10 presents the estimated coefficient for rainy season onset on log yields from the preferred specification (Table 1, Column 3) under several alternative clustering assumptions. These include clustering at finer ($0.1° \times 0.1°$ grid cell) and coarser ($1° \times 1°$ grid cell) spatial units, as well as two-way clustering by spatial unit and planting year.

Table B10: Alternative clustering of standard errors

Dependent variable:	**Log yields (2020 USD per hectare)**					
Level of clustering	None	0.1°×0.1°	0.5°×0.5°	1°×1°	0.5°×0.5° × planting year	1°×1° × planting year
	(1)	(2)	(3)	(4)	(5)	(6)
Onset of the rainy season	-0.021	-0.021	-0.021	-0.021	-0.021	-0.021
	(0.004)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
	[0.000]	[0.001]	[0.002]	[0.004]	[0.023]	[0.025]
Standardised effect	-0.033	-0.033	-0.033	-0.033	-0.033	-0.033
Mean (dep. var.)	5.744	5.744	5.744	5.744	5.744	5.744
Identifying observations	129,919	129,919	129,919	129,919	129,919	129,919
Singleton observations	180	180	180	180	180	180
Countries	6	6	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Each column uses a different clustering level for standard errors. Standard errors are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

B.4 Stability of impacts on productivity over time

This section tests whether the impact of a delayed rainy season onset is stable over time. To do this, I leverage the temporal dimension of the data for the subset of countries with a sufficient number of survey waves to identify within-location changes between two distinct periods. Thus, identification of the interaction terms involving the *Later waves* variable relies solely on the variation within Ethiopia, Malawi, and Nigeria, which are the only countries in the sample with four or more waves allowing for the definition of distinct early and late periods.

I create an indicator variable, *Later waves*, which takes the value 1 for more recent survey periods. The definition of this variable is country-specific to reflect the different survey timings: for Ethiopia, *Later waves* = 1 for Waves 3-5 (covering years 2015-2022) and 0 for Waves 1-2 (2011-2013); for Malawi, it equals 1 for Waves 3-4 (2016-2019) and 0 for Waves 1-2 (2010-2013); and for Nigeria, it equals 1 for Waves 3-4 (2015-2018) and 0 for Waves 1-2 (2010-2012). For all other countries (Mali, Niger, Tanzania), this variable effectively remains constant within the available waves, and thus observations from these countries do not contribute to identifying the coefficients on terms interacted with *Later waves*.

Table B11 tests for a change in the average impact over time by interacting the onset week with this *Later waves* indicator. Table B12 expands this to a triple-interaction model to test whether the buffering or vulnerability effects of farmer characteristics have also changed over time.

Table B11: Dynamic effect of the onset on agricultural productivity

Dependent variable:	**Log yields (2020 USD per hectare)**		
	(1)	(2)	(3)
Onset of the rainy season	-0.018	-0.018	-0.019
	(0.007)	(0.007)	(0.006)
	[0.005]	[0.011]	[0.004]
× Later waves	-0.014	-0.013	-0.011
	(0.005)	(0.005)	(0.004)
	[0.005]	[0.008]	[0.011]
Mean (dep. var.)	5.744	5.744	5.744
Identifying observations	130,556	130,556	129,919
Singleton observations	182	182	180
Countries	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022
Weather controls	No	Yes	Yes
Demographic controls	No	No	Yes

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. The regressions are run on the full sample, but the identification of the interaction term × *Later waves* relies solely on variation within Ethiopia, Malawi, and Nigeria. *Later waves* is an indicator variable equal to 1 for the following periods: Ethiopia (2015-2022), Malawi (2016-2019), and Nigeria (2015-2018); and 0 otherwise. The main effect of *Later waves* is absorbed by the fixed effects. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. A full list of controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. *p*-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Table B12: Dynamic heterogeneous effects of the onset on agricultural productivity

Dependent variable:	**Log yields (2020 USD per hectare)**				
Indicator:	Female-managed plot	Educated plot manager	Below median assets	Irrigated	Improved seeds
	(1)	(2)	(3)	(4)	(5)
Onset of the rainy season	-0.018	-0.018	-0.014	-0.019	-0.007
	(0.006)	(0.007)	(0.006)	(0.007)	(0.009)
	[0.005]	[0.005]	[0.021]	[0.004]	[0.453]
× Later waves	-0.008	-0.018	-0.008	-0.010	-0.030
	(0.004)	(0.005)	(0.005)	(0.004)	(0.011)
	[0.054]	[0.000]	[0.081]	[0.014]	[0.006]
× Indicator	-0.001	0.002	-0.005	0.011	0.008
	(0.002)	(0.002)	(0.002)	(0.009)	(0.006)
	[0.741]	[0.409]	[0.009]	[0.234]	[0.201]
× Indicator × Later waves	-0.007	0.009	-0.004	0.012	-0.003
	(0.003)	(0.003)	(0.003)	(0.010)	(0.007)
	[0.014]	[0.001]	[0.169]	[0.233]	[0.646]
Mean (dep. var.)	5.744	5.744	5.743	5.749	5.705
Identifying observations	129,919	129,919	129,886	128,486	105,671
Singleton observations	180	180	180	177	215
Countries	6	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022	2011–2022

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. The regressions are run on the full sample, but the identification of interaction terms involving *Later waves* relies solely on variation within Ethiopia, Malawi, and Nigeria. *Later waves* is an indicator variable equal to 1 for the following periods: Ethiopia (2015-2022), Malawi (2016-2019), and Nigeria (2015-2018); and 0 otherwise. The *Indicator* variables in each column are defined as follows: (1) *Female-managed plot*: an indicator equal to 1 if the plot manager is female; (2) *Educated plot manager*: an indicator equal to 1 if the plot manager has any formal education; (3) *Below median assets*: an indicator equal to 1 if the household's asset index is below the sample median; (4) *Irrigated*: an indicator equal to 1 if the plot is irrigated; (5) *Improved seeds*: an indicator equal to 1 if improved seeds were used. All specifications include the relevant lower-order terms (e.g., *Onset* × *Indicator*, *Indicator* × *Later waves*) and the main effect of the *Indicator*. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

B.5 Supplementary results on adaptation strategies

This appendix provides additional results for adaptation margins, supplementing the summary findings presented in Section 5.3 and Table 4. The tables below present these supplementary findings: Table B13 details responses in the use of various on-farm inputs including seeds, fertilisers, pesticides, and on-farm labour; Table B14 covers crop choice and plot composition; Table B15 shows changes in land use along the extensive and intensive margin; and Table B16 examines sectoral labour allocation.

Table B13: Onset of the rainy season and input use

Dependent variable:	Seed value	Inorganic fertiliser quantity	Inorganic fertiliser value	Days from family	Days from hired workers
	(1)	(2)	(3)	(4)	(5)
Onset of the rainy season	0.005	0.002	-0.014	2.204	-5.854
	(0.009)	(0.013)	(0.009)	(24.010)	(4.320)
	[0.591]	[0.880]	[0.125]	[0.927]	[0.176]
Standardised effect	0.008	0.003	-0.020	3.539	-9.388
Mean (dep. var.)	2.953	3.906	4.787	1070.531	102.819
Identifying observations	121,001	44,516	46,098	117,596	117,145
Singleton observations	198	192	195	143	143
Countries	6	6	6	5	5
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variables are: *Seed value*, the value of seeds used, measured in 2020 US dollars per hectare; *Inorganic fertiliser quantity*, the natural logarithm of the quantity of inorganic fertiliser applied, measured in kilograms per hectare; *Inorganic fertiliser value*, the natural logarithm of the value of inorganic fertiliser used, measured in 2020 US dollars per hectare; *Days from family*, the total number of days per hectare family members worked on the plot; and *Days from hired workers*, the total number of days per hectare hired workers worked on the plot. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1^\circ \times 0.1^\circ$ grid-cell fixed effects, climatic area by country by year fixed effects, weather controls, and plot-level controls as listed in Section 3. Standard errors, clustered at the $0.5^\circ \times 0.5^\circ$ grid-cell level, are reported in parentheses; p-values are in brackets. Appendix A.1 provides further details on variable construction.

Table B14: Onset of the rainy season and crop choice

Dependent variable:	Share of plot value attributed to			Plot contains		
	Cereals	Tubers	Legumes	Cereals	Tubers	Legumes
	(1)	(2)	(3)	(4)	(5)	(6)
Onset of the rainy season	0.001	-0.002	0.000	-0.003	-0.002	-0.001
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
	[0.596]	[0.013]	[0.690]	[0.098]	[0.044]	[0.623]
Standardised effect	0.001	-0.004	0.001	-0.004	-0.004	-0.001
Mean (dep. var.)	0.408	0.056	0.126	0.645	0.076	0.234
Identifying observations	136,855	136,855	136,855	136,855	136,855	136,855
Singleton observations	175	175	175	175	175	175
Countries	6	6	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variables are: *Share of plot value attributed to cereals/tubers/legumes*, the proportion of the total production value of the plot derived from the respective crop category. *Plot contains cereals/tubers/legumes* are indicator variables taking value 1 if the plot contains any crop from the respective category, and 0 otherwise. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1^\circ \times 0.1^\circ$ grid-cell fixed effects, climatic area by country by year fixed effects, weather controls, and plot-level controls as listed in Section 3. Standard errors, clustered at the $0.5^\circ \times 0.5^\circ$ grid-cell level, are reported in parentheses; p-values are in brackets. Appendix A.1 provides further details on variable construction.

Table B15: Onset of the rainy season and land use

Dependent variable:	No. fallow plots	No. cultivated plots	Log total cultivated area	Log total farm size
	(1)	(2)	(3)	(4)
Onset of the rainy season	0.004	0.016	-0.003	0.002
	(0.003)	(0.032)	(0.007)	(0.006)
	[0.128]	[0.610]	[0.631]	[0.729]
Standardised effect	0.007	0.026	-0.005	0.003
Mean (dep. var.)	0.153	4.451	-0.134	0.055
Identifying observations	30,881	30,881	30,613	30,696
Singleton observations	313	313	320	320
Countries	6	6	6	6
Survey year range	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variables are household-level measures: *No. fallow plots*, the number of plots left unused by the household; *No. cultivated plots*, the number of plots used in agriculture by the household; Column (3) *Log Total Cultivated Area*, the natural logarithm of total area in hectares actively cultivated by the household; and Column (4) *Log total farm size*, the natural logarithm of total area in hectares to which the household has planting rights. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects, climatic area by country by year fixed effects, weather controls, and household-level controls (a full list of these controls is presented in Section 3). Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are reported in parentheses; *p*-values are reported in brackets. Appendix A.1 provides further details on variable construction, selected surveys, and weighting procedures. Data are from the LSMS-ISA household-level sample.

Table B16: Onset of the rainy season and sector choice

Dependent variable:	Last 12 months			Last 7 days	
	Any wage work	Agriculture	Services	Any wage work	Household business
	(1)	(2)	(3)	(4)	(5)
A. Working-age individuals					
Onset of the rainy season	-0.001	-0.001	-0.000	-0.002	-0.001
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
	[0.029]	[0.254]	[0.131]	[0.000]	[0.158]
Standardised effect	-0.002	-0.001	-0.001	-0.003	-0.002
Mean (dep. var.)	0.050	0.018	0.025	0.054	0.127
Identifying observations	224,735	224,735	224,735	223,915	224,140
Singleton observations	8	8	8	8	9
Countries	6	6	6	6	6
Survey year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022
B. Household head					
Onset of the rainy season	-0.002	-0.001	-0.001	-0.003	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	[0.167]	[0.464]	[0.438]	[0.005]	[0.136]
Standardised effect	-0.003	-0.001	-0.001	-0.005	-0.004
Mean (dep. var.)	0.121	0.032	0.069	0.130	0.197
Identifying observations	47,653	47,653	47,653	47,402	47,582
Singleton observations	353	353	353	353	355
Countries	6	6	6	6	6
Survey year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variables are indicator variables (1 if the condition is met, 0 otherwise) for individuals of working age: *Any wage work*, reports working for a wage in the past twelve months; *Agriculture*, reports working for a wage in the agricultural sector in the past twelve months; *Services*, reports working for a wage in the service sector in the past twelve months; *Any wage work*, reports working for a wage in the past seven days; *Household business*, reports working for a household business in the past seven days. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects, climatic area by country by year fixed effects, weather controls, and individual-level controls (sex, age, age squared, marital status, formal education, urban residence, household size) as specified in Section 4. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are reported in parentheses; *p*-values are in brackets. Appendix A.1 provides further details on variable construction.

Table B17: Lagged onset on adaptation

Dependent variable:	Timing			Inputs			Main crop is		
	Planting (DOY)	Planting duration	Harvest (DOY)	Seeds	Fertiliser	Pesticide	Cereals	Tubers	Legumes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Onset of the rainy season, $t-1$	0.036	0.001	-0.075	-0.004	-0.003	0.002	-0.003	0.001	0.000
	(0.208)	(0.003)	(0.999)	(0.009)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
	[0.862]	[0.769]	[0.940]	[0.639]	[0.100]	[0.098]	[0.047]	[0.175]	[0.760]
Standardised effect	0.062	0.002	-0.126	-0.007	-0.005	0.004	-0.005	0.002	0.001
Mean (dep. var.)	188.934	1.105	250.370	3.453	0.516	0.066	0.572	0.069	0.122
Identifying observations	125,609	125,609	106,970	120,669	135,081	135,306	136,855	136,855	136,855
Singleton observations	74	74	117	197	173	173	175	175	175
Countries	5	5	5	6	6	6	6	6	6
Planting year range	2009–2022	2009–2022	2009–2022	2009–2022	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variables are defined as follows. *Planting (DOY)* is the approximated first planting day of the year; this approximation assumes planting occurred on the 15th of the reported month. *Planting duration* is the number of distinct months planting occurred on the plot. *Harvest (DOY)* is the approximated first harvest day of the year, assuming harvest occurred on the 15th of the reported month. *Seed* is the log of seed quantity used on the plot, measured in kilograms per hectare. *Fertiliser* is an indicator variable taking the value 1 if any fertiliser, whether organic or inorganic, was used, and 0 otherwise. *Pesticide* is an indicator variable that equals 1 if any pesticide was used and 0 otherwise. *Cereals*, *Tubers*, and *Legumes* are indicator variables taking the value 1 if the main crop by production value on the plot belongs to the respective category, and 0 otherwise. *Onset of the rainy season,* $t-1$ is the week of the year when conditions specified in Section 3 in the year before the current season. All specifications include $0.1^\circ \times 0.1^\circ$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors, which are clustered at the $0.5^\circ \times 0.5^\circ$ grid-cell level, are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides further details on variable construction, selected surveys, and weighting procedures.

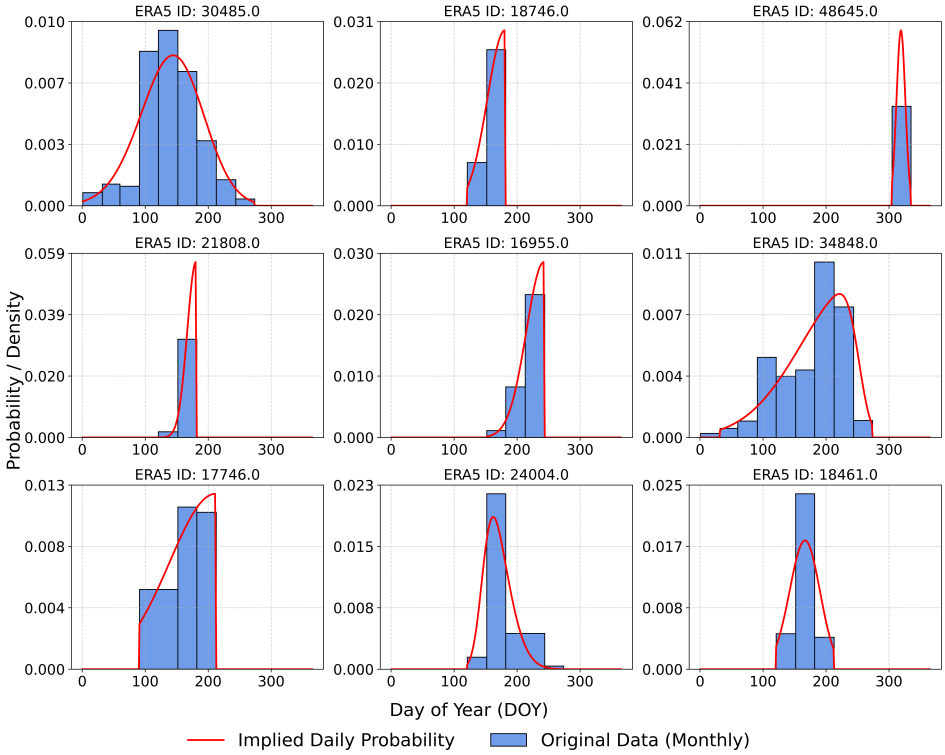
B.6 Simulation methodology to correct for aggregation bias

The use of monthly-level planting data may introduce aggregation bias when I estimate the effect of daily-level weather shocks. To address this, I implement a simulation exercise designed to recover the true daily-level coefficient from my coarse monthly data. The simulation involves two main stages: first, generating daily planting probability distributions for each location, and second, using these distributions in an iterative search algorithm to find the unbiased daily coefficient.

Generating daily planting probability distributions. To disaggregate the observed monthly planting dates into daily probabilities, I generate a daily probability mass function (PMF) for each unique ERA5 location in my sample. This PMF represents the likelihood of planting occurring on any given day of the year. I employ two distinct methods to construct these distributions as a robustness check. For each method, I generate 20 unique *drawing pools*, one for each iteration of the main simulation. Each drawing pool contains 100 simulated planting dates for each plot observation, providing a rich set of potential outcomes for the simulation. To reduce noise from outlier responses, I only consider planting months that, for a given location, account for at least 1% of the total observations for that location.

The first method is a parametric approach that fits a three-parameter skew-normal distribution to the sequence of observed planting *months*. This captures the timing and potential asymmetry of the planting season. To correctly model seasons that cross over the new year (e.g., November–February), I transform the month data onto a continuous axis where January becomes month 13 and February becomes month 14, allowing the distribution to be fitted over an uninterrupted domain. The resulting continuous probability density function is then discretised to generate daily probabilities. Figure B10 provides a visual representation of this method for a random sample of locations.

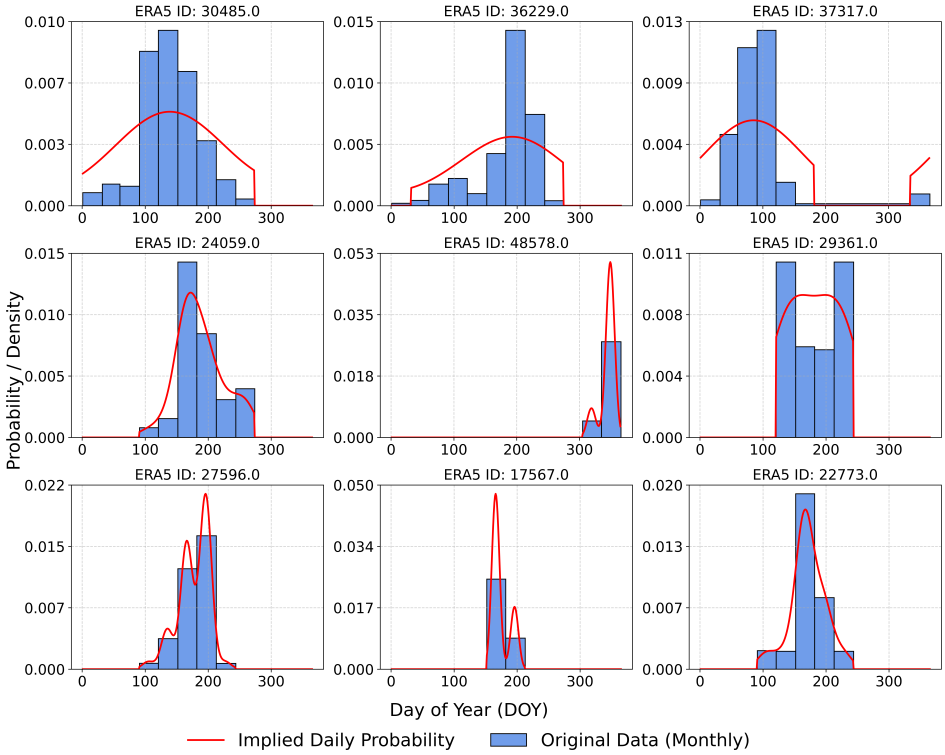
Figure B10: Implied daily planting probabilities from kernel density estimation



Note. This figure displays a random sample of nine ERA5 locations to illustrate the fit of the skew-normal distribution. The blue bars represent the density of observed monthly planting dates, with the day-of-year (DOY) approximated as the 15th of the reported month. The red line is the implied daily planting probability, derived by fitting a three-parameter skew-normal distribution to the sequence of observed months for each location. To reduce noise from outlier responses, only planting months that account for at least 1% of the total observations for a given location are considered. The model accounts for year-crossing seasons by transforming the month variable onto a continuous axis (e.g., January becomes month 13) before fitting the distribution.

As an alternative, I use a non-parametric Kernel Density Estimation (KDE). This method is more flexible and fits a Gaussian KDE directly to the day-of-year (DOY) of each planting observation, approximated as the 15th of the reported month. To handle the circular nature of annual data, I employ a *mirroring* technique. Planting dates from early in the year (e.g., day 15) are duplicated and shifted forward by a year (to day 380), while points from late in the year (e.g., day 350) are duplicated and shifted backward (to day -15). This ensures the KDE treats the year as a continuous loop, correctly modeling the density across the year-end boundary. Figure B11 illustrates the fit of the KDE method.

Figure B11: Implied daily planting probabilities from kernel density estimation



Note. This figure displays a random sample of nine ERA5 locations to illustrate the fit of the Kernel Density Estimator (KDE). The blue bars represent the density of observed monthly planting dates, with the day-of-year (DOY) approximated as the 15th of the reported month. The red line is the implied daily planting probability, derived by fitting a Gaussian KDE to the DOY data for each location. To reduce noise from outlier responses, only planting months that account for at least 1% of the total observations for a given location are considered. The model accounts for the circular nature of annual data using a *mirroring* technique, where data points near the start and end of the year are duplicated and shifted to ensure a continuous density across the year-end boundary.

Simulation design. The simulation exercise uses an iterative search algorithm to find the true daily-level coefficient (β_{daily}) that, when subjected to the same temporal aggregation as my observed data, reproduces the coefficient estimated from the original monthly data. The process is as follows:

1. **Establish target coefficient.** First, I establish a benchmark by running my main regression specification. I convert the observed planting months to a day-of-year (approximated as the 15th of the month) and regress this on the onset of the rainy season. The resulting coefficient serves as the target for the simulation to match, denoted as β_{target}.

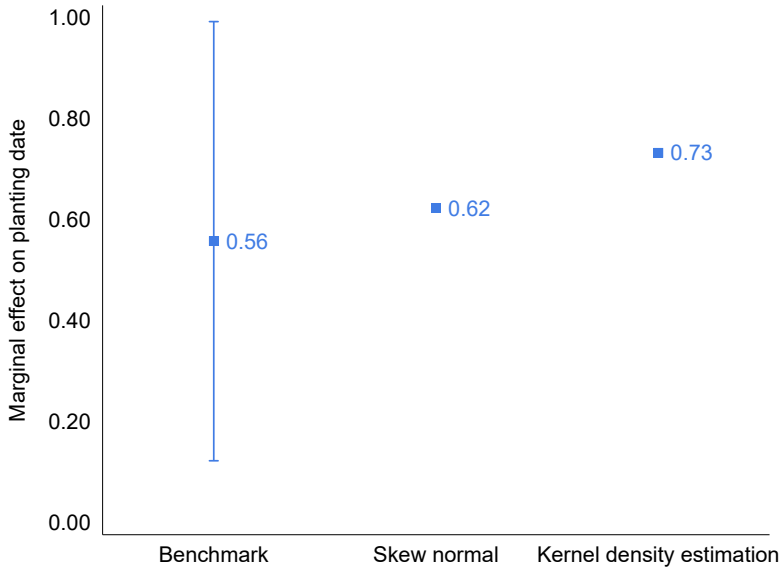
2. **Generate stable residuals.** Before the main simulation, I run a preliminary 10-iteration loop to generate a stable counterfactual residual for my treatment variable. This step purges the residual of any confounding effects that might arise from shifts in the year-level fixed effects, which can occur when the treatment effect is applied to the daily data.

3. **Iterative search for β_{daily}.** The core of the simulation is a `while` loop that searches for the true daily coefficient. The search begins with an initial guess where $\hat{\beta}_{daily} = \beta_{target}$.

In each iteration, I simulate a daily planting date for each observation by taking a random draw from the appropriate daily probability distribution and adding the treatment effect, which is calculated as the stable residual multiplied by the current guess for $\hat{\beta}_{daily}$. This new dataset of simulated daily planting dates is then aggregated back to the monthly level, mimicking the coarseness of the original survey data. I re-run the benchmark regression on this simulated monthly data to obtain a new coefficient, $\beta_{simulated}$. The algorithm compares $\beta_{simulated}$ to the target β_{target}, adjusts the guess for $\hat{\beta}_{daily}$ based on the difference, and the loop repeats. The process continues until $\beta_{simulated}$ converges to β_{target} (i.e., the difference is within a pre-defined tolerance).

The final, converged value of $\hat{\beta}_{daily}$ is my estimate of the true, unbiased daily-level coefficient. This entire procedure is performed independently for both the Skew-Normal and KDE-based drawing pools.

Figure B12: Correcting for aggregation bias in planting dates



Note. Comparison of regression coefficients from observed, monthly-reported planting data with two simulated point estimates of the underlying true daily coefficient. The *Benchmark* is the coefficient from a regression of planting day on the week of rainy season onset using the coarse monthly data; the vertical line represents its 95% confidence interval. This benchmark coefficient serves as the target for the simulation. The *Skew Normal* and *Kernel Density Estimation* points are the results of an iterative search for the true daily-level coefficient that, after being coarsened to a monthly frequency, reproduces the benchmark estimate. The y-axis represents the number of days planting is delayed for a one-week delay in the rainy season onset. All regressions are based on equation (1). See Appendix B.6 for further details on the simulation.

B.7 The role of adjusting planting dates

As established in the main text, farmers adapt to a delayed rainy season by shifting their planting dates. While this low-cost adaptation strategy is effective—a finding supported by the instrumental variable (IV) analysis presented in this appendix—quantifying the magnitude of the effect is challenging. The primary difficulty is the endogeneity of a farmer's planting decision, which biases simple OLS estimates. For instance, the OLS estimate in Column 1 of Table B18 suggests only a minimal mitigating effect .

To obtain a causal estimate, I therefore employ an instrumental variable (IV) strategy that leverages local planting norms. The core intuition is twofold. First, for the instrument to be relevant, it must predict a farmer's own planting date. I argue that the planting decisions of farmers in a broad surrounding area, driven by shared agro-climatic conditions and cultural norms, will be highly correlated with an individual farmer's timing. Second, for the instrument to be valid, it must satisfy the exclusion restriction—meaning it should not directly affect a plot's yield other than through the farmer's own planting choice. A potential threat to this restriction would be an unobserved, spatially correlated shock like a local weather event that affects both neighbours' planting and the farmer's own yields. To address this, the model conditions on weather outcomes measured at the farmer's own high-resolution $0.1° \times 0.1°$ cell. Conditional on these local weather controls, the planting dates of neighbors should not have a direct impact on a farmer's own productivity. Based on this logic, I construct a leave-one-out mean instrument. I first create an adjusted planting date for each plot by taking the residual based on equation (1). This regression includes location fixed effects at the $0.5° \times 0.5°$ grid-cell level and the benchmark local trend, which effectively de-means and de-trends the planting date. The instrument is then the average of these residuals for all other plots within a broad $1° \times 1°$ region, excluding those from the plot's own immediate $0.1° \times 0.1°$ cell.

The results of this IV strategy are presented in Columns (2) and (3) of Table B18. The reduced-form estimate in Column (2) is positive and statistically significant, confirming that during a delayed onset, crop yields are higher in regions where neighbouring farmers also delayed their own planting. Column (3) presents the main two-stage least squares (2SLS) estimate. This estimate must be interpreted with caution, as the instrument is weak by conventional standards (significant only at the 10% level). With this caveat in mind, the coefficient provides our best available, though suggestive, estimate of the causal effect. The magnitude of the coefficients implies that a strategic delay in planting of approximately two weeks is sufficient to offset half of the negative productivity shock from a one-week delay in the rainy season. This greater than one-to-one tradeoff can be understood through two complementary factors. First, the estimated benefit of delaying planting is likely an underestimate due to measurement error (Section 1.6.3); the monthly reporting of planting dates attenuates the coefficient, which in turn inflates the calculated number of weeks needed for the trade-off. Second, the IV strategy identifies the

benefit of delaying planting relative to the local average. In a context where the rainy season is systematically shifting later, this local average itself may be sub-optimal. If the entire local planting calendar is lagging behind the shifting climate, the measured coefficient only captures the benefit of a marginal adjustment, not a shift to a truly optimal schedule.

Table B18: Shifting dates as a strategy

Dependent variable:	**Log yields (2020 USD per hectare)**		
Estimator	OLS	OLS (reduced form)	2SLS
	(1)	(2)	(3)
Onset of the rainy season	-0.026	-0.027	-0.026
	(0.009)	(0.009)	(0.010)
	[0.003]	[0.003]	[0.009]
× Planting date (week of the year)	0.000		0.006
	(0.000)		(0.003)
	[0.028]		[0.055]
× Planting date in neigh. cells (week of the year)		0.002	
		(0.001)	
		[0.031]	
Mean (dep. var.)	5.781	5.782	5.782
Identifying observations	118,933	116,058	116,058
Singleton observations	73	75	71
Countries	5	5	5
Interview year range	2009–2022	2009–2022	2009–2022

Note. OLS and 2SLS regression estimates based on Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. This table presents three specifications. Column (1) shows the OLS estimate of the interaction between onset week and the plot's own planting date. Column (2) presents the reduced-form OLS estimate, using the interaction with the instrument. Column (3) presents the two-stage least squares (2SLS) instrumental variable estimate. The instrument, *Planting date in neigh. cells*, is the leave-one-out mean of adjusted planting dates of other plots in the surrounding $1° \times 1°$ grid cell. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects, climatic area by country by year fixed effects, and a full set of weather and plot-level controls. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

B.8 Additional results for nutritional outcomes

This section presents supplementary analyses for the nutritional outcomes discussed in Section 5.2. First, it presents the results for children. Second, it tests the sensitivity of the main findings for women to the specification of the location fixed effects. Finally, it examines the results for women on a broader sample including those outside agricultural households.

Table B19: Onset of the rainy season and children's nutrition

Dependent variable:	**Weight-for-height**	**Wasted**	**Underweight**
	(1)	(2)	(3)
Onset of the rainy season	0.002	0.000	-0.000
	(0.002)	(0.000)	(0.001)
	[0.444]	[0.671]	[0.857]
Standardised effect	0.004	0.000	-0.000
Mean (dep. var.)	-0.424	0.079	0.291
Identifying observations	53,047	53,047	52,628
Singleton observations	14	14	14
Countries	6	6	6
Interview year range	1990–2018	1990–2018	1990–2018

Note. OLS regression estimates of Equation 1. The data are from the DHS (Croft et al., 2018) for children in agricultural households. The dependent variables are defined as follows. *Weight-for-height* is a z-score, representing the standard deviation from the median of the DHS reference population. *Wasted* is an indicator for a weight-for-height z-score below -2. *Underweight* is an indicator for a BMI below 18.5. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include weather controls, individual-level controls, climatic area by country by year fixed effects, and location fixed effects at the $0.5° \times 0.5°$ grid-cell level. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides further details.

Table B20: Table 3 – Finer location fixed effects

Dependent variable:	Weight-for-height	Wasted	Underweight
	(1)	(2)	(3)
Onset of the rainy season	-0.005	0.000	0.001
	(0.006)	(0.002)	(0.002)
	[0.350]	[0.924]	[0.660]
Standardised effect	-0.005	0.000	0.001
Mean (dep. var.)	-0.850	0.110	0.142
Identifying observations	58,670	58,670	51,322
Singleton observations	360	360	426
Countries	6	6	6
Interview year range	1995–2018	1995–2018	1992–2022

Note. OLS regression estimates of Equation 1. The data are from the DHS (Croft et al., 2018) for women in agricultural households. The dependent variables are defined as follows. *Weight-for-height* is a z-score, representing the standard deviation from the median of the DHS reference population. *Wasted* is an indicator for a weight-for-height z-score below -2. *Underweight* is an indicator for a BMI below 18.5. *Onset of the rainy season* is the week of the year of the onset. All specifications include weather controls, individual-level controls, climatic area by country by year fixed effects, and location fixed effects at the $0.1° \times 0.1°$ grid-cell level. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides further details.

Table B21: Table 3 – All women

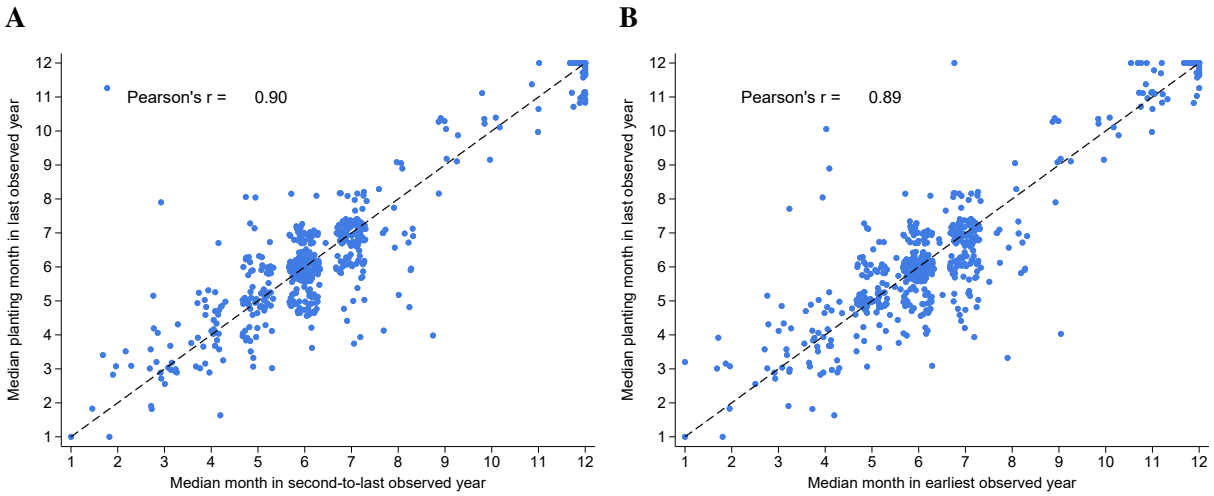
Dependent variable:	Weight-for-height	Wasted	Underweight
	(1)	(2)	(3)
Onset of the rainy season	-0.003	0.000	-0.000
	(0.001)	(0.000)	(0.000)
	[0.016]	[0.528]	[0.902]
Standardised effect	-0.007	0.000	-0.000
Mean (dep. var.)	-0.666	0.098	0.151
Identifying observations	255,422	255,422	229,479
Singleton observations	0	0	0
Countries	6	6	6
Interview year range	1995–2018	1995–2018	1992–2022

Note. OLS regression estimates of Equation 1. The data are from the DHS (Croft et al., 2018). The dependent variables are defined as follows. *Weight-for-height* is a z-score, representing the standard deviation from the median of the DHS reference population. *Wasted* is an indicator for a weight-for-height z-score below -2. *Underweight* is an indicator for a BMI below 18.5. All specifications include weather controls, individual-level controls, climatic area by country by year fixed effects, and location fixed effects at the $0.1° \times 0.1°$ grid-cell level. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides further details.

B.9 Stability of planting months

Figure B13 illustrates the stability of the median planting month over time at the $0.5° \times 0.5°$ level. Panel A plots the median planting month from the most recent survey wave against the median from the prior wave, which are separated by approximately three years on average. Panel B presents the same comparison but uses the earliest available survey wave, showing stability over an average period of approximately six years. In both panels, the points cluster tightly around the 45-degree line, demonstrating a high correlation between the median planting months across survey waves.

Figure B13: Stability of the median planting month



Note. Median planting month over time at the $0.5° \times 0.5°$ level. For locations observed in at least two survey waves, Panel A plots the median planting month from the most recent wave against the median from the prior wave. Panel B plots the most recent wave against the earliest available wave to show stability over a longer period. Points are jittered to show density. The 45-degree dashed line represents perfect stability, where the median planting month does not change between waves. The Pearson's correlation coefficient is reported in the top-left.

B.10 False onsets: Robustness and alternative agronomical pathways

This section tests whether the main findings are sensitive to the parameters used to define a false onset. Table 6 re-estimates the analysis using alternative definitions that vary the length of the search window around the usual onset date and set a more stringent threshold for the dry spell.

Table B22: Table 6 – Different definitions of false onsets

Dependent variable:	Planting (DOY)	Seeds	Log yields (2020 USD per hectare)
	(1)	(2)	(3)
A. 70-day search window and 7-day dry spell			
Onset of the rainy season	0.430	0.031	-0.025
	(0.228)	(0.012)	(0.007)
	[0.060]	[0.012]	[0.001]
× False onset	-0.989	0.086	-0.066
	(0.377)	(0.054)	(0.028)
	[0.009]	[0.108]	[0.018]
Mean (dep. var.)	188.916	3.453	5.744
Identifying observations	125,588	120,650	129,897
Singleton observations	95	216	202
Countries	5	6	6
Planting year range	2009–2022	2009–2022	2008–2022
B. 60-day search window and 10-day dry spell			
Onset of the rainy season	0.450	0.030	-0.023
	(0.230)	(0.012)	(0.007)
	[0.051]	[0.014]	[0.001]
× False onset	-0.714	0.077	-0.036
	(0.316)	(0.018)	(0.011)
	[0.024]	[0.000]	[0.001]
Mean (dep. var.)	188.923	3.453	5.744
Identifying observations	125,598	120,658	129,907
Singleton observations	85	208	192
Countries	5	6	6
Planting year range	2009–2022	2009–2022	2008–2022
C. 70-day search window and 10-day dry spell			
Onset of the rainy season	0.495	0.030	-0.023
	(0.223)	(0.012)	(0.007)
	[0.027]	[0.012]	[0.001]
× False onset	-0.702	0.078	-0.036
	(0.314)	(0.018)	(0.011)
	[0.026]	[0.000]	[0.001]
Mean (dep. var.)	188.923	3.453	5.744
Identifying observations	125,598	120,658	129,905
Singleton observations	85	208	194
Countries	5	6	6
Planting year range	2009–2022	2009–2022	2008–2022

Note. OLS regression estimates. The dependent variable for each column is listed at the top. The analysis examines the interaction between the rainy season onset week and a dummy variable indicating a *False onset*. Benchmark fixed effects are interacted with the false onset dummy. All specifications include a full set of weather and plot-level controls. Standard errors, clustered at the $0.5^\circ \times 0.5^\circ$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides further details.

Table B23: False onsets: Alternative pathways

Dependent variable:	Fertiliser	Pesticide	Days from hired workers
	(1)	(2)	(3)
Onset of the rainy season	-0.000	-0.001	-6.801
	(0.002)	(0.002)	(4.498)
	[0.862]	[0.573]	[0.131]
× False onset	-0.071	-0.015	8.740
	(0.034)	(0.028)	(5.504)
	[0.040]	[0.580]	[0.113]
Mean (dep. var.)	0.516	0.066	102.816
Identifying observations	135,059	135,284	117,139
Singleton observations	195	195	149
Countries	6	6	5
Planting year range	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates based on Equation (1). *Fertiliser* is an indicator variable taking the value 1 if any fertiliser, whether organic or inorganic, was used, and 0 otherwise. *Pesticide* is an indicator variable that equals 1 if any pesticide was used and 0 otherwise. *Days from hired workers*, the total number of days per hectare hired workers worked on the plot. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. The analysis examines the interaction between the rainy season onset week and a dummy variable indicating a *False onset*. All specifications include the main effect of the false onset dummy (absorbed) and fully flexible fixed effects (all benchmark fixed effects are interacted with the false onset dummy). All specifications include a full set of weather and plot-level controls. Standard errors, clustered at the $0.5° \times 0.5°$ grid-cell level, are in parentheses. p-values are in brackets. Appendix A.1 provides further details.

B.11 Additional results on heterogeneity

Table B24 presents additional results on heterogeneous impacts on productivity, extending the analysis in Section 5.4. Table B25, instead, presents results on heterogeneous impacts on planting dates. Table B26 presents results on heterogeneous impacts on other adaptation measures, such as the use of seeds and fertilisers.

Table B24: Additional results on heterogeneous impacts on productivity

Dependent variable:	Log yields (2020 USD per hectare)				
	(1)	(2)	(3)	(4)	(5)
Onset of the rainy season	-0.021	-0.021	-0.020	-0.024	-0.024
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
	[0.002]	[0.002]	[0.004]	[0.001]	[0.001]
× Manager has worked (past 12 m)	0.003				
	(0.002)				
	[0.204]				
× Household head has worked (past 12m)		0.002			
		(0.002)			
		[0.233]			
× Household head is female			-0.006		
			(0.001)		
			[0.000]		
× Household head is educated				0.006	
				(0.002)	
				[0.001]	
× Sold produce					0.009
					(0.002)
					[0.000]
Mean (dep. var.)	5.749	5.742	5.743	5.743	5.743
Identifying observations	128,186	129,611	129,670	129,612	129,805
Singleton observations	177	179	180	179	180
Countries	6	6	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Table B25: Heterogeneous effects on planting dates

Dependent variable:	**Planting DOY**				
	(1)	(2)	(3)	(4)	(5)
Onset of the rainy season	0.542	0.574	0.555	0.600	0.736
	(0.223)	(0.224)	(0.227)	(0.215)	(0.276)
	[0.015]	[0.011]	[0.015]	[0.005]	[0.008]
× Manager is female	0.093				
	(0.083)				
	[0.264]				
× Manager has formal education		-0.034			
		(0.060)			
		[0.570]			
× Below median assets			0.008		
			(0.069)		
			[0.902]		
× Irrigated plot				-0.832	
				(0.430)	
				[0.054]	
× Improved seeds					0.361
					(0.122)
					[0.003]
Mean (dep. var.)	188.934	188.934	188.968	188.753	179.593
Identifying observations	125,609	125,609	125,533	124,531	108,328
Singleton observations	74	74	74	74	99
Countries	5	5	5	5	5
Planting year range	2009–2022	2009–2022	2009–2022	2009–2022	2011–2022

Note. OLS regression estimates of Equation 1. *Planting (DOY)* is the approximated first planting day of the year. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Table B26: Heterogeneous impacts on other adaptation measures

Dependent variable:	**Seeds**		**Fertiliser**		**Pesticide**	
	(1)	(2)	(3)	(4)	(5)	(6)
Onset of the rainy season	0.029	0.027	-0.001	-0.000	-0.001	0.000
	(0.011)	(0.011)	(0.002)	(0.002)	(0.002)	(0.002)
	[0.011]	[0.015]	[0.605]	[0.960]	[0.462]	[0.895]
× Manager has formal education	0.007		0.000		0.001	
	(0.003)		(0.001)		(0.000)	
	[0.011]		[0.553]		[0.000]	
× Below median assets		0.004		-0.001		-0.001
		(0.003)		(0.000)		(0.000)
		[0.140]		[0.002]		[0.000]
Mean (dep. var.)	3.453	3.453	0.516	0.516	0.066	0.066
Identifying observations	120,558	120,511	135,081	135,007	135,306	135,231
Singleton observations	197	197	173	173	173	173
Countries	6	6	6	6	6	6
Planting year range	2009–2022	2009–2022	2008–2022	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. *Seeds* is the log of seed quantity used on the plot, measured in kilograms per hectare. *Fertiliser* is an indicator variable taking the value 1 if any fertiliser, whether organic or inorganic, was used, and 0 otherwise. *Pesticide* is an indicator variable that equals 1 if any pesticide was used and 0 otherwise. *Days from hired workers* is the total number of days per hectare hired workers worked on the plot. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. All specifications include $0.1° \times 0.1°$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5° \times 0.5°$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

Table B27: Heterogeneous impacts by information access

Dependent variable:	**Log yields (2020 USD per hectare)**		
Indicator:	Owns mobile phone	Owns radio	Owns television
	(1)	(2)	(3)
Onset of the rainy season	-0.021	-0.020	-0.021
	(0.007)	(0.007)	(0.007)
	[0.002]	[0.004]	[0.004]
× Indicator	0.004	0.003	0.000
	(0.002)	(0.002)	(0.004)
	[0.039]	[0.104]	[0.921]
Mean (dep. var.)	5.769	5.750	5.710
Identifying observations	118,329	114,985	110,279
Singleton observations	179	181	182
Countries	6	6	6
Planting year range	2008–2022	2008–2022	2008–2022

Note. OLS regression estimates of Equation 1. The dependent variable is the log of yield, measured in 2020 US dollars per hectare. *Onset of the rainy season* is the week of the year when conditions specified in Section 3 are met. The regressions are run on the full sample, but the identification of interaction terms involving *Later waves* relies solely on variation within Ethiopia, Malawi, and Nigeria. The *Indicator* variables in each column are binary: (1) *Owns mobile phone* equals 1 if the household owns at least one mobile phone; (2) *Owns radio* equals 1 if the household owns at least one radio; (3) *Owns television* equals 1 if the household owns at least one television set. Each specification interacts the onset week with the respective indicator. All specifications include the main effect of the *Indicator* and control for an indicator variable equal to 1 if the household's asset index is below the sample median. All specifications include $0.1^\circ \times 0.1^\circ$ grid-cell fixed effects and climatic area by country by year fixed effects. All specifications also include weather and plot-level controls; a full list of these controls is presented in Section 3. Standard errors clustered at the $0.5^\circ \times 0.5^\circ$ grid-cell level are reported in parentheses. p-values are reported in brackets. Appendix A.1 provides detailed information on variables, selected surveys, and weighting procedures.

C Methodology for Projecting Economic Damages

This appendix details the methodology used to calculate the Net Present Value (NPV) of future economic damages from delayed rainy season onsets, as presented in the main text. The calculation combines empirical estimates from this paper with country-specific projections of climate and economic variables.

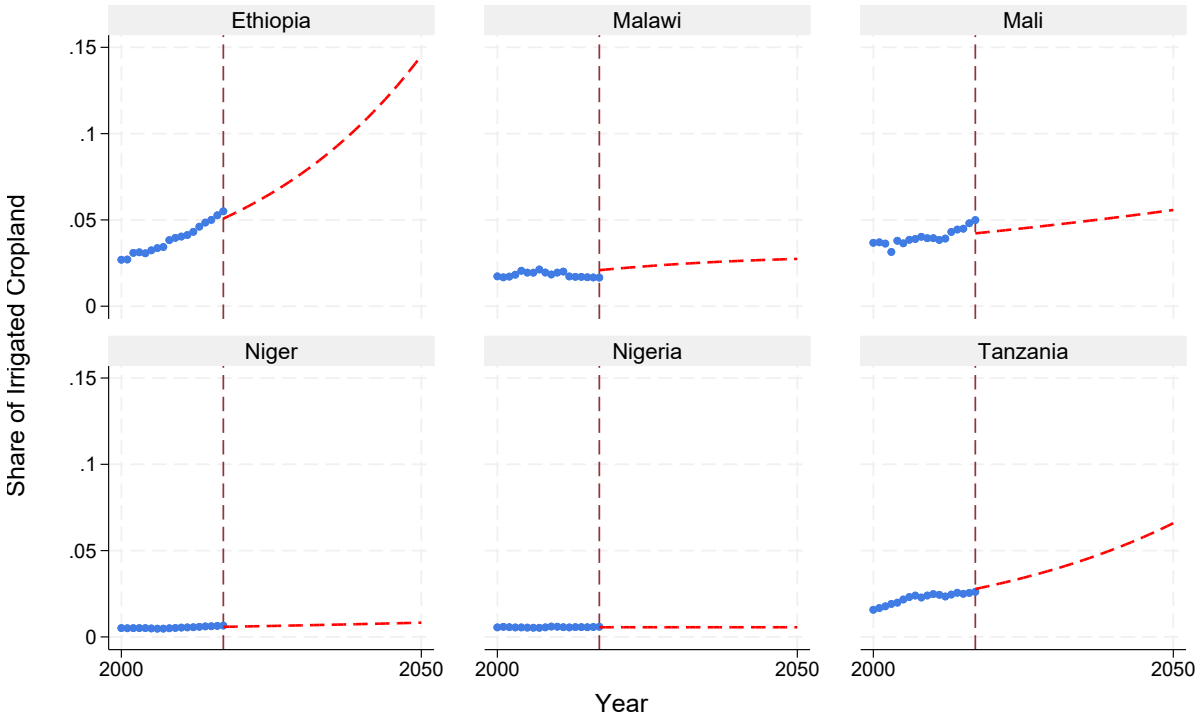
Annual GDP loss. The annual loss in GDP for a given country c in year t is calculated in absolute monetary terms (constant 2015 USD). It is the product of the total agricultural yield loss and the projected real value of agricultural GDP. The full calculation for the annual loss is:

$$\text{Annual Loss}_{ct} = \underbrace{(D_{ct} \times \beta_{\text{yield}})}_{\text{Gross Yield Loss}} \times \underbrace{(1 - I_{ct})}_{\text{Irrigation Adjustment}} \times \underbrace{(A_{ct} \times G_{ct})}_{\text{Real Agri. Value Added}} \tag{2}$$

In this equation, D_t represents the cumulative delay in the rainy season onset in weeks, and β_{yield} is the semi-elasticity of crop yield to this delay (a 2.11% yield loss per week). This gross loss is adjusted by I_t, the projected share of irrigated agricultural land. To forecast I_t through 2050, several functional forms (linear, quadratic, exponential, logistic, Gompertz) are fitted to the historical trend from the HYDE 3.2 dataset (Klein Goldewijk et al., 2017). The model with the best fit for each country, determined by the lowest Bayesian Information Criterion (BIC), is used to extrapolate the trend.

The final two components, A_t and G_t, represent the projected real share of agriculture in the economy and the projected real GDP (in constant 2015 USD) for year t. The historical portion of these series (up to 2024) is constructed using World Bank indicators. To derive the historical real agricultural share (A_t), an implicit GDP deflator is calculated by dividing nominal GDP by real GDP. The real agricultural share is then the nominal agricultural value added divided by this deflator, all divided by real GDP. To address missing data for Ethiopia's nominal agricultural value for 2023 and 2024, values are imputed by applying the country's 10-year average growth rate from 2013-2022. Finally, to project both A_t and G_t forward from 2025 to 2050, the annual growth rates implied by the FAO's long-term agricultural and macroeconomic scenarios are applied to the final historical values.

Figure C1: Historical and projected share of irrigated cropland



Note. The figure shows the historical and projected share of total cropland that is irrigated for each country. The historical data are from the HYDE 3.2 dataset (Klein Goldewijk et al., 2017). The projection from 2018 to 2050 is a forecast based on the best-fitting time series model (selected via BIC) for each country's historical trend.

Scenario framework. To quantify the economic benefits of climate mitigation, the analysis compares two distinct future pathways from 2024 to 2050. These are adapted from the Food and Agriculture Organization's (FAO) long-term global scenarios (FAO, 2018), which are built upon the Shared Socio-economic Pathways (SSPs) framework, formulated by the Intergovernmental Panel on Climate Change Sixth Assessment Report (IPCC, 2021). The scenarios differ in their assumptions for future economic development and for the magnitude of the climate shock itself.

1. **Business as Usual (BAS):** This scenario is adapted from the IPCC's middle of the road pathway (corresponding to SSP2), which assumes that socio-economic and technological trends continue along historical patterns with uneven progress toward sustainability goals. For this scenario, I assume the delay in the onset of the rainy season is based on the projected delay shown in Figure B1.

2. **Toward Sustainability (TS):** This scenario is adapted from the IPCC's sustainability pathway (corresponding to SSP1), which assumes a global shift toward more equitable development and respect for environmental boundaries. To reflect the success of these mitigation efforts, the expected delay in the rainy season onset is assumed to be halved relative to the BAS scenario.

In Equation 2, these scenarios determine the future values for the cumulative delay (D_t), the agricultural share (A_t), and real GDP (G_t), allowing for a direct comparison of their economic consequences.

Social discount rate and NPV. To calculate the Net Present Value (NPV) of this stream of future damages, annual losses must be discounted to the present day. The social discount rate, r_t, is not assumed to be constant but is derived from the DICE-2023 model's Ramsey-rule framework, which accounts for time-varying consumption growth and climate-related capital risks. The formula for the discount rate in year t is:

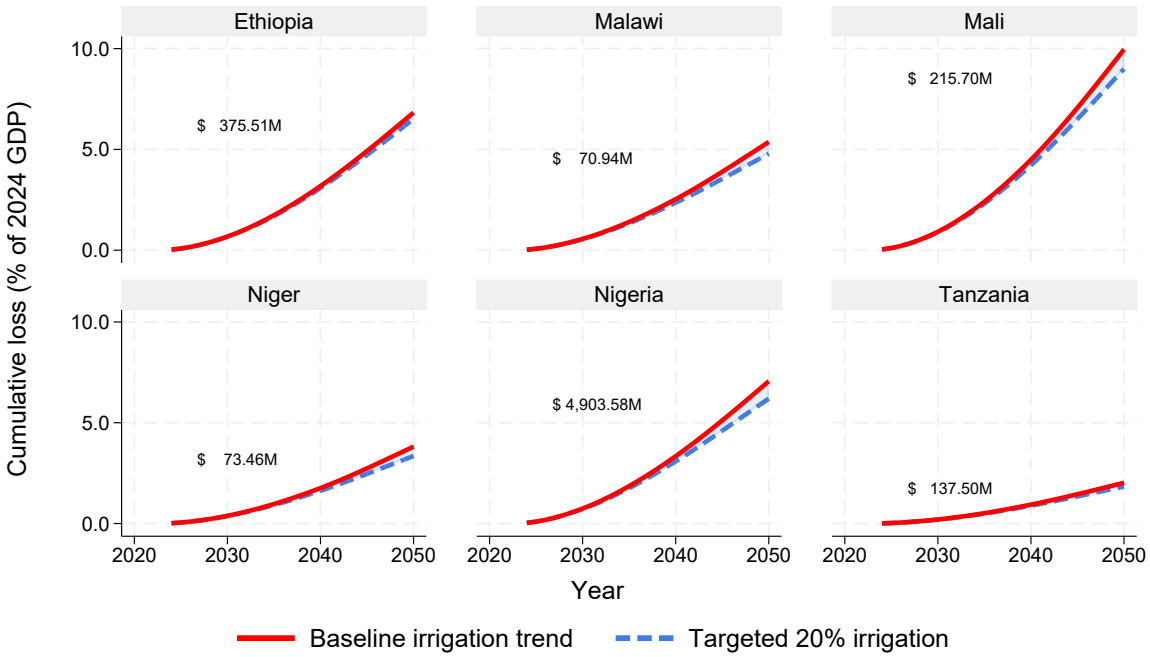
$$r_t = \rho + \phi g_t - \frac{1}{2}\phi^2 \sigma_c^2 t + \beta_{clim} \pi_{capital} \tag{3}$$

The specific parameter values from the DICE-2023 framework are as follows: the rate of pure time preference $\rho = 0.001$; the elasticity of the marginal utility of consumption $\phi = 0.95$; the variance of the logarithm of consumption growth $\sigma_c^2 = (0.01)^2$; a parameter for the systematic risk of climate damages $\beta_{clim} = 0.5$; and the risk premium on capital $\pi_{capital} = 0.05$. The model also incorporates a time-varying growth rate of per capita consumption, g_t, which is assumed to decline linearly from 1.9% to 1.7% over the projection period.

The NPV of damages is then the sum of all discounted annual GDP losses from the start year ($t_0 = 2024$) to the end year ($T = 2050$):

$$NPV = \sum_{t=t_0}^{T} \frac{\text{Annual Loss}_t}{\prod_{i=t_0}^{t}(1 + r_i)} \tag{4}$$

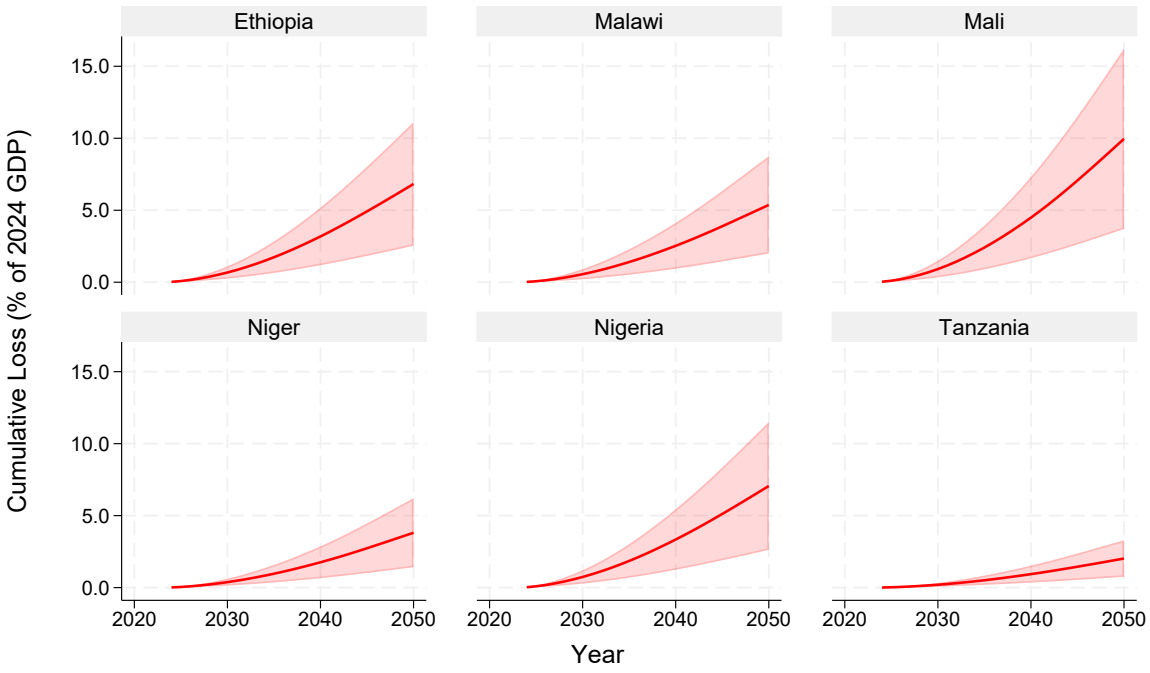
Figure C2: The mitigating effect of irrigation expansion



Note. Benefit of future irrigation expansion under the Business as Usual scenario. The dashed line shows an optimistic counterfactual where the share of irrigated land increases linearly to 20% of agricultural land in all sampled countries. The solid line shows the main projection, which includes a data-driven forecast of irrigation expansion based on historical trends. The shaded area between the lines represents the irrigation benefit—the total value of damages averted due to this specific form of adaptation, quantified in millions of US dollars.

Uncertainty analysis. To provide a sense of the uncertainty surrounding these projections, two sensitivity analyses are conducted. First, uncertainty in the damage coefficient is assessed by recalculating the NPV of damages using the lower and upper bounds of the 95% confidence interval from the benchmark regression. Second, to model uncertainty in future adaptation, an optimistic scenario is run where the damage coefficient, β_{yield}, is assumed to decrease linearly to zero between 2024 and 2050.

Figure C3: Sensitivity of future damages to coefficient uncertainty



Note. Statistical uncertainty of the damage projection under the Business as Usual scenario. The solid line is the central projection, calculated using the point estimate of a 2% yield loss per week of onset delay. The shaded area represents the 95% confidence interval, with the upper and lower bounds calculated by re-running the projection using the corresponding bounds of the yield loss coefficient from the benchmark regression in Table 1.

Table C1: Uncertainty around estimates

Country	Lower Bound (95% CI)	Point Estimate	Upper Bound (95% CI)
Ethiopia	2.53%	6.83%	11.12%
Malawi	1.99%	5.37%	8.75%
Mali	3.70%	9.96%	16.23%
Niger	1.42%	3.81%	6.21%
Nigeria	2.62%	7.06%	11.50%
Tanzania	0.75%	2.02%	3.29%

Note. The table shows the projected NPV of damages by 2050 as a percentage of 2024 GDP under the Business as Usual scenario. Each column uses a different value for the yield loss per week of onset delay, corresponding to the point estimate and the 95% confidence interval from the benchmark regression.