Do ultra-poor graduation programs build resilience against droughts? Evidence from rural Ethiopia

Hirvonen, Kalle (IFPRI); Gilligan, Daniel O. (IFPRI); Leight, Jessica (IFPRI); Tambet, Heleene (IFPRI); Villa, Victor (CIAT)

Abstract

We study the role of a multifaceted ultra-poor graduation program in protecting household wellbeing and women's welfare from the effects of localized droughts in Ethiopia. Using data from a large experimental trial, we assess an integrated livelihood and nutrition intervention that supplemented the consumption support provided by Ethiopia's Productive Safety Net Program (PSNP). We match three rounds of household survey data to detailed weather data to identify community-level exposure to droughts. Exploiting the random assignment to the graduation program, we assess whether the impacts of droughts on household food security, livestock holdings, women's diets and nutritional status, and the prevalence of intimate partner violence (IPV) differ between households in treated and control clusters. We find that droughts have substantial negative effects on these outcomes, but the intervention serves to consistently moderate these effects, and for some outcomes (particularly diets and nutrition and IPV), the intervention fully protects households from any adverse drought effects. A further analysis exploits variation across treatment arms that received different program elements and suggests that the primary mechanism is enhanced household savings.

Key words: resilience; weather shocks; climate change; graduation model programs; social safety nets

JEL Codes: I38, O12, O15

Acknowledgements: This research was funded through the CGIAR Research Program on Harnessing gender and social equality for resilience in agrifood systems (HER+). The original trial was funded by the United States Agency for International Development (USAID) through a subaward to World Vision under Cooperative Agreement No AID-FFP-A-16-00008 (2016-2021). We also thank World Vision, CARE, and ORDA for their contributions to the original impact evaluation. The ethical approvals for the trial were obtained from Hawassa University in Ethiopia and from IFPRI. We have benefitted from discussions with and comments from Jenny Aker, Leah Bevis, Martina Björkman Nyqvist, Ben D'Exelle, John Giles, Melissa Hidrobo, Namrata Kala, David McKenzie, Michael Mulford, Shalini Roy, Carly Trachtman, Kelsey Wright, and participants at the IPA and the Global Poverty Action Lab's 2024 Annual Researcher Gathering, 2023 DIME-KDI conference, 2023 Nordic Conference in Development Economics in Gothenburg, and UNU-WIDER brownbag seminar for useful comments. The opinions expressed here belong to the authors and do not necessarily reflect those of the funders or the CGIAR.

1. Introduction

Evidence amassed by the Intergovernmental Panel on Climate Change (IPCC) now suggests with high confidence that global warming increases the risk of harmful weather events such as droughts, floods, and tropical cyclones (Seneviratne et al., 2021). A substantial literature in economics and other disciplines has demonstrated that such shocks are largely uninsured across rural areas in low- and middle-income countries,[1] and thus adverse rainfall and temperature shocks often force households to cut back on consumption or sell assets, worsening their food security and increasing their vulnerability to chronic poverty.[2] Beyond economic hardship, adverse weather events may disproportionately harm women due to their reduced intrahousehold access to resources and greater health disparities (Björkman-Nyqvist, 2013; Corno et al., 2020; Dercon and Krishnan, 2000; van Daalen et al., 2020). Moreover, recent research links adverse weather events to heightened risk of intimate partner violence, a form of violence experienced almost exclusively by women (Abiona and Koppensteiner, 2018; Díaz and Saldarriaga, 2023; Epstein et al., 2020). Therefore, as climate change intensifies, one of the most urgent global development challenges is ensuring that households can withstand these shocks – particularly in poor rural areas where formal insurance markets are typically absent (Collins et al., 2009).

Addressing these risks requires scalable interventions that not only alleviate poverty in the short term but also enhance resilience against future shocks. Over the past two decades, safety net programs providing recurring cash or in-kind payments have become an increasingly common response to address chronic poverty in low- and middle-income countries (Beegle et al., 2018; World Bank, 2018). When carefully designed and implemented, these programs can be effective in reducing poverty, improving food security, and facilitating household asset accumulation (Andrews et al., 2018; Crosta et al., 2023; Hidrobo et al., 2018; Leight et al., 2024). However, these programs have rarely been able to deliver sustainable improvements in livelihoods large

---

[1] Seminal empirical papers in this area include, but are not limited to: Fafchamps et al. (1998), Dercon (2004), Alderman et al. (2006), and Maccini and Yang (2009).

[2] Droughts, floods, and other adverse weather events have also been found to be harmful to children in low-income households, negatively affecting their physical, cognitive, and non-cognitive development as well as educational outcomes (Chang et al., 2022; Cooper et al., 2019; Hoddinott and Kinsey, 2001; Webb, 2023) with long-term consequences that shape welfare and health outcomes in adulthood (Alderman et al., 2006; Dercon and Porter, 2014; Dinkelman, 2017; Maccini and Yang, 2009).

enough to enable a substantial share of beneficiaries to sustainably exit poverty (Sabates-Wheeler et al., 2021). These shortcomings have motivated the development of so-called graduation model programs, which combine asset transfers, technical and business skills training, access to savings groups or microfinance, and regular coaching or mentoring. Unlike traditional safety nets that primarily provide short-term consumption support, graduation programs seek to generate long-term shifts in income and economic self-sufficiency. While multi-faceted graduation model programs show promise in improving livelihood outcomes (Balboni et al., 2022; Bandiera et al., 2017; Banerjee et al., 2015; Bossuroy et al., 2022; Brune et al., 2022), evidence remains limited on their ability to build resilience and provide protection during droughts and other adverse weather events.

We contribute to this literature by providing the first evidence on whether a graduation program can enhance resilience to adverse weather events. We use data from a cluster randomized controlled trial (RCT) that evaluated a multifaceted graduation model program embedded within Ethiopia's Productive Safety Net Programme (PSNP) – a large-scale social protection program in a country with a long history of devastating droughts (De Waal, 2017) that are expected to become more frequent due to global warming (Price et al., 2022; World Bank, 2021). The *Strengthen PSNP4 Institutions and Resilience* (SPIR) program was a five-year graduation model program designed to complement PSNP by providing additional gender-sensitive livelihood and nutrition interventions. It aimed to strengthen economic opportunities and improve household well-being through a combination of financial literacy training, savings groups, and the promotion of new income-generating activities. Both women and men were encouraged to participate in these activities, which were delivered through Village Economic and Social Associations (VESAs), serving as platforms for both livelihoods and nutrition programming. In some study arms, SPIR also provided targeted one-time livelihood transfers to the poorest households, either in the form of an improved poultry production package or an unconditional cash transfer. Importantly, the set of livelihood interventions was smaller in scope, and the transfer amounts lower in value, relative to well-known graduation model programs originally designed by BRAC (an international non-governmental organization) and evaluated in Banerjee et al. (2015) and Bandiera et al. (2017).

The multi-arm RCT was designed to evaluate combinations of four interventions: core livelihood activities and core nutrition activities as well as enhanced versions of both types of activities. Prior reporting of results from this experiment showed that SPIR had positive impacts on domains such as livestock-related production, financial inclusion, and access to health services, but no significant impacts on consumption or poverty (Alderman et al., 2023; Leight et al., 2023), suggesting a more limited response to this ‘light-touch graduation approach’ on average than in graduation programs providing larger transfers and more intensive livelihood support. There was also, in general, no evidence of significant effects on children’s anthropometrics in a context of widespread stunting (Alderman et al., 2023). However, the impacts on savings and access to credit were comparable to the literature evaluating BRAC-implemented programs (Bandiera et al., 2017; Banerjee et al., 2015) and visible both at the midline and endline (Leight et al., 2023). This finding is promising in terms of resilience given that higher levels of savings among low-income households can impact their ability to manage risks more effectively (Beaman et al., 2014; Cui and Tang, 2024; Karlan et al., 2017; Pomeranz and Kast, 2024).

We map high-resolution gridded climate data to the experiment’s household survey data to understand whether and how this light-touch graduation program protected households against the effects of droughts. Our identification relies on exogenous variation in program exposure induced by the cluster RCT coupled with both longitudinal and cross-sectional variation in drought incidence over the three survey rounds conducted between 2016 and 2021.[3] Our primary measure of drought is constructed using the Standardised Precipitation-Evapotranspiration Index (SPEI)[4] developed by Vicente-Serrano et al. (2010). Our drought indicator captures relative dryness experienced during the cropping season as any negative deviation relative to the long-term mean in the same locality. In other words, we examine even moderate fluctuations in dryness that represent typical variation across cropping seasons, distinct from rarely occurring extreme weather anomalies.

Our findings suggest that even these moderate changes in relative dryness during the main cropping season cause statistically significant and economically meaningful fluctuations in household welfare as well as women’s health and welfare, specifically. To assess the magnitude

[3] Related work in this area has primarily relied on cross-sectional variation in weather or other type of shocks (Emerick et al., 2016; Karlan et al., 2017; Macours et al., 2022; Pomeranz and Kast, 2024), but given high spatial correlation in weather shocks, an analysis relying solely on cross-sectional variation may confront limited power. Accordingly, we argue that exploiting substantial longitudinal variation is a meaningful strength of our empirical design.
[4] We follow a number of recent papers in using the SPEI as a primary measure to capture weather shocks, including Couttenier and Soubeyran (2014); Harari and La Ferrara (2018); Maystadt et al. (2015); Webb (2023).

of the coefficients of interest, we evaluate the effects of a 0.25 standard deviation (SD) increase in relative dryness defined relative to the locality-specific long-term mean; this corresponds roughly to the median increase in relative dryness observed in drought-affected areas, 0.23 SD. In communities in the control arm, a drought of this magnitude leads to a 36 percent increase in self-reported food insecurity (the food gap) and a 35 percent decrease in aggregated livestock holdings. Focusing on women's health and human capital specifically, there is also some suggestive evidence that droughts have an adverse effect on women's dietary diversity and BMI, and substantial evidence of increased risk of intimate partner violence: a 0.25 SD increase in relative dryness increases the risk of intimate partner violence by 21 percent, consistent with previous literature (Abiona and Koppensteiner, 2018; Díaz and Saldarriaga, 2023; Epstein et al., 2020). While we lack consumption data from all survey rounds, evidence based on the endline data only shows that a past drought in the magnitude of 0.25 SD results in a five percent decline in consumption in control households.

Strikingly, the corresponding effects of these drought shocks are either partially or completely muted in households that resided in clusters randomly assigned to receive SPIR programming. In drought-affected clusters served by SPIR, the increase in the food gap is roughly half the magnitude of the increase observed in drought-affected clusters in the control arm, while the decline in livestock holdings is less than a third of the corresponding decline in drought-affected control clusters; and drought does not lead to any significant increase in IPV at all in clusters served by SPIR. For women living in drought-affected clusters, both dietary diversity and BMI are significantly larger in SPIR communities than in the control group. Similarly, drought impacts on household consumption seem completely muted in treated households. These findings indicate that the livelihood interventions protected households and especially women during droughts, over and above the support received through the existing safety net program, the PSNP.

Further analysis suggests little evidence of heterogeneity across different SPIR treatment arms, despite the different combinations of interventions offered. Moreover, we find no indication that the SPIR intervention led households to adjust their farming practices. While the role of livestock in total income and asset portfolios increased, there was no meaningful diversification into non-agricultural activities. The primary mechanism appears to be a substantial relative increase in savings, attributable to the establishment of VESAs and observed consistently among all households served by SPIR (including those who did not receive one-time cash or poultry transfers). In the face of a drought shock, treated households then use this enhanced stock of savings to mitigate the impact of the shock, allowing them to smooth consumption and partially protect productive (livestock) assets. Together, these results suggest that while the graduation program did not lead to higher levels of aggregate consumption or income, it facilitated the establishment of a buffer stock of savings that safeguarded households during droughts. These findings align with existing literature on savings group promotion campaigns, which demonstrate modest increases in savings (less than $15, on average) that then protect households from income shocks and seasonal income fluctuations (Beaman et al., 2014; Karlan et al., 2017).

Our study expands upon the extensive body of literature examining the negative impact of adverse weather events on household wellbeing. We add evidence for an extremely poor population (nearly two thirds of households were below the extreme poverty line) where severe droughts have contributed to starvation and long-term poverty in the past (De Waal, 2017; Dercon and Porter, 2014; Webb and von Braun, 1994) but during a period when droughts were less severe, reflecting a more typical year-to-year variation in drought stress in this context. To date, a considerably smaller literature has tested interventions to build household resilience against adverse weather events. [5] A small but growing body of research shows that safety net programs can at least partly mitigate the negative impacts of adverse weather events and other natural disasters (Adhvaryu et al., 2024; Asfaw et al., 2017; Christian et al., 2019; De Janvry et al., 2006; Hou, 2010; Knippenberg and Hoddinott, 2017; Premand and Stoeffler, 2020).[6]

Importantly, our study is the first to test whether increasingly popular graduation model programs make households resilient to drought shocks, building on Macours et al. (2022), who study a program that offered vocational training or a $200 investment grant to short-term safety net program beneficiaries in rural Nicaragua. Our findings are similar to those from the cash-plus program analyzed by Macours et al. in that we also find that simple cash transfers (offered to the control arm via the PSNP itself) are insufficient to buffer households against droughts, while a

---

[5] Apart from social protection programs, other active research in this area include the promotion of climate-smart agriculture policies (e.g., drought/flood resistant crop varieties) (Emerick et al., 2016; Lipper et al., 2017), weather index insurance programs targeted to agricultural households (Carter et al., 2017; Jensen and Barrett, 2017; Karlan et al., 2014), anticipatory cash transfers to areas forecasted to experience weather shocks (Pople et al., 2021) and emergency loans to affected households (Lane, 2024).

[6] In line with Knippenberg and Hoddinott (2017), evidence from the control households provided in this paper indicates that the PSNP is not able to fully insure poor households against droughts.

bundled model of interventions effectively assists households in coping with shocks. However, the relevant mechanisms are notably different. In their context, the complementary interventions – vocational training or productive grants – facilitated household diversification into non-agricultural activities, and thus protected them against the negative impacts of weather shocks on agricultural income. In our context, diversification into non-agricultural activities is virtually non-existent: less than five percent of the households in the control arm report any non-agricultural business or regular wage employment at endline, and there is no evidence of any meaningful treatment effect of SPIR on these outcomes (Leight et al., 2023). Instead, we observe some realignment within agriculture, as livestock income increases while crop agriculture income remains unchanged. Yet, the primary mechanism appears to be buffer stock savings. Overall, these findings are encouraging because they suggest that even relatively light-touch interventions (without larger asset or cash transfers) can promote resilience within the context of an existing safety net program, and even in a context with minimal off-farm income-generating opportunities.

Finally, we contribute to the growing body of literature unpacking the differential effects of drought shocks and other weather shocks on women, and particularly the literature analyzing the impact of negative and positive income shocks on intimate partner violence. While the differential vulnerability of women to climate-related risks is of high interest to policymakers (World Bank, 2015), high-quality evidence that empirically demonstrates this vulnerability is more limited, partly because measuring individual-level consumption or economic status in extremely poor, rural subsistence production households is challenging. Our paper addresses this challenge by drawing on nutritional and health measures (dietary diversity and BMI) measured specifically for women, thus joining an emerging literature that uses individual-specific outcomes to document the disproportionate vulnerability of women to weather shocks (Fruttero et al., 2023), and particularly expanding the literature focusing on the vulnerability of women to shocks experienced during adulthood, as distinct from shocks experienced in childhood.[7]

For IPV specifically, Díaz and Saldarriaga (2023), Epstein et al. (2020), and Abiona and Koppensteiner (2018) use exogenous variations in rainfall as a proxy for income shocks and

[7] The literature reviewed in Fruttero *et al.* encompasses papers analyzing the effects of both shocks during childhood and during adulthood, but the papers reporting on health-related outcomes are primarily analyzing early life shocks, while we also analyze the relationship between weather shocks and adult women's nutritional status.

demonstrate that dry spells result in large increases in the risk of IPV. A sizable existing literature has also documented that positive income shocks, in the form of randomly assigned cash or in-kind transfers, significantly reduce the risk of intimate partner violence (Buller et al., 2018; Haushofer et al., 2019; Heath et al., 2020; Roy et al., 2019). Here we bring these two strands of literature together by providing evidence that a (gender sensitive) graduation program can offset the heightened risk of IPV during droughts, potentially by weakening a pathway in which droughts increase poverty-related stress levels and thus IPV for households in the control arm (while treated households are less affected by drought shocks). Our further analysis provides suggestive support to the oft-cited stress-pathway in the IPV literature (Aizer, 2010; Fox et al., 2002): droughts increase stress levels among male household members while exposure to the SPIR program weakens this relationship.

2. Context and the intervention

The Productive Safety Net Program (PSNP)

This analysis draws on data from a trial conducted in the context of the PSNP, a safety net program launched in 2005 and designed to provide a more sustainable response mechanism in areas of Ethiopia that have been historically vulnerable to droughts, as opposed to recurring *ad hoc* emergency appeals for food aid and famine relief (Wiseman et al., 2010). Within the districts served by the PSNP, beneficiary households receive payments (in the form of cash and/or food) for six months in exchange for performing labor-intensive public works, while poor and chronically food-insecure households with limited labor capacity receive unconditional, direct transfers. Communities themselves select beneficiaries by applying a proxy means testing strategy. The program reaches eight million people, rendering it one of the largest safety net programs in Africa and, outside India, the largest public works program in the world (Beegle et al., 2018). PSNP is implemented by the government of Ethiopia, and largely funded by its international partners.

Evaluations based on quasi-experimental methods show that the PSNP has been successful in improving household food security and creating community assets (Hirvonen et al., 2022). However, PSNP households remain vulnerable to droughts, despite recovering from these shocks more rapidly when compared to poor non-PSNP households (Knippenberg and Hoddinott, 2017).

In addition, there has been negligible exit from poverty for PSNP beneficiaries (Sabates-Wheeler et al., 2021).

SPIR program

The SPIR Development Food Security Activity (DFSA) in Ethiopia was a five-year program (2016–2021) providing complementary livelihood, nutrition, and gender activities intended to strengthen the PSNP program and expand its impacts. Funded by USAID and implemented by World Vision, CARE and ORDA in close collaboration with the Government of Ethiopia, SPIR was organized around a core set of livelihood and nutrition activities delivered to nearly 500,000 beneficiaries. The core livelihood activities (L) included the formation of Village Economic and Social Associations (VESAs), the primary platform supporting financial literacy training and the promotion of new income generation activities. Both women and men were encouraged to join VESAs. They were also used as a platform for the core nutrition activities (N), including behavior change communication (BCC) around nutrition and water, sanitation, and hygiene (WASH).

SPIR also introduced enhanced models of livelihoods and nutrition activities. The enhanced livelihoods model (L*) included all core livelihood activities, supplemented with a targeted transfer provided to 10 out of 18 households in each kebele (sub-district)[8] that were classified as extremely poor according to a baseline asset index.[9] The transfers were provided based on a kebele-level randomization either as an improved poultry production package ($200 in value) including 16 pullets, poultry feed, support for veterinary services and training; or a one-time unconditional cash transfer of equivalent value of $200 in Ethiopian birr. The transfer was also formally targeted and disbursed to the female spouse. The enhanced nutrition activities (N*) included a more targeted program of nutrition BCC including home visits, a 2-week community-based participatory nutrition promotion activity for caregivers of underweight children to learn and practice improved child feeding practices, male's engagement groups, and an invitation to participate in weekly Interpersonal Psychotherapy in Groups (IPT-G) for women and men who were screened for

[8] Administratively, Ethiopia is divided into regions, zones, woredas (districts) and kebeles (sub-districts). Kebele is the lowest-level administrative unit and the community targeting of the PSNP as well as many key government services such as agricultural and health extension are organized and provided at this level.

[9] The L* activities also included a one-time aspirations promotion event in the form of inspirational documentary films, but analysis of the impact of this aspirations intervention found null effects (Leight et al., 2021).

elevated depressive symptoms. (Male engagement groups and IPT-G for men were rolled out only following the midline survey.)

These packages were combined into multisectoral graduation model programs and randomized into four treatment arms: T1: L*+N*; T2: L*+N; T3: L+N*; and T4: PSNP only. Figure 1 summarizes the contents of each study arm and provides the timing of the interventions and evaluation surveys.

[Figure 1 here]

Average treatment effects

The impacts of the SPIR program on a set of pre-specified economic outcomes have been documented in Leight et al. (2023), and we summarize the key findings here. The enhanced livelihoods interventions, which included asset transfers targeted at extremely poor households, led to moderate improvements in financial inclusion and increased cash income from livestock. However, these interventions resulted in only modest asset accumulation, and there were no statistically significant effects on household income or consumption. Notably, there was no meaningful difference in outcomes between households that received poultry transfers and those that received an equivalent cash transfer.

The most pronounced effects were observed in financial inclusion. Access to credit increased by 8–10 percentage points compared to a control mean of 45 percent, while the probability of reporting any savings rose by more than 30 percentage points from a baseline level of 40 percent. Moreover, both poultry and cash recipients experienced a 25 percent increase in past-year income from livestock. However, since livestock earnings accounted for only about 11 percent of total household consumption on average, this income gain had little impact on overall household well-being; no significant changes were detected in total consumption or food security.

Among extremely poor households who received only savings groups and financial literacy training but no poultry or cash transfers, the only measurable impact was an increase in reported savings. The same pattern emerged among less poor households who participated in savings groups and training across all treatment arms but received no additional transfers. In both cases, participation in the program increased the likelihood of reporting any savings by approximately 30 percentage points.

Importantly, there is no evidence that the SPIR program influenced participation in non-agricultural activities. Across both midline and endline rounds, engagement in non-agricultural employment remained minimal: only three percent of households reported operating a non-agricultural business (rising to six percent among less poor households), approximately four percent engaged in regular wage labor, and around 25 percent participated in casual wage labor. These low participation rates remained stable throughout the study period, suggesting that the intervention did not facilitate a shift toward non-farm economic activities.

In this paper, we assess how these treatment effects evolve in response to a shock; whether households exposed to the SPIR program were better able to withstand the impact of droughts.

3. Data

Surveys

The study took place in 13 woredas (districts) and 192 kebeles across the Amhara and Oromia regions of Ethiopia (Figure A1 in Appendix A). The randomization into treatment was stratified at the woreda level and the treatments were randomly assigned at the level of clusters, or kebeles. At the household level, eligibility criteria required being current PSNP beneficiaries, having at least one child aged 0-35 months at baseline, and having the mother or primary female caregiver of the same child also living in the household at baseline.

Three rounds of data were collected, with a baseline survey between February and April 2018, a midline survey between July and October 2019, and an endline survey originally planned for 2020 but delayed due to COVID-19 and undertaken 36 months after the baseline between February and April 2021. The baseline sample included 3,314 households with a child under three years of age, or just over 17 PSNP beneficiary households in each kebele. The midline survey achieved an overall sample of 3,220 of the original households, yielding an attrition rate of 2.8 percent. Of the 3,246 households eligible for the endline survey (after removing those who had permanently moved), 3,094 were able to be located and interviewed, leading to an attrition rate of 4.7 percent relative to the target sample. A large portion of the attrition (80 households) at the endline was due to conflict in northern Amhara (linked to the emerging conflict in Tigray), where four kebeles were rendered permanently inaccessible to the survey team.

Outcome variables

The variables of interest for this analysis are outcomes that are plausibly responsive to drought and can capture its effects on welfare at both the household and individual level. At the household level, food security, consumption, and/or assets are typically employed to analyze the effects of weather-related shocks such as droughts on economic welfare (Carpena, 2019; Kazianga and Udry, 2006; Macours et al., 2022). However, to capture whether the adverse effects of droughts are uniquely salient for women, it is arguably most informative to use measures of women's nutritional and health status, given that measuring individual-level consumption in extremely poor rural households is challenging. In addition, we generally constrain our analysis to focus on variables that were consistently measured in all three survey rounds, though we will present some supplementary analysis around one key variable (consumption) that was not measured at midline and another one that was not measured at baseline (savings).

First, as a measure of household level food security, we use a construct called the "food gap". Employed as the main food security indicator in PSNP evaluations, the food gap is measured by asking survey participants to report the number of months, out of the preceding 12 months, that they had "problems satisfying the food needs of the household".[10] Knippenberg and Hoddinott (2017) show how a self-reported drought shock increases food gap by 1.6 months among PSNP households. In our sample, the mean food gap in control households at the baseline was 2.1 months.

As a second measure of household economic welfare, we use an index of livestock assets. Livestock in this context has several purposes. First, large livestock (bulls, oxen) can be used as draft power during ploughing and threshing (Mekuriaw and Harris-Coble, 2021). Second, livestock products (e.g., dairy, eggs, meat) provide additional income to agrarian households. Third, as access to formal savings institutions remains highly limited in rural areas, livestock is an important form of savings that can be liquidated during droughts and other shocks (Fafchamps et al., 1998). Our measure of livestock holdings is based on tropical livestock units (TLUs).[11] The

[10] A month in which the household had "problems satisfying food needs" is defined as one where the household experienced hunger for five or more days.

[11] The standard measure of a TLU is one cattle with a body weight of 250 kg (Jahnke, 1982). TLU are expressed as ratios relative to this standard unit, where the ratios are based on metabolic weights. So, for example, six sheep have the same energy requirements as one cattle and so six sheep translate into one TLU. Consequently, 1 sheep equals 0.15 TLU while 1 chicken is 0.01 TLU.

average household in the control cluster owned 0.97 TLUs at the baseline, including on average 0.6 heads of cattle (bull, oxen, or cow), 0.3 pack animals (donkey or mule), 0.5 sheep/goat and 1.5 chickens.

To measure women's individual exposure to drought shocks, we use three variables. First, we assess dietary diversity among women.[12] The survey instrument was administered to the primary female in the household, who was identified as the mother or primary caregiver of an index child aged 0-35 months. Following the FAO and FHI 360 (2016) guidelines, we construct a diet diversity index that counts the number of food groups consumed by the primary female in the 24 hours prior to the interview. In this index, foods with similar nutritional contents have been categorized into 10 food groups. Since the indicator has been validated for women who are 15 to 49 years of age, we restrict the sample to primary women in this age range when analyzing this outcome. After this restriction, the final data has 9,087 diet diversity observations across the three survey rounds, excluding 16 percent of observations. At the baseline, the average primary female in control households consumed only from 2.1 food groups.

As a second measure of women's individual-level welfare, we use data collected on the height and weight of the primary female to construct a body-mass index (BMI), which is computed by dividing weight in kilograms by the square of body height in meters. The BMI is considered a reasonable proxy for adult health risks (Fogel, 1994, Waaler, 1984) and shown to be responsive to drought and other shocks (Dercon and Krishnan, 2000; Hoddinott and Kinsey, 2000). For the BMI analysis, we restrict the sample to women who are 15 to 49 years and exclude women with implausible BMI values (BMI below 15 or above 50). After these restrictions, we are left with 8,100 BMI observations across the three survey sounds, excluding 16 percent of observations. The mean BMI among primary females in control households at the baseline is 20.2. About 26 percent of the women in our sample are categorized as underweight (BMI<18.5) at the baseline.

As a third individual-level variable, we also analyze women's recent experience of intimate partner violence, motivated in part by recent research examining the influence of drought or rainfall shocks on the risk of IPV (Díaz and Saldarriaga 2023, Epstein et al. 2020, Abiona and Koppensteiner

[12] There is also some previous evidence that droughts reduce dietary diversity measured at the household level (Carpena, 2019).

2018). All three survey rounds contained an IPV module, administered to the primary female in accordance with the WHO protocol on ethical guidelines for conducting research on IPV (WHO, 2016). As the module was administered only if the respondent reported living with her husband in the last 12 months and if she was alone or with a child less than 36 months at the time of the interview,[13] the sample size is smaller than for the other outcomes.[14] At baseline, 17 percent of the women in the control clusters reported to have experienced physical, sexual and/or emotional violence in the previous 12 months. While this pooled variable is our primary indicator relevant to IPV, we report our results separately for the three forms of IPV that were collected in the surveys (see Appendix B for definitions).

Drought indicator

We focus on meteorological droughts that are characterized by precipitation (rainfall) deficiencies that persist across a large geographical area for an extended duration (Van Loon, 2015).[15] These deficiencies are linked with heightened potential evapotranspiration, in which reduced rainfall and higher temperatures contribute to greater evaporation of moisture from the air and more rapid transpiration of water from the soil by plants. Related empirical work in economics has traditionally used variation in local rainfall patterns as a proxy for droughts. However, temperature is a critical factor in meteorological drought, exacerbating water stress and evapotranspiration (Vicente-Serrano et al., 2010). Consequently, contemporary drought indicators integrate both precipitation and temperature data, thus accounting for temperature's influence in drought dynamics in a context of accelerating global warming.

Our primary drought indicator is the Standardised Precipitation-Evapotranspiration Index (SPEI) developed by Vicente-Serrano et al. (2010). The SPEI quantifies changes in climatic water balance over a given period considering the temporal difference between precipitation and potential

[13] Very few eligible women refused to respond to the IPV questions.

[14] At the baseline, 2,136 women responded to the IPV module, at the midline 1,648 women, and at the endline 2,161 women. An error in the CAPI program in the midline survey led to the IPV questions being excluded from the survey administered to about 600 women. All women who reported violence were given the option to be referred to the Women's Affairs Committee in her district.

[15] There are at least three other inter-related drought concepts (Van Loon, 2015). Hydrological drought refers to deficits in both surface and subsurface water resources. Soil moisture (or agricultural) drought is linked to reduced supply of moisture to plants. Socio-economic drought captures the socio-economic impacts of droughts. An important distinction here is that meteorological, hydrological or soil moisture droughts do not necessarily translate into socio-economic droughts if households and communities are able to successfully cope with the negative impacts caused by meteorological, hydrological or soil moisture droughts.

evapotranspiration influenced by temperature. The SPEI has been widely used by economists to analyze socio-economic impacts of droughts in various settings (e.g., Couttenier and Soubeyran, 2014; Harari and La Ferrara, 2018; Maystadt et al., 2015; Webb, 2023) and has been shown to be a more accurate predictor of crop yield than other drought indicators (Vicente-Serrano et al., 2012). We computed the SPEI for our study clusters using a user-written R-routine (Beguería and Vicente-Serrano, 2017). As inputs we used monthly precipitation accumulation (PR) and potential evapotranspiration (PET) from TerraClimate (Abatzoglou et al., 2018), a database with a timespan from 1990 to the present and a 4-km spatial resolution. In addition, we will also demonstrate that our primary findings are robust to alternate definitions of drought.

The SPEI is a standardized variable with a zero mean and standard deviation of one, quantifying the water balance in terms of standard deviations from the long-run average (here 1990-2020) in the locality, and thus a negative SPEI value indicates a deficit in water availability relative to the long-term average. We use this as our key definition: a locality experiences a drought if the SPEI value is negative, with more negative values indicating a more severe and prolonged drought event. While this definition agrees with a qualitative definition of a drought event: 'a deficit of water compared with normal conditions' (Van Loon, 2015), it is important to note that there is no agreed universal definition of drought that is quantifiable (Lloyd-Hughes, 2014). In standardized drought indicators such as the SPEI, drought intensity can categorized into mild (negative values up to -1), moderate (-1 to -1.5), severe (-1.5 to -2) and extreme (less than -2) (McKee et al., 1993).

We focus on the meher season: the main cropping season in the majority of Ethiopia, spanning between June and September.[16] We therefore compute the 4-month lag SPEI (SPEI-4) for each cluster at the end of each September, calculating the location specific changes in the climatic water balance from the long-run average during the whole meher season.

Figure 2 shows the distribution of SPEI during the meher seasons in the past 16 years, between 2005 and 2020. The 2016 meher season was characterized by an extreme drought in the study localities with the median SPEI value close to -2 standard deviations – a rare event that is statistically expected to occur approximately once every 44 years in normally distributed data. In

[16] Meher is by far the most important agricultural season in Ethiopia with more than 90 percent of the total cereal production in the country taking place during this season (Taffesse et al., 2012). In our sample of households, the belg harvest contributed only about 5% to the total annual crop production (in terms of total value measured at the baseline).

other years, the average (median) SPEI value fluctuates between -1 and +1 standard deviations, and this is also the case for the 2017-2020 period during which the study took place. Focusing on this period, we see that the 2017 and 2018 meher seasons were characterized by more relative dryness than the 2019 and 2020 seasons. As noted above, we primarily consider drought conditions in 2019 and 2020 meher seasons that are the relevant seasons for the midline (in 2019) and endline (in 2021) survey rounds, respectively. During the 2019 meher season, about one-third of the study clusters experienced drought-like conditions (i.e., negative SPEI), while in the 2020 meher only 5 percent of the clusters recorded a negative SPEI value.

[Figure 2 here]

To focus on drought events, we create a drought shock variable by setting positive SPEI values to zero, a procedure that is appropriately described as generating a positive rectified SPEI variable. Then, to facilitate interpretation of the regression coefficients, we multiply this positive rectified SPEI variable with -1 so that larger positive values indicate worsening drought conditions. Conditional on being positive (i.e., non-zero), the mean as well as median value of this drought shock variable is 0.23 units of standard deviation (SD) and maximum 0.47 SD in the 2019 and 2020 meher seasons (where the standard deviation refers to the normalized deviation relative to the long-term mean). Therefore, we will interpret the impact of drought shocks at 0.25 SD, a benchmark level that falls at the center of the spectrum of variation we see in our drought indicator during the study period. Importantly, this is a shock that is considerably milder than what is commonly understood to be a severe drought (i.e., the 2016 drought in Ethiopia, as noted above, has a median SPEI score close to -2, or roughly eight times more severe than the typical shock observed in our study period). In Appendix C, we show how the value of cereal crop harvests is negatively affected by these relatively mild drought shocks both in the control and treated households in our sample. Based on the estimates reported in Table C1, a 0.25 SD increase in relative dryness during the cropping season results in an approximate $27 PPP – or about six percent – drop in the total value of cereal crops harvested in the cropping season.[17] Later, we will

[17] Each survey round included a harvest module, but its structure was slightly different across rounds. At baseline and midline, the modules asked the respondent to report on all crops harvested by the household during the most recent harvest season. At endline, the module asked about the five most important crops by area. To maximize comparability, we focus on cereal crops, which are the most important crop category in this context. At baseline, nearly 80 percent of the total harvest value consisted of cereal crops.

demonstrate that these income shocks translate into sizable adverse impacts on household and individual wellbeing among the poor PSNP households in our sample.

Our econometric analyses measure the impact of drought shocks on outcomes of interest measured at the midline and endline. The endline took place in the post-harvest period in February-April 2021. For the endline observations, we therefore define the drought indicator based on the previous meher season that occurred between June and September 2020. The midline took place during the 2019 meher (between July and October). It is therefore not *a priori* clear whether the 2019 or 2018 drought conditions are most relevant for the midline observations. To address this, we seek guidance from our data by restricting the sample to control clusters and then regressing the outcome variables separately on two types of drought indicators. The first indicator defines drought using the 2019 meher (i.e., contemporaneous) conditions for the midline and the second indicator defines drought using the 2018 meher (i.e., previous year's meher) conditions for the midline. The regression results reported in Table C2 in Appendix C indicate that the majority of the outcomes respond to the drought conditions during the ongoing meher season in the midline. The only clear exception is women's BMI for which the previous meher season in the midline seems a stronger predictor. Therefore, when we measure the impacts on women's body masses, we define the drought variable using the 2018 meher conditions if the measurements were taken in the midline survey. For all other indicators measured at the midline, we consider the 2019 meher that overlapped with the midline survey round. However, we will also subsequently demonstrate that our primary findings are robust to alternate strategies to identify the appropriate drought shock at midline.

Balance

The statistical tests reported in Table D1 in Appendix D show that the sample is balanced across the study arms based on these outcome variables. The table also shows that the households are equally exposed to drought shocks across the four study arms: the SPEI values for each meher season during the study period are balanced across the study arms. This is expected given that the randomization was stratified at the woreda level and weather outcomes for a single season are highly spatially correlated within a woreda (as we show later).

Another potential source of bias in this analysis would be a high level of serial correlation in drought exposure, such that areas identified as experiencing droughts during the study period were already characterized by higher levels of poverty driven by past exposure to droughts. We can assess this hypothesis by regressing the drought shock indicator for midline and endline on various household welfare indicators measured at the baseline. None of the coefficients are statistically significant at the 5%-level and all are close to zero, indicating that there is no correlation between baseline poverty levels and drought exposure during the study period (Table D2 in Appendix D).

4. Econometric approach

We model the outcomes of interest (food security, livestock holdings, women's diet diversity, women's body mass index and intimate partner violence) for household *i* located in kebele *k* and woreda d at the time t ($Y_{i,k,d,t}$) as a function of drought conditions in the most relevant meher season ($Drought_{k,\tilde{t}}$):

$$(1) \qquad Y_{i,k,d,t} = \beta Drought_{k,\tilde{t}} + \vartheta T_{k,d} + \delta\left(Drought_{k,\tilde{t}} * T_{k,d}\right) + R_t + \theta_d + Y_{i,k,d,t=0} + \varepsilon_{i,k,d,t},$$

where t=1 if the household is observed in the midline survey and 2 if in the endline survey. $Drought_{k.\tilde{t}}$ is based on the SPEI measured in the relevant meher season, which would have affected the most recent major harvest as well as availability of grazing land and fodder for feeding livestock. For the endline survey that took place in 2021 during the harvest period, the relevant meher season is the one that took place in 2020. For the midline survey that took place during the 2019 meher season, we define the drought shock based on the 2019 meher conditions. However, as discussed in Section 3, women's anthropometry (BMI) seems to respond to drought shocks with a lag. Therefore, for this outcome we define the drought variable for the midline using the 2018 meher conditions.

To understand the degree to which the SPIR programming mitigated adverse impacts of droughts, we interact the drought variable with a binary variable (T) capturing the clusters that were exposed to any SPIR treatments. The equation also includes the non-interacted treatment variable. Initially, we pool the three treatment arms (implying that 75 percent of the sample clusters were exposed to SPIR, while only 25 percent are control clusters), but we will later consider impacts by treatment arm. The SPIR treatment variable is defined on the basis of the randomization and not on the basis of actual compliance, and thus δ and ϑ represent intention-to-treat (ITT) effects.

Given this specification, the impact of droughts in control localities is quantified by β. Again, the drought variable has been rectified so that it only obtains positive values, reflecting drought conditions, and thus when the outcome variable is the food gap or the risk of IPV, we expect β>0: when drought intensity increases, the period during which households have challenges in satisfying their food needs grows and the risk of IPV increases. For all other outcomes, we expect β<0: when drought intensity increases, households own fewer heads of livestock and women consume less diverse diets and have lower weight. The impact of droughts in treatment clusters is quantified by $\beta + \delta$. If SPIR treatments mitigate the adverse impacts of droughts, δ should be of opposite sign to β.

Additional control variables in equation (1) include the binary variable R_t equal to one for the endline survey and zero for the midline survey, woreda fixed effects θ_d corresponding to the randomization strata, and the outcome level at baseline ($Y_{i,k,d,t=0}$); the error term is captured in $\varepsilon_{i,k,d,t}$. The inclusion of baseline controls renders our estimation an analysis of covariance (ANCOVA) approach, shown to be more efficient than difference in differences particularly when the autocorrelation in the outcome variable is low (McKenzie, 2012). For the IPV outcomes, we replaced any missing baseline observation with zero and appended the estimated equation with a binary variable capturing these households for which the baseline value was missing.

The recommendation in the analyses of RCTs is to cluster the standard errors at the level of the unit of treatment assignment (Abadie et al., 2022). Here, however, the use of weather data creates strong spatial dependencies across study clusters (for more details, see Appendix E) in which case clustered standard errors are no longer valid (Barrios et al., 2012). Therefore, we compute Conley (1999) standard errors that are robust to both spatial autocorrelation and heteroskedasticity.[18] The Conley approach is based on a weighing matrix that places more weight on observations located closer to each other. We set the cut-off to 600 km; the weights are set to zero after this point. As shown in Figure E2, there is a negative and relatively linear association between the correlation of weather outcomes and distance between clusters. We therefore apply a Bartlett spatial weighting matrix in which the weights linearly decay as distance between the clusters grows (Conley, 1999).

[18] We used user-written *acreg* command in Stata to estimate these regressions (Colella et al., 2019).

5. Results

Main results

The regression results based on the estimation of Equation 1 are provided in Table 1. Figure 3 summarizes these results for household food security (Panel A) and livestock holdings (Panel B). In each panel, the top part of the graph provides the average effect of droughts in control clusters (i.e., β in Equation 1) while the corresponding estimate in treatment clusters (i.e., $\beta+\delta$) is presented below.[19] The difference in the drought impact between control and SPIR households is reported and labeled δ. The regression estimates have been divided by the baseline mean in the control group and then multiplied by 100, thus capturing the drought impacts in terms of percent of the baseline mean in the control group. As noted above, we interpret the impact of drought shocks in terms of 0.25 SD increase in relative dryness, roughly equivalent to the median drought shock within the sample areas experiencing any drought during the study period.[20]

[Table 1 here]

Interpreting the findings in Column 1 of Table 1 and Panel A of Figure 3, we can observe that control households exposed to droughts report considerably higher food gap: a 0.25 SD increase in relative dryness leads to a 0.75 month increase in the food gap, on average ($p<0.01$). Considering the baseline mean of 2.1 months in the control group, this translates into a 36 percent increase in the household food gap. The corresponding drought impact in treated clusters is less than half of this: 0.32 months or 15 percent, and this effect ($\beta+\delta$) is only weakly statistically significant ($p=0.06$). The coefficient on the interaction term (δ) reported in column 1 of Table 1 shows the positive impact of the graduation program on mitigating the drought effect on the food gap, and this impact is statistically significant ($p<0.01$).

[Figure 3 here]

Moving on to Column 2 of Table 1 and Panel B of Figure 4, a 0.25 SD increase in relative dryness decreases livestock holdings by 35 percent (0.34 TLU units), on average. (This is roughly equivalent to the loss of two goats and four chickens.) The corresponding impact in SPIR clusters is negative and statistically significant, indicating that treated households are also negatively

---

[19] The confidence intervals for the joint estimates were calculated using the *lincom* command in Stata 18.

[20] As noted previously, the median drought shock within drought-affected areas is 0.23 standard deviations.

affected by drought shocks. However, the average drought impact is one third of this magnitude in treated households, where a 0.25 SD increase in relative dryness decreases livestock holdings only by nine percent (0.09 TLU units), on average. The difference between control and treated households, again captured by the interaction term δ, is statistically different from zero (p<0.01) (see Column 2 of Table 1).

Columns 3 to 4 in Table 1 and Panels A and B of Figure 4 summarize the results on women's diet diversity and body-mass index. A 0.25 SD increase in relative dryness decreases women's diet diversity by 9 percent (0.18 food groups), on average. However, this estimate is not statistically different from zero at conventional levels (p-value = 0.164). The estimated effect of droughts for treated households is positive but not statistically different from zero (p = 0.232); however, the hypothesis that the effects of drought are equivalent for treatment and control households can be rejected at the one percent level (column 3 in Table 1). A very similar pattern is observed for BMI, where in control households, the effect of a 0.25 SD increase in relative dryness is less than -1 percent (-0.04 m^2/kg) and marginally significant (p=0.089). The corresponding estimate in treated households is positive and insignificant, but again statistically different vis-à-vis the effect estimated for control households (column 4 in Table 1). Overall, droughts seem to exert a small negative impact on women's diets and BMI in control households, and there is some evidence that women in treated households were insulated from these negative drought impacts.

Column 5 of Table 1 and Panels C of Figure 4 report the results on the impact of droughts on risk of IPV in the past 12 months. In line with the previous literature (Abiona and Koppensteiner, 2018; Díaz and Saldarriaga, 2023; Epstein et al., 2020), droughts increase the risk of IPV in control clusters. Here, a 0.25 SD increase in SPEI increases the risk of women experiencing intimate partner violence by 21 percent (or 1.4 percentage points). The corresponding impacts in treated households are considerably smaller in magnitude, and we cannot reject the null hypothesis that droughts have no impact on IPV in treated clusters. The difference between the estimated drought impacts between control and treated households is statistically significant at the one percent level, suggestive of a meaningful effect of SPIR in buffering the effects of drought. In Table H1, we examine the impacts of droughts and SPIR programming on different forms of IPV and find that sexual violence is particularly sensitive to drought shocks in control households. The drought impacts on physical and emotional violence are positive, but somewhat smaller and less precisely estimated. For all forms of IPV, the estimated drought impacts in SPIR households are not statistically significant.

[Figure 4 here]

Thus far, we have assessed the heterogeneity in program impact with respect to drought exposure by focusing on the signs and magnitudes of β and δ estimates in Equation (1). We could also use our regression results to quantify the program's (pooled midline and endline) ITT effects at the mean level of our drought indicator using the β, δ and ϑ estimates. Using this approach, the ITT estimates for food gap and livestock reported in Table 1 are not statistically different from zero[21], in line with the estimated program impacts on livelihood outcomes as reported in Leight et al. (2023).[22] This is all broadly consistent with the stylized fact reported above that only a minority of households are exposed to drought as defined in this analysis: about one-third of study clusters experienced drought in 2019 and only 5 percent in 2020, implying that the positive protective effects of SPIR were experienced by a minority of the sample. In the analysis pooling across drought-affected and non-drought-affected areas, these positive effects are attenuated toward zero.

Heterogeneity across treatment arms

Next, we explore the heterogeneity across treatment arms to assess whether these protective effects are driven by the more intensive treatment packages, or the livelihood transfers given to the poorest households.

Table H3 in Appendix H provides the regression results with β, δ, and ϑ in equation 1 estimated for each of the three treatment arms. All estimates of δ are of opposite sign to the estimated β coefficients indicating that all treatment arms provided at least some protection against droughts.

---

[21] These are calculated by comparing the drought impacts at the mean level of relative dryness in treated and control households: [β*0.04+ δ *0.04+ ϑ] - β*0.04, where 0.04 is the mean level of relative dryness across midline and endline calculated for all households. Using the *lincom* command in Stata we estimate the ITT impacts for food gap as -0.007 (p = 0.923) and for TLU as -0.003 (p = 0.858).

[22] Comparing the estimates discussed here to those reported in Leight et al. (2023) is not straightforward because the outcome variables are not identical (e.g., livestock assets are measured in terms value while we use TLU and the food security indicators are not the same) and because here we pool midline and endline survey rounds while in Leight et al. the treatment effects are estimated separately for midline and endline (as is standard in impact evaluations). Yet, the overall narratives reported in this paragraph and in Leight et al. are similar: the SPIR treatment effects on livestock assets and food security are modest when the full sample of households is considered. Leight et al find positive effects on livestock holdings when the sample is restricted to the poorest households that received the livelihood transfer, but not for the less poor households that did not receive the transfer.

The magnitudes of these coefficients are also broadly similar across the treatment arms, implying that increasing the treatment intensity did not result in any additional protective benefits.[23]

To understand the role of the livelihood transfers given to the poorest households, we next restrict the sample to eligible households.[24] We then re-estimate Table H3 for this sub-sample. Overall, the estimates on the interaction term reported in Table H4 are similar to the ones reported in Table H3. The impacts of drought on livestock holdings are smaller in absolute terms but not in relative terms because the baseline livestock holdings are considerably more modest among these households (and livestock holdings were part of the index that determined the eligibility for the livelihood transfers). The corresponding impacts on the IPV outcomes reported in Table H4 show that the drought estimates for the control households seem larger than those reported in Table H3 indicating that the risk of IPV increases relatively more in poorer households during droughts.[25] Treatment arm T3 did not include a livelihood transfer but was still effective in reducing the negative impact of droughts on sexual and emotional violence. Treatment arm T2 (combining livelihoods training with grants and the core nutrition intervention) seems consistently to have been most effective in protecting women against the increased risk of IPV due to droughts.

Other outcomes: household per capita consumption

We also consider household per capita consumption based on standard household food and non-food consumption expenditure modules (Deaton and Grosh, 2000). However, these modules were only administered at baseline and endline, not at midline, and thus we cannot directly estimate equation (1). Moreover, because the meher season prior to the endline had limited drought exposure (only about 5 percent of the clusters recorded a negative SPEI value during 2020 meher), we assess the impact of the 2019 drought conditions during which about one-third of the study clusters recorded negative SPEI. Figure 5 summarizes the impacts on log per capita (adult equivalent) consumption, benchmarked against 0.25 SD increase in relative dryness. The

[23] There is some suggestive evidence that the intensive nutrition interventions (N*) were more effective in protecting self-reported food security against droughts. Meanwhile, the intensive livelihood interventions (L*) were more effective in protecting livestock during droughts. In contrast, the Wald tests for diet diversity and BMI show no statistically significant differences across treatment arms.

[24] The eligibility assessment for livelihood transfers was conducted in all treatment clusters, including the control and T3 clusters where none of the households received these transfers.

[25] This was further confirmed with a simple interaction model when the sample was restricted to the control clusters: droughts did not lead to increases in the IPV risk in less poor households. Results available upon request.

corresponding regression table is reported in Appendix H (Table H2). In control households, we see that a drought in the magnitude of 0.25 SD during the 2019 meher season results in about 5.5% drop in household per capita consumption observed almost two years later in 2021 ($p = 0.065$). The corresponding impact in SPIR households is close to zero and not statistically significant ($p = 0.427$) while the difference vis-à-vis the estimate in the control households is highly significant. These results suggest that control households are in fact forced to cut down their consumption in the face of droughts while treated households are able to smooth their consumption.

[Figure 5 here]

Other outcomes: self-reported stress levels

Previous research on the impact of weather shocks on the risk of IPV has suggested that increased poverty-related stress levels, particularly among males, are a key factor driving the heightened risk of IPV following such shocks (Díaz and Saldarriaga, 2023; Epstein et al., 2020), but rarely report any data on stress. Our endline survey asked both male and female primary respondents to rate their current stress levels using a scale ranging between 1 (not stressed at all) and 10 (extremely stressed).[26] To explore whether droughts increase stress levels and whether SPIR weakens this relationship, we regressed the stress indicator measured in 2021 on a drought indicator capturing the weather conditions in the 2019-meher season, the SPIR treatment indicator and their interaction while controlling for woreda fixed effects.[27] The results, reported in Table H5, indicate that droughts increase stress levels for both males (column 1) and females (column 2). The negative and statistically significant ($p<0.10$ for males; $p<0.01$ for females) coefficients on the interaction terms suggest that the SPIR intervention slightly weakens this relationship. Specifically, a 0.25-SD drought is associated with a 30 percent rise in male stress levels in control clusters and 26 percent rise in SPIR clusters. For females, a drought of the same magnitude is linked to a 27 percent increase in stress levels in control clusters and 22 percent increase in SPIR clusters.

[26] This relatively simple stress indicator has been employed in the 'Stress in America' surveys conducted by the American Psychological Association during the COVID-19 pandemic, see https://www.apa.org/news/press/releases/stress/2021/stress-snapshot-january.pdf.

[27] We cannot directly estimate equation (1) because the outcome data (stress indicators) were only collected at endline. Consequently, we do not include a control for baseline value or survey round fixed effects. In addition, because the meher season prior to endline was relatively good in terms of weather (only 5% of study clusters had negative SPEI), we assess the impact of the 2019 drought conditions (in which about 30% of study clusters had negative SPEI).

Robustness

We performed a range of robustness checks to explore the sensitivity of these findings. First, previous work examining the impacts of droughts in Ethiopia and elsewhere in Africa have often quantified droughts in terms of deviation from long-term precipitation during the cropping season (Hirvonen et al., 2020; Shively, 2017; Thiede, 2014). We obtained monthly precipitation data from the TerraClimate database (Abatzoglou et al., 2018) and computed Z-score deviations of precipitation (rainfall) during the meher season. As with the SPEI indicator, we created a rainfall shock variable by rectifying positive Z-score values to zero and multiplied this positive rectified rainfall variable with -1 so that larger positive values indicate worsening conditions. Table F1 in Appendix F shows that this alternative drought indicator yields qualitatively similar results: rainfall shocks are highly damaging in control clusters and less so in treated clusters.

Second, a potential concern in including the woreda fixed effects in the estimated equation is that they may absorb a large amount of variation in the drought indicator (Fisher et al., 2012). To check this, we regressed the drought indicator on the woreda dummies and the survey round dummy. This regression yields an R^2 of 0.36, indicating that there remains considerable variation in the drought indicator after controlling for woreda and survey round fixed effects. In Table F2, we further show that our results are robust to replacing woreda fixed effects with a binary indicator capturing the region in which the study cluster is located.

Third, the survey round dummy aims to control for shocks that are common to all households surveyed at the same time (e.g., the COVID-19 pandemic). However, in Ethiopian context, covariate shocks can be highly region-specific raising a concern that our drought shock variable could be capturing the impact of some other covariate shocks instead. To address this, we replace the survey round dummy with region-specific time fixed effects. As before, the results are robust to this adjustment (Table F3 in Appendix F).

Fourth, the midline survey took place during the 2019 meher season. In the analyses we considered the drought conditions during this season for the outcomes measured in the midline survey, except for BMI that we hypothesized to have longer response time to droughts. Table F4 in Appendix F shows the results when we switch to using the 2018 meher season conditions for food security, livestock holdings, women's diet diversity and IPV outcomes. For BMI, we now used the 2019 meher drought indicator value. For all outcomes, the magnitudes of the coefficients and/or their precision decrease, but the coefficients on the drought indicator are always of the opposite sign to those on the interaction term. This corresponds to the finding that SPIR offered protection during droughts.

Fifth, the 2020-2021 period is characterized with considerable volatility in Ethiopia. Apart from the COVID-19 pandemic, part of the study area was attacked by desert locusts that decimated harvests (Alderman et al., 2020). Meanwhile, a 2-year civil war broke in November 2020 wreaking havoc in the SPIR study areas located in the Amhara region (Alderman et al., 2021). If drought shocks are correlated with these other covariate shocks, then our drought impact estimates may be biased. To explore this, we merged the household survey data with remote sensed data on locust invasions (FAO, 2023) and conflict (Raleigh et al., 2010). Table F5 in Appendix F then replicates the main regressions, adding controls for locust and conflict shocks. The estimates are nearly identical to those reported in Table 1, indicating that the story documented here is not driven by other major shocks that occurred in these localities during the study period.

Sixth, the 2016 drought was particularly severe in our study clusters, raising a question of whether our results are driven by households' exposure to the 2016 drought through serially correlated shocks. In Appendix G, we show that the drought shocks that occurred during the study period in 2019-2021 are only weakly correlated with the variation in household's exposure to the 2016 drought. Moreover, our results are robust to controlling for drought conditions in the 2016 meher season (Table G1 in Appendix G).

Seventh, previous research has found that drought shocks shape migration patterns in rural Ethiopia (Gray and Mueller, 2012). To explore whether changes in household composition (e.g., through in- or out-migration) could explain our findings, we re-estimated equation (1) using household size as the dependent variable. As Table F6 shows, we cannot reject the null hypothesis that drought shocks do not change household composition in control clusters, when measured in terms of number of members (column 1) or in terms of adult equivalents (column 3). Moreover, the coefficient on the SPIR interaction term is not statistically different from zero (columns 2 and 4), indicating no differential drought impacts between treatment and control clusters.

Finally, we used a 600 km distance cut-off to calculate Conley (1999) standard errors. Our main findings – that droughts have a negative effect, particularly on livestock holdings and food security, and increase the risk of IPV, while SPIR helps to mitigate these effects – remain robust to alternative distance cut-offs (50, 100, …, 700 km). Panels A to E in Figure F1 report the p-values for the key estimates based on these alternative cut-offs.

6. Mechanisms

The foregoing results demonstrate that SPIR effectively protected household food security, livestock assets, women's dietary diversity and BMI against localized droughts. Moreover, droughts substantially increased the risk of IPV, but only in study clusters where SPIR was not operational.

Prior research identifies several strategies that rural households can use to manage income shocks in the absence of formal insurance and credit markets. One strategy is adjusting crop production – either by investing in improved technologies, such as improved seeds, or diversifying the crop mix (Boucher et al., 2024; Emerick et al., 2016; Michler and Josephson, 2017). Another is accumulating liquid buffer stocks, such as cash reserves, which can be drawn upon during financial distress. Productive assets such as livestock can also serve as a buffer (Rosenzweig and Wolpin, 1993). However, in poorly integrated markets asset prices tend to co-move with incomes (Lange and Reimers, 2021), making asset accumulation a highly imperfect form of self-insurance. Finally, households may reduce their drought risk by diversifying beyond crop agriculture. Some increase their involvement in livestock farming, which is less sensitive to rainfall variability than crop production (Toulmin, 1986), while others shift into non-agricultural activities (Macours et al., 2022) that are not directly dependent on rainfall.[28]

Building on these insights, we next examine whether SPIR influenced households' ability to adopt these coping strategies. Like many multifaceted interventions, SPIR likely operates through multiple channels, supporting resilience by easing financial constraints, expanding livelihood opportunities, and strengthening social and economic networks. This aligns with the theory of change in graduation programs, which aims to address the structural barriers that keep households trapped in poverty.

[28] Even non-agricultural livelihoods may still be affected by rainfall fluctuations through local general equilibrium effects in agrarian economies.

Shifts in cropping patterns

We begin by examining whether SPIR led households to adjust their farming practices. Specifically, we estimate treatment effects on investments in seeds, inorganic fertilizers, pesticides, and hired labor. We also assess its impact on the number of crops cultivated. To do so, we restrict the sample to households that cultivated crops in the previous cropping season and use an ANCOVA specification, regressing each outcome on the SPIR treatment indicator, the baseline value of the outcome, the survey round, and woreda fixed effects. We use data from both the midline and endline rounds when available.[29]

In Table 2, we assess the impact of SPIR on expenditures for crop seeds, inorganic fertilizers, pesticides, and hired labor at the internal and external margins, as well as on the number of different crops cultivated by households (crop diversity). The estimated treatment effects are uniformly statistically insignificant and mostly small in magnitude, suggesting that shifts in crop agriculture are not a channel for the overall protective effects of SPIR observed.

[Table 2 here]

Accumulation of buffer stocks

Most survey respondents cited protecting themselves against unexpected income losses or health emergencies as their primary reason for saving, highlighting the importance of precautionary savings in this context. Building on this insight as well as on the finding reported in Leight et al. (2023) that SPIR led to increases in savings, we next assess whether SPIR households tap into their savings when faced with a drought shock. To explore this, we re-estimate equation (1) using total savings as the dependent variable.

The coefficient on the uninteracted SPIR variable shows that among households not affected by a drought shock, SPIR led to a $20 PPP increase in household savings relative to control households. The coefficients on the drought variables and the interaction terms capturing SPIR households affected by a drought confirm our hypothesis: treated households draw down their savings in response to shocks, whereas savings levels among control households remain unchanged. A

[29] The crop module in the endline questionnaire only asked whether the household spent money on inputs (except seeds) but did not ask how much was spent.

drought of 0.25 SD reduces savings in treated households by approximately $10 PPP – about half the magnitude of SPIR's overall effect on savings ($20 PPP) in this specification.

These cash buffers help cushion the impact of shocks, reducing the need to liquidate livestock or cut consumption – coping strategies that can be costly in the long term, as observed among control households (Figure 3, Panel A, and Figure 5, respectively). Column 2 of Table 3 further supports this interpretation, showing that SPIR households sell fewer TLU after a shock compared to control households.

[Table 3 here]

Income diversification

The results reported in Leight et al. (2023) show that SPIR households accumulate more livestock and generate higher cash income from livestock than control households. At the same time, SPIR has no impact on crop incomes, suggesting that livestock began playing a larger role in SPIR households' income portfolios. Unlike Macours et al. (2022) in the context of rural Nicaragua, there is no evidence of meaningful diversification out of agriculture in this setting. At endline, fewer than five percent of households report owning a non-agricultural business, and fewer than five percent report engaging in regular wage labor. This may reflect the limited demand for non-agricultural goods and services in a context where widespread poverty constrains spending largely to food.

7. Conclusions

The combination of high levels of poverty (Beegle et al., 2016), reliance on rainfed agriculture (You et al., 2012), and warm-to-hot climates make sub-Saharan Africa one of the world's most vulnerable regions to climate change (IMF, 2020). The impacts of global warming are already visible; in East Africa, droughts are now occurring more frequently than in the past (Haile et al., 2019). Therefore, one of the most urgent global development questions is how to improve resilience of poor rural households as the negative effects of global warming intensify.

Our research explores the effectiveness of light-touch graduation model programming in protecting households against local droughts. Droughts result in large increases in food insecurity and sizable fluctuations in household livestock holdings in control households participating in Ethiopia's flagship safety net program, the PSNP. The average drought impacts on women's diet diversity and body masses are relatively smaller in magnitude and less precisely estimated. However, the risk of intimate partner violence increases dramatically during droughts in the control clusters, which is likely linked to the increased male stress levels after these shocks. Assessing the impacts of identical droughts in localities that were exposed to the SPIR program, we find that these relatively light-touch livelihood interventions were highly effective in mitigating the negative effects of droughts on household food security and productive assets (livestock), indicating that the program substantially improved resilience among poor households. Previous work in this area finds that women often bear the brunt of weather and other shocks within households (Abiona and Koppensteiner, 2018; Dercon and Krishnan, 2000; Díaz and Saldarriaga, 2023; Epstein et al., 2020). To this end, we find that the SPIR interventions protected women's diet diversity and physical health (as measured by body mass) against droughts. Moreover, the impact on IPV was entirely muted in clusters that benefitted from the SPIR livelihood interventions.

As documented in Leight et al. (2023), the SPIR program led to substantial increases in accumulation of livestock assets as well as increased income from livestock products. Households exposed to the intervention were also considerably more likely to save and access formal credit than the control households only receiving the PSNP. These increases in productive assets and improvements in financial inclusion likely served as the mechanism that provided protection during droughts. Our further analysis suggests that varying the intensity of the livelihood or nutrition interventions did not bring about additional protective impacts. Moreover, the livelihood grants were not a necessary component: comparable poor households that did not receive the grant but benefitted from basic livelihood programming were similarly protected against droughts as were the other treated households.

Due to the experiment and the exogenous nature of drought shocks we consider the internal validity of these findings strong. However, we should be cautious when generalizing these results outside of the study population (Banerjee et al., 2017) – or outside the study period within the same population (Rosenzweig and Udry, 2019). While our study period was characterized by typical weather fluctuations during the main cropping season (see Figure 2), it did not contain a major drought. We find that these typical year-to-year variations in drought stress are highly damaging for households and women in control areas, but not in areas exposed to the SPIR program. Yet, we cannot be sure whether the program would have offered similar protection against a major 2015/16 – style drought that occurred across large part of Ethiopia (Hirvonen et al., 2020; NDRMC, 2016).

The most successful graduation model programs pioneered by BRAC have shown large increases in household consumption, food security and asset levels, both in the short and long-term (Balboni et al., 2022; Bandiera et al., 2017; Banerjee et al., 2015). The BRAC model involves large asset transfers (in the region of 500 to 2,000 USD-PPP) and requires intensive implementation raising questions about its scalability into existing large scale safety net programs. The SPIR intervention was designed as a light-touch graduation model program, embedded into the PSNP, one of the largest safety net programs in Africa. The main evaluation of the program documented positive impacts on productive assets, incomes, and financial inclusion but it did not result in similar transformation in households' economic trajectories as the BRAC model evaluated by Banerjee et al. (2015). However, the results presented here suggest that a light-touch graduation program can be highly effective in protecting households against droughts.

References

Abadie A, Athey S, Imbens GW, Wooldridge J. When should you adjust standard errors for clustering? *Quarterly Journal of Economics* 2022;138; 1–35.

Abatzoglou JT, Dobrowski SZ, Parks SA, Hegewisch KC. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific data* 2018;5; 1-12.

Abiona O, Koppensteiner MF. The impact of household shocks on domestic violence: Evidence from Tanzania. IZA Discussion Paper, vol. No. 11992. IZA – Institute of Labor Economics: Bonn, Germany; 2018.

Adhvaryu A, Nyshadham A, Molina T, Tamayo J. Helping children catch up: Early life shocks and the progresa experiment. *The Economic Journal* 2024;134; 1-22.

Aizer A. The gender wage gap and domestic violence. *American Economic Review* 2010;100; 1847-1859.

Alderman H, Billings L, Gilligan DO, Hidrobo M, Leight J, Taffesse AS, Tambet H. Impact evaluation of the strengthen PSNP4 institutions and resilience (SPIR) development food security activity (DFSA): Endline report. International Food Policy Research Institute (IFPRI): Washington D.C.; 2021.

Alderman H, Gilligan D, Hidrobo M, Leight J, Mulford M. Including Scalable Nutrition Interventions in a Graduation Model Program: Evidence from an Impact Evaluation in Ethiopia. IFPRI Discussion Paper, vol. 02208. International Food Policy Research Institute (IFPRI): Washington D.C.; 2023.

Alderman H, Gilligan DO, Hidrobo M, Leight J, Tambet H. Short-term evidence on wellbeing of rural Ethiopian households during the COVID-19 pandemic. SPIR learning brief. International Food Policy Research Institute (IFPRI): Washington D.C.; 2020.

Alderman H, Hoddinott J, Kinsey B. Long term consequences of early childhood malnutrition. *Oxford Economic Papers-New Series* 2006;58; 450-474.

Andrews C, Hsiao A, Ralston L. Social Safety Nets Promote Poverty Reduction, Increase Resilience, and Expand Opportunities. In: Beegle K, Coudouel A, Monsalve E (Eds), Realizing the Full Potential of Social Safety Nets in Africa. The World Bank: Washington, D.C.; 2018.

Asfaw S, Carraro A, Davis B, Handa S, Seidenfeld D. Cash transfer programmes, weather shocks and household welfare: evidence from a randomised experiment in Zambia. *Journal of Development Effectiveness* 2017;9; 419-442.

Balboni C, Bandiera O, Burgess R, Ghatak M, Heil A. Why do people stay poor? *The Quarterly Journal of Economics* 2022;137; 785-844.

Bandiera O, Burgess R, Das N, Gulesci S, Rasul I, Sulaiman M. Labor markets and poverty in village economies. *The Quarterly Journal of Economics* 2017;132; 811-870.

Banerjee A, Duflo E, Goldberg N, Karlan D, Osei R, Parienté W, Shapiro J, Thuysbaert B, Udry C. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* 2015;348; 1260799.

Banerjee AV, Chassang S, Snowberg E. Decision theoretic approaches to experiment design and external validity. Handbook of Economic Field Experiments, vol. 1. Elsevier; 2017. p. 141-174.

Barrios T, Diamond R, Imbens GW, Kolesár M. Clustering, spatial correlations, and randomization inference. *Journal of the American Statistical Association* 2012;107; 578-591.

Beaman L, Karlan D, Thuysbaert B. Saving for a (not so) rainy day: A randomized evaluation of savings groups in Mali. NBER Working Paper, vol. 20600. National Bureau of Economic Research (NBER); 2014.

Beegle K, Christiaensen L, Dabalen A, Gaddis I. Poverty in a rising Africa. The World Bank; 2016.

Beegle K, Coudouel A, Monsalve E. Realizing the Full Potential of Social Safety Nets in Africa. The World Bank: Washington D.C.; 2018.

Beguería S, Vicente-Serrano SM. SPEI: calculation of the standardised precipitation-evapotranspiration index. *R package version* 2017;1.

Berhane G, Gilligan DO, Hoddinott J, Kumar N, Taffesse AS. Can Social Protection Work in Africa? The Impact of Ethiopia's Productive Safety Net Programme. *Economic Development and Cultural Change* 2014;63; 1-26.

Björkman-Nyqvist M. Income shocks and gender gaps in education: Evidence from Uganda. *Journal of Development Economics* 2013;105; 237-253.

Bossuroy T, Goldstein M, Karimou B, Karlan D, Kazianga H, Parienté W, Premand P, Thomas CC, Udry C, Vaillant J. Tackling psychosocial and capital constraints to alleviate poverty. *Nature* 2022;605; 291-297.

Boucher SR, Carter MR, Flatnes JE, Lybbert TJ, Malacarne JG, Mareyna PP, Paul LA. Bundling Genetic and Financial Technologies for More Resilient and Productive Small-Scale Farmers in Africa. *The Economic Journal* 2024; ueae012.

Brune L, Karlan D, Kurdi S, Udry C. Social protection amidst social upheaval: Examining the impact of a multi-faceted program for ultra-poor households in Yemen. *Journal of Development Economics* 2022;155; 102780.

Buller AM, Peterman A, Ranganathan M, Bleile A, Hidrobo M, Heise L. A mixed-method review of cash transfers and intimate partner violence in low-and middle-income countries. *The World Bank Research Observer* 2018;33; 218-258.

Carpena F. How do droughts impact household food consumption and nutritional intake? A study of rural India. *World Development* 2019;122; 349-369.

Carter M, de Janvry A, Sadoulet E, Sarris A. Index insurance for developing country agriculture: a reassessment. *Annual Review of Resource Economics* 2017;9; 421-438.

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

Christian P, Kandpal E, Palaniswamy N, Rao V. Safety nets and natural disaster mitigation: evidence from cyclone Phailin in Odisha. *Climatic Change* 2019;153; 141-164.

Colella F, Lalive R, Sakalli SO, Thoenig M. Inference with arbitrary clustering. IZA Discussion Paper, vol. 12584. 2019.

Collins D, Morduch J, Rutherford S, Ruthven O. Portfolios of the Poor: How the World's Poor Live on $2 a Day. Princeton University Press: Princeton, NJ; 2009.

Conley TG. GMM estimation with cross sectional dependence. *Journal of Econometrics* 1999;92; 1-45.

Cooper MW, Brown ME, Hochrainer-Stigler S, Pflug G, McCallum I, Fritz S, Silva J, Zvoleff A. Mapping the effects of drought on child stunting. *Proceedings of the National Academy of Sciences* 2019;116; 17219-17224.

Corno L, Hildebrandt N, Voena A. Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica* 2020;88; 879-915.

Couttenier M, Soubeyran R. Drought and civil war in sub-saharan africa. *The Economic Journal* 2014;124; 201-244.

Crosta T, Karlan D, Ong F, Ruschenpohler J, Udry C. Unconditional Cash Transfers: A Bayesian Meta-Analysis of 50 Randomized Evaluations in 26 Low and Medium Income Countries. Unpublished manuscript. 2023.

Cui X, Tang Q. Extreme heat and rural household adaptation: Evidence from Northeast China. *Journal of Development Economics* 2024;167; 103243.

De Janvry A, Finan F, Sadoulet E, Vakis R. Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks? *Journal of Development Economics* 2006;79; 349-373.

De Waal A. Mass starvation: The history and future of famine. Polity Press: Cambridge, UK; 2017.

Deaton A, Grosh M. Consumption. In: Grosh M, Glewwe P (Eds), Designing Household Survey Questionaires for Developing Countries: Lessons from 15 Years of Living Standards Measurement Study, vol. 1. World Bank: Washington D.C.; 2000. p. 91-133.

Dercon S. Growth and shocks: evidence from rural Ethiopia. *Journal of Development Economics* 2004;74; 309-329.

Dercon S, Krishnan P. In sickness and in health: Risk sharing within households in rural Ethiopia. *Journal of Political Economy* 2000;108; 688-727.

Dercon S, Porter C. Live Aid Revisited: Long-term Impacts of the 1984 Ethiopian Famine on Children. *Journal of the European Economic Association* 2014;12; 927-948.

Díaz J-J, Saldarriaga V. A drop of love? Rainfall shocks and spousal abuse: evidence from rural Peru. *Journal of Health Economics* 2023;89; 102739.

Dinkelman T. Long-run health repercussions of drought shocks: evidence from South African homelands. *The Economic Journal* 2017;127; 1906-1939.

Emerick K, De Janvry A, Sadoulet E, Dar MH. Technological innovations, downside risk, and the modernization of agriculture. *American Economic Review* 2016;106; 1537-1561.

Epstein A, Bendavid E, Nash D, Charlebois ED, Weiser SD. Drought and intimate partner violence towards women in 19 countries in sub-Saharan Africa during 2011-2018: a population-based study. *PLoS Medicine* 2020;17; e1003064.

Fafchamps M, Udry C, Czukas K. Drought and saving in West Africa: are livestock a buffer stock? *Journal of Development Economics* 1998;55; 273-305.

FAO. FAO Locust Hub FAO's Desert Locust Information Service. Food and Agriculture Organization of the United Nations (FAO): Rome; 2023.

FAO, FHI 360. Minimum Dietary Diversity for Women: A Guide to Measurement. Food and Agriculture Organization (FAO) of the United Nations and USAID's Food and Nutrition Technical Assistance III Project (FANTA), managed by FHI 360: Rome; 2016.

Fisher AC, Hanemann WM, Roberts MJ, Schlenker W. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather: comment. *The American Economic Review* 2012;102; 3749-3760.

Fox GL, Benson ML, DeMaris AA, Van Wyk J. Economic distress and intimate violence: Testing family stress and resources theories. *Journal of Marriage and Family* 2002;64; 793-807.

Fruttero A, Halim D, Broccolini C, Coelho B, Gninafon H, Muller N. Gendered Impacts of Climate Change: Evidence from Weather Shocks. World Bank Policy Research Working Papers, vol. 10442. The World Bank: Washington D.C.; 2023.

Gray CL, Mueller V. Drought and Population Mobility in Rural Ethiopia. *World Development* 2012;40; 134-145.

Haile GG, Tang Q, Sun S, Huang Z, Zhang X, Liu X. Droughts in East Africa: Causes, impacts and resilience. *Earth-Science Reviews* 2019;193; 146-161.

Harari M, La Ferrara E. Conflict, Climate and Cells: A disaggregated analysis. *Review of Economics and Statistics* 2018;100; 594-608.

Haushofer J, Ringdal C, Shapiro JP, Wang XY. Income changes and intimate partner violence: Evidence from unconditional cash transfers in Kenya. National Bureau of Economic Research; 2019.

Heath R, Hidrobo M, Roy S. Cash transfers, polygamy, and intimate partner violence: Experimental evidence from Mali. *Journal of Development Economics* 2020;143; 102410.

Hidrobo M, Hoddinott J, Kumar N, Olivier M. Social Protection, Food Security, and Asset Formation. *World Development* 2018;101; 88-103.

Hirvonen K, Machado E, Simons A, Taraz V. Ethiopia's flagship public works program increases tree cover in rural areas and sloped terrains. *Global Environmental Change* 2022;75.

Hirvonen K, Sohnesen TP, Bundervoet T. Impact of Ethiopia's 2015 drought on child undernutrition. *World Development* 2020;131.

Hoddinott J, Berhane G, Gilligan DO, Hirvonen K, Kumar N, Lind J, Sabates-Wheeler R, Taffesse AS. Securing food, building livelihoods? A 15-year appraisal of Ethiopia's Productive Safety Net Programme. UNU-WIDER Working Paper. UNU-WIDER: Helsinki; 2024.

Hoddinott J, Kinsey B. Adult health in the time of drought. FCND Discussion paper, vol. 79. International Food Policy Research Institute (IFPRI): Washington D.C.; 2000.

Hoddinott J, Kinsey B. Child Growth in the Time of Drought. *Oxford Bulletin of Economics and Statistics* 2001;63; 409-436.

Hou X. Can drought increase total calorie availability? The impact of drought on food consumption and the mitigating effects of a conditional cash transfer program. *Economic Development and Cultural Change* 2010;58; 713-737.

IMF. Chapter 2. Adapting to Climate Change in Sub-Saharan Africa. In: (IMF) IMF (Ed), Regional Economic Outlook Sub-Saharan Africa. International Monetary Fund (IMF): Washington, D.C.; 2020.

Jahnke HE. Livestock production systems and livestock development in tropical Africa. Kieler Wissenschaftsverlag Vauk Kiel: Germany; 1982.

Jensen N, Barrett C. Agricultural index insurance for development. *Applied Economic Perspectives and Policy* 2017;39; 199-219.

Karlan D, Osei R, Osei-Akoto I, Udry C. Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics* 2014;129; 597-652.

Karlan D, Savonitto B, Thuysbaert B, Udry C. Impact of savings groups on the lives of the poor. *Proceedings of the National Academy of Sciences* 2017;114; 3079-3084.

Kazianga H, Udry C. Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso. *Journal of Development economics* 2006;79; 413-446.

Knippenberg E, Hoddinott J. Shocks, social protection, and resilience: Evidence from Ethiopia. ESSP Working Paper vol. 109. International Food Policy Research Institute (IFPRI): Addis Ababa; 2017.

Lane G. Adapting to climate risk with guaranteed credit: Evidence from Bangladesh. *Econometrica* 2024;92; 355-386.

Lange S, Reimers M. Livestock as a buffer stock in poorly integrated markets. *Economic Development and Cultural Change* 2021;69; 727-764.

Leight J, Alderman H, Gilligan D, Hidrobo M, Mulford M. Can a light-touch graduation model enhance livelihood outcomes? Evidence from Ethiopia. IFPRI Discussion Paper, vol. 2203. International Food Policy Research Institute (IFPRI): Washington D.C.; 2023.

Leight J, Gilligan DO, Mulford M, Taffesse AS, Tambet H. Aspiring to more? New evidence on the effect of a light-touch aspirations intervention in rural Ethiopia. IFPRI Discussion Paper, vol. 2070. International Food Policy Research Institute (IFPRI): Washington D.C.; 2021.

Leight J, Hirvonen K, Zafar S. The effectiveness of cash and cash plus interventions on livelihoods outcomes: Evidence from a systematic review and meta-analysis. IFPRI Discussion Paper, vol. 2262. International Food Policy Research Institute (IFPRI): Washington D.C.; 2024.

Lipper L, McCarthy N, Zilberman D, Asfaw S, Branca G. Climate smart agriculture: building resilience to climate change. Springer Nature; 2017.

Lloyd-Hughes B. The impracticality of a universal drought definition. *Theoretical and Applied Climatology* 2014;117; 607-611.

Maccini S, Yang D. Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review* 2009;99; 1006-1026.

Macours K, Premand P, Vakis R. Transfers, Diversification and Household Risk Strategies: Can productive safety nets help households manage climatic variability? *The Economic Journal* 2022;132; 2438–2470.

Maystadt J-F, Calderone M, You L. Local warming and violent conflict in North and South Sudan. *Journal of Economic Geography* 2015;15; 649-671.

McKee TB, Doesken NJ, Kleist J. The relationship of drought frequency and duration to time scales. Proceedings of the 8th Conference on Applied Climatology, vol. 17. California; 1993. p. 179-183.

McKenzie D. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics* 2012;99; 210-221.

Mekuriaw Z, Harris-Coble L. Ethiopia's Livestock Systems: Overview and Areas of Inquiry. In: Systems FtFILfL (Ed). Gainesville, FL, USA; 2021.

Michler JD, Josephson AL. To specialize or diversify: Agricultural diversity and poverty dynamics in Ethiopia. *World Development* 2017;89; 214-226.

NDRMC. Ethiopia: Humanitarian Requirements Document 2016. Joint Government and Humanitarian Partners' Document. National Disaster Risk Management Commission (NDRMC): Addis Ababa; 2016.

Pomeranz D, Kast F. Savings accounts to borrow less: experimental evidence from Chile. *Journal of Human Resources* 2024;59; 70-108.

Pople A, Hill R, Dercon S, Brunckhorst B. Anticipatory cash transfers in climate disaster response. CSAEWorkingPaper, vol. WPS/202107. Oxford; 2021.

Premand P, Stoeffler Q. Do cash transfers foster resilience? Evidence from rural Niger. World Bank Policy Research Working Paper. The World Bank: Washington D.C.; 2020.

Price J, Warren R, Forstenhäusler N, Wallace C, Jenkins R, Osborn TJ, Van Vuuren D. Quantification of meteorological drought risks between 1.5° C and 4° C of global warming in six countries. *Climatic Change* 2022;174; 12.

Raleigh C, Linke A, Hegre H, Karlsen J. Introducing ACLED: an armed conflict location and event dataset: special data feature. *Journal of Peace Research* 2010;47; 651-660.

Rosenzweig M, Udry C. External validity in a stochastic world. *The Review of Economic Studies* 2019;87; 343–381.

Rosenzweig MR, Wolpin KI. Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in India. *Journal of Political Economy* 1993; 223-244.

Roy S, Hidrobo M, Hoddinott J, Ahmed A. Transfers, behavior change communication, and intimate partner violence: Postprogram evidence from rural Bangladesh. *Review of Economics and Statistics* 2019;101; 865-877.

Sabates-Wheeler R, Lind J, Hoddinott J, Tefera Taye M. Graduation after 10 years of Ethiopia's Productive Safety Net Programme: Surviving but still not thriving. *Development Policy Review* 2021;39; 511-531.

Seneviratne SI, Zhang X, Adnan M, Badi W, Dereczynski C, Luca AD, Ghosh S, Iskandar I, Kossin J, Lewis S, Otto F, Pinto I, Satoh M, Vicente-Serrano SM, Wehner M, Zhou B. Weather and Climate Extreme Events in a Changing Climate. In: Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, Zhou B (Eds), Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA; 2021. p. 1513–1766.

Shively GE. Infrastructure mitigates the sensitivity of child growth to local agriculture and rainfall in Nepal and Uganda. *Proceedings of the National Academy of Sciences* 2017;114; 903-908.

Taffesse AS, Dorosh P, Gemessa SA. Crop Production in Ethiopia. In: Dorosh P, Rashid S (Eds), Food and Agriculture in Ethiopia. University of Pennsylvania Press: Philadelphia; 2012.

Thiede BC. Rainfall shocks and within-community wealth inequality: evidence from rural Ethiopia. *World Development* 2014;64; 181-193.

Toulmin C. Drought and the farming sector: Loss of farm animals and post-drought rehabilitation. *ILCA ALPAN Network Paper* 1986.

van Daalen K, Jung L, Dhatt R, Phelan AL. Climate change and gender-based health disparities. *The Lancet Planetary Health* 2020;4; e44-e45.

Van Loon AF. Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water* 2015;2; 359-392.

Vicente-Serrano SM, Beguería S, López-Moreno JI. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate* 2010;23; 1696-1718.

Vicente-Serrano SM, Beguería S, Lorenzo-Lacruz J, Camarero JJ, López-Moreno JI, Azorin-Molina C, Revuelto J, Morán-Tejeda E, Sanchez-Lorenzo A. Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interactions* 2012;16; 1-27.

Webb D. Critical Periods in Cognitive and Socioemotional Development: Evidence from Weather Shocks in Indonesia. *The Economic Journal* 2023;134; 1637–1665.

Webb P, von Braun J. Famine and food security in Ethiopia: lessons for Africa. John Wiley & Sons Ltd: UK; 1994.

WHO. Ethical and safety recommendations for intervention research on violence against women: building on lessons from the WHO publication putting women first: ethical and safety recommendations for research on domestic violence against women. World Health Organization (WHO): Geneva; 2016.

Wiseman W, Van Domelen J, Coll-Black S. Designing and implementing a rural safety net in a low income setting: Lessons learned from Ethiopia's Productive Safety Net Program 2005-2009. The World Bank: Washington D.C.; 2010.

World Bank. World Bank Group gender strategy (FY16-23): Gender equality, poverty reduction and inclusive growth. The World Bank Group: Washington D.C.; 2015.

World Bank. The State of Social Safety Nets 2018. World Bank: Washington D.C.; 2018.

World Bank. Climate Risk Profile: Ethiopia. The World Bank Group: Washington D.C.; 2021.

You L, Ringler, C., Wood-Sichra U, Robertson R, Wood S, Zhou T, Nelson G. What Is the Irrigation Potential for Africa? A Combined Biophysical and Socioeconomic Approach. *Food Policy* 2012;36; 770-782.

Tables and Figures

Figure 1. SPIR interventions and surveys

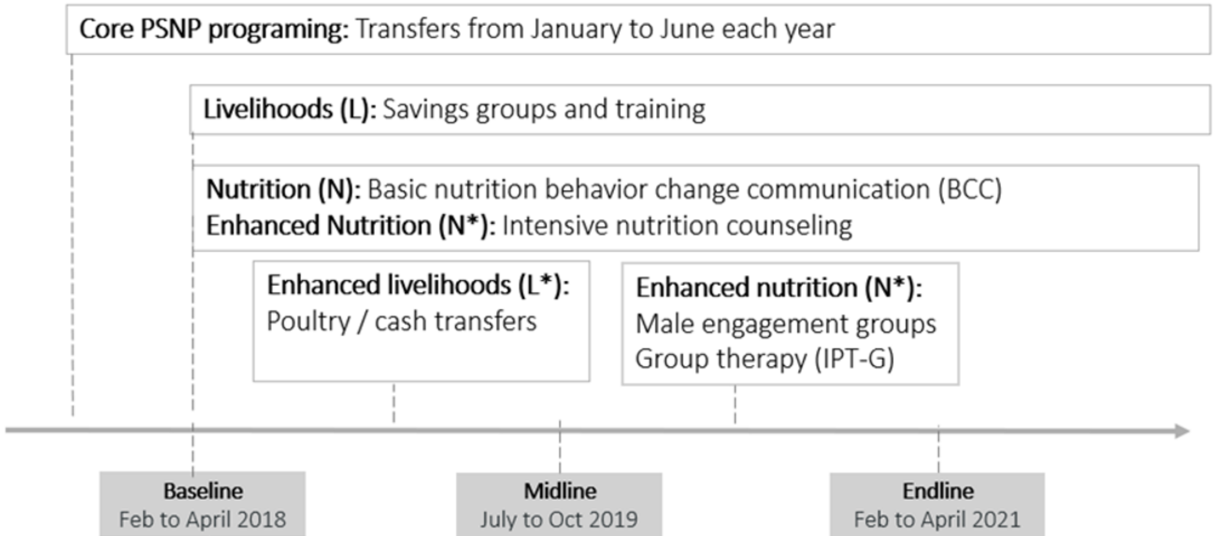
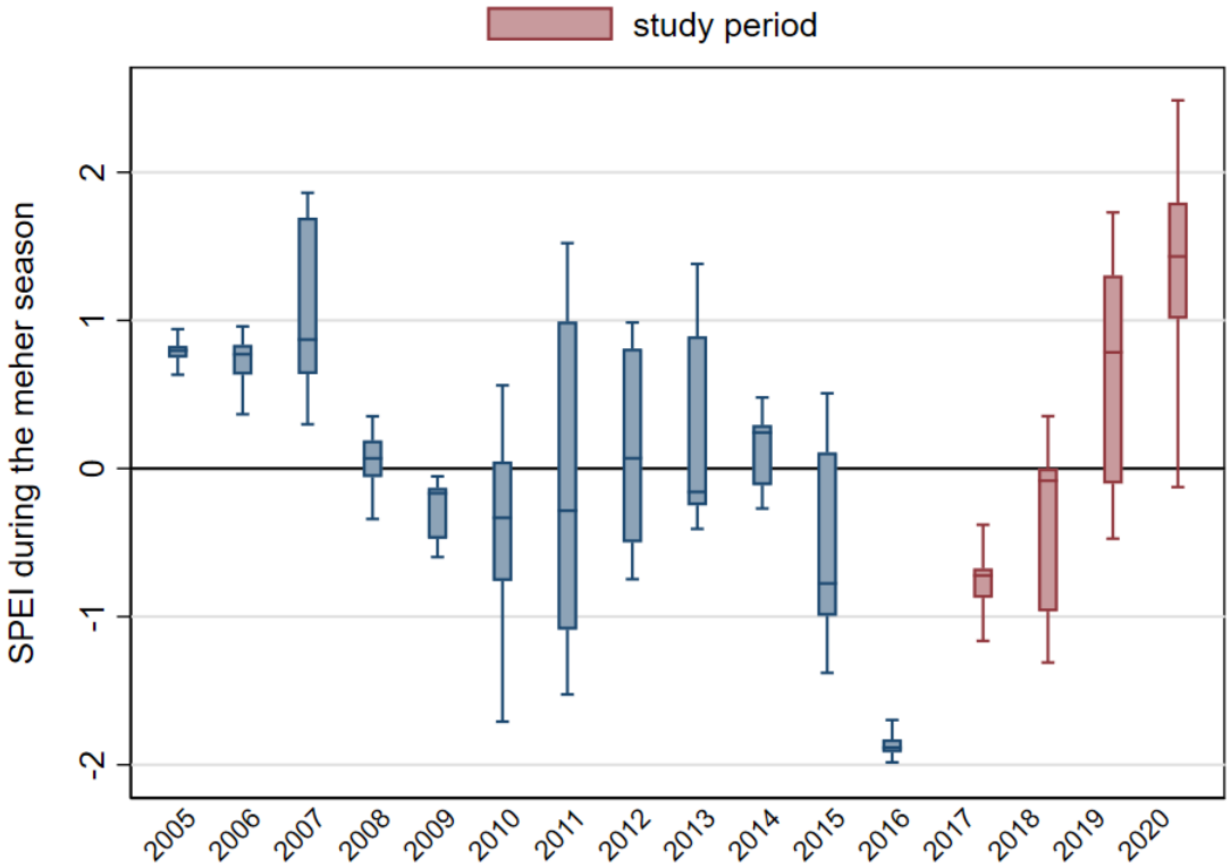


Figure 2. SPEI during the meher seasons in 2005–2020



Note: Box plot. The size of the box indicates the difference between the 25th percentile (the bottom part of the box) and the 75th percentile (the top part of the box) of the Standardized Precipitation-Evapotranspiration Index (SPEI) distribution. The bottom and top rule marks the bottom 5th and top 5th percentiles, respectively. The vertical bar rule inside the box shows the median value. N = 192 clusters (kebeles).

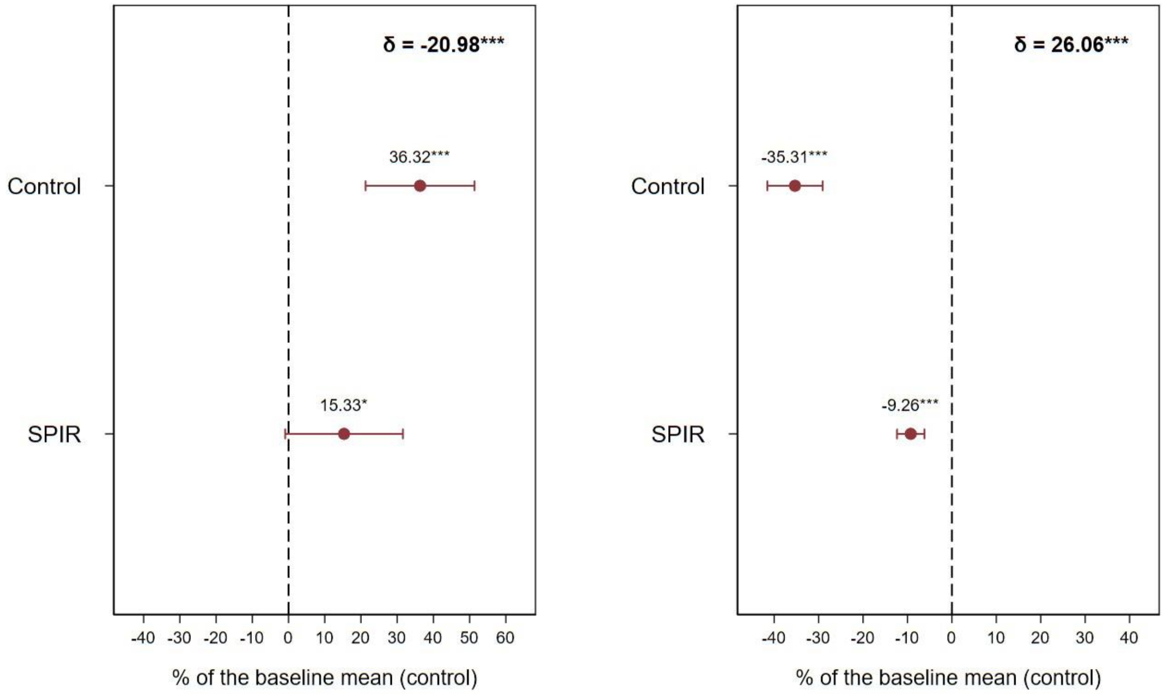
Table 1. Impact of graduation programming and drought conditions on household and individual level outcomes

	(1)	(2)	(3)	(4)	(5)
Outcome:	Household food gap	Tropical livestock units owned by household	Primary woman's diet diversity	BMI of primary woman	Primary woman experienced IPV
Drought	2.997***	-1.375***	-0.706	-0.158*	0.144***
	(0.633)	(0.123)	(0.507)	(0.094)	(0.056)
Drought X SPIR	-1.732***	1.015***	1.294***	0.213***	-0.161***
	(0.453)	(0.104)	(0.168)	(0.038)	(0.048)
SPIR	0.067	-0.046***	0.033	-0.032	0.007
	(0.061)	(0.017)	(0.053)	(0.039)	(0.006)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	2.06	0.97	2.02	20.12	0.17
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.064	p < 0.001	p = 0.232	p = 0.489	p = 0.637
Normalized drought impact in control households	0.75***	-0.34***	-0.18	-0.04*	0.04***
As % of the baseline control mean	*36.32****	*-35.31****	*-8.72*	*-0.20**	*20.67****
Normalized drought impact in SPIR households	0.32*	-0.09***	0.15	0.01	-0.00
As % of the baseline control mean	*15.33**	*-9.26****	*7.26*	*0.07*	*-2.42*
Number of observations	6,212	6,274	5,877	4,888	3,808

Note: BMI = Body mass index. IPV = Intimate Partner Violence. ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. In column 5, the estimated equation includes a binary variable capturing households for which the baseline value was missing (these missing baseline values were set to zero). The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * p < 0.10, ** p < 0.05, *** p < 0.01. Normalized drought impact refers to the estimated drought impact to a 0.25 SD increase in relative dryness.

Figure 3. Impact of a 0.25 SD increase in relative dryness on household level outcomes, by treatment status

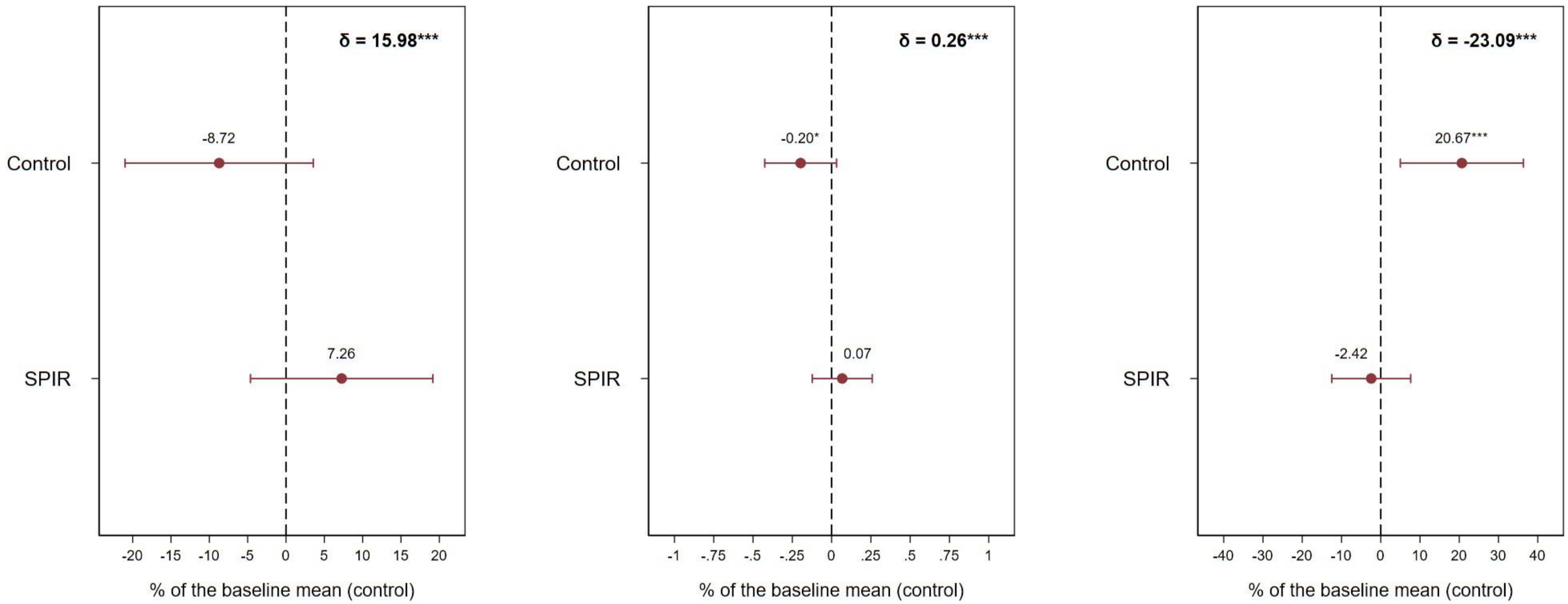
A. Food gap **B. Tropical livestock units**



Note: δ = difference between the estimates (coefficient on the interaction term). Based on equation 1 with underlying regression results reported in column 1 and 2 of Table 1. The capped lines represent 95-% confidence intervals of the point estimates (marked with a solid dot). The number of observations in Panel A is 6,212 and in Panel B 6,274. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

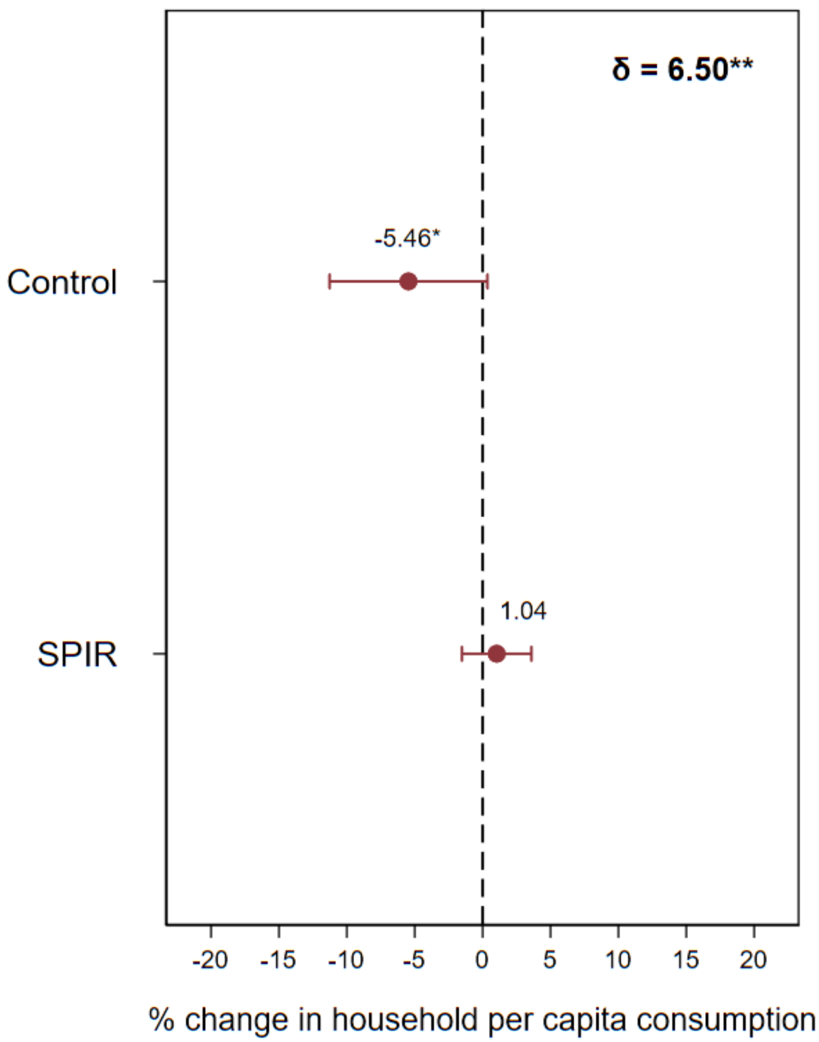
Figure 4. Impact of a 0.25 SD increase in relative dryness on individual level outcomes, by treatment status

A. Diet diversity **B. Body-mass index** **C. Intimate Partner Violence**



Note: δ = difference between the estimates (coefficient on the interaction term). Based on equation 1 with underlying regression results reported in columns 3-5 of Table 1. The capped lines represent 95-% confidence intervals of the point estimates (marked with a solid dot). The number of observations in Panel A is 5,877, in Panel B 4,888 and Panel C 3,808. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 5. Impact of a 0.25 SD increase in relative dryness in 2019-meher on household per capita consumption in 2021, by treatment status



Note: δ = difference between the estimates (coefficient on the interaction term). The underlying regression result are reported in Table H2. The outcome variable is log household consumption in adult equivalents. The capped lines represent 95-% confidence intervals of the point estimates (marked with a solid dot). Endline data only: the number of observations is 2,962. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2. Impact of graduation programming on agricultural investments and practices

Outcome:	(1) Seeds (0/1)	(2) Inorganic fertilizer (0/1)	(3) Pesticides (0/1)	(4) Hired labor (0/1)	(5) Crop diversity (count)
SPIR	0.003	0.031	-0.028	-0.004	-0.035
	(0.032)	(0.031)	(0.027)	(0.018)	(0.056)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey rounds	ML	ML & EL	ML & EL	ML & EL	ML & EL
Survey fixed effects?	No	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome variable in the control group	0.229	0.357	0.037	0.078	2.06
Number of observations	2,574	5,077	2517	5,087	5,124
Outcome:	**(6) Seeds ($PPP)**	**(7) Inorganic fertilizer ($PPP)**	**(8) Pesticides ($PPP)**	**(9) Hired labor ($PPP)**	
SPIR	1.67	-0.12	-0.50	-3.16	
	(1.71)	(2.19)	(0.54)	(11.14)	
Woreda fixed effects?	Yes	Yes	Yes	Yes	
Survey rounds	ML	ML	ML	ML	
Survey fixed effects?	No	No	No	No	
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	
Baseline mean of the outcome variable in the control group	7.27	21.77	0.44	3.85	
Number of observations	2574	2560	2517	2576	

Note: An ANCOVA specification. The sample is restricted to households that cultivated crops during the previous cropping season. The outcome variable in columns 1 to 4 is binary, obtaining value 1 if the household reported expenditures on the input category in the previous cropping season, zero otherwise. The outcome variable in column 5 is a count capturing the number of different crops cultivated by the household in the previous cropping season. The outcome variables in columns 6 to 9 are continuous, capturing the input expenditure amounts in 2017 $PPP. These outcomes were not measured at the endline. The standard errors are reported in parentheses and clustered at the kebele level (unit of treatment). ML = Midline survey; EL = Endline survey. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. Impact of graduation programming and drought conditions on household savings and livestock sales

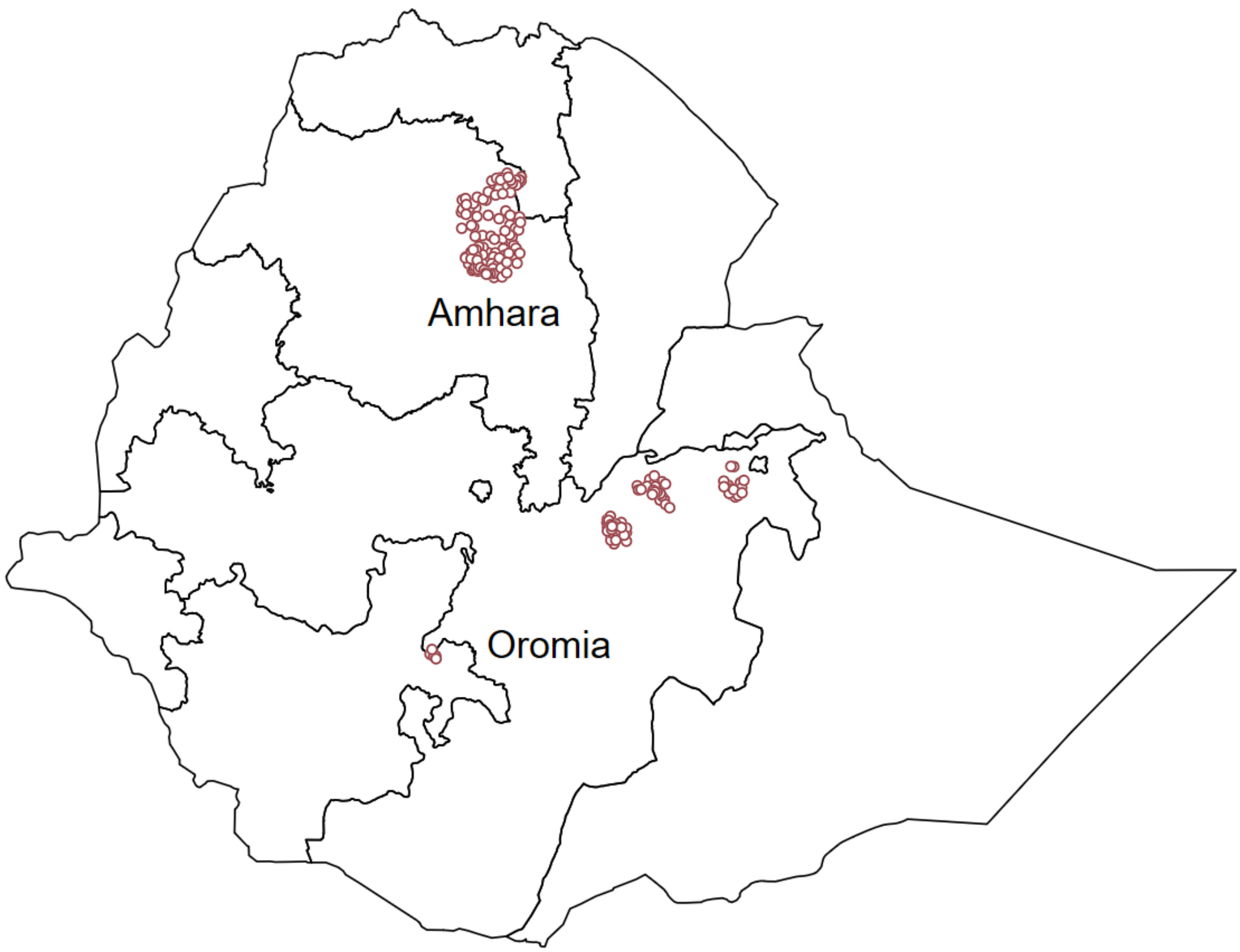
	(1)	(2)
Outcome:	Total amount saved ($PPP)	Livestock sales (TLU)
Drought	14.35	0.074
	(16.86)	(0.084)
Drought X SPIR	-52.89**	-0.290***
	(24.95)	(0.039)
SPIR	20.27***	0.018***
	(4.16)	(0.005)
Woreda fixed effects?	Yes	Yes
Survey round fixed effects?	Yes	Yes
Outcome variable at the baseline?	No	Yes
Standard errors?	Conley	Conley
Mean of the outcome in control group	42.53	0.245
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.049	0.008
Normalized drought impact in control households	3.59	0.02
As % of the control mean	8.44	7.59
Normalized drought impact in SPIR households	-9.64**	-0.05***
As % of the control mean	-22.66**	-22.07***
Number of observations	5,568	6,274

Note: Columns 1 is not based on an ANCOVA specification (savings were not asked at the baseline). ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * p < 0.10, ** p < 0.05, *** p < 0.01. Normalized drought impact refers to the estimated drought impact to a 0.25 SD increase in relative dryness.

Appendices

Appendix A. Map of the study locations

Figure A1. Location of study clusters



Note: 192 study clusters (kebeles): 112 in the Amhara region and 80 in the Oromia region of Ethiopia. Solid black lines represent regional or national boundaries.

Appendix B. Definition of IPV categories

Physical spousal violence: Husband/partner pushed you, shook you, or threw something at you; slapped you; twisted your arm or pulled your hair; punched you with his fist or with something that could hurt you; kicked you, dragged you, or beat you up; tried to choke you or burn you on purpose; or threatened or attacked you with a knife, gun, or any other weapon.

Sexual spousal violence: Husband/partner physically forced you to have sexual intercourse with him even when you did not want to; physically forced you to perform any other sexual acts you did not want to; forced you with threats or in any other way to perform sexual acts you did not want to.

Emotional spousal violence: Husband/partner said or did something to humiliate you in front of others; threatened to hurt or harm you or someone close to you; insulted you or made you feel bad about yourself.

Source: Alderman et al. (2021).

Appendix C. Additional drought regressions

Table C1. Impact of drought on cereal crop production

	(1) Cereal crop production ($PPP)	(2) Cereal crop production ($PPP)
Sample:	*Control clusters*	*SPIR clusters*
Drought	-105.95**	-107.50***
	(41.63)	(27.17)
Woreda fixed effects?	Yes	Yes
Survey round fixed effects?	Yes	Yes
Outcome variable at the baseline?	Yes	Yes
Mean of the outcome variable	445.06	465.07
Normalized drought impact	-26.50	-26.88
Number of observations	1,221	3,943

Note: ANCOVA specification with midline and endline survey rounds. The sample is restricted to households that reported cultivating crops during the cropping season. The variable 'Drought' is positive rectified SPEI (Standardised Precipitation-Evapotranspiration Index), multiplied by -1 so that larger positive values reflect increases in relative dryness (i.e., worse drought conditions). The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * p < 0.10, ** p < 0.05, *** p < 0.01. Normalized drought impact refers to the estimated drought impact to a 0.25 SD increase in relative dryness.

Table C2. Impact of drought on household and women's outcomes, control clusters only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Household food gap		Tropical livestock units		Primary woman's diet diversity		BMI of primary female		Experienced IPV	
Drought, ongoing meher for the midline	2.285***		-1.346***		-0.515**		-0.421		0.104*	
	(0.409)		(0.256)		(0.201)		(0.587)		(0.062)	
Drought, previous meher for the midline		0.678**		-0.277***		-0.056		-0.311***		0.026**
		(0.289)		(0.046)		(0.068)		(0.070)		(0.012)
Observations	1503	1503	1505	1505	1392	1392	1168	1168	905	905

Note: 'Drought' is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The regressions include baseline value of the outcome variable, survey round dummy and woreda fixed effects. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix D. Balance checks

Table D1. Balance in baseline outcomes and drought conditions across study arms

Variable	Study arm:	(1) T1 Mean/*N*	(2) T2 Mean/*N*	(3) T3 Mean/*N*	(4) Control Mean/*N*	t-test p-val (1)-(2)	t-test p-val (1)-(3)	t-test p-val (1)-(4)	t-test p-val (2)-(3)	t-test p-val (2)-(4)
Household food gap		2.094	2.200	2.386	2.063	0.692	0.292	0.907	0.502	0.602
		810	*836*	*850*	*791*					
Tropical livestock units owned by household		0.912	0.912	0.884	0.973	0.997	0.735	0.488	0.705	0.438
		812	*857*	*853*	*791*					
Primary woman's diet diversity		2.103	1.944	2.091	2.024	0.079*	0.905	0.457	0.110	0.390
		775	*817*	*827*	*757*					
BMI of primary female		20.169	20.024	19.993	20.117	0.334	0.276	0.724	0.825	0.475
		766	*814*	*821*	*755*					
Experienced intimate partner violence (IPV)		0.153	0.195	0.182	0.174	0.234	0.422	0.490	0.746	0.554
		524	*564*	*549*	*499*					
Drought during 2017 meher season [a)]		0.798	0.829	0.809	0.822	0.545	0.821	0.651	0.696	0.903
		812	*858*	*853*	*791*					
Drought during 2018 meher season [a)]		0.389	0.453	0.436	0.454	0.531	0.650	0.531	0.868	0.989
		812	*858*	*853*	*791*					
Drought during 2019 meher season [a)]		0.083	0.065	0.081	0.078	0.489	0.946	0.861	0.552	0.606
		812	*858*	*853*	*791*					
Drought during 2020 meher season [a)]		0.008	0.009	0.008	0.006	0.903	0.989	0.830	0.889	0.749
		812	*858*	*853*	*791*					

Note: The value displayed for t-tests are p-values based on standard errors clustered at the cluster (kebele) level. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The numbers in italics are the number of non-missing observations. T1, T2 and T3 are SPIR treatment arms where T1 received the intensive livestock (L*) and intensive nutrition intervention (N*), T2 received the intensive livestock intervention (L*) and standard nutrition intervention (N), and T3 received standard livestock (L) and intensive nutrition intervention (N*). Control are households located in clusters that did not receive any SPIR interventions.
[a)] 'Drought' is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions.

Table D2. Balance in baseline characteristics with respect to drought shocks prior midline and endline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log household per capita consumption at the baseline	0.000						-0.001
	(0.001)						(0.001)
Household size at the baseline		-0.002*					-0.002*
		(0.001)					(0.001)
Household head had some education at the baseline			0.003*				0.003*
			(0.001)				(0.001)
Household head's age at the baseline				-0.000			0.000*
				(0.000)			(0.000)
Food gap (months) at the baseline					0.000		0.000
					(0.000)		(0.000)
Tropical livestock units owned at the baseline						-0.000	0.000
						(0.000)	(0.001)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of the outcome variable	0.042	0.042	0.042	0.042	0.042	0.042	0.042
Observations	6550	6628	6622	6624	6574	6626	6526

Note: OLS regression. Outcome variable is drought shock prior midline and endline. The drought shock variable is based on positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * p < 0.10, ** p < 0.05, *** p < 0.01.

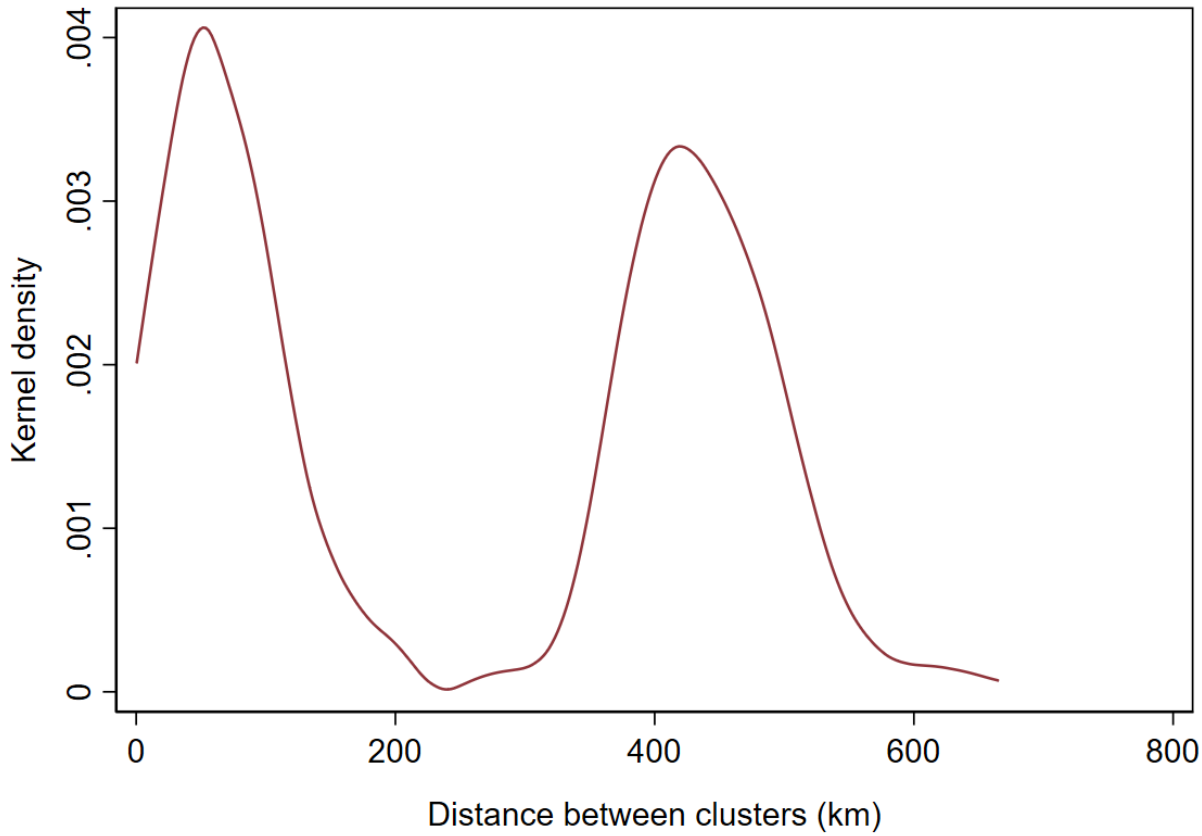
Appendix E. Exploring spatial correlation across study clusters

To explore spatial dependencies in our data, we first calculated the Euclidean distances between all cluster pairs. The mean distance is 256 km, and the median is 308 km. However, these averages hide the fact that the density distribution of the distances is bimodal (Figure E1). Given that our study clusters are in two regions, the cluster pairs are located either between 0 and 200 km or between 300 and 600 km from each other.

We expect that the weather outcomes in each agricultural season are highly correlated between clusters that are located close to each other.[30] To verify this, we computed the SPEI for each meher season between 1990 and 2021 for all clusters and used these data to calculate the correlation coefficient of the meher season SPEI between all cluster-pairs. We then used these dyadic data to explore the relationship between these correlation coefficients and distances between clusters. Figure E2 shows the results based on a local polynomial regression that regresses the correlation coefficient on the distance variable. As expected, the correlation coefficient is close to 1 for clusters located very close to each other. As we move to clusters that are farther apart, the correlation decreases. For clusters located 600 km apart from each other, the correlation coefficient is 0.2, on average, suggesting a relatively weak spatial correlation in meher season weather outcomes.

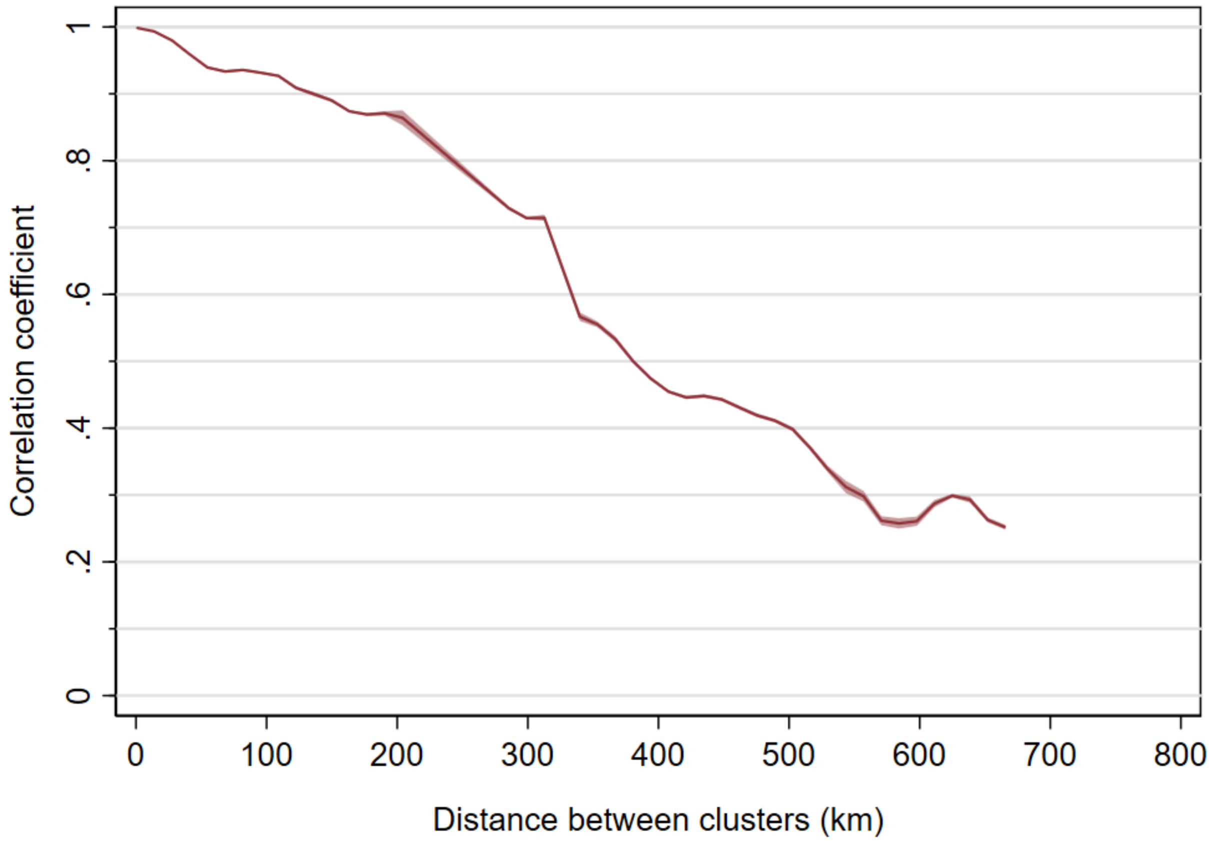
[30] Day-to-day weather outcomes can vary within small geographical areas and with high resolution climate data one may be able to detect this variation. However, when aggregated over an entire agricultural season, the spatial autocorrelation tends to be very high.

Figure E1. Density distribution of distances between study clusters



Note: Kernel density estimator. Based on dyadic data with 18,336 observations from 192 clusters (192*191/2=18,336). The vertical axis measures kernel density. The horizontal axis measures the Euclidean distance in km between the cluster dyads.

Figure E2. Relationship between meher season SPEI correlation coefficients and distance between study clusters



Note: Local polynomial regression. Shaded areas represent 95-% confidence intervals. Based on dyadic data with 18,336 observations from 192 clusters (192*191/2=18,336). The vertical axis measures the correlation coefficient in meher season SPEI in 1990-2021 between all possible cluster dyads. The horizontal axis measures the Euclidean distance in km between the cluster dyads.

Appendix F. Robustness checks

Table F1. Replicating Table 1, but using rainfall Z-score instead of SPEI to define droughts

	(1)	(2)	(3)	(4)	(5)
Outcome:	Household food gap	Tropical livestock units owned by household	Primary woman's diet diversity	BMI of primary woman	Primary woman experienced IPV
Drought	2.460***	-1.399***	-1.106***	-0.149	0.327***
	(0.281)	(0.478)	(0.318)	(0.099)	(0.092)
Drought X SPIR	-1.135	1.105***	0.963***	0.197***	-0.328***
	(0.723)	(0.347)	(0.360)	(0.052)	(0.114)
SPIR	0.021	-0.035*	0.062	-0.032	0.009
	(0.067)	(0.018)	(0.060)	(0.040)	(0.006)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	2.06	0.97	2.02	20.12	0.17
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.037	p = 0.048	p = 0.824	p = 0.530	p = 0.985
Number of observations	6,212	6,274	5,877	4,888	3,808

Note: ‘Drought’ is rainfall z-score multiplied by -1 and then negative rectified at zero so that larger (positive) values reflect worse rainfall conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * p < 0.10, ** p < 0.05, *** p < 0.01.

Table F2. Replicating Table 1, but using region fixed effects instead of woreda fixed effects

	(1)	(2)	(3)	(4)	(5)
Outcome:	Household food gap	Tropical livestock units owned by household	Primary woman's diet diversity	BMI of primary woman	Primary woman experienced IPV
Drought	3.743***	-1.290***	-0.908**	-0.075	0.189***
	(0.548)	(0.148)	(0.405)	(0.077)	(0.065)
Drought X SPIR	-1.685***	0.991***	1.186***	0.215***	-0.140***
	(0.510)	(0.114)	(0.229)	(0.036)	(0.035)
SPIR	0.039	-0.046***	0.052	0.014	0.002
	(0.050)	(0.016)	(0.057)	(0.046)	(0.007)
Region fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	2.06	0.97	2.02	20.12	0.17
‘Drought’ + ‘Drought X SPIR’ = 0	$p < 0.001$	$p = 0.001$	$p = 0.218$	$p = 0.054$	$p = 0.299$
Number of observations	6,212	6,274	5,877	4,888	3,808

Note: ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F3. Replicating Table 1, but using region-by-survey round fixed effects instead of survey round fixed effects

	(1)	(2)	(3)	(4)	(5)
Outcome:	Household food gap	Tropical livestock units owned by household	Primary woman's diet diversity	BMI of primary woman	Primary woman experienced IPV
Drought	1.530**	-1.369***	-0.386	-0.092	0.173**
	(0.661)	(0.136)	(0.610)	(0.336)	(0.068)
Drought X SPIR	-1.716***	1.015***	1.286***	0.213***	-0.162***
	(0.436)	(0.104)	(0.165)	(0.037)	(0.049)
SPIR	0.071	-0.046***	0.032	-0.033	0.007
	(0.062)	(0.017)	(0.053)	(0.038)	(0.006)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Region X survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	2.06	0.97	2.02	20.12	0.17
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.782	p < 0.001	p = 0.119	p = 0.696	p = 0.806
Number of observations	6212	6274	5877	4888	3808

Note: ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * p < 0.10, ** p < 0.05, *** p < 0.01.

Table F4. Replicating Table 1, but defining drought differently for the midline

	(1)	(2)	(3)	(4)	(5)
Outcome:	Household food gap	Tropical livestock units owned by household	Primary woman's diet diversity	BMI of primary woman	Primary woman experienced IPV
Midline drought conditions based on:	*2018 meher*	*2018 meher*	*2018 meher*	*2019 meher*	*2018 meher*
Drought	0.877***	-0.275***	-0.198**	-0.285	0.039*
	(0.296)	(0.030)	(0.083)	(0.612)	(0.021)
Drought X SPIR	-0.226*	0.172***	0.031	0.250	-0.035***
	(0.126)	(0.023)	(0.081)	(0.319)	(0.012)
SPIR	0.047	-0.042**	0.084	0.006	0.004
	(0.091)	(0.019)	(0.093)	(0.021)	(0.008)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	2.06	0.97	2.02	20.12	0.17
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.001	p < 0.001	p = 0.005	p = 0.917	p = 0.806
Number of observations	6,212	6,274	5,877	4,888	3,808

Note: ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The estimated equation includes a binary variable capturing households for which the baseline value was missing (these missing baseline values were set to zero). The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F5. Replicating Table 1, but controlling for locust and conflict shocks

	(1)	(2)	(3)	(4)	(5)
Outcome:	Household food gap	Tropical livestock units owned by household	Primary woman's diet diversity	BMI of primary woman	Primary woman experienced IPV
Drought	3.045***	-1.397***	-0.590	-0.136	0.126***
	(0.618)	(0.115)	(0.483)	(0.100)	(0.019)
Drought X SPIR	-1.737***	1.022***	1.266***	0.212***	-0.100***
	(0.453)	(0.106)	(0.168)	(0.038)	(0.028)
SPIR	0.067	-0.046***	0.035	-0.031	0.003
	(0.060)	(0.017)	(0.056)	(0.039)	(0.002)
Controls for locust and conflict shocks?	Yes	Yes	Yes	Yes	Yes
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	2.06	0.97	2.02	20.12	0.04
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.050	p < 0.000	p = 0.149	p = 0.370	p = 0.210
Number of observations	6,212	6,274	5,877	4,888	3,808

Note: ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The controls for locust and conflict are binary variables obtaining value 1 if there were conflict events or locust swarms within 20 km distance from the community, and zero otherwise. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * p < 0.10, ** p < 0.05, *** p < 0.01.

Table F6. Impact of graduation programming and drought conditions on household composition

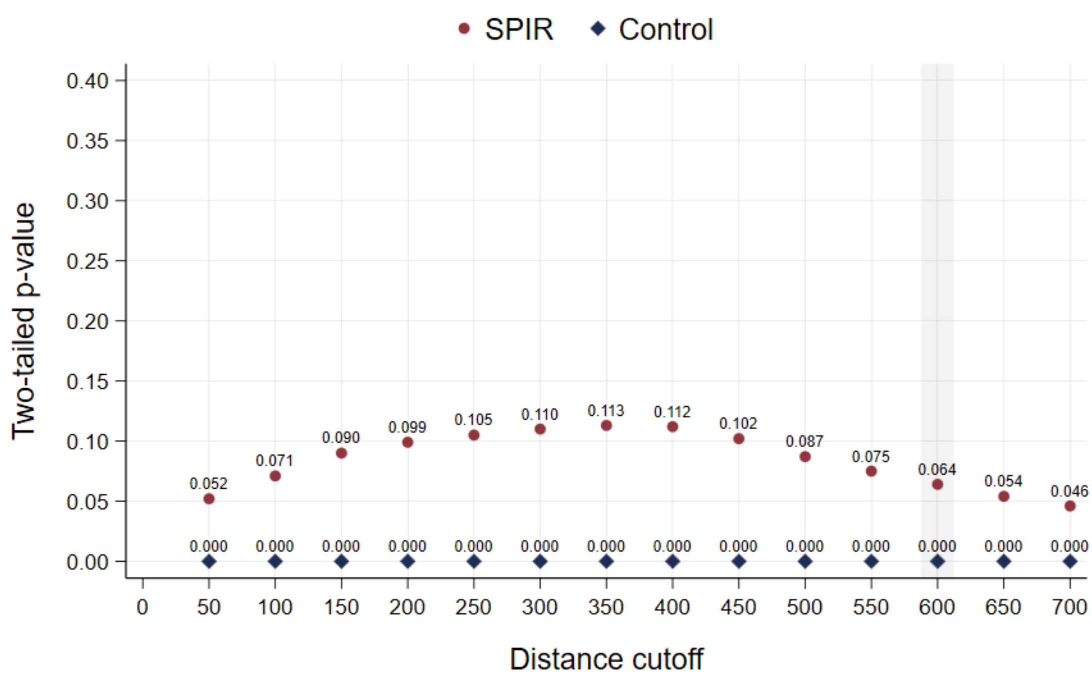
	(1)	(2)	(3)	(4)
Outcome:	Household size	Household size	Household size in adult equivalent units	Household size in adult equivalent units
Sample:	*Control households*	*All households*	*Control households*	*All households*
Drought	-0.039	-0.121	0.075	-0.132
	(0.545)	(0.357)	(0.404)	(0.255)
Drought X SPIR		-0.080		0.107
		(0.205)		(0.126)
SPIR		0.076		0.057*
		(0.053)		(0.032)
Woreda fixed effects?	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes
Baseline mean of the outcome variable in the control group	5.75	5.75	4.41	4.41
Number of observations	1,517	6,314	1,517	6,312

Note: 'Drought' is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

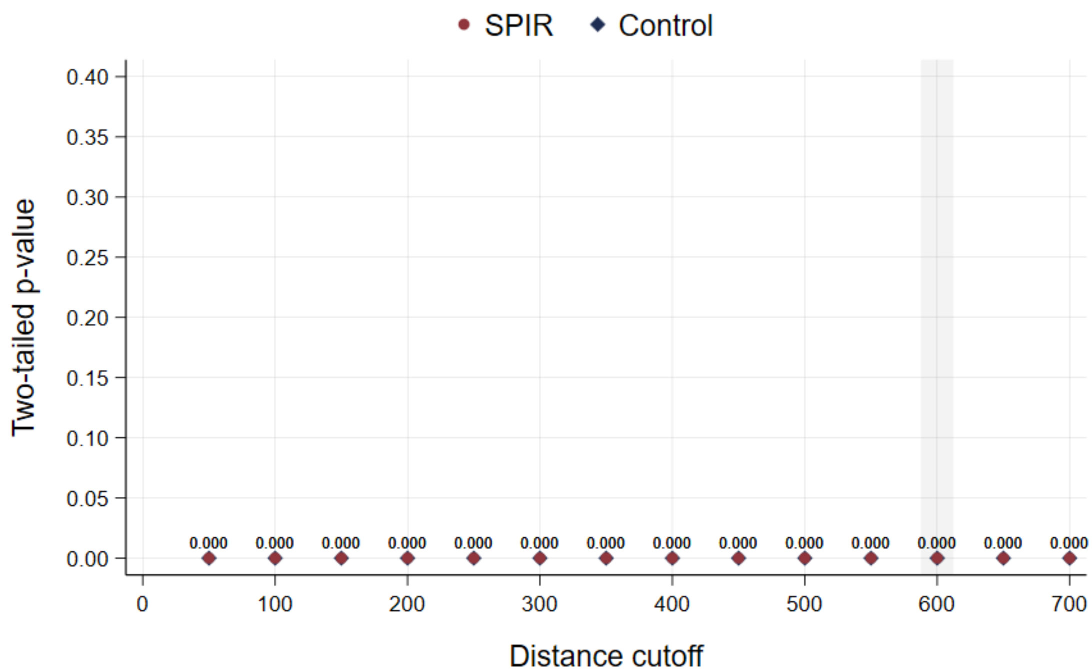
Figure F1. Impact of increase in relative dryness on household and individual level outcomes, by treatment status: p-values based on Conley (1999) standard errors with different distance cut-offs

Note: 'SPIR' is the p-value for the test of $\beta + \delta = 0$ in Eq. (1), capturing the effect of drought in SPIR households based on a joint hypothesis test. 'Control' is the p-value of $\beta = 0$ in Eq. (1), capturing the effect of drought in control households based on a single coefficient test. The shaded distance cut-off, 600 km, is used throughout the paper.

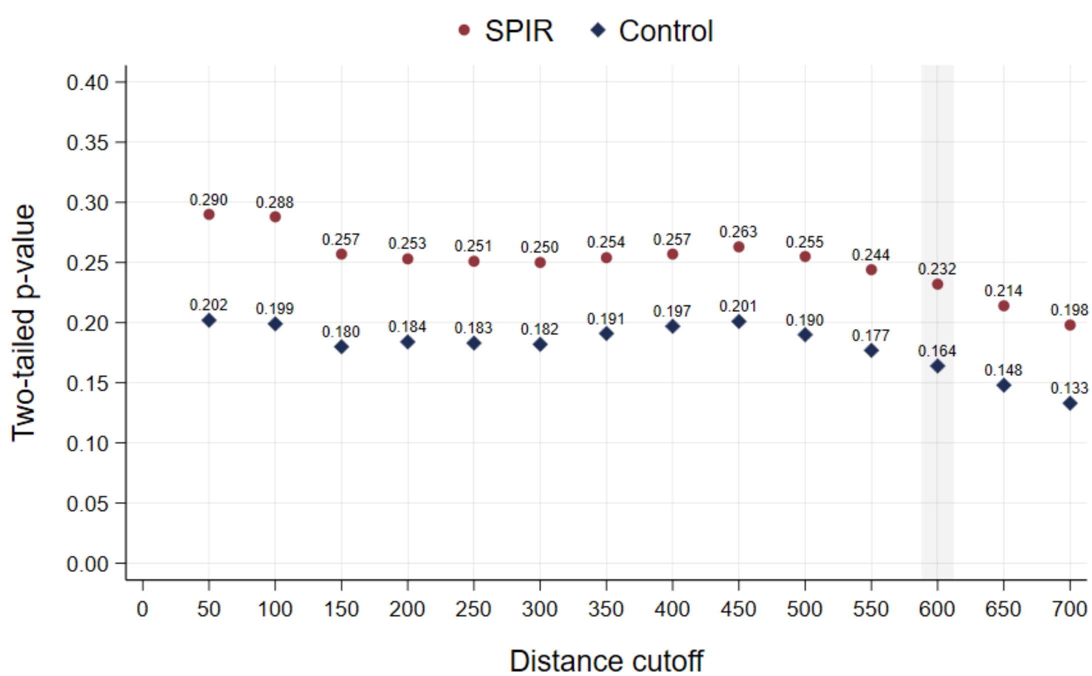
A) Household food gap



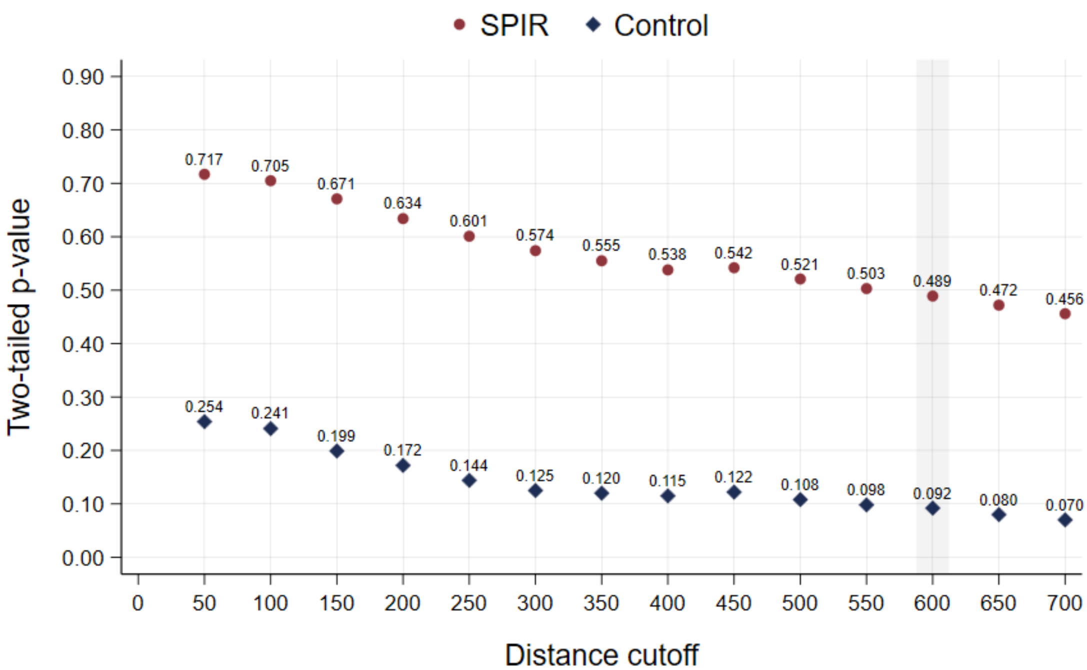
B) Tropical livestock units owned by household



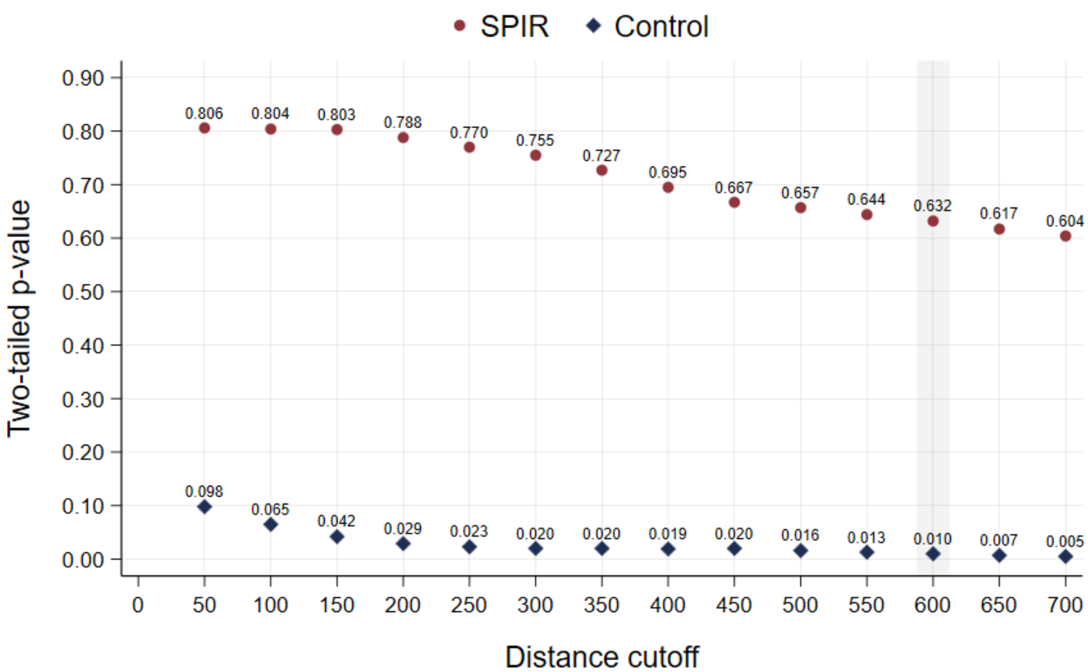
C) Primary woman's diet diversity



D) BMI of primary woman



E) Primary woman experienced IPV



Appendix G. Exposure to the 2016 drought

The droughts that occurred in 2015 and 2016 in Ethiopia were one of the most severe meteorological droughts the country had witnessed in many decades. In our study clusters, the drought in 2016 was particularly severe, raising a question of whether the findings documented here could be attributed to the households' exposure to the 2016 drought. We begin by acknowledging that nearly all the study clusters were similarly exposed to this drought (as shown in Figure 2), with the minimum and maximum SPEI value in our sample ranging between -1.7 and -1.9. Moreover, the local polynomial regressions reported in Figures G1 to G3 below indicate that SPEI value in 2016 meher season is not correlated with SPEI values during the study period. This is also evident when we look at the raw correlation coefficients: the correlation coefficient between 2016 meher SPEI and 2018 meher SPEI is 0.068; between 2016 and 2019 is 0.135, and between 2016 and 2020 is 0.048. Therefore, it is unlikely that the exposure to the 2016 drought is driving our results. This is further confirmed in Table G1 where we replicate Table 1 by controlling for the 2016 drought conditions in the study cluster: the estimated β, ϑ and δ coefficients are very similar to those reported in Table 1.

Figure G1. The relationship between SPEI in 2018 meher and SPEI in 2016 meher

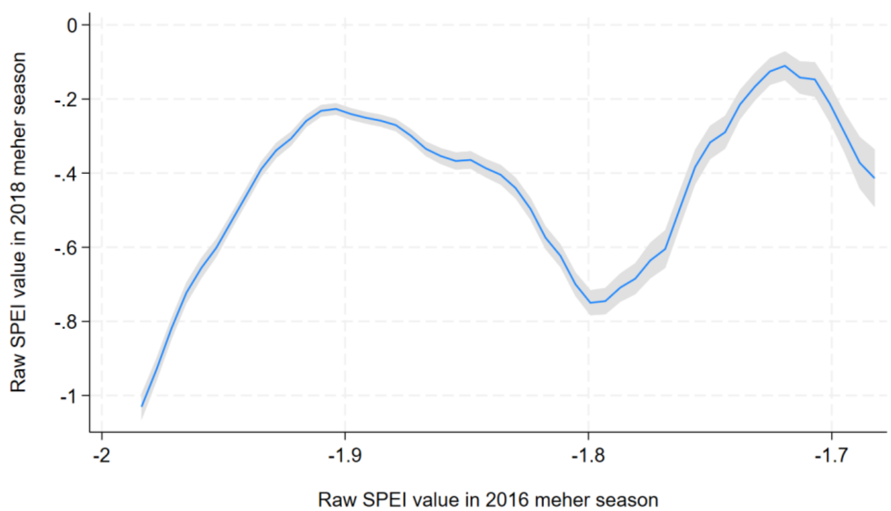


Figure G2. The relationship between SPEI in 2019 meher and SPEI in 2016 meher

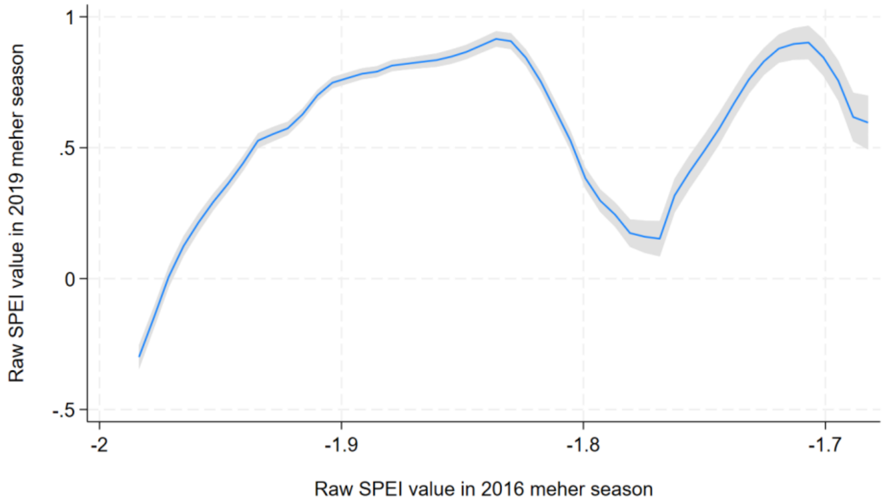


Figure G3. The relationship between SPEI in 2020 meher and SPEI in 2016 meher

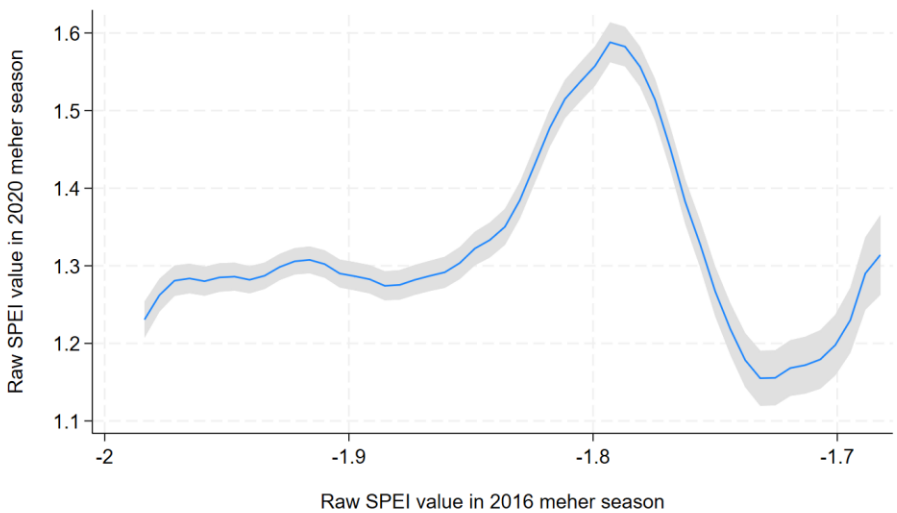


Table G1. Replicating Table 1 but controlling for 2016 drought conditions

	(1)	(2)	(3)	(4)	(5)
Outcome:	Household food gap	Tropical livestock units owned by household	Primary woman's diet diversity	BMI of primary woman	Primary woman experienced IPV
Drought	3.443***	-1.294***	-1.025**	-0.143*	0.144**
	(0.582)	(0.074)	(0.411)	(0.087)	(0.059)
Drought X SPIR	-1.786***	1.003***	1.333***	0.207***	-0.161***
	(0.418)	(0.098)	(0.162)	(0.039)	(0.049)
SPIR	0.068	-0.045**	0.034	-0.032	0.007
	(0.066)	(0.018)	(0.055)	(0.041)	(0.006)
Drought in 2016?	Yes	Yes	Yes	Yes	Yes
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	2.06	0.97	2.02	20.12	0.17
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.005	p = 0.001	p = 0.450	p = 0.381	p = 0.649
Number of observations	6,212	6,274	5,877	4,888	3,808

Note: IPV = Intimate Partner Violence. ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. In column 5, the estimated equation includes a binary variable capturing households for which the baseline value was missing (these missing baseline values were set to zero). The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix H. Extensions

Table H1. Impact of graduation programming and drought conditions on different forms of IPV

	(1)	(2)	(3)	(4)
Outcome:	Any IPV	Physical	Sexual	Emotional
Drought	0.144***	0.068	0.128***	0.086*
	(0.056)	(0.044)	(0.020)	(0.046)
Drought X SPIR	-0.161***	-0.080*	-0.099***	-0.144**
	(0.048)	(0.044)	(0.028)	(0.061)
SPIR	0.007	-0.003	0.003	0.012**
	(0.006)	(0.004)	(0.002)	(0.006)
Woreda fixed effects?	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes
Outcome variable at the baseline?	Yes	Yes	Yes	Yes
Baseline mean of the outcome in control group	0.17	0.06	0.04	0.13
‘Drought’ + ‘Drought X SPIR’ = 0	p = 0.637	p = 0.374	p = 0.139	p = 0.150
Normalized drought impact in control households	0.04***	0.02	0.03***	0.02*
Normalized drought impact in SPIR households	-0.00	-0.00	0.01	-0.01
Number of observations	3,808	3,808	3,808	3,808

Note: ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The estimated equation includes a binary variable capturing households for which the baseline value was missing (these missing baseline values were set to zero). The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Normalized drought impact refers to the estimated drought impact to a 0.25 SD increase in relative dryness.

Table H2. Impact of graduation programming and drought conditions on household per capita consumption

	(1)
Outcome:	Log household consumption in adult equivalents
Drought (in 2019)	-0.219*
	(0.119)
Drought X SPIR	0.260**
	(0.114)
SPIR	-0.042
	(0.029)
Woreda fixed effects?	Yes
Survey round fixed effects?	No
Outcome variable at the baseline?	Yes
'Drought' + 'Drought X SPIR' = 0	p = 0.427
Normalized drought impact in control households	-0.055*
Normalized drought impact in SPIR households	0.010
Number of observations	2,962

Note: ANCOVA specification without midline data. 'Drought' is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Normalized drought impact refers to the estimated drought impact to a 0.25 SD increase in relative dryness.

Table H3. Impact of graduation programming and drought conditions on household and individual level outcomes, by treatment arm

	(1)	(2)	(3)	(4)	(5)
Outcome:	Food gap in months	Tropical livestock units	Diet diversity	Body-mass index	IPV
Drought	3.007***	-1.366***	-0.709	-0.158*	0.143**
	(0.632)	(0.119)	(0.503)	(0.094)	(0.056)
Drought X T1: L* + N*	-1.918***	1.259***	1.571***	0.281***	-0.216***
	(0.402)	(0.156)	(0.346)	(0.062)	(0.037)
Drought X T2: L* + N	-0.891	1.048***	0.988***	0.182***	-0.141
	(0.788)	(0.174)	(0.175)	(0.069)	(0.093)
Drought X T3: L + N*	-2.201***	0.774***	1.209***	0.200***	-0.145**
	(0.454)	(0.128)	(0.179)	(0.064)	(0.058)
T1: L* + N*	0.180***	0.003	0.055	0.029	0.019**
	(0.050)	(0.026)	(0.039)	(0.049)	(0.008)
T2: L* + N	0.018	-0.052**	-0.052	-0.065	-0.010**
	(0.065)	(0.020)	(0.077)	(0.089)	(0.005)
T3: L + N*	-0.000	-0.090***	0.097*	-0.062*	0.016*
	(0.143)	(0.019)	(0.051)	(0.036)	(0.009)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome variable in the control group	2.06	0.97	2.02	20.12	0.17
‘Drought X T1’ = ‘Drought X T2’	p = 0.060	p = 0.211	p = 0.177	p = 0.212	p = 0.386
‘Drought X T1’ = ‘Drought X T3’	p = 0.585	p = 0.030	p = 0.219	p = 0.388	p = 0.123
‘Drought X T2’ = ‘Drought X T3’	p = 0.024	p = 0.065	p = 0.267	p = 0.850	p = 0.963
Number of observations	6,212	6,274	5,877	4,888	3,808

Note: ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. T1, T2 and T3 are SPIR treatment arms where T1 received the intensive livestock (L*) and intensive nutrition intervention (N*), T2 received the intensive livestock intervention (L*) and standard nutrition intervention (N), and T3 received standard livestock (L) and intensive nutrition intervention (N*). The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table H4. Impact of graduation programming and drought conditions on household and individual level outcomes in poorest households eligible for livelihood transfers, by treatment arm

	(1)	(2)	(3)	(4)	(5)
Outcome:	Food gap in months	Tropical livestock units	Diet diversity	Body-mass index	IPV
Drought	3.050***	-0.972***	-0.763	-0.120	0.359***
	(0.770)	(0.149)	(0.598)	(0.092)	(0.050)
Drought X T1: L* + N*	-2.439***	0.855***	1.422***	0.371***	-0.235***
	(0.549)	(0.208)	(0.423)	(0.118)	(0.077)
Drought X T2: L* + N	-1.356*	0.737***	1.130***	0.173*	-0.435***
	(0.731)	(0.195)	(0.245)	(0.091)	(0.062)
Drought X T3: L + N*	-2.207***	0.407***	1.005***	0.281***	-0.270***
	(0.669)	(0.060)	(0.181)	(0.042)	(0.102)
T1: L* + N*	0.303***	0.077*	0.092	0.007	0.025**
	(0.087)	(0.042)	(0.082)	(0.094)	(0.012)
T2: L* + N	0.006	0.015	-0.020	-0.028	0.006
	(0.131)	(0.016)	(0.129)	(0.086)	(0.010)
T3: L + N*	0.074	-0.110***	0.062	-0.056	0.026**
	(0.191)	(0.018)	(0.040)	(0.036)	(0.013)
Woreda fixed effects?	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects?	Yes	Yes	Yes	Yes	Yes
Outcome variable at baseline?	Yes	Yes	Yes	Yes	Yes
Baseline mean of the outcome variable in the control group	2.13	0.65	1.95	20.18	0.18
'Drought X T1' = 'Drought X T2'	p = 0.057	p = 0.622	p = 0.582	p = 0.087	p = 0.004
'Drought X T1' = 'Drought X T3'	p = 0.608	p = 0.035	p = 0.294	p = 0.457	p = 0.608
'Drought X T2' = 'Drought X T3'	p = 0.078	p = 0.077	p = 0.685	p = 0.278	p = 0.011
Number of observations	3,553	3,590	3,374	2,826	2,015

Note: Sample restricted to the poorest 55 % of households at the baseline that were eligible for the livelihood transfer. 'Drought' is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. T1, T2 and T3 are SPIR treatment arms where T1 received the intensive livestock (L*) and intensive nutrition intervention (N*), T2 received the intensive livestock intervention (L*) and standard nutrition intervention (N), and T3 received standard livestock (L) and intensive nutrition intervention (N*). The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table H5. Impact of graduation programming and drought conditions on male and female stress levels

	(1)	(2)
Outcome:	Male stress levels	Female stress levels
Drought in 2019	6.331***	5.532***
	(0.910)	(1.035)
Drought in 2019 X SPIR	-0.725*	-0.976***
	(0.432)	(0.370)
SPIR	0.135*	0.158***
	(0.075)	(0.044)
Woreda fixed effects?	Yes	Yes
Survey round fixed effects?	No	No
Outcome variable at the baseline?	No	No
Mean of the outcome in control group	5.32	5.19
‘Drought’ + ‘Drought X SPIR’ = 0	p<0.001	p<0.001
Normalized drought impact in control households	1.583***	1.383***
As % of the control mean	29.78***	26.67***
Normalized drought impact in SPIR households	1.402***	1.139***
As % of the control mean	26.37***	21.96***
Number of observations	1,981	3,011

Note: Not an ANCOVA specification (endline data only). ‘Drought’ is positive rectified SPEI multiplied by -1 so that larger (positive) values reflect worse drought conditions. The standard errors are reported in parentheses and based on Conley (1999) with a 600 km distance cut-off. Statistical significance is denoted with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The number of observations for males is lower because they were not always present in the household during the interview.