50x2030
DATA-SMART AGRICULTURE

SUBSTANTIAL IMPACTS OF CLIMATE SHOCKS IN AFRICAN SMALLHOLDER AGRICULTURE

Philip Wollburg
Yannick Markhof
Thomas Bentze
Giulia Ponzini

SEPTEMBER 2024
50x2030 WORKING PAPER SERIES

nature sustainability

Analysis

https://doi.org/10.1038/s41893-024-01411-w

Substantial impacts of climate shocks in African smallholder agriculture

Received: 24 January 2024

Accepted: 17 July 2024

Published online: 17 September 2024

Check for updates

Philip Wollburg[1,2] ✉, Yannick Markhof[1,3], Thomas Bentze[1] & Giulia Ponzini[1]

Climate change is affecting the frequency and severity of extreme weather events, such as droughts or floods, which result in loss and damage to people, crops and infrastructure. Global data on loss and damage used in research, policy and media primarily come from macrostatistics based on disaster inventories. Here, we propose a different approach, based on survey microdata. We harmonize data from 120,000 agricultural fields in six African countries for a period from 2008 to 2019 and quantify crop production losses related to climate shocks. We find substantial damages which affect around 35% of plots and reduce national crop production by 29% on average. The economic impacts are greater than the global disaster data suggest. The economic losses resulting from droughts and flood alone are US$5.1 billion higher than reported in disaster inventories, affecting between 145 and 170 million people. The difference stems mostly from smaller and less severe but frequent adverse events that go under-reported or undetected in disaster inventories and therefore elude macrostatistics and reporting. The findings have implications for measurement and policies related to loss and damage and disaster risk reduction.

Large-scale environmental disasters, such as cyclone Freddy in Malawi, Mozambique and Madagascar in 2023, the devastating floods in Pakistan in 2022 or the severe and prolonged drought in the Horn of Africa, affect millions of people, routinely capturing news headlines and trigger national and international responses[1–3]. Anthropogenic climate change probably contributes to the frequency and severity of environmental disasters, a trend which is set to accelerate as global warming progresses[4–6]. This, and the scale of their impacts, makes large disasters a natural focal point of discussions and advocacy around the loss-and-damage fund, established at COP28 in Dubai in November 2023 (ref. 7). In contrast, smaller disaster events and climatic shocks rarely receive widespread attention but the limited evidence suggests that their cumulative effects can be substantial[8–10].

Reporting on disaster impacts relies predominantly on country-level macrodata. A key data source is the Emergency Events Database (EM-DAT), which is a publicly available global inventory of disaster impacts that is widely used in media[11], research[12] and international policy reports[13]. Inventories such as EM-DAT use a specific set of criteria for what constitutes a 'disaster', in terms of people affected and of damage to assets, which, individually, smaller events and climactic shocks may not meet. The implication is that these events and their impacts go undetected or under-reported in global macrodatabases[14].

Here, we offer a different approach to capturing the impacts of disasters and climate shocks, based instead on survey microdata. We quantify the value of crop production losses due to climatic shocks on more than 120,000 fields across six African countries and study their impacts on African agriculture, rural populations and the national economies. Agriculture is a key sector which employs more than half the workforce in the region and on which most poor and rural households depend for their livelihoods[15,16]. The impacts of climate change and environmental disasters are expected to be especially severe in this region[17,18], whereas smallholder agriculture is highly exposed as it remains predominantly rainfed and the adoption of drought- or heat-resistant seeds or other such climate-smart technologies is limited[19].

Our analysis of microdata offers an important complementary perspective to existing analyses based on macrostatistics derived from disaster inventories. Aggregate statistics are critical to the study of disaster impacts, providing annual data at a global scale. They are

[1]The World Bank, Living Standards Measurement Study, Development Economics Data Group, Rome, Italy. [2]Wageningen University, Development Economics Group, Wageningen, the Netherlands. [3]United Nations University, UNU-MERIT, Maastricht, the Netherlands. ✉e-mail: pwollburg@worldbank.org

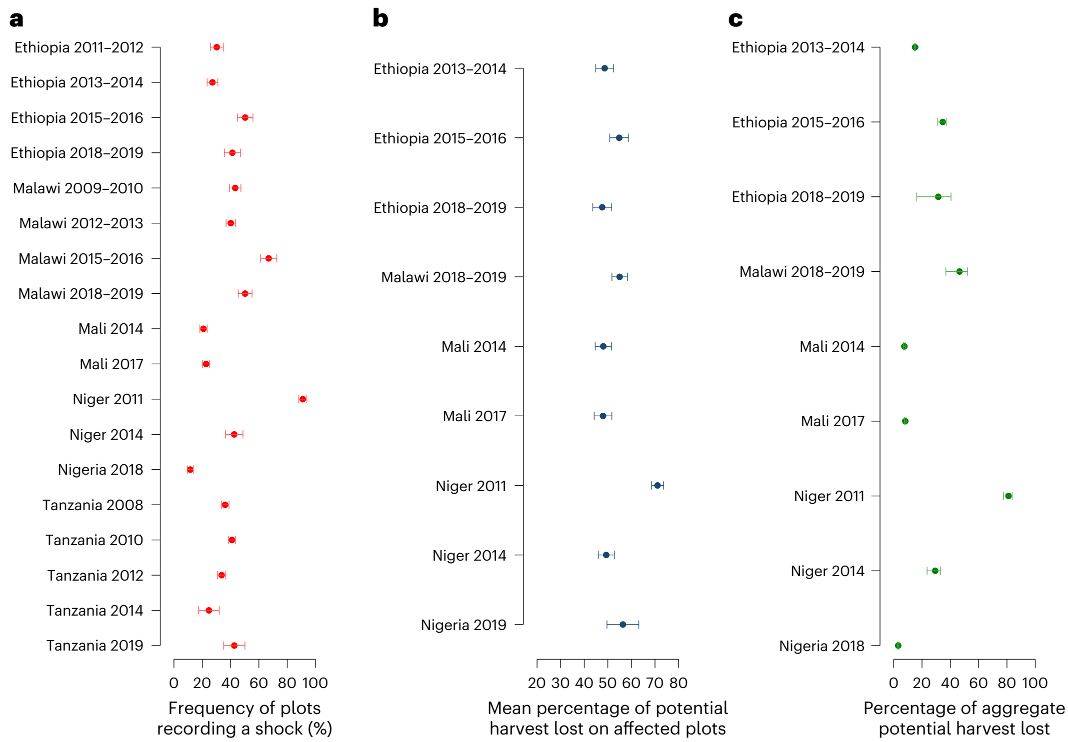

Fig. 1 | Frequency of crop losses due to adverse events, value lost and aggregate production loss. a, Prevalence of crop shocks on plots across country-waves, presented as mean values and their associated 95% confidence interval (CI). **b**, Percentage of potential harvest value lost on plot, by country-wave, presented as mean values and their associated 95% CI. **c**, Percentage of aggregate potential harvest lost (valued with present prices), per country-wave, presented as the ratio of total estimated losses to total estimated potential output and the ratios of their associated 95% CIs. Estimates use population sampling weights. Sample sizes (*n*) for **a** are the following: *n* = 3,613 (Ethiopia 2011–2012); *n* = 14,625 (Ethiopia 2013–2014); *n* = 14,405 (Ethiopia 2015–2016); *n* = 7,795 (Ethiopia 2018–2019); *n* = 5,032 (Malawi 2009–2010); *n* = 5,855 (Malawi 2012–2013); *n* = 3,669 (Malawi 2015–2016); *n* = 5,057 (Malawi 2018–2019); *n* = 8,979 (Mali 2014); *n* = 23,799 (Mali 2017); *n* = 5,792 (Niger 2011); *n* = 4,106 (Niger 2014); *n* = 6,665 (Nigeria 2018); *n* = 2,918 (Tanzania 2008); *n* = 3,646 (Tanzania 2010); *n* = 4,484 (Tanzania 2012); *n* = 793 (Tanzania 2014); *n* = 878 (Tanzania 2019). Sample sizes (*n*) for **b** are the following: *n* = 4,187 (Ethiopia 2013–2014); *n* = 7,717 (Ethiopia 2015–2016); *n* = 3,161 (Ethiopia 2018–2019); *n* = 2,180 (Malawi 2018–2019); *n* = 1,595 (Mali 2014); *n* = 5,794 (Mali 2017); *n* = 5,191 (Niger 2011); *n* = 1,346 (Niger 2014); *n* = 691 (Nigeria 2018). Sample sizes (*n*) for **c** are the following: *n* = 14,317 (Ethiopia 2013–2014); *n* = 14,379 (Ethiopia 2015–2016); *n* = 7,767 (Ethiopia 2018–2019); *n* = 5,017 (Malawi 2018–2019); *n* = 8,921 (Mali 2014); *n* = 23,777 (Mali 2017); *n* = 5,711 (Niger 2011); *n* = 4,048 (Niger 2014); *n* = 6,406 (Nigeria 2018).

less well-suited to capture the differential impacts of disasters on different population groups, especially poor and vulnerable people and are not designed to record smaller climatic shocks[20,21]. They account primarily for damages to assets and losses in agricultural production whose value is greater and better documented amongst richer households and in richer countries. For instance, between 2003 and 2022, of the disasters recorded in Africa by EM-DAT, only 12% contained information on total economic damages. For the same period, just under half the recorded economic losses occurred in the Americas, compared to 1% in Africa[22]. A recent study using the same data source concluded that disaster impacts do not affect poor people as much as the general population[23]. In contrast, evidence from survey microdata suggests that poorer households and individuals are more exposed and less resilient to climatic shocks and suffer disproportionately greater well-being losses than do better-off households[17,20,24]. Our analysis suggests that production losses due to climatic shocks are meaningful not only for the well-being of low-income households individually but, because of how many households are affected, they are important also for the whole economies of our study countries and on a global scale.

Results

Prevalence of crop losses among African farmers

The data used in this analysis are from the Living Standards Measurement Study-Integrated Survey on Agriculture (LSMS-ISA) in Ethiopia, Malawi, Mali, Niger, Nigeria and Tanzania. The data were harmonized across countries and cover close to 120,000 fields on around 30,000 farms. The data show that crop losses due to climatic shocks are widespread and substantial in African smallholder agriculture. Farmers report crop losses on between 12% (Nigeria 2018/19) and 91% of plots (Niger 2011), depending on country and year (Fig. 1a and Supplementary Table 1). Overall, 35% of plots report a crop loss. Farmers reported losing, on average, 53% of their harvest on plots affected by crop shocks (Fig. 1b and Extended Data Table 1). Mean plot losses vary across countries and years, ranging from 48% of harvest (Ethiopia 2018/19) to 71% of harvest (Niger 2011). Crop losses due to climatic shocks have also become more common over time (Supplementary Table 2).

In aggregate, crop losses due to adverse climatic events reduce the total national crop production by between 3% in Nigeria 2018–2019 and 81% in Niger in 2011. An average of 29% of potential harvest value is lost across the countries and agricultural seasons observed in our dataset (Fig. 1b and Extended Data Table 2). In some cases, there is a divergence between aggregate production loss and average plot-level losses, which may mean lower value crops are more likely to suffer losses.

We further quantify the average impact of crop losses on household welfare. We provide indicative evidence of consumption losses amounting to US$35 on average (7.4% of median and 17.6% of bottom quintile consumption among unaffected households) based on a matching exercise (Supplementary Table 3).

Crop production impacted by multiple shocks

Farmers face a diversity of adverse climatic shocks. Several shocks are recorded to affect agricultural production in each year and across all

https://doi.org/10.1038/s41893-024-01411-w

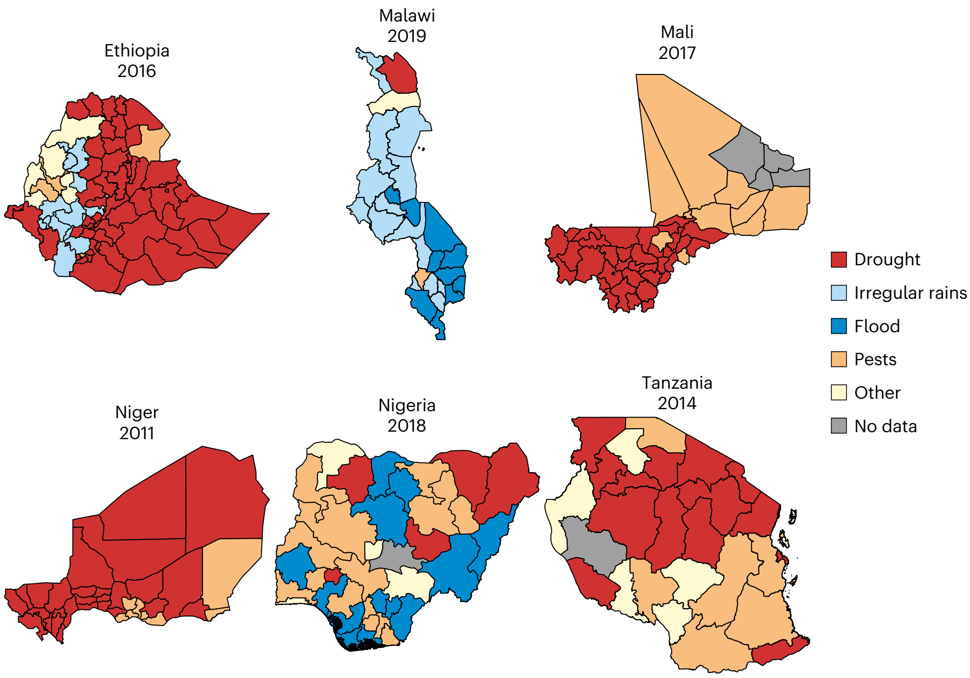

Fig. 2 | Most common climatic shocks by administrative unit, selected countries and years. This figure shows the most frequently recorded climatic shocks that caused crop losses for subnational administrative divisions in selected countries and years. Map created with GADM (https://gadm.org/license.html).

countries (Supplementary Table 1). There are also some instances of several shocks affecting the same farm in a given agricultural season (Supplementary Table 4). This ranges from 1.5% of farms (Tanzania 2014) to 21% of farms (Ethiopia 2018–2019).

Overall, drought is the most common shock and is recorded on an estimated 19% of plots (Supplementary Table 1). More than one in ten plots records losses due to irregular rains, meaning erratic rainfall at unusual times in the agricultural season. Pests are also widespread across our sample, affecting 7% of all plots. Still, there is substantial variation across countries and years. The severity of the damages caused varies between different events (Extended Data Table 1 and Supplementary Table 5). Floods in particular cause more damage than other shocks, reducing crop production per plot on average by 62%. Losses from pests and irregular rains tend to be smaller.

Which shocks are the most prevalent varies between and within countries. Figure 2 illustrates this for selected countries and years, showing the most reported events by subnational administrative divisions. There is some geographical clustering but we commonly see different events accounting for most of the impacted plots in different areas of the same country in the same year. This is true even in years with exceptionally severe events such as the droughts in Niger in 2011 and Ethiopia in 2015–2016 where many, but not all, areas of the country record drought as the primary loss reason.

Local crop losses and farmer characteristics

Not all farmers and plots are equally affected. Some are less likely to experience a loss even in the face of an adverse climatic event. Here, we show that shock exposure and impacts can differ even between neighbouring plots in the same area. We limit this analysis to droughts and floods. Given the nature of droughts, all plots in the same small geographic cluster should be faced with the same drought shock—but the impacts of that drought can differ. Indeed, in 41% of the geographical clusters in our sample, some but not all plots report being affected by a drought, even when they grow the same crops (Supplementary Table 6). Flood losses are (even) more idiosyncratic than drought losses (Extended Data Table 3). Conditional on flood losses being reported on at least one plot in the cluster, only 15% of plots within the same geographical cluster record a flood loss—compared to 35% of plots in the case of drought.

The result extends to plots on the same farm (Supplementary Table 7). Conditional on one maize (sorghum) plot being affected by drought, 67% (80%) of maize (sorghum) plots on the same farm record a drought shock as well.

These findings suggest that climate shock impacts are highly localized, consistent with the high spatial concentration that meteorological events can have[25]. Further, idiosyncratic factors, such as land characteristics and management practices and happenstance play a role in determining whether and how much production is affected. Elevation is negatively associated with the likelihood of experiencing losses and the size of the losses incurred (an effect almost twice as strong for floods compared to other climatic shocks), whereas smaller plots are less likely to suffer losses but record higher losses when they are affected (Extended Data Tables 4 and 5). Losses on intercropped plots are 7.5 percentage points lower than on monocropped plots, although intercropped plots are more likely to experience a loss in the first place (+3.6 percentage points). Plots farmed in more input- and technology-intensive ways appear more resilient to crop shocks.

Shock exposure and impact also vary according to who manages the plot. Plots managed by women are more often affected by crop losses due to climatic shocks (+2.2 percentage points; Extended Data Table 6) than plots managed by men and their losses are also larger on average (+4.4 percentage points; Extended Data Table 7). This may be because plots managed by women are endowed and farmed differently than plots managed by men, which in turn may follow from differential access to inputs and land between women and men[26–28]. We conduct a mediation analysis for these results[29] testing variables capturing potentially differential plot endowments of women and men (Supplementary Table 8). We find that 17% of the difference in shock incidence is explained by women's plots being located at lower elevation than men's plots. For loss size, about 41% of the effect is related to women farming smaller plots. The remaining differences are probably context-dependent and may also relate to the complex interplay of economic with social factors[30].

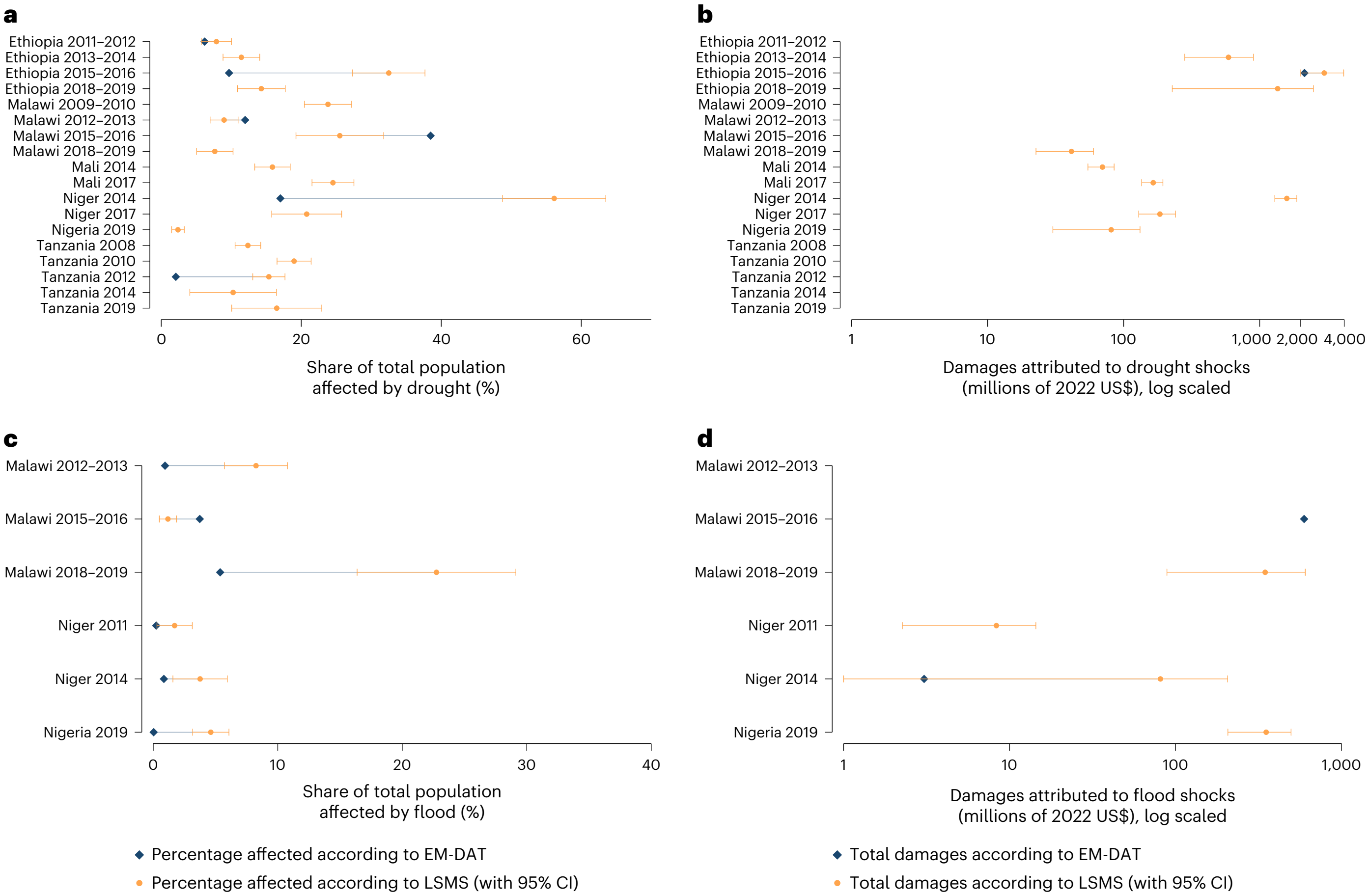

Fig. 3 | Comparison of shock prevalence and impact between EM-DAT and LSMS-ISA data. a, Comparison of the total estimated individuals affected by droughts between EM-DAT (in blue) and LSMS-ISA data (in orange). **b**, Comparison of the estimated damages (in millions of 2022 US$), in years where damages could be estimated in the LSMS-ISA surveys. **c**, Comparison for floods, in years where floods are listed as a potential shock in the LSMS-ISA data. **d**, Comparison of estimated damages from floods. Blue bars plot the total percentage of the population (**a**,**c**) and total damages (**b**,**d**) in EM-DAT, while orange bars plot the percentage affected (**a**,**c**), presented as estimated totals of affected individuals, along with their 95% CIs, divided by the country population and estimated log total damages (**b**,**d**), along with 95% CIs, from the LSMS-ISA data. CIs for **b** and **d** were calculated before log-transformation and are hence asymmetrically situated around log-scaled point estimates. Sample sizes (*n*) for **a** are the following: *n* = 1,548 (Ethiopia 2011–2012); *n* = 2,849 (Ethiopia 2013–2014); *n* = 2,746 (Ethiopia 2015–2016); *n* = 1,902 (Ethiopia 2018–2019); *n* = 2,535 (Malawi 2009–2010); *n* = 2,981 (Malawi 2012–2013); *n* = 1,858 (Malawi 2015–2016); *n* = 2,286 (Malawi 2018–2019); *n* = 2,234 (Mali 2014); *n* = 6,254 (Mali 2017); *n* = 2,226 (Niger 2011); *n* = 1,738 (Niger 2014); *n* = 3,047 (Nigeria 2018); *n* = 1,743 (Tanzania 2008); *n* = 2,025 (Tanzania 2010); *n* = 2,467 (Tanzania 2012); *n* = 438 (Tanzania 2014); *n* = 551 (Tanzania 2019). Sample sizes (*n*) for **b** are the following: *n* = 2,873 (Ethiopia 2013–2014); *n* = 2,761 (Ethiopia 2015–2016); *n* = 1,916 (Ethiopia 2018–2019); *n* = 2,308 (Malawi 2018–2019); *n* = 2,234 (Mali 2014); *n* = 6,254 (Mali 2017); *n* = 2,226 (Niger 2011); *n* = 1,740 (Niger 2014); *n* = 3,382 (Nigeria 2018). Sample sizes (*n*) for **c** are the following: *n* = 2,981 (Malawi 2012–2013); *n* = 1,858 (Malawi 2015–2016); *n* = 2,286 (Malawi 2018–2019); *n* = 2,226 (Niger 2011); *n* = 1,738 (Niger 2014); *n* = 3,059 (Nigeria 2019). Sample sizes (*n*) for **d** are the following: *n* = 2,308 (Malawi 2018–2019); *n* = 2,226 (Niger 2011); *n* = 1,752 (Niger 2014);); *n* = 3,403 (Nigeria 2019).

Finally, we also find losses to be more prevalent and higher as a share of total potential harvest among less well-off households based on a wealth index. Households in the wealthiest decile are about 4.3 percentage points less likely to record crop losses than those in the least wealthy decile and lose about 6.6 percentage points less of their harvest if they are affected (Extended Data Table 8).

Taken together, these findings emphasize that climate shock impacts arise as a result of the confluence of hazard and vulnerability[31,32].

Underestimation of shock impacts in aggregate data sources

How do climatic shock impacts as captured in the survey data compare to estimates from other commonly used data sources? Here, we contrast the results from the survey microdata with publicly available estimates of disaster impacts from the EM-DAT. The EM-DAT aggregates reports from UN agencies, governments, insurance companies, research institutes and the media into a global inventory of disaster impacts[33]. EM-DAT is, to our knowledge, the pre-eminent and only publicly available data source of this kind, used widely in reporting and research on the impacts of climate shocks and disasters[34]. We focus on two types of climatic shocks, droughts and floods, and compare two estimates: the number of people affected and the total economic damages caused in the years that the microdata cover. We create aggregate figures from the microdata using population sampling weights, counting as affected any household with a production loss.

We first examine drought and flood shocks in our study countries that were recorded as 'disasters' in EM-DAT. For this, shocks need to meet a minimum set of criteria for inclusion in the database; that is, at least 10 deaths or at least 100 affected (people affected, injured or homeless) or an emergency declaration or a call for international assistance[33]. This is the case for six drought and six flood events across Ethiopia, Malawi, Niger, Nigeria and Tanzania (Fig. 3). The microdata estimates of people affected exceed the EM-DAT estimates in all but three cases (Supplementary Tables 9 and 10). Estimates of the economic value of damages of the recorded disasters are missing in the EM-DAT data in nine cases, while the microdata document substantial losses. In one case, the Malawi 2015–2016 floods, there are no damage estimates in the microdata, while EM-DAT records damages of US$595 million. For events with damage estimates in both sources, the 2015–2016 drought in Ethiopia and the 2014 flood in Niger, the microdata estimates exceed the EM-DAT estimates (Supplementary Tables 11 and 12).

Moreover, there are many years in which the microdata document damages due to droughts and floods but EM-DAT records no impacts at all. For example, drought shocks are prevalent to some degree across every country–year combination covered in the microdata, while EM-DAT records droughts affecting the population in only a third of country–year combinations. The events that go unreported in EM-DAT are smaller, on average, in terms of the population affected and the damages caused. As such, these events may not be severe enough to be considered 'disasters' and may not meet the minimum requirements for inclusion in EM-DAT. However, we show that such smaller, under-covered events have substantial impacts on the livelihoods of farmers and the economies of the study countries. For example, droughts in Malawi in 2009–2010 and Mali in 2014 affected the production and incomes of more than a fifth of the respective populations. The value of damages during droughts in Niger in 2011 and Ethiopia in 2018–2019 amounted to US$1.6 billion and US$1.4 billion, respectively. The total number of people affected by droughts or floods in all years covered by the microdata is between 145 and 170 million, more than four times higher than what is recorded in EM-DAT, while the microdata estimates of drought and flood damages exceed the EM-DAT data by US$5.1 billion (Extended Data Table 9).

We recompute the microdata estimates applying various different inclusion thresholds in terms of the severity of a drought or flood shock (at least 25%, 50%, 75% of potential harvest lost; at least 25%, 50%, 75% of plots affected per region; Supplementary Tables 13 and 14). We find that even with these higher thresholds for inclusion, the microdata estimates still significantly exceed the EM-DAT data.

What explains these discrepancies? The first reason is that the microdata capture some climate shocks that probably do not meet the criteria for inclusion in a disaster inventory such as EM-DAT but which are still causing substantial damages. But disaster inventories such as EM-DAT and survey microdata differ in other meaningful ways. Disaster inventories do not measure shock impacts themselves but instead aggregate data from government sources, humanitarian organizations, the media and others. They therefore rely on the comprehensiveness and accuracy with which climatic shocks are covered by one or more of these sources[14,34]. Less salient events, as well as those affecting marginalized population groups, are less likely to be reported on and less likely to have detailed information on the affected population or economic and welfare impacts[14,35,36]. This is particularly acute in the context of low and lower-middle income countries (LMICs), such as our study countries, which are more likely to have incomplete coverage or inaccurate information in disaster inventories[14,34,37,38].

By comparison, the survey microdata capture shock impacts by directly asking farmers. But the microdata suffer from some similar drawbacks and limitations. Whether and how well climatic shock impacts are captured depends on the survey design. For example, there are cases in which we only capture the population affected but limitations in survey design impedes calculation of the value of damages. Microdata in LMICs rarely have annual coverage and often lack full comparability between countries. Finally, survey data rely on respondent recall of climatic shock impacts. Human recall and reporting have been shown to suffer from cognitive biases and be susceptible to respondents' incentives, misreporting and misperceptions, which could cause classical and non-classical measurement error in survey estimates[39–42]. We discuss the implications of these differences between microdata and aggregate sources for the internal and external validity of climate shock impact estimates in detail in Supplementary Text A. The discussion is summarized in Extended Data Table 10.

Discussion

We explore the crop production impacts of climatic shocks on 120,000 fields on 30,000 smallholder farms in Sub-Saharan Africa. Smallholder agriculture is of special interest for achieving sustainable development goals SDG 1 and 2 as it remains the primary means of livelihood for many of the world's poor[16].

Our findings advance our understanding of the natural hazards and crop losses that smallholder farmers suffer. They relate to research examining the phenomenon of 'small disasters'. Case studies from Colombia, Mali and Senegal found that smaller events not included in EM-DAT caused considerable damage, on par with larger events[8,42]. Another set of studies have used survey microdata to investigate the vulnerability specifically of smallholder farmers to climatic shocks[43–45] but without quantifying production losses. Other studies have relied on macrodata from global disaster inventories to assess and quantify the impact of disasters on agriculture[12,46]. Here, we offer systematic, cross-country evidence based on survey microdata, which allows us to value crop production losses due to climatic shocks, to link them to individual farmers and to assess their economic importance more broadly. The analysis shows that disaster-related crop production losses among African smallholder farmers are frequent and important both to individual farms and for the entire agriculture sectors and economies. We further show that the EM-DAT disaster inventory misses out on a meaningful share of disaster impacts in the agricultural sector in Africa when compared to the microdata analysis, which is mostly due to smaller events not included in the database.

The findings have implications for several global policy debates around climate change and disaster risk resilience. The evidence and insights we present on damages associated with climatic shocks and affecting some of the world's poorest populations are relevant to the discussions and advocacy around the loss-and-damage fund that was agreed at COP28. Key flashpoints relate to how to allocate funds to the most vulnerable and how to measure the losses they suffered[7,47,48]. As our analysis and comparison with macrodata show, loss measurement critically depends both on the data used and the availability of data in the first place—and data gaps are most acute in vulnerable countries and among vulnerable populations[7]. Granular microdata allow capturing heterogeneous losses within countries and across different populations but more work needs to be done to extend the analysis to capture non-economic losses in well-being[21] or harm to human rights related to climate change[48]. Another issue is the attribution of climatic shocks and their impacts on climate change to calculate climate change-related loss and damage[5,47]. However, attribution exercises in the most vulnerable countries including those that we study are hampered by limited data[49]. Our analysis further suggests that a substantial portion of damages comes from many frequent but smaller climatic shocks, which will probably render attribution exercises even more complicated.

Our findings also contribute to the policy discourse surrounding disaster risk reduction, with implications for understanding disaster risk (Sendai priority 1) and strengthening disaster risk governance (Sendai priority 2). Climatic shocks cause substantial damage to a vulnerable group such as smallholder farmers but these shocks are under-reported in important data sources. This impedes action to support affected groups[31]. Disaster risk governance would benefit from explicitly accounting for the nature and impacts of the kinds of climatic shocks whose prevalence we document. Although our results indicate differing level of resilience, the effects of climatic shocks could be mitigated by proactive social protection schemes or by insuring crop losses[50]. More research is needed on appropriate responses and risk management instruments.

The findings also have implications for data and measurement. Pre-eminent databases on disaster impacts such as EM-DAT have incomplete data in LMICs and do not include smaller events by construction, probably missing impacts in poorer countries and among poorer people. Survey microdata such as the LSMS-ISA have much more limited country and temporal coverage. Combining both sources promises to yield a more complete and nuanced understanding of the issue that will promote more effective policy designs. Improving microdata systems in countries vulnerable to damages and losses from climate change is key to systematically using microdata for monitoring and reporting of emergency events. More flexible and higher frequency data collection is needed to provide better temporal coverage and account for disaster impacts when they occur. Phone surveys more widely adopted in low-income settings during the COVID-19 emergency may provide this function, for instance as part of mixed-mode survey systems that combine traditional in-person surveys with data collection over the phone[51]. Data scarcity also affects other key data sources, including hydrometeorological data[47]. Integration of survey with such geospatial data promises to improve spatial coverage and identification of natural hazards and shock occurrence, facilitating better responses[52]. The integration of survey and geospatial data relies on geolocations of survey households and communities. Too few surveys capture and disseminate geolocations. For privacy reasons, those that do, provide geolocations with some imprecision, which could in some cases hamper analyses. Privacy-conserving mechanisms to allow access to true geolocations should be elaborated.

Our study faces several limitations. The valuation of losses relies on human reporting which is subject to human error, respondents' incentives, misreporting and misperceptions[39,40]. Data quality and availability may also be affected by climate shocks (Supplementary Text A). The data allow for a detailed analysis of damage to crop production but other damages, for example, damages to agricultural assets, storage losses or impacts on livestock, are missing. As a consequence, the full extent of damages and losses to smallholder agriculture is probably underestimated. We also do not discuss damages incurred in other sectors. Future research should aim to quantify the aggregate welfare impact of climate shocks on the basis of credible causal identification.

Methods

Survey microdata

We use plot-level survey data from the LSMS-ISA in Ethiopia, Malawi, Mali, Niger, Nigeria and Tanzania. The LSMS-ISA comprise a series of harmonized, national, multitopic household panel surveys with a focus on agriculture. We created a harmonized dataset on crop losses due to climatic shocks, as well as other relevant information such as agricultural outputs, inputs and plot characteristics for close to 30,000 households over the six countries for a total of 18 survey waves collected between 2008 and 2019.

The combined dataset contains over 120,000 plot observations. More specifically, the dataset includes data from the Ethiopian Social Survey (waves 1 to 4), Malawi's Integrated Household Panel Survey (waves 1 to 4), Mali's Enquête Agricole de Conjoncture Intégrée (waves 1 and 2), Niger's Enquête National sur les Conditions de Vie des Ménages et Agriculture (waves 1 and 2), Nigeria's General Household Survey (wave 4) and Tanzania's National Panel Survey (waves 1 to 5). Waves 1 to 3 of the Nigeria general household survey were not included, as respondents were only asked to report the cause of their losses on plots with full crop failure. While we include a total of 18 rounds in this analysis, we are only able to compute the value of crop damages or the share of harvest lost in nine rounds because of survey design limitations.

Households in these surveys are selected to be representative of the population at the national and subnational level and are sampled using a two-stage stratified sampling design with census enumeration areas as primary sampling units and households as secondary sampling units. Households are then tracked through time, except for Mali which only tracks enumeration areas. Each survey wave covers an agricultural production cycle or season.

Emergency event database

We compare survey estimates to country-level data from the EM-DAT. The EM-DAT is the pre-eminent public database taking stock of shocks on a global scale and is widely used for research and to inform policy[53]. Both natural (for example, geophysical and meteorological) and technological (for example, industrial accidents) events are recorded, along with information on disaster damages valued in 2022 US$. The EM-DAT compiles information from a broad range of sources including insurance companies, international organizations, press agencies and governmental agencies. Disasters are recorded if they provoke ten or more deaths, affect 100 or more people (injured/ homeless/ in need of immediate assistance) or are accompanied by an official declaration of emergency or appeal for international assistance[33].

Variable construction

Identification of crop losses due to climatic shocks. Identification of disaster events is based on farmers reporting crop production losses before harvest for each crop on each of their plots. Specifically, for each crop cultivated on each plot, farmers are asked whether the area harvested was less than the area planted, that is if some of their crop has been lost, along with the cause of the loss in harvested area. In Ethiopia, farmers are further asked whether the crops they harvested had any damage on them and what the cause of damage was. We define crop losses due to climatic shocks as any loss in crop area or any damage on the crops harvested due to climatological (drought, irregular rain, hail and wildfire), hydrological (flood) or biological (insect infestation and disease) reasons. We denote any plot for which at least one crop on the plot had crop losses for one of these reasons as having had climate shock-related crop losses. Notably, this does not include losses due conflict, unavailability of inputs or other household or socio-economic events. Identifying climate shock exposure based on information from the shocks module (part of the LSMS-ISA household questionnaire) and imposing minimum disaster impact thresholds paints a similar picture of the number of affected households (Supplementary Table 15).

Loss size as percentage of potential harvest lost due to climatic shocks. To quantify the impact of climatic shocks on production, we calculate the share (as percentage) of potential harvest lost due to climatic shocks. To determine the potential harvest that could have been achieved in the absence of losses due to climatic shocks, we follow a methodology proposed by the Food and Agriculture Organization[41,54]. We rely on farmers' reports of the harvest quantities of different plots, the share of the planted area lost due to climatic shocks and, where available, the percentage of damage on crops that were harvested.

Equation (1) formalizes the construction of the plot-level loss aggregate following this methodology.

$$\begin{aligned} L_{i,s} &= \sum_{j=1}^{N} p_j \times \left(Y^{\mathrm{p}}_{i,j,s} - Y^{\mathrm{r}}_{i,j,s}\right) \\ &= \sum_{j=1}^{N} p_j \times \left(\left(Y^{\mathrm{r}}_{i,j,s} \times \frac{1}{1-l_{i,j,s}} \times \frac{1}{1-d_{i,j,s}}\right) - Y^{\mathrm{r}}_{i,j,s}\right) \end{aligned} \quad (1)$$

where the harvest loss on plot I in agricultural season s is equal to the difference between the potential harvest in the absence of disasters, Y^{p}, and the realized harvest reported by the farmer, Y^{r}, in kg. The realized harvest for crop j is based on farmers' reports on how much of crop j they harvested in season s on plot i. The potential harvest of crop j is calculated by scaling up the realized harvest in proportion to the share of the planted area of crop j lost to climatic shocks, $l_{i,j,s}$, and the percentage of damage on crop j, $d_{i,j,s}$.

To aggregate across different crops grown on the same plot, we use a set of price weights p_j which are constant for each crop in each country. Specifically, each price weight corresponds to the median crop sale price in kg calculated in one survey round in each country, which is converted and adjusted to 2022 US$ using exchange rates and a consumer price index drawn from a library of World Bank Development Indicators[55]. This allows us to express the harvest loss $L_{i,s}$ in a common unit across crops and plots. Losses $L_{i,s}$ are then winsorized at the 99th percentile to correct for outliers. We also value and aggregate realized harvest using the same set of prices, such that $y^{r}_{i,s} = \sum_{j=1}^{N} p_j \times Y^{r}_{i,j,s}$. Realized harvest is winsorized at the 1st and 99th percentiles, while allowing full losses (full crop failure) to be equal to 0.

To derive the plot-level percentage of potential harvest lost due to climatic shocks, we simply compute the ratio between losses and potential harvest:

$$\delta_{i,s} = \frac{L_{i,s}}{y^{r}_{i,s} + L_{i,s}} \tag{2}$$

Valuation of crop losses. To estimate the value of losses perceived by farmers, as well the value of potential harvest reported in the analysis, we use a 'current' set of prices that is closer to those faced by farmers in each agricultural season. To do so, we calculate realized harvest using cluster- and round-specific median farmer-reported sale prices. Plot-level realized harvest in present values can thus be written as $y^{o}_{i,s} = \sum_{j=1}^{N} p_{j,s,c} \times Y^{r}_{i,j,s}$, where clusters are denoted by c. This harvest estimate is then winsorized following the same procedure as described for $y^{r}_{i,s}$ above. Assuming that $y^{op}_{i,s}$ denotes potential present harvest value, we can state that:

$$y^{o}_{i,s} = y^{op}_{i,s}\,(1 - \delta_{i,s}) \tag{3}$$

Losses faced by farmers are therefore calculated as:

$$L^{o}_{i,s} = \frac{y^{o}_{i,s}}{1 - \delta_{i,s}} - y^{o}_{i,s} \tag{4}$$

Both losses and harvest values are then converted to US$ using exchange rates drawn from a library of World Bank Development Indicators[55].

Imputation of full losses. In case the harvest for a crop on a plot is fully lost ($l_{i,j,s} = 1$ or $d_{i,j,s} = 1$), equation (1) is not defined. Instead, we estimate the quantity lost in these cases by imputing potential harvest values using a Gaussian normal regression imputation method[56]. To this end, we define a model in which potential harvest is the outcome variable, regressed on the set of explanatory variables, along with country and crop fixed effects. The explanatory variables used in the imputation are the following: (1) agricultural input variables, specifically, plot area, non-hired labour days spent working on the plot (for example, family labour), as well as hired labour value, inorganic fertilizer value and seed value. Inputs are valued in a similar fashion to the production values described above, a constant set of prices was computed within each country, based on median purchase prices. These input variables are all expressed in per hectare terms, winsorized and logged; (2) an agricultural asset index was computed using a principal component analysis based on an inventory of household assets; (3) plot-level dummy variables were included to indicate if a plot is irrigated, pesticides are used, organic fertilizers are applied, the plot is intercropped and if the plot is owned by the household; (4) gender of the primary decision-makers on each plot; (5) household-level variables, household size and dummies for livestock ownership, electricity access and urban/rural residence; and (6) a set of geophysical variables consisting of plot elevation, a topographic wetness index and the distance of the household from the closest population centre and closest road.

Our final imputed value is obtained by calculating the mean of 100 imputations.

Estimation

We rely on different estimation methods in each part of the analysis as described next.

'Prevalence of crop losses among African farmers'. Our main descriptive analysis of climatic shock prevalence and intensity is conducted at the plot level and involves the estimation of means, proportions and frequencies at the national level as well as pooled across countries. These estimates, as well as any household-level estimates of disaster exposure formed by aggregating across plots belonging to the same farm, are weighted using the probability weights described in the section on 'Sampling weights' that follows.

To estimate the average effect of climate shocks on total annual household consumption (in 2020 US$) we match household farms that incurred a climate shock to households of similar wealth, farm size and household size that did not incur a shock. Farm households are matched to the five nearest neighbours within the same survey wave on the basis of Mahalanobis distance and the following variables: an agricultural asset index, a household asset index, whether the household engages in livestock farming, whether the household has access to electricity, household size and total farm size (in ha). Estimates are adjusted to correct for large-sample bias[57].

'Crop production impacted by multiple shocks'. Similarly, our analysis of the prevalence of different shock types is conducted at the plot level and involves the estimation of frequencies at the national level and pooled across countries using the household sampling weights. Our estimates of the most common shock type within enumeration areas are based on simple, unweighted frequencies.

'Local crop losses and farmer characteristics'. Our multivariate analysis focuses on two main outcome variables: a binary variable indicating any disaster crop loss and a continuous variable denoting the percentage share of the total potential harvest that was lost to disasters. We estimate all models with the binary crop loss indicator as outcome variable via maximum likelihood using logistic regression. Models with the percentage share of harvest lost as outcome variable are estimated by means of ordinary-least-squares regression. Our independent variables for these multivariate regressions are comprised of plot characteristics (plot size, elevation, a topographical wetness index, an indicator for ownership of the plot and main crop fixed effects), as well as plot management (hired labour and fertilizer input use, irrigation and intercropping), plot manager (age, gender and education) and household characteristics (urban/rural residence, an indicator for livestock farming and electricity access). Models pooling the sample across countries further include country fixed effects. We also conduct multivariate analysis using a binary variable capturing the gender of the plot manager as outcome variable and plot characteristics and plot-management characteristics as independent variables. As before, all multivariate regressions are weighted using the sampling weights.

To study the factors that drive the gender gap in climate shock exposure, we conduct a mediation analysis of the effect of plot manager gender on extensive and intensive margin crop losses[29]. Effects are decomposed into the direct effect of plot manager gender and the mediated (indirect) effect through the key mediators of interest. We express the indirect (mediated) effect as a proportion of the total effect. In the outcome model, we control for plot manager characteristics (age and education), inputs (any hired labour, any inorganic fertilizer and any organic fertilizer), plot characteristics and production technology (log plot size, plot owned by manager, elevation, slope, soil fertility, potential wetness index, irrigation and intercropping), as well as crop and country fixed effects. For the respective mediator of interest, the models do not include that mediator among the controls. The mediation model controls for the same variables.

Our analysis of differences in drought impact within the same enumeration areas first determines whether some, but not all, plots belonging to the same enumeration area recorded a drought loss and then estimates the simple, unweighted proportion of enumeration areas for which this is the case. Further, we calculate the share of plots affected by drought or floods, respectively, within an enumeration area conditional on at least one plot recording a drought or flood. We exclude enumeration areas with less than ten plots and calculate the average across enumeration areas for each survey round and overall. Our analysis of within-household differences in shock impacts first limits the sample to households with several maize plots and where at least one of the household's maize plots recorded a drought shock. We then calculate the share of remaining maize plots belonging to the same household that also record a drought loss and report a simple, unweighted average of this share across households for each country and year.

'Underestimation of shock impacts in aggregate data sources'. Our estimates of the number of people affected by disasters and aggregate economic losses, are totals at the national level and use the household-level sampling weights.

To compare drought and flood impacts using LSMS-ISA data with those using EM-DAT data, we use two metrics: the share of individuals 'affected' by the shock and the estimated total value of damages. Since the LSMS-ISA surveys run every 2 to 3 years, we only retain events in the EM-DAT database for which the start or end date is within a year containing LSMS-ISA data. The comparison for flood shocks is possible in fewer countries and years because of the limitations in scope of the survey questionnaire in some cases (Supplementary Table 16).

To compute the total number of individuals impacted by a shock within a specified period in the EM-DAT database, we aggregate the total number of people 'affected' by the shock in the macrodata. Affected people are those who are reportedly injured, homeless or otherwise in need of 'immediate assistance'. To estimate the total number of individuals affected by a shock in the LSMS-ISA microdata, we construct population weights by multiplying household weights by household size. These weights are then used as expansion factors, which we multiply by a dummy variable equal to one in the household reporting a shock on any of its cultivated plots. We then add up this product to calculate an expansion estimator[58,59].

To obtain shares of the total population, the numbers of individuals affected in both EM-DAT and the LSMS-ISA data are divided by total yearly population estimates drawn from a library of World Bank Development Indicators[55].

We then compute the estimated damages from both droughts and floods in the periods and years covered by LSMS-ISA data. We first aggregate the 'total damages' estimated in the EM-DAT database, defined as the values of total losses 'directly or indirectly related to the disaster', in 2022 US$ values. Using LSMS-ISA microdata, we aggregate the estimated value of crop losses for each household. In this case, the value of losses is calculated by multiplying the potential value of total plot output using present prices ($y_{i,s}^{op}$) by the estimated percentage of output lost at the plot level ($\delta_{i,s}$). The loss values are then converted to 2022 US$ values, to allow comparison with EM-DAT data. As above, we use population weights as expansion factors, which we multiply with our loss value estimates[58,59].

Sampling weights

In the survey data, household sampling weights are used to compute estimates that are representative of the national or subnational population. These reflect the inverse probability of selection into the sample, are adjusted to account for non-response and survey design choices and are post-stratified to ensure that they sum to known household population totals[60].

Moreover, a further adjustment was made in estimations using plot-level data. In those cases, we divide household weights by the number of plots in the household.

The resulting weights were used to compute estimates in our analysis. Wherever population means are estimated, the standard errors provided with the estimate take into account the clustered and stratified sampling design[58]. All regression models with continuous variables as outcomes are estimated using ordinary-least-squares regression, while regressions with binary outcomes are estimated by means of maximum likelihood using a logistic regression model.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw data used in this study as well as the full questionnaires are publicly available through the World Bank online resources. LSMS-ISA datasets, questionnaires and documentation are publicly accessible through the following link: https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA. Links to the various datasets used in this study are provided in the 'dissemination' tabs. More specifically, the dataset includes data from the Ethiopian Social Survey (waves 1 to 4), Malawi's Integrated Household Panel Survey (waves 1 to 4), Mali's Enquête Agricole de Conjoncture Intégrée (waves 1 and 2), Niger's Enquête National sur les Conditions de Vie des Ménages et Agriculture (waves 1 and 2), Nigeria's General Household Survey (wave 4) and Tanzania's National Panel Survey (waves 1 to 5). The EM-DAT data were downloaded for free on the following public website: https://public.emdat.be/. Data were downloaded on 29 November 2023. The following data filters were applied—Classification: Natural; Countries: Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania; Time range: 2008–2020. The analysis dataset is available from Zenodo at https://doi.org/10.5281/zenodo.12667754 (ref. 61). Shapefiles and other raw geodata required to produce Fig. 2 were downloaded from the GADM database, which can be accessed for free at https://gadm.org/download_country.html. Population and other aggregate statistics were downloaded from the World Bank Development Indicators Database, directly into Stata using the wbopendata package. Source data are provided with this paper.

Code availability

The code for the analysis is available from Zenodo at https://doi.org/10.5281/zenodo.12667754 (ref. 61).

References

1. Harvey, F. Human-driven climate crisis fuelling Horn of Africa drought—study. *The Guardian*www.theguardian.com/environment/2023/apr/27/human-driven-climate-crisis-fuelling-horn-of-africa-drought-study (2023).
2. Janjua, H. A year on, the devastating long-term effects of Pakistan's floods are revealed. *The Guardian* www.theguardian.com/global-development/2023/aug/05/a-year-on-the-devastating-long-term-effects-of-pakistans-floods-are-revealed (2023).
3. Matonga, G. & Chutel, L. Cyclone Freddy brings mudslides and floods, leaving nearly 200 dead in Malawi. *The New York Times* www.nytimes.com/2023/03/14/world/africa/cyclone-freddy-malawi-mudslides-flooding.html (2023).
4. Otto, F. E. L. et al. Climate change increased extreme monsoon rainfall, flooding highly vulnerable communities in Pakistan. *Environ. Res. Clim.* **2**, 025001 (2023).
5. Clarke, B., Otto, F., Stuart-Smith, R. & Harrington, L. Extreme weather impacts of climate change: an attribution perspective. *Environ. Res. Clim.* **1**, 012001 (2022).
6. Newman, R. & Noy, I. The global costs of extreme weather that are attributable to climate change. *Nat. Commun.* **14**, 6103 (2023).
7. Vaidyanathan, G. A giant fund for climate disasters will soon open. Who should be paid first? *Nature* www.nature.com/articles/d41586-024-00149-x (2024).
8. Brennan, M. E. & Danielak, S. Too small to count? The cumulative impacts and policy implications of small disasters in the Sahel. *Int. J. Disaster Risk Reduct.* **68**, 102687 (2022).

9. Walecki, N. Tiny climate crises are adding up to one big disaster. *The Atlantic* www.theatlantic.com/science/archive/2023/11/climate-disasters-low-intensity/675864/ (2023).
10. *Report of the Open-ended Intergovernmental Expert Working Group on Indicators and Terminology relating to Disaster Risk Reduction* (UN, 2016).
11. Ritchie, H. Opinion | 3 charts show how better buildings save lives in earthquakes. *Washington Post* www.washingtonpost.com/opinions/2023/02/13/earthquake-deaths-chile-japan-turkey-compare/ (2023).
12. Lesk, C., Rowhani, P. & Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* **529**, 84–87 (2016).
13. Pirlea, A. F., Serajuddin, U., Wadhwa, D. & Walsh, M. (eds) *Atlas of Sustainable Development Goals 2023* (World Bank, 2023).
14. Gall, M., Borden, K. A. & Cutter, S. L. When do losses count? Six fallacies of natural hazards loss data. *Bull. Am. Meteor. Soc.* **90**, 799–810 (2009).
15. *World Bank Open Data* (World Bank, accessed 12 January 2024).
16. *Poverty and Shared Prosperity 2022: Correcting Course* (World Bank, 2022); https://doi.org/10.1596/978-1-4648-1893-6
17. Park, J., Bangalore, M., Hallegatte, S. & Sandhoefner, E. Households and heat stress: estimating the distributional consequences of climate change. *Environ. Dev. Econ.* **23**, 349–368 (2018).
18. Tol, R. S. J. The economic impacts of climate change. *Rev. Environ. Econ. Policy* **12**, 4–25 (2018).
19. Arslan, A., Floress, K., Lamanna, C., Lipper, L. & Rosenstock, T. S. A meta-analysis of the adoption of agricultural technology in Sub-Saharan Africa. *PLoS Sustain. Transform.* **1**, e0000018 (2022).
20. Hallegatte, S., Vogt-Schilb, A., Rozenberg, J., Bangalore, M. & Beaudet, C. From poverty to disaster and back: a review of the literature. *Econ. Disaster Clim. Change* **4**, 223–247 (2020).
21. Markhvida, M., Walsh, B., Hallegatte, S. & Baker, J. Quantification of disaster impacts through household well-being losses. *Nat. Sustain.***3**, 538–547 (2020).
22. *2023 Disasters in Numbers. A Significant Year of Disaster Impact* (CRED, 2024); https://files.emdat.be/reports/2023_EMDAT_report.pdf
23. Song, Z., Hochman, G. & Timilsina, G. R. *Natural Disaster, Infrastructure, and Income Distribution: Empirical Evidence from Global Data* (World Bank, 2023); https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099312406272326097/IDU018bdbef90c42104d730a46700d5b8c4c1c83
24. Hallegatte, S., Vogt-Schilb, A., Bangalore, M. & Rozenberg, J. *Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters* (World Bank, 2016).
25. Philippon, N., Mougin, E., Jarlan, L. & Frison, P.-L. Analysis of the linkages between rainfall and land surface conditions in the West African monsoon through CMAP, ERS-WSC, and NOAA-AVHRR data. *J. Geophys. Res. Atmos.* https://doi.org/10.1029/2005JD006394 (2005).
26. *The Status of Women in Agrifood Systems* (FAO, 2023); https://doi.org/10.4060/cc5343en
27. Aguilar, A., Carranza, E., Goldstein, M., Kilic, T. & Oseni, G. Decomposition of gender differentials in agricultural productivity in Ethiopia. *Agric. Econ.* **46**, 311–334 (2015).
28. Singbo, A., Njuguna-Mungai, E., Yila, J. O., Sissoko, K. & Tabo, R. Examining the gender productivity gap among farm households in Mali. *J. Afr. Econ.* **30**, 251–284 (2021).
29. Imai, K., Keele, L. & Tingley, D. A general approach to causal mediation analysis. *Psychol. Methods* **15**, 309–334 (2010).
30. Ali, D., Bowen, D., Deininger, K. & Duponchel, M. Investigating the gender gap in agricultural productivity: evidence from Uganda. *World Dev.* **87**, 152–170 (2016).
31. Kelman, I., Gaillard, J. C., Lewis, J. & Mercer, J. Learning from the history of disaster vulnerability and resilience research and practice for climate change. *Nat. Hazards* **82**, 129–143 (2016).
32. Oliver-Smith, A. in *The Angry Earth* (eds Oliver-Smith, A. et al.) Ch. 4 (Routledge, 1999).
33. *EM-DAT Documentation* (EM-DATA, accessed 4 April 2024); https://doc.emdat.be/docs/
34. Jones, R. L., Kharb, A. & Tubeuf, S. The untold story of missing data in disaster research: a systematic review of the empirical literature utilising the Emergency Events Database (EM-DAT). *Environ. Res. Lett.* **18**, 103006 (2023).
35. Panwar, V. & Sen, S. Disaster damage records of EM-DAT and DesInventar: a systematic comparison. *Econ. Disaster Clim. Change* **4**, 295–317 (2020).
36. Wirtz, A., Kron, W., Löw, P. & Steuer, M. The need for data: natural disasters and the challenges of database management. *Nat. Hazards* **70**, 135–157 (2014).
37. Osuteye, E., Johnson, C. & Brown, D. The data gap: an analysis of data availability on disaster losses in sub-Saharan African cities. *Int. J. Disaster Risk Reduct.* **26**, 24–33 (2017).
38. Harrington, L. J. & Otto, F. E. L. Reconciling theory with the reality of African heatwaves. *Nat. Clim. Change* **10**, 796–798 (2020).
39. Beegle, K., Carletto, C. & Himelein, K. Reliability of recall in agricultural data. *J. Dev. Econ.* **98**, 34–41 (2012).
40. De Weerdt, J., Gibson, J. & Beegle, K. What can we learn from experimenting with survey methods? *Annu. Rev. Resour. Econ.* **12**, 431–447 (2020).
41. Markhof, Y. V., Ponzini, G. & Wollburg, P. R. *Measuring Disaster Crop Production Losses Using Survey Microdata: Evidence from Sub-Saharan Africa* (World Bank, 2022); http://documents.worldbank.org/curated/en/324181647280329139/Measuring-Disaster-Crop-Production-Losses-Using-Survey-Microdata-Evidence-from-Sub-Saharan-Africa
42. Marulanda, M. C., Cardona, O. D. & Barbat, A. H. Revealing the socioeconomic impact of small disasters in Colombia using the DesInventar database. *Disasters* **34**, 552–570 (2010).
43. Derbile, E. K., Chirawurah, D. & Naab, F. X. Vulnerability of smallholder agriculture to environmental change in North-Western Ghana and implications for development planning. *Clim. Dev.* **14**, 39–51 (2022).
44. Chandra, A., McNamara, K. E., Dargusch, P., Caspe, A. M. & Dalabajan, D. Gendered vulnerabilities of smallholder farmers to climate change in conflict-prone areas: a case study from Mindanao, Philippines. *J. Rural Stud.* **50**, 45–59 (2017).
45. Harvey, C. A. et al. Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar. *Philos. Trans. R. Soc. B* **369**, 20130089 (2014).
46. Coulibaly, T., Islam, M. & Managi, S. The impacts of climate change and natural disasters on agriculture in African countries. *Econ. Disaster Clim. Change* **4**, 347–364 (2020).
47. Clarke, R. H. et al. Climate loss-and-damage funding: a mechanism to make it work. *Nature* **623**, 689–692 (2023).
48. McNamara, K. E. et al. Using a human rights lens to understand and address loss and damage. *Nat. Clim. Change* **13**, 1334–1339 (2023).
49. King, A. D., Grose, M. R., Kimutai, J., Pinto, I. & Harrington, L. J. Event attribution is not ready for a major role in loss and damage. *Nat. Clim. Change* **13**, 415–417 (2023).
50. Jensen, N. & Barrett, C. Agricultural index insurance for development. *Appl. Econ. Perspect. Policy* **39**, 199–219 (2017).
51. Gourlay, S., Kilic, T., Martuscelli, A., Wollburg, P. & Zezza, A. Viewpoint: high-frequency phone surveys on COVID-19: good practices, open questions. *Food Policy* **105**, 102153 (2021).
52. Carletto, G. & Banerjee, R. in *Capitalism, Global Change and Sustainable Development* (ed. Paganetto, L.) 87–96 (Springer International, 2020).

53. Jones, R. L., Guha-Sapir, D. & Tubeuf, S. Human and economic impacts of natural disasters: can we trust the global data? *Sci. Data* **9**, 572 (2022).
54. Conforti, P., Markova, M. & Tochkov, D. *FAO's Methodology for Damage and Loss Assessment in Agriculture* (FAO, 2020); http://www.fao.org/documents/card/en/c/ca6990en
55. *World Development Indicators* (World Bank, accessed 12 November 2023).
56. Gelman, A. *Bayesian Data Analysis* (CRC Press, 2014).
57. Abadie, A. & Imbens, G. W. Bias-corrected matching estimators for average treatment effects. *J. Bus. Econ. Stat.* **29**, 1–11 (2011).
58. Heeringa, S., West, B. T & Berglund, P. A. *Applied Survey Data Analysis* (Chapman & Hall/CRC, 2010).
59. Haziza, D. & Beaumont, J.-F. Construction of weights in surveys: a review. *Stat. Sci.* **32**, 206–226 (2017).
60. Himelein, K. Weight calculations for panel surveys with subsampling and split-off tracking. *Stat. Public Policy* **1**, 40–45 (2014).
61. Bentze, T., Wollburg, P. & Markhof, Y. Data and code for 'Substantial impacts of climate shocks in African smallholder agriculture'. *Zenodo* https://doi.org/10.5281/zenodo.12667753 (2024).

Acknowledgements

This study received funding support from the 50×2030 Initiative to Close the Agricultural Data Gap (P.W., Y.M., T.B. and G.P.) and the World Bank Research Support Budget grant 'On the Measurement of Agricultural Productivity Trends in Africa' (P.W.). The authors are grateful to D. Gollin, E. Bulte, G. Carletto, R. Hill, S. Hallegatte, T. Kilic, T. Lybbert and participants of the CSAE 2023 Conference, the ICAS IX Conference, the EAAE 2023 Congress and at the World Bank for their comments and feedback.

Author contributions

P.W., Y.M. and G.P. conceived the ideas. P.W., Y.M. and T.B. designed the methodology. T.B. and Y.M. curated and analysed the data. P.W., Y.M. and T.B. wrote the original draft. P.W., G.P., Y.M. and T.B. edited and reviewed the draft.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41893-024-01411-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41893-024-01411-w.

Correspondence and requests for materials should be addressed to Philip Wollburg.

Peer review information *Nature Sustainability* thanks Mook Bangalore, Guy Jackson and Sandy Tubeuf for their contribution to the peer review of this work.

 https://doi.org/10.1038/s41893-024-01411-w

Extended Data Table 1 | Mean fraction of potential harvest lost at the plot level, by country-year

Country	Year	Mean fraction of potential harvest lost, on all plots with climate shock losses	Mean fraction of potential harvest lost, on plots affected by drought	Mean fraction of potential harvest lost, on plots affected by floods	Mean fraction of potential harvest lost, on plots affected by irregular rains	Mean fraction of potential harvest lost, on plots affected by pests
Ethiopia	2013 - 2014	48.66 % [45%;52%] n = 4187	53.45 % [45%;62%] n = 1598	missing	49.36 % [44%;54%] n = 1414	45.22 % [39%;51%] n = 994
	2015 - 2016	54.84 % [51%;59%] n = 7717	59.71 % [55%;65%] n = 5952	missing	55.90 % [50%;62%] n = 2332	47.35 % [40%;55%] n = 401
	2018 - 2019	47.63 % [44%;52%] n = 3161	46.90 % [40%;54%] n = 1511	missing	47.65 % [42%;53%] n = 1065	42.36 % [36%;48%] n = 474
Malawi	2018 - 2019	55.00 % [52%;58%] n = 2180	57.17 % [52%;62%] n = 259	62.65 % [58%;67%] n = 1048	49.79 % [46%;53%] n = 872	49.26 % [44%;55%] n = 364
Mali	2014	48.08 % [45%;52%] n = 1595	48.37 % [45%;52%] n = 1212	missing	38.11 % [28%;48%] n = 144	40.73 % [31%;51%] n = 87
	2017	47.96 % [44%;52%] n = 5794	47.50 % [44%;51%] n = 5411	missing	35.75 % [22%;50%] n = 108	66.73 % [52%;82%] n = 182
Niger	2011	71.09 % [69%;74%] n = 5191	71.79 % [69%;75%] n = 3764	77.53 % [66%;89%] n = 37	missing	66.96 % [63%;71%] n = 1544
	2014	49.33 % [46%;53%] n = 1346	52.90 % [49%;57%] n = 825	56.79 % [44%;70%] n = 103	missing	40.97 % [36%;46%] n = 379
Nigeria	2019	56.38 % [50%;63%] n = 691	36.38 % [28%;45%] n = 115	61.87 % [50%;74%] n = 263	missing	30.71 % [24%;37%] n = 219
All countries	2011 – 2019	53.49 % [52%;55%] n = 31862	56.75 % [54%;60%] n = 20647	62.28 % [56%;68%] n = 1451	51.52 % [48%;55%] n = 5935	46.52 % [44%;49%] n = 4644

Note: Mean shares, 95% confidence intervals (rounded to the nearest integer) and sample sizes are provided for each estimated parameter. Plots with 0 losses are excluded. Sample weights are used to calculate estimates. Results are 'missing' when information was not captured by the survey in question. Some questionnaires (the IHPS, for example) allow respondents to report multiple shock types on a single plot.

 https://doi.org/10.1038/s41893-024-01411-w

Extended Data Table 2 | Total fraction of aggregate potential harvest lost, by country-year

Country	Year	Fraction of total potential harvest lost in shocks
Ethiopia	2013 - 2014	15.14 % [14.53%;15.50%] n = 14829
Ethiopia	2015 - 2016	34.63 % [31.12%;37.09%] n = 14501
Ethiopia	2018 - 2019	31.50 % [16.28%;40.54%] n = 7904
Malawi	2019	46.52 % [36.79%;52.50%] n = 5106
Mali	2014	7.57 % [6.85%;8.10%] n = 8979
Mali	2017	8.22 % [7.61%;8.74%] n = 23799
Niger	2011	81.11 % [77.50%;83.68%] n = 5792
Niger	2014	29.34 % [23.58%;33.09%] n = 4110
Nigeria	2018 - 2019	3.18 % [2.51%;3.67%] n = 7865
Mean of country-waves	2011-2019	28.58 %

Note: Mean shares, 95% confidence interval and sample sizes are provided for each estimated parameter. Observations in country-years during which surveys did not capture partial losses were dropped. Sample weights are used to calculate estimates. Current prices are used to value losses and attainable harvest. The 'pooled' row is a simple average of the point estimates above.

 https://doi.org/10.1038/s41893-024-01411-w

Extended Data Table 3 | Heterogeneity in climate shock exposure within the same clusters

Survey	Drought	N	Flood	N
Ethiopia 2011-12	0.325	84		
Ethiopia 2013-14	0.267	134		
Ethiopia 2015-16	0.547	218		
Ethiopia 2018-19	0.381	107		
Malawi 2009-10	0.384	113		
Malawi 2012-13	0.135	111	0.197	81
Malawi 2015-16	0.391	79	0.044	23
Malawi 2018-19	0.087	67	0.241	88
Mali 2014	0.426	135		
Mali 2017	0.432	365		
Niger 2011	0.658	77	0.048	12
Niger 2014	0.235	65	0.077	29
Nigeria 2019	0.082	77	0.101	111
Tanzania 2008	0.230	96		
Tanzania 2010	0.288	136		
Tanzania 2012	0.230	162		
Tanzania 2014	0.191	17		
Tanzania 2019	0.294	26		
Total	0.346	2069	0.152	344

Note: The table shows the average share of shock-affected plots within the same geographical clusters, separately for drought and flood shocks, if at least one plot in the cluster was affected. Values close to zero indicate that shocks were very idiosyncratic, affecting only a small share of plots in the same cluster whereas values close to one indicate that most plots in the cluster recorded crop losses due to the same climate shock. We drop clusters with fewer than 10 plots and clusters in which no plot recorded the shock in question.

Extended Data Table 4 | Heterogeneity in climate shock exposure by plot characteristics

Dependent variable: Any climate shock loss on plot	(1) Pooled	(2) Ethiopia	(3) Malawi	(4) Mali	(5) Niger	(6) Nigeria	(7) Tanzania
Any hired labor on plot	-0.0151	-0.00713	-0.0514***	-0.000181	-0.0288	-0.00267	-0.0128
	(0.0111)	(0.0242)	(0.0168)	(0.0143)	(0.0204)	(0.0215)	(0.0151)
	0.173	*0.768*	*0.00240*	*0.990*	*0.159*	*0.901*	*0.397*
Any inorganic fertilizer used	-0.0447***	-0.0482**	0.0217	-0.0620***	-0.0532*	-0.0358**	-0.0885***
	(0.0109)	(0.0187)	(0.0162)	(0.0171)	(0.0290)	(0.0153)	(0.0292)
	4.01e-05	*0.0103*	*0.182*	*0.000305*	*0.0684*	*0.0202*	*0.00262*
Any organic fertilizer used	0.0284***	0.0665***	0.0376**	-0.00259	-0.00807	-0.0114	-0.0129
	(0.0105)	(0.0195)	(0.0159)	(0.0160)	(0.0255)	(0.0187)	(0.0266)
	0.00707	*0.000713*	*0.0190*	*0.871*	*0.752*	*0.541*	*0.630*
Plot is irrigated	-0.00336	-0.000139	-0.0650	-0.201***	-0.173***	0.00273	0.0375
	(0.0211)	(0.0333)	(0.0592)	(0.0461)	(0.0540)	(0.0305)	(0.0501)
	0.874	*0.997*	*0.274*	*1.43e-05*	*0.00165*	*0.929*	*0.454*
Plot is intercropped	0.0490***	0.0352	0.146***	-0.0338	-0.0258	-0.0224	0.0419**
	(0.0117)	(0.0242)	(0.0182)	(0.0292)	(0.0365)	(0.0161)	(0.0210)
	2.78e-05	*0.146*	*0*	*0.247*	*0.481*	*0.166*	*0.0467*
Plot is owned	-0.000240	0.0200	-0.00114	-0.0540**	-0.0526	-0.0299	-0.00931
	(0.0121)	(0.0199)	(0.0215)	(0.0246)	(0.0330)	(0.0184)	(0.0234)
	0.984	*0.316*	*0.958*	*0.0287*	*0.113*	*0.105*	*0.690*
Log plot area (ha)	0.00870***	0.00984*	0.0220***	0.000677	-0.0264***	-0.0126*	0.0338***
	(0.00332)	(0.00504)	(0.00570)	(0.00562)	(0.00840)	(0.00677)	(0.00807)
	0.00887	*0.0513*	*0.000152*	*0.904*	*0.00199*	*0.0636*	*3.36e-05*
Plot topographic wetness index	0.00732***	0.00315	0.00499	0.00217	0.00351	0.00132	0.0121***
	(0.00165)	(0.00407)	(0.00345)	(0.00246)	(0.00267)	(0.00161)	(0.00225)
	9.79e-06	*0.439*	*0.150*	*0.378*	*0.191*	*0.414*	*1.14e-07*
Plot elevation (m)	-9.13e-05***	-8.39e-05***	-0.000372***	-0.000299**	-0.000177	-2.21e-05	-6.54e-05***
	(1.60e-05)	(2.64e-05)	(3.41e-05)	(0.000131)	(0.000186)	(4.59e-05)	(2.39e-05)
	1.21e-08	*0.00157*	*0*	*0.0228*	*0.345*	*0.630*	*0.00646*
Urban household	-0.0212	0.00309	-0.116***	-0.0358	0.0449	-0.0213	0.0272
	(0.0206)	(0.0529)	(0.0410)	(0.0485)	(0.0474)	(0.0250)	(0.0291)
	0.304	*0.953*	*0.00529*	*0.460*	*0.345*	*0.393*	*0.350*
HH engaged in livestock farming	0.00548	0.0210	-0.00549	0.0271	-0.145***	-0.0155	0.0397**
	(0.0101)	(0.0211)	(0.0133)	(0.0203)	(0.0466)	(0.0182)	(0.0173)
	0.587	*0.321*	*0.679*	*0.181*	*0.00214*	*0.394*	*0.0219*
HH has access to electricity	0.0711***	0.120***	0.0176	-0.0242	-0.0300	0.00185	0.0198
	(0.0161)	(0.0241)	(0.0283)	(0.0167)	(0.0493)	(0.0197)	(0.0358)
	1.08e-05	*8.96e-07*	*0.535*	*0.147*	*0.543*	*0.925*	*0.580*
Observations	116,773	39,801	17,900	31,257	9,004	6,623	12,187
Crop FE	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	NO	NO	NO	NO	NO	NO

Note: Average marginal effects from multivariate logit regressions. Base category for crop fixed effects is 'other crop'. The estimates are weighted to be nationally representative. 2-sided t-tests to determine if each coefficient is significantly different from zero: *** p<0.01, ** p<0.05, * p<0.1. P values are in italics.

Extended Data Table 5 | Heterogeneity in loss size by plot characteristics

Dependent variable: Percent of harvest lost	(1) Pooled	(2) Ethiopia	(3) Malawi	(4) Mali	(5) Niger	(6) Nigeria
Any hired labor on plot	-5.121***	-3.631	-3.207	-0.777	-4.192***	-9.213*
	(1.764)	(2.235)	(3.235)	(1.967)	(1.461)	(5.107)
	0.00374	*0.105*	*0.324*	*0.693*	*0.00470*	*0.0724*
Any inorganic fertilizer used	-8.082***	-7.749***	-2.178	-15.48***	-4.372**	-17.52***
	(1.508)	(1.881)	(1.891)	(2.564)	(1.974)	(4.426)
	9.73e-08	*4.45e-05*	*0.252*	*2.96e-09*	*0.0283*	*9.70e-05*
Any organic fertilizer used	0.670	3.486*	-3.168**	-4.188**	-6.626***	-9.890**
	(1.314)	(1.844)	(1.500)	(2.106)	(1.180)	(3.813)
	0.610	*0.0593*	*0.0371*	*0.0472*	*9.17e-08*	*0.0100*
Plot is irrigated	3.486	1.095	9.082	-20.05***	4.904	13.26
	(3.079)	(3.651)	(14.02)	(6.362)	(4.218)	(8.228)
	0.258	*0.765*	*0.519*	*0.00172*	*0.247*	*0.108*
Plot is intercropped	-7.527***	-3.779	-1.495	-6.408*	-6.364***	-36.63***
	(1.752)	(2.327)	(2.460)	(3.554)	(1.770)	(4.863)
	1.84e-05	*0.105*	*0.545*	*0.0720*	*0.00044*	*0*
Plot is owned	-1.935	-2.290	6.044*	-2.886	-2.496	-2.970
	(1.557)	(2.019)	(3.117)	(3.705)	(1.642)	(4.409)
	0.214	*0.257*	*0.0552*	*0.436*	*0.131*	*0.501*
Log plot area (ha)	-2.586***	-2.788***	-4.575***	-0.229	-0.595	0.0895
	(0.410)	(0.498)	(1.116)	(0.685)	(0.429)	(1.393)
	3.67e-10	*3.52e-08*	*8.18e-05*	*0.738*	*0.167*	*0.949*
Plot topographic wetness index	0.265	0.577*	0.0898	0.952***	0.0433	0.0349
	(0.206)	(0.301)	(0.435)	(0.335)	(0.250)	(0.436)
	0.199	*0.0558*	*0.837*	*0.00463*	*0.863*	*0.936*
Plot elevation (m)	-0.00728***	-0.00650***	-0.0254***	-0.0179	-0.00679	0.00354
	(0.00213)	(0.00227)	(0.00402)	(0.0186)	(0.0109)	(0.00882)
	0.000665	*0.00443*	*6.93e-09*	*0.339*	*0.536*	*0.689*
Urban household	4.414	12.24***	-3.757	4.937	1.414	-6.727
	(2.785)	(2.838)	(3.384)	(6.567)	(2.665)	(5.651)
	0.113	*1.97e-05*	*0.270*	*0.452*	*0.597*	*0.235*
Household engaged in livestock farming	-2.419	1.268	0.185	2.999	-8.985***	-2.777
	(1.735)	(2.354)	(1.831)	(3.219)	(2.200)	(4.421)
	0.163	*0.590*	*0.920*	*0.352*	*7.16e-05*	*0.530*
Household has access to electricity	2.397	1.592	3.470	-1.246	-2.511	6.016
	(1.948)	(2.220)	(4.478)	(2.334)	(2.214)	(4.297)
	0.219	*0.473*	*0.440*	*0.594*	*0.259*	*0.163*
Constant	64.26***	53.20***	62.65***	56.16***	83.58***	106.1***
	(6.171)	(7.216)	(10.07)	(10.07)	(6.959)	(10.92)
	0	*0*	*1.01e-08*	*3.96e-08*	*0*	*0*
Observations	30,845	14,896	2,059	7,130	6,071	689
Crop FE	YES	YES	YES	YES	YES	YES
Country FE	YES	NO	NO	NO	NO	NO

Note: Results from OLS regressions with the share of potential harvest lost as outcome variable. Base category for crop fixed effects is 'other crop'. The estimates are weighted to be nationally representative. Clustered standard errors at the EA level in parentheses. 2-sided t-tests to determine if each coefficient is significantly different from zero: *** p<0.01, ** p<0.05, * p<0.1. P values are in italics.

 https://doi.org/10.1038/s41893-024-01411-w

Extended Data Table 6 | Heterogeneity in climate shock exposure by plot manager characteristics

	Pooled		Ethiopia		Malawi		Mali		Niger		Nigeria		Tanzania	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Dependent variable: Any climate shock loss on plot	No control	Control	No control	Control	No control	Control	No control	Control	No control	Control	No control	Control	No control	Control
Female plot manager	0.0222**	0.0267***	0.00350	0.0126	0.0529***	0.0272**	0.0843***	0.0614***	0.0503*	0.0475*	0.0386*	0.00467	0.0138	0.0355*
	(0.00939)	(0.00904)	(0.0171)	(0.0169)	(0.0137)	(0.0123)	(0.0237)	(0.0215)	(0.0293)	(0.0280)	(0.0207)	(0.0165)	(0.0201)	(0.0186)
	0.0182	*0.00314*	*0.838*	*0.456*	*0.000148*	*0.0273*	*0.000403*	*0.00481*	*0.0878*	*0.0951*	*0.0624*	*0.745*	*0.492*	*0.0569*
Age of plot manager (decades)	0.00727***	0.00708***	0.00829*	0.00750*	0.00112	0.00407	0.0124***	0.0124***	-0.0137***	-0.00896*	0.00774	0.00564	0.0119**	0.00878*
	(0.00249)	(0.00237)	(0.00457)	(0.00418)	(0.00446)	(0.00416)	(0.00443)	(0.00429)	(0.00525)	(0.00517)	(0.00542)	(0.00533)	(0.00523)	(0.00481)
	0.00354	*0.00285*	*0.0704*	*0.0734*	*0.802*	*0.326*	*0.00538*	*0.00377*	*0.00984*	*0.0886*	*0.154*	*0.280*	*0.0235*	*0.0699*
Plot manager has primary educ	-0.0140	-0.0138	-0.0176	-0.0258	-0.0467***	-0.00906	-0.0386*	-0.0178	-0.0609	-0.0594	0.0186	-0.00118	-0.0216	-0.00847
	(0.0125)	(0.0124)	(0.0300)	(0.0285)	(0.0145)	(0.0140)	(0.0224)	(0.0216)	(0.0480)	(0.0473)	(0.0155)	(0.0179)	(0.0314)	(0.0336)
	0.264	*0.266*	*0.558*	*0.367*	*0.00147*	*0.522*	*0.0845*	*0.402*	*0.206*	*0.213*	*0.229*	*0.986*	*0.493*	*0.806*
Observations	115,041	115,041	39,289	39,289	17,123	17,123	31,019	31,019	8,911	8,911	6,591	6,591	12,107	12,107
Controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Crop FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

Note: Average marginal effects from multivariate logit regressions. Controls include dummy variables for input use (any hired labour, any inorganic fertilizer, any organic fertilizer), plot characteristics (plot area, irrigation, intercropping, plot ownership, elevation, and a topographic wetness index), household characteristics (urban/rural residence, livestock farming, and electricity access) as well as main crop and country fixed effects. The estimates are weighted to be nationally representative. 2-sided t-tests to determine if each coefficient is significantly different from zero: *** p < 0.01, ** p < 0.05, * p < 0.1. P values are in italics.

 https://doi.org/10.1038/s41893-024-01411-w

Extended Data Table 7 | Heterogeneity in loss size by plot manager characteristics

	Pooled		Ethiopia		Malawi		Mali		Niger		Nigeria	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent variable: Percent of harvest lost on plot	No controls	Controls	No controls	Controls	No controls	Controls	No controls	Controls	No controls	Controls	No controls	Controls
Female plot manager	4.405***	2.325**	3.345**	1.970	3.952	2.295	5.167	0.985	5.025***	3.730**	10.76**	-0.120
	(1.158)	(1.138)	(1.448)	(1.387)	(2.383)	(2.134)	(3.452)	(2.894)	(1.742)	(1.749)	(5.463)	(6.270)
	0.000147	*0.0412*	*0.0213*	*0.156*	*0.100*	*0.285*	*0.135*	*0.734*	*0.00450*	*0.0346*	*0.0499*	*0.985*
Age of plot manager (decades)	0.231	0.282	0.373	0.408	-0.995	0.0292	-0.537	-0.417	-0.382	-0.0255	0.974	-0.221
	(0.320)	(0.315)	(0.384)	(0.377)	(0.871)	(0.798)	(0.663)	(0.541)	(0.481)	(0.473)	(1.837)	(1.328)
	0.472	*0.372*	*0.331*	*0.281*	*0.256*	*0.971*	*0.419*	*0.441*	*0.428*	*0.957*	*0.596*	*0.868*
Plot manager has primary educ	-0.0274	-1.658	-1.508	-3.204	0.497	1.950	-4.344	-1.227	-5.307*	-3.517	5.493	-1.366
	(1.831)	(1.682)	(2.362)	(2.192)	(2.893)	(2.228)	(2.950)	(2.777)	(2.759)	(2.516)	(4.983)	(5.250)
	0.988	*0.325*	*0.524*	*0.145*	*0.864*	*0.383*	*0.141*	*0.659*	*0.0563*	*0.164*	*0.271*	*0.795*
Constant	49.31***	62.74***	48.94***	51.44***	55.91***	61.52***	49.98***	57.82***	65.99***	82.37***	46.19***	107.3***
	(1.922)	(6.401)	(2.152)	(7.448)	(4.869)	(10.21)	(3.617)	(10.70)	(2.491)	(7.160)	(10.01)	(15.45)
	0	*0*	*0*	*0*	*0*	*2.80e-08*	*0*	*1.01e-07*	*0*	*0*	*6.24e-06*	*0*
Observations	30,230	30,230	14,680	14,680	1,739	1,739	7,100	7,100	6,027	6,027	684	684
Crop FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

Note: Results from OLS regressions with the share of potential harvest lost as outcome variable. Controls include dummy variables for input use (any hired labour, any inorganic fertilizer, any organic fertilizer), plot characteristics (plot area, irrigation, intercropping, plot ownership, elevation, and a topographic wetness index), household characteristics (urban/rural residence, livestock farming, and electricity access) as well as main crop and country fixed effects. The estimates are weighted to be nationally representative. Standard errors in parentheses. 2-sided t-tests to determine if each coefficient is significantly different from zero: *** p<0.01, ** p<0.05, * p<0.1. P values are in italics.

 https://doi.org/10.1038/s41893-024-01411-w

Extended Data Table 8 | Loss incidence and size by wealth deciles

	(1) Any loss	(2) Any loss	(3) Any loss	(4) Any loss	(5) Share lost	(6) Share lost	(7) Share lost	(8) Share lost
Decile of wealth index	-0.00194	-0.0027	-0.00486***	-0.00434**	-0.471*	-0.684***	-0.799***	-0.660**
	(0.00189)	(0.00169)	(0.00179)	(0.00175)	(0.246)	(0.240)	(0.250)	(0.259)
	0.305	*0.11*	*0.00655*	*0.0132*	*0.0562*	*0.00447*	*0.00143*	*0.0111*
Observations	112,882	112,882	112,500	108,565	30,267	30,267	30,138	29,433
Crop FE	NO	YES	YES	YES	NO	YES	YES	YES
Country-wave FE	NO	YES	YES	YES	NO	YES	YES	YES
Household controls	NO	NO	YES	YES	NO	NO	YES	YES
Input plot and production tech controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: Average marginal effects from multivariate logit regressions (columns 1-4: any crop loss due to climate shock) or OLS regressions (columns 5-8: share of potential harvest lost), pooling across countries and time. Controls include inputs (any hired labour, any inorganic fertilizer, any organic fertilizer), plot and production technology characteristics (irrigation, intercropping, plot ownership, log plot area, potential wetness index, elevation), household characteristics (urban area, livestock farming, electricity access), crop fixed effects and country-wave fixed effects. The estimates are weighted to be nationally representative. Wealth index is calculated by country and year based on a wide range of household assets and dwelling characteristics that may vary from survey to survey based on the information available. Clustered standard errors at the EA level in parentheses. 2-sided t-tests to determine if each coefficient is significantly different from zero: 2-sided t-tests to determine if each coefficient is significantly different from zero: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. P values are in italics.

 https://doi.org/10.1038/s41893-024-01411-w

Extended Data Table 9 | Estimated combined impacts of drought or flood events captured in LSMS-ISA

Country	Year	Number of individuals affected by droughts or floods	Share of total population affected by droughts or floods	Value of damages from droughts and floods	Number of individuals affected by droughts or floods	Share of total population affected by droughts or floods	Value of damages from droughts and floods
		LSMS-ISA	LSMS-ISA	LSMS-ISA	EM-DAT	EM-DAT	EM-DAT
Ethiopia	2011 - 2012	7.4 million [5.4 ; 9.5]	7.9 % [5.7 % ; 10.0 %]	*Not possible to estimate losses*	5.846 million	6.3%	*No damage estimate*
	2013 - 2014	11.4 million [8.8 ; 14.0]	11.4 % [8.8 % ; 14.1 %]	589.8 million [281.1 ; 898.5]	52 thousand	0.1%	3.5 million
	2015 - 2016	34.2 million [28.8 ; 39.6]	32.5 % [27.4 % ; 37.6 %]	2,981.5 million [1,999.3 ; 3,963.7]	10.904 million	10.4%	2,134.4 million[±]
	2018 - 2019	16.3 million [13.1 ; 19.5]	14.3 % [11.5 % ; 17.1 %]	1,355.5 million [227.5 ; 2,483.5]	200 thousand	0.2%	*No damage estimate*
Malawi	2009 - 2010	3.5 million [3.0 ; 4.0]	23.8 % [20.5 % ; 27.1 %]	*Not possible to estimate losses*	38 thousand	0.3%	*No damage estimate*
	2012 - 2013	2.6 million [2.2 ; 3.1]	16.5 % [13.5 % ; 19.5 %]	*Not possible to estimate losses*	2.05 million	13.0%	*No damage estimate*
	2015 - 2016	4.6 million [3.5 ; 5.7]	26.4 % [20.1 % ; 32.8 %]	*Not possible to estimate losses*	7.341 million	42.4%	594.6 million
	2018 - 2019	5.4 million [4.1 ; 6.7]	28.5 % [21.5 % ; 35.4 %]	379.9 million [122.9 ; 636.9]	1.001 million	5.4%	*No damage estimate*
Mali	2014	2.8 million [2.4 ; 3.2]	15.9 % [13.4 % ; 18.4 %]	69.8 million [54.6 ; 85.1]	*No event recorded*	*No event recorded*	*No damage estimate*
	2017	4.7 million [4.2 ; 5.3]	24.5 % [21.7 % ; 27.4 %]	164.9 million [135.6 ; 194.2]	*No event recorded*	*No event recorded*	*No damage estimate*
Niger	2011	9.8 million [8.5 ; 11.1]	56.7 % [49.1 % ; 64.4 %]	1,589.2 million [1,297.7 ; 1,880.6]	3.041 million	17.4%	*No damage estimate*
	2014	4.7 million [3.6 ; 5.8]	24.1 % [18.4 % ; 29.8 %]	265.5 million [129.4 ; 401.6]	166 thousand	0.9%	3.1 million
Nigeria	2019	13.6 million [10.5 ; 16.7]	6.8 % [5.3 % ; 8.4 %]	427.9 million [276.2 ; 579.6]	71 thousand	< 0.1 %	*No damage estimate*
Tanzania	2008	5.3 million [4.5 ; 6.1]	12.4 % [10.6 % ; 14.2 %]	*Not possible to estimate losses*	10 thousand	< 0.1 %	*No damage estimate*
	2010	8.6 million [7.5 ; 9.7]	19.0 % [16.5 % ; 21.4 %]	*Not possible to estimate losses*	50 thousand	0.1%	*No damage estimate*
	2012	7.3 million [6.3 ; 8.4]	15.4 % [13.1 % ; 17.7 %]	*Not possible to estimate losses*	1.0 million	2.1%	*No damage estimate*
	2014	5.2 million [2.1 ; 8.4]	10.3 % [4.1 % ; 16.5 %]	*Not possible to estimate losses*	40 thousand	0.1%	3.1 million
	2019	9.9 million [6.2 ; 13.6]	16.5 % [10.4 % ; 22.6 %]	*Not possible to estimate losses*	5 thousand	0.0%	*No damage estimate*
All countries	2008 - 2019	157.0 million [144.6 ; 170.3]	-	7,823.9 million [6,168.5 ; 9,479.4]	31.814 million	-	2,738.6 million

Missing entries correspond to cases where either no event was reported or information on the population affected was missing in the EM-DAT. 95% confidence intervals for estimated totals are calculated with LSMS-ISA survey micro-data. The statistics above are drawn from a sample of 27,375 households with valid flood or shock information in columns 1 and 2, and 21,187 household in column 3. [±] Damage estimated from some of the floods and/or droughts listed in the time period were missing.

Extended Data Table 10 | Comparison of micro-data with disaster inventories

Dimension	Disaster Inventories	Micro-data
Coverage		
Shock coverage (Threshold and hazard bias)	• Subject to reporting in aggregate data sources and reports • Salient shocks likely covered but data less sensitive to idiosyncratic or small shocks • Coverage varies by shock type	• Based on "grassroots" reports elicited from those affected by shocks • Granular and able to cover small and localized shocks • Shock recording subject to questionnaire design (list of shocks, number of shocks recorded)
Population coverage (Population coverage biases)	• Coverage depending on comprehensiveness of coverage in underlying data sources (e.g. news reports) but not limited to a specific population of interest • under-coverage of poor and marginalized population groups within countries likely	• Limited to (stratified) sample of target population • Potential for gaps in coverage wherever shock impacts not well represented by sample • Poor and marginalized population groups explicitly covered
Temporal coverage (Temporal coverage biases)	• Continuous coverage but subject to improvements in quality of reports in the long run	• Intermittent coverage limited to years in which survey was conducted and/or recall period of survey questions
Detail and accuracy		
Detail of available information (Missing data biases)	• Dependent on information reported in sources, limited detail and frequently missing information in some dimensions of shock impacts (e.g. economic and welfare impacts) • Detail and completeness of information related to size and salience of shock • No or limited ability to disaggregate information • Incomplete recording of uninsured damages • Lack of information to quantify impact on (asset-)poor but highly vulnerable population groups	• High level of detail and completeness of data collected, even for small and idiosyncratic shocks as well as poor or vulnerable population groups • Information collected at highly disaggregated level but can be aggregated • Available information subject to survey design
Accuracy of data (Accounting biases)	• Regularly relying on one or few sources per event and exposed to measurement error therein • Usually based on approximations without ability to quantify uncertainty or accuracy	• Estimation based on (large) sample from population of interest • Explicit quantification of uncertainty in estimates • Subject to systematic (non-classical) measurement error
Comparability of source data	• Broad coverage across geographies and time • Exposed to idiosyncrasies in reporting protocols between different data sources without them typically being explicit	• Increased scope for harmonized data collection across contexts • Idiosyncrasies in data collection between different micro-data sources are explicit
Geographical and income differences		
Geographical and income differences	• Lower information density in LMICs and lower ability to draw on ancillary data sources for shock recording • Systematic under-coverage of events and marginalized population groups in LMICs • Lower accuracy of data due to lower density of (independent) sources of information • Underestimation of impacts in LMICs due to higher share of uninsured damages and low value of affected goods and assets	• Potentially greater accuracy of micro-data in HICs due to use of more sophisticated measurement approaches and higher levels of education among respondents

The table compares survey microdata with disaster inventories such as EM-DAT with regard to their coverage, detail, and accuracy with which they capture climatic shocks as well as geographical and income differences.

nature portfolio

Corresponding author(s): Philip Wollburg

Last updated by author(s): 05/07/2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Survey data was collected using the World Bank's software Survey Solutions.
Data analysis	Authors used Stata/MP 18.0 for data analysis. No custom code is necessary to replicate the results, but code to exactly replicate all tables and figures is available on Zenodo at the following link: 10.5281/zenodo.12667754

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The raw data used in this study as well as the full questionnaires are publicly available through the World Bank's on-line resources. LSMS-ISA datasets, questionnaires and documentation are publicly accessible via the following link: https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA. Links to the various datasets used in this study are provided in the "dissemination" tabs. More specifically, the dataset includes data from the Ethiopian Social Survey (waves 1 to 4), Malawi's Integrated Household Panel Survey (waves 1 to 4), Mali's Enquête Agricole de Conjoncture Intégrée (waves 1 and 2), Niger's Enquête National sur les Conditions de Vie des Ménages et Agriculture (waves 1 and 2), Nigeria's General Household Survey (wave 4) and Tanzania's National Panel Survey (waves 1 to 5).

The EM-DAT data was downloaded for free on the following public website: https://public.emdat.be/. Data was downloaded on the 29th of November 2023. The following data filters were applied - Classification: Natural. Countries: Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania. Time range: 2008 – 2020

The analysis dataset is available from Zenodo using the following link: 10.5281/zenodo.12667754

Shapefiles and other raw geodata required to produce figure 2 were downloaded from the GADM database, which can be accessed for free at the following link: https://gadm.org/download_country.html.

Population and other aggregate statistics were downloaded from the World Bank Development Indicators Database, directly into Stata using the "wbopendata" package.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Information on sex and gender was collected based on self-reported information.
Population characteristics	See below
Recruitment	The authors were not involved in data collection but used publicly available survey data from the World Bank. In these datasets, households were selected to be representative of the population at the national and sub-national level using a two-stage stratified sampling design with census enumeration areas (EAs) as primary sampling units and households as secondary sampling units. Households are then tracked through time, except for Mali which only tracks enumeration areas (EAs). For this analysis, only households engaged in agriculture were retained.
Ethics oversight	The authors were not involved in data collection but used publicly available survey data from the World Bank. The World Bank surveys were implemented by the respective national statistical office (NSO). The NSO conducts the survey as the sole official statistical authority in the country and in accordance with the respective National Statistical Act, which exempts the NSO from institutional ethics approvals. Informed consent was received from all survey respondents in each country. The World Bank does not require institutional ethics approval for household surveys that are partly or fully financed by the World Bank, including the surveys in Ethiopia, Malawi, Mali, Niger, Nigeria and Tanzania that inform our research.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study uses quantitative, observational data from surveys conducted in six countries: Ethiopia, Malawi, Mali, Niger, Nigeria, and Tanzania.
Research sample	In this analysis, our research sample consists of all households which have engaged in agriculture within the LSMS, which are nationally representative surveys that already existed prior to this research. Households are selected to be representative of the population at the national and sub-national level using a two-stage stratified sampling design with census enumeration areas (EAs) as primary sampling units and households as secondary sampling units. Households are then tracked through time, except for Mali which only tracks enumeration areas (EAs).
Sampling strategy	This study uses data that already existed, and the researchers did not guide the sampling strategy. The surveys use a stratified two-stage sampling procedure, with census enumeration areas (EAs) as primary sampling units and households as secondary sampling units. Surveys are representative at the national and sub-national level and are stratified by administrative division and urban/rural levels.
Data collection	This study uses data that already existed, and the researchers did not have guide data collection. Data was collected in in-person interviews, with the World Bank's Survey Solutions CAPI software. For more information about specific survey rounds, information can be found in Basic Information Documents (BIDs) that accompany surveys, and that are made available in the World Bank's microdata library: https://microdata.worldbank.org/index.php/home

Timing	In Ethiopia, data from the Ethiopian Social Survey (ESS) were assembled across four survey periods: 2010/2011, 2012/2013, 2014/2015 and 2017/2018. In Malawi, data from the Integrated Household Panel Survey (IHPS) were assembled across four periods: 2009/2010, 2012/2013, 2015/2016 and 2018/2019. In Mali, data from the Enquête Agricole de Conjoncture Intégrée (EACI) was assembled from two periods: 2014 and 2017. In Niger, data were drawn from the Enquête National sur les Conditions de Vie des Ménages et Agriculture - ECVM/A) across two periods: 2011 and 2014. In Nigeria, data were assembled from the General Household Survey (GHS) across four periods: 2010/2011, 2012/2013, 2015/2016 and 2018/2019. In Tanzania, data were assembled from the National Panel Survey (NPS) across five periods: 2008/2009, 2010/2011, 2012/2013, 2014/2015 and 2018/2019.
Data exclusions	For this analysis, only households engaged in crop agriculture were retained. This means that 25,201 households across the six countries of study were excluded from the analysis.
Non-participation	This information can be found in Basic Information Documents (BIDs) that accompany surveys, and that are made available in the World Bank's microdata library: https://microdata.worldbank.org/index.php/home
Randomization	Our study is observational and experimental randomization does not apply. Randomization in the selection of the sample is discussed above.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging