**Title**
Natural Disasters and Livelihoods: Evidence from Pests, Floods, and Disease in African Countries



**Author**
Biscaye, Pierre E

**Publication Date**
2024

Natural Disasters and Livelihoods:
Evidence from Pests, Floods, and Disease in African Countries

by

Pierre E Biscaye

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ethan Ligon, Chair
Professor Edward Miguel
Professor Maximilian Auffhammer

Spring 2024

Natural Disasters and Livelihoods:
Evidence from Pests, Floods, and Disease in African Countries

Abstract

Natural Disasters and Livelihoods:
Evidence from Pests, Floods, and Disease in African Countries

by

Pierre E Biscaye

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Ethan Ligon, Chair

Natural disasters and plagues have marked the course of human history, and remain among the greatest challenges and threats to economies and society. This dissertation uses a number of publicly-available datasets and natural experiments to explore how exposure to natural disasters affect households and societies in low- and middle-income African countries, with a particular focus on effects over time in agricultural communities. The disasters I study include desert locust swarms, floods, and the COVID-19 pandemic. In Chapter 1, I study the long-term effects of transitory agricultural shocks on the risk of violent conflict across Africa and the Arabian Peninsula using local variation in the areas exposed to desert locust swarms. In Chapter 2, I explore issues of measurement in determining what communities have been exposed to floods and how the use of different flooding definitions affects conclusions about the medium-term impacts of flood exposure on the incomes and activities of households in agricultural communities in Nigeria. In Chapter 3, I use a natural experiment during the COVID-19 pandemic to analyze how changes in household childcare burdens affect adults' labor supply in Kenya.

My first chapter studies how transitory agricultural shocks affect the local risk of violent conflict over time. I answer this question using data on conflict events desert locust swarms across 0.25° grid cells in Africa and the Arabian peninsula from 1997-2018. Using modern difference-in-differences and event study approaches, I compare areas exposed to a desert locusts swarm–a localized agricultural disaster–at some point during this period to nearby areas with similar underlying risk of exposure. Local variation in exposure is caused by swarm flight patterns along their migratory paths during major outbreaks, characterized

by daily downwind flights of over 100km. I find that having been exposed to a locust swarm significantly increases the average annual probability of violent conflict in a cell by 0.8 percentage points (43%). This is a very large effect, equivalent to the effect of a 1.6°C higher temperature in the same year. Effects persist for at least 14 years and are driven by swarms arriving in crop cells during the main growing season.

I interpret the results using a model of occupational choice to explore income-related mechanisms. Persistent increases in conflict following a localized disaster suggest a decrease in the opportunity cost of fighting in affected areas. In line with this, I find that swarm exposure significantly reduces cereal yields in subsequent years, building on previous studies finding persistent effects of locust exposure on measures of household well-being and agricultural profits. Such effects indicate a permanent income mechanism for this severe transitory shock. This mechanism is not sufficient to explain the pattern of long-term impacts on violent conflict risk: the feasibility of engaging in violent conflict also matters. Increases in conflict risk are concentrated in years with active fighting groups in neighboring areas who may recruit or coerce individuals to join in violent conflict. Conflict therefore increases only when the reduced opportunity cost of fighting is combined with opportunities to fight. Patterns of long-term impacts on violent conflict are similar for severe droughts, indicating the mechanisms are not specific to locust shocks. Long-term impacts of transitory economic shocks on conflict risk add further motivation for policies mitigating the risk of such shocks and promoting household resilience and long-term recovery.

My second chapter explores how extreme weather events affect livelihood decisions of agricultural households over time, focusing on the devastating 2012 floods in Nigeria. I use nationally-representative panel household survey data together with satellite imagery to analyze how exposure to the 2012 floods affects household labor supply and income in subsequent years. I first show that identification of flooding exposure varies depending on whether survey reports or MODIS satellite imagery are used. Differences in characteristics of communities identified as flooded by only one definition imply specific measurement issues for each. These results indicate that more attention is needed to considering how flooding is defined and measured, and whether there are different impacts of exposure to different types of floods.

As in Chapter 1, I use a difference-in-differences design comparing changes in outcomes over time for households in communities exposed to flooding in 2012 against households in non-exposed communities that had a similar risk of flooding. I find no effects of flood exposure on engagement in wage employment or on total household income but a small significant increase in the probability of experiencing food insecurity using both definitions. The value of crop production falls, driven by a decrease in the value of commercial crops, while staple crop production either increases or remains stable. Using a MODIS-based flooding definition, I find that households exposed to floods significantly increase non-farm enterprise income relative to non-exposed households, as they reallocate labor from crop production to existing businesses. Using a survey-based definition, I find that flooding causes some farm households to become engaged in non-farm enterprise but without significantly increasing enterprise income while some non-farm households begin production of subsistence crops, with limited returns. The results emphasize that survey and satellite measures capture different types of flooding-related events which accordingly have different effects on exposed households. Studies of the impacts of floods should carefully consider their choice of flooding measure and the type of flooding they are most interested in analyzing.

My third chapter, coauthored with Dennis Egger and Utz Pape, identifies the impact of a shock to childcare and child labor on adult labor supply and intra-household allocation of productive activities in the context of COVID-19-related school closures in Kenya. Using nationally-representative bi-monthly panel data, we compare changes in labor supply after schools partially reopened in October 2020 for adults with children in a grade eligible to return against adults with children in adjacent grades. We find that a child returning to school increases adults' weekly work hours by 29% in the short run, concentrated among the most flexible margins of adjustment, particularly household agriculture.

Contrary to evidence from high-income settings, overall effects are not gendered. However, equal average labor supply responses for women and men are driven by different mechanisms particular to low- and middle-income settings. Women free up more time than men when childcare burdens fall but specialize more in childcare when returning students were net caregivers to younger siblings during school closures. Women also shoulder more of the reduction in child agricultural labor when students return to school, and shift from non-agricultural work into household agriculture (more easily combined with care of younger children). A back-of-the-envelope calculation suggests that school closures account for at least 40% of the overall drop in labor supply during the pandemic in Kenya, and a fall in GDP of 2.6%. Our results suggest policies increasing childcare access could substantially increase adult labor supply in low- and middle-income countries.

Climate change is increasing not only global temperatures but also the frequency and severity of extreme weather events. Globalization has made societies around the world more connected but also more vulnerable to the spread of pests and disease. Yet there is little evidence on how such disasters affect households in low-income countries over time, and how they affect labor and livelihood decisions of poor farm households in particular. The three chapters of my dissertation combine publicly-available data, economic theory, and rigorous identification based on modern econometric methods to contribute to policy discussions around mitigation and adaptation to climate change, structural transformation, and civil conflict, among the greatest concerns for policymakers in Africa and around the world.

# Contents

# List of Figures

# List of Tables

# Acknowledgments

I am would first like to express my deep gratitude to the members of my dissertation committee: Ethan Ligon, Edward Miguel, and Maximilian Auffhammer. Their consistent support, regular availability, and invaluable feedback greatly improved the quality of my research and the production of this dissertation. I am extremely grateful to the 2018 UC Berkeley ARE PhD cohort—Daniel Agness, Qianmiao Chen, Joel Ferguson, Wei Guo, Qingyang Huang, and Jed Silver—for their fellowship, support, and feedback over the course of my PhD. I could not have asked for a better group of friends and colleagues. I thank Sofia Villas-Boas for her guidance and encouragement on the job market. I also thank the Berkeley ARE support staff, and Carmen Karahalios and Diana Lazo in particular, for their helpful support and advice with the financial and administrative side of the PhD.

Chapter 2 of this dissertation grew out of initial discussions with Abdulrasheed Isah. The material is based on work supported by the UC Berkeley Center for African Studies and the Structural Transformation and Economic Growth initiative (under Grant No. STEG SRG LOA_1144). I am grateful to both of these organizations. All errors are my own. I co-authored Chapter 3 of this dissertation with Dennis Egger and Utz Pape, and I am particularly grateful to Dennis for his support across multiple rounds of review and revisions. I am grateful to the IPA Women's Work, Entrepreneurship, and Skilling (WWES) Initiative and the Department of Economics at UC Berkeley for generous financial support. All errors are our own.

I owe a tremendous debt of gratitude to C. Leigh Anderson for encouraging me to pursue a PhD and for her continued mentorship, advice, guidance, and insight into academia, development, mentoring, research, and life. I am grateful to Leigh and to Mary Kay Gugerty for enabling my first academic fieldwork experience in Tanzania, and to them and Travis Reynolds for introducing me to the process of academic publishing and empowering me to co-author papers with them before starting my PhD. I am also grateful to Catherine Wolfram for giving me the opportunity to work with her on research projects in Ghana for multiple years during my PhD at Berkeley and to Susanna Berkouwer and Steve Puller for incredible collaboration on these projects, even though I do not include the resulting papers in this dissertation.

Finally and most of all, I thank my family for their love and continuous support. I thank my wife Linda for embarking with me on my PhD journey, for giving me the space and time I needed to complete this dissertation, and for her unfailing encouragement. I thank my daughters Évelyne and Éloïse for bringing me joy every day. Last, I thank my parents Patricia Ayité and Timothy Biscaye for modeling hard work, dedication, and balance, and for my childhood in Africa which has defined the focus of my research and work. None of this would have been possible without you all.

# Chapter 1

# Agricultural shocks and long-term conflict risk: Evidence from desert locust swarms

## 1.1 Introduction

Violent civil conflict—between actors within states and sometimes spilling over—has been increasing in recent decades, especially in Africa and the Middle East.[1] Such conflict harms lives, health, and living standards of affected populations and can slow or reverse economic growth and development as time and investment are allocated to destruction rather than production (Blattman and Miguel, 2010; Fang et al., 2020), with potentially broader impacts through economic instability and migrant flows. Understanding the drivers of civil conflict therefore has important implications for policy. A large economic literature explores the impacts on conflict risk of transitory agricultural shocks which do not permanently affect potential land productivity. This is a primary concern given the prominence of agricultural livelihoods in many of the areas most affected by civil conflict and the threat to agriculture posed by climate change. Studies of this relationship focus on short-term impacts, and those that analyze shocks to agricultural production are limited in their ability to identify causal mechanisms.[2]

[1] See Cederman and Pengl (2019) for a review of recent conflict trends using data from the Uppsala Conflict Data Program.

[2] Several studies have found that shocks to agricultural prices increase conflict incidence (e.g., (Dube and Vargas, 2013; Fjelde, 2015; McGuirk and Burke, 2020; Ubilava et al., 2022). Impacts on agricultural productivity are speculated to explain the widely-studied relationship between climate or weather deviations and conflict risk (see Burke et al. (2015); Carleton et al. (2016); Dell et al. (2014); Hsiang and Burke (2013);

This paper analyzes the dynamic impact of a severe transitory shock to agricultural production—exposure to a desert locust swarm—on conflict risk and tests for evidence of a long-run permanent income or wealth mechanism. Desert locusts are the world's most dangerous and destructive migratory pest (Cressman et al., 2016; Lazar et al., 2016) and effectively constitute an agriculture-specific natural disaster. The arrival of a locust swarm often leads to complete destruction of agricultural production and other vegetation (Symmons and Cressman, 2001; Thomson and Miers, 2002), without the effects on infrastructure or human physiology which may result from weather shocks. Desert locusts swarms—with billions of locusts covering tens of square kilometers—make daily flights as a cohesive unit, typically downwind. Swarm movements can be predicted but with a great deal of uncertainty and there are no effective means for farmers to prevent damages even with awareness of the potential threat of locust swarms. This creates quasi-random variation in the areas exposed to agricultural destruction in a swarm's migratory path. Swarms continue migrating and reproducing until they are controlled by pesticides or reach areas with scarce vegetation, where they die out. They do not increase long-term risk from pests in affected areas or otherwise affect agricultural production fundamentals. Climate change is creating conditions more conducive to swarm formation (McCabe et al., 2021), potentially undoing progress from increased international monitoring and control efforts in recent decades. These characteristics make locust swarms a useful natural experiment for analyzing how transitory agricultural production shocks affect the risk of conflict.

Using data on the location and timing of desert locust swarm observations from the Food and Agricultural Organization of the United Nations (FAO) and of conflict events from the Armed Conflict Location & Event Data Project (ACLED) and Uppsala Conflict Data Program (UCDP) I estimate a model of conflict at the annual level for 0.25° (around 28×28km) grid cells between 1997-2018 across Africa and the Arabian peninsula.[3] As severe agricultural shocks may have persistent effects on wealth and productivity which could affect conflict risk, I define exposure to a locust swarm as a permanent treatment. I estimate dynamic impacts of swarm exposure on different types of conflict using event study designs from the recent literature on difference-in-differences with staggered treatment timing (Borusyak et al., 2021; Callaway and Sant'Anna, 2021; Roth et al., 2023). Regressions include cell and country-by-

Koubi (2019); Mach et al. (2019) for reviews), though weather may affect conflict through mechanisms other than agriculture and some studies find results that are not consistent with effects through agricultural productivity (e.g., (Bollfrass and Shaver, 2015; Sarsons, 2015).

[3]I include all countries where at least 10 locust swarms are reported during the sample period. Torngren Wartin (2018), an unpublished master's thesis from Stockholm University, estimates short-term impacts of desert locusts on conflict in Africa using similar data. Torngren Wartin (2018) focuses on potential measurement issues around short-term impacts which I discuss in Section 1.8. He does not consider long-term impacts of locust swarms or mechanisms that are the main contributions of this paper.

year fixed effects and control for annual rainfall, temperature, and population to account for potential time-varying differences in cells affected by locust swarms, and I test the sensitivity of the results to weighting observations based on their estimated propensity to have been exposed to any locust swarm (Stuart et al., 2014). I explore patterns in calendar time using an event study of the 2003-2005 major desert locust upsurge, which accounts for 73% of swarm exposure events in the sample period, and estimate average long-term impacts on and heterogeneity by different cell characteristics using in two-way fixed effects regressions.

Locust swarms increase the annual probability of any violent conflict event occurring in a 0.25° grid cell by 0.8 percentage points (43%) on average in years after exposure to the swarm, compared to unaffected areas. The event studies show no significant impacts of locust swarms on violent conflict in the year of exposure but increases in all following years up to 14 years after exposure. Impacts are entirely driven by cells with crop or pasture land, and by swarms arriving in crop cells during the main growing season in particular. The estimated effects of having been affected by a locust swarm are large relative to the effects of annual weather deviations. Experiencing 100mm higher rainfall than average in a year increases the risk of violent conflict by 0.3 percentage points, and a 1°C higher temperature increases it by 0.5 percentage points. The results are robust to a variety of alternative specifications.

I interpret the results through the lens of a commonly-used model of individual occupation choice between production and conflict (Chassang and Padró i Miquel, 2009; Dal Bó and Dal Bó, 2011). In the model, transitory agricultural shocks affect the short-term risk of conflict by changing both the returns to engaging in agricultural production—the *opportunity cost* mechanism—and the returns to fighting over agricultural output—the *predation* mechanism. The predation mechanism is particularly relevant for violent conflict over outputs (i.e., banditry), and less so for conflict over factors of production (land in particular) or non-violent conflict such as on protests. I extend the model to allow past agricultural production shocks to affect conflict risk through a permanent income or *wealth* mechanism. Adoption of insurance against negative agricultural shocks is very low in the study area, leading households to undertake costly consumption smoothing strategies and reducing household wealth (financial, physical, and human capital).[4] This wealth effect can decrease household permanent income as households are less productive in the long-run, leading to persistent reductions in the opportunity cost of fighting.

Null effects of locust swarms on violent conflict in the same year contrast with effects of rainfall and temperature increases, though other studies similarly report delayed impacts of

[4]See for example Alderman et al. (2006); de Janvry et al. (2006); Dercon (2004); Dercon and Hoddinott (2004); Dinkelman (2017); Fafchamps et al. (1998); Hallegatte et al. (2020); Hoddinott and Kinsey (2001); Hoddinott (2006); Maccini and Yang (2009); Townsend (1995) on coping strategies LMIC agricultural households use to respond to uninsured shocks being largely uninsured and their impacts on household wealth/assets, including long-term consequences.

weather shocks on conflict (Crost et al., 2018; Harari and La Ferrara, 2018). There are a few potential explanations for this difference. First, weather deviations affect conflict risk through mechanisms other than agricultural productivity such as impacts on physiology, psychology, and infrastructure.[5] I find that the association between rainfall and violent conflict does not vary significantly by agricultural land cover, suggesting these channels may be more important than effects on agricultural production. Second, the level of destruction of locust swarms reduces the resources available for fighting and the returns to fighting at the same time as it decreases the opportunity costs of fighting. I find similar null effects of locusts on factor conflict and protests in the same year indicating it is not just the predation mechanism and reduced returns to banditry that explains the effect. Third, the fall in opportunity costs following locust exposure may be limited by out-migration (a common response to locust destruction and by relief efforts. I find evidence that international agricultural aid flows increase to countries after exposure to locust swarms, which could dissuade affected individuals from engaging in violent conflict.

Long-term increases in conflict risk following exposure to a locust swarm are consistent with predictions in the model based on a wealth shock decreasing permanent income and productivity. Using spatial data from the Demographic and Health Survey (DHS) AReNA database (IFPRI, 2020), I show that NDVI falls in the first year of swarm exposure and that cereal yields decrease by 108kg/ha (6%) on average in years after exposure relative to unaffected areas, indicating a persistent decrease in agricultural productivity. Swarm exposure leads to large increases in the long-term risk of violent conflict involving non-state actors (such as rebel groups, identity militias, and terrorist organizations) and of protests, which would be expected to be affected by changes in individuals' opportunity costs, but has smaller and inconsistent effects on violent conflict only involving state actors (governments), which would not. This evidence supports the argument that long-term impacts of locust swarm exposure on conflict risk are due to a persistent reduction in the opportunity cost of fighting.

Long-term conflict risk does not increase uniformly, with the largest effects on violent conflict risk coming 7-14 years after swarm exposure and a similar pattern for the likelihood of protests. Long-term impacts on violent conflict are driven entirely by cells and periods where there are groups actively engaged in conflict in the surrounding area. The prevalence of violent conflict in the sample countries is relatively low until around 2011 when it begins increasing significantly. The majority of swarm exposure comes from the 2003-2005 upsurge, explaining the lag between exposure and the largest increases in violent conflict in the event studies. Violent conflict is generally an organized group activity, as groups

[5] Dell et al. (2012, 2014); Mellon (2022) document a wide variety of channels through which temperature and rainfall can affect the economy and society.

reduce the cost to individuals of engaging in fighting and increase the probability of success. Fighting groups formed following some precipitating event[6] can more easily recruit individuals in areas affected by locusts as their opportunity costs are lower and they will therefore demand less returns to join. Lower agricultural productivity of the levels observed in this study are unlikely to make fighting optimal for affected individuals in most periods absent such recruitment. Locust swarm exposure thus affects conflict incidence over the long-term but does not cause the onset of new conflicts.

Finally, I show that the patterns in impacts on violent conflict over time I observe are not unique to desert locust swarms. Severe droughts also consistently increase long term violent conflict, with the annual risk 1.0 percentage points higher than in unaffected areas on average over the following 12 years. As with locust swarms, impacts of droughts are driven entirely by exposure in agricultural cells and the long-term effects are only significant in years with groups engaged in violent conflict in surrounding cells, indicating the same mechanisms are involved. The long-term increases in conflict risk imply that analyses defining shock exposure as temporary and estimating short-term impacts using fixed effects (the main method in studies of agricultural shocks and conflict) are misspecified. I show that such analyses result in downward-biased estimates of the short-term impacts of both locust swarms and severe drought on violent conflict, affecting the policy implications.

This paper makes several contributions to the literature. First, while a relationship between weather shocks and conflict has been repeatedly demonstrated the mechanisms driving this impact are not fully understood (Mach et al., 2020). Weather shocks affect a variety of economic and social outcomes in addition to reducing agricultural labor productivity and agricultural output (Dell et al., 2012, 2014; Mellon, 2022), and some papers cast doubt on the agricultural channel by finding that impacts of rainfall on conflict do not vary by sensitivity or presence of agricultural production (Bollfrass and Shaver, 2015; Sarsons, 2015). This paper addresses these limitations by testing the impacts on conflict risk of locust swarms, an agricultural shock that does not affect other economic outcomes. I compare impacts by land cover and timing of swarm exposure and show further evidence that impacts of rainfall shocks do not vary by agricultural land cover. The results indicate that the opportunity cost of fighting mechanism alone does not explain impacts of locust swarms and drought on conflict risk and highlight the importance of the predation mechanism and its influence on opportunities for fighting.[7]

[6]Global food price shocks, the Arab Spring, and the spread of terrorist organization in many sample countries are among the proximate causes of conflict onset in the study period.

[7]Little attention is given to the predation mechanism in the climate-conflict literature. Studies showing evidence of opportunity cost and predation mechanisms in agriculture have primarily explored impacts on conflict risk of changes in global prices of agricultural goods (e.g., (Dube and Vargas, 2013; Fjelde, 2015; McGuirk and Burke, 2020) rather than shocks to local agricultural production. McGuirk and Nunn (2021)

Second, I analyze long-term impacts of an economic shock on conflict to further our understanding of the drivers of conflict (Bazzi and Blattman, 2014; Blattman and Miguel, 2010; Collier and Hoeffler, 1998; Dube and Vargas, 2013; Grossman, 1999; Hodler and Raschky, 2014; McGuirk and Burke, 2020; Miguel et al., 2004), and the role of agricultural production in particular (Crost et al., 2018; Harari and La Ferrara, 2018; Iyigun et al., 2017; McGuirk and Nunn, 2021; Von Uexkull et al., 2016). Studies of the impacts of agricultural shocks on conflict have focused on the short-term—impacts in the same year and potentially the following year.[8] I consider the possibility of long-term impacts through a permanent income or wealth mechanism, and test for this channel by examining long-term impacts of desert locust swarms on measures of agricultural productivity. Significant increases in long-term conflict risk following locust and drought exposure suggest that studies focusing on short-term impacts of severe economic shocks may be misspecified if they define the shock 'treatment' as temporary.

Third, I add to a broader literature on the impacts of environmental shocks and natural disasters. Many papers have explored how environmental shocks can have persistent effects on poverty and well-being (e.g., (Baseler and Hennig, 2023; Carter and Barrett, 2006; Carter et al., 2007; Lybbert et al., 2004), but these mechanisms have not been related to conflict risk. More generally, the evidence on long-term impacts of disasters such as hurricanes and droughts is limited, inconclusive, and focused on a small number of outcomes (see Botzen et al. (2019); Klomp and Valckx (2014) for reviews). I study how the impacts of desert locusts swarms—an extreme shock to agricultural production akin to a natural disaster—on conflict risk evolve over time and test whether patterns are consistent with a wealth mechanism.

Finally, this paper also contributes to a small literature studying economic impacts of desert locusts beyond immediate crop destruction and costs of control operations (see e.g., (Thomson and Miers, 2002), and to a slightly larger literature on the long-term impacts of agricultural pests (Baker et al., 2020; Banerjee et al., 2010). The range of many agricultural pests is expanding due to climate change and globalization, and though locust outbreaks have become less frequent in recent decades due to increased monitoring desert locusts are ideally situated to benefit from climate change (McCabe et al., 2021; Qiu, 2009; Youngblood et al., 2023). Marending and Tripodi (2022) use panel data from the Ethiopia Socioeconomic Survey and find that exposure to desert locust swarms decreases farm profits by 20-48% two harvest seasons after swarm arrival. Several papers use Demographic and Health Survey

is an exception, analyzing impacts of drought on conflict between pastoralists and farmers.

[8] Crost et al. (2018); Harari and La Ferrara (2018) estimate effects of weather shocks on conflict 1 and 4 years afterward. Iyigun et al. (2017) is an exception, considering long-run effects on conflict risk of a *positive and permanent* agricultural productivity shock from the introduction of the potato to the Eastern Hemisphere. To my knowledge, no study has explored long-term impacts on conflict risk of a transitory negative shock to agricultural production.

(DHS) data and variation in swarm exposure across birth cohorts and over space to show negative impacts on school enrollment and educational attainment (De Vreyer et al., 2015) and on child height-for-age or stunting (Conte et al., 2023; Le and Nguyen, 2022; Linnros, 2017). These papers illustrate how locust swarm exposure can adversely affect long-term productivity and human capital, but I am not aware of any study considering the impacts of a pest shock on conflict. The impacts of locust swarms on long-term agricultural productivity and conflict risk should be considered in determining policy around desert locust prevention and control.

The remainder of the paper is organized as follows. Section 1.2 provides background on desert locusts and summarizes the literature on agricultural shocks and conflict. Section 1.3 presents a model of how agricultural shocks affect occupational choice and the decision to fight over time through income-related mechanisms. Section 1.4 describes the data used in the analyses and Section 1.5 outlines the empirical approach. Section 1.6 shows the results for the impacts of locust swarm exposure on violent conflict. Section 1.7 discusses the results in light of the model and presents additional analyses testing the mechanisms behind the estimated effects. Section 1.8 compares impacts of exposure to locust swarms and severe drought and consider implications for analyses of short-term impacts of agricultural shocks. Section 1.9 concludes.

## 1.2 Background

### 1.2.1 Desert locusts

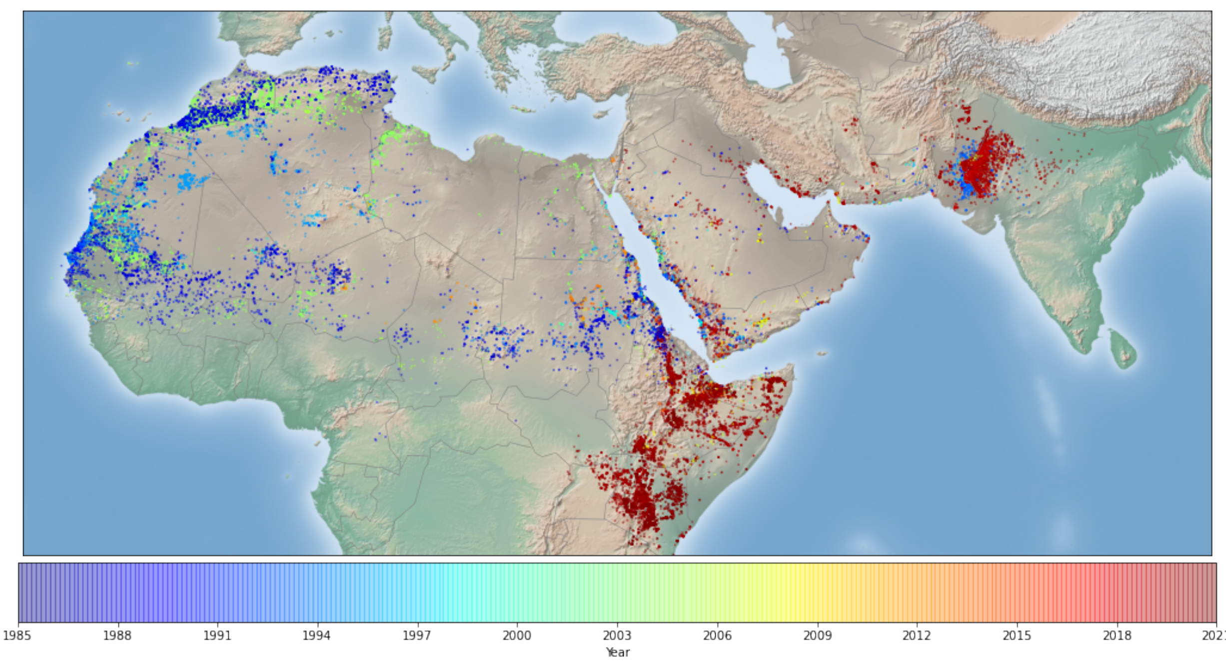
Desert locusts (*Schistocerca gregaria*) are a species of grasshopper always present in small numbers in desert 'recession' areas from Mauritania to India.[9] They usually pose little threat to livelihoods but favorable climate conditions in breeding areas—periods of repeated rainfall and vegetation growth overlapping with the breeding cycle—can lead to exponential population growth. Unique among grasshopper species, after reaching a particular population density locusts undergo a process of 'gregarization' wherein they mature physically and begin to move as a cohesive unit (Symmons and Cressman, 2001). In the adult locust stage—lasting 2-4 months—after developing wings, gregarization leads to the formation of 'swarms'. In this paper I focus exclusively on locust swarms, which are much more mobile and destructive than other groupings of desert locusts. Climate change is expected to increase the risk of locust swarm formation, as desert locusts can easily withstand elevated temperatures and the increased frequency of extreme weather events can create conditions

[9]Additional detail on desert locusts is included in Appendix A.2. Any time I use 'locusts' in this paper I am referring exclusively to desert locusts.

conducive to exponential population growth (McCabe et al., 2021; Qiu, 2009; Youngblood et al., 2023).

Figure 1.1 displays the locations of all desert locust swarm observations recorded in the FAO Locust Watch database by year. As illustrated by the figure, locust swarms are not observed with any regularity over time and only in a few countries have locations been exposed to more than one swarm. Which locations are exposed to a locust swarm during an upsurge—a major outbreak of locust swarms from breeding areas which affects multiple countries—depends on which breeding areas fostered initial swarm formation and on wind patterns in the months following swarm formation. The countries affected by the 2003-2005 upsurge (in green in Figure 1.1), which originated from multiple small outbreaks in Summer 2003 in the Western Sahel, are not the same as those exposed to the 2019-2021 upsurge (in red), which originated in the southern Arabian Peninsula. Conditional on being in the migratory path of a locust swarm, swarm flight patterns create further variation in exposure as some areas in the flight path are flown over and spared any damages.[10]

Figure 1.1: FAO Locust Watch swarm observations by year




Note: Map created by authors using all recorded locust swarms in the FAO Locust Watch database.

[10]Figure A.2.3 shows an example of local variation in exposure to the 2003-2005 upsurge in Mali.

Desert locusts are migratory rather than endemic, moving on after consuming available vegetation without permanently changing local agricultural pest risk. Locust swarm migration follows wind patterns which tend to bring them to breeding areas in time for seasonal breeding. Swarms generally fly downwind and can easily move 100km or more in a day even with minimal wind (FAO and WMO 2016).[11] Locusts live 2-6 months and swarms continue breeding and migrating until dying out from a combination of migration to unfavorable habitats, limited vegetation in breeding areas, and control operations (Symmons and Cressman, 2001).

Locust swarms vary in density and extent. The average swarm includes around 50 locusts per $m^2$ and can cover tens of square kilometers, including billions of locusts (Symmons and Cressman, 2001). About half of swarms exceed 50km$^2$ in size (FAO and WMO 2016). Locusts consume any available vegetation without preference for different types of crops (Lecoq, 2003). A small swarm covering one square kilometer consumes as much food in one day as 35,000 people and the median swarm consumes 8 million kg of vegetation per day (Food and Agriculture Organization of the United Nations (FAO), 2023a).

The arrival of a swarm can lead to the total destruction of local agricultural output (Symmons and Cressman, 2001; Thomson and Miers, 2002)). During the 2003-2005 locust upsurge in North and West Africa, 100, 90, and 85% losses on cereals, legumes, and pastures respectively were recorded, affecting more than 8 million people and leading to 13 million hectares being treated with pesticides (Showler, 2019). Chatterjee (2022) finds that wheat yields are 12% lower on average in Indian districts typically affected by desert locusts in years of locust outbreaks, suggesting large decreases in the specific areas exposed to locust swarms in those years. Over 25 million people in 23 countries were affected during the most recent 2019-2021 upsurge and damages were estimated to reach $1.3 billion (Green, 2022), with control efforts—including treating over 2 million hectares with pesticides—estimated to have prevented over $1 billion in damages (Newsom et al., 2021).

An important result of the local variation in locust swarm damages during upsurges is that macro level impacts may be muted, since upsurges occur in periods of positive rainfall shocks in breeding areas which tend to be associated with better agricultural conditions in surrounding countries. Several studies find that impacts of locust upsurges on national agricultural output and on prices are minimal, despite devastating losses in affected areas (Joffe, 2001; Krall and Herok, 1997; Showler, 2019; Zhang et al., 2019).

Farmers have no proven effective recourse when faced with the arrival of a locust swarm (Dobson, 2001; Hardeweg, 2001; Thomson and Miers, 2002). The only current viable method

[11] Swarms do not always fly with prevailing winds and may wait for warmer winds which lead to seasonal breeding areas (FAO and WMO 2016). Small random deviations in the positions of individuals in the swarm can also sometimes alter the course of the swarm's flight.

of swarm control is direct spraying with pesticides, which can take days to have effects as well as being slow and costly to organize and requiring robust locust control infrastructure (Cressman and Ferrand, 2021). Knowledge of locust breeding patterns and swarm flight characteristics inform efforts to predict locust swarm formation and movements, but forecasts remain highly imprecise (Latchininsky, 2013). Farmers in affected areas report viewing locust swarms as an unpredictable natural disaster that is the government's responsibility to address (Thomson and Miers, 2002). Households use a variety of measures to cope with short-term food security and livelihood effects of locust outbreaks. In addition to seeking help from social networks and food aid, households commonly report selling animals and other assets, consuming less food, sending household members away, taking loans and cutting expenses, and consuming seed stocks as coping strategies (Thomson and Miers, 2002). A swarm exposure shock therefore represents a shock to income and household wealth as well as a shock to agricultural productivity in the year of exposure.

The characteristics of desert locust swarms make them a useful natural experiment for analyzing the impacts of agricultural production shocks on the risk of conflict. The arrival of a swarm is effectively a locally and temporally concentrated natural disaster where all crops and pastureland are at risk (Hardeweg, 2001), but other aspects of the economy are importantly unaffected. Whereas impacts of temperature or rainfall deviations on conflict might go through channels other than agricultural production, such as effects on infrastructure or physiology, impacts of locust swarms should be restricted to this channel. The level of damage to agriculture from swarms arriving during the crop growing season and lack of tools for farmers to prevent damages imply more severe reductions in agricultural production than moderate weather deviations. Decreased wealth following such a catastrophic shock may be more likely to persist and affect agricultural labor productivity in following seasons, increasing the potential for long-term impacts on conflict through the opportunity cost channel.

### 1.2.2 Agricultural shocks and conflict

A growing literature explores the impacts of climate or weather on conflict (see Burke et al. (2015); Carleton et al. (2016); Dell et al. (2014); Hsiang and Burke (2013); Koubi (2019); Mach et al. (2019) for reviews), primarily analyzing impacts of deviations of rainfall or temperature from historical norms. Most studies find that weather shocks increase short-term conflict risk, with meta-analyses finding a mean increase in the risk of conflict of between 5-10% for a one standard deviation increase in a weather shock variable and more consistent positive effects of temperature increases (Burke et al., 2015; Carleton et al., 2016).[12] These results have important implications for conflict risk as climate change increases the frequency and severity of weather shocks.

The majority of papers in the climate-conflict literature focus on income-related mechanisms, following early work in Miguel et al. (2004). Arguments typically follow the models in Chassang and Padró i Miquel (2009); Dal Bó and Dal Bó (2011), discussing how weather affects agricultural labor productivity and therefore the opportunity cost of engaging in conflict for agricultural producers—a large share of the population in the largely low- and middle-income countries studied. Many studies use variation by land cover or timing relative to show support for this mechanism (e.g., (Caruso et al., 2016; Crost et al., 2018; Gatti et al., 2021; Harari and La Ferrara, 2018; McGuirk and Nunn, 2021; Von Uexkull, 2014). Discussion of offsetting changes in the returns to conflict from capture of agricultural output is typically limited. If the predation or rapacity mechanism is mentioned it is presented as being outweighed by the opportunity cost mechanism.

Weather affects the economy and society through multiple channels besides agricultural production (Dell et al., 2012, 2014; Mellon, 2022), and studies have pointed to physiological, psychological, and infrastructural effects of weather shocks in explaining impacts on conflict (Baysan et al., 2019; Carleton et al., 2016; Chemin et al., 2013; Dell et al., 2014; Hsiang and Burke, 2013; Sarsons, 2015; Witsenburg and Adano, 2009). A few studies have also cast doubt on the importance of the agriculture channel. Sarsons (2015) shows that rainfall shocks in India predict riot incidence similarly across districts where rainfall has different effects on income due to access to irrigation dams. Bollfrass and Shaver (2015) find that positive temperature deviations in sub-Saharan Africa increase conflict even in areas without significant agricultural production. Reviews of the climate-conflict literature agree that the mechanisms remain unclear and deepening insight into them is highlighted as a priority for future climate-conflict research in Mach et al. (2020). An important advantage of analyzing the effects of locust swarm exposure is that the only active channel should be agricultural destruction, allowing a cleaner identification of the importance of income-related mechanisms.

Income mechanisms have been discussed and tested in the literature on the drivers of conflict more generally including studies looking at changes in the value of agricultural labor or output not related to production shocks. Several studies have shown that plausibly exogenous changes in prices of agricultural commodities affect the risk local conflict in areas producing the affected goods (Bazzi and Blattman, 2014; Dube and Vargas, 2013; Fjelde, 2015; McGuirk and Burke, 2020; Ubilava et al., 2022). A recent literature explores how the

[12]Most studies reporting null effects use country-level data (e.g., (Buhaug et al., 2015; Ciccone, 2013; Couttenier and Soubeyran, 2014; Klomp and Bulte, 2013).

onset of harvest season in agricultural areas affects economic incentives to fight (Guardado and Pennings, 2021; Hastings and Ubilava, 2023). Koren (2018) finds that increased cereal yields in Africa, instrumented by a measure of drought, are associated with an increased risk of conflict.

These studies illustrate different ways in which agricultural shocks affect both the opportunity cost of fighting related to agricultural labor productivity and the returns to predatory capture of agricultural output. In some cases the opportunity cost mechanism appears to dominate (Bazzi and Blattman, 2014; Dube and Vargas, 2013; Fjelde, 2015; Guardado and Pennings, 2021), while in others the predation mechanism is decisive (Koren, 2018; Ubilava et al., 2022), sometimes within the same context (Hastings and Ubilava, 2023; McGuirk and Burke, 2020). Studies identifying results consistent with predation or banditry emphasize the role of armed groups such as militias or insurgents which coordinate such attacks. It is not clear whether the opportunity cost or predation mechanism would dominate for a shock to agricultural production. Locust swarms can entirely destroy agricultural output, sharply reducing both agricultural labor productivity and the returns to predation over agricultural output.

The majority of studies on agricultural shocks and conflict focus on impacts within the same time period, with a few exceptions. Crost et al. (2018); Harari and La Ferrara (2018) find that growing season weather shocks have inconsistent effects on conflict in the same year, but consistently increase conflict risk the following year. Harari and La Ferrara (2018) finds some persistence of impacts up to 4 years afterward. Crost et al. (2018) argue that lagged effects on conflict could be due to storage and savings offsetting effects in the same year. This appears inconsistent with effects driven by reduced opportunity costs as the largest impacts on agricultural labor productivity should be realized in the same season as the rainfall shock, unless wealth effects reduce subsequent productivity. To my knowledge, only Iyigun et al. (2017) consider how a permanent agricultural productivity shock impacts conflict in the long term. They find that introducing potatoes to Europe, the Near East, and North Africa led to a large and persistent reduction in the risk of conflict in subsequent centuries by comparing changes in areas with different suitability for potato cultivation. A paucity of evidence on long-term impacts is a limitation of the literature on natural disasters more generally (Botzen et al., 2019), and particularly in low-income countries (Baseler and Hennig, 2023).

## 1.3 Model

The standard models discussed in the literature on agricultural shocks and conflict are models of occupational choice (as in (Roy, 1951; Heckman and Honore, 1990; French and Taber, 2011) where actors allocate their labor between productive activities and fighting.[13] In particular, Chassang and Padró i Miquel (2009) develop a bargaining model of conflict where groups allocate labor to crop production or fighting over land, and Dal Bó and Dal Bó (2011) model individuals choosing between a labor-intensive sector, a capital-intensive sector, and an 'appropriation' sector fighting over output.

The models illustrate how agricultural shocks affect the risk of conflict through two main, opposing, mechanisms. First, negative shocks such as low prices or drought reduce the returns to agricultural labor. This means the *opportunity cost* of fighting is lower: producers have less to lose by engaging in conflict. At the same time, lower agricultural prices or output reduce the returns to *predation*: bandits or looters have less to gain from fighting. These mechanisms are not unique to agricultural shocks, as they are also discussed in earlier work on the economic drivers of conflict more generally (e.g., (Collier and Hoeffler, 2004; Grossman, 1999). For transitory agricultural shocks, the value of output available to capture falls but the value of factors of production—land in particular—is less affected. In line with this, the literature generally finds that the opportunity cost mechanism dominates for shocks that temporarily reduce agricultural returns, increasing conflict risk.

Prior research models transitory agricultural shocks as having only temporary effects on conflict. But a transitory income shock can also reduce permanent income when consumption smoothing decreases assets. Most agricultural households in developing countries lack insurance and have constrained access to credit. Strategies to smooth consumption following an income shock, such as selling animals and other assets, taking loans, reducing food, health, and education spending, and sending members away reduce household physical and human capital (e.g., (de Janvry et al., 2006; Dercon and Hoddinott, 2004; Dinkelman, 2017). The resulting reductions in wealth mean temporary shocks can have persistent impacts on productivity (Dercon, 2004; Hallegatte et al., 2020; Hoddinott, 2006; Karim and Noy, 2016). In the context of an occupational choice model of conflict, this permanent income or *wealth* effect increases the long-term risk of conflict by reducing the long-term opportunity cost of fighting.

I present a streamlined model of occupational choice as in Chassang and Padró i Miquel (2009); Dal Bó and Dal Bó (2011) including this long-term wealth mechanism to build intuition and generate hypotheses about the effect of agricultural shocks on conflict. In the model, individuals in each time period allocate one unit of labor $L$ to either agricultural production, non-agricultural work, or violent conflict to maximize total net income $I$. Returns to all activities are affected by individual and location characteristics $X$ such as land quality, level of education, and fighting ability.

Net returns to agricultural production $F^A(L^A, S, W, X)$ are affected by agricultural shocks

[13] Becker (1968) uses a similar setup to model interpersonal conflict such as theft.

$S$, which may vary across individuals over space and in intensity with a larger $S$ indicating a more severe negative shock. Agricultural production is a concave function of $S$, with $\frac{\partial F^A}{\partial S} < 0$ and $\frac{\partial^2 F^A}{\partial S^2} < 0$. A larger $S$ therefore reduces agricultural labor productivity—the *opportunity cost mechanism*.

Agricultural production also depends on wealth $W$ with $\frac{\partial F^A}{\partial W} > 0$, where wealth broadly includes human, physical, and financial capital. Wealth in period $t$ is weakly increasing in income $I$ from activities in period $t-1$. As agricultural shocks decrease income, this creates a relationship between past agricultural shocks $S_{t-s}$ and agricultural production in period $t$, where $s \in [1, \tau]$ for some $\tau$. We can write $F_t^A = F^A(L_t^A, S_t, W_t(\{S_{t-s}\}_{s=1}^{\tau}), X_t))$, with $\frac{\partial F_t^A}{\partial W_t} > 0$ and therefore $\frac{\partial F_t^A}{\partial S_{t-s}} < 0$. This is the permanent income or *wealth mechanism*.[14]

Net returns to non-agricultural work $F^N(L^N, X, W)$ are based on the most productive activity available outside of own agricultural production. The highest returns available depends on individual and location characteristics $X$ and wealth $W$. As a simplifying assumption, I suppress the direct dependence of non-agricultural returns on $S$. Returns to non-agricultural work thus set a lower bound on how far the opportunity cost of fighting may fall following a negative agricultural shock in the short term. $F^N$ will be weakly smaller for individuals primarily engaged in agriculture that experienced a past agricultural shock due to the wealth mechanism.

Individual $i$ can also decide to engage in predatory violent conflict with a set of potential targets $J$ near $i$ that are feasible to attack within the time period. The potential net returns $F^C(L_i^C, X_i, \{I_j, W_j, X_j\}_{j \in J})$ depend on the incomes (production output), wealth (factors of production), and characteristics of the individuals in $J$. Wealth is valued based on the expected discounted stream of returns. Agricultural shocks $S_j$ for individuals $j \in J$ engaged in agriculture affect $i$'s returns to fighting by decreasing the income available to capture: $\frac{\partial F_t^C}{\partial S_{j,t}} < 0$. This is the *short-term predation mechanism*. Past agricultural shocks to individuals $j \in J$ will reduce both the output and the factors that $i$ can capture through the wealth mechanism, meaning $\frac{\partial F_t^C}{\partial S_{j,t-s}} < 0$. This is the *long-term predation mechanism*.

The probability of success and costs of fighting depend on characteristics $X_{i,t}$ and $X_{j,t}$—some targets will be farther away or have better fighting ability. An important variable in $X_{i,t}$ is whether there are groups engaged in armed conflict nearby, as joining such a group will reduce the costs of fighting for the individual and increase the probability of success. Costs of fighting are incurred with certainty and include economic, social, and emotional costs as well as risk of injury or death. These costs make fighting sub-optimal for most individuals in most time periods.

[14] Note that wealth in period $t$ is a function of many factors other than past agricultural shocks, including other types of shocks as well as past conflict, but I focus on the role of past agricultural shocks to build intuition with this model.

The individual's problem in period $t$ can be presented as choosing their labor allocation $L_{i,t}$ to maximize income $I_{i,t}$ given some current and past shock realizations $S_i, S_j$. For simplicity and intuition I ignore uncertainty in returns and suppose that decisions are made (or equivalently, updated) after the agricultural shocks in the period are realized.

$$\begin{aligned}\max_{L^A_{i,t},L^N_{i,t},L^C_{i,t}} \quad I_{i,t} =& F^A(L^A_{i,t}, S_{i,t}, W_{i,t}(\{S_{i,t-s}\}^{\tau}_{s=1}), X_{i,t}) + F^N(L^N_{i,t}, W_{i,t}(\{S_{i,t-s}\}^{\tau}_{s=1}), X_{i,t}) \\ &+ F^C(L^C_{i,t}, X_{i,t}, \{I_{j,t}, W_{j,t}(\{S_{j,t-s}\}^{\tau}_{s=1}), X_{j,t}\}_{j\in J}) \\ \text{subject to } & L^O_{i,t} \in \{0,1\}, \quad F^O(0,.) = 0, \quad \text{and } \sum_O L^O_{i,t} = 1 \text{ for } O \in \{A,N,C\} \\ & \frac{\partial F^A}{\partial S_{i,t}} < 0; \quad \frac{\partial F^A}{\partial S_{i,t-s}} < 0; \quad \frac{\partial F^C}{\partial S_{j,t}} < 0; \quad \frac{\partial F^C}{\partial S_{j,t-s}} < 0\end{aligned}$$

This yields

$$\begin{aligned}L^C_{i,t} = 1 \text{ iff } & F^C(1, X_{i,t}, \{I_{j,t}, W_{j,t}(\{S_{j,t-s}\}^{\tau}_{s=1}), X_{j,t}\}_{j\in J}) \\ & \geq \max(F^A(1, S_{i,t}, W_{i,t}(\{S_{i,t-s}\}^{\tau}_{s=1}), X_{i,t}), F^N(1, W_{i,t}(\{S_{i,t-s}\}^{\tau}_{s=1}), X_{i,t}))\end{aligned}$$

In words: actor $i$ chooses to engage in violent conflict if the net returns from fighting exceed their opportunity cost: the highest net returns they could receive from choosing another occupation.

Conflict occurs in the locations of the individuals being attacked. In this paper I analyze conflict at the level of grid cells which contain many individuals. I test the sensitivity of the results to using larger grid cells to capture spillovers of conflict outside the areas affected by agricultural shocks.

### 1.3.1 Effects of agricultural shocks on conflict

The effect of an agricultural shock on the decision to fight in the same time period is ambiguous, particularly if there is a strong positive correlation between shocks over space, as in most agricultural shocks. When the shocks are correlated, larger $S_{i,t}$ will decrease $i$'s returns to agricultural production in the same period but also be associated with a decrease in output available to capture from nearby targets for predatory attacks. This makes conflict over output (i.e., banditry) less attractive, but has minimal effect on the returns to conflict over factors of production.[15] We would therefore expect a larger impact of a transitory agricultural shock on conflict over factors of production in the same period than on conflict over

[15] I focus on transitory agricultural shocks which do not have a permanent direct effect on local agricultural productivity. Transitory shocks may have some effect on the returns to fighting over factors if they affect

output. Overall impacts on violent conflict depend on whether outputs or factors make up a greater share of the returns fighting.

Reductions in agricultural output available to capture are particularly severe for desert locust swarm shocks. Swarms consume all types of vegetation and there are no effective methods for farmers to limit damage to their output, meaning levels of agricultural destruction will therefore be similar across individuals with different characteristics in the affected area. Shocks to agricultural production from weather deviations (lower rainfall, higher temperatures) may have more heterogeneous and less severe effects on production depend on the extent of the deviation.

The long-term effects of past agricultural shocks $S_{i,t-s}$ on the decision to engage in conflict in period $t$ also involve offsetting mechanisms. Because of impacts on wealth as a result of consumption smoothing, the returns to both agricultural production and to fighting will be lower than before the shock in affected areas relative to unaffected areas, though higher than in the period of the shock. As with short-term impacts of an agricultural shock, long-term impacts should be smaller for conflict over output than over factors of production. Land is likely to be the main factor of production available to capture and its value should not be much affected by a transitory agricultural shock, though increased conflict risk following the shock could reduce land values.

The wealth effect is likely to be particularly strong following desert locust swarm exposure due to the severity of the income shock relative to small decreases in rainfall or increases in temperature from local averages. If the wealth mechanism is important, this should be observable in long-term impacts on measures of assets and agricultural productivity.

Whether a long-term productivity reduction increases the risk of violent conflict depends on the relative magnitudes of the reductions in productivity and in returns to fighting, but also on the absolute returns to fighting. Decreases in agricultural productivity will not be expected to increase violent conflict if the returns to fighting are very low (or negative). Since violent conflict is not the norm in most locations and periods, it implies the returns are generally low. A persistent decrease in agricultural productivity would therefore be expected to increase conflict risk primarily in locations or periods where the returns to fighting are higher.

One factor affecting these returns is the possibility of forming or joining an armed group, as this this reduces the costs of fighting and increases the probability of success. In practice, individuals are unlikely to engage in violent conflict alone, as such fighting generally involves organized armed groups which recruit members and pay them a wage or share of the returns

individuals' ability to productively utilize factors or if they affect expectations about future productivity. Shocks that have direct permanent productivity effects, for example through soil erosion or other land degradation, would have larger effects on the returns to capturing factors of production.

from victory (Collier and Hoeffler, 2004; Grossman, 1999).

While there is evidence that agricultural shocks motivate the formation of fighting groups and cause the onset of new violent conflict immediately following the shock (Harari and La Ferrara, 2018; McGuirk and Burke, 2020), it is not clear whether lower productivity in the long term after a shock would lead fighting groups to form in the absence of some precipitating event. One proxy for the possibility of joining an armed group in a particular location and period could be the activity of groups engaged in violent conflict nearby. Given some proximate cause for groups to form and engage in violent conflict, individuals with persistently lower agricultural productivity should be easier to recruit as their opportunity cost is lower.

While the model focuses on violent conflict with the objective of capturing outputs or factors of production, agricultural shocks could also affect non-violent forms of conflict such as protests (Hastings and Ubilava, 2023). Such activity does not target the capture of agricultural output or wealth and thus will not be affected by the predation mechanism, but would still be affected by a decrease in agricultural production as the opportunity cost of participating falls. Individuals may also derive value from expressing their grievances or motivating increased attention and relief. Effects of an agricultural shock should be larger for protests than for violent conflict, since there is no offsetting effect through the predation mechanism.

To summarize, the model informs a set of testable predictions for the impacts of a transitory agricultural shock on conflict risk:

1. The impact on the likelihood of violent conflict in the year of the shock is ambiguous, but if the predation mechanism offsets the opportunity cost mechanism, a shock should increase the short-term likelihood of factor conflict and protests by more than the likelihood of output conflict.

2. If a transitory but severe shock has long-term effects through a wealth mechanism, this should be observable through long-term reductions in measures of productivity.

3. The impact of a shock on the long-term likelihood of violent conflict is ambiguous, but if opportunity costs of fighting fall and the returns to predation are an important mechanism we should observe different impacts across types of conflicts. Impacts on the likelihood of protests should be larger than on the likelihood of violent conflict, and larger for conflict over factors of production than for conflict over output.

4. If there are long-term effects of a severe shock on opportunity costs of fighting through the wealth mechanism and the returns to predation are an important mechanism, the

long-term risk of violent conflict should increase by more in periods when fighting groups are active.

## 1.4 Data

The Locust Watch database (FAO 2022b) reports observations of desert locust swarms as well as smaller concentrations of locusts from 1985 to the present. I consider only data on locust swarms which pose the greatest threat to agriculture and whose flight patterns create local variation in exposure. The Locust Watch data include latitude, longitude, and date of swarm observations. Locust observations are recorded by national locust control and monitoring officers on the ground, but incorporate reports from agricultural extension agents, government officials, and other sources. Local farmer scouts are also often trained in locust monitoring and reporting (Thomson and Miers, 2002).

A concern might be that locust reporting is correlated with violent conflict. Showler and Lecoq (2021) review how insecurity has affected national and international desert locust control operations from 1985-2020 across countries where locusts are active. They mention Chad, Mali, Somalia, Sudan, Western Sahara, and Yemen as countries with areas where insecurity has constrained locust control operations in certain periods since 1997. This concern is the focus of Torngren Wartin (2018)'s analysis of the impact of locusts on conflict, which uses similar data but focuses on the short-term, modeling locusts as temporary shocks.

Insecurity is likely less of a constraint for locust monitoring than for control operations. FAO locust monitoring guidelines discuss conducting aerial surveys and using reports from local scouts, agricultural extension agents, security forces, and other sources (Cressman, 2001), which would allow reporting even in insecure areas. The Locust Watch data includes observations of locust swarms even in countries and periods where Showler and Lecoq (2021) indicate control operations have not been possible. For example, the authors mention that control operations in Western Sahara have been largely infeasible due to Polisario activity over the whole sample period, but 166 swarms have been recorded there in 9 different years from 1996-2018. None of the monthly FAO locust swarm bulletins published during the 2003-2005 upsurge—the major locust event in the sample period—mention issues related to insecurity affecting locust monitoring efforts. The share of cells within 50km of a locust swarm observation in a given year that have reports of both violent conflict and a locust swarm in the cell is 27% in the set of countries Showler and Lecoq (2021) indicate pose challenges for locust control, similar but below the 34% in all other countries. I test the sensitivity of the results to excluding the countries listed in the report as potential locations of locust swarm under-reporting, and to imputing 'missing' locust swarms near the locations of reported swarms.

Data on conflict events come from the Armed Conflict Location & Event Data Project (ACLED) database (Raleigh et al., 2010). The database records the location, date, actors, and nature of conflict events globally starting from 1997 by compiling and validating reports from traditional media at different levels, from institutions and organizations, from local partners in each country, and from verified new media sources. The analysis focuses on events categorized by ACLED as "violent conflict," which includes battles, explosions, and violence against civilians. I also test impacts on protest and riot events recorded by ACLED and on larger-scale violent conflicts from the Uppsala Conflict Data Program (UCDP; (Sundberg and Melander, 2013) and distinguish between conflict that does and does not involve a non-state actor, as these types of conflict will involve different mechanisms. The UCDP database is similar to the ACLED database but goes back to 1989 and only records conflicts involving at least one "organized actor" and resulting in at least 25 battle-related deaths in a calendar year. The ACLED database has no organized actor or minimum death threshold requirements. McGuirk and Burke (2020) characterize UCDP events as more likely to represent conflict over territory and factors of production, and I follow them in constructing a measure of output conflict (i.e., banditry) using ACLED records of violence against civilians, rioting, and looting.

I collapse the data to a raster grid with annual observations for cells with a 0.25° resolution (15 arcminutes, approximately 28×28km). Analyzing impacts at this spatial level reduces potential measurement error about the specific areas affected by swarm and conflict events and allows me to leverage local variation in swarm presence created by their flight patterns. Nearly all swarms will be contained within 0.25° cells ($\sim$784km$^2$), except those near cell boundaries. I test for robustness to analyzing data at the level of 0.5° and 1° cells, which will also capture potential spillovers from swarm exposure (McGuirk and Nunn, 2021). In each cell and year I measure whether any locust swarm and conflict event was recorded.

I determine the country and highest sub-national administrative level in which each cell centroid lies using country boundaries from the Global Administrative Areas (2021) database v3.6. I use sub-national boundaries to create a set of 170 regions, all of which include at least 32 individual grid cells. These regions are either existing sub-national administrative units or combinations of adjacent units within the same country. I cluster standard errors at the level of these regions.

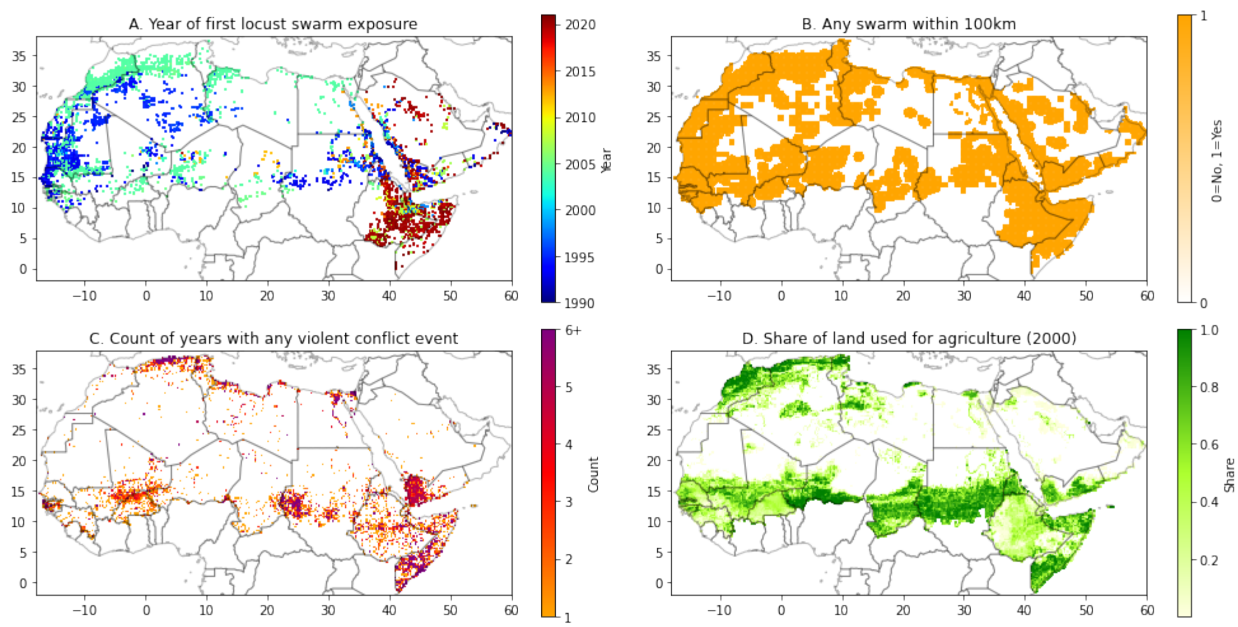
Given the role of weather in desert locust biology, its importance in determining agricultural production, and the well-documented relationship between weather variations and conflict, all analyses control for local weather to isolate the impact of locust swarm exposure. I measure total annual precipitation (in mm) and maximum temperature (in °C) using high-resolution monthly data from WorldClim available through 2018.[16] I also incorporate

[16]CRU-TS 4.03 (Harris et al., 2014) downscaled with WorldClim 2.1 (Fick and Hijmans, 2017).

raster population data for every 5 years from CIESIN (2018), linearly interpolating within cells between years where the population is estimated, and raster data on land cover in 2000 from CIESIN, giving the share of land cover in each cell that is cropland and pasture (Ramankutty et al., 2010). I include additional cell characteristics from the PRIO-GRID dataset (Tollefsen et al., 2012), assigning all 0.25° cells the values for the 0.5° cell in which they are located.

Since ACLED records conflicts beginning in 1997 and the weather data are available until 2018, the analysis sample includes observations from 1997 to 2018. I restrict the analysis to countries with at least 10 locust swarm observations in this period. These countries include all of North Africa, most of the Arabian Peninsula and West Africa, and the Horn of Africa. The resulting analysis sample covers 22 years across 24,460 cells, for a total of 538,086 observations. Figure 1.2 visualizes swarm exposure, violent conflict incidence, and agricultural land cover for the sample countries. Summary stats are included in Table A.1.1.

Figure 1.2: Swarm exposure, violent conflict incidence, and land cover in sample countries




Note: Land used for agriculture includes crop land and pasture land. Panels B and D show most clearly which countries in West, Central, and East Africa are excluded from the study sample.

A locust swarm is recorded in the Locust Watch database between 1985-2021 for 18% of cells, and 63% of cells are within 100km of a locust swarm report (Figure 1.2 Panel B). For each cell, I identify the first year after 1989 in which a locust swarm is recorded (Figure 1.2

Panel A), and define a cell as exposed to a locust swarm in each following year and not exposed in all other years or if no locust swarm is ever observed.[17] Locations where locust swarms are observed in more than one year (12.6% of exposed cells) are not distinguished from those where they are observed only once. Cells exposed to a swarm before 1997 (in dark blue in Figure 1.2 Panel A) are treated during the entire sample period and are therefore dropped from the analyses, while cells exposed to a swarm after 2018 (in red) are considered not treated during the sample period. 7.5% of cells are exposed to a swarm during the sample period, including 5.4% exposed during the 2003-2005 upsurge (in teal).

Fourteen percent of cells experienced at least one violent conflict event, in 3.4 different years on average (Figure 1.2 Panel C). Protests and riot events and large-scale violent conflict events as defined in the UCDP are both recorded at least once in 8% of cells. About half the cells (54%) in the sample include some agricultural land: 53% have pasture land while 29% have crop land. Across all cells, the mean share of land allocated to agriculture is 24% (Figure 1.2 Panel D), with 19% pasture land and 5% crop land.

## 1.5 Empirical approach

I estimate the causal impacts of locust swarms on conflict using a difference-in-differences approach allowing for long-term effects of this transitory agricultural shock. Cells are defined exposed to (or affected by) a locust swarm in all years starting from the first year a swarm is observed in a cell, and not exposed before or if no locust swarm is ever observed.

I estimate both static average impacts of locust swarm exposure using two-way fixed effects (TWFE) models and dynamic impacts over time using event study approaches. The TWFE linear probability models take the form:

$$Conflict_{ict} = \alpha + \beta Exposed_{ict} + \delta X_{ict} + \gamma_{ct} + \mu_i + \epsilon_{ict} \tag{1.1}$$

where $i$ indexes cells, $c$ countries, and $t$ years. $Conflict$ is a dummy variable for observing any conflict event and $Exposed$ is a dummy variable for having been exposed to a locust swarm. The primary specifications focus on impacts on violent conflict using the ACLED data. I consider effects on other measures of conflict in tests of the impact mechanisms. Analyzing conflict as a binary variable at an annual level reduces potential measurement error and is the main approach in the literature. $\gamma_{ct}$ are country-by-year fixed effects, and $\mu_i$

[17] A major locust upsurge occurred from 1985-1989. 1,195 cells where a locust swarm was recorded between 1985 and 1989 but in no year afterward are treated as not exposed to locust swarms during the 1997-2018 sample period. Treatment timing for another 997 cells where a locust swarm was recorded between 1985-1989 is based on the first year after this period that a locust swarm was recorded. The results are robust to basing treatment timing on the first year any swarm is recorded in a cell in the Locust Watch database.

are cell fixed effects. $X_{ict}$ is a vector of time-varying controls at the cell level. Standard errors (SEs) are clustered at the sub-national region level (170 clusters) to allow for correlation in the errors within nearby areas over time.[18]

I estimate event study models considering the 14 years before and after locust swarm exposure using the staggered treatment difference-in-differences methods developed in Borusyak et al. (2021)—which I will refer to as BJS—and Callaway and Sant'Anna (2021)—which I will refer to as CS.[19] I include the same fixed effects and clustering as in Equation 1.1. These two event study approaches deal with concerns with TWFE estimators when there is heterogeneity in treatment effects by time since treatment or across treatment cohorts, which can lead to 'forbidden' comparisons between late- and early-treated groups and negative weighting of effects for certain treatment groups or periods (Goodman-Bacon, 2021). Both methods effectively estimate an average treatment effect on the treated in each time period separately for groups exposed in different years $g$. Event study estimates are then calculated by taking weighted averages of these treatment effects. The main difference between the two approaches is that the CS method takes as a base value the year prior to treatment while the BJS imputation estimator makes comparisons against the average over all pre-treatment periods (Roth et al., 2023). The methods imply different assumptions about parallel trends but generally give quite similar results. I primarily present results using the BJS method.

In addition to the analyses considering all swarm exposure events I also specifically estimate impacts of exposure to the 2003-2005 major desert locust upsurge, dropping cells exposed to swarms in other years. This is the only major locust upsurge in the sample period (1997-2018), and accounts for 72% of first cell locust swarm exposure. This is a 'canonical' difference-in-differences analysis with the upsurge 'treatment' occurring in the same period for all treated units and a comparison group that never receives this treatment. I treat 2004 as the year of first exposure as the small number of swarms in 2003 at the beginning of the upsurge arrived in the last months of the year. Trends for violent conflict events and locust swarms prior to the upsurge are similar across cells that were and were not affected by the 2003-2005 upsurge supporting the parallel trends assumption, and locust swarm presence is similar following the upsurge (Figure A.1.1).

To test for heterogeneity in the impacts of swarms, I estimate Equation 1.1 fully interacting the right-hand side variables with another variable of interest. I test robustness of the

[18]This is likely more restrictive than necessary and will lead to a conservative interpretation of the results. Patterns of statistical significance are largely unchanged when using two-way clustered errors at the year and region level and using Conley (1999) Heteroskedasticity and Autocorrelation-Consistent (HAC) SEs allowing for spatial correlation over 100 and 500km and serial correlation over 0 or 10 time periods, following Hsiang (2010)'s approach.

[19]Specifically, I use the Stata packages *did_imputation* (Borusyak et al., 2021) and *csdid* (Callaway and Sant'Anna, 2021), and generate event study plots using Stata's *event_plot* package (Borusyak et al., 2021).

results to different controls and fixed effects, restrictions of the analysis sample, cell sizes, and clustering of SEs. Results of robustness tests are included in Appendix A.3.

The key identifying assumption of this design is that trends in conflict risk would be parallel over time in affected and unaffected areas within the same country in the absence of locust swarm exposure, after controlling for effects of weather, population. I find no significant differences in the probability of violent conflict by swarm exposure in the pre-exposure periods, indicating no differential conflict risk pre-trends (Figure 1.3). Though this does not preclude the possibility that trends would differ in the years after swarm exposure for reasons unrelated to agricultural destruction, it is an encouraging sign that the parallel trends assumption may be likely to hold.

Cells exposed to a locust swarm during the sample period have different baseline characteristics than unexposed cells which are largely consistent with desert locusts rarely being observed in the interior of the Sahara desert (as shown in Figure 1.2); among cells with agricultural land only mean annual rainfall and the share of crop land differ with swarm exposure status (Table A.1.2). Differences are similar when comparing cells exposed to the 2003-2005 locust upsurge to unexposed cells.

Baseline differences are not a concern if they only affect levels of civil conflict and not trends. Cell fixed effects control for time invariant cell characteristics that might affect the risk of conflict such as distance to major cities or country boundaries, topography, and agricultural suitability, as well as factors affecting risk of locust exposure such as distance from locust breeding areas and typical seasonal wind patterns. Country-by-year fixed effects flexibly control for factors varying over time at the country level that might affect conflict risk and the impact of agricultural shocks, such as food price shocks, weather patterns, the policy environment and national economic and social conditions. Importantly, they also control for trends in violent conflict incidence, which increases over the sample period. To isolate impacts of locust swarm exposure from other potentially time-varying factors affecting conflict risk, I include controls for characteristics that differ between exposed and unexposed cells. In particular, my preferred specification includes total annual precipitation (in mm), the maximum annual temperature (in °C), and annual population.[20]

I test the robustness of the results to efforts to increase the plausibility of the parallel trends assumption by imposing different constraints on the areas included in the comparison sample and to using inverse propensity weights based on the estimated probability of locust swarm exposure (Stuart et al., 2014).[21] Exposed and non-exposed cells are well-balanced on

[20]Results are robust to including squared temperature and rainfall terms or used logged versions and to including 1 year weather lags. I do not have time-varying data on land cover, but results are not sensitive to accounting for changes in crop land over time as documented in Xiong et al. (2017).

[21]I calculate propensity scores using a logit regression with a dummy for being exposed to a locust swarm on baseline cell characteristics and mean weather and country fixed effects. I calculate inverse propensity

baseline characteristics after include these inverse propensity weights (Table A.3.1), suggesting the parallel trends assumption may be more likely to hold.

Another identification assumption relevant to event study designs with staggered treatment timing is the no anticipation assumption: knowledge of future treatment timing does not affect current outcomes (Roth et al., 2023). Populations may expect a higher probability of swarm exposure in years of major upsurges but cannot perfectly anticipate timing of exposure. For example, the FAO Desert Locust Watch publishes monthly forecasts of areas predicted to be at risk of locust swarm exposure but the predictions include a great deal of uncertainty due to unpredictable variation in swarm flight patterns. Consequently, areas forecast to be at risk are generally quite large, the majority of which end up not being affected by locusts.[22] Anticipation may also have limited effects as there are no effective methods of defending vegetation against locust swarms, and farmers in at-risk areas typically describe locust prevention and control as out of their hands and the responsibility of governments (Thomson and Miers, 2002).

## 1.6 Results

### 1.6.1 Dynamic impacts of swarm exposure on violent conflict

Figure 1.3 Panel A presents event study estimates of the impacts of swarm exposure on the risk of violent conflict over time using the staggered treatment timing difference-in-differences approach developed in Borusyak et al. (2021). The pattern of results is similar using the Callaway and Sant'Anna (2021) method (Figure A.3.1). There is no significant difference in conflict risk between cells that are and are not exposed to locusts in the years prior to exposure, supporting the parallel trends assumption.[23]

The point estimate for the effect on violent conflict risk in the year locusts arrive is -0.002 but not statistically significant. All other estimated treatment effects are positive

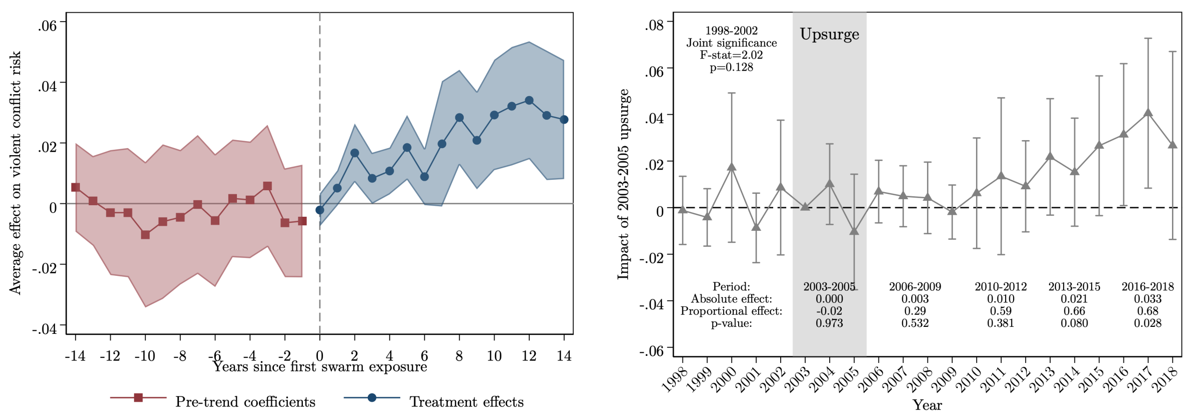
weights as $\frac{1}{p}$ for cells that were exposed and $\frac{1}{1-p}$ for cells that were not, where $p$ is the estimated probability of swarm exposure. I assign cells with estimated probabilities outside the range of common support a weight of 0. I separately calculate inverse propensity weights for any swarm exposure and for exposure during the 2003-2005 upsurge.

[22]Monthly forecasts during the major upsurge in 2004 on average covered 40.6% of 0.25° cells in sample countries. 78% of recorded swarms across were in areas forecast to be at risk in the month the swarm was observed, and being mentioned in a monthly forecast is associated with a significant but small 2% increase in the likelihood of being exposed to a locust swarm.

[23]The average difference in the pre-exposure period is -0.005. None of the 14 pre-exposure period differences are significantly different from 0 and I fail to reject that pre-period differences are jointly equal to 0 ($p$=0.236).

Figure 1.3: Impacts of exposure to locust swarms on violent conflict risk over time

A) Exposure to any swarm　　B) Exposure to a 2003-2005 upsurge swarm




Note: The dependent variable is a dummy for any violent conflict event observed.
Panel A shows event study results for impacts of exposure to any locust swarm on violent conflict risk over time using the Borusyak et al. (2021) approach. Estimated impacts in each time period are weighted averages across effects for swarm exposure in particular years. Time period 0 is the year of swarm exposure. A joint test that the pre-exposure coefficients equal 0 gives $p = 0.236$.
Panel B shows coefficients for the interaction of a dummy for being exposed to a 2003-2005 upsurge swarm with year using 2003 as the reference period. In the top left is the result from a joint test of the hypothesis that all pre-upsurge coefficients equal 0. Estimates from the same specification with binned years are reported at the bottom of the figure. Proportional effects are relative to the probability of observing any violent conflict during the particular time period. $p$-values are for tests of the null of 0 impact of the upsurge in each period.
Shading and bars represent 95% confidence intervals using SEs clustered at the sub-national region level. All regressions include country-by-year and cell fixed effects. The analyses in Panel B also includes controls for rainfall, temperature, and population and inverse propensity weights. Observations are grid cells approximately 28×28km by year. Locust swarm exposure is based on the first year a locust swarm is observed in a cell. Cells with locust swarms observed before the start of the sample period are not included in the analyses, and cells exposed outside of 2003-2005 are not included in Panel B.

and statistically significant. Impacts in years 1-6 post-exposure are relatively smaller—the average effect is a 1.2 percentage point increase in conflict risk—and for 3 of years the effects are only marginally statistically significant. The impacts are larger in years 7-14 post-exposure and are highly significant after year 7 even as the standard errors are larger with more time since exposure. Swarm exposure causes an average increase in violent conflict risk in a given year of 2.8 percentage points over these periods.

Figure 1.3 Panel B shows the results of an event study analysis of the 2003-2005 upsurge. I include inverse propensity weights to account for differences in which countries in the sample were affected by this upsurge. Patterns are similar but estimated effects are larger and more frequently statistically significant when not including inverse propensity weights

(Figure A.3.4). Effects are all estimated relative to 2003, when a few locust swarms were observed near the end of the year. There are no significant differences in the risk of violent conflict between areas exposed to locust swarms during this upsurge and areas that were not in the years preceding the upsurge, and I cannot reject the joint hypothesis that the pre-trend coefficients are equal to 0 ($p$=0.128). This again supports the assumption of parallel trends between these areas if not for the locust swarm exposure.

There is no significant impact of locust exposure on the risk of conflict in the main years of the upsurge, 2004 and 2005, as in the exposure year estimates in Figure 1.3 Panel A. Estimated impacts of the 2003-2005 upsurge on conflict in the following years are almost all positive (the coefficient for 2009 is slightly negative), and generally become larger in magnitude over time starting after 2009. The estimated effects for individual years are noisier and only statistically significant at a 90% confidence level or greater for 2013 and 2015-2017.

Average estimates over bins of years are shown at the bottom of the figure and are more precise. Being exposed to a 2003-2005 upsurge swarm does not significantly affect violent conflict risk from 2006-2012, but increases the annual risk by 2.1 percentage points on average between 2013-2015 and by 3.3 percentage points between 2016-2018, and these effects are statistically significant. An advantage of the upsurge event study is the ability to compare effects against violent conflict incidence in control areas at different points in time. The effect of the upsurge represents a 67% increase in violent conflict risk relative to the mean in unaffected areas from 2013-2018, as violent conflict in the sample countries increased over time due to a variety of factors (Figure A.1.1 Panel A).

In summary, after null effects on violent conflict risk in the short-term immediately following swarm exposure, the annual probability of any violent conflict event is significantly higher than in unaffected cells in the years post-exposure. The largest impacts are realized around 7-14 years after swarm exposure. Patterns are similar when considering impacts of the exposure to the 2003-2005 locust upsurge in particular, though treatment effects are not significant for the first several years after exposure. The large and statistically significant increases in violent conflict risk in areas affected by the upsurge from 2013-2018 (8-14 years after exposure) align with the periods of highest estimated effects in the main event study and also with the timing of a general increase in violent conflict in the study sample. This suggests heterogeneity in impacts by the activity of fighting groups discussed in the model, which I return to in Section 1.7.

### 1.6.2 Average impacts of swarm exposure on violent conflict

I now estimate average long-term impacts of swarm exposure on the annual risk of violent conflict to compare estimated effects against those of annual weather deviations and test heterogeneity by land cover. Figure 1.4 Panel A shows that on average cells exposed to locust swarms are 2.1 percentage points (pp) more likely to experience any violent conflict in a given year in the period after swarm exposure than cells not exposed. This estimate is very close to the average of the treatment effects in Figure 1.3—a 1.9pp increase—indicating limited bias in the TWFE estimates from staggered timing of swarm exposure. A 2.1pp increase on the probability of any violent conflict in a year represents an 88% increase over the mean for cells not exposed to locusts.

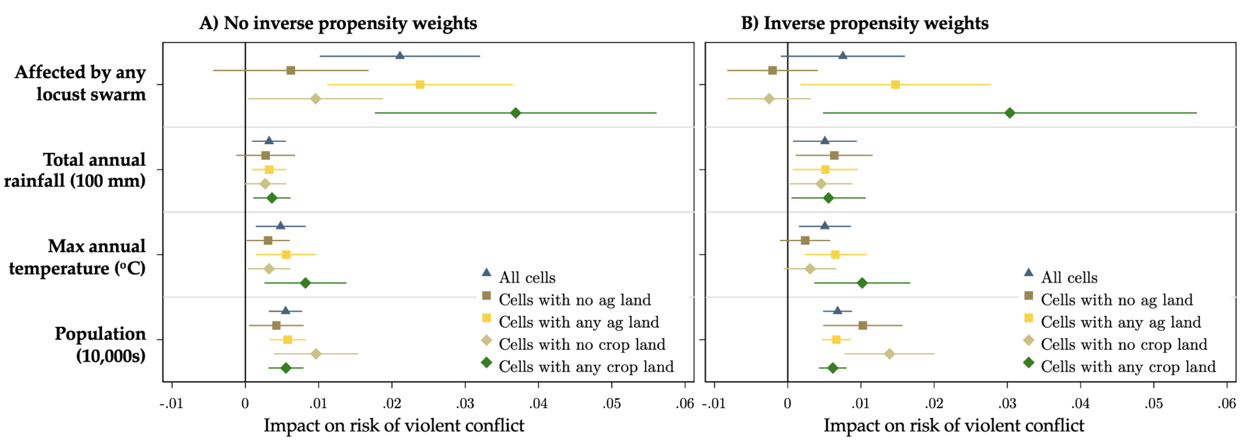
This effect is large compared to the effects of weather deviations. An increase in annual precipitation of 100mm relative to the average in the cell during the sample period increases the probability of violent conflict in the same year by 0.3pp (13%), while an increase in the maximum annual temperature of 1°C relative to the cell average increases conflict risk in the same year by 0.5pp (21%). These effects fall in the upper middle range of estimates reported in Burke et al. (2015)'s meta-analysis of the impacts of weather deviations on conflict. Cell population is also positively associated with conflict risk, with an increase of 10,000 people associated with a 0.5pp (21%) increase in the probability of any violent conflict event in a year.

The average impact of locust exposure remains statistically significant at the 99% confidence level under other forms of standard error clustering, including two-way clustering at the region and year level and using Conley (1999) SEs allowing for spatial correlation within 100 and 500km and serial correlation over 0 or 10 years (Figure A.3.5). Associations between weather deviations and population and violent conflict risk are also statistically significant at the 95% confidence level or greater under all clustering choices. Clustering at the sub-national region level consistently leads to SEs at least as large as Conley SEs allowing for spatial correlation within 500km and serial correlation over 10 years, implying the main SEs I report with region clustering are conservative and may understate statistical significance of certain estimates.

I test for differences in the impacts of swarm exposure and weather by whether a cell has any agricultural (crop or pasture) land or any crop land using measures of land cover from 2000 (Ramankutty et al., 2010).[24] Locust swarm exposure in non-agricultural cells (46% of the sample) has no significant effect while in agricultural cells annual violent conflict risk increases by 2.4pp. Locust swarms increase annual violent conflict risk by 3.7pp in crop

[24]Southward expansion of the Sahara desert, anti-desertification efforts, deforestation, changing seasonal distribution of precipitation, and expansion of farming in traditional pastureland have all contributed to changing land cover over the study period (Davis, 2022; Liu and Xue, 2020; Rahimi et al., 2021). Xiong et al. (2017) report that from 2003-2014 croplands increased by 1 Mha per year on average. As a result, some cells with crop land may be inaccurately classified as non-crop cells in this analysis based on land cover in 2000, which would reduce the estimated difference in impacts by cell land cover. Results are similar when defining crop cells as those with any cropland in either Ramankutty et al. (2010) or Xiong et al. (2017).

Figure 1.4: Average impacts of exposure to locust swarms on violent conflict risk by land cover




Note: The dependent variable is a dummy for any violent conflict event observed. Coefficients and 95% confidence intervals are from three separate regressions in each panel: one with no land cover interactions and the other two interacting all right-hand side variables with cell land cover dummies. The coefficients and standard errors for cells with any agricultural (pasture or crop) land and with any crop land are calculated using Stata's *xlincom* command based on the sums of the coefficients for the baseline effect and the interaction term. Inverse propensity weights included in the regressions in Panel B are calculated based on the estimated propensity for cells to have been exposed to a swarm. Cells are regarded as 'affected' by locust swarms for all years starting from the first year in which a swarm is recorded in the cell. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses. Results from the regressions are reported in Table A.1.3.

cells (29% of the sample) compared to 1.0pp in non-crop cells (35% of which have pasture land), indicating that impacts on crop production drive most of the effect on conflict risk but impacts on pasture also explain part of the total effect. This aligns with desert locusts consuming all forms of vegetation.

There is no significant difference in the effects of higher rainfall by land cover, echoing previous work questioning whether agricultural mechanisms explain the relationship between rainfall and conflict (Bollfrass and Shaver, 2015; Sarsons, 2015). Impacts of higher temperatures are larger in by cells with crop land, where a 1°C hotter year increases annual violent conflict risk by 0.8pp. But higher temperatures also increase conflict in non-agricultural cells. The association between population and conflict risk does not vary significantly with land cover.

Panel B of Figure 1.4 shows the results from the same regressions weighting observations by the inverse of the estimated propensity to have been exposed to a locust swarm during the study period. Estimated magnitudes for the impacts of swarm exposure are smaller than in Panel A with no weights, suggesting part of those estimated effects are due to comparisons between cells affected by locust swarms and dissimilar unaffected cells with different trends in violent conflict risk. But the patterns and qualitative conclusions are the same. On average, locust swarm exposure increases annual violent conflict risk by 0.8pp in subsequent years. This is significant at the 90% confidence level and represents a 43% increase relative to the mean in unaffected cells. Estimated effects in cells without pasture or crop land are close to zero and not statistically significant. Impacts are driven particularly by cells with any crop land where the annual risk of violent conflict increases by 3pp, compared to an increase of 1.5pp in cells with either crop or pasture land.

The pattern of swarm exposure impacts is very similar when specifically analyzing the 2003-2005 locust upsurge with upsurge-specific inverse propensity weights (Table A.1.3). This is not surprising since the 2003-2005 upsurge was the main swarm exposure event in the sample period. Upsurge exposure is associated with a 1.1pp (60%) increase in violent conflict risk over the following years on average. This is again driven exclusively by cells with crop or pasture land.

The heterogeneity in TWFE results by land cover is consistent with locust swarms representing shocks that affect the economy and society solely through their impacts on agricultural production, and particularly through damages to crop land. Effects also vary by timing of swarm exposure relative to the local crop calendar in ways consistent with this channel (Table A.1.4). I categorize swarms as arriving during particular stages of the crop production cycle by matching the month in which a swarm is observed to country-level crop calendars from The United States Department of Agriculture (USDA) (2022).[25] The off season—between harvesting and planting—lasts between 3 and 6 months in most of the sample countries, with an average of slightly over 4 months.[26] Focusing on the 2003-2005 upsurge I find that impacts are driven by exposure to upsurge swarms that arrived in crop cells during the country's crop growing season when damages to agriculture would be greatest. Exposure to upsurge swarms during the off-season or in cells with no crop land does not significantly affect conflict risk, though point estimates for impacts of off season swarms in crop and non-crop cells are relatively large, indicating effects through damages to pasture land and other vegetation. The finding relates to studies showing that the impact of weather shocks on conflict risk varies depending on whether the timing of the shock is such that it

[25]Figure A.1.2 shows example crop calendars from Libya and Mali. In countries with different agricultural cycles by crop, I identify the crop activity associated with the most commonly grown crops each month.

[26]Figure A.1.3 shows the share of sample cells at different stages of agricultural cycle by month and the counts of locust swarms during the sample period observed by season and region. I do not distinguish between planting, growing, and harvest season swarms in the analysis as the months assigned to these periods are approximate and effects should largely be similar. Most 2003-2005 upsurge swarms were recorded during months of the main growing season, which covers a greater share of the year, and the majority of swarms were in the 71% of cells with no crop land. A small number of cells were exposed to both off-season and growing season swarms.

is likely to decrease agricultural productivity (Caruso et al., 2016; Crost et al., 2018; Harari and La Ferrara, 2018).

### 1.6.3 Robustness

I test the sensitivity of the results to various alternative specifications and estimate similar impacts of locust swarm exposure on violent conflict risk. The results are included in Appendix A.3. In this section I summarize the main sensitivity checks and discuss any differences.

Estimated impacts are largely unchanged when varying the set of control variables included. Dropping weather and population controls leads to slightly larger magnitude effects on violent conflict risk and greater statistical significance in the main event study but do not affect the upsurge event study. Dynamic effects of upsurge exposure are also very similar when interacting the weather and population controls with year. TWFE estimates with and without inverse propensity weights are similar without controls and when including 1 year lags of rainfall and maximum temperature as additional controls.

Across all analyses the estimated treatments effect sizes are slightly larger and standard errors are smaller when using sub-national region-by-year fixed effects instead of country-by-year fixed effects. More local time fixed effects control for the possibility that swarm migratory paths may overlap with parts of countries with greater future risk of violent conflict. The similarity of the results increases confidence that the estimates reflect true long-term causal effects of swarm exposure. Using year fixed effects results in a similar pattern of treatment effects as in Figure 1.3, with slightly larger estimated magnitudes. TWFE estimates are smaller with year fixed effects and lose significance when also including inverse propensity weights, though impacts in crop cells remain significant. These differences across fixed effects indicate that impacts of swarm exposure are most notable when comparing exposed cells to nearby unexposed cells.

Results are similar when restricting the sample of control cells to those within 100km of any locust swarm exposure (shown in Figure 1.2). This ensures that the results are not driven by comparisons between exposed cells and unexposed desert cells with low risk of either swarm exposure or violent conflict, though the inverse propensity weights serve a similar function. I also obtain similar results when systematically omitting countries from certain regions (North, West, and East African and the Arabian Peninsula). This addresses concerns that the long-term impacts on violent conflict may be spurious and due to swarm exposure during the sample period being correlated with factors driving later conflict emergence. For example, dropping North Africa ensures that results are not driven by the Arab Spring and dropping Arabia ensures results are not driven by the civil war in Yemen.[27]

Estimated impacts on violent conflict risk of locust swarm exposure events are slightly smaller when dropping countries listed in Showler and Lecoq (2021) as areas where insecurity limited desert locust control operations during the sample period to addresses concerns about possible unreported locust swarms correlated with conflict risk.[28] This could indicate that some swarms were not reported in insecure areas and the cells are classified as not exposed to a swarm, biasing the estimated effect of exposure downward. In general, unreported swarms would bias the estimates downward if they also increase violent conflict risk on average, but they would bias estimates upward if swarms are more likely to be unreported in areas with persistently low risk of violent conflict. Both types of bias may exists and partially offset. The similarity of the results when excluding the most insecure countries from the sample and robustness to including inverse propensity weights (which exclude cells with low estimated probability of any locust swarm being reported) increase confidence that bias from unreported swarms is unlikely to be driving the estimated impacts of exposure on violent conflict risk.

Finally, I consider differences in estimates at the 0.5° and 1° cell levels. Using larger grid cells addresses several potential measurement issues. First, it minimizes the possibility that the area exposed to a locust swarm recorded in a cell exceeds the boundaries of the cell. Second, it reduces concerns about nearby areas that might have been affected by unreported swarms since the entire cell is considered exposed if any swarm is reported within it. Third, it limits the potential for conflict spillovers outside the cell. Downsides to analyzing impacts in a more coarse grid are dilution of treatment intensity (as the share of the cell affected by swarms weakly decreases with cell size) and the loss of local variation in swarm exposure, which the previous robustness checks show is important for estimating the impacts on violent conflict risk.

The pattern of dynamic swarm exposure effects over time is similar at different levels. Estimated magnitudes increase with cell size (though not in proportion to mean conflict risk) but the standard errors do as well. In the full event study, the negative effect in the year of exposure becomes statistically significant while some of the later treatment effects

[27] Cells in the Arabian peninsula first exposed to locust swarms during an outbreak in 2007 experienced a much larger increase in subsequent conflict risk than other cells first exposed to swarms during the sample period. The 2003-2005 locust upsurge largely did not affect Arabian cells, so estimates from the analysis of the impact of locust exposure during that upsurge will not be affected by the factors influencing the differential effect of locust exposure in Arabia.

[28] This is true when considering all locust swarm exposure events. Estimated impacts of the 2003-2005 upsurge are slightly larger when dropping these countries. The difference is likely because two of these countries, Yemen and Somalia, are already excluded from the upsurge analysis as the 2003-2005 upsurge did not affect these countries.

lose significance. Average impacts increase from a 2.1 percentage point rise in annual violent conflict risk at the 0.25° cell level to a 2.4pp increase at the 0.5° level and 4.1pp at the 1° levels, with the largest impacts concentrated in years 7-10 post-exposure. With inverse propensity weights, impacts of swarm are not significant at the 0.5 or 1° levels on average but are significant and larger for cells with crop land. Changes are similar for average impacts of the 2003-2005 locust upsurge at different scales. Larger estimated magnitudes when using larger grid cells despite dilution of treatment intensity suggest violent conflict spills over outside exposed areas. The consistent negative point estimate in the year of swarm exposure indicate this is not driven by spillover conflict outside affected cells. Nearby cells not exposed to locust swarms could be targets for predatory attacks as they are likely to have greater wealth in the long-term, having been spared the major income shock. Focusing on 0.25° cells may therefore understate the full effect of swarm exposure on violent conflict around affected areas.

## 1.7 Mechanisms

### 1.7.1 Short-term

The small and non-significant impacts of locust swarms in the year of exposure provide evidence that negative agricultural production shocks need not increase the risk of conflict in the same year. This contrasts with much of the literature on climate and conflict. One explanation for the difference is that the effect of weather shocks on violent conflict may primarily go through channels other than agricultural production. I find no significant difference in the effects of rainfall deviations by land cover while higher temperatures significantly increase conflict risk in both agricultural and non-agricultural areas (Figure 1.4), echoing previous work finding impacts of rainfall on conflict that cannot be explained by agricultural mechanisms (Bollfrass and Shaver, 2015; Sarsons, 2015). But some studies also find that growing season weather shocks have null or inconsistent effects on conflict in the same year, notably Crost et al. (2018); Harari and La Ferrara (2018). These studies emphasize increases in conflict risk the year after a shock, which I also find for locust swarms, and suggest a general mechanism suppressing immediate impacts of a shock on violent conflict.

A non-significant short-term effect of a negative agricultural shock on conflict risk is not surprising in the context of the model, as the decrease in opportunity cost of fighting is offset by a decrease in the returns to predation when the value of output and wealth available to capture is smaller. Studies of shocks to agricultural prices have shown instances where the predation mechanism outweighs the opportunity cost mechanism: McGuirk and Burke

(2020); Ubilava et al. (2022) both document increases in violent conflict in cells producing agricultural goods following increases in the global price of these goods.

If the null effect in the year of exposure results from these offsetting mechanism, we should observe larger short-term impacts on conflict over factors and on protests—whose returns are less affected by agricultural destruction—than on conflict over output. Figure 1.5 shows results from event studies of the impact of swarm exposure on alternative measures of conflict. I follow McGuirk and Burke (2020) in using reports of violence against civilians, riots, and looting from ACLED as representing conflict over output (Panel A) and violent conflict events reported in the UCDP database, which must include at least one "organized actor" and result in at least 25 battle-related deaths in a calendar year, as more likely to represent conflict over factors of production (Panel B). I compare these to reports of protests from ACLED (Panel C). Effects in the year of exposure are all close to 0 and not statistically significant and I cannot reject that impacts in the year of swarm exposure are the same across all types of conflict.[29]

Null effects on factor conflict and protests in the year of swarm exposure indicate mechanisms other than reductions in the returns to predation explain the null effect on violent conflict. There are several possible factors that could suppress violent conflict in the short term but not the long term.

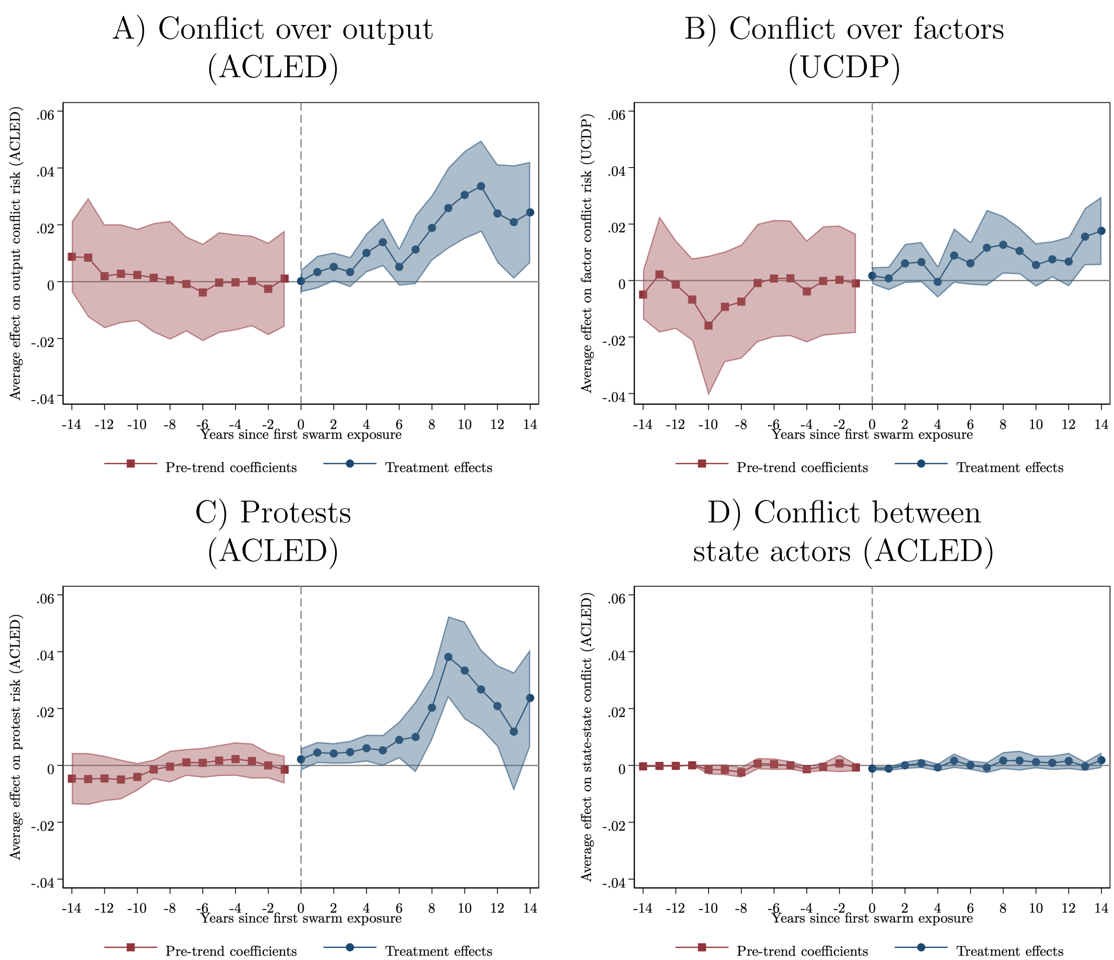
Psychological mechanisms may contribute the null short-term impact of locust swarms on violent conflict. The dominant religion in the sample countries is Islam, where locusts are mentioned as both a punishment from Allah and as a sign of Judgment Day. If locust swarms temporarily increase religiosity this may affect the perceived returns to fighting by increasing social, emotional, and supernatural costs, suppressing immediate violent conflict.

Temporary migration following an agricultural shock may be preferred to fighting. This could decrease the short-term risk of conflict as there are fewer people to potentially engage in fighting, given that cell population is positively and significantly correlated with the risk of violent conflict (Figure 1.4). Out-migration may be greater for more severe agricultural shocks such as the arrival of a locust swarm. Leaving to search of work is a common response to locust crop destruction (Thomson and Miers, 2002) and over 8 million people were displaced across East Africa as a result of the 2019-2021 locust outbreak (The World Bank, 2020).[30] Impacts of swarm exposure on violent conflict may then be realized with a

[29]For event studies of the 2003-2005 upsurge, the estimated effect of exposure in 2005 is -0.010 for violent conflict risk compared to +0.011 for protests, the latter of which is statistically significant in line with the model predictions. I can reject that the effects are the same ($p$=0.007). But I cannot reject that effects in 2004 are the same.

[30]Out-migration may increase the likelihood of conflict in nearby areas if it increases competition over local output and resources (as in McGuirk and Nunn (2021)), to the extent locusts are driving out-migration the evidence indicates that this is not leading to short-term conflict spillovers as the estimated impact in the

Figure 1.5: Impacts of swarm exposure on conflict risk over time, by conflict type




Note: The dependent variables are dummies for any conflict event being observed in a cell in a year, with the conflict type specified in the panel title. Each panel replicates Figure 1.3 Panel A for a different conflict outcome. See the figure note for Figure 1.3 for more detail. Shading represents 95% confidence intervals using SEs clustered at the sub-national region level.

delay when displaced populations return and are faced with the effects of the severe income shock.

Another possibility is that temporary relief efforts increase the opportunity cost of fighting by providing another source of livelihood that is presumably not available to fighters and

year of exposure remains negative when analyzing impacts in 0.5 and 1° cells.

deters individuals from engaging in violent conflict. On the other hand, aid may also increase incidence of violent conflict if it serves as a target for predation or banditry (Nunn and Qian, 2014), so the net effect of any increase in aid on violent conflict risk following a shock is unclear. Country-level annual data on international development flows to agriculture (which includes food aid) from FAOSTAT (2023b) provide some indication of temporarily increased relief following exposure to desert locust swarms, though with a short delay (Table A.1.5). Agricultural development flows increase by 185 million USD (20%) the year after a country experiences any locust swarm, and by 313 million USD (32%) the year after swarms arrived during the major 2003-2005 locust upsurge which drew significant international attention and support. An increase in international aid with a 1 year delay does not explain null effects on violent conflict in the year of exposure, but more local aid flows may arrive more quickly.

Finally, the null effects on violent conflict in the year of exposure may be due to costs of fighting being too high. Locust swarms are highly destructive and cause severe food insecurity so affected individuals may lack the resources needed to fight. Affected individuals could also have had few opportunities to join fighting groups in years of locust exposure. The majority of locust exposure occurred in 2004 and 2005, when violent conflict incidence in the sample countries was relatively low. On average 4.8% of cells exposed to locust swarms during this upsurge were adjacent to another cells that experienced any violent conflict during the period of exposure, compared to 13.1% in the following years. I return to differences by activity of fighting groups in the following subsection.

In summary, the predation mechanism does not appear to explain null effects of locust swarms on violent conflict risk in the year of exposure. Psychological effects, out-migration, and aid relief may also suppress the short-term likelihood of fighting, and affected households may also have lacked the resources and opportunity to engage in violent conflict.

### 1.7.2 Long-term

Locust swarm exposure causes a significant and persistent increase in the long-term risk of violent conflict. In the simple model of occupational choice, the likelihood of conflict increases if either the opportunity cost of fighting falls or if the returns to fighting increase. Locust swarms do not have any effect that could lead to long-term increases agricultural production or wealth, which means the impact on violent conflict risk must be driven by decreases in opportunity costs of fighting.

Event study results for different types of conflict in Figure 1.5 are consistent with such a mechanism. If individuals are deciding to fight when their opportunity cost is low this is most likely to occur in groups such as identity/political militias, rebel groups, and terrorist organizations as opposed to through joining state armed forces. In line with this, impacts on

ACLED conflict not involving non-state actors (Panel D) are smaller and less consistently statistically significant. On the other hand, long-term impacts on protests—which should be affected by opportunity costs of participation but not potential resources available to capture—are large. Swarm exposure significantly increases the long-term risk of protests with a similar pattern of smaller impacts in the first 7 years and much larger increases afterward, though the estimates for year 13 is not statistically significant.

Comparing impacts on different conflict types suggests a role of the predation mechanism as well, in line with model predictions. The average treatment effect is a 1.5pp increase in the annual risk of protests, a 138% increase relative to the mean in unaffected cells. The impact on the ACLED-based measure of output conflict is a 1.5pp (92%) average increase, and the impact on the UCDP-based measure of factor conflict is a 0.8% (69%) average increase. A larger impact on protests than output conflict is consistent with predictions of the model based on the predation mechanism offsetting some of the effect of lower opportunity costs on the risk of violent conflict. A smaller impact on factor conflict does not align with predictions, as the value of land in particular should not be affected by locust exposure. But land values may have fallen for other reasons, and another possibility is that factor conflict is more costly to engage in than output conflict and therefore increases by less in response to a decrease in opportunity costs.

If long-term impacts on conflict risk are due to persistent decreases in agricultural productivity reducing the opportunity cost of fighting, this should be observable in effects on productivity in affected areas. Marending and Tripodi (2022) find that agricultural profits of households in parts of Ethiopia exposed to locust swarms in 2014 were 20-48% lower two harvest seasons after swarm arrival. This indicates that impacts on agricultural productivity are not limited to the year of swarm exposure.

Data from the DHS AReNA database (IFPRI 2020) provide additional evidence of long-term reductions in agricultural productivity following swarm exposure. The DHS AReNA database includes geolocated data on cereal yields at the level of household survey clusters for 40 surveys from 9 countries in the study sample conducted between 1992 and 2018, as well as monthly Normalized Difference Vegetation Index (NDVI) measures at the cluster locations over this time frame. NDVI is a commonly-used satellite-based measure of vegetation greenness which in crop land can be considered a rough proxy for agricultural production. I collapse these data to the level of annual 0.25° cells.

Table 1.1 Column 1 shows that in cells where DHS surveys have been conducted in the sample countries between 2000-2019, NDVI falls significantly in years where a locust swarm is observed in a cell. The magnitude of the effect is similar to the effect of a 1°C increase in annual maximum temperature. The effect is likely an underestimate of the level of local agricultural destruction as not all parts of a 0.25° (28×28km) cell will be affected by a locust swarm, and because locust swarms may arrive and damage crops after peak crop growth and NDVI, unlike a temperature increase which would prevent crops reaching peak growth over a broader area. The average effect of locust swarm exposure on maximum NDVI in subsequent years is a fairly precise 0, consistent with locust swarms not affecting agricultural productivity fundamentals (Column 2).

Table 1.1: Average impacts of locust swarm exposure on measures of productivity

| | (1) Max annual NDVI (10,000s) | (2) Max annual NDVI (10,000s) | (3) Cereal yield, kg/ha | (4) Cereal yield, kg/ha | (5) Gross cell product, 1990 USD PPP millions | (6) Gross cell product, 1990 USD PPP millions |
|---|---|---|---|---|---|---|
| Any swarm in cell | -0.011*** | | 12.1 | | 0.59 | |
| | (0.003) | | (56.9) | | (2.44) | |
| Any swarm in cell previous year | 0.001 | | -92.3 | | -6.13** | |
| | (0.004) | | (63.4) | | (2.86) | |
| Affected by any locust swarm | | 0.000 | | -108.4* | | -5.98* |
| | | (0.002) | | (60.0) | | (3.10) |
| Total annual rainfall (100 mm) | -0.000 | -0.000 | 13.1** | 17.2** | 5.55* | 5.76* |
| | (0.001) | (0.001) | (5.9) | (7.0) | (3.11) | (3.17) |
| Total annual rainfall previous year (100 mm) | 0.002*** | 0.002*** | 11.9 | 7.6 | 0.59 | 0.62 |
| | (0.000) | (0.000) | (8.9) | (8.9) | (1.06) | (1.12) |
| Max annual temperature (deg C) | -0.012*** | -0.012*** | -61.2 | -86.4* | -4.08* | -4.45 |
| | (0.002) | (0.002) | (47.3) | (48.3) | (2.45) | (2.78) |
| Max annual temperature previous year (deg C) | -0.001 | -0.001 | 41.2 | 57.6 | 4.01 | 4.00 |
| | (0.002) | (0.002) | (45.1) | (43.3) | (2.85) | (2.90) |
| Population (10,000s) | -0.000*** | -0.000*** | 0.0 | 0.1 | 8.17*** | 8.09*** |
| | (0.000) | (0.000) | (1.2) | (1.4) | (2.00) | (2.06) |
| Observations | 54412 | 51961 | 4181 | 3750 | 41108 | 38734 |
| Outcome mean, no swarms | 0.537 | 0.551 | 1942.5 | 1987.0 | 59.93 | 60.25 |
| Country-Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Cell FE | Yes | Yes | Yes | Yes | Yes | Yes |

Note: Outcome variables in columns 1-4 are from the DHS AReNA database (IFPRI 2020) at the level of the clusters where DHS surveys are conducted. Maximum NDVI is measured based on satellite data for the location of each survey cluster each month from 2000-2019; I take the maximum value each year. Cereal crop yield in kg/ha is measured based on survey reports and averaged at the cluster level for years in which surveys are conducted in specific clusters. I collapse these data to the level of 28×28km grid cells by year taking means of the outcome variables. Gross cell product in columns 5-6 is from Nordhaus (2006) as included in the PRIO-GRID database. All regressions include country-by-year and location fixed effects. SEs are clustered at the sub-national region level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

While this might suggest no persistent impacts on agricultural labor productivity, Column 4 shows that household reports of cereal yields fall by a statistically significant 108 kg/ha (5.4%) on average in years after locust swarm exposure. This is larger than the effect of a decrease of 100mm of rainfall in the same year or a 1°C increase in maximum temperature. This decrease emerges from the first year after locust swarm exposure (though it is not statistically significant), but is not observed in the year of exposure where there is no significant effect on yield (Column 3). No effect on yield in the year of exposure may be a result of two factors. First, in some countries and years survey timing will involve asking about cereal yields over a season that ended before swarms arrived. Second, yield is typically reported as output over area harvested, and consequently will not capture damages on plots where crops are completely destroyed so there is no harvest.

Columns 5 and 6 show how locust swarm destruction affects income using measures of gross cell product—estimates of total income at the cell level based on population and output in agricultural and non-agricultural activities—from Nordhaus (2006) as included in the PRIO-GRID database. These estimates are only available for 2000 and 2005 in the sample period, allowing an analysis of the immediate impact of exposure to the 2003-2005 locust upsurge. Total income falls by 6 million USD (in 1990 PPP terms)—nearly 10%—in cells affected by locust swarms, in the year after swarm is exposure, consistent with when yields are affected.

A significant immediate decrease in average incomes in areas exposed to locust swarms and a persistent decrease in cereal yields in the following years relative to unexposed areas are consistent with the wealth mechanism described in the model. Locust swarms do not affect local agricultural productivity fundamentals, but damages to agriculture from locust exposure result in a permanent income and wealth shock. Adoption of agricultural insurance is very low in the sample countries, access to credit is constrained, and local risk sharing networks may be insufficient to help affected households recover from the extent of damage caused by locust swarms. Household coping strategies involve drawing down wealth to deal with the income and food security shock. The severity of this shock has been demonstrated by studies showing long-term negative impacts of locust swarm exposure on children's education (De Vreyer et al., 2015) and health (Conte et al., 2023; Le and Nguyen, 2022; Linnros, 2017). Reduced household wealth would decrease access to productive inputs in subsequent years, which would explain a persistent decrease in cereal yields. This implies lower agricultural productivity and therefore a lower opportunity cost of fighting in areas exposed to locust swarms, which would increase the probability of violent conflict.

The wealth mechanism alone would suggest that impacts of swarm exposure on conflict risk should fall over time if households gradually recover or should be fairly stable if households reach a new productivity equilibrium. The fact that long-term impacts of swarm exposure on conflict are not consistent over time and are largest in a period starting around 8 years after exposure suggests some heterogeneity in impacts by time-varying conditions affecting either opportunity costs or returns to fighting.

The largest effects of upsurge swarm exposure coincide with the years when the general risk of conflict increased across the sample countries (Figure A.1.1), indicating heterogeneity by local conflict activity. Proximate causes for the increase in conflict include food price shocks, the Arab Spring, the spread of Islamic militant groups, and multiple civil wars. Recruitment of fighters to armed groups following the onset of conflict will be easier in areas where opportunity costs of fighting are lower (Collier and Hoeffler, 2004). But individuals in those areas may otherwise not find switching to fighting optimal as violent conflict is generally a collective activity. Without a group, the social, emotional, and monetary costs of fighting are likely to be high and the probability of victory low, as discussed in Section 1.3.

Table 1.2 shows the estimates from testing whether long-term impacts swarm exposure vary by the presence of groups active in violent conflict the previous year or in the surrounding cells. Results are similar when considering any swarm exposure and when focusing on the 2003-2005 upsurge. Impacts of swarm exposure on violent conflict risk are driven entirely by effects in areas and periods with groups engaging in violent conflict, as predicted. This is consistent with Figure 1.3 showing significant impacts of upsurge exposure are delayed until years when violent conflict incidence was greatest in the sample countries, and similarly explains the delay in the largest effects on protests in Figure 1.5.

Table 1.2: Average impacts of exposure to locust swarms on violent conflict risk, by presence of groups active in civil conflict

| Outcome: Any violent conflict event | (1) Any exposure | (2) 2003-2005 exposure | (3) Any exposure | (4) 2003-2005 exposure |
|---|---|---|---|---|
| Exposed to locust swarm | 0.002 | 0.003 | 0.001 | 0.002 |
| | (0.003) | (0.004) | (0.003) | (0.004) |
| Any violent conflict in cell previous year | 0.299 | 0.275 | | |
| | (0.218) | (0.274) | | |
| Exposed to locust swarm × Any violent conflict in cell previous year | 0.081*** | 0.160*** | | |
| | (0.028) | (0.031) | | |
| Any violent conflict elsewhere in 1 degree cell | | | 0.132 | 0.210** |
| | | | (0.092) | (0.105) |
| Exposed to locust swarm × Any violent conflict elsewhere in 1 degree cell | | | 0.049*** | 0.056*** |
| | | | (0.014) | (0.017) |
| Observations | 452218 | 400668 | 452218 | 400668 |
| Outcome mean post-2005, no swarm exposure | 0.018 | 0.019 | 0.018 | 0.019 |
| p-value: coefficients shown all =0 | 0.031 | <0.001 | 0.007 | 0.016 |
| Country-Year FE | Yes | Yes | Yes | Yes |
| Cell FE | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes |
| IPW | Yes | Yes | Yes | Yes |

Note: The dependent variable is a dummy for any violent conflict event observed. Cells are regarded as 'affected' by locust swarms for all years starting from the first year in which a swarm is recorded in the cell. Analyses of the 2003-2005 upsurge exclude cells exposed in other years from the sample. Inverse propensity weights are applied in all regressions and are calculated separately based on estimates of the propensities of cells to have been exposed to any swarm and to the 2003-2005 upsurge. Controls include annual total rainfall, maximum temperature, and population, and their interactions with the conflict activity variables. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Swarm exposure magnifies the correlations in violent conflict activity over time and space.

Cells exposed to any swarm (the 2003-2005 upsurge) are 8.1 percentage points (16.0pp) more likely to have any violent conflict if there was violent conflict in the cell the previous year, relative to unexposed cells, and 4.9pp (5.6pp) more likely if there is violent conflict in any of the cells in the surrounding 1° cell.[31] Being exposed to a swarm thus makes cells significantly more likely to engage in violent conflict when it emerges, even years after exposure, but not otherwise.

In summary, long-term impacts of locust swarm exposure on conflict risk can be explained by persistent decreases in opportunity costs of fighting. These decreases are observed in a cereal yields, a measure of agricultural productivity, and are consistent with an initial wealth shock affecting long-term access to productive resources. The decrease in opportunity cost does not cause outbreaks of new violent conflicts but rather affects the locations and duration of violent conflicts after these emerge due to other proximate causes, when net returns to fighting are high enough to make a switch to fighting optimal.

## 1.8 Comparing locust swarms and severe drought

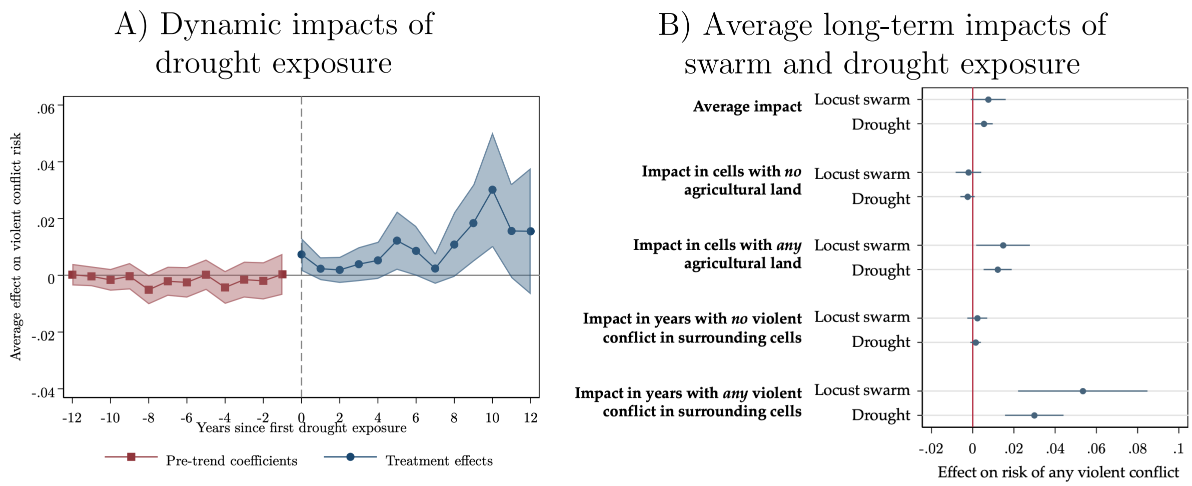
Locust swarms are a unique and catastrophic agricultural shock, but the model described in Section 1.3 is general and predicts similar patterns of impacts on the risk of violent conflict for other severe shocks to agricultural production. In this section I compare the impacts of locust swarm exposure to the impacts of exposure to a severe drought. Following Harari and La Ferrara (2018) and others I use the Standardized Precipitation and Evapotranspiration Index (SPEI) which combines both rainfall and the ability of the soil to retain water. Grid cell SPEI data from 1996-2014 are included in the PRIO-GRID database (Begueria et al., 2014). The units for the measure of SPEI are standard deviations from the cell's historical average. I define a cell as experiencing a severe drought shock in a particular year if there are at least 4 consecutive months where the SPEI is below -1.5 (deviations within 1 indicate near normal conditions). As with locust swarm exposure I identify the first year in which a cell experiences such a drought and consider cells to be 'affected' in all subsequent years.

Figure 1.6 Panel A shows the results from an event study of drought exposure using the BJS method. Pre-exposure coefficients are small in magnitude and not statistically significant, with the exception of the period 8 years before exposure to a drought. Conflict risk increases by a statistically significant 0.9 percentage points in the year of exposure.

[31] Because swarm exposure affects long-term conflict risk it is not orthogonal to conflict in the previous year nor likely to surrounding conflict in the same year, though I reject that the interacted terms are jointly equal to 0 with a high level of confidence. These results should therefore be considered as illustrative of differences in upsurge impacts by the broader conflict environment rather than accurately estimating differences in the causal impacts.

Treatment effect estimates are positive and statistically significant for years 5-6 and 8-11 post-exposure, with the largest effects in years 9-11. The average effect over the 12 years post-exposure is a 1.0pp increase in the annual risk of violent conflict.

Figure 1.6: Long-term impacts of drought and locust swarm exposure on violent conflict risk




Note: The dependent variable is a dummy for any violent conflict event observed. Panel A shows the event study results for impacts of exposure to any drought (4 consecutive months with SPEI<-1.5) on violent conflict risk over time using the Borusyak et al. (2021) approach. Estimated impacts in each time period are weighted averages across effects for drought exposure in particular years. The time period 0 is the year of drought exposure. The shaded areas represent 95% confidence intervals. Treatment effects are based on comparisons against the pre-exposure period.
Panel B shows coefficients and 95% confidence intervals are from three separate TWFE regressions for the average impacts of swarm and drought exposure: one with no interactions and the others interacting all right-hand side variables with a dummy for any agricultural (pasture or crop) land and a dummy for any violent conflict in the 15 other cells in the broader 1 degree cell in which the cell is located. Coefficients and standard errors are calculated using Stata's *xlincom* command based on the sums of the coefficients for the baseline effect and the interaction term. Inverse propensity weights are included in the regressions of swarm exposure.
In both panels, cells are regarded as 'affected' by locust swarms or drought for all years starting from the first year in which the shock is recorded in the cell. All regressions include country-by-year and cell fixed effects. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses. Regression results are shown in Table A.1.6.

Panel B shows TWFE estimates of the average long-term impacts of swarm and drought exposure by cell land cover and activity of armed groups in the surrounding cells. Impacts of locust swarms are the same as those reported previously including inverse propensity weights. The TWFE estimate of the average impact of drought is a significant 0.5pp increase in the annual risk of violent conflict. slightly smaller than the average of the event study treatment period effects.

The patterns in the impacts of swarm and drought exposure are very similar with slightly larger effects of swarm exposure, consistent with both of these transitory agricultural shocks affecting the risk of violent conflict through the same mechanisms. Impacts of both shocks are concentrated in cells with any agricultural land, and in years where there is any violent conflict in surrounding cells. The impact of past swarm exposure in years where there is conflict activity near a cell is particularly large for locust swarms—a 5.3pp increase in the risk of violent conflict in the cell compared to 2.9pp for drought—likely because the sample for the analysis of swarm exposure includes more high-conflict years from 2015-2018 that are not included in the drought analysis sample.

The results show that exposure to two different severe but transitory agriculture shocks cause large and persistent long-term increases in violent conflict risk. The patterns indicate that the mechanism for both shocks is a persistent decrease in the opportunity cost of fighting making affected areas more likely to engage in future conflict when they emerge.

This finding has implications for previous research on the short-term impacts of economic shocks on conflict which also use grid cell panel data.[32] This literature has overwhelmingly focused on the short-term and assumes effects of shocks are temporary. A common empirical approach is a distributed lag two-way fixed effects model which takes the form:

$$Conflict_{ict} = \alpha + \beta_1 Shock_{ic,t} + \beta_2 Shock_{ic,t-1} + \delta X_{ict} + \gamma_{ct} + \mu_i + \epsilon_{ict} \tag{1.2}$$

This follows the persistent effects model in Equation 1.1 with the exception that instead of the *Shock* variable representing a permanent treatment status over subsequent years, in this temporary effects model the years following a shock are regarded as unaffected except as captured by the one year lag. This lag allows for limited delays or persistence in impacts of the shock (Burke et al., 2015).

With cell fixed effects the short-term impacts in the temporary effects model are estimated relative to conflict risk in other years in the same cell where a shock is not observed, including years after exposure to a shock. For shocks that cause persistent increases in conflict, this implies that the temporary effects estimate will be biased downward as a result of comparing conflict risk in the year a shock is observed against later years with no shock but higher conflict risk *caused* by the shock. I show that this is the case for locust swarms and severe drought, comparing estimates from the temporary effects approach to the event study estimates treating shock effects as permanent Table A.1.7.

In the case of the locust swarms, the temporary effects model estimates a highly significant 1.5pp *decrease* in the probability of any violent conflict in the year of exposure relative to unaffected cells. I can reject that this is the same as the event study estimate of -0.2pp

[32]See for example Fjelde (2015); Harari and La Ferrara (2018); McGuirk and Burke (2020); McGuirk and Nunn (2021); Ubilava et al. (2022). These studies analyze the impact of various shocks on conflict in Africa, and vary in their choice of controls and in the size of grid cells they analyze but all use a similar econometric specification.

with high confidence ($p$=0.009), consistent with downward bias of the temporary effects estimate. In the case of severe drought, the temporary effects estimate is a non-significant 0.4pp increase in violent conflict risk the year of a drought, compared to a highly significant 0.9pp increase in the event study estimate. While I cannot reject that the estimates for impact of drought on violent conflict risk in the year of exposure are the same ($p$=0.198), the two approaches would yield very different conclusions with different policy implications.

The temporary effects estimate for locust swarms—a 1.5pp decrease in violent conflict the year locusts are observed—is very close to Torngren Wartin (2018)'s estimate of a 1.3pp decrease using a similar method.[33] He interprets the result as suggesting endogenous under-reporting of locust swarm presence correlated with violent conflict, and does not consider potential bias from ignoring long-term impacts of swarm exposure on conflict risk. The much larger event study estimate for the impact of swarms on conflict in the same year suggests the large share negative estimate in the temporary effects regression can instead be attributed to downward bias from ignoring long-term impacts.

These results provide some evidence of a potential misspecification of studies analyzing short-term impacts on conflict of transitory economic shocks that are severe enough to decrease permanent income and reduce wealth and future productivity. Studies of such shocks using specifications similar to Equation 1.2 and ignoring possible long-term effects will generate downward-biased short-term impact estimates to the extent the shocks increase long-term conflict risk.

## 1.9 Conclusion

Violent conflict and environmental shocks can have devastating consequences for economic and human development which are the subject of significant study even beyond the economics literature. This paper shows that exposure to a severe agricultural shock—both desert locust swarms and drought—significantly increases long-term conflict risk. The mechanism is a persistent reduction in the opportunity cost of fighting, as shown in a long-term decrease in cereal yields following locust swarm exposure. This effect implies that short-term efforts to provide food aid and other support to populations affected by severe agricultural shocks are insufficient to facilitate full recovery, in line with many studies showing natural disasters have long-term impacts on poverty and well-being.

[33] Torngren Wartin (2018) uses a similar sample and the analysis is at the level of 0.5° and 0.1° cells with somewhat different controls for lagged locust presence and weather but employs the same general distributed lag specification with cell and country-by-year fixed effects. He considers locust swarms and bands together while I focus on more destructive swarms alone, and includes some African countries with very few locust swarm observations over time while excluding Arabian countries with extensive locust activity.

Locust swarms have no significant effect on violent conflict in the year of exposure. I show evidence that relief efforts may also play a role, as international agricultural aid increases in the year after countries are affected by locust swarms. Geographically-disaggregated data on relief could be useful to test whether such support deters short-term conflict and limits the extent of the wealth shock and long-term increase in conflict risk, though prior studies suggest aid may increase conflict risk (Crost et al., 2014; Nunn and Qian, 2014). More generally, additional research could explore whether policies that can promote resilience to agricultural shocks, such as cash transfers (Crost et al., 2016; de Janvry et al., 2006; Garg et al., 2020), livelihood graduation programs (Hirvonen et al., 2023), improved infrastructure (Gatti et al., 2021), and work programs (Fetzer, 2020) also reduce conflict risk.

The lack of immediate impacts of swarm exposure on violent conflict and the long-term decrease in the opportunity costs of fighting mean locust swarms do not increase conflict onset but rather make affected areas more likely to become involved in periods of elevated civil conflict. This contrasts with prior studies finding price (McGuirk and Burke, 2020) and drought (Harari and La Ferrara, 2018) shocks cause the onset of new conflict, but is consistent with Bazzi and Blattman (2014), who find that commodity price shocks primarily affect conflict incidence and not conflict onset. Many factors have contributed to the increased onset and incidence of violent conflict in the sample countries in this period, including the Arab Spring, several civil wars, insurgencies, and the spread of terrorist organizations. The results illustrate how failing to support communities affected by disasters to fully recover can increase the incidence of future conflict and influence which areas become involved.

The long-run results are consistent with a wealth mechanism where the initial shock to agricultural production income and coping strategies lead to persistently lower wealth and productivity. Further research on impacts of economic shocks on household measures of productivity (as in Marending and Tripodi (2022)), labor supply, food security, and wealth would help further explore how temporary shocks to productivity can have persistent effects on the opportunity cost of fighting by reducing household productive assets (including human capital).

The findings suggest future avenues of research in the literature on climate and conflict. I show that the methods typically used in this literature, which treat shocks as only affecting conflict risk in the short-term, can result in downward-biased estimates of short-term effects when the shocks have long-term impacts. Event study analyses of severe weather shocks such as droughts could show the extent and patterns of long-term impacts on conflict risk and how this affects estimates of short-term effects. Although not a focus of this paper, the lack of variation in the impacts of rainfall and temperature deviations on violent conflict risk by land cover cast further doubt on whether effects on agricultural production are the primary mechanism. The association between climate and conflict has been demonstrated in a wide variety of settings but the mechanisms remain unclear (Mach et al., 2020). A better understanding of the different mechanisms is essential to determining the appropriate policy responses.

The results also have implications for estimates of the economic and social costs of desert locust outbreaks. In addition to short-term impacts on agricultural production, incomes, and food security, previous studies have found that children exposed to locust swarms have worse education and health outcomes (Conte et al., 2023; De Vreyer et al., 2015; Le and Nguyen, 2022; Linnros, 2017). This study shows that swarm exposure also increases the long-term risk of conflict, which imposes further costs.

Past research on desert locusts has argued that limited impacts of outbreaks on aggregate national measures of agricultural production may mean expensive locust monitoring and control operations have limited net economic benefits (Joffe, 2001; Krall and Herok, 1997), though others have argued that local damages are extensive and motivate continued proactive locust control efforts (Showler, 2019; Zhang et al., 2019). Decisions about proactive locust control versus potential alternative responses—for example, increasing adoption of agricultural insurance or relying entirely on disbursement of relief after locust damages (Hardeweg, 2001)—should take the broader economic and social impacts of agricultural destruction by locusts into consideration. These costs should also motivate increased cross-country communication and collaboration in response to threats of locust swarms, particularly in light of the recent 2019-2021 outbreak..

Beyond contributing to our understanding of the relationship between agricultural productivity shocks and conflict risk, the findings are also relevant for considering multilateral policy around climate change mitigation and adaptation. Climate change is increasing the frequency and severity of agricultural shocks, including by creating conditions suitable for desert locust swarm formation. These shocks impose additional costs on society through their impacts on conflict risk which should be considered when weighing the costs and benefits of potential actions to reduce and respond to risks from agricultural shocks.

# Chapter 2

# Flooding and livelihood diversification in Nigeria: the view from the sky and the view from the ground

## 2.1 Introduction

Climate change is predicted to reduce agricultural productivity in many parts of the world due to increases in temperatures, increased variability of rainfall, and associated increases in the severity and frequency of extreme weather events. A possible response in agricultural areas is reallocation of labor to non-farm activities. Environmental shocks such as droughts and floods could facilitate or hinder the process of structural change in low-income agricultural settings. For example, weather-driven reductions in agricultural productivity could accelerate exit of labor from agriculture conditional on the ability of non-agricultural sectors to absorb the displaced workers (Barrett et al., 2021; Colmer, 2021). But income losses due to adverse weather could lead to low demand for non-agricultural goods and services, limiting opportunities for labor reallocation (Liu et al., 2023). Local demand for agricultural labor following poor weather might also temporarily increase in attempts to offset production risk (Mueller et al., 2020a). Migration is a particularly important potential response to climate change and weather shocks (Auffhammer and Kahn, 2018; Cruz and Rossi-Hansberg, 2021), as households update their beliefs about future agricultural productivity and risk and some decide to give up agriculture. But households may face challenges migrating and finding non-farm employment. In particular, reductions in wealth following catastrophic weather shocks could prevent household exit from agriculture by reducing the capital needed to diversify their livelihoods by investing in skills, businesses, or migrating (Carter et al., 2007).

The impact of severe weather-related shocks for labor reallocation is therefore an empirical question with important implications as such shocks become more frequent as a result of climate change.

This paper analyzes the medium-term labor supply impacts of exposure to flooding shocks in low- and middle-income countries (LMICs) by analyzing the impacts of a major flooding event in Nigeria on measures of subsequent livelihood diversification. Flooding can completely destroy agricultural production in affected areas but may also cause important damages to infrastructure, housing, and other assets. Consequently it is unclear whether exposure to flooding would result in a persistent shift toward non-agricultural activities for farm households. In 2012, excessive rainfall between September-November caused widespread flooding across much of Nigeria. The flood caused an estimated \$16.9 billion in damages and killed 431 people (Unah, 2021). Using four rounds of the nationally-representative General Household Survey-Panel (GHSP), I test whether community-level exposure to the 2012 floods leads households in agricultural areas to persistently diversify their sources of income between subsistence farming, commercial farming, non-farm household enterprise, and wage employment.

The empirical approach compares households in communities that experienced flooding in 2012 to households in communities with similar predicted flood risk in a difference-in-differences design using the 2010-2011 survey round as the pre-exposure period. I consider all households in a community where flooding was identified in 2012 as 'treated' by flood exposure in the 2012-13 (almost entirely post-flooding), 2015-16, and 2018-19 survey rounds. Because flooded and non-flooded communities have different geographic characteristics and flood risk, I estimate 2012 flood exposure propensity scores using baseline and time-invariant community characteristics, including prior flooding history. I use these to construct inverse propensity weights and apply these in all of my analyses to compare communities with similar flooding risk and increase the plausibility of the parallel trends assumption.

An initial challenge is how to identify flooding exposure (see also Guiteras et al. (2015); Patel (2023)). I consider two different measures of GHSP community flood exposure, one based on survey reports and one based on MODIS satellite imagery, and explore whether the choice of how to define flooding affects conclusions about the effects of exposure on households. The survey-based definition of flooding measures whether a flood was reported by any resident household in 2012. The MODIS-based definition using MODIS satellite imagery to identify areas that were flooded at any point in 2012 and defining communities with a centroid within 5km of such areas as exposed to flooding. These definitions capture different subsets of communities: 42 are considered exposed to flooding in 2012 according to both definitions compared to 86 using the survey-based definition only and 59 using the MODIS-based definition only, with 278 not considered flooded by either.

Part of the difference can be attributed to the offsetting of the community coordinates in the public use GHSP data, but measurement issues with both definitions also play a role and result in the communities identified as flooded under different different having different characteristics. Communities where flooding is identified by MODIS only are less rural and agricultural and have higher average annual rainfall, and communities where flooding is identified by the survey only are more agricultural and have less average annual rainfall but had higher than usual wettest quarter rainfall in 2012. These suggest that flooding may not cause damage and therefore go unreported 'on the ground' in communities with certain characteristics, and that measures 'from the sky' fail to capture certain types of flooding events—in particular those caused by heavy rainfall where cloud cover prevents satellite observation.

Despite these differences some of the estimated effects on measures of household engagement in and income from livelihood activities are similar across the two flooding definitions. Both definitions show no significant effects on average total household income for households in flooded communities relative to non-flooded communities in the years after 2012. The probability of being under the poverty line and the value of household assets also do not change. These suggest limited long-term effects of exposure to the 2012 floods, though the probability of experiencing any food insecurity in the past 12 months increases by 5 to 7 percentage points indicating exposed households are worse off in at least some respects.

I find no effects of flood exposure on engagement in wage employment or wage income. This implies that opportunities for wage employment—which some other studies have found increases for agricultural households immediately following exposure to a flood as they strive to smooth their consumption—are limited in flood-affected areas in Nigeria, and that any potential immediate increases in wage work do not persist in this setting.

Both flooding definitions show evidence of decreases in net crop production income driven by significant 21-23% decreases in the value of crop production and particularly of non-staple more commercially-oriented crops. The results suggest flood exposure pushes households toward subsistence crop production for own consumption and reduces commercialization. Resource constraints following a flood shock may prevent some households from both producing crops for sale and meeting household food consumption needs. The value of crop production and of commercial crops falls by significantly more for households below the poverty line in 2010-11 who likely face the strictest resource constraints, indicates a shift toward subsistence crops for the most vulnerable households. Using the survey-based floods definition only I find that some households not cultivating crops in 2010-11 began doing so after the floods, largely growing staple crops with very little net income contribution. This drives an increase in the probability that a household engages in crop production following flood exposure which is significant using this definition only.

The two flooding measures also both identify some increases in household non-enterprise activity and income following flood exposure, though with differences in underlying mechanisms. The MODIS-based flood definition shows a significant average increase in household non-farm enterprise income concentrated among households with active businesses before the 2012 floods and driven particularly by less wealthy households. Increased enterprise income appears to come at the expense of crop production value, suggesting households are resource-constrained and divert resources toward non-farm enterprises. There is limited evidence of the creation of new non-farm enterprises among households not previously engaged in such activities. Only less wealthy households change their income shares from different activities following flood exposure. If changes in livelihood decisions are a response to a flood shock, this suggests that it is less wealthy households perhaps more vulnerable to flood exposure that are most likely to respond by shifting their income activities whereas wealthier households may have been more resilient to flooding and therefore less affected.

The survey-based flood definition identifies a relative increase in household enterprise income only among households engaged in crop production before the 2012 floods. There is no significant difference by whether households were previously engaged in non-farm enterprise meaning some farm households exposed to flooding started new non-farm enterprises, despite the null average effect on the probability of households engaging in such activity. This indicates an effort to diversify household income sources, but I find no significant changes in the shares of total income from enterprise or from crop production according to the survey-based flooding definition. New non-farm enterprises do not appear to have become a significant net contributor to household incomes.

The two flooding definitions therefore lead to different conclusions about the effects of flood exposure on household livelihoods in agricultural areas. The MODIS-based definition indicates that some households, particularly those previously engaged in non-farm household enterprises, reallocate resources from crop production to enterprise and grow these businesses. The survey-based definition indicates that farm households reduce production of commercial crops in favor of staples, reducing the total value of crop production, and that there is some diversification of households into household farm or non-farm activities they were not previously engaged in but resulting in limited net incomes.

Increases in non-farm enterprise income are largest for areas where flooding was identified only by MODIS meaning these effects may be driven by community characteristics that allow flooding to occur without causing enough damage to be reported in surveys, such as adaptation following prior floods. To the extent this is true the results suggest that transitioning away from agriculture appears to be an important effect of flooding in Nigeria. Where flood exposure is more of an adverse shock leading to survey reports the results indicate that households undertake various coping strategies to minimize decreases in incomes and well-being but potentially increasing vulnerability to future shocks and without meaningfully reducing reliance on household agriculture. Studies of the impacts of flooding should carefully consider what types of flood exposure their measures allow them to capture and how this might affect the effects they observe.

This paper contributes to our understanding of the impacts of weather-related disasters, particularly floods, on labor supply and livelihoods in agricultural communities in sub-Saharan Africa, adding to a rich literature on the impact of climate shocks in LMICs. Several studies have analyzed the effects of flooding on migration, labor supply, and other outcomes in LMICs, though studies of its impacts on labor supply have focused primarily on South Asia (Chen et al., 2017; Mueller et al., 2014; Gray and Mueller, 2012b; Mueller and Quisumbing, 2011) and few studies of natural disasters in agricultural settings have analyzed effects on measures of livelihood or labor supply other than migration beyond the short term (Albert et al., 2021; Efobi, 2022; Kirchberger, 2017)). I find that short-term increases in wage employment following floods documented in other settings do not occur in agricultural areas in Nigeria. Livelihood responses involve reducing commercial crop cultivation in favor of staple crops and increased emphasis on household non-farm enterprises.

I add to a growing literature using remote sensing to measure economic shocks, including several papers estimating flood exposure (Brakenridge, 2023; Guha-Sapir et al., 2023; Guiteras et al., 2015; Patel, 2023; Tellman et al., 2021), and particularly considering tradeoffs across different measures. I find large differences in the communities identified as affected by the 2012 floods in Nigeria according to survey reports ‘on the ground’ and to MODIS satellite imagery 'in the sky.' The differences in characteristics of identified communities indicate that each source suffers from specific measurement issues and thus presents a limited picture of flooding. I show how estimated effects of flooding differ according to the two definitions, in line with each measure capturing a different subset of flood-related events. These differences imply that studies of the impact of flooding should consider what types of exposure their flood measure allows them to observe and whether this is the exposure of economic interest.

Finally, I contribute to the literature on structural change and on factors encouraging or impeding this process in LMICs. Global reductions in agricultural productivity growth over the past few decades due to climate change (Hultgren et al., 2022; Ortiz-Bobea et al., 2021) have important implications for structural transformation (Liu et al., 2023) and may be complicated by increasingly frequent and severe extreme weather events. While a large empirical literature studies the consequences of climate change (Pörtner et al., 2022) and many studies have analyzed short-term labor supply responses to extreme weather events,[1] there is limited evidence on the relationship between extreme weather events and livelihood

[1] See for example Afridi et al. (2022); Branco and Féres (2021); Chuang (2019); Colmer (2021); Dillon et al. (2011); Emerick (2018); Franklin and Labonne (2019); Grabrucker and Grimm (2021); Gröger and Zylberberg (2016); Ito and Kurosaki (2009); Jayachandran (2006); Kijima et al. (2006); Kleemans and Magruder (2018); Kochar (1995, 1999); Kubik and Maurel (2016); Matsuura et al. (2023); Maystadt et al. (2016); Mueller et al. (2020b); Noack et al. (2019); Rose (2001).

diversification or structural change over the longer term in sub-Saharan Africa (Barrett et al., 2021). I show that flooding exposure may be both a constraint and a 'push' factor in the process of households transitioning away from agriculture, depending on the flooding definition used. I find evidence that floods lead households to shift from commercial to staple crop cultivation and that entry into non-farm enterprise is usually a low-productivity endeavor. To the extent households reallocate labor out of crop production it is to non-farm household enterprises, but show that the narrative of transitioning away from agriculture following flood exposure as captured by satellites may be biased.

## 2.2 Background

Floods are one of the most common types of natural disasters experienced globally, accounting for 44% of all disaster events from 1970-2019 (World Meteorological Association et al., 2021). The Intergovernmental Panel on Climate Change Sixth Assessment Report finds that there is high confidence that climate change is increasing the risk and severity of flooding, due largely to increases in extreme precipitation events (Caretta et al., 2022). Tellman et al. (2021) use high-frequency and high-resolution satellite data to identify 913 large flood events from 2000-2018 affecting 2.23 million km$^2$ and over 255 million people. They estimate that the proportion of the global population exposed to floods has increased by 20-24 percent from 2000 to 2015, and find that this will increase further by 2030 under current climate change projections. Projected increases in flood risk are not distributed evenly over space; the largest projected increases are concentrated in South and Southeast Asia and Sub-Saharan Africa (Caretta et al., 2022).

Nigeria is one of the countries with the largest change in the proportion of the population exposed to floods, with this proportion nearly doubling from 2000 to 2015 and expected to continue increasing (Tellman et al., 2021). The Nigeria Hydrological Services Agency (NIHSA) reports that the frequency of flooding has intensified in recent years, disrupting agricultural production and livelihoods but also causing death, population displacement, and destruction of housing and other infrastructure (NIHSA, 2021). An analysis of NIHSA annual reports of flooding events at the Local Government Area (LGAs) level shows that fewer than 1 in 10 LGAs across the country experiencing flooding in a given year from 2013-2017 compared to more than 1 in 4 from 2018-2022 (Figure B.1.1. Flooding is most common in coastal areas and along major rivers though heavy rainfall also causes flooding in areas farther from bodies of water (Figure B.1.2). Widespread floods in 2022 are considered the worst in Nigeria's history, affecting 33 of 36 states and the Federal Capital Territory. These floods displaced over 1.3 million people and killed over 603 (Oguntola, 2022) while also destroying nearly 1,000,000 hectares of farmland and over 350,000 homes and causing an estimated $176 million of damages to transportation infrastructure (NIHSA, 2023). The previous worst flooding occurred in 2012 as a result of excessive rainfall between September and November across much of the country, affecting 30 of 36 states. The flood caused an estimated $16.9 billion in damages, displaced over 2.1 million people, and killed 431 people (Unah, 2021).

Increasing risk of flooding and the recent occurrence of particularly severe national flood disasters make Nigeria a relevant context for studying the economic impacts of flood exposure and labor supply. Nigeria is also an important setting to study labor reallocation and structural transformation, with the largest population and economy in Africa. While 76% percent of Nigeria's GDP comes from non-agricultural sectors, around 70% of Nigeria's population is engaged in agriculture (Statista, 2023). Most of agricultural production is at a subsistence level and productivity is low such that Nigeria relies on imports to feed its population (FAO, 2022).

Several studies have analyzed impacts of weather shocks on labor supply in African countries. In Nigeria, Dillon et al. (2011) find that men from agricultural households practice short-term migration in response to temperature shocks, while Efobi (2022) reports that women exposed to drought early in life have worse labor market outcomes as adults, pointing to reduced education and earlier marriage as important channels. Takeshima et al. (2018) find that higher agricultural productivity in Nigeria increases agriculture-oriented non-farm activities. More generally, many studies find that agricultural households respond to extreme weather events by increasing labor supply to non-agricultural activities in the same year,[2] in order to offset lost agricultural income and smooth household consumption. But there is less evidence on whether exposure to such shocks leads to persistent changes in labor supply and income diversification of agricultural households.[3]

[2]See Afridi et al. (2022); Branco and Féres (2021); Chuang (2019); Colmer (2021); Dillon et al. (2011); Jayachandran (2006); Kijima et al. (2006); Kochar (1995); Kubik and Maurel (2016); Matsuura et al. (2023); Mueller et al. (2020b); Noack et al. (2019); Rose (2001). Emerick (2018); Grabrucker and Grimm (2021) also document increases in non-agricultural labor following *positive* rainfall shocks.

[3]On migration: Gray and Mueller (2012a) on drought in Ethiopia, Gray and Mueller (2012b) on floods in Bangladesh, Mueller and Osgood (2009a,b) on droughts in Brazil, and Mueller et al. (2014) on flooding and heat stress in Pakistan. On labor supply: Albert et al. (2021) on drought in Brazil, Efobi (2022) on drought in Nigeria, Kirchberger (2017) on earthquakes in Indonesia, Mueller and Quisumbing (2011) on flooding in Bangladesh, and Liu et al. (2023) on temperature in India.

## 2.3 Data

### 2.3.1 Survey data

The main data source for the analysis is Nigeria's General Household Survey Panel (GHSP). The GHSP is a nationally-representative survey including approximately 5,000 households conducted by the Nigeria National Bureau of Statistics, and is part of the Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA). In each round, households are surveyed twice: once post-planting and once post-harvest.[4] Some information, such as labor supply, is recorded in each survey while other modules are only included in either the post-planting or post-harvest surveys. Households are tracked over time, including if they move to a new location, but individual household members are not tracked. In 2018-19 the panel sample was partially refreshed, with 1425 households from the original panel retained and 3551 new panel households added to the sample.

The household survey includes detailed modules on household agricultural production, labor supply, income, assets, and socio-demographics. The survey also includes data on individual-level labor supply; I aggregate these variables up the household level for analysis. The primary household-level outcomes are measures of total annual income by activity and participation in different types of work. I also clean data on agricultural production, crop and livestock product sales, household and agricultural assets, household work hours by activity, migration, food security, and household sociodemographic characteristics. For certain variables—notably the measures of total household income by activity—I modify code developed by EPAR (2023)[5] to construct variables in consistent ways accounting for differences in the survey instruments across rounds.

Basic cleaning decisions include replacing missing values with 0 where appropriate (i.e., for income in an activity the household was not engaged in) and replacing impossible values (such as more than 24 hours per day) with missing values. I winsorize continuous variables by replacing values above the 99th percentile with the value at the 99th percentile. For incomes I also winsorize values at the 1st percentile.

Household wage income is based on individual earnings over the previous 12 months. Net household non-farm enterprise income is based on revenues from all enterprises less expenses

[4]Post-planting surveys took place in September-October 2010, September-October 2012, August-September 2015, and July-August 2018. Post-harvest surveys took place in February-April of 2011, 2013, and 2016 and in January-February 2019.

[5]The Evans School Policy Analysis & Research Group (EPAR) curates a set of agricultural development indicators using Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) data from several countries including Nigeria. The code is available on GitHub: https://github.com/EvansSchoolPolicyAnalysisAndResearch/LSMS-Agricultural-Indicators-Code.

over the previous 12 months. Net household crop income is based on the value of crop production less explicit expenses, which do not include household labor or other inputs that were not purchased. The value of crop production is based on crop sales revenue plus the value of unsold production, calculated as the quantity not sold times the median price for sales of that crop in the community (or state if there are fewer than 3 instances of sales of a crop within a community). Net household livestock income is calculated similarly by taking the total value of livestock products (e.g., animal labor, eggs, milk, meat, live animals) and subtracting livestock expenditures. I also calculate total income from other sources, such as investments, remittances, pensions, etc.

Because household enterprise, crop, and livestock incomes are net of expenses, some households have negative incomes from these categories. Income shares for different activities are therefore calculated only for households with positive incomes. The shares for activities with negative net income among these household are set to 0. I use these shares to calculate a Herfindahl-Hirschman Index of income concentration by taking the sum of squared income shares across activities, where a sum closer to 1 indicates income is more concentrated in a single activity.

Each household is also associated with community-level data based on surveys with a group of community informants conducted at the same time as the household surveys and on spatial characteristics such as distances to markets and administrative centers, local climate and weather, and geographic characteristics.

### 2.3.2 Flooding data

An important challenge in analyzing floods is measuring which areas have been exposed to flooding. One possibility involves relying on survey reports at either the household or community level (e.g., Freudenreich and Kebede (2022); Gray and Mueller (2012b); Mueller and Quisumbing (2011); Stein and Weisser (2022)). Survey reports identify areas and households directly affected by flooding and can allow analysis of direct flooding as opposed to intent to treat analyses based on being located in an area where floods were reported or observed. Self-reports at both the household and community level are subject to measurement error based on different definitions of what constitutes flooding exposure, recall errors, and potential strategic misreporting.

Local rainfall is a poor proxy for flooding as many flooding events result from the overflow of rivers and other bodies of water. Whether this occurs is a function of upstream rainfall rather than local rainfall. Guiteras et al. (2015) find that the locations where flooding is reported in Bangladesh is not correlated with rainfall at those locations.

In some cases, administrative records may be used to identify flooded areas (e.g., Baez et al. (2020); Berlemann et al. (2023)), though in developing countries the spatial resolution of such records may be limited. For example, the Nigeria Hydrological Services Agency only maintains records of flooding at the Local Government Area level (roughly equivalent to United States counties). These records may also be incomplete. The Nigerian Emergency Management Authority keeps records of communities where flooding was report but fieldwork in Jigawa State revealed that many communities flooded in 2022 were not included in the official records. Other studies (e.g., Djoumessi Tiague (2023); Salazar et al. (2019)) use international databases such as the Dartmouth Flood Observatory Archive (Brakenridge, 2023) and the EM-DAT International Disaster Database (Guha-Sapir et al., 2023). The primary sources for these databases are government or news reports (though remote sensing is incorporate for more recent flooding events), and as such only approximate the spatial extent of flooding and are not comprehensive. Patel (2023) finds that these two sources "agree on the total number of flooding events [in a country-year] just 26.61 percent of the time, despite relying on the same source material" (p. 6), with each database identifying large numbers of floods not recorded in the other. The same study finds that these databased are more likely to underreport flooding "in poorer and more remote parts of the world" (p. 8).

Recent studies take advantage of the availability of high-frequency and high-resolution satellite imagery to measure flood exposure. The main data source is NASA's MODIS (Moderate Resolution Imaging Spectroradiometer).[6] The MODIS instrument is located on NASA's Terra and Aqua satellites and covers Earth's surface every 1-2 days, collecting data on a variety of spectral bands at up to a 250m resolution. These spectral bands are used to identify surface water, and pixel-level flooding is identified based on deviations in the presence of surface water relative to adjacent days and to the same day in other years.

MODIS data informs several databases used to measure flooding. The Dartmouth Flood Observatory uses MODIS data to outline a convex hull of areas where flooding was reported or detected during a given event, but as a result misses masks spatial detail in flood exposure. NASA's Near Real Time (NRT) Global Flood Mapping division records daily flooding at the pixel level (NASA, 2024). The NRT flooding product is increasingly used in the economics literature (see e.g., Akter (2021); Chen et al. (2017); Sajid and Bevis (2021); Vitellozzi and Giannelli (2023)), though some studies work directly with the MODIS data (e.g., Guiteras et al. (2015); Sajid and Bevis (2021)). Satellite data allows the identification of flooded areas at a high spatial and temporal resolution but may also be affected by measurement error, particularly due to the inability of the MODIS instrument to see through clouds. Though the NRT algorithm uses data from multiple adjacent days to identify flooding on cloudy days, it may fail to capture floods that are of short duration and caused by heavy rainfall

[6]Some other sources are used, such as the Joint Research Centre's Global Surface Water Dataset which used NASA's Landsat satellite imagery (Pekel et al., 2016).

meaning that cloud cover will mask the flooding extent.

I use two definitions of flood exposure for this study, both focused on flooding in 2012: the worst year of flooding in Nigeria before 2022. First, I use a *survey-based* definition using household reports of floods in the household shocks module of the 2012-13 GHSP post-harvest round (conducted in February-April 2013). In each survey round, households are asked to report recent years in which they experienced flooding that caused harvest failure or loss of property due to flood (as well as other types of shocks). I define a community as having been exposed to flooding in 2012 if any household within a community reported experiencing flooding. I also record data on flood exposure at the household and community level for other years using the same questions for all rounds of the GHSP (2010-11, 2012-13, 2015-16, and 2018-19; Figure B.1.3).

Second, I use a *MODIS-based* definition using 2-day 500m resolution MODIS data from NASA's NRT flooding product. I identify pixels as having been exposed to flooding if there is any flooding recorded in that pixel in 2012. I match GHSP communities to MODIS flooded pixels based on the community coordinates. These coordinates are randomly offset from the original community centroids, by 0-2km in urban areas and by 0-5km in rural areas (and up to 10km for a random 1% of rural communities). I therefore define community flooding status using MODIS data based on whether the community's offset coordinates are within 5km of a flooded pixel.

### 2.3.3 Sample

Because I am primarily interested in effects on agricultural households I drop communities with no agricultural land in the surrounding $1\text{km}^2$. Household 2012 flooding status, the main treatment variable, is based on the community in which they resided in 2012 rather than their location at the time of the survey in the case of movers (7% of households moved after the 2012-13 survey round). The treatment variable therefore represents whether the community in which the household resided during the 2012-13 survey round was flooded in 2012. I drop households not observed in the 2012-13 survey. This results in a panel sample of 3,769 households in 465 communities. Of these households, 1,071 are observed in all four rounds of the GHSP (2010-11, 2012-13, 2015-16, and 2018-19), 2,482 are observed only in the first three rounds, and 216 are only observed in the first two rounds. Table 2.1 presents summary statistics for the analysis sample.

Three-quarters of sample households reside in a community defined as rural, and 36% of the land in the $1\text{km}^2$ surrounding the community centroid is classified as agricultural on average. The median household resides in a community with no reported flooding from 2007-2011, but 49% of households reside in communities that experienced at least one year of flooding, and 22% were exposed to at least two years of flooding. 31% of households

Table 2.1: Sample summary statistics

| | Mean | SD | Min | $25^{th}$ | $50^{th}$ | $75^{th}$ | Max | N |
|---|---|---|---|---|---|---|---|---|
| *Community characteristics* | | | | | | | | |
| Rural | 0.75 | 0.43 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 3769 |
| Percent ag land w/in approx 1km | 36.36 | 26.57 | 0.0 | 13.0 | 35.0 | 56.0 | 100.0 | 3769 |
| Years community reported flooding from 2007-2011 | 0.85 | 1.12 | 0.0 | 0.0 | 0.0 | 1.0 | 6.0 | 3769 |
| Mean annual share of months with drought 1998-2014 | 0.09 | 0.02 | 0.0 | 0.1 | 0.1 | 0.1 | 0.1 | 3769 |
| Exposed to 2012 flooding: survey-based | 0.31 | 0.46 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 3741 |
| Exposed to 2012 flooding: MODIS-based | 0.20 | 0.40 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 3769 |
| *Household characteristics* | | | | | | | | |
| Female-headed HH | 0.13 | 0.33 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 3769 |
| Count of HH members | 5.83 | 3.16 | 1.0 | 4.0 | 6.0 | 8.0 | 31.0 | 3769 |
| Livestock holdings (TLU) | 1.06 | 2.86 | 0.0 | 0.0 | 0.0 | 0.9 | 20.1 | 3769 |
| Value of HH assets (USD) | 1377.91 | 2995.60 | 0.0 | 176.3 | 500.7 | 1243.8 | 24916.3 | 3769 |
| Months of HH food insecurity in last year | 0.40 | 1.02 | 0.0 | 0.0 | 0.0 | 0.0 | 6.0 | 3741 |
| HH under 1.90 USD PPP pc daily poverty line | 0.36 | 0.48 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 3769 |
| HH uses formal financial services | 0.28 | 0.45 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 3769 |
| *Household productive activities* | | | | | | | | |
| Total annual HH net income (USD) | 3209 | 6723 | -8884 | 71 | 1322 | 3779 | 56833 | 3769 |
| Any HH crop production activity | 0.76 | 0.43 | 0.0 | 1.0 | 1.0 | 1.0 | 1.0 | 3728 |
| Any HH livestock activity | 0.52 | 0.50 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 | 3728 |
| Any HH non-farm enterprise activity | 0.73 | 0.44 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 | 3728 |
| Any HH wage work activity | 0.47 | 0.50 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 3728 |
| Total crop and pasture holdings (ha) | 1.57 | 2.90 | 0.0 | 0.0 | 0.4 | 1.7 | 14.9 | 3769 |
| HH uses any inorganic fertilizer | 0.45 | 0.50 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 2604 |
| Value of crop sales (USD) | 1441.23 | 2866.34 | 0.0 | 0.0 | 486.5 | 1692.6 | 33966.4 | 3769 |
| Prop. of value of crop production sold | 0.23 | 0.29 | 0.0 | 0.0 | 0.1 | 0.4 | 1.0 | 2421 |

Note: This table summarizes the baseline (2010-11) characteristics of sample households and their communities. Values are in constant 2016 USD. The household income Herfindahl-Hirschman Index is calculated as the sum of the squared shares of total income from farming, household non-farm enterprise, and wage work, among households with non-negative household net income.

were exposed to flooding in 2012 according to the survey-based definition compared to 20% according to the MODIS-based definition.

The median household in the 2010-11 survey round was male-headed with 6 members, no livestock, and USD 501 worth of household assets. Households experienced 0.4 months of food insecurity in the past year, though 36% are under the USD 1.90 PPP per capita daily poverty line based on their total annual income. Around one quarter of households use any formal financial services.

The median household had total annual net income of USD 1,322, but 15% of households have negative estimated annual incomes due to negative profits in household farm and enterprise activities. Over three-quarters of households are engaged in any crop production, similar to the 73% engaged in any non-farm enterprise activity. Just over half rear livestock and 47% had any household member engaged in any wage employment. The median household has 0.4 ha of crop and pasture land and sells 10% of the value of their crop production.

The large share of households with any member engaged in any wage employment in these largely rural and agricultural settings suggests some baseline level of baseline income diversification. Participating in wage work alongside household farm and non-farm activities is fairly common: 34% of households engage in both crop production and some wage employment, and 35% engage in both non-farm enterprise and wage employment. But the mean household income Herfindahl-Hirschman Index (HHI)—the sum of squared income shares for household farm income (crop and livestock), non-farm household enterprise income, and wage income among households with non-negative net income—is 0.87 and the median is 1. This implies that households typically rely on only one activity for the large majority of their annual income.[7]

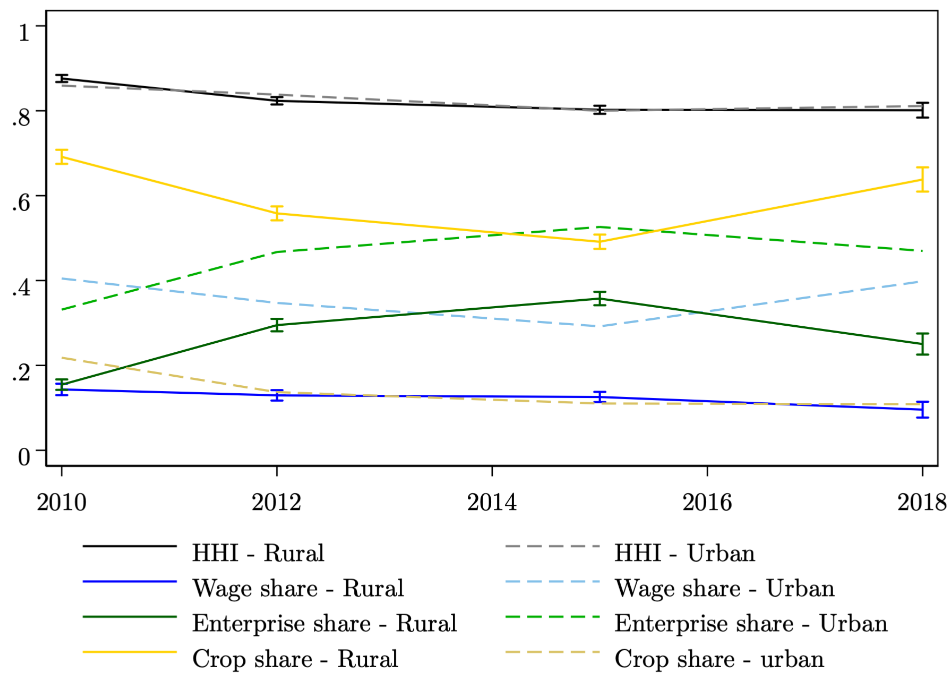
Figure 2.1 shows that income concentration has decreased slightly over time in Nigeria but remains high, with no difference between rural and urban communities. The share of total household net income from wage work has similarly stayed relatively stable at around 15% in rural areas and between 35-40% in urban areas. The share of income from household non-farm enterprise increased in both rural and urban areas from 2010 to 2015 before falling somewhat in 2018. Non-farm enterprise accounts for around 45% of urban income and 25% of rural income on average during the sample period. The residual share is nearly household farm income, as mean incomes from other sources are very small in this sample. These patterns suggest limited livelihood diversification or structural transformation on average at the household level in Nigeria from 2010-2018.

## 2.4 Results: 2012 flood exposure as seen from the sky and on the ground

Table 2.1 shows that the share of households identified as exposed to flooding in 2012 differs significantly between the survey-based and MODIS-based definitions, with a larger share classified as exposed to the 2012 floods according to the survey-based definition. In this section I evaluate the alignment between different survey-based measures of community flood exposure in 2012 and the MODIS-based definition, and explore reasons for differences in which communities are identified as flooded under the main survey-based definition and the MODIS-based definition.

[7] A HHI value of 1 may also indicate households engaged in multiple activities where only one of them yielded a positive net income after accounting for production expenses.

Figure 2.1: Mean household income shares by survey round and rural status




Note: The figure plots the mean across households in each survey round by rural/urban status for each variable. Surveys were conducted in 2010-11, 2012-13, 2015-16, and 2018-19. The Herfindahl-Hirschman Index (HHI) is the sum of squared income shares for household farm income (crop and livestock), non-farm household enterprise income, and wage income and can take a value between 0 and 1, where 1 indicates that all household income comes from a single activity. The wage and enterprise shares are the share of total household income represented by household wage earnings and household non-farm enterprise earnings, respectively. All variables are only measured among households with non-negative net income (91% of households across rounds).

### 2.4.1 Differences across survey flood measures

The main survey-based flooding definition uses information from the post-harvest household shocks survey module, but households can also indicate experiences with flooding in other survey modules. In the 2012-13 post-planting agricultural survey, households can list flooding as the main cause of loss of stored crops since the beginning of the new year for each cultivated crop. In the post-harvest agricultural survey, households can list flooding as a reason for not harvesting a particular crop. In both the post-planting and post-harvest household surveys, households can list flooding as a cause of household food insecurity in the past 12 months. These questions only capture specific aspects of flood exposure but can be used to check household recall of overall flood exposure in the household shocks module.

Seven percent of GHSP sample households report having been affected by floods in 2012 (Figure B.1.3), with nearly all of these reporting harvest failure caused by floods and 17% reporting loss of property. Among households affected by floods, 56% list floods as the most severe shock they faced in the past 5 years. The most common responses are receiving assistance or borrowing from friends and family, selling livestock and other assets, and reducing food and non-food consumption.

Related to this, 3% of households in both the post-planting and post-harvest survey rounds in 2012-13 report experiencing food insecurity over the past 12 months that was caused primarily by flooding. But just 53 of 155 households reporting food insecurity caused by floods in the 12 months preceding the post-harvest survey (which took place in February-April 2013) also report being affected by floods in 2012. This highlights a potential shortcoming of using the shocks module to measure flood exposure: respondents are only asked about floods that caused harvest failure or that caused loss of property.[8] The questions may therefore fail to capture floods that damaged but did not destroy property or that had effects outside of harvest failure and property loss.

In the agricultural survey modules, 1.2% of households report losing stored crops due to flooding since the start of 2012 at the time of the post-planting visit in September-October (12.7% report any such losses), and 3.1% report not harvesting any of at least one crop as a result of floods at the time of the post-harvest visit (20.3% report any such harvest failure). The overlap here with floods reported in the shocks module is stronger: 96 of 140 household reporting flood-related harvest failure in the agricultural module also report being affected by floods in shocks module. The fact that more than twice as many households report harvest failure caused by floods in the shocks module as report any complete failure to harvest a crop due to floods suggests the definition of harvest failure in the shocks module is more general than the total failure asked about in the agricultural module. But the 44 households reporting total harvest failure due to floods for at least one crop and not reporting any flooding-related harvest failure in the shocks module indicates important measurement errors in the household survey.

Aggregating across all the different flood-related questions, the share of households reporting begin affected by floods at any point in the 2012-13 survey rises to 12.3%, compared to 7.4% reporting floods in the shocks module which I use for the main survey-based flooding definition I apply in this paper. The above analysis indicates that household flooding reported in the shocks module may miss some floods that did not cause harvest failure or property damage but affected household food insecurity, and may also miss other flooding that the respondent failed to recall or report.

[8] Respondents have the option to specify other shocks, but none of the responses explicitly mention floods.

How do these household flooding reports compare to what is reported in the GHSP community surveys? In each post-harvest community survey (in 2009, 2013, 2016, and 2019), a group of community informants is asked to recall major community events over the last few years, including floods. Twenty-nine percent of households reside in communities where flooding was reported in 2012 in the community survey. Yet 34% of households reporting being affected by flooding in the shocks module reside in communities where no flooding was reported. This rises to 41% if I consider the broader measure of household flooding reports across all survey modules.

As a result, even though the shares of communities identified as flooded in 2012 using the different measures are similar—29% reporting flooding in the community survey, 27% with at least one household reporting flooding in the shocks module, and 40% with at least one household reporting flooding anywhere during the survey—the measures identify different communities as flooded. While community survey- and household survey-based measures of community flood exposure agree around 80% of the time about which communities were not flooded in 2012, agreement about which communities *were* flooded is only around 50-60% and there is little improvement if I allow for household survey error by requiring at least two sample households in the community to report flooding for the community to be considered exposed (Table B.2.1).

An obvious reason for the community survey to report flooding in 2012 even if no sample households report it is that the flooding affects community households other than those included in the GHSP sample. Households reporting flooding not captured in the community survey may be driven by lack of complete information by the group of community survey informants, or by error in the household surveys. But even for the case where at least three separate sample households report flooding in the shocks module there are 13 communities in which no flooding was reported in the community survey. This suggests important measurement errors in the community survey and motivates my reliance on the household reports to define survey-based community flood exposure for the main analyses rather than the community reports.

### 2.4.2 Identifying flooding using surveys and satellites

How do survey-based community flooding definitions align with flooding identified using the MODIS-based definition? Table 2.2 presents an overview, focusing on the community survey and household shocks module reports of flooding—results are similar considering the measure of household flood reports across the entire survey. I also include a broader survey-based measure of 2012 flooding exposure defining communities as exposed to flooding if *either* the community survey or any household survey reports flooding, and a narrower measure of cross-validated flooding exposure including only communities where flooding is reported in *both* the community survey and at least one household survey.

Table 2.2: Alignment of survey- vs. MODIS-based definitions of 2012 GHSP community flood exposure

| | | MODIS-based flooding definition | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | Row total | Share agreed |
| Survey-based flood definition | | 364 | 101 | 465 | |
| Any household with | 0 | 278 | 59 | 337 | 0.82 |
| shocks module flooding | 1 | 86 | 42 | 128 | 0.33 |
| At least 2 households with | 0 | 313 | 77 | 390 | 0.80 |
| shocks module flooding | 1 | 51 | 24 | 75 | 0.32 |
| Community survey | 0 | 265 | 66 | 331 | 0.80 |
| reported flooding | 1 | 99 | 35 | 134 | 0.26 |
| Either any household or | 0 | 220 | 55 | 275 | 0.80 |
| community reported flooding | 1 | 144 | 46 | 190 | 0.24 |
| Both any household and | 0 | 319 | 75 | 394 | 0.81 |
| community reported flooding | 1 | 45 | 26 | 71 | 0.37 |

Communities considered not flooded in 2012 according to GHSP reports are also considered not flooded by the MODIS-based definition between 80-82% of the time for all survey-based measures, but the level of agreement about which communities *were* flooded is much lower. Just 24% of communities identified as exposed to flooding by the broadest survey-based measure are also identified as flooded by the MODIS-based measure. This increases to 37% for the narrower definition with both community and household flood reports, close to the 33% considering just the household reports. Overlap is worse for community survey flood reports than for household survey flood reports, potentially indicating lower quality data in the community surveys. As the count of communities identified as flooded by both MODIS and the GHSP is higher with the survey-based measure using any household report of flooding in the household shocks module than with the narrower measure, the former is the only survey-based measure I consider for the remainder of this paper.[9]

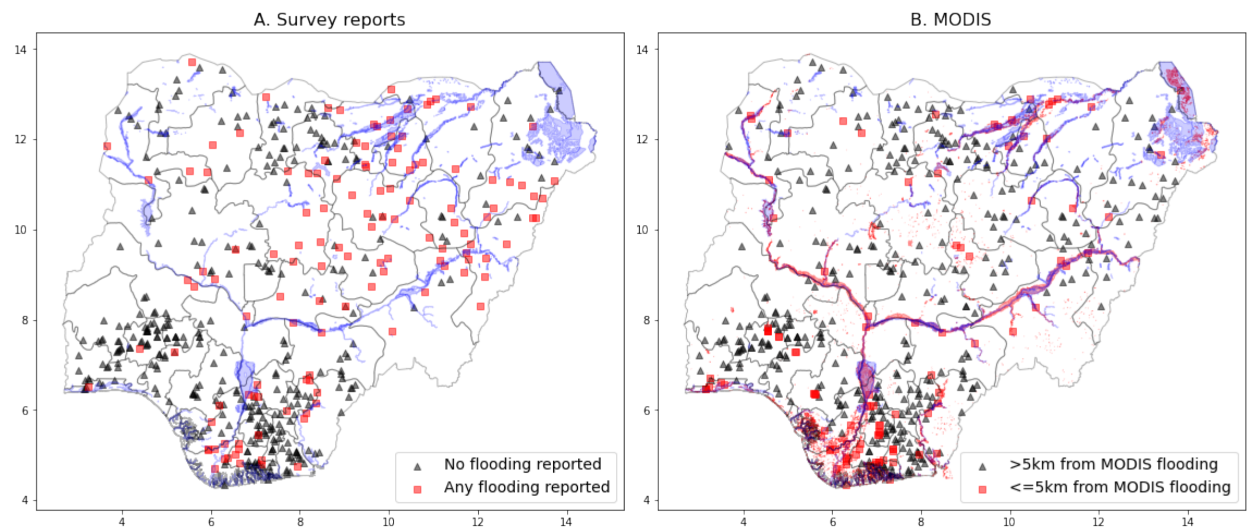
That the MODIS-based definition does not identify flooding in around two-thirds of GHSP communities where there are multiple separate reports of flooding—reducing concerns about misreported flood exposure—strongly indicates that the algorithms using MODIS to detect floods are missing a large share of events perceived by households and communities as damaging floods. Equally important is that the MODIS-based measure is also detecting many water-related events that are not perceived by local communities as having sufficient negative effects to be reported in the various GHSP questions about flooding. Figure 2.2 maps the locations of GHSP communities and their 2012 flood exposure status based on these two definitions. The figure illustrates how widespread flooding was across the country but also highlights differences in community flood exposure between the two definitions. Of 465 GHSP communities, 128 had at least one household report flooding in 2012 and 101 are within 5km of a MODIS flooded pixel, but just 42 communities are considered flooded according to both definitions.

This is similar to the finding of Guiteras et al. (2015), who show that household reports of flooding in Bangladesh in 2004 are poorly correlated with exposure measured using MODIS data. They argue that this is caused by measurement issues in the household reports due to recall error, reference dependence, and adaptation.

The types of issues they discuss could help explain the 59 communities that are within 5km of a flooded pixel but where no flooding was recorded in the GHSP. Recall error seems unlikely to be a meaningful explanation as all surveyed households (around 10 per community) would have had to fail to recall a recent high-profile flooding event, and the number changes little if I also include community survey reports. Reference dependence and adaptation may be more important. The binary satellite-based measure of flood exposure captured by the NRT algorithm does not measure flood intensity such as number of days flooded or depth of inundation. If the flooding in these communities was relatively mild relative to their experiences and expectations or if they had taken measures to adapt, the floods might not have adversely affect residents and therefore not been reported in the surveys.

[9] I also consider agreement between the survey-based and MODIS-based flooding definitions and a third measure based on the Nigerian Emergency Management Authority's records of flooding. These records include a list of Local Government Areas (LGAs) where flooding was reported in 2012, drawing from both media reports and local government officials. NEMA records 2012 flood reports in the LGAs of 94 of the 465 GHSP communities, but overlap with communities identified as flooded by the other measures is limited. Just 21% of communities exposed to flooding according to the MODIS-based definition are in LGAs included in NEMA's 2012 flooding records, compared to 28% for the survey-based definition. This implies that the NEMA records are missing the majority of flood occurrences in Nigeria in 2012, in line with other studies finding that media-based flood event records are more likely to not report small floods and floods in less urban and wealthy areas (Patel, 2023).

Figure 2.2: 2012 community flooding exposure, by definition




Note: Each panel maps the locations of communities surveyed in the Nigeria GHSP, colored by their flood exposure in 2012. In Panel A, flood exposure is based on survey reports: any flood event reported by any household in the community. In Panel B, flood exposure is based on being within 5km of a pixel where flooding was identified using MODIS satellite data. MODIS flooded pixels are shown in red. In both panels, rivers and bodies of water are shown in blue and state boundaries are shown in gray.

I test whether there are systematic differences in the characteristics of communities defined as flooded according to the two definitions, and summarize the results in Table 2.3. Relative to communities defined as flooded according to both the survey and MODIS, communities defined as flooded according to MODIS only are 40 percentage points less likely to be rural, have 33% less agricultural land, report fewer floods over the 5 years before 2012, have 174mm (14%) more annual rainfall on average, and experienced 10% more rainfall in 2012.

Some of these differences could have affected survey flood reporting. These communities are used to higher levels of rainfall and are less reliant on agriculture, so may not have perceived conditions in 2012 as abnormal and may have been more prepared for or less affected by any flooding. On the other hand, these communities had less recent exposure to floods so we might have expected the 2012 floods to be more salient and likely to be reported.

Another potential reason for the 59 communities within 5km of MODIS-observed floods but with no survey-reported flooding is noise in the community coordinates. Flooding may

Table 2.3: Differences in GHSP community characteristics by 2012 flooding status

| | N | Mean: Both definitions N=42 | $\beta_1$: Survey only N=86 | $\beta_2$: MODIS only N=59 | $\beta_3$: No flooding N=278 | $\beta_1 = \beta_2$ (*p-val*) | $\beta_1 = \beta_3$ (*p-val*) | $\beta_2 = \beta_3$ (*p-val*) |
|---|---|---|---|---|---|---|---|---|
| Rural | 465 | 0.79 | 0.06 | -0.40*** | -0.09 | 0.00*** | 0.01*** | 0.00*** |
| | | (0.42) | (0.08) | (0.09) | (0.07) | | | |
| Percent ag land w/in approx 1km | 465 | 34.14 | 13.60*** | -11.19** | -2.86 | 0.00*** | 0.00*** | 0.02** |
| | | (26.37) | (4.67) | (5.01) | (4.11) | | | |
| Years community reported flooding from 2007-2011 | 465 | 1.21 | 0.20 | -0.64*** | -0.72*** | 0.00*** | 0.00*** | 0.49 |
| | | (1.09) | (0.19) | (0.20) | (0.17) | | | |
| Distance to nearest MODIS flooded pixel (km) | 465 | 1.80 | 16.30*** | 0.71 | 16.61*** | 0.00*** | 0.84 | 0.00*** |
| | | (1.61) | (2.03) | (2.18) | (1.79) | | | |
| Distance to nearest water area (km) | 465 | 11.89 | 13.45** | 10.66* | 31.22*** | 0.50 | 0.00*** | 0.00*** |
| | | (21.62) | (5.41) | (5.80) | (4.76) | | | |
| Average 12-month total rainfall (mm) | 465 | 1253.86 | -150.67* | 173.60** | 71.79 | 0.00*** | 0.00*** | 0.10 |
| | | (496.57) | (82.00) | (87.94) | (72.11) | | | |
| 12-month total rainfall (mm) in 2012 | 465 | 1301.24 | -71.02 | 135.76* | 66.06 | 0.00*** | 0.00*** | 0.19 |
| | | (443.43) | (69.50) | (74.54) | (61.12) | | | |
| Total rainfall in wettest quarter in 2012 | 465 | 664.90 | 28.32 | -9.02 | 4.13 | 0.03** | 0.11 | 0.46 |
| | | (152.69) | (23.27) | (24.96) | (20.46) | | | |

Note: This table presents the results from separate regressions of community characteristics on dummies for exposure to flooding in 2012 according to different definitions. Data are from the GHSP 2012-13 survey round, except distance to a water area which is calculated using community coordinates and a shapefile of water bodies in Nigeria. The reference group is communities considered flooded under both the survey-based and MODIS-based definitions. The survey-based definition considers flooding as any 2012 flood event reported by any household in the community. The MODIS-based definition considers flooding as being within 5km of a pixel where flooding was identified in 2012 using MODIS satellite data.

have taken place very near to these communities without damaging households, farms, or infrastructure, and therefore without being reported in the community or household surveys.

The offsetting of GHSP community coordinates will also explain part of the 86 communities where flooding was reported in 2012 but not detected by MODIS. Twenty-five of these communities are within 10km of a MODIS flooded pixel. The true locations of these communities may in fact overlap with the flooded area since the community coordinate is a centroid of sample household coordinates and many communities will extend across several kilometers. But indiscriminately using a 10km buffer around MODIS flooded pixels to define community flood exposure would also include 72 more communities where no flooding was reported in 2012 in the GHSP. Increasing the buffer for the MODIS-based flooding definition will therefore not improve the overlap with the survey-based definition.

The 61 communities where flooding was reported in the surveys but that are more than 10km for a MODIS flooded pixel could result from different sources of survey measurement error. First, there may be strategic misreporting if respondents hope that stating they were affected by floods will lead to the receipt of government funding in the community. This explanation does not seem particularly likely as on average 24% of households (around 10 are sampled in each community) in these 61 communities report being affected by flooding, not far below the 32% on average in the 67 communities within 10km of a MODIS flooded pixel where flooding was reported in the survey (though the difference is statistically significant).

Second, the survey reports may refer to flooding that occurred outside the community but that nevertheless affected community members. Floods outside the community could have been reported in the household shocks module if households have agricultural plots or other property far away from the community center. Other types of recall error, such as reporting floods occurring nearby that did not directly affect the household, are unlikely as the household shocks module only asks about floods that caused household harvest failure or loss of property.

Third, flood reports in the surveys only could also reflect flooding or other water-related damages that the NRT flood detection algorithm cannot identify. Survey respondents may report water-related damages from heavy rainfall that did not involve inundation.[10] In addition, the MODIS satellite instrument cannot see through clouds so rapid flooding caused by heavy rainfall may not be detected.

Communities defined as flooded according to the survey only have a greater share of agricultural land in the surrounding area than communities defined as flooded according to both the survey and MODIS, are 13.5km farther on average from a body of water, and experience 151mm (12%) less annual rainfall on average. These differences and those for the floods identified by MODIS only suggest the NRT flood detection algorithm may be more effective at identifying flooding in more built-up urban areas where heavy rainfall is more likely to pool rather than be absorbed by the soil, and potentially less effective at identifying flooding in some agricultural areas far from bodies of water. In addition, areas where floods are reported in the surveys only experience less total rainfall on average but experienced significantly more rainfall in the wettest quarter of 2012 than areas identified as flooded by MODIS only. This suggests their rainfall in 2012 may have been particularly concentrated, leading to reports of flooding that MODIS might have missed due to cloud cover.

In summary, the two main flooding definitions identify different sets of flooded communities. While the MODIS-based definition is more objective it may capture some flooding that was not damaging to communities and households and may fail to capture certain types of flooding. The survey-based definition is subject to different types of measurement error but may more accurately reflect exposure to flooding or other water-related shocks that negatively affected community members. While Guiteras et al. (2015) argue that survey re-

[10]During interviews with farmers in Nigeria's Jigawa State the year after severe floods in 2022, in some cases farmers reported experiencing flooding and when pressed to describe it related stories of heavy rainfall damaging buildings made from mud rather than inundation by water.

ports of flooding in Bangladesh have little value relative to satellite-based measures because households "perceive exposure relative to their average environment" (p. 235), if we are interested in analyzing the effect of floods as an economic shock it is unclear whether one source should be preferred to the other. In the remainder of the paper, I estimate impacts of flood exposure on outcomes for GHSP households using both flooding definitions and discuss differences in the conclusions that follow.

## 2.5 Empirical approach: Impacts of flood exposure

I analyze the impacts of community exposure to flooding in 2012 on household income diversification in subsequent years, considering treatment definitions based on both flooding definitions. As not all households in flooded communities were themselves flooded, this specification estimates an intent-to-treat effect of having been in a community that experienced flooding in 2012. This estimator will capture direct effects from households that were themselves flooded as well as any indirect impacts from proximity to flooding.[11]

The major flood events in 2012 largely occurred between August and November, at the same time as the 2012 post-planting survey. The 2012-13 round (along with the 2015-16 and 2018-19 rounds) is therefore considered post-flooding even though some income activities will have been realized prior to the flooding. I use data from 2010-11 survey round to serve as a pre-flooding comparison period and to test for baseline balance in household and community characteristics.

I estimate impacts of the 2012 flood using a Difference-in-Differences setup. Specifically, I estimate two-way fixed effects regressions

$$Y_{icst} = \alpha + \beta Flood_c \times Post_t + \mu_i + \gamma_{st} + \epsilon_{icst} \tag{2.1}$$

where outcome $Y$ varies across households $i$, communities $c$, states $s$, and years $t$. *Floods* is an indicator for having been a resident in 2012 of a community which experienced flooding in 2012 and *Post* is an indicator for being observed in a survey round after 2011. I include household fixed effects ($\mu$) to control for time-invariant household characteristics which might affect livelihood decisions and 2012 flood exposure, and state-by-year fixed effects ($\gamma$) to control for common changes over time across broad geographic areas. The effects of experiencing flooding in 2012 are therefore identified using variation within households over time and across communities within states and survey rounds. I test the robustness of the

[11]Because household outcomes may be affected by proximity to flooding even without direct losses—through effects on local infrastructure or expectations of flood risk, for example—I do not use community-level flooding to instrument for household-level flooding due to concerns about violations of the exclusion restriction.

results to including time-varying household and community controls—including household- and community-reported exposure to flooding in other years—and varying the set of fixed effects. I cluster standard errors at the level of the community of residence in 2012 since this is the level at which the flooding treatment is assigned.

This is a 'canonical' difference-in-differences model where all units receive the treatment—being a resident in 2012 of a community that flooded—at the same time and the control group never receives this treatment. The key assumption of this model is the parallel trends assumption, that households that resided in communities which flooded in 2012 would have experienced similar trends over time as households that resided in communities which did not flood, if it had not been for the flooding.

If the occurrence of flooding in 2012 were random over space, this assumption would hold in expectation. As I include state-by-year fixed effects in my regression models the assumption is effectively that flooding is random within states, but this is unlikely to be true, particularly as rivers overflowing accounts for a large share of floods. Communities that flooded in 2012 are likely to have different geographic characteristics than communities that did not. As a result, baseline household productive activities and trends over time are also likely to differ, to the extent that geographic attributes associated with flood risk also affect economic outcomes. Indeed, I find that characteristics of communities and households differ significantly by flooding status on a number of baseline characteristics (as measured in 2010-11) according to both the survey-based and MODIS-based flood exposure definitions (Table B.2.2, Table B.2.3).

These differences in baseline characteristics indicate a need to identify a group of communities not exposed to flooding in 2012 that represents a better counterfactual for the flooded communities. I accomplish this by estimating the probability that a community was exposed to flooding in 2012 and using these estimated probabilities to construct inverse propensity weights. This approach puts a greater weight on communities that did not flood but are estimated to have been at high risk of flooding, and therefore might be better controls for flooded communities. For each of the two flooding definitions (survey-based and MODIS-based), I estimate a logit regression for the likelihood that a community was exposed to floods in 2012 based on geographic characteristics and prior flooding history,[12] predict a flooding probability, and construct inverse propensity weights based on the estimated probability. I set the weight to 0 for communities whose estimated probability is outside the common support range. Each household is assigned a weight based on the community where it resided in 2012. The resulting samples are well-balanced in terms of community and household charac-

[12] I use Stata's *lasso2* package to select the set of variables to include in the logit model. The set of included variables differs for the two flooding definitions but in both cases survey reports of flooding between 2005 and 2011 are included.

teristics after including the estimated weights (Table B.2.2, Table B.2.3). Importantly, the samples also have similar trends in survey flooding reports each year from 2008-2018 (excluding 2012), indicating the weights create balance in risk of flooding (Figure B.1.4). This balance in baseline characteristics and flooding risk increases the plausibility of the parallel trends assumption.

The main specifications I report include the full sample of communities and weight observations using the 2012 flooding inverse propensity weights (IPW) according to both the survey-based and MODIS-based flooding definitions. I test the sensitivity of the results to dropping these IPW and restricting the sample to drop control communities that were never within 10km of a flooded community between 2008-2018. I also match communities by predicted 2012 flooding risk within strata of years of flood exposure between 2005 and 2011 and include strata-by-round fixed effects. Results are similar in these alternative specifications (Figure B.1.6). I do not directly control for estimated flooding risk as this is absorbed by household or community fixed effects.

## 2.6 Results: Impacts of flood exposure

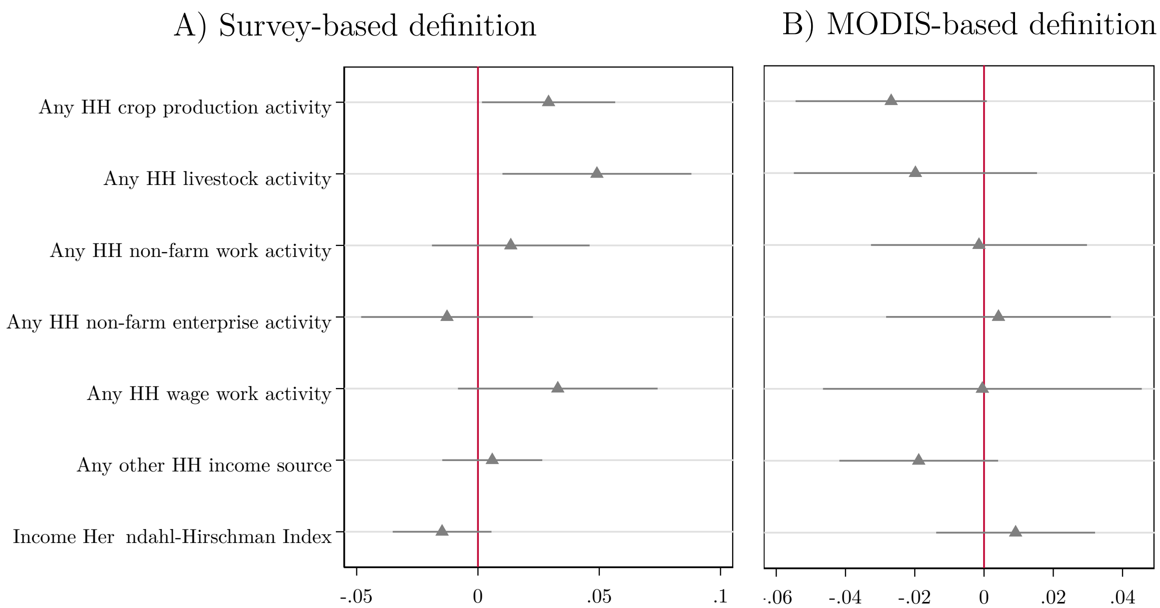
Panel A of Figure 2.3 shows that households exposed to flooding in 2012 according to the survey-based measure are 3-5 percentage points more likely in household crop and livestock production than unexposed households in the periods after exposure, though only the effect on livestock activity remains statistically significant after a false discovery rate (FDR) adjustment for multiple hypothesis testing. This increased engagement is reflected in a 0.23 ha increase in crop area planted—68% of the mean in non-exposed communities in 2010-11—driven by cultivation of staple crops,[13], as well as increases in household crop labor and in crop production expenditures (Figure B.1.7 Panel A).

In contrast, Panel B shows that when we consider the MODIS-based measure of 2012 flood exposure the effect on engagement in crop production becomes *negative*, though the significance of this effect does not survive FDR adjustment. Average crop area planted does not change significantly though the point estimate is negative, but hired crop labor days decrease by 52% (Figure B.1.7 Panel B).

There is no effect of 2012 flood exposure on engagement in non-farm work activities or on the concentration of household income sources (measured by the HHI), according to either flood measure.

[13]The main staple crops in Nigeria include maize, millet, sorghum, rice, cassava, yam, cowpea, and groundnut. I define all other crops, including cocoa, sweet potato, soybeans, kola nuts, oil palm, sesame, fruit crops, and vegetables, as 'commercial' crops.

Figure 2.3: Average impacts of flood exposure in 2012 on household engagement in different livelihood activities, by flooding definition




Note: The figure shows the coefficient and 95% confidence interval from separate regressions of engagement in different livelihood activities on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. Household engagement each survey round is based on any household or household member activity in a given area in either the post-planting or post-harvest survey. The Herfindahl-Hirschman Index (HHI) is the sum of squared income shares for household farm income (crop and livestock), non-farm household enterprise income, and wage income and can take a value between 0 and 1, where 1 indicates that all household income comes from a single activity. All regressions include household and state-by-year fixed effects. Standard errors are clustered at the level of the community of residence in 2012. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. In Panel A, flood exposure and the sample weights are defined based on survey reports for the community of residence in 2012. In Panel B, they are defined based on the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data. Results and FDR-adjusted q-values are shown in Table B.2.4.

Figure 2.4 shows that both flood definitions indicate 2012 flooding exposure does not significantly affects total household income in subsequent years. There are also no effects of flood exposure on livestock income, wage income, or income from other sources using either flood measure. In line with this I find no effects on two measures of household well-being: whether household daily per capita consumption is below the international poverty line of USD 1.90 PPP and the total value of household assets (Figure B.1.5). In contrast, exposure to the 2012 floods increases the probability of the household having experienced any food insecurity over the past 12 months by between 0.15 and 0.2 standard deviations—5 to 7 percentage points—according to both flooding definitions, though the intensity of this food insecurity is low, with flood exposure increasing months of food insecurity by around 0.2 months. These results indicate that households in communities exposed to the 2012 floods do not become significantly worse off on average over the following 4-7 years relative to households in non-flooded communities, with the exception of a small but significant increase in food insecurity.

Although flood exposure does not significantly change total income, net crop production income decreases by USD 307 on average over the following survey rounds according to the survey-based definition. This represents a 29% decrease relatively to the mean for control communities before the 2012 floods. While the point estimate for the effect of flooding on net crop income is also negative for the MODIS-based definition it is not statistically significant.

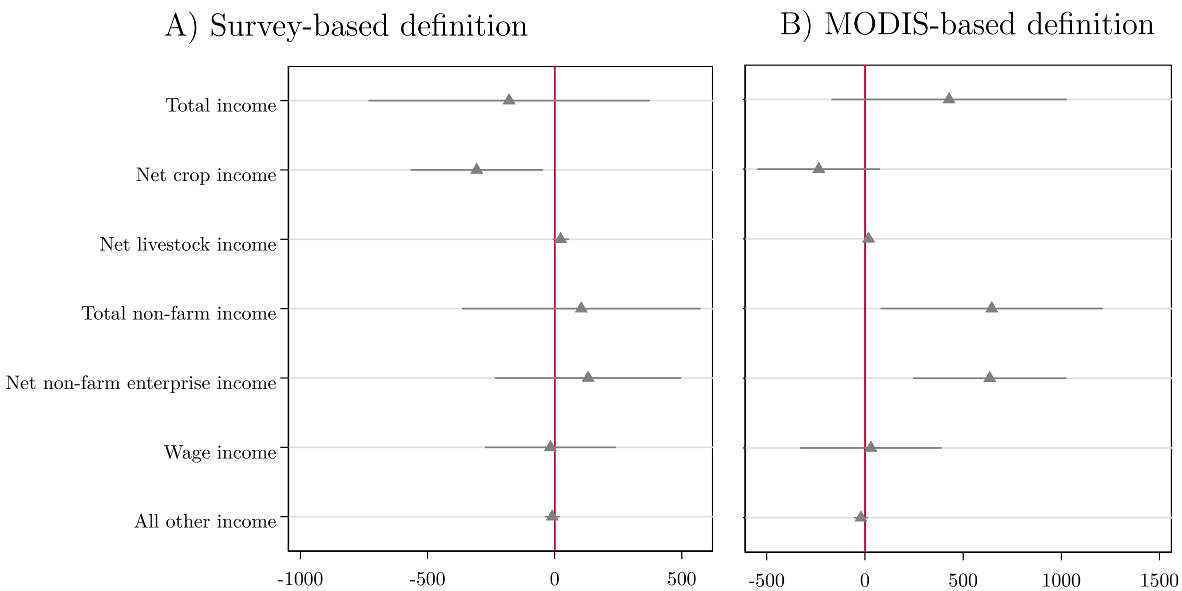
The calculation of crop 'income' include not just realized income from crop sales but also unrealized income from the value of crops produced for household consumption. I find that the value of crop sales falls significantly according to the survey-based definition, but the decrease in net crop income is driven by a significant fall in the value of crop production in flooded communities relative to non-flooded communities. This 21-23% decrease relative to 2010-11 levels is statistically significant according to both flooding definitions, and is driven by decreases in the value of more commercially-oriented crops of USD 391 and 452 according to the MODIS- and survey-based definitions, respectively (Figure B.1.7). The decrease captured by the MODIS-based definition may reflect at 37% decrease in area planted with these crops ($p = 0.102$) relative to non-flooded communities, but there is no change in their area planted under the survey-based definition. This suggests that flood exposure led to decreases in either productivity in these crops or in their prices.[14]

Although the large and negative average point estimate on net crop income using the MODIS-based definition is not statistically significant, crop income decreases by significantly more following flood exposure for households that were engaged in non-farm enterprise at baseline (Figure B.1.8 Panels B and D). Households engaged in non-farm enterprise at baseline increased household crop labor hours by less and drive the decrease in the value of crop production after flood exposure (Figure B.1.9). Households below the poverty line—which increase labor to non-farm enterprises—also have significantly larger decreases in crop production value. I observe the same patterns using the survey-based flooding definition.

I do not find evidence of changes in the count household members working in household agriculture in any group of households following the 2012 floods. The exception is that the survey-based flooding measure shows that the count of members working in household agriculture increase significantly after flood exposure in 2012 for households not engaged in

[14] I find no evidence of effects on prices of bananas or cocoa, the two most commonly cultivated non-staple crops.

Figure 2.4: Average impacts of flood exposure in 2012 on total net household income (in 2016 USD) from different livelihood activities, by flooding definition




Note: The figure shows the coefficient and 95% confidence interval from separate regressions of total household net income from different livelihood activities on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. The units for all income variables are constant 2016 USD. Net income from household farm activities is calculated by taking the value of all household farm output (including output produced for own consumption) and subtracting expenditures on all purchased inputs over the prior agricultural year. Net income from household non-farm enterprise is calculating as total revenues minus total expenses over the prior 12 months. Income from wage work is calculated as total earnings over the prior 12 months. All regressions include household and state-by-year fixed effects. Standard errors are clustered at the level of the community of residence in 2012. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. In Panel A, flood exposure and the sample weights are defined based on survey reports for the community of residence in 2012. In Panel B, they are defined based on the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data. Results and FDR-adjusted q-values are shown in Table B.2.5.

crop production in 2010-11, relative to similar households not exposed to the floods. This is consistent with the probability of having any household farm activity (crop and livestock) increasing on average in the years following the 2012 floods in flooded communities relative to non-flooded communities (Figure 2.3). As households not previously engaged in crop production did not significantly increase their net crop income in the post-floods period—the point estimate is negative—this implies that any entry into crop production was at a small scale and not very productive.

In contrast to the decrease in net crop production income, with the MODIS-based definition I find that flood exposure causes a highly significant USD 636 (74%) increase in annual net non-farm enterprise income (Figure 2.4 Panel B). This measures the sum of revenues minus operating costs from household enterprises not related to crops or livestock. The increase in enterprise profits despite no change in the probability of households engaging in such activities indicates either an increase in profitability or an increase in scale of existing enterprises over time in flooded relative to non-flooded communities. There is no effect on net non-farm income using the survey-based definition.

Effects on enterprise income using the MODIS-based definition are driven exclusively by households already engaged in such activity at baseline in 2010-11 and by households with below the median value of assets (Figure B.1.8 Panels A and C). The count of household members working in non-farm enterprise does not change on average after flood exposure in households previously engaged in this activity, but the point estimate is positive and close to marginally significant for households with no non-farm enterprise activity in 2010-11. Less wealthy households at baseline in terms of both asset value and consumption per capita increase the number of household members working in household enterprises—by around 0.25 members (18%) on average—while this decreases for wealthier households.

The increase in non-farm enterprise workers among less wealthy households, including some with no previously active businesses, implies a reallocation of household labor to household enterprises following flood exposure. But the results show that only households with active businesses in 2010-11 increased their enterprise incomes after being exposed to flooding relative to households in non-flooded communities, indicating the productivity of new businesses is low. Increases in net non-farm enterprise incomes following flood exposure must therefore be due to intensification of existing non-farm activities on margins other than the count of household members involved, such as the hours of household member labor and investment in business assets and inventory.

These patterns in enterprise outcomes are not as evident using the survey-based measure of flood exposure. Similar to the MODIS-based results I find that the count of household members engaged in this activity increases by significantly more for households below the poverty line, and does not change for households with existing enterprises at baseline. Interestingly, I find that non-farm enterprise income increases by significantly more for households engaged in crop production at baseline, while the point estimate is negative but not statistically significant for households that were not engaged in crop production. This increase in enterprise income for crop-producing households is not associated with an increase the count of members working in household enterprises. This nevertheless implies a reallocation of resources to non-farm enterprises for households engaged in crop production following flood exposure.

### 2.6.1 Robustness

The findings for net crop income and net non-farm enterprise income are robust to a variety of alternative specifications, including adding time-varying community controls, replacing state-by-round with round or 1 degree cell-by-round fixed effects, replacing household with community fixed effects, dropping control communities more than 10km from any flooded community in the study period, and dropping the 2018-19 survey round (Figure B.1.6). The results are also similar if I drop the inverse propensity weights (IPW) and instead restrict the sample by dropping communities more than 10km from any flood realization during the study period or by matching communities by predicted 2012 flooding risk within strata of years of flood exposure between 2005 and 2011. Decreases in crop income with the MODIS-based measure are marginally significant in some of these alternative specifications, but there is no effect on non-farm enterprise income in any specification with the survey-based measure.

If we think of each flooding definition as measuring 'true' flood exposure with error, this suggests the possibility of using one flood measure to instrument for the other. With this IV approach I find significant effects of 2012 flood exposure on the probability a household experienced any food insecurity in the last 12 months using both measures, consistent with Figure B.1.5, but the effects on net non-farm enterprise income and net crop production income differ across measures and relative to the OLS estimates (Table B.2.7). When instrumenting MODIS-based flood exposure with survey-based flood exposure, the effect on non-farm enterprise income is no longer statistically significant (as in Figure 2.4) but I find a very large decrease in net crop production income. When instrumenting survey-based flood exposure with MODIS-based flood exposure, the effect on crop production income is no longer statistically significant but I find a very large increase in net non-farm enterprise income. Effectively, the pattern of IV results mirrors the OLS effects of the flooding definition I use as the instrument, but with larger magnitudes, consistent with the identifying variation being driven by the instrument.

Interpretation of these IV estimates is subject a weak instruments concern.[15] Overlap between the survey-based and MODIS-based definitions is limited. Communities identified as flooded by one definition are between 25-30 percentage points more likely to be identified as flooded by the other and the first stage Kleibergen-Paap Wald F-statistics are 10.8 and 12.3, just above the rule of thumb threshold for weak instruments. This emphasizes that the two flooding definitions appear to be capturing quite different flooding-related phenomena.

Finally, I find no consistent heterogeneity in effects by previous exposure to flooding, either at the household or community level, or by community survey reports of the share of

[15] Conditional on household and state-by-round fixed effects, the exclusion restrictions for both instruments should hold.

households affected by the 2012 floods. Effects are also not significantly different by the sex of the head of household and by whether the household has 5 or more members.[16]

### 2.6.2 Differences and agreement across flooding definitions

Taken together, the results indicate that flood exposure increases household efforts to diversify their incomes or to shift from relying on crop production to relying on non-farm enterprise. Households with members already active in non-farm enterprise activities at baseline grew these businesses following the 2012 flood exposure by more than similar households in non-flooded communities, without involving additional household members. Negative differential effects on the change in net crop income for these household engaged in non-farm enterprises implies that the increased emphasis on enterprise came at the cost of resources that might have been allocated to crop production. These households see a significantly larger increase in the share of household income from non-farm enterprise (Figure B.1.8 Panel E).

The MODIS-based flooding definition indicates that the shift from crop production to non-farm enterprise is concentrated entirely among less wealthy households (Figure B.1.8 Panels E and F). Households with below the median value of assets at baseline increase the non-farm enterprise share of income by around 10 percentage points and decrease the crop production share by an equivalent amount after flood exposure relative to similar households in non-flooded communities. Households with above-median asset value have no change in the income shares of these activities. I do not observe differences by wealth with the survey-based flooding measure.

Null effects on the household income Herfindahl-Hirschman Index—0.87 at baseline indicating most income comes from just one activity—according to both measures and across baseline characteristics indicate that any average shifts in income shares do not involve a more even diversification of household income sources. Instead, they result from a subset of households effectively flipping the shares coming from crop production and non-farm enterprise, since household income remains similarly highly concentrated post-flood exposure. Increased labor supply to non-farm activities following a flood would be consistent with several studies of effects of exposure to agricultural shocks on labor supply in the same and following year (e.g., Branco and Féres (2021); Macours et al. (2022); Mueller and Quisumbing (2011)).

I test whether impacts of flood exposure vary by agreement in the two flooding definitions. Table 2.4 shows results from estimating Equation 2.1 but including three $Flood \times Post$ terms capturing floods identified by both definitions, by the survey-based definition only, and by

[16]Results for these specific heterogeneity checks are available upon request.

the MODIS-based definition only. Differences in estimated effects are consistently largest when considering whether MODIS-based flood exposure coincides with survey-based flood exposure. Though the magnitudes differ, estimated effects of 2012 survey-based flooding are generally not statistically different by whether it coincides with MODIS-based flooding. This suggests that the nature of the flooding identified by the survey-based definition is more internally consistent while the MODIS-based measure captures some potentially less damaging floods that have different effects.

Table 2.4: Impacts of different 2012 flood exposure measures on household outcomes

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Net non-farm enterprise income (USD) | | Net crop production income (USD) | | Any food insecurity last 12 months | |
| $\beta_1$: Both flooding definitions × Post | 339.8 | 186.0 | -541.1 | -831.9 | 0.092*** | 0.070*** |
| | (278.7) | (268.5) | (367.4) | (567.4) | (0.023) | (0.024) |
| $\beta_2$: Survey-based flooding only × Post | 184.9 | 555.2 | -216.0 | 204.9 | 0.029 | 0.027 |
| | (222.4) | (338.4) | (132.6) | (523.7) | (0.019) | (0.028) |
| $\beta_3$: MODIS-based flooding only × Post | 590.8* | 805.9*** | -9.1 | 5.5 | -0.004 | 0.023 |
| | (301.7) | (276.4) | (151.2) | (141.0) | (0.006) | (0.029) |
| Observations | 10365 | 11054 | 10365 | 11054 | 10375 | 11068 |
| $p-$value, $\beta_1 = \beta_2$ | 0.625 | 0.297 | 0.398 | 0.301 | 0.020 | 0.147 |
| $p-$value, $\beta_1 = \beta_3$ | 0.462 | 0.064 | 0.093 | 0.149 | 0.002 | 0.141 |
| $p-$value, $\beta_2 = \beta_3$ | 0.252 | 0.515 | 0.220 | 0.677 | 0.256 | 0.891 |
| Baseline mean, non-flooded | 415.5 | 303.3 | 1119.1 | 928.8 | 0.175 | 0.192 |
| State-Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Household FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Inv. Prop. Weight | Survey | MODIS | Survey | MODIS | Survey | MODIS |

Note: This table shows the results from separate regressions of household outcomes on the interactions of dummies for different categories of 2012 flood exposure and a dummy for being observed after 2012. The excluded category is no 2012 flood exposure according to either the survey-based or MODIS-based definition. Survey-based flood exposure is defined based on survey reports for the community of residence in 2012, while MODIS-based exposure is defined as the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data. I include $p$-values for tests of the equality of coefficients. All incomes are in constant 2016 USD. Net income from household farm activities is calculated by taking the value of all household farm output (including output produced for own consumption) and subtracting expenditures on all purchased inputs over the prior agricultural year. Net income from household non-farm enterprise is calculating as total revenues minus total expenses over the prior 12 months. All regressions include household and state-by-year fixed effects. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012 using either the survey-based or MODIS-based definition. Standard errors are clustered at the level of the community of residence in 2012. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The point estimates for the effect of flooding on net household non-farm enterprise income are large and positive for all flood exposure measures but are only statistically significant for households in communities identified as flooded in 2012 by the MODIS-based definition alone. This pattern is not driven by the greater share of urban communities among those identified as flooded only by MODIS (Table 2.3). The increase in non-farm enterprise income is in fact concentrated among rural communities within this set. Significant effects in communities where no flooding was reported but MODIS identifies nearby flooding may imply that increased non-farm enterprise activity is most pronounced in areas where flooding was less severe. It also indicates that changes in non-farm enterprise income could result from whatever community characteristics lead flooding to be less damaging, including adaptation following prior flood exposure.

The decreases in net crop production income over time in flooded relative to non-flooded communities identified with the survey-based definition of the 2012 floods in Figure 2.4 are no longer significant when controlling for the overlap in exposure with the MODIS-based definition. The decreases are largest in magnitude for communities where both the surveys and MODIS identify flooding. If the alignment of these two definitions indicates flooding was more severe, this would indicate that household agricultural decisions are more strongly affected by more damaging floods. Table 2.3 also showed that this set of communities is relatively more agricultural and closer to bodies of water, meaning farm households might be more likely to updated their expectations about agricultural risk and productivity following a severe agricultural shock.

Estimated effects on household food insecurity align with these interpretations. Average increases in the probability of food insecurity are greatest and only statistically significant for communities where flooding in 2012 was identified according to both definitions. The smallest magnitudes are for communities where only MODIS identified flooding, consistent with the flooding detected in these areas not causing meaningful harm or damage to communities and households.

## 2.7 Conclusion

Identification of areas affected by the widespread and highly damaging floods in Nigeria in 2012 differs depending on the flooding definition used. The overlap in the communities identified as exposed to flooding according to both survey reports and MODIS satellite imagery is relatively small, and each definition identifies many communities as flooded that the other does not.

Flooding identified by MODIS only may indicate adaptation or reference dependence such that communities are not severely affected by the floods. It is important to include such flooding in an analysis of its economic impacts in order to capture how adaptation may mitigate potential effects. Flooding identified by survey only could be due to errors in reporting but also shows limitations in the types of flooding that satellite-based methods are able to detect. Relying on satellites may result in misclassifying some communities adversely affected by flood damages as not flooded, biasing estimates. A careful consideration of possible measurement issues is necessary for the interpretation of any studies of the effects of flood exposure.

The choice of flooding definition matters for estimates of their economic impacts. Results using the MODIS-based definition of the 2012 floods support a relatively simple narrative of flood exposure causing households to reduce their engagement in crop production in favor of non-farm enterprise on average, but especially for households already active in such businesses before 2012. Non-farm enterprise income increases relative to households in non-flooded communities, implying that flood exposure may promote exit from agriculture and potentially make households better off, though total income and measures of poverty or assets are not affected.

Results using the survey-based definition are more nuanced. Some farm households form new non-farm businesses while some households not cultivating crops in 2010-11 begin doing so, but neither group is productive and these new activities do not increase net incomes relative to changes in non-flooded communities. Households exposed to the 2012 floods also reduce cultivation of higher-value crops to focus more on staples, particularly cereals, increasing household labor to crop production while decreasing the total value of crop production. This results in significant decreases in net crop incomes relative to households in non-flooded communities, and negative but noisy effects on total income though again without affecting poverty or assets.

Increases in non-farm enterprise income are largely driven by areas where flooding was identified by MODIS only. If this mismatch in flooding identification is due to adaptation of communities to floods in these areas, this may indicate that prior adaptation involves transitioning out of agriculture putting these communities on different paths than non-flooded communities. This would imply that analyses of a single major flood event will fail to capture longer-term effects of changes in flood risk and resulting adaptation unless prior flooding was limited in frequency and severity.

Though studies such as this one analyzing effects individual flooding events may provide insight into short-term effects, future work on the long-term economic impacts of flood exposure should consider effects of cumulative flood exposure and changes in flood risk and exposure over longer time frames. Satellite-based measures will likely be the best source of data for such analysis, but researchers should seek additional data sources with which to check for possible types of flooding events that cannot be seen from space.

# Chapter 3

# Balancing work and childcare: Evidence from COVID-19 school closures and reopenings in Kenya

## 3.1 Introduction

The availability and cost of childcare have been shown to significantly affect adult labor supply in high-income countries, particularly for women. But there is less evidence on this relationship in low- and middle-income countries (LMICs), particularly in Sub-Saharan Africa (Halim et al., 2023). Yet, a historical perspective highlights the important role of increases in women's labor supply in driving economic development (Boserup et al., 2013). Understanding how childcare and adult labor supply interact is therefore crucial in these settings.

Sub-Saharan African countries differ from high-income countries in many ways relevant to this question. Female labor force participation is high but concentrated in informal activities (ILO, 2017). Family farm or non-farm enterprise work is widespread, and may be more accommodating of childcare needs than wage employment. Women are more likely than men to be engaged in such own-account or contributing family work—less likely to have access to social protection and more exposed to economic cycles—and particularly so if they have children (Lo Bue et al., 2021). Households have more children but also more adults on average (UN, 2020). Formal early childhood care availability is increasing, but from a low base and there are concerns around quality and cost (Samman et al., 2016). Critically, older children play an important role in household productive activities (Kielland and Tovo, 2006; Levison and Moe, 1998) and sibling childcare (Jakiela et al., 2020; Qureshi, 2018), meaning they are not just childcare recipients within the household. It is not clear *a priori* how this complex and interdependent intra-household allocation of activities and decision-making affects the aggregate relationship between household childcare arrangements and adult labor supply decisions.

An important factor influencing household childcare needs is the availability of low- or no-cost schooling. In 2020, countries around the world closed schools in response to the COVID-19 pandemic. This paper leverages school closure policies in Kenya as exogenous shocks to provide empirical estimates of the impact of childcare responsibilities on adult labor supply and intra-household allocation of productive activities in a LMIC setting. Kenya closed all schools nationwide after its first COVID-19 cases in March 2020, partially reopened schools for specific grades in October 2020, and fully reopened for all grades in January 2021. Household childcare needs increase during school closures creating trade-offs for adults' time allocation across childcare, work within and outside the homestead, and other activities.

We exploit quasi-random variation in when children enrolled in different grades were eligible to return to school to implement a difference-in-differences analysis comparing changes in labor supply after the October partial reopening for adults in households with children in grades 4 or 8—eligible to return (99% did)—against those with children in adjacent grades. Our data consists of a nationally representative bi-monthly panel from the Kenya COVID-19 Rapid Response Phone Survey, including information on household composition, adult labor, childcare, and child labor. The timing of surveys is ideally suited for our study, with two rounds conducted when schools were fully closed, one round exactly after the partial reopening, and further rounds after all schools reopened.

Labor supply effects of the partial reopening are concentrated on the intensive margin: weekly work hours increase by 4.7 (29%) after the partial reopening for adults with a child eligible to return to school, driven by a 33% increase in hours worked in household agriculture. We find no effects on wage employment hours, household enterprise hours, or on the extensive margin of employment in the weeks following the reopening.[1] Household agriculture is likely more flexible in allowing adults to adjust working hours after a change in the childcare burden; participation in household agriculture was also less affected by the pandemic. In line with this, average labor impacts are driven by agricultural households and adults that were able to continue working during the school closures period. Less wealthy households—based on an index of housing and assets—increased work hours by significantly more than wealthier households. Poorer households are more likely to engage in agriculture and may have fewer resources to deal with shocks to childcare needs and child labor. Consistent with the strongly increasing returns to childcare observed in our data, the partial reopening does

[1] Analysis of longer-term impacts is complicated by the fact that schools fully reopened in January 2021, meaning our comparison group also becomes 'treated.'

not affect labor supply in households with both children eligible to return to school *and* children in adjacent grades who do not return—changes in the childcare burden and in child labor are muted in these households.

Surprisingly, the average impacts of the reopening are not significantly different by sex, contrasting with evidence on pandemic labor supply changes from high-income contexts (see e.g., Alon et al. (2022); Amuedo-Dorantes et al. (2023); Collins et al. (2021); Hansen et al. (2022); Heggeness (2020)) and expectations based on women's role as primary caregivers in most Kenyan households. One reason for the lack of difference is that in our sample, both sexes contribute substantially to childcare. Women report spending 40 hours per week on childcare on average before the pandemic while men report 25, and both increased childcare hours during school closures by approximately 15 hours per week. The similar average labor supply increases for men and women after schools reopen are also driven by different mechanisms that may be particular to lower- and middle-income settings.

First, school-age children in Kenya (and many LMICs) are both *receivers* of childcare and *contributors* of childcare to younger siblings. Moreover, men and women engage differently in the care of children in different age groups. The combination of these factors leads to gendered responses. Men's childcare engagement focuses relatively more on school-aged children. Estimated effects on childcare hours are noisy due to limitations in the childcare measure, but suggest that men benefit from more consistent reductions in childcare hours when a student returns to school, and are able to increase labor supply across all activities. Women also care for school-aged children but in addition have the primary care responsibility for children below school-age, which is supplemented by older siblings. While women benefit more than men from childcare reductions after the reopening when there are no young children present, in other households they lose childcare support they had been receiving from the returning student. In households with below-school-age children, women actually *increase* their childcare hours when older children return to school. The results suggests women in these households specialize more in childcare while men reduce childcare engagement. Women shift from non-agricultural work into household agriculture (more easily combined with care of younger children), and their total labor supply does not increase leading to null average effects of the partial reopening in households with children below school-age. These results highlight the importance of sibling childcare in this setting, and demonstrate how women with young children could increase their labor supply with access to similar childcare while their older children are in school.

Second, school-age children in many households contribute to productive activities, including household agriculture. The partial reopening coincided with the main harvest season for most of Kenya, and adult hours in household agriculture increase by more after schools reopen in households where children were engaged in household agricultural labor during school closures. Though estimated decreases in total child agricultural labor after a child returns to school are not statistically significant (perhaps due to increased labor from other children), the evidence indicates that part of the increase in adult agricultural hours is driven by efforts making up for reduced child labor. In households where children work in agriculture, women pick up most of the slack, and—mirroring what we find for households with below-school-age children—substitute labor supply away from potentially more productive wage employment and self-employment in household enterprise. One reason for this is that work on the household farm is more compatible with contemporaneous childcare requirements for younger children. Together, these mechanisms indicate that even if overall labor supply responses to the reopening are similar for men and women, childcare still represents a constraint that affects women differently from men.

This paper makes three main contributions. First, we consider how formal childcare for school-age children (through schooling) affects households through changes in both childcare burdens and availability of child labor in a context where labor in home production is common, providing evidence on an under-explored dimension of the relationship between childcare and adult labor supply (e.g., Browning (1992); Connelly (1992); Ribar (1992)). Moreover, we highlight how differing gender roles in the care of children of different ages lead to potentially different labor supply responses by men and women to childcare shocks affecting only some children. The current literature largely studies childcare for below-school-age children and treats children solely as childcare recipients, while focusing on settings dominated by wage employment (Morrissey, 2017). These characteristics do not generalize to many LMIC contexts, where the impact of the more complex patterns of intra-household allocation of activities on childcare arrangements and adult labor supply is not well understood. Among studies of childcare and labor supply in Africa (Bjorvatn et al., 2022; Clark et al., 2019; Delecourt and Fitzpatrick, 2021; Heath, 2017; Lokshin et al., 2000; Martinez et al., 2012; Quisumbing et al., 2007), causal identification is limited,[2] only two include rural areas, and none consider the role of children as household labor providers. This paper documents childcare patterns using a nationally-representative sample of Kenyan households with mobile phones, where most adults are engaged in household farm and non-farm enterprise rather than wage work. We estimate causal impacts of a change in household childcare needs and child labor availability using a natural experiment affecting formal care provision for school-age children to shed light on the role of child household labor in the relationship between childcare needs and adult labor supply.

Second, we demonstrate that a change in children's labor availability around harvest time due to the school calendar impacts parents' labor supply. This result is consistent with

[2] In a systematic review of the literature identifying causal impacts of childcare on mothers' labor market outcomes in LMICs, Halim et al. (2023) identify 22 studies, but just 6 are from outside Latin America and only one reports on an African country.

evidence on the importance of the timing of school breaks for both child and household outcomes in LMICs (Admassie, 2003; Allen, 2024; Duryea and Arends-Kuenning, 2003; Ito et al., 2020; Kadzamira and Rose, 2003). Much of this literature has focused on impacts of school calendar policy on outcomes for children. We show that adults in Kenya make up for reductions in child labor when a student returns to school by increasing their time allocation to household agriculture, sometimes substituting away from other occupations.[3] More generally, we contribute to a broad literature on child labor in low-income settings (see e.g., Basu and Van (1998); Beegle et al. (2009); Bharadwaj et al. (2020); Udry (2006)), further demonstrating how children can play an important role in household agricultural production in particular. The substitution between child and adult labor supply at peak agricultural times has important implications for both children's education and household agricultural production as Kenya's school calendar shifts in the years following the COVID-19 pandemic to return back to its pre-pandemic schedule.

Finally, we also contribute to understanding labor impacts of pandemics and pandemic-related policies—school closures in particular. Many studies have analyzed the gendered effects of the COVID-19 pandemic on childcare and employment (see e.g., Alon et al. (2022); Amuedo-Dorantes et al. (2023); Collins et al. (2021); Del Boca et al. (2020); Furman et al. (2021); Giurge et al. (2021); Hansen et al. (2022); Heggeness (2020); Liu et al. (2021); Prados and Zamarro (2021); Zamarro and Prados (2021)). These primarily report on high-income settings and fairly consistently find that increased childcare burdens contributed to greater adverse labor effects for mothers during the pandemic. In particular, Prados and Zamarro (2021) find that school closures in the United States increased childcare responsibilities for mothers relative to fathers and that this persisted after schools began to reopen, while transitions out of employment for working mothers appear more persistent than for working fathers. Descriptive evidence on the gendered impacts of COVID-19 in LMICs (Casale and Posel, 2020; Chauhan, 2021; Deshpande, 2020; Grantham et al., 2021; Kugler et al., 2021; Torres et al., 2021)[4] suggests that women increased domestic work and reduced their labor supply more than men, but causal estimates demonstrating the role of childcare in labor supply changes are lacking.[5] We show that time spent on childcare increased by about the same amount for both women and men in Kenya during the COVID-19 school closures in 2020, and returned to pre-pandemic levels after schools reopened in 2021. Contrary to evidence from other contexts, we find that the partial reopening of schools in October 2020

[3] Allen (2024) reports that households in Malawi respond to harvest-period overlap with the school calendar by increasing expenditure on hired labor, which is not measured in the RRPS.

[4] Wenham et al. (2020) report on impacts of the Ebola crisis in Sierra Leone and Liberia for women and men.

[5] Ma et al. (2020) analyzes the impacts of school closures on parents' labor in Shaanxi province, China, but we are not aware of similar studies in an African country.

significantly increased adults' labor supply with no significant difference by sex, driven by different, but offsetting, mechanisms for women vs. men. Building on this result, we calculate that school closures account for at least 40% of the initial drop in average working hours during the COVID-19 pandemic and resulted in 2 billion fewer hours of work across Kenya in 2020. At the average hourly earnings in the data, this represents a cost of USD 2.4 billion (2.6% of 2019 GDP).

That the policy responses to the COVID-19 pandemic tended to decrease labor supply more for women than men on average globally re-focused attention on how the burden of balancing childcare needs with labor supply falls primarily on women. It has long been known that reducing household childcare burdens can increase female labor force participation, and the pandemic has shown how the inverse is also true. By further demonstrating how older siblings play an important role in household childcare and household labor in Kenya, this paper emphasizes how policies improving childcare availability for younger children —such that their older children are not parents' main alternatives—could broadly increase adults' labor participation in an African LMIC context. We show that benefits would extend to rural agricultural households, and particularly to women as the primary caregivers for children below school-age.

The remainder of the paper is organized as follows. In Section 3.2 we describe the context around COVID-19 and school closures in Kenya, discuss the data used in this paper, and present descriptive statistics on childcare during the pandemic in Kenya. Section 3.3 presents the empirical approach. Section 3.4 presents overall results on the impacts of school reopenings, while Section 3.4.1 analyzes heterogeneity and Section 3.4.2 explores mechanisms. Section 3.5 discusses the findings and Section 3.6 concludes.

## 3.2 Context and data

This section summarizes Kenyan COVID-19 school closure policies, the data we use to analyze their impacts on labor supply, and information on childcare arrangements.

### 3.2.1 Kenyan school system and COVID-19 closure policies

Public primary and secondary education in Kenya is free for all children starting around age 6, and education is compulsory for the first nine years. Pre-school has also become broadly available, particularly for age 5 children. Academic years in Kenya begin in January and end in late October, and consist of three terms.

Schools in Kenya closed on March 16, 2020, days after the first reported COVID-19 cases, as part of a broad set of national restrictions to reduce the risk of disease transmission. The rest of academic Term 1 was cancelled. Figure C.1.1 shows a timeline of school closures and reopenings, other key pandemic-related policy changes, and weekly confirmed COVID-19 cases in Kenya.[6] Pandemic school closure policy in Kenya was decided nationally. Top-down changes in policies thus represent exogenous shocks to households, unrelated to local economic or health conditions.

On September 15, the Ministry of Education released guidelines for safe reopening of schools, but the timing and nature of reopening remained uncertain until October 6, when the Ministry announced that students in grades 8 and 12—those sitting national exams—along with students in grade 4 should return to school on October 12 for Term 2 of 2020. This announcement was presented in the media as "a shocking move that caught parents and candidates off guard" (The Star, 2020). On November 4, the President announced that schools would reopen fully for all students on January 4, 2021 to complete the 2020 school year. There were no additional fees incurred when schools reopened as parents had already paid fees for the 2020 school year, but some parents may have been asked to pay outstanding bills from before the school closures and others may have paid for new materials or extra lessons.[7]

Students in grades 4, 8, and 12 returned for Term 3 from January-March 2021 while all other students returned for Term 2; their Term 3 was shifted to May-July 2021. Grade 8 and 12 students sat national exams in March-April. 2021 Term 1 for all students began in late July 2021. Terms and breaks for the 2021-2023 academic calendars were shortened to allow a gradual return to the standard pre-pandemic term schedule (running from January-October) in time for the 2024 academic year.

We focus on the impacts of the partial school reopening for several reasons. First, unlike initial school closures, the partial reopening did not coincide with other pandemic-related policies (see Appendix B.2), allowing for cleaner identification. Second, we can exploit discontinuities in the timing that children enrolled in different grades were eligible to return to school to isolate the effect of the childcare shock. Further, because households vary in whether the students eligible for the partial reopening are net contributors or receivers of childcare—depending on the presence of younger siblings—this shock sheds light on the importance of sibling-provided childcare.

### 3.2.2 Data

Data come from the Kenya COVID-19 Rapid Response Phone Survey (RRPS) panel, collected by the World Bank in collaboration with the Kenya National Bureau of Statistics and

[6] An overview of specific pandemic-related policies is presented in Appendix B.2.

[7] There are some additional costs associated with national exams, but these were paid by the government for all candidates at the time of the exams in Spring 2021.

the University of California at Berkeley (Pape, 2021).[8] The main sample ($\sim 80\%$) is drawn from the nationally-representative Kenya Integrated Household Budget Survey conducted in 2015-2016, and this sample is supplemented by random digit dialing. The sample is intended to be representative of the population of Kenya using cell phones—80% of households nationally report owning a mobile phone, and these have better socioeconomic conditions on average than households that do not (Pape et al., 2021). We use data from the first four survey rounds, covering May 2020-March 2021. In addition, we construct measures for February 2020, before the first COVID-19 cases in Kenya, using recall questions from the first round.

The outcomes of interest are measures of labor supply.[9] The extensive margin is measured by participation in the last 7 days in three activities: employed/wage labor, household non-farm enterprise, and household agriculture. The intensive margin is captured using hours of work by activity in the last 7 days; an individual not working in a given activity is coded as working 0 hours.

The survey also includes data on total child hours spent working in household agriculture. Child agricultural labor is reported by 42.7% of agricultural households.

Household roster data includes the age and sex of all household members, as well as school enrollment information for all children. Information on what grades children were enrolled in prior to the initial school closures allows us to identify households affected by the partial reopening. Nearly 99% of eligible students are reported to have returned to school.[10]

We define 'treatment' households as those with children enrolled in grades 4 or 8 prior to the pandemic (eligible for the partial reopening) while 'control' households have children in grades 3, 5, 6, 7, or 9, but not in grade 4 or 8.[11] We separate 'mixed' households with children in both grade groups from 'treatment' households as they might experience different effects when not all children in the relevant grade range return to school. The main analysis sample includes 335 treatment, 361 mixed, and 948 control households.

Finally, the data include questions on household childcare arrangements, including childcare hours though this is only reported for the survey respondent over the time period of the partial reopening. The survey asks about time spent on childcare in the last 7 days, but

[8] See Appendix B.3 for more detail.

[9] We use the term labor 'supply' to refer to equilibrium outcomes, acknowledging that individuals may have been willing to supply additional labor but faced limited demand.

[10] A survey of 3,000 grade 8 students in Busia County, Kenya similarly shows that 97% reported back to school after the partial reopening (Bonds, 2021). Across all grades, 97% of previously enrolled students in the RRPS returned to school after the full reopening in January 2021.

[11] Few households report any children in grade 12. Results are robust to varying the grades included in the definition of control households, and to including grade 12 students in the treatment definition and grade 10 and 11 students in the control definition (Table C.5.3).

does not distinguish between time actively spent caring for a child and time spent on other activities while responsible for a child.[12] In households with older children in particular, much of what is reported as childcare hours likely falls into the latter category. We topcode reported childcare hours at 140, or 20 hours a day.[13]

### 3.2.3 Childcare arrangements

At least 93% of children at each age from 6-16 in the RRPS are reported to have been enrolled in school in February 2020. In addition, over 90% of children age 5 were also enrolled in preschool. After the March closures these children were all home requiring care and supervision during the working day, representing a large and unexpected shock to household childcare needs. Children primarily stayed at home with a parent during the closures (Figure C.1.2), including in situations where parents were simultaneously working. Almost no households report their children spending time with childcare providers outside the home or with a maid/domestic helper at home, and this does not vary by rural/urban setting or change as pandemic restrictions were relaxed and case numbers fluctuated. Adults with schoolchildren at home will have faced trade-offs in their allocation of time across childcare, work in different sectors, and other activities given a limited time budget to accommodate increased childcare burdens.

Figure 3.1 Panel A shows how hours of childcare from different providers (excluding schools) vary with the number of household children among analysis sample households (with at least one child in any grade from 3 to 9). We present data from the January-March 2021 survey round after schools fully reopened, as previous survey rounds only include data on the respondent's childcare hours.

Respondents in the analysis sample provide 30-35 hours of childcare on average per week,[14] and this increases very little after the first child. This is likely due to a combination of how childcare is measured in the data and different dynamics of childcare by child age. The survey measure of childcare includes 'passive' childcare when the adult is responsible for the child but not actively providing care, meaning that adults responsible for their children for all (or most) of the day will report spending the same number of hours on childcare regardless of their number of children. But the median for reported childcare hours over

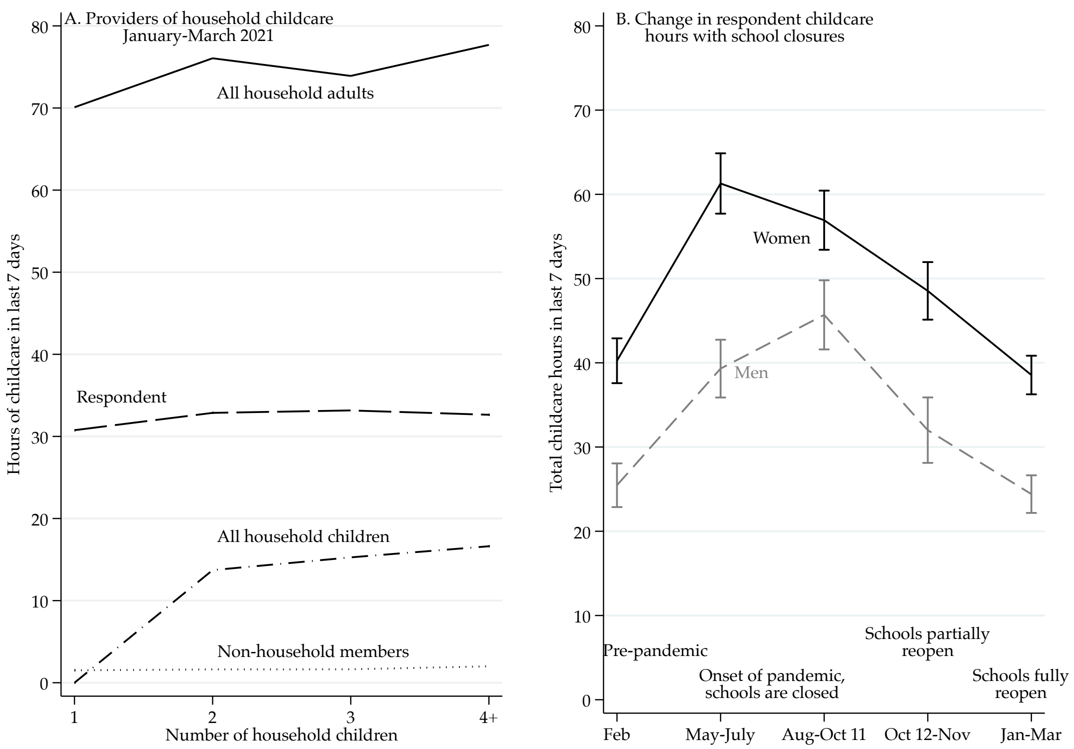
[12]The survey explicitly asks respondents to consider time spent providing childcare alongside other activities, asking "In the past 7 days, how many hours have you spent doing childcare for your household, even if it overlapped with other tasks?"

[13]This recognizes that, for young children especially, care responsibilities spill into sleeping hours. Results are qualitatively the same when topcoding childcare hours at 112, or 16 hours a day.

[14]Respondents provide 46% of total household childcare on average, which aligns with expectations of parents being the primary care providers, with both parents involved.

the last 7 days is 20 hours in round 4, suggesting overly broad measurement is not solely responsible for the limited increase in childcare hours as the number of children increases. Households with more children may benefit from some economies of scale in childcare on the extensive margin, even if more children require more active attention during a given hour of childcare on the intensive margin. More generally, with additional children some will be older and thus require less care or supervision, and some will be net childcare providers.

Figure 3.1: Childcare hours in the last 7 days among analysis sample households, by provider of care and school closure status




Note: The figures show mean childcare hours in the last 7 days among analysis sample households (with at least one child in any grade from 3 to 9), by number of household children (ages 0-17) and by time period. Patterns are very similar when considering all households with at least one school-age child (age 5-17) (Figure C.5.1).
Panel A presents data from RRPS round 4 (January-March 2021) which asks about childcare hours for each household adult, for all children in total, and for all non-household members in total. Previous rounds only ask about childcare hours for the respondent. The hours for 'all household adults' include the respondent's hours.
Panel B presents data for female (black) and male (gray) *respondent* childcare hours over time as school closure policies changed. Data on childcare hours before and during the school closures period for other care providers are not available.

The survey rounds during school closures do not ask about childcare provided by siblings, but round 4 after the full reopening asks about total sibling-provided childcare hours and for which children provided care. Consistent with prior studies in Kenya (Jakiela et al., 2020) and other LMICs (Levison and Moe, 1998; Qureshi, 2018), siblings in the sample play an important role in childcare. In households with at least 2 children 59% of children age 5-17 provided childcare to siblings in the last 7 days, for 15-20 hours on average in total, demonstrating how childcare burdens can also fall on household children. For children in grades 3-9 after schools fully reopened, 66% provided care for a sibling in the last 7 days for an average of 6.4 hours each. Contrary to other studies and to norms around childcare provision, reported engagement in sibling childcare (at least on the extensive margin) does not vary by sex of the child in this age range.[15]

Sibling care hours may overlap with adult care hours, if for example siblings are providing childcare while an adult is present with overall care responsibility. To the extent that this is true, changes in sibling care provision would affect the intensity of adult childcare provision (which we do not measure) more than the number of adult childcare hours. Sibling childcare provision as reported by a household adult may primarily represent 'active' childcare since that would be most easily observed and recalled by the respondent, and since they might report supervised sibling care-giving under their own childcare time rather than the child's. If this is the case, total hours of sibling childcare may be higher than what is reported in the survey. Sibling childcare hours were likely higher during school closures as school-age children were at home, but we only measure sibling childcare after the full reopening. The importance of sibling childcare in this setting suggests that a student returning to school might increase rather than decrease parents' childcare burden, in situations where they were net childcare providers while at home.

Other adults besides the parents are present in 37% of households with children, and on average provide around 10 hours per week of childcare. Non-household members provide very little childcare on average—86% of households with children report 0 hours of care from non-household members in the last 7 days. Reported use of childcare providers outside the home does not increase significantly over the rest of 2021 as reported in later rounds of the RRPS, suggesting this is not driven by pandemic concerns or restrictions. While childcare availability has been increasing in Kenya (particularly in urban areas), affordability remains a challenge for most households (Clark et al., 2021; Murungi, 2013). Such sources of childcare would therefore not have provided much or any relief to increased childcare burdens during the school closures period (as shown also in Figure C.1.2).

Figure 3.1 Panel B shows that while female respondents provide around 15-20 more

[15] Girl children may provide more childcare hours than boys despite similar engagement in childcare, but the survey does not measure individual childcare hours for children.

childcare hours than men outside the school closures period, men still contribute around 25 hours on average. This contrasts with the image of fathers in African countries as primarily providing economic support and little childcare, but is consistent with recent evidence (Clark et al., 2015; Kah, 2012). We note, however, that the nature of care provided in these childcare hours might vary by adult sex, given the broad definition of childcare hours in the data.

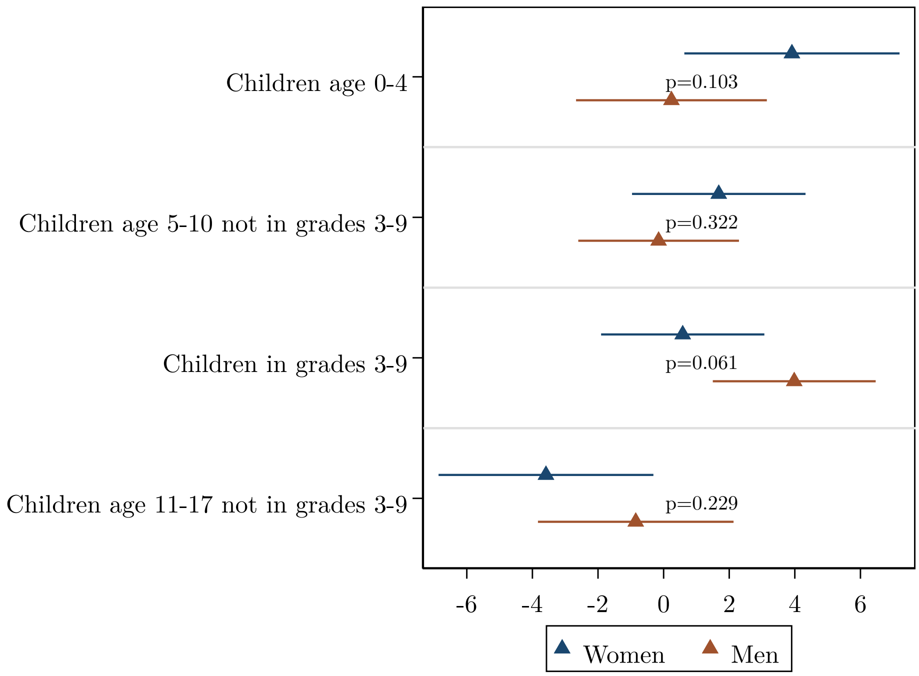
Childcare hours increased significantly for both women and men during the school closures period, with a larger initial increase for women before the gender gap returned to pre-pandemic levels. This pattern is similar to findings of increases in domestic work during the pandemic for both men and women, but larger for women, in India (Deshpande, 2020), South Africa (Casale and Posel, 2020), and many higher-income countries (see e.g., Andrew et al. (2020); Del Boca et al. (2020); Farré et al. (2020); İlkkaracan and Memiş (2021)).

Childcare hours began to drop in October-November 2020, coinciding with the partial reopening of schools. In survey rounds 1 and 2 (May-September), respondents were asked to report only their own childcare hours. In survey rounds 3 and 4 (October-March), they were also asked to report childcare hours for other household adults. This change may explain part of the drop observed from August-October 11 to October 12-November, if it led respondents to reassess downward their own childcare hours in the context of total adult childcare hours. After schools fully reopened, respondent childcare hours returned almost exactly to pre-pandemic levels from the year before.

Figure 3.2 illustrates how the composition of household children affects respondent childcare hours by sex during the school closures period, among analysis sample households. Although additional children among these households do not significantly increase childcare hours on average (Figure 3.1 Panel A), there is considerable heterogeneity. First, childcare responsibilities are gendered. As a baseline, in analysis households with one child (necessarily in grade 3-9), women report spending 53 hours on childcare in the last 7 days compared to 33 for men. But the care of additional children is also gendered. Women are relatively more responsible for the care of young children, and particularly children age 0-4. We consider this set of 'below-school-age' children in particular as over 90% of children age 5 in the RRPS were enrolled in pre-school before the pandemic. On the other hand, men are relatively more engaged in care of older children, and particularly for children enrolled in grades 3-9 before the pandemic. Second, the figure highlights older siblings' role as childcare providers. Having additional children ages 11-17 not in grades 3-9 (either older or not enrolled in school before the pandemic) significantly decreases women's childcare hours. Additional children in grades 3-9 are neutral for women's childcare: they either require no additional care time beyond what is already being provided, or they offset it by also providing care themselves.

These patterns suggest there may be important heterogeneity in how a child in grades 3-9 returning to school would affect childcare burdens, by both the sex of the respondent and the presence of other children of different ages in the household. Impacts for women would be

Figure 3.2: Change in childcare hours during school closures with additional children, female and male analysis sample respondents




Note: The figures show coefficients and 95% confidence intervals for the correlation between an additional child and respondent childcare hours, by respondent sex and child age category, from a single regression. $p$-values are for tests of difference in coefficients by respondent sex. Effects are relative to mean childcare hours of 53.5 in the last 7 days for women and 32.7 for men with one child in grade 3-9. Data are for May-November 2020 while schools were closed.

limited in households with other children, though men in these households would experience a reduction in childcare hours. Women would primarily benefit from a child returning to school if they are not caring for any other children. We explore this heterogeneity in our analysis of mechanisms for impacts of the partial school reopening on parents' labor supply outcomes.

## 3.3 Empirical approach

We identify the effect of partial school reopenings through a difference-in-differences analysis comparing outcomes before and after the reopening between households with and without children who are eligible to return to school.[16] We estimate two-way fixed effects regressions of the form

$$y_{iht} = \alpha + \beta_1 \cdot Post_t \times Treat_h + \beta_2 \cdot Post_t \times Mixed_h + \mu_h + \tau_t + \epsilon_{iht} \tag{3.1}$$

where $y_{iht}$ are outcomes for adults (age 18-64) $i$ in household $h$ at time $t$, or household-level outcomes. In some cases, outcomes are only available for the household respondent. $Post_t$ is an indicator for observations after the partial reopening on 12 October 2020. We include observations from May-November, omitting data from after schools fully reopened.[17] $Treat_h$ is an indicator for whether all household children in grades 3-9 were eligible to return to school. $Mixed_h$ is an indicator for whether the household has both eligible and ineligible children in this grade range. The omitted reference category is control households with children in this grade range but none eligible for the partial reopening. Household fixed effects $\mu_h$ absorb time invariant characteristics of households which may affect labor supply outcomes. Month fixed effects $\tau_t$ control for common shocks affecting households over time. We test robustness of the main results to specifications with different fixed effects and additional time-varying household- and individual-level controls and find that the coefficients generally remain stable (see Appendix B.5). We cluster standard errors at the household level.

Our empirical strategy exploits quasi-random discontinuities in which households are affected by the partial reopening. We compare households with children eligible to return to school (i.e., those in grades 4 and 8) to control group households with children in adjacent grades (i.e., those in grades 3, 5, 6, 7 and 9).[18] Causal identification is based on the argument that unobserved factors that could affect labor supply outcomes are continuous around the thresholds of children being in adjacent grades, and did not differentially affect treated and control households around the time of schools reopening.

Respondent and household characteristics are balanced across treatment and control households during the school closures period (Table C.1.1).[19] Mixed households with children in both treated and adjacent grades have more children: 4.1 compared to 2.9 children in

[16]Since schools partially reopened all at once, our two-way fixed effects estimator should not suffer from the negative weighting issues that arise in staggered difference-in-difference designs (see e.g. Callaway and Sant'Anna (2021); Goodman-Bacon (2021)).

[17]We do not show the estimated coefficients for *Post* as they are absorbed by the month fixed effects, where we include separate dummies for October 1-11 and 12-31.

[18]Results are robust to various modifications in which grades are included in the control definition (Table C.5.3), including focusing on just the immediately adjacent grades for the control group.

[19]The household head and respondent are slightly older in treatment households. Treatment respondents are also 5% more likely to have been working in February 2020 and to have worked in a household non-farm enterprise during the school closures period.

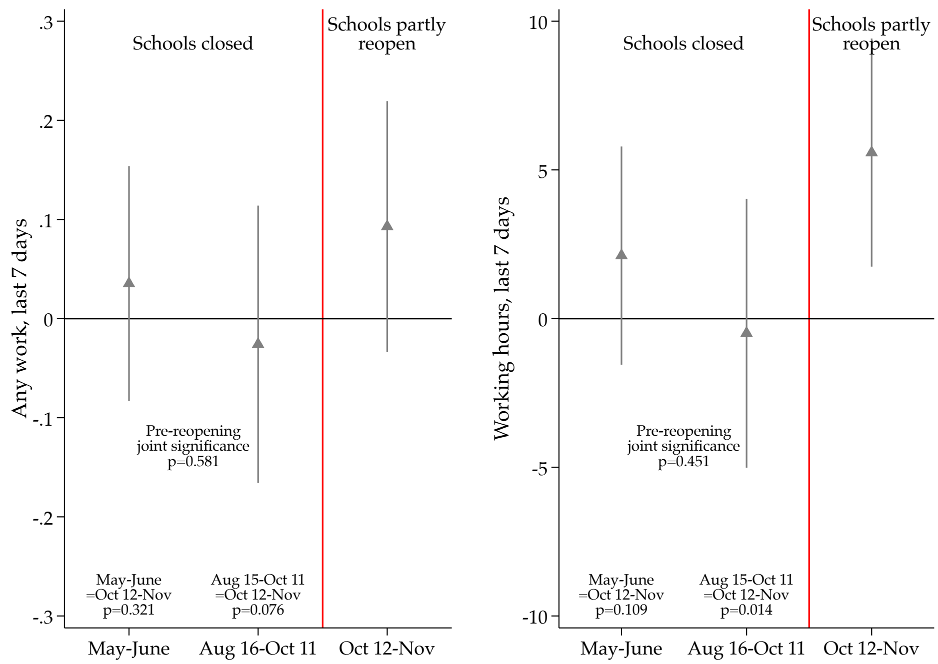
control households and 2.8 in treatment households. This difference is mechanical, as mixed households must have at least one child in both the treatment and control grades. Mixed households have 2.6 children in grades 3-9 on average, compared to 1.4 for control households and 1.2 for treatment households, accounting for the entire difference. Respondents in mixed households are more likely to be married and have lower levels of education than control of treatment households, indicating some differences correlated with having multiple children in this grade range. The difference in household composition and the fact that not all children in the target grade range in mixed households return to school leads to predictions of muted impacts of reopenings on this group of households. We therefore present only the main results for mixed households, and focus on fully treated vs. control households only for mechanisms. The main results are robust to dropping mixed households from the analysis (Table C.5.3), and the results on mechanisms are similar when including mixed households (Tables C.5.6 and C.5.7).

Beyond being similar in levels, work hours trend almost identically for adults in treatment and control households from February to early October 2020. Differences emerge following the partial reopening but are eliminated after schools fully reopen, when all households become 'treated' (Figure C.1.3). Figure 3.3 shows that there are no significant differences between treatment and control adults in overall labor supply in the periods when schools were fully closed, There is also no evidence of anticipation effects in the period from August 16 to 11 October.[20] The patterns are similar when considering women and men separately (Figure C.1.4) and when looking at work in different sectors. We can also reject any differences in the school closures period for respondent childcare hours and for total child hours in household agriculture. Taken together, these patterns support the validity of the parallel trends assumption in our setting.

To test for heterogeneous impacts of school reopenings, and to investigate underlying mechanisms, we interact all regressors in Equation 3.1 (except household fixed effects) with household or individual characteristics, corresponding to stacking separate regressions for each group. Moreover, we separately estimate treatment effects for households with a child in either grade 4 or grade 8 returning to school to test for differential impacts by the age of the child returning to school.

[20]This is not surprising, as although there were indications from mid-September that schools could begin reopening in October, the specific timing and the partial nature of reopening was not announced until the week before students were invited to return to school.

Figure 3.3: Impact of treatment on labor participation in the last 7 days, by time period




Note: The figures show estimated coefficients and 95% confidence intervals for the interaction between *Treat* and time period from Equation 3.1, where *Post* is replaced with time period dummies, which also enter separately into the equation. Time periods prior to the partial school reopening are constructed to have roughly equal sample sizes. The reference period is July-August 15, while schools were closed and before the partial reopening was announced. The red bar indicates the timing of Kenya's partial school reopening. Outcomes are any work engagement (left) and total work hours (right) in the 7 days prior to the interview. Treatment households have a child enrolled in grades 4 or 8 eligible to return to school on October 12, and control households have a child enrolled in grades 3, 5, 6, 7, or 9. We do not show coefficients for mixed households with children in both grade groups. $p$-values for the test of joint significance of the pre-reopening coefficients and for tests of equality between the pre-reopening coefficient and the post-reopening coefficient are shown.

## 3.4 Results

Table 3.1 presents results for the impacts of partial reopening on labor supply. We include all adults aged 18-64 present in the household at the time of the survey. Results are similar if we include only survey respondents, whom we focus on in our analysis of mechanisms (Table C.5.1). Among control households, 59% of adults were working during the school closures period. Mean work hours of 16.5 reflect that many workers were not working 'full-time'. Work is concentrated in household agriculture, despite 46% of the sample being classified as 'urban'—the categorization of urban locations in the RRPS includes city peripheries where agriculture remains common.

Table 3.1: Impacts of partial school reopening on adult labor supply

| | N | Control Mean (SD) | Post (SE) | Post x Treat (SE) | Post x Mixed (SE) |
|---|---|---|---|---|---|
| Engaged in any work in last 7 days | 8717 | 0.588 | 0.010 | 0.035 | 0.028 |
| | | (0.492) | (0.042) | (0.036) | (0.034) |
| Engaged in wage employment in last 7 days | 8717 | 0.062 | 0.002 | 0.004 | -0.014 |
| | | (0.241) | (0.018) | (0.012) | (0.012) |
| Engaged in HH agriculture in last 7 days | 8717 | 0.511 | 0.025 | 0.033 | 0.015 |
| | | (0.500) | (0.038) | (0.037) | (0.032) |
| Engaged in HH non-ag enterprise in last 7 days | 8717 | 0.073 | -0.008 | 0.008 | 0.020 |
| | | (0.260) | (0.019) | (0.017) | (0.015) |
| Total work hours, last 7 days | 8717 | 16.504 | 0.438 | 3.410** | -0.830 |
| | | (20.139) | (2.047) | (1.526) | (1.507) |
| Wage hours, last 7 days | 8717 | 2.077 | 0.364 | -0.080 | -0.686 |
| | | (9.934) | (0.780) | (0.562) | (0.587) |
| Ag hours, last 7 days | 8717 | 11.921 | 0.588 | 2.888** | -0.394 |
| | | (15.465) | (1.552) | (1.261) | (1.248) |
| Enterprise hours, last 7 days | 8717 | 2.562 | -0.334 | 0.603 | 0.073 |
| | | (10.669) | (0.734) | (0.649) | (0.657) |

Note: This table presents estimates of Equation 3.1 for individual labor supply. Individuals not working in a given sector are coded as working 0 hours. From left to right, the columns show the dependent variable, number of observations, the control mean prior to the partial reopening, and the impacts of being in the partial reopening period for treatment households (Post x Treat) and mixed households (Post x Mixed). Impacts for control households are absorbed by month fixed effects. Control households have a child in grades 3, 5, 6, 7, or 9, treatment households have a child in grades 4 or 8, and mixed households have both. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. Standard errors are clustered at the household level. Data include observations for adults age 18-64 from May to November 2020. Significant treatment impacts on total and agricultural work hours are robust to multiple testing adjustment using FDR q-values.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We find sizeable and borderline significant effects of treatment on the extensive margin of labor supply. Households with children returning to school are 8.3 percentage points (or 14%) more likely to engage in any work in the last 7 days ($p = 0.105$) relative to control households after the partial reopening. Adult labor supply responses to schools partly reopening are even larger on the intensive margin. Work hours in the last 7 days increase by 4.7 (28.8%) relative to adults in control households ($p < 0.001$). This is not sufficient to increase hours to pre-pandemic levels for treatment households, but instead offsets a fall in hours worked in this time period, potentially linked to the end of the main harvest season. Hours of work in the sample do not recover to pre-pandemic levels until after schools fully reopen 2023 (Figure C.1.3).

The increases in labor supply are driven by increased engagement in household agriculture: adults in treatment households increase their hours in household agriculture by 33.4%. Increases in wage labor and non-agricultural self-employment hours are 22% and 13% respectively, but neither is statistically significant. Greater impacts on household agricultural hours than in wage employment are not surprising given that we estimate short-term impacts in the weeks following the partial school reopening. Wage employment opportunities and hours may be constrained in the short run and take longer to adjust, while hours in household agriculture are more flexible. As we discuss in the mechanisms section, increased adult agricultural labor may also be a response to reduced child labor due to children who had been helping on the household farm returning to school.

The smaller 15% and non-significant impact on the probability of adults engaging in any household agriculture indicate that hours increase primarily among those already working in agriculture during the closure period. Household agriculture engagement was affected less than other work activities by school closures and other pandemic restrictions. Adults may have been more likely to pause their engagement in household enterprise—more exposed to infections and pandemic restrictions as well as potentially more challenging to combine with childcare—and slower to resume these activities after reopening.

'Mixed' households with children eligible to return to school as well as children in adjacent grades do not change labor supply following the partial reopening. This is consistent with the economies of scale in childcare hours we observe in our data: one child returning to school while another of a similar age stays home is unlikely to meaningfully reduce adult childcare burdens, and the child remaining at home may be able to help make up for reductions in sibling childcare or agricultural labor provided by the child returning to school. In our analysis of mechanisms we thus drop mixed households and focus on 'pure' treatment vs. control households for a cleaner comparison.

We conduct a variety of robustness tests, focusing on the main impacts on total working hours in the last 7 days (Appendix B.5). Estimated impacts of the partial reopening on treated households are nearly identical when using individual rather than household fixed effects and county-by-month rather than month fixed effects and are not sensitive to the inclusion of individual or time-varying household controls (Table C.5.2). Results are also not sensitive to focusing on sub-samples of adults more likely to be parent caregivers or engaged in work, to omitting 'mixed' households from the analysis, or to varying which grades are included in the definition of treatment and control households (Table C.5.3). Finally, results are robust to defining *Post* by the date the potential reopening was announced rather than the actual reopening date (Table C.5.4). Slightly smaller estimated impacts in this specification together with no differential pre-trends suggest limited anticipation effects.

In addition to effects of schools reopening on individual labor supply we also consider short-term effects on household-level earnings by activity over the last 14 days (Table C.1.2). We see no statistically significant effects of the partial reopening on household income, and the point estimate for the impact on total household income is close to zero for treatment households.[21] Earnings data are limited—for all activities the 90th percentile of household earnings in the analysis sample is 0—in part due to a focus on the last 14 days, which does not accommodate seasonality or other variability in earnings, limiting our ability to detect impacts on income. Moreover, returns to agricultural labor hours may take until after harvest to materialize and the positive impacts of the partial reopening on agricultural work hours indicate treatment households are still engaged in harvest activities. The survey measures of income cannot be used to test whether these increased work hours led to increased incomes post-harvest. Despite these limitations, economically small and statistically insignificant impacts of the reopening on reported income suggest efforts to acquire funds to pay costs associated with the return to school are not an important mechanism for the labor supply effects we observe. This is consistent with no school fees being due at the time of the partial reopening, though some households may have needed to purchase school materials and pay other school-related expenses.

### 3.4.1 Heterogeneity

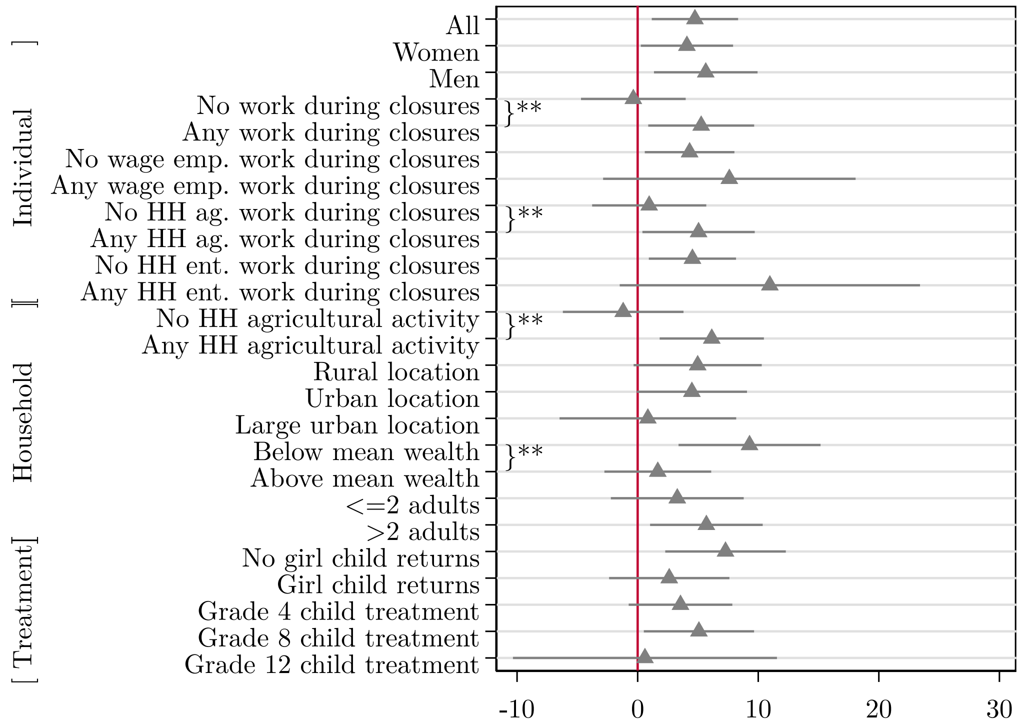
We next look at heterogeneous effects of schools reopening on total adult working hours based on individual, household, and treatment characteristics. Estimated effects for each subgroup are reported in Figure 3.4. We focus on differences for treatment households, as effects of the partial reopening for mixed households are not statistically significant.

Somewhat surprisingly, impacts on work hours are not significantly different for women (54% of the sample) relative to men. This contrasts with evidence from high-income countries, which consistently shows larger effects of the pandemic on mothers' labor supply relative to fathers' and other women's (e.g., Alon et al. (2022); Collins et al. (2021)). But this result aligns somewhat with the data on childcare hours in our context: responsibilities prior to the pandemic are less gendered than expected and both women's and men's hours increased by around 15 hours during school closures. The partial reopening shock thus affects both men's and women's labor supply, but the similar average increases in work hours for women and men partly reflect different mechanisms we discuss further in subsection 3.4.2.

The impact of partial reopening on work hours is large and positive for adults in agricultural households (61% of the sample, defined as households with any agricultural activity

[21] Estimates remain statistically insignificant using logged income (adding 1 for households with no income).

Figure 3.4: Heterogeneity in impacts of partial school reopening on adult work hours




Note: The figure summarizes estimated coefficients and 95% confidence intervals for the effect of $Post * Treat$ from Equation 3.1 for sub-samples with specified characteristics. Only coefficients for treatment households are shown. The outcome is total work hours in the 7 days prior to the interview. Data include observations from May-November 2020. Household characteristics are from the first time they are observed. Wealth is measured by an index based on housing and asset ownership. Brackets indicate significant differences between pairs of characteristics. Results showing tests of significance of the difference in impact by specified characteristics are reported in Table C.1.3 and Table C.1.4.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

during any school closures) but small and not statistically significant in non-agricultural households, consistent with effects on total hours being driven by household agriculture. This may be a result of household agriculture being the most flexible margin of labor supply in the short run, as well as adults compensating for potential shortfalls in on-farm child labor as children return to school.

In further support of the hypothesis that workplace flexibility is a key driver of heterogeneous impacts, we find that labor supply responses are concentrated among adults who worked at some point during the period of school closures (66%) and particularly by those that remained engaged in household agriculture (56%). Treated adults that worked while schools were closed increase their work time by 5.4 hours after one of their children returns to school, while those not working during the closures period do not significantly change their work hours after the partial reopening. The increase in the impact of the partial reopening is larger across adults continuing work in all sectors, though the difference compared to those that did not continue working is statistically significant only for work in household agriculture. This may be due to a lack of power, resulting from the small share of people working for either a wage (11%) or in a household enterprise (8%) during school closures. It may also be due to the fact that most adults that did not work in these activities engaged in household agriculture, while most of those that did not engage in agriculture did not work. The fact that adults who were able to continue working in wage employment or household enterprise during the school closures increase work hours by 8.6 and 8.5 hours per week respectively after the partial reopening—more than the 5.6 hours for those working in household agriculture—indicates how strongly labor supply to these activities can be constrained by childcare responsibilities.[22] Taken together, these results suggests that households adjusted their labor supply in the short run on the most flexible margins by increasing hours in activities they were already engaged in, particularly in household agriculture.

We observe no differences in impacts between urban (46%) and rural households. Part of this stems from the fact that the definition of 'urban' in the data includes many peri-urban areas; over 35% of household classified as urban are engaged in agriculture. Low take-up of formal childcare services and low wage employment in the sample may also limit urban/rural heterogeneity. The difference in impacts by location remains non-significant when restricting urban households to those in counties with the largest cities in Kenya ('large urban' locations), though the point estimate for these counties is lower than the effect in rural areas. This is consistent with impacts driven by agricultural work.

Turning next to differences by household wealth, measured as an index based on housing quality and asset ownership prior to the pandemic, we find that adults in poor households (39.4%) increase work hours by 9.3 hours compared to 1.7 hours for households with wealth above the mean. Below-mean-wealth households are around 30% more likely to engage in agriculture than wealthier households, and the wealth difference in impacts is substantially smaller, and no longer statistically significant, after conditioning on engagement in household agriculture. The difference by wealth remains negative for both agricultural and non-agricultural households separately, however, suggesting other contributing factors. In particular, poorer households may have had fewer resources to absorb increased childcare burdens during school closures and thus been more affected by the reopening. As no fees were associated with the return of grade 4 and 8 students to school, differences in the need

[22]These increases are marginally statistically significant Table C.1.3.

for income to pay school costs are unlikely to explain different impacts by household wealth.

While there is no significant difference in the average impact of the reopening by sex, the differences by household wealth are driven by men (Table C.1.5). Male survey respondents work 22 additional hours after the partial reopening in households with below-mean-wealth but do not change work hours in above-mean-wealth households. For female respondents there is no difference by household wealth, with an increase of 5.9 hours in below-mean- and 4.8 hours in above-mean-wealth households. This could imply that men are relatively more involved in childcare in less wealthy households than in wealthier ones, while women's responsibilities are similar. Differences in impacts on respondent childcare hours by household wealth and sex are consistent with this explanation, though differences may also relate to a need to make up for lost child agricultural labor. We further explore differences by household wealth in section 3.4.2.

Hours increase by slightly more in households with more than 2 adults (44.2%), but the difference is not statistically significant. We return to other differences by household composition—notably the presence of young children—in our exploration of mechanisms, as the differences in impacts between pure treatment and mixed households (with children in both treated and control grades) suggest the presence of siblings is critical in determining the effects of the partial school reopening on parents' labor supply.

Finally, we consider heterogeneity by characteristics of the child returning to school: their sex and age. Adult labor supply increases by 7.3 hours after the partial reopening in treated households when the returning child is a boy (43.7%), compared to 2.6 hours when it is a girl. Although only marginally significant (p=0.149), this economically meaningful difference aligns with expectations about gender differences in sibling childcare provision (Jakiela et al., 2020; Levison and Moe, 1998). If girls were providing more childcare to siblings than boys while schools were closed, the partial reopening may have been a less positive shock in households with a girl returning than those with a boy returning. The difference could also reflect different contributions by child sex to household agriculture, however, if boys worked more while schools were closed.

Children may also differ in their net demand for childcare and in their net contribution to household economic activities depending on their age. Figure 3.4 thus presents results separately for the sub-samples of households with a child around grade 4, around grade 8, and around grade 12.[23]

Work hours significantly increase by 30% for adults in treatment households with students in grade 8, around age 13-14, driven by increased hours in household agriculture. The impact is similar, at 21%, and borderline significant at the 90% level for households with

[23]Full results are presented in Table C.1.4. The patterns are similar if we use a categorical treatment grade variable interacted with a dummy for 'mixed' over the full sample of analysis households.

grade 4 students around age 9-10 ($p = 0.105$), and we cannot reject that the effects are the same across these two grades. The estimated impact is close to zero and not statistically significant for households with a grade 12 student. This is consistent with there being no childcare advantage to children at this age returning to school, though the estimate is highly imprecise due to the small number of grade 12 students in the sample and we consequently cannot reject that the effect is the same as the effect of a grade 8 student returning to school.

### 3.4.2 Mechanisms

We now turn to the mechanisms driving increased labor supply and the observed heterogeneity. A simple model of household production, labor supply, and the intra-household allocation of different activities guides our analysis (We describe the model in more detail in Appendix B.4). Adults in the household supply labor to home production as well as in the form of wage employment. In addition, they provide childcare to children present in the household and consume leisure. Adults get utility from consuming the returns to household labor, from leisure, and from altruistically caring about the welfare of their children.

Child well-being is subject to a constraint that total care received must not be below childcare needs, which vary by age and school closures. In line with the patterns in Figure 3.1 and Figure 3.2, we assume adult childcare provision is decreasing in child age and exhibits increasing returns to scale as a function of the number of children receiving care. Children present in the household are both recipients of childcare as well as potential contributors to home production and of childcare to other siblings. Whether they are net contributors or recipients of childcare depends on the child's age and the presence of other children demanding childcare.

In the model, a child returning to school decreases the childcare demanded by that child while also reducing their provision of sibling childcare (if they have younger siblings), so the effect on the household childcare constraint depends on the returning child's net childcare demand. A child returning to school also provides less household labor in households engaged in home production, which increases adult returns in home production relative to work outside the household. Which effects dominate depend on the age distribution of household children and the economic activities the household engages in (which—in line with our main results—we consider fixed in the short run).[24]

[24]In theory, school-related costs for a child returning to school may also lead to an income effect, increasing overall labor supply. But there were no additional fees incurred when schools reopened in Kenya. Though some parents may have paid outstanding bills and/or purchased school materials, this is unlikely to drive the labor supply effects we observe. Further, we do not find clear impacts of the partial reopening on measures of household income. We therefore concentrate our discussion of mechanisms for the increase in treated households' labor supply on changes in childcare and child labor.

Our model also differentiates by adult sex. When children (and in particular young children) benefit more from female care, or when social norms are such that the social cost of engaging in childcare relative to other activities is greater for men, women will engage in more childcare overall. Women are also more likely to engage in home production, since this work can be partially combined with childcare while wage work cannot. The model also predicts that women's labor supply effects after schools reopen will be skewed towards home production and more muted relative to men's when younger children remain in the household.

Our analysis of mechanisms is focused on impacts on household survey respondents. These are the only individuals for whom we observe childcare hours during the period around the partial school reopening and are also likely have lower measurement error in work hours. We exclude mixed households with children both in treated grades and in adjacent grades from our analysis of mechanisms. Separating treated households into pure treatment and mixed households is already implicitly testing for different effects by presence of additional adolescents in treated households. As shown above, we observe no significant impacts of the partial reopening on mixed households. The presence of other adolescent children remaining in the household after schools reopen—demanding both similar levels of childcare and providing similar levels of sibling childcare and household productive labor as their siblings who return to school—mutes the potential mechanisms by which one adolescent returning to school could affect parent labor supply. Results for treated households are similar when including mixed households in the analyses (see Tables C.5.6 and C.5.7).

#### 3.4.2.1 Childcare

Schools reopening unambiguously decreases the childcare burden for household adults when there are no other children who do not return to school. Figure 3.1 documents substantial increasing returns scale in to childcare: reported adult childcare hours are similar for households with one vs. those with multiple children. In the presence of other children, therefore, demand for adult childcare may not decrease as much. In fact, it may even increase, as older siblings returning to school may have taken part in the care of their younger siblings. This is common in our setting: while we do not measure sibling childcare during the study period, older siblings provided between 15-20 hours of childcare to younger siblings on average after schools fully reopened (Figure 3.1), with children in grades 3-9 providing 6.4 hours per week on average. Sibling childcare hours would likely have been greater while schools were closed.

Null effects of schools reopening on labor supply for mixed households in the previous section provide a first indication that this mechanism is important. When other children of similar age ranges remain in the household, the net childcare burden is unlikely to change substantially (Figure 3.2). Our results are consistent with this.

To further investigate the childcare mechanism, we test for differences in impacts of the partial reopening by the presence of below-school-age children in the household. We consider impacts on hours of work in household agriculture compared to hours in household enterprise and wage employment together, based on their differential flexibility in accommodating childcare needs. We also directly test for impacts on respondent childcare hours as a proxy for the household childcare burden,[25] and analyze whether the direction and magnitude of changes in childcare hours align with changes in labor supply.

Respondent childcare hours may be a reasonable proxy for the household childcare burden to the extent that respondents (including men) are key providers of childcare in the sample households, and that childcare hours capture meaningful variation in the childcare burden. Childcare hours can be thought of as a measure of the extensive margin of the childcare burden, as they combine time spent actively caring for children as well as passive time the adult is responsible for children but doing other activities. Variation on the intensive margin is not captured but is critical for the ability of adults to combine work and childcare, which may be most important for women. Children in grades 4 or 8 returning to school may not require significant active childcare but still require oversight and support.

We estimate impacts on these outcomes for the pooled sample of all respondents first, and then separately for women and men and discuss differences in the childcare and work hours responses by sex. Table 3.2 summarizes the results of these tests. The first panel shows average impacts of schools reopening on respondent labor supply and childcare hours for all households. We see again that both women and men increase their work hours and that this is driven primarily by increases in agricultural hours. This is not matched by a fall in respondent childcare hours on average: the point estimate is positive but not statistically significant.

The second panel of Table 3.2 presents heterogeneous impacts by the presence of younger children aged 0-4 in the household.[26] These below-school-age children are the most important source of childcare demand in most households, and also the children most likely to receive care from older siblings. Though impacts on agricultural hours are similar, the effect on overall labor supply is driven by respondents without below-school-age children in the household (51% of the sample). Their weekly work hours increase by 8.3 hours (or 48%) after older children return to school, while those of respondents with below-school-age children to

[25] Data on other adults' childcare is not available until after schools fully reopened.

[26] We run a single regression interacting a dummy for the presence of young children with treatment and month fixed effects, and use the *xlincom* command in Stata to calculate and present the treatment effect and standard error for households with young children, rather than showing the estimate for the interaction term. The results are identical to what we would obtain from running separate regressions on subsamples with and without very young children. We follow the same approach for Table 3.4.

Table 3.2: Impacts of partial school reopening by presence of children age 0-4 and respondent sex

| Respondent last 7 days hours in | HH agriculture work | | | HH enterprise or wage work | | | Childcare | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post × Treat, All | 4.177** | 3.950** | 4.452 | 1.717 | 1.265 | 3.161 | 1.980 | 3.280 | -5.463 |
| | (1.642) | (2.004) | (2.841) | (1.739) | (2.153) | (2.609) | (5.027) | (6.904) | (7.390) |
| Post × Treat, No child age 0-4 | 4.568** | 3.831 | 4.769 | 3.758* | 4.263 | 4.068 | -5.937 | -10.545 | 0.349 |
| | (2.195) | (2.570) | (4.158) | (2.094) | (2.848) | (3.073) | (6.213) | (8.302) | (9.572) |
| Post × Treat, Any child age 0-4 | 3.905 | 4.614 | 4.270 | -0.941 | -3.008 | 2.409 | 11.672 | 23.172** | -11.656 |
| | (2.450) | (3.097) | (3.901) | (2.916) | (3.178) | (4.410) | (8.171) | (11.779) | (11.674) |
| Observations | 2288 | 1331 | 902 | 2288 | 1331 | 902 | 2284 | 1328 | 901 |
| Mean, pre-reopen control | 12.947 | 11.702 | 14.827 | 6.325 | 5.509 | 7.651 | 52.747 | 60.059 | 41.175 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men |
| $p$-value, diff. by Any child age 0-4 | 0.840 | 0.846 | 0.930 | 0.191 | 0.089 | 0.758 | 0.087 | 0.020 | 0.427 |

Note: This table estimates impacts of the partial school reopening on respondent hours spent in the last 7 days on household agricultural labor (columns 1-3), wage employment and household non-agricultural enterprise labor (columns 4-6), and childcare (columns 7-9). Table C.1.6 includes results for total work hours. Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample, though results are similar when they are included (Table C.5.6). 'Post' is a dummy for being observed on or after the partial school reopening on October 12. The first panel shows average estimated impacts, and the second panel shows estimated impacts from regressions fully interacted with a dummy for the presence of children age 0-4. Coefficients for 'Any child age 0-4' represent the sum of the $Post \times Treat$ and $Post \times Treat \times$ *Any child age 0-4* terms. We include $p$-values for tests of whether the interaction term is equal to 0. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

take care of increase by a much lower and statistically insignificant 3.0 hours.[27]

Impacts of the partial reopening on childcare hours by presence of very young children are again not statistically significant but align with the impacts on work hours. The point estimate for childcare hours after children return to school is a decrease of 5.9 hours (or 11%) only for households without below-school-age children present, a similar magnitude to the increase in work hours in these households. For households with young children, the point estimate for childcare hours is an *increase* by 11.7 hours, a change that is in line with the average childcare hours provided by older siblings in our data, suggesting substitution of parent childcare for reduced sibling childcare.[28] These patterns are consistent with increases

[27]Impacts on total work hours are calculated by summing impacts on HH agriculture work and HH enterprise or wage work, and are shown in Table C.1.6.

[28]Impacts are similar for households where the child returning to school is in grade 4 and those where the child is in grade 8 (around ages 10 and 14 respectively), suggesting they have similar net childcare demands (Table C.5.5). We cannot reject that the coefficients are the same across grades, but the differences in magnitude for impacts on childcare hours align with the intuition that grade 4 students would demand more

in work hours after the partial reopening being mainly driven by households without below-school-age children where the reopening reduces the household childcare burdens on average, although we cannot reject that respondent childcare hours did not change in these households.

Next, we test whether the childcare mechanism is gendered (with the caveat that we are running into slightly low power for these tests).[29] Changes in household childcare needs are likely to affect female adults more than men, though our descriptive statistics suggest this may be less true in our context than elsewhere. Female respondents in control households report 60 hours of childcare in the last week on average during the school closures period, while men report 41 hours. Moreover, overall labor supply effects of school openings are similar for women and men, and both women and men reported childcare hours increasing substantially during the school closures (Figure 3.1).

Differences by the presence of young children may be important, as Figure 3.2 shows that, controlling for women's overall higher childcare burden, women are more responsible for care of children age 0-4 while men are more responsible for care of children in grades 3-9 during the school closures period for analysis sample households. As men's childcare responsibilities focus relatively more on monitoring older children, the partial reopening may particularly reduce their childcare burden while the impact on women—as primary caregivers for more demanding young children—would vary by whether there are younger children in the household which would have received care from their older sibling during the school closures.

The results in Columns 8 and 9 in Table 3.2 are broadly in line with this expectation, though due to sample size limitations and noisiness in the childcare hours measure we cannot reject the null of no impacts of the partial reopening on respondent childcare hours in most cases. Across all households the point estimate for the effect of the partial reopening on men's childcare hours is negative and of a similar magnitude as their increase in work hours, while the point estimate for women is positive. Women's weekly childcare time *increases* by 23.2 hours (38.6%) when below-school-age children are present. This is larger than the average hours of childcare provided by older siblings, suggesting women are not simply making up for reduced sibling childcare but taking on more of the household childcare burden. In contrast, women's childcare hours decrease by 10.5 hours (17.6%) when no young children are present. This effect is not statistically significant, but is significantly different from the effect in households with young children.

---

childcare and provide less sibling childcare than grade 8 students. There is no difference in the change in respondent childcare by whether the returning child is a girl or boy.

[29]Differences by sex of the respondent are not driven by differences between female- and male-headed households, as the survey respondent is not necessarily the household head. In 30% of households, including 49% of households with female respondents, the respondent is not the head.

Women's labor supply effects are qualitatively in line with these different childcare effects. They increase work hours by 8.1 hours after the partial reopening when below-school-age children are not present, but do not work more when they are: the point estimate is 0.9 hours. In such households women appear able to absorb additional childcare demands without reducing their labor supply, potentially because they are primarily engaged in household agriculture. The increase in women's work hours in households with no children age 0-4 is very close to the point estimate for their decrease in childcare hours: though that decrease is not statistically significant, it suggests that at least some of the increase in labor supply reflects a reallocation of time away from childcare.

The difference in labor supply impacts for women is driven by a large and significant effect of the presence of very young children on work hours in household enterprise and wage employment. Women in treatment households increase agricultural work hours relative to control households after the partial reopening regardless of the presence of below-school-age children, but only increase non-agricultural work hours if they have no such young children. If they do, they work three *fewer* hours outside of household agriculture on average, and this difference is statistically significant. This pattern suggests a substitution of labor supply hours from household enterprise and wage employment to household agriculture for women when the partial reopening increases the childcare burden: when below-school-age children are present. Such a substitution is consistent with household agriculture being more accommodating of multitasking with childcare. The consistent increase in household agriculture hours for women in treatment households can potentially be explained by the need to make up for reduced childcare agricultural labor, which we discuss in the next subsection.

The effects of the partial reopening on men's childcare hours are not statistically significant, and unlike for women they do not vary significantly by whether below-school-age children are present. This is consistent with men having a more limited role in care of the youngest children. Together with the significant increase in women's childcare in households with children age 0-4, the negative point estimate for men suggests that the partial reopening may lead to increased specialization in childcare within the household. This could also explain why the point estimate for men in households with no very young children is close to 0: there is a limited effect of the child returning to school and no change in women's specialization in household childcare. In line with the effects on childcare hours, the labor supply effects of the partial school reopening for men are similarly positive whether below-school-age children are present or not.

Overall these results do not provide clear evidence that increases in work hours after the partial school reopening are driven by reduced childcare burdens, but the analysis of impacts on childcare hours is limited by the lack of data on adults other than the respondent and the broad definition of childcare in the survey, leading to noisy estimated effects. Though we cannot conclude definitively, the patterns in the point estimates in Table 3.2 are consistent with net reductions in childcare for men on average and for women in households without younger children being a main mechanism driving increased labor supply by adults. We next turn to changes in child agricultural labor as a potential alternative channel.

#### 3.4.2.2 Child agricultural labor

The school reopenings in Kenya coincided with the main harvest season for most of the country. Children in agricultural households in the analysis sample worked an average of 6 hours per week in household agriculture while schools were closed, which rises to 10.3 hours (21.6% of the household total) in the 44% of agricultural households reporting some child agricultural labor during this period.[30] Child agricultural labor hours trend similarly across treatment and control households while schools are closed and are nearly identical after school fully reopened. On average, they fall by 2.6 hours per week in treatment households in the period after schools partly reopened compared to 0.9 hours in control households. This difference appears consistent with a reduction in child labor availability when a grade 4 or 8 child returns to school even if other children may be making up some of the shortfall, which we cannot distinguish given we only observe total agricultural hours for all children.

Adults in agricultural households may thus have increased agricultural work hours after the partial reopening in part to make up for reduced child labor. This mechanism could also contribute to the difference in impacts of the reopening on work hours between agricultural and non-agricultural households (Figure 3.4), as child labor is not (as) relevant for the latter group.

We explore this hypothesis by testing for differences in labor supply impacts by child engagement in household agriculture, and by estimating impacts on total child agricultural labor. We compare results between all households and agricultural households (those engaged in agriculture in 2020, which should be the only ones affected by this mechanism), and test for differences by whether they reported children contributing labor to household agriculture during the school closures period. As with the analysis of the childcare mechanisms, we look separately at impacts on work hours in household agriculture and in household enterprise or wage employment, as substitution for reduced child agricultural labor should lead to an increase in labor supply to household agriculture only, and possibly to a decrease in labor supply to other activities if respondents' time is constrained. We focus first on impacts for the pooled sample of all respondents, then analyze differences in impacts by respondent

[30]Moyi (2011) reports that around 30% of Kenyan children ages 6-14 engage in farm work (often alongside attending school). Child farm labor may be under-reported due to social desirability bias in answers, but is not correlated with treatment prior to the partial reopening (Table C.1.1).

sex and household wealth (which could be expected to be correlated with child agricultural labor).

Table 3.3 shows again that impacts of the partial reopening are concentrated in household agriculture hours, with no significant effect on hours in non-agricultural activities. The increase in respondent work hours is larger when considering the subset of households who engaged in any agriculture in 2020, as shown in Figure 3.4. Surprisingly, this is true for both agricultural and non-agricultural work hours.

Table 3.3: Impacts of partial school reopening by child engagement in household agriculture, survey respondents

| Last 7 days hours in | HH agriculture work (Respondent) | | | HH enterprise or wage work (Respondent) | | | Household agriculture (All household children) | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post x Treat | 4.177** | 5.244** | 2.019 | 1.717 | 3.366* | 5.281** | -0.668 | -1.009 | -0.046 |
| | (1.642) | (2.163) | (2.774) | (1.739) | (1.922) | (2.566) | (1.003) | (1.320) | (0.948) |
| Post x Treat × Any child ag labor before reopening | | | 7.076* | | | -4.994 | | | -1.945 |
| | | | (4.272) | | | (3.809) | | | (2.428) |
| Observations | 2288 | 1768 | 1768 | 2288 | 1768 | 1768 | 2288 | 1768 | 1768 |
| Mean, pre-reopen control | 12.947 | 16.806 | 16.806 | 6.325 | 5.325 | 5.325 | 4.167 | 5.409 | 5.409 |
| Sample of households | All | Ag. HH | Ag. HH | All | Ag. HH | Ag. HH | All | Ag. HH | Ag. HH |

Note: This table estimates impacts of the partial school reopening on respondent hours of work in the last 7 days for household agriculture (columns 1-3) and other activities (wage employment and household non-agricultural enterprise, columns 4-6), and on total child agricultural labor hours in the last 7 days (columns 7-9). Columns with the 'Ag. HH' sample include households engaged in agricultural production in 2020. Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample, though results are similar when they are included (Table C.5.7). 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The concentration of labor supply impacts in household agriculture could reflect that engagement in this activity is more flexible in the short-term than engagement in wage employment or household enterprise, but may also reflect substitution for reduced child agricultural labor. The point estimates for the impacts on reported child agricultural labor hours (Columns 7-9) are negative and larger in households reporting child agricultural labor before schools reopened, as expected, but they are not statistically significant. The magnitudes are similar to the raw differences in means in child agricultural hours by treatment status over the start of the partial reopening.

The failure to reject the null that child agricultural hours do not change differentially in treatment households after the partial reopening may partly reflect other household children increasing their labor hours when one child returns to school (we do not measure individual child labor hours). The small estimated magnitudes and lack of statistical significance may also be due to issues with measuring child hours in household agricultural labor in the data. First, social desirability bias may lead respondents to understate child labor hours. Second, some respondents may have misinterpreted the question about child agricultural hours, which asks for the total across all children (after first asking about each adult individually). In survey round 3, respondents report agricultural labor hours for a randomly selected child as well as the total for all children, and these are the same for a large share of households with multiple school-age children, suggesting some households may have interpreted the child agricultural labor question as asking for the average per child rather than the total.

Despite the non-significant estimated impacts of the partial reopening on child agricultural hours, patterns in the impacts on respondent work hours suggest this mechanism may yet explain part of the increase in agricultural work hours after the reopening. Column 3 of Table 3.3 shows that respondents in agricultural households reporting no child agricultural labor before the reopening increase agricultural work hours by a non-significant 2.0 hours, while those in households reporting some child agricultural labor work a statistically significant 9.1 additional hours in household agriculture after the partial school reopening. This is much larger in magnitude than the point estimate for total child agricultural labor, indicating that this mechanism likely cannot explain a large share of the increase in work hours observed in treatment households relative to control households after the partial reopening. Even if we assume respondents were reporting average hours per child rather than the total and we multiply this by the number of school-age children in the household, the estimated change in total child agriculture hours remains less than half of the estimated change in the respondent's agriculture hours. Yet the significant differences in impacts on adult labor supply by prior child agricultural labor indicate that the child labor mechanism does play a role—the extensive margin of child engagement in household agriculture is likely subject to less measurement error than the intensive margin of child agriculture hours.

The impacts of the partial reopening by prior child engagement in agricultural labor are the opposite for household enterprise or wage work hours compared to household agriculture hours. Respondents in agricultural households double their hours worked outside of household agriculture after the reopening—an increase of 5.3 hours—when they do not have children engaged in agricultural labor, increasing these hours by more than their agricultural work hours. But for respondents with children working in agriculture, non-agricultural work hours remain unchanged. A possible interpretation of these results is that when the partial school reopening is not a shock to child agricultural labor—and is therefore primarily a change in household childcare needs—adults respond by increasing labor supply to potentially higher-productivity non-agricultural activities. When the partial school reopening marks a shock to child labor as well as childcare, adults are prevented from increasing their non-agricultural labor supply by the need to make up for reduced child agricultural labor.

The opposite effects on hours spent in household agriculture and on other work activities of having children engaged in agricultural labor on the impact of the partial reopening are largely driven by women. Table 3.4 shows that both women and men increase household agriculture work after the reopening by more if they have children working in household agriculture, but this is significant only for women for whom the difference is stark. Women do not increase agricultural hours if they have no children that provided agricultural labor during school closures, but work 16.1 more hours in household agriculture after the partial reopening if they do—more than double the hours worked during school closures. The point estimates for work hours outside household agriculture are positive for both women and men in households with no child agricultural labor. While the estimated impact become smaller for men in households with children engaged in agriculture, they become *negative* for women. This pattern suggests that while all adults increase agricultural work hours in part to make up for reduced child agricultural labor after the partial reopening, women are more responsible for this than men.[31]

Impacts of child agricultural labor also differ by household wealth (Table 3.4 columns 3-4 and 7-8). Child engagement in household agriculture is more common in households with below-mean wealth (40.3% compared to 26.3%), and it is only in these households that we find differences in the impact of the reopening by child labor. Adults in low-wealth households supply 19.5 hours more to household agriculture after the reopening if children are engaged in agriculture, compared to 2.5 if they are not. There is no difference for adults in high-wealth households. Less wealthy households may be particularly dependent on agricultural production and have fewer alternatives to replace lost child agricultural labor.[32]

A potential concern with Table 3.3 is that child engagement in household agricultural labor may be correlated with other household characteristics which drive the significant difference in labor supply impacts. Table 3.4 shows that effects are not driven by respondent sex or household wealth—there are large and statistically significant differential impacts on labor supply by child agricultural labor for women and for low-wealth households. Differences in the impact of the partial reopening by child agricultural labor also persist when interacting this with the presence of additional household adults, having an older or female household head, and having any household enterprise.

In short, though impacts of the partial reopening on total reported child agriculture hours are not statistically significant there is suggestive evidence that part of the labor supply increase is driven by adults compensating for lost child labor on their homestead. While

[31] Table C.1.6 shows impacts on total work hours.

[32] These differences by wealth do not drive the differences by respondent sex: about the same share of female and male respondents are in low-wealth households.

Table 3.4: Impacts of partial school reopening by child engagement in household agriculture, survey respondents

| Respondent last 7 days hours in | HH agriculture work | | | | HH enterprise or wage work | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Post × Treat, All | 3.950** | 4.452 | 1.012 | 9.402*** | 1.265 | 3.161 | 1.311 | 2.201 |
| | (2.004) | (2.841) | (1.980) | (2.823) | (2.153) | (2.609) | (2.489) | (2.081) |
| Post × Treat, No child ag labor before reopening | -0.453 | 2.541 | 0.474 | 2.458 | 2.404 | 4.350 | 1.786 | 2.700 |
| | (2.167) | (3.368) | (2.253) | (3.114) | (2.759) | (3.275) | (3.136) | (2.449) |
| Post × Treat, Any child ag labor before reopening | 16.068*** | 6.854 | 2.241 | 19.455*** | -2.462 | 1.680 | -0.564 | 1.131 |
| | (4.430) | (4.787) | (4.068) | (5.539) | (2.165) | (4.050) | (4.014) | (4.111) |
| Observations | 1331 | 902 | 1385 | 903 | 1331 | 902 | 1385 | 903 |
| Mean, pre-reopen control | 11.702 | 14.827 | 11.223 | 15.500 | 5.509 | 7.651 | 7.718 | 4.262 |
| Sample | Women | Men | Above mean wealth | Below mean wealth | Women | Men | Above mean wealth | Below mean wealth |
| $p$-value, diff. by Any child ag labor | 0.001 | 0.462 | 0.704 | 0.008 | 0.166 | 0.609 | 0.645 | 0.743 |

Note: This table estimates impacts of the partial school reopening on respondent work hours in the last 7 days in household agriculture (columns 1-4) and in household non-agricultural enterprise or wage employment (columns 5-8). Table C.1.6 includes results for total work hours. Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Sub-samples based on wealth are determined based on an index of asset ownership and housing characteristics during the school closures period. The first panel shows average estimated impacts, and the second panel shows estimated impacts from regressions fully interacted with a dummy for any reported child agriculture labor before schools reopened. Coefficients for 'Any child ag labor' represent the sum of the $Post \times Treat$ and $Post \times Treat \times Any\ child\ ag\ labor$ terms. We include $p$-values for tests of whether the interaction term is equal to 0. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

present for all adults, this mechanism is particularly important for women, who may be able to combine increased agricultural work hours with childcare, thus more flexibly adjusting to the school reopening. This comes at a cost, however: adults in households with children engaged in agriculture lower their labor supply response in other potentially more productive sectors following the partial reopening.

## 3.5 Discussion

We find large and positive impacts of Kenya's partial school reopening on adult labor supply and suggest these impacts are driven by both changes in the household childcare burden and by reductions in child agricultural labor. Evidence for both mechanisms is stronger when considering differences in labor supply effects by household characteristics expected to determine how much the household childcare burden and child agricultural work hours change than when testing how schools reopening impacts these variables directly. Even though issues with accurately measuring the household childcare burden and children's agricultural labor in the RRPS lead to noisy estimates on our measures of these variables, the patterns in the point estimates match the predictions of a simple model of household labor and childcare decisions.

Changes in the childcare burden when a child returns to school are a primary driver of the increases in adult work hours. These increases are concentrated in households with no below-school-age children who receive high levels of childcare from parents and older siblings. In these households the child returning to school meaningfully reduces adult childcare hours, most strongly for women, who bear more of the responsibility for care of young children. Men see large average reductions in childcare hours and are less affected by the presence of young children, in line with more of their childcare being oriented toward older children such as those returning to school with the partial reopening. One interpretation of these patterns is men returning closer to their pre-pandemic childcare levels after the partial reopening, while women lose the childcare support they had been receiving from older children while schools were closed and specialize more in childcare. These differences in childcare impacts helps to explain why we do not observe differences in labor supply responses to the partial reopening by sex.

Women's specialization in childcare can also be seen in the substitution of work hours from wage employment and household enterprise to household agriculture. This is consistent with household agriculture being more flexible in accommodating childcare needs—including working while caring for children—than these other potentially more remunerative activities. This result highlights how childcare can be an important constraint preventing women from engaging in wage employment or non-agricultural self-employment in this setting (Delecourt and Fitzpatrick, 2021). During school closures, older siblings provided increased childcare to younger siblings at no (monetary) cost. Data from the RRPS suggest households lack access to affordable alternative childcare sources that might similarly free up women to engage in work outside of household agriculture (Figure C.1.2). Policies to increase access to affordable childcare might therefore free up women's labor supply.

The impact of the partial school reopening on total adult work hours is driven primarily by hours in household agriculture. Along with changes in the childcare burden, a second mechanism explaining these impacts is changes in child agricultural labor. Grade 4 and 8 students returning to school likely reduce child hours on the household farm, though we do not observe significant differences in total child agriculture hours. Since the partial reopening coincides with the main harvest season in most of Kenya, part of adults' increase in agricultural work hours appears to be motivated by a need to make up for this reduction in child work hours. Again, this mechanism appears to be more important for women, who increase agricultural work hours only when the household reports children engaged in agricultural labor. Women may have benefited more than men from the increase in child agricultural labor during the school closures since women have lower access to agricultural inputs than men on average in low- and middle-income countries such as Kenya (Anderson et al., 2021), and therefore been more adversely affected by the partial reopening. Substitution for child agricultural labor is also more important for low-wealth households, which may lack access to non-household-labor agricultural inputs relative to wealthier households. These results are consistent with child agricultural labor being an important input for many households in low-income countries.

School closures affected the availability of child labor, but school breaks more generally determine when children are available to work on the household farm. Changes in the Kenyan academic calendar following the COVID-19 school closures varied the timing of school breaks relative to periods of peak agricultural labor demand. In Kenya the main harvest season is from late September to November, though this differs somewhat by region. In 2020 due to school closures, all children except those in grades 4, 8, and 12 were out of school at this time. In 2021 all children were back in school and a one week holiday between terms fell in October, the beginning of the main harvest season for much of Kenya, but children were otherwise in school throughout the harvest period. In 2022 that holiday came two weeks sooner, potentially too soon for the harvest for most households, and children were again in school throughout the main harvest. In 2023 schools return to the pre-pandemic academic calendar, with a holiday in August and the last term ending in early November such that all children except those taking national exams are released to potentially contribute on the household farm. Allen (2024) finds that decreasing overlap between school days and peak agricultural periods in Malawi increases children's school participation and advancement; our study suggests that such a policy change could also benefit households by increasing parents' flexibility to continue work outside of household agriculture.

In addition to shedding light on the role of childcare burdens and child participation in household agriculture on adults' labor supply decisions in Kenyan households, the study also highlights the labor supply impacts of Kenya's pandemic school closures. Here, we present a back-of-the-envelope answer to the following questions: How do the labor supply effects of the positive shock of partial school reopenings compare to the initial labor supply reductions at the onset of the COVID-19 pandemic in Kenya, and what can we learn from our results about the the labor supply impact of the COVID-19 school closures? We apply survey weights generated by Pape et al. (2021) to estimate mean changes in labor supply nationally during the pandemic. We summarize these estimated parameters in Table 3.5.

Using recall data on respondents' pre-pandemic labor supply, we find that labor market participation across adults ages 18-64 in Kenya fell from 76% in February 2020 before the

Table 3.5: Summary of estimated parameters

| Value | Population | Estimate | Source |
|---|---|---|---|
| Change in work hours from Feb. to May-July 2020 | All Kenyan adults age 18-64 | 7.0 hours per week | RRPS, authors' calculations |
| Change in work hours from Feb. to May-July 2020 | Kenyan adults age 18-64 in analysis sample (with children in grades 3-9) | 11.1 hours per week | RRPS, authors' calculations |
| Change in working hours after partial school reopening in Oct. 2020 | Kenyan adults age 18-64 in analysis sample (with children in grades 3-9) | 4.7 hours per week | Table 1 |
| Change in work hours from Feb. to May-July 2020 if schools did not close | All Kenyan adults age 18-64 | 4.1 hours per week | RRPS, authors' calculations |
| Individuals in Kenya's labor force | All Kenyan working adults | 23.7 million | ILO (2021) |
| Weeks affected by school closures in 2020 | - | 29 weeks | Authors' calculations |
| Average hourly income in Kenya | All Kenyan households | USD 1.23 per hour | RRPS, authors' calculations |

Note: This table summarizes parameters used to estimate the impact of COVID-19 pandemic school closures in Kenya in 2020. RRPS is the Kenya Rapid Response Phone Survey (Pape et al., 2021). All parameters calculated from the RRPS are estimated using survey weights to generate nationally-representative figures.

pandemic to 59% in May-July, similar to what is reported in Pape et al. (2021) for all adults.Average working hours in the last 7 days among *all* Kenyan adults also fell, from 23.9 hours in February to 16.9 hours in May-July.[33] This 7 hour drop in labor supply represents the total effect of changes at the onset of the pandemic, of which school closures are one component.[34] For adults in our analysis sample (i.e., those with children in grades 3-9), average weekly hours fell from 30.4 in February to 19.3 in May-July.[35] As all of the analysis sample households include children, part of this 11.1 hour decrease in labor supply will reflect the effect of school closures. The fact that labor supply reductions were substantially larger on average among households with school-age children is a first indicator that school closures

[33] Among those working, hours fell from 39.7 to 29.2.

[34] The fall in work hours is similar for both agricultural and non-agricultural households, suggesting seasonality of labor does not account for a large share of the change.

[35] These values differ from those shown in Figure C.1.3 due to the use of survey weights, which we do not use in the main analysis.

may have played a substantial role in the aggregate fall in labor supply. Next, we discuss what the causal effects of schools reopening imply – under admittedly strong assumptions and extrapolations – for the likely contribution of school closures to the overall fall in labor supply during the pandemic.

The partial reopening only affected a subset of children older than those with the greatest childcare needs, whereas initial closures affected all school-age children. Reducing labor supply is also likely easier than increasing it, reflected in the lack of significant impacts of the partial reopening outside of household agriculture hours. Impacts of the partial school reopening on labor supply are thus likely to provide a conservative estimate of the reverse effect of initial school closures. We find that work hours in the last 7 days for adult household members increased by 4.7 after the partial school reopening (Table 3.1. This corresponds to 42.3% of the observed 11.1 hour reduction reduction in work hours in the analysis sample at the start of the pandemic.

To estimate how aggregate labor supply among *all* households—those with and without school-age children—would have changed if schools did not close at the onset of the pandemic, we assume work hours for all adults in households with school-age children (66.4% of households) increase by the same 4.7 hours per week regardless of which grades the children are in. Since children in grades 4 and 8 are relatively representative of school-aged children, and since we do not find differences in labor supply impact by the grade of the returning child, extrapolating our results to all households with school-aged children seems justifiable. We assume that households without school-age children would be unaffected, thus abstracting from any general equilibrium effects in the labor market.[36] Combining these assumptions, we estimate that school closures led to a drop in average weekly work hours from February to May-July among *all* adults nationally of 2.9 hours. In other words, total labor hours would have fallen by 4.1 hours per adult instead of 7.0 had schools remained open. We therefore estimate that school closures account for (at least) 40% of the decrease in work hours in the first few months of the COVID-19 pandemic in Kenya, a very substantial effect. We next turn to the implied aggregate implications of school closures.

Across Kenya's labor force of 23.7 million (ILO, 2021), a reduction of 2.9 work hours per week over the 29 weeks during which schools were closed from 16 March 2020 to 4 January 2021 (excluding 12 weeks during which schools would have been on break) adds up to just under 2 billion hours in total. If we value these hours at the average hourly income observed in the data (USD 1.23 per hour[37]), this suggests that school closures reduced incomes nationally

[36]Workers in the education sector will have been affected regardless of whether they have school-age children; ignoring how their labor supply would also have been higher if schools had not closed during the pandemic is one reason to believe our estimates of the contribution of school closures to pandemic changes in labor supply are conservative.

[37]If we take Kenya's GDP per capita and divide that by the average weekly hours worked in the data

by USD 2.45 billion—2.6% of Kenya's 2019 GDP. This number may be overstated, if labor supply reductions due to school closures were concentrated in the lowest-paying activities.

This is a simplified back-of-the-envelope calculation making admittedly strong assumptions. In particular, it focuses on labor supply only, and abstracts from any counterfactual labor demand changes during the pandemic or general equilibrium effects. That said, it provides a rough ballpark estimate of the magnitude of the labor supply impact of Kenya's school closures during COVID-19. Importantly, we do not draw any conclusion about the net benefits of school closures in Kenya. Our calculation considers just one aspect: their impact on market labor for adults in households with schoolchildren. It ignores the value of increased non-market provided childcare to children and adults themselves. Moreover, we do not estimate costs for teachers or the potential loss of learning of the affected children due to school closures. We also do not assess the public health benefits of school closures—the primary motivation—nor increased income for households having older children contributing more to agriculture or sibling care. While these are key components for fully evaluating the impacts of school closures, a better understanding of their labor supply impacts in a context with more informal and flexible employment may inform the discussion of school closures as a policy response to a resurgent COVID-19 or a future pandemic.

Although the household labor supply shock we analyze takes place in the context of a global pandemic, the results on labor supply impacts will continue to have relevance. Further school closures appear unlikely at this stage in the pandemic but may occur in response to future outbreaks of COVID-19 variants or other diseases. In addition, although some pandemic-related restrictions were still in effect at the time schools partly reopened in Kenya in October 2020, many had been relaxed. The impacts we estimate may therefore generalize to policies affecting household childcare needs and child labor supply in similar settings with some ongoing COVID-19 caseloads and basic government policies around public health and safety, which appear to be the new normal moving forward. They also shed light on likely effects of school closures for reasons unrelated to public health events, such as teacher strikes (Jaume and Willén, 2021).

## 3.6 Conclusion

We present nationally-representative estimates of the impacts of a shock to childcare burdens and child labor on adult labor supply in a lower-middle-income African country using pandemic-related school closure policy changes in Kenya as exogenous childcare shocks. Having a child eligible to return increases adult work hours in the weeks after schools partially

prior to the pandemic times the number of weeks in a year, we obtain a similar hourly income estimate of USD 1.45.

reopen, primarily in more flexible household agriculture. Though estimated effects on childcare hours and child agricultural labor are noisy, the results indicate that both changes in childcare burdens and a need to make up for reduced child labor play a role in adult labor supply changes. Increases are driven by less wealthy households who are more engaged in agriculture and may have fewer resources to cope with childcare and labor shocks.

Unlike studies of pandemic school closures in high-income countries, impacts are not concentrated primarily among women. Both men and women increase overall work hours by a substantial 29% after schools partially reopen, relative to adults in households with children in grades adjacent to those affected by the reopening. Although overall labor supply effects are similar by sex, we show that women in particular remain constrained by having to care for younger children when older siblings return to school. In households with below-school-age children where students returning to school were likely net childcare providers, women bear more of the increased childcare burden after schools reopen, preventing increases in labor supply observed by women in households where the reopening was a positive childcare shock. Men's childcare responsibilities concentrate relatively more on school-age children and and benefit more from childcare reductions after the reopening. Both sexes increase agricultural work hours by more in households with children engaged in household agriculture, though the difference is larger for women. Some studies of changes in childcare availability or cost in Africa similarly report significant impacts for men as well as women (Bjorvatn et al., 2022; Martinez et al., 2012), but most focus exclusively on women.Considering how childcare burdens are allocated across *all* household members is critical for understanding the intra-household distributional impacts of childcare shocks.

Our study generates three main policy-relevant takeaways. First, parents in Kenya appear to have limited options for coping with increased childcare burdens beyond reducing work hours or combining work and childcare. We show that women in particular substitute away from wage employment or household enterprise work into household agriculture when their childcare responsibilities require such multitasking. This is despite many households having additional adults, many parents being engaged in potentially more flexible household farm work, and adults working just 24 hours a week on average (less than 'full-time') before the pandemic. Older siblings are an important source of (unpaid) childcare in this context, and the partial reopening increases childcare burdens for parents with younger children when sibling caregivers return to school. This indicates that households lack alternative childcare options or that they cost more than adults could earn by working instead of caring for children themselves. Households in the data report almost no childcare provision by non-household members in 2020 after the start of the pandemic, but even in 2021 as concerns and restrictions eased use of non-household childcare remained very uncommon. Several studies point to high costs as a main constraint to using formal childcare centres in Kenya and advocate for public subsidies to facilitate access (Clark et al., 2021; Murungi, 2013).

Policies aiming to increase childcare availability may therefore be less effective if they are not complemented by policies to reduce cost.

Second, the timing of when children are in school affects some households through child agricultural labor. Older children contribute significantly to household agriculture when not in school, and after the partial reopening adults in households with children returning to school increased agriculture work hours partly to make up for reduced child farm labor.The 2020 school closures disrupted academic calendars with implications for the timing of school terms and breaks relative to the agricultural cycle over 2021-2023 and thereafter. This affected whether children were in school during labor-intensive agricultural periods in Kenya. Given the important role of children in agricultural production for many households, future work could consider how these changes affect children's school attendance and achievement (Allen (2024) explores this question in Malawi) and household production decisions including adults' labor supply.

Lastly, our results highlight childcare gradients by age of the child and sex of the caregiver that have have relevance to other policies affecting household childcare. If we expect that childcare needs decrease with child age, we would expect the estimated average impacts on labor supply to be lower bounds on the impact of policies increasing childcare availability for younger children, such as providing free full-day childcare for young children during the working week (as schools implicitly provide to students). Clark et al. (2019) show that subsidies for childcare centres increase labor supply for women in an informal settlement in Nairobi. Our results indicate such policies could have positive effects outside urban settings, and that women should be particularly strongly affected given their primary role as caregivers for children below school-age. In addition, older children might also benefit from reduced need to care for younger siblings.

# Bibliography


Admassie, A. (2003). Child labour and schooling in the context of a subsistence rural economy: can they be compatible? *International Journal of Educational Development*, 23(2):167–185.

Afridi, F., Mahajan, K., and Sangwan, N. (2022). The gendered effects of droughts: Production shocks and labor response in agriculture. *Labour Economics*, 78:102227.

Akter, S. (2021). Do catastrophic floods change the gender division of labor? panel data evidence from pakistan. *International Journal of Disaster Risk Reduction*, 60:102296.

Albert, C., Bustos, P., and Ponticelli, J. (2021). The effects of climate change on labor and capital reallocation. National Bureau of Economic Research WP 28995.

Alderman, H., Hoddinott, J., and Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxford Economic Papers*, 58(3):450–474.

Allen, J. I. (2024). Double-booked: Effects of overlap between school and farming calendars on education and child labor. IFPRI WP02235.

Alon, T., Coskun, S., Doepke, M., Koll, D., and Tertilt, M. (2022). From mancession to shecession: Women's employment in regular and pandemic recessions. *NBER Macroeconomics Annual*, 36(1):83–151.

Amuedo-Dorantes, C., Marcén, M., Morales, M., and Sevilla, A. (2023). Schooling and parental labor supply: Evidence from COVID-19 school closures in the united states. *ILR Review*, 76(1):56–85.

Anderson, C. L., Reynolds, T. W., Biscaye, P., Patwardhan, V., and Schmidt, C. (2021). Economic benefits of empowering women in agriculture: Assumptions and evidence. *The Journal of Development Studies*, 57(2):193–208.

Andrew, A., Cattan, S., Dias, M. C., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A., and Sevilla, A. (2020). How are mothers and fathers balancing work and family under lockdown? Institute for Fiscal Studies Briefing Note 290.

ASU (2020). The global locust initiative. Arizona State University, https://sustainability.asu.edu/global-locust-initiative/.

Auffhammer, M. and Kahn, M. E. (2018). The farmer's climate change adaptation challenge in least developed countries. In *Handbook of environmental economics*, volume 4, pages 193–229. Elsevier.

Baez, J. E., Caruso, G., and Niu, C. (2020). Extreme weather and poverty risk: evidence from multiple shocks in mozambique. *Economics of Disasters and Climate Change*, 4:103–127.

Baker, R. B., Blanchette, J., and Eriksson, K. (2020). Long-run impacts of agricultural shocks on educational attainment: Evidence from the boll weevil. *The Journal of Economic History*, 80(1):136–174.

Banerjee, A., Duflo, E., Postel-Vinay, G., and Watts, T. (2010). Long-run health impacts of income shocks: Wine and phylloxera in nineteenth-century france. *The Review of Economics and Statistics*, 92(4):714–728.

Barrett, C., Ortiz-Bobea, A., and Pham, T. (2021). Structural transformation, agriculture, climate and the environment. STEG Pathfinding Paper 01.

Baseler, T. and Hennig, J. (2023). Disastrous displacement: the long-run impacts of landslides. World Bank Policy Research Working Paper 10535.

Basu, K. and Van, P. H. (1998). The economics of child labor. *American Economic Review*, pages 412–427.

Baysan, C., Burke, M., González, F., Hsiang, S., and Miguel, E. (2019). Non-economic factors in violence: Evidence from organized crime, suicides and climate in mexico. *Journal of Economic Behavior & Organization*, 168:434–452.

Bazzi, S. and Blattman, C. (2014). Economic shocks and conflict: Evidence from commodity prices. *American Economic Journal: Macroeconomics*, 6(4):1–38.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Beegle, K., Dehejia, R., and Gatti, R. (2009). Why should we care about child labor? the education, labor market, and health consequences of child labor. *Journal of Human Resources*, 44(4):871–889.

Begueria, S., Vicente-Serrano, S. M., Reig, F., and Latorre, B. (2014). Standardized precipitation evapotranspiration index (spei) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10):3001–3023.

Berlemann, M., Methorst, J., and Thum, M. (2023). Do floods scare off residents? *Economics Letters*, 222:110942.

Bharadwaj, P., Lakdawala, L. K., and Li, N. (2020). Perverse consequences of well intentioned regulation: Evidence from india's child labor ban. *Journal of the European Economic Association*, 18(3):1158–1195.

Bjorvatn, K., Ferris, D., Gulesci, S., Nasgowitz, A., Somville, V., and Vandewalle, L. (2022). Childcare and cash grants for labor supply and wellbeing: experimental evidence from uganda. CEPR Discussion Paper No. DP17243.

Blattman, C. and Miguel, E. (2010). Civil war. *Journal of Economic Literature*, 48(1):3–57.

Bollfrass, A. and Shaver, A. (2015). The effects of temperature on political violence: global evidence at the subnational level. *PLoS One*, 10(5):e0123505.

Bonds, S. (2021). Information, student-parent communication, and secondary school choice: Experimental evidence from kenya. Working Paper, University of California at Berkeley.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint 2108.12419*.

Boserup, E., Tan, S. F., and Toulmin, C. (2013). *Woman's role in economic development*. Routledge: London.

Botzen, W. W., Deschenes, O., and Sanders, M. (2019). The economic impacts of natural disasters: A review of models and empirical studies. *Review of Environmental Economics and Policy*.

Brader, L., Djibo, H., Faye, F., Ghaout, S., Lazar, M., Luzietoso, P., and Babah, M. O. (2006). Towards a more effective response to desert locusts and their impacts on food security, livelihoods and poverty. Multilateral evaluation of the 2003–05 Desert locust campaign. Food and Agriculture Organisation, Rome.

Brakenridge, G. (2023). Global active archive of large flood events. Technical report, Dartmouth Flood Observatory, University of Colorado.

Branco, D. and Féres, J. (2021). Weather shocks and labor allocation: Evidence from rural brazil. *American Journal of Agricultural Economics*, 103(4):1359–1377.

Browning, M. (1992). Children and household economic behavior. *Journal of Economic Literature*, 30(3):1434–1475.

Buhaug, H., Benjaminsen, T. A., Sjaastad, E., and Theisen, O. M. (2015). Climate variability, food production shocks, and violent conflict in sub-saharan africa. *Environmental Research Letters*, 10(12):125015.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). Climate and conflict. *Annual Review of Economics*, 7(1):577–617.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Caretta, M., Mukherji, A., Arfanuzzaman, M., Betts, R., Gelfan, A., Hirabayashi, Y., Lissner, T., Liu, J., Lopez Gunn, E., Morgan, R. andMwanga, S., and Supratid, S. (2022). Water. In *In: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner et al. (eds.)]*, page 551–712. Cambridge University Press.

Carleton, T., Hsiang, S. M., and Burke, M. (2016). Conflict in a changing climate. *The European Physical Journal Special Topics*, 225(3):489–511.

Carter, M. R. and Barrett, C. B. (2006). The economics of poverty traps and persistent poverty: An asset-based approach. *The Journal of Development Studies*, 42(2):178–199.

Carter, M. R., Little, P. D., Mogues, T., and Negatu, W. (2007). Poverty traps and natural disasters in ethiopia and honduras. *World Development*, 35(5):835–856.

Caruso, R., Petrarca, I., and Ricciuti, R. (2016). Climate change, rice crops, and violence: Evidence from indonesia. *Journal of Peace Research*, 53(1):66–83.

Casale, D. and Posel, D. (2020). Gender and the early effects of the COVID-19 crisis in the paid and unpaid economies in South Africa. *NIDS-CRAM Policy Paper*, 18.

Cederman, L.-E. and Pengl, Y. (2019). Global conflict trends and their consequences. ETH Zürich Working Paper.

Chassang, S. and Padró i Miquel, G. (2009). Economic shocks and civil war. *Quarterly Journal of Political Science*, 4(3):211–28.

Chatterjee, S. (2022). How hard did that sting? estimating the economic costs of locust attacks on agricultural production. *Applied Economic Perspectives and Policy*, 44(1):434–459.

Chauhan, P. (2021). Gendering COVID-19: Impact of the pandemic on women's burden of unpaid work in India. *Gender Issues*, 38:395–419.

Chemin, M., De Laat, J., and Haushofer, J. (2013). Negative rainfall shocks increase levels of the stress hormone cortisol among poor farmers in kenya. *Available at SSRN 2294171*.

Chen, J. J., Mueller, V. A., Jia, Y., and Tseng, S. K.-H. (2017). Validating migration responses to flooding using satellite and vital registration data. *American Economic Review*, 107(5):441–445.

Chuang, Y. (2019). Climate variability, rainfall shocks, and farmers' income diversification in india. *Economics Letters*, 174:55–61.

Ciccone, A. (2013). Estimating the effect of transitory economic shocks on civil conflict. *Review of Economics and Institutions*, 4(2):Article 1.

CIESIN (2018). Gridded population of the world, version 4 (gpwv4): Population count, revision 11. Center for International Earth Science Information Network - CIESIN - Columbia University. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4JW8BX5.

Clark, S., Cotton, C., and Marteleto, L. J. (2015). Family ties and young fathers' engagement in Cape Town, South Africa. *Journal of Marriage and Family*, 77(2):575–589.

Clark, S., De Almada, M., Kabiru, C. W., Muthuri, S., and Wanjohi, M. (2021). Balancing paid work and child care in a slum of nairobi, kenya: the case for centre-based child care. *Journal of Family Studies*, 27(1):93–111.

Clark, S., Kabiru, C. W., Laszlo, S., and Muthuri, S. (2019). The impact of childcare on poor urban women's economic empowerment in Africa. *Demography*, 56(4):1247–1272.

Collier, P. and Hoeffler, A. (1998). On economic causes of civil war. *Oxford Economic Papers*, 50(4):563–573.

Collier, P. and Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4):563–595.

Collins, C., Landivar, L. C., Ruppanner, L., and Scarborough, W. J. (2021). COVID-19 and the gender gap in work hours. *Gender, Work & Organization*, 28:101–112.

Colmer, J. (2021). Temperature, labor reallocation, and industrial production: Evidence from india. *American Economic Journal: Applied Economics*, 13(4):101–24.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Connelly, R. (1992). The effect of child care costs on married women's labor force participation. *The Review of Economics and Statistics*, pages 83–90.

Conte, B., Piemontese, L., and Tapsoba, A. (2023). The power of markets: Impact of desert locust invasions on child health. *Journal of Health Economics*, 87:102712.

Couttenier, M. and Soubeyran, R. (2014). Drought and civil war in sub-saharan africa. *The Economic Journal*, 124(575):201–244.

Cressman, K. (2001). Desert locust guidelines: Survey. Food and Agriculture Organization of the United Nations (FAO).

Cressman, K. and Ferrand, C. (2021). Signs of hope in east africa, as control campaign tames locust upsurge. FAO Newsroom, https://www.fao.org/news/story/en/item/1393635/icode/.

Cressman, K. and Stefanski, R. (2016). Weather and desert locusts. Food and Agriculture Organization of the United Nations (FAO).

Cressman, K., Van der Elstraeten, A., and Pedrick, C. (2016). elocust3: An innovative tool for crop pest control. Food and Agriculture Organization of the United Nations (FAO).

Crost, B., Duquennois, C., Felter, J. H., and Rees, D. I. (2018). Climate change, agricultural production and civil conflict: Evidence from the philippines. *Journal of Environmental Economics and Management*, 88:379–395.

Crost, B., Felter, J., and Johnston, P. (2014). Aid under fire: Development projects and civil conflict. *American Economic Review*, 104(6):1833–1856.

Crost, B., Felter, J. H., and Johnston, P. B. (2016). Conditional cash transfers, civil conflict and insurgent influence: Experimental evidence from the philippines. *Journal of Development Economics*, 118:171–182.

Cruz, J.-L. and Rossi-Hansberg, E. (2021). The economic geography of global warming. Technical report, National Bureau of Economic Research WP 28466.

Dal Bó, E. and Dal Bó, P. (2011). Workers, warriors, and criminals: social conflict in general equilibrium. *Journal of the European Economic Association*, 9(4):646–677.

Davis, M. (2022). Pastoralism, farming and a changing climate in the sahel region. Stockholm Environment Institute (SEI). https://www.sei.org/featured/pastoralism-farming-climate-in-sahel/.

de Janvry, A., Finan, F., Sadoulet, E., and Vakis, R. (2006). 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*, 79(2):349–373.

De Vreyer, P., Guilbert, N., and Mesple-Somps, S. (2015). Impact of natural disasters on education outcomes: evidence from the 1987–89 locust plague in mali. *Journal of African Economies*, 24(1):57–100.

Del Boca, D., Oggero, N., Profeta, P., and Rossi, M. (2020). Women's and men's work, housework and childcare, before and during COVID-19. *Review of Economics of the Household*, 18(4):1001–1017.

Delecourt, S. and Fitzpatrick, A. (2021). Childcare matters: Female business owners and the baby-profit gap. *Management Science*, 67(7):4455–4474.

Dell, M., Jones, B. F., and Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, 4(3):66–95.

Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? the new climate-economy literature. *Journal of Economic literature*, 52(3):740–798.

Dercon, S. (2004). Growth and shocks: evidence from rural ethiopia. *Journal of Development Economics*, 74(2):309–329.

Dercon, S. and Hoddinott, J. (2004). Health, shocks and poverty persistence. *Insurance against poverty*, pages 123–136.

Deshpande, A. (2020). The COVID-19 pandemic and lockdown: First order effects on gender gaps in employment and domestic time use in India. GLO Discussion Paper No. 607.

Dillon, A., Mueller, V. A., and Salau, S. (2011). Migratory responses to agricultural risk in northern nigeria. *American Journal of Agricultural Economics*, 93(4):1048–1061.

Dinkelman, T. (2017). Long-run health repercussions of drought shocks: evidence from south african homelands. *The Economic Journal*, 127(604):1906–1939.

Djoumessi Tiague, B. (2023). Floods, agricultural production, and household welfare: Evidence from tanzania. *Environmental and Resource Economics*, 85(2):341–384.

Dobson, H. M. (2001). Desert locust guidelines: control. Food and Agriculture Organization of the United Nations.

Dube, O. and Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from colombia. *The Review of Economic Studies*, 80(4):1384–1421.

Duryea, S. and Arends-Kuenning, M. (2003). School attendance, child labor and local labor market fluctuations in urban Brazil. *World Development*, 31(7):1165–1178.

Efobi, U. (2022). The long-term labor market effect of drought exposure: Evidence from nigeria. *The Journal of Development Studies*, pages 1–19.

Emerick, K. (2018). Agricultural productivity and the sectoral reallocation of labor in rural india. *Journal of Development Economics*, 135:488–503.

Fafchamps, M., Udry, C., and Czukas, K. (1998). Drought and saving in west africa: are livestock a buffer stock? *Journal of Development Economics*, 55(2):273–305.

Fang, X., Kothari, S., McLoughlin, C., and Yenice, M. (2020). The economic consequences of conflict in sub-saharan africa. IMF Working Paper 20/221.

FAO (2022). Nigeria agriculture at a glance. Technical report, Food and Agriculture Organization of the United Nations, https://www.fao.org/nigeria/fao-in-nigeria/nigeria-at-a-glance/en/.

Farré, L., Fawaz, Y., González, L., and Graves, J. (2020). How the COVID-19 lockdown affected gender inequality in paid and unpaid work in Spain. IZA Discussion Paper No. 13434.

Fetzer, T. (2020). Can workfare programs moderate conflict? evidence from india. *Journal of the European Economic Association*, 18(6):3337–3375.

Fick, S. E. and Hijmans, R. J. (2017). Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12):4302–4315.

Fjelde, H. (2015). Farming or fighting? agricultural price shocks and civil war in africa. *World Development*, 67:525–534.

Food and Agriculture Organization of the United Nations (FAO) (2022a). Desert locust upsurge may be declining but remaining swarms require vigilance in east africa and yemen. FAO Newsroom, https://www.fao.org/newsroom/detail/desert-locust-upsurge-may-be-declining-but-remaining-swarms-require-vigilance-in-east-africa-and-yemen/en.

Food and Agriculture Organization of the United Nations (FAO) (2022b). Locust watch: Desert locust. FAO Locust Watch, http://www.fao.org/ag/locusts/en/info/info/index.html.

Food and Agriculture Organization of the United Nations (FAO) (2023a). Desert Locust. FAO, https://www.fao.org/locusts/en.

Food and Agriculture Organization of the United Nations (FAO) (2023b). Development flows to agriculture. FAOSTAT, https://www.fao.org/faostat/en/#data/EA/visualize.

Food and Agriculture Organization of the United Nations (FAO) and World Meteorological Organization (WMO) (2016). Weather and desert locusts.

Franklin, S. and Labonne, J. (2019). Economic shocks and labor market flexibility. *Journal of Human Resources*, 54(1):171–199.

French, E. and Taber, C. (2011). Identification of models of the labor market. In *Handbook of Labor Economics*, volume 4, pages 537–617. Elsevier.

Freudenreich, H. and Kebede, S. W. (2022). Experience of shocks, household wealth and expectation formation: Evidence from smallholder farmers in kenya. *Agricultural Economics*, 53(5):756–774.

Furman, J., Kearney, M. S., and Powell, W. (2021). The role of childcare challenges in the us jobs market recovery during the COVID-19 pandemic. National Bureau of Economic Research WP 28934.

Garg, T., McCord, G. C., and Montfort, A. (2020). Can social protection reduce environmental damages? IZA DP No. 13247.

Gatti, N., Baylis, K., and Crost, B. (2021). Can irrigation infrastructure mitigate the effect of rainfall shocks on conflict? evidence from indonesia. *American Journal of Agricultural Economics*, 103(1):211–231.

Giurge, L. M., Whillans, A. V., and Yemiscigil, A. (2021). A multicountry perspective on gender differences in time use during COVID-19. *Proceedings of the National Academy of Sciences*, 118(12):e2018494118.

Global Administrative Areas (GADM) (2021). Database of global administrative boundaries v3.6. GADM, https://gadm.org/.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Grabrucker, K. and Grimm, M. (2021). Is there a rainbow after the rain? how do agricultural shocks affect non-farm enterprises? evidence from thailand. *American Journal of Agricultural Economics*, 103(5):1612–1636.

Grantham, K., Rouhani, L., Gupta, N., Melesse, M., Dhar, D., Mehta, S. K., and Kingra, K. J. (2021). Evidence review of the global childcare crisis and the road for post-COVID-19 recovery and resilience. Technical report, International Development Research Centre.

Gray, C. and Mueller, V. A. (2012a). Drought and population mobility in rural ethiopia. *World development*, 40(1):134–145.

Gray, C. and Mueller, V. A. (2012b). Natural disasters and population mobility in bangladesh. *Proceedings of the National Academy of Sciences*, 109(16):6000–6005.

Green, J. (2022). The unique challenges of responding to desert locust outbreaks. Entomology Today. https://entomologytoday.org/2022/01/11/unique-challenges-integrated-pest-management-desert-locust-outbreaks/.

Gröger, A. and Zylberberg, Y. (2016). Internal labor migration as a shock coping strategy: Evidence from a typhoon. *American Economic Journal: Applied Economics*, 8(2):123–153.

Grossman, H. I. (1999). Kleptocracy and revolutions. *Oxford Economic Papers*, 51(2):267–283.

Guardado, J. and Pennings, S. (2021). The seasonality of conflict. World Bank Policy Research Working Paper 9373.

Guha-Sapir, D., Below, R., and Hoyos, P. (2023). Em-dat: The CRED/OFDA international disaster database. Technical report, Universite Catholique de Louvain.

Guiteras, R., Jina, A., and Mobarak, A. M. (2015). Satellites, self-reports, and submersion: exposure to floods in bangladesh. *American Economic Review*, 105(5):232–236.

Halim, D., Perova, E., and Reynolds, S. (2023). Childcare and mothers' labor market outcomes in lower-and middle-income countries. *The World Bank Research Observer*, 38(1):73–114.

Hallegatte, S., Vogt-Schilb, A., Rozenberg, J., Bangalore, M., and Beaudet, C. (2020). From poverty to disaster and back: A review of the literature. *Economics of Disasters and Climate Change*, 4:223–247.

Hansen, B., Sabia, J. J., and Schaller, J. (2022). Schools, job flexibility, and married women's labor supply: Evidence from the COVID-19 pandemic. National Bureau of Economic Research WP 29660.

Harari, M. and La Ferrara, E. (2018). Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics*, 100(4):594–608.

Hardeweg, B. (2001). A conceptual framework for economic evaluation of desert locust management interventions. *Pesticide Policy Project Publication Series*, Special Issue No. 5.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations–the cru ts3. 10 dataset. *International Journal of Climatology*, 34(3):623–642.

Hastings, J. and Ubilava, D. (2023). Agricultural shocks and social conflict in southeast asia. *arXiv preprint arXiv:2304.10027*.

Heath, R. (2017). Fertility at work: Children and women's labor market outcomes in urban ghana. *Journal of Development Economics*, 126:190–214.

Heckman, J. J. and Honore, B. E. (1990). The empirical content of the roy model. *Econometrica: Journal of the Econometric Society*, pages 1121–1149.

Heggeness, M. L. (2020). Estimating the immediate impact of the COVID-19 shock on parental attachment to the labor market and the double bind of mothers. *Review of Economics of the Household*, 18(4):1053–1078.

Hirvonen, K., Gilligan, D. O., Leight, J., Tambet, H., and Villa, V. (2023). Do ultra-poor graduation programs build resilience against droughts? evidence from rural ethiopia. IFPRI Discussion Paper 02206.

Hoddinott, J. (2006). Shocks and their consequences across and within households in Rural Zimbabwe. *Journal of Development Studies*, pages 301–321.

Hoddinott, J. and Kinsey, B. (2001). Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, 63(4):409–436.

Hodler, R. and Raschky, P. A. (2014). Economic shocks and civil conflict at the regional level. *Economics Letters*, 124(3):530–533.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of Sciences*, 107(35):15367–15372.

Hsiang, S. M. and Burke, M. (2013). Climate, conflict, and social stability: what does the evidence say? *Climatic Change*, 123(1):39–55.

Hultgren, A., Carleton, T., Delgado, M., Gergel, D. R., Greenstone, M., Houser, T., Hsiang, S., Jina, A., Kopp, R. E., Malevich, S. B., et al. (2022). Estimating global impacts to agriculture from climate change accounting for adaptation. *Available at SSRN 4222020*.

İlkkaracan, İ. and Memiş, E. (2021). Transformations in the gender gaps in paid and unpaid work during the COVID-19 pandemic: Findings from Turkey. *Feminist Economics*, 27(1-2):288–309.

ILO (2017). World employment social outlook: Trends for women. Technical report, International Labour Organization.

ILO (2021). ILOSTAT database. International Labour Organization.

International Food Policy Research Institute (IFPRI) (2020). AReNA's DHS-GIS database v1. Harvard Dataverse, https://doi.org/10.7910/DVN/OQIPRW.

Ito, S., Shonchoy, A., et al. (2020). Seasonality, academic calendar and school drop-outs in developing countries. Florida International University Working Paper.

Ito, T. and Kurosaki, T. (2009). Weather risk, wages in kind, and the off-farm labor supply of agricultural households in a developing country. *American Journal of Agricultural Aconomics*, 91(3):697–710.

Iyigun, M., Nunn, N., and Qian, N. (2017). The long-run effects of agricultural productivity on conflict, 1400-1900. National Bureau of Economic Research WP24066.

Jakiela, P., Ozier, O., Fernald, L., and Knauer, H. (2020). Big sisters. World Bank WP 9454.

Jaume, D. and Willén, A. (2021). The effect of teacher strikes on parents. *Journal of Development Economics*, 152:102679.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Joffe, S. (2001). Economic and policy issues in desert locust management: A preliminary analysis. FAO/EMPRES Workshop on Economics in Desert Locust Management.

Kadzamira, E. and Rose, P. (2003). Can free primary education meet the needs of the poor? evidence from malawi. *International Journal of Educational Development*, 23(5):501–516.

Kah, H. K. (2012). Husbands in wives' shoes: Changing social roles in child care among cameroon's urban residents. *Africa Development*, 37(3):101–114.

Karim, A. and Noy, I. (2016). Poverty and natural disasters: a regression meta-analysis. *Review of Economics and Institutions*, 7(2):26.

Kielland, A. and Tovo, M. C. (2006). *Children at work: Child labor practices in Africa*. Lynne Rienner Publishers Boulder, CO.

Kijima, Y., Matsumoto, T., and Yamano, T. (2006). Nonfarm employment, agricultural shocks, and poverty dynamics: evidence from rural uganda. *Agricultural Economics*, 35:459–467.

Kirchberger, M. (2017). Natural disasters and labor markets. *Journal of Development Economics*, 125:40–58.

Kleemans, M. and Magruder, J. (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal*, 128(613):2032–2065.

Klomp, J. and Bulte, E. (2013). Climate change, weather shocks, and violent conflict: A critical look at the evidence. *Agricultural Economics*, 44(s1):63–78.

Klomp, J. and Valckx, K. (2014). Natural disasters and economic growth: A meta-analysis. *Global Environmental Change*, 26:183–195.

Kochar, A. (1995). Explaining household vulnerability to idiosyncratic income shocks. *The American Economic Review*, 85(2):159–164.

Kochar, A. (1999). Smoothing consumption by smoothing income: hours-of-work responses to idiosyncratic agricultural shocks in rural india. *Review of Economics and Statistics*, 81(1):50–61.

Koren, O. (2018). Food abundance and violent conflict in africa. *American Journal of Agricultural Economics*, 100(4):981–1006.

Koubi, V. (2019). Climate change and conflict. *Annual Review of Political Science*, 22:343–360.

Krall, S. and Herok, C. (1997). Economics of desert locust control. In *New Strategies in Locust Control*, pages 401–413. Springer.

Kubik, Z. and Maurel, M. (2016). Weather shocks, agricultural production and migration: Evidence from tanzania. *The Journal of Development Studies*, 52(5):665–680.

Kugler, M., Viollaz, M., Duque, D., Gaddis, I., Newhouse, D. L., Palacios-Lopez, A., and Weber, M. (2021). How did the COVID-19 crisis affect different types of workers in the developing world? World Bank Policy Research Working Paper 9703.

Latchininsky, A. V. (2013). Locusts and remote sensing: a review. *Journal of Applied Remote Sensing*, 7(1):075099–075099.

Lazar, M., Piou, C., Doumandji-Mitiche, B., and Lecoq, M. (2016). Importance of solitarious desert locust population dynamics: lessons from historical survey data in Algeria. *Entomologia Experimentalis et Applicata*, 161(3):168–180.

Le, K. and Nguyen, M. (2022). Desert locust swarms and child health. *Economics & Human Biology*, 44:101094.

Lecoq, M. (2003). Desert locust threat to agricultural development and food security and fao/international role in its control. In *Presented at the Eighth Arab Congress of Plant Protection*.

Levison, D. and Moe, K. (1998). Household work as a deterrent to schooling: An analysis of adolescent girls in peru. *The Journal of Developing Areas*, 32(3):339–356.

Linnros, E. (2017). Plant pests and child health: evidence from locust infestations in west africa. Working Paper, Expert Group for Aid Studies (EBA).

Liu, M., Shamdasani, Y., and Taraz, V. (2023). Climate change and labor reallocation: Evidence from six decades of the indian census. *American Economic Journal: Economic Policy*, 15(2):395–423.

Liu, Y., Wei, S., and Xu, J. (2021). COVID-19 and women-led businesses around the world. *Finance Research Letters*, 43:102012.

Liu, Y. and Xue, Y. (2020). Expansion of the sahara desert and shrinking of frozen land of the arctic. *Scientific Reports*, 10(1):1–9.

Lo Bue, M. C., Le, T. T. N., Silva, M. S., and Sen, K. (2021). Gender and vulnerable employment in the developing world: Evidence from global micro-data. WIDER Working Paper 154/2021.

Lokshin, M. M., Glinskaya, E., and Garcia, M. (2000). The effect of early childhood development programs on women's labor force participation and older children's schooling in kenya. Technical report, The World Bank.

Lybbert, T. J., Barrett, C. B., Desta, S., and Layne Coppock, D. (2004). Stochastic wealth dynamics and risk management among a poor population. *The Economic Journal*, 114(498):750–777.

Ma, S., Sun, Z., and Xue, H. (2020). Childcare needs and parents' labor supply: Evidence from the COVID-19 lockdown. Working Paper, Harvard University.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–1026.

Mach, K. J., Adger, W. N., Buhaug, H., Burke, M., Fearon, J. D., Field, C. B., Hendrix, C. S., Kraan, C. M., Maystadt, J.-F., O'Loughlin, J., et al. (2020). Directions for research on climate and conflict. *Earth's Future*, 8(7):e2020EF001532.

Mach, K. J., Kraan, C. M., Adger, W. N., Buhaug, H., Burke, M., Fearon, J. D., Field, C. B., Hendrix, C. S., Maystadt, J.-F., O'Loughlin, J., et al. (2019). Climate as a risk factor for armed conflict. *Nature*, 571(7764):193–197.

Macours, K., Premand, P., and Vakis, R. (2022). Transfers, diversification and household risk strategies: Can productive safety nets help households manage climatic variability? *The Economic Journal*, 132(647):2438–2470.

Marending, M. and Tripodi, S. (2022). Gone with the wind: the welfare effect of desert locust outbreaks. Copenhagen Business School Working Paper.

Martinez, S., Naudeau, S., and Pereira, V. (2012). The promise of preschool in africa: A randomized impact evaluation of early childhood development in rural Mozambique. Technical report, 3ie Impact Evaluation 001.

Matsuura, M., Luh, Y.-H., and Islam, A. H. M. S. (2023). Weather shocks, livelihood diversification, and household food security: Empirical evidence from rural bangladesh. *Agricultural Economics*, 54(4):455–470.

Maystadt, J.-F., Mueller, V. A., and Sebastian, A. (2016). Environmental migration and labor markets in nepal. *Journal of the Association of Environmental and Resource Economists*, 3(2):417–452.

McCabe, B., Barboza, S., Basu, M., ohmann, L., Mwangi, E., Arango, M., Ambani, M., and Abdillahi, H. S. (2021). Desert locust and climate: A weather and bio-climatic case study of desert locust conditions in Northern Kenya - Kenya. ReliefWeb, https://reliefweb.int/report/kenya/technical-paper-desert-locust-and-climate-weather-and-bio-climatic-case-study-desert.

McGuirk, E. and Burke, M. (2020). The economic origins of conflict in africa. *Journal of Political Economy*, 128(10):3940–3997.

McGuirk, E. F. and Nunn, N. (2021). Transhumant pastoralism, climate change, and conflict in africa. National Bureau of Economic Research WP28243.

Mellon, J. (2022). Rain, rain, go away: 192 potential exclusion-restriction violations for studies using weather as an instrumental variable. Available at SSRN 3715610.

Miguel, E., Satyanath, S., and Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112(4):725–753.

Morrissey, T. W. (2017). Child care and parent labor force participation: a review of the research literature. *Review of Economics of the Household*, 15(1):1–24.

Moyi, P. (2011). Child labor and school attendance in Kenya. *Educational Research and Reviews*, 6(1):26–35.

Mueller, V. A., Gray, C., and Hopping, D. (2020a). Climate-induced migration and unemployment in middle-income Africa. *Global Environmental Change*, 65:102183.

Mueller, V. A., Gray, C., and Kosec, K. (2014). Heat stress increases long-term human migration in rural Pakistan. *Nature Climate Change*, 4(3):182–185.

Mueller, V. A. and Osgood, D. E. (2009a). Long-term consequences of short-term precipitation shocks: evidence from Brazilian migrant households. *Agricultural Economics*, 40(5):573–586.

Mueller, V. A. and Osgood, D. E. (2009b). Long-term impacts of droughts on labour markets in developing countries: Evidence from Brazil. *The Journal of Development Studies*, 45(10):1651–1662.

Mueller, V. A. and Quisumbing, A. (2011). How resilient are labour markets to natural disasters? the case of the 1998 Bangladesh flood. *Journal of Development Studies*, 47(12):1954–1971.

Mueller, V. A., Sheriff, G., Dou, X., and Gray, C. (2020b). Temporary migration and climate variation in eastern Africa. *World Development*, 126:104704.

Murungi, C. G. (2013). Reasons for low enrolments in early childhood education in Kenya: The parental perspective. *International Journal of Education and Research*, 1(5):1–10.

NASA (2024). NRT MODIS/AquaTerra Global Flood Product MCDWD_L3_NRT distributed from NASA LANCE. Technical report, Available on-line [https://www.earthdata.nasa.gov/global-flood-product]. DOI: 10.5067/MODIS/MCDWD_L3_NRT.061.

Newsom, A., Koli, M., and Sebesvari, Z. (2021). Locust outbreak 2019-2021. United Nations University Technical Report.

Nigerian Hydrological Services Agency (2021). Annual flood outlook. Technical report, Abuja, Nigeria.

Nigerian Hydrological Services Agency (2023). Annual flood outlook. Technical report, Abuja, Nigeria.

Noack, F., Riekhof, M.-C., and Di Falco, S. (2019). Droughts, biodiversity, and rural incomes in the tropics. *Journal of the Association of Environmental and Resource Economists*, 6(4):823–852.

Nordhaus, W. D. (2006). Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Sciences*, 103(10):3510–3517.

Nunn, N. and Qian, N. (2014). Us food aid and civil conflict. *American Economic Review*, 104(6):1630–1666.

Oguntola, T. (2022). 2022 flood: 603 dead, 1.3m displaced across nigeria – federal govt. Technical report, https://leadership.ng/2022-flood-603-dead-1-3m-displaced-across-nigeria-federal-govt/.

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., and Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4):306–312.

Pape, U. J. (2021). Kenya COVID-19 rapid response phone survey households 2020-2021, panel. The World Bank.

Pape, U. J., Baraibar Molina, J., Delius, A. J. S., Gitau, C. L., and Rios Rivera, L. A. (2021). Socio-economic impacts of COVID-19 in Kenya on households: Rapid response phone survey round 1. The World Bank.

Patel, D. (2023). Floods. Technical report, Harvard University Working Paper.

Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633):418–422.

Pörtner, H.-O., Roberts, D. C., Adams, H., Adler, C., Aldunce, P., Ali, E., Begum, R. A., Betts, R., Kerr, R. B., Biesbroek, R., et al. (2022). Climate change 2022: Impacts, adaptation and vulnerability. *IPCC Sixth Assessment Report*.

Prados, M. J. and Zamarro, G. (2021). School re-openings, childcare arrangements, and labor outcomes during COVID-19. Working Paper, University of Southern California.

Qiu, J. (2009). Global warming may worsen locust swarms . *Nature*.

Quisumbing, A. R., Hallman, K., and Ruel, M. T. (2007). Maquiladoras and market mamas: Women's work and childcare in Guatemala City and Accra. *The Journal of Development Studies*, 43(3):420–455.

Qureshi, J. A. (2018). Additional returns to investing in girls' education: Impact on younger sibling human capital. *The Economic Journal*, 128(616):3285–3319.

Rahimi, J., Haas, E., Grote, R., Kraus, D., Smerald, A., Laux, P., Goopy, J., and Butterbach-Bahl, K. (2021). Beyond livestock carrying capacity in the sahelian and sudanian zones of west africa. *Scientific Reports*, 11(1):1–15.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: an armed conflict location and event dataset: special data feature. *Journal of Peace Research*, 47(5):651–660.

Ramankutty, N., Evan, A. T., Monfreda, C., and Foley, J. A. (2010). Global agricultural lands: Croplands/pastures, 2000. Technical report, Socioeconomic Data and Applications Center (SEDAC), http://sedac.ciesin.columbia.edu/es/aglands.html.

Renier, C., Waldner, F., Jacques, D. C., Babah Ebbe, M. A., Cressman, K., and Defourny, P. (2015). A dynamic vegetation senescence indicator for near-real-time desert locust habitat monitoring with modis. *Remote Sensing*, 7(6):7545–7570.

Ribar, D. C. (1992). Child care and the labor supply of married women: Reduced form evidence. *Journal of Human Resources*, pages 134–165.

Rose, E. (2001). Ex ante and ex post labor supply response to risk in a low-income area. *Journal of Development Economics*, 64(2):371–388.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Sajid, O. and Bevis, L. E. (2021). Flooding and child health: Evidence from pakistan. *World Development*, 146:105477.

Salazar, C., Ayalew, H., and Fisker, P. (2019). Weather shocks and spatial market efficiency: evidence from mozambique. *The Journal of Development Studies*, 55(9):1967–1982.

Samil, H. M. O. A., Martin, A., Jain, A. K., Amin, S., and Kahou, S. E. (2020). Predicting regional locust swarm distribution with recurrent neural networks. *arXiv preprint arXiv:2011.14371*.

Samman, E., Presler-Marshall, E., Jones, N., Stavropoulou, M., and Wallace, J. (2016). Women's work: Mothers, children and the global childcare crisis. report. Technical report, Overseas Development Institute (ODI).

Sarsons, H. (2015). Rainfall and conflict: A cautionary tale. *Journal of Development Economics*, 115:62–72.

Showler, A. T. (2019). Desert locust control: the effectiveness of proactive interventions and the goal of outbreak prevention. *American Entomologist*, 65(3):180–191.

Showler, A. T. and Lecoq, M. (2021). Incidence and ramifications of armed conflict in countries with major desert locust breeding areas. *Agronomy*, 11(1):114.

Statista (2023). Nigeria: Distribution of gross domestic product (gdp) across economic sectors from 2011 to 2021. Technical report, statista.com/statistics/382311/nigeria-gdp-distribution-across-economic-sectors/.

Stein, W. and Weisser, R. A. (2022). Direct shock experience vs. tangential shock exposure: Indirect effects of flood shocks on well-being and preferences. *The World Bank Economic Review*, 36(4):909–933.

Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M. E., and Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services and Outcomes Research Methodology*, 14:166–182.

Sundberg, R. and Melander, E. (2013). Introducing the UCDP georeferenced event dataset. *Journal of Peace Research*, 50(4):523–532.

Symmons, P. and Cressman, K. (2001). Desert locust guidelines: biology and behaviour. Food and Agriculture Organization of the United Nations.

Takeshima, H., Amare, M., Mavrotas, G., et al. (2018). The role of agricultural productivity in non-farm activities in nigeria: effects on sector orientation and factor intensity. IFPRI Discussion Papers 1761.

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., and Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature*, 596(7870):80–86.

The Star (2020). Short notice: Rush against time for parents and candidates as cs orders phased resumption of classes next week.

The United States Department of Agriculture (USDA) (2022). Crop calendar charts. Foreign Agricultural Service, International Production Assessment Division. https://ipad.fas.usda.gov/ogamaps/cropcalendar.aspx.

The World Bank (2020). The locust crisis: The world bank's response. Fact-sheet. https://www.worldbank.org/en/news/factsheet/2020/04/27/the-locust-crisis-the-world-banks-response.

Thomson, A. and Miers, H. (2002). Assessment of the socio-economic impact of desert locusts and their control. *UK Department for International Development: London, UK*, page 37.

Tollefsen, A. F., Strand, H., and Buhaug, H. (2012). Prio-grid: A unified spatial data structure. *Journal of Peace Research*, 49(2):363–374.

Torngren Wartin, A. S. (2018). The sound of their wings: Desert locusts and conflicts in africa. Stockholm University Working Paper. Available at SSRN: https://ssrn.com/abstract=4467022.

Torres, J., Maduko, F., Gaddis, I., Iacovone, L., and Beegle, K. (2021). The impact of the COVID-19 pandemic on women-led businesses. World Bank Policy Research Working Paper 9817.

Townsend, R. M. (1995). Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *Journal of Economic Perspectives*, 9(3):83–102.

Ubilava, D., Hastings, J. V., and Atalay, K. (2022). Agricultural windfalls and the seasonality of political violence in africa. *American Journal of Agricultural Economics*, DOI: 10.1111/ajae.12364.

Udry, C. (2006). Child labor. *Understanding poverty*, 4:3607–3709.

Unah, L. (2021). Flood-ridden nigeria farmers need more help adapting to climate change. Technical report, The New Humanitarian.

United Nations Population Division (2020). Database on household size and composition 2019. United Nations.

Vitellozzi, S. and Giannelli, G. C. (2023). Thriving in the rain: natural shocks, time allocation, and empowerment in bangladesh. Technical report, IZA DP No. 16030h.

Von Uexkull, N. (2014). Sustained drought, vulnerability and civil conflict in sub-saharan africa. *Political Geography*, 43:16–26.

Von Uexkull, N., Croicu, M., Fjelde, H., and Buhaug, H. (2016). Civil conflict sensitivity to growing-season drought. *Proceedings of the National Academy of Sciences*, 113(44):12391–12396.

Wenham, C., Smith, J., Davies, S. E., Feng, H., Grépin, K. A., Harman, S., Herten-Crabb, A., and Morgan, R. (2020). Women are most affected by pandemics—lessons from past outbreaks. Nature Publishing Group, Comment.

Witsenburg, K. M. and Adano, W. R. (2009). Of rain and raids: Violent livestock raiding in northern kenya. *Civil Wars*, 11(4):514–538.

World Meteorological Association et al. (2021). Wmo atlas of mortality and economic losses from weather, climate and water extremes (1970–2019). Technical report, World Meteorological Association.

Xiong, J., Thenkabail, P. S., Gumma, M. K., Teluguntla, P., Poehnelt, J., Congalton, R. G., Yadav, K., and Thau, D. (2017). Automated cropland mapping of continental africa using google earth engine cloud computing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 126:225–244.

Youngblood, J. P., Cease, A. J., Talal, S., Copa, F., Medina, H. E., Rojas, J. E., Trumper, E. V., Angilletta, M. J., and Harrison, J. F. (2023). Climate change expected to improve digestive rate and trigger range expansion in outbreaking locusts. *Ecol. Monogr.*, 93(1):e1550.

Zamarro, G. and Prados, M. J. (2021). Gender differences in couples' division of childcare, work and mental health during COVID-19. *Review of Economics of the Household*, 19(1):11–40.

Zhang, L., Lecoq, M., Latchininsky, A., Hunter, D., et al. (2019). Locust and grasshopper management. *Annual Review of Entomology*, 64(1):15–34.

# Appendix A

# Agricultural shocks and long-term conflict risk: Evidence from desert locust swarms – Appendix

## A.1 Additional figures and tables

## Table A.1.1: Summary statistics

Panel A: Yearly variables

| | Mean | SD | Min | $25^{th}$ | $50^{th}$ | $75^{th}$ | Max | N |
|---|---|---|---|---|---|---|---|---|
| Any violent conflict event - ACLED | 0.02 | 0.14 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 538086 |
| Any protest or riot event - ACLED | 0.01 | 0.10 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 538086 |
| Any violent conflict event - UCDP | 0.01 | 0.10 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 538086 |
| Any swarm in cell | 0.00 | 0.07 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 538086 |
| Any swarm within 100km outside cell | 0.04 | 0.21 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 538086 |
| Any swarm within 100-250km of cell | 0.11 | 0.31 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 538086 |
| Population (10,000s) | 1.70 | 9.10 | 0.0 | 0.0 | 0.2 | 1.0 | 749.8 | 538083 |
| Total annual rainfall (100 mm) | 2.47 | 3.80 | 0.0 | 0.3 | 0.9 | 3.0 | 43.4 | 532498 |
| Max annual temperature (deg C) | 37.54 | 5.17 | 12.4 | 33.8 | 38.1 | 41.4 | 49.0 | 532498 |

Panel B: Fixed variables

| | Mean | SD | Min | $25^{th}$ | $50^{th}$ | $75^{th}$ | Max | N |
|---|---|---|---|---|---|---|---|---|
| Any ACLED violent conflict event in cell in any year | 0.14 | 0.35 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 24460 |
| Any protest/riot event in cell in any year | 0.08 | 0.27 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 24460 |
| Any UCDP violent conflict event in cell in any year | 0.08 | 0.27 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 24460 |
| Any locust swarm reported, 1985-2021 | 0.18 | 0.38 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 24460 |
| Any swarms within 100 km in any year | 0.63 | 0.48 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 | 24460 |
| First exposed to locust swarm in sample period (1997-2018) | 0.07 | 0.26 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 24460 |
| First exposed to locust swarm in 2003-2005 upsurge | 0.05 | 0.23 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 24460 |
| Any cropland or pasture in cell | 0.54 | 0.50 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 | 23923 |
| Share of crop and pasture land in cell | 0.24 | 0.32 | 0.0 | 0.0 | 0.0 | 0.5 | 1.0 | 23923 |
| Any pasture in cell | 0.53 | 0.50 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 | 23923 |
| Share of pasture in cell | 0.19 | 0.27 | 0.0 | 0.0 | 0.0 | 0.3 | 1.0 | 23923 |
| Any cropland in cell | 0.29 | 0.45 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 | 23923 |
| Share of cropland in cell | 0.05 | 0.13 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 23923 |

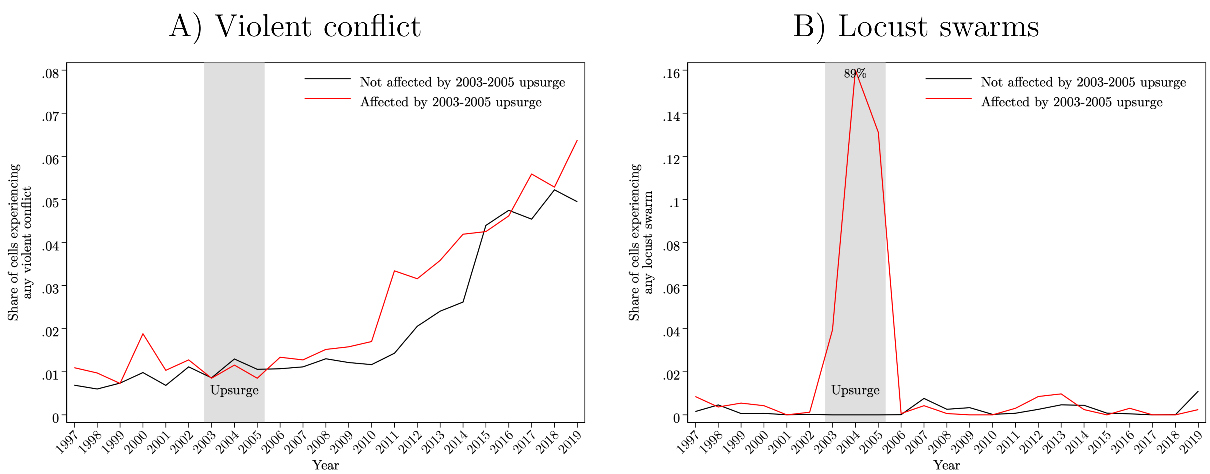
Note: Observations are grid cells approximately 28×28km by year.

Table A.1.2: Balance by exposure to any locust swarm during full sample period and during 2003-2005 upsurge

| | N | No sample period exposure Mean (SD) | Any sample period exposure Difference (SE) | N | No sample period exposure Mean (SD) | 2003-2005 swarm exposure Difference (SE) |
|---|---|---|---|---|---|---|
| Panel A: All cells | | | | | | |
| Population in 2000 | 23242 | 1.30 | 1.22** | 22748 | 1.30 | 1.07 |
| (10,000s) | | (6.39) | (0.61) | | (6.39) | (0.86) |
| Gross cell product | 21931 | 0.25 | 0.13 | 21513 | 0.25 | 0.16 |
| in 2005 (USD PPP) | | (0.95) | (0.09) | | (0.95) | (0.11) |
| Mean of cell | 22982 | 0.05 | 0.01* | 22490 | 0.05 | 0.01 |
| nightlights 1996-2012 (0-1) | | (0.04) | (0.00) | | (0.04) | (0.01) |
| First month rainy | 22982 | 6.29 | 0.40 | 22490 | 6.29 | 0.43 |
| season | | (3.33) | (0.26) | | (3.33) | (0.35) |
| Percent of cell | 22714 | 4.80 | 0.90 | 22225 | 4.80 | 1.41 |
| covered by crop land in 2000 | | (13.33) | (0.79) | | (13.33) | (0.99) |
| Percent of cell | 22714 | 17.95 | 10.68*** | 22225 | 17.95 | 11.66*** |
| covered by pasture land in 2000 | | (27.26) | (2.62) | | (27.26) | (2.98) |
| Mean annual rainfall | 22988 | 2.50 | 0.08 | 22494 | 2.50 | 0.06 |
| 1997-2018 (dm) | | (3.83) | (0.29) | | (3.83) | (0.36) |
| Mean annual max | 22988 | 37.60 | -1.41*** | 22494 | 37.60 | -1.21** |
| temperature 1997-2018 (deg C) | | (5.19) | (0.47) | | (5.19) | (0.60) |
| Joint significance | | | $F = 8.01$<br>$p < 0.001$ | | | $F = 3.29$<br>$p < 0.001$ |
| Panel B: Cells with agricultural land | | | | | | |
| Population in 2000 | 12093 | 2.30 | 0.83 | 11729 | 2.30 | 0.62 |
| (10,000s) | | (8.39) | (0.73) | | (8.39) | (1.05) |
| Gross cell product | 11170 | 0.32 | 0.10 | 10878 | 0.32 | 0.13 |
| in 2005 (USD PPP) | | (1.09) | (0.09) | | (1.09) | (0.12) |
| Mean of cell | 12077 | 0.05 | 0.01 | 11713 | 0.05 | 0.01 |
| nightlights 1996-2012 (0-1) | | (0.04) | (0.00) | | (0.04) | (0.01) |
| First month rainy | 12077 | 6.34 | 0.08 | 11713 | 6.34 | 0.13 |
| season | | (2.96) | (0.28) | | (2.96) | (0.38) |
| Percent of cell | 12093 | 9.35 | -1.88* | 11729 | 9.35 | -1.28 |
| covered by crop land in 2000 | | (17.43) | (1.08) | | (17.43) | (1.31) |
| Percent of cell | 12093 | 35.02 | 2.55 | 11729 | 35.02 | 3.49 |
| covered by pasture land in 2000 | | (29.18) | (2.18) | | (29.18) | (2.47) |
| Mean annual rainfall | 12093 | 4.35 | -1.21*** | 11729 | 4.35 | -1.28*** |
| 1997-2018 (dm) | | (4.55) | (0.34) | | (4.55) | (0.44) |
| Mean annual max | 12093 | 34.89 | 0.65 | 11729 | 34.89 | 1.15* |
| temperature 1997-2018 (deg C) | | (5.06) | (0.42) | | (5.06) | (0.61) |
| Joint significance | | | $F = 3.99$<br>$p < 0.001$ | | | $F = 2.25$<br>$p = 0.026$ |

Note: The table shows results from separate bivariate regressions of baseline or mean cell outcomes on locust swarm exposure. The rows indicate which dependent variable is used. The first set of columns compares cells where a swarm was first observed between 1998-2018 ('Any sample period exposure') to cells where no swarm was observed from 1990-2018. The second set of columns compares cells where a swarm was first observed during the major 2003-2005 locust upsurge ('2003-2005 swarm exposure') to the same control cells. Panel A includes all cells while Panel B includes only cells with agricultural land. I include results of joint tests that there is no relationship between any of the characteristics and swam exposure. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A.1.1: Trends in swarm and violent conflict events over time, by experience of 2003-2005 locust upsurge



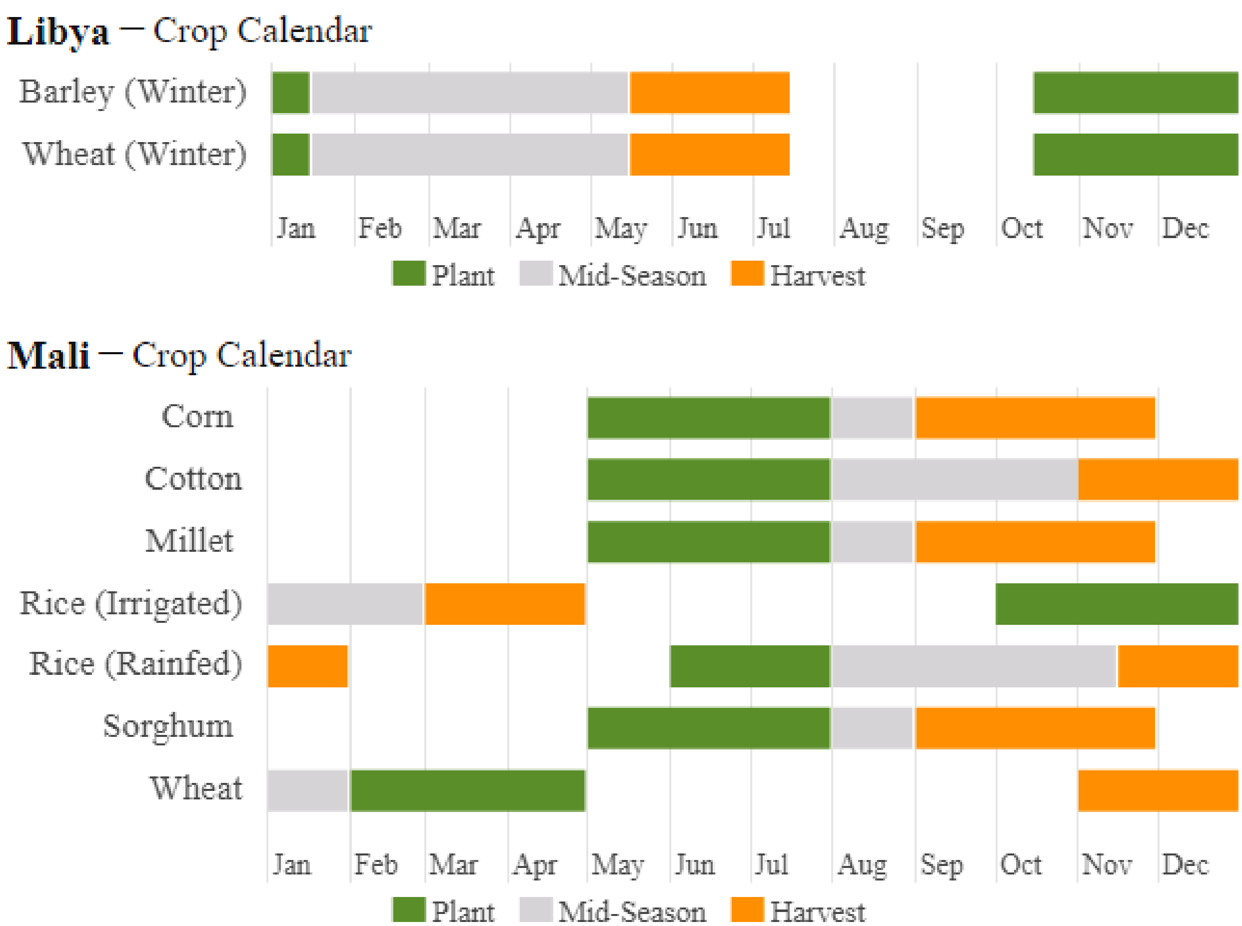

Note: The figures shows the share of cells experiencing any locust swarm or violent conflict event by year, separately for cells that did and did not experience any locust swarms during the 2003-2005 upsurge. Observations are grid cells approximately 28×28km by year.

Table A.1.3: Average impacts of exposure to locust swarms on violent conflict risk by land cover and specification

| Outcome: Any violent conflict event | All land | | | Land = Any crop/pasture land | | | Land = Any crop land | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Swarm exposure | Any | Any | Upsurge | Any | Any | Upsurge | Any | Any | Upsurge |
| Affected by any locust swarm | 0.021*** | 0.008* | 0.011* | 0.006 | -0.002 | 0.002 | 0.010** | -0.003 | -0.001 |
| | (0.006) | (0.004) | (0.006) | (0.005) | (0.003) | (0.007) | (0.005) | (0.003) | (0.005) |
| Total annual rainfall (100 mm) | 0.003*** | 0.005** | 0.004* | 0.003 | 0.006** | 0.006 | 0.003* | 0.005** | 0.004 |
| | (0.001) | (0.002) | (0.002) | (0.002) | (0.003) | (0.006) | (0.001) | (0.002) | (0.003) |
| Max annual temperature (deg C) | 0.005*** | 0.005*** | 0.007*** | 0.003** | 0.002 | 0.006** | 0.003** | 0.003* | 0.006** |
| | (0.002) | (0.002) | (0.003) | (0.001) | (0.002) | (0.003) | (0.001) | (0.002) | (0.003) |
| Population (10,000s) | 0.005*** | 0.007*** | 0.007*** | 0.004** | 0.010*** | 0.001 | 0.010*** | 0.014*** | 0.004 |
| | (0.001) | (0.001) | (0.001) | (0.002) | (0.003) | (0.002) | (0.003) | (0.003) | (0.003) |
| Affected by any locust swarm | | | | 0.018** | 0.017** | 0.014* | 0.027*** | 0.033** | 0.032** |
| × Land==1 | | | | (0.007) | (0.007) | (0.008) | (0.009) | (0.014) | (0.013) |
| Total annual rainfall (100 mm) | | | | 0.000 | -0.001 | -0.003 | 0.001 | 0.001 | -0.001 |
| × Land==1 | | | | (0.002) | (0.003) | (0.006) | (0.002) | (0.004) | (0.003) |
| Max annual temperature (deg C) | | | | 0.002 | 0.004** | 0.002 | 0.005** | 0.007** | 0.003 |
| × Land==1 | | | | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) | (0.004) |
| Population (10,000s) | | | | 0.002 | -0.004 | 0.006** | -0.004 | -0.008*** | 0.003 |
| × Land==1 | | | | (0.002) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Observations | 505679 | 473754 | 419748 | 499651 | 473754 | 419748 | 499651 | 473754 | 419748 |
| Outcome mean, no swarms & Land=1 | 0.024 | 0.018 | 0.019 | 0.042 | 0.033 | 0.033 | 0.056 | 0.046 | 0.046 |
| Proportional impact of exposure, Land=1 | 0.893 | 0.430 | 0.597 | 0.561 | 0.451 | 0.479 | 0.654 | 0.653 | 0.683 |
| Country-Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Cell FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Inverse propensity weights | No | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes |

Note: The dependent variable is a dummy for any violent conflict event observed. Cells are regarded as 'affected' by locust swarms for all years starting from the first year in which a swarm is recorded in the cell. In columns labeled 'Upsurge' only cells exposed during the 2003-2005 locust upsurge are included as treated; other exposed cells are excluded from the sample. The 'Land=1' rows show the coefficients for the interaction of right-hand side variables with cell land cover dummies indicated in the column heading. Inverse propensity weights are calculated separately based on estimates of the propensities of cells to have been exposed to any swarm and to the 2003-2005 upsurge. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

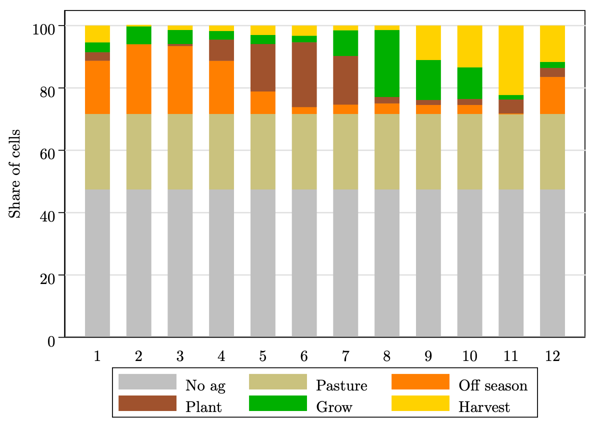
Figure A.1.2: Example crop calendars



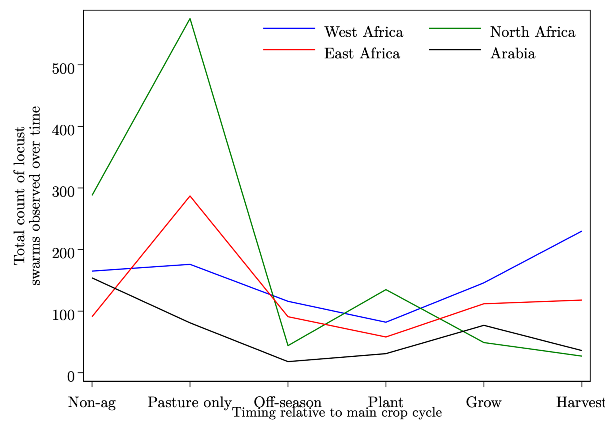

Source: U.S. Department of Agriculture Foreign Agricultural Service (USDA 2022).
Note: The Libya crop calendar is fairly representative of other North African countries, and the Mali crop calendar is fairly representative of other West African countries.

Figure A.1.3: Timing of locust swarm arrival by phase of crop calendar and region

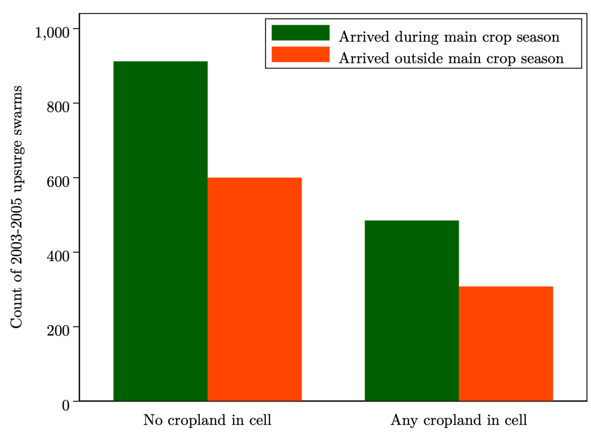
A) Agricultural activities by month of year




B) Timing of locust swarm arrival




C) Timing and locations of 2003-2005 upsurge swarms




Note: Agricultural activities by month are determined by assigning each cell with any crop land the primary activity for that month in the country in which it is located, using USDA 2022 crop calendars. Land cover in the year 2000 is from Ramankutty et al. (2010). Locust swarm observations are matched to agricultural activities based on the location and month of their arrival.

Table A.1.4: Average impacts of exposure to 2003-2005 upsurge swarms on conflict risk, by timing of upsurge swarms and land cover

| Outcome: Any violent conflict event | (1) All cells | (2) W/in 100km of upsurge |
|---|---|---|
| Affected by off season upsurge swarm | 0.010 | 0.009 |
| Non-crop cell | (0.010) | (0.010) |
| Affected by off season upsurge swarm | 0.016 | 0.016 |
| Crop cell | (0.013) | (0.013) |
| $p$, diff. in off season upsurge swarm effect | 0.604 | 0.526 |
| Affected by growing season upsurge swarm | -0.004 | -0.005 |
| Non-crop cell | (0.004) | (0.004) |
| Affected by growing season upsurge swarm | 0.031** | 0.040*** |
| Crop cell | (0.013) | (0.012) |
| $p$, diff. in growing season upsurge swarm effect | 0.012 | 0.001 |
| $p$, non-crop off season = growing season | 0.176 | 0.171 |
| $p$, crop off season = growing season | 0.343 | 0.145 |
| Observations | 400668 | 180999 |
| Outcome mean post-2005, no upsurge | 0.019 | 0.023 |
| Country-Year FE | Yes | Yes |
| Cell FE | Yes | Yes |
| Controls | Yes | Yes |
| Inverse propensity weights | Yes | Yes |

Note: The results in each column are from a single regression of a dummy for any violent conflict event on indicators of 2003-2005 upsurge swarm exposure by season and controls all interacted with a dummy for any crop land cover in a cell. The columns indicate the subset of cells considered. For each interaction, I show the coefficient for the swarm variable when there is no cropland, the sum of this coefficient and the interaction with land cover (with standard errors calculated using Stata's *xlincom* command), and the $p$-value for the test that the coefficient on the interaction is equal to 0. I also include $p$-values for the tests that the effects of upsurge swarm exposure in different seasons are the same by land cover. Cells are regarded as 'affected' by a 2003-2005 upsurge swarm for all years starting from the first year in which a swarm is recorded in the cell. Cells exposed in other years are not included. Inverse propensity weights calculated based on estimates of the propensities of cells to have been exposed to any swarm during the upsurge are included. Controls include annual total rainfall, maximum temperature, and population, and their interactions with the crop land indicator. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.1.5: Correlations between locust swarm exposure and international development flows to agriculture

| Outcome: Development flows to agriculture (2021 USD millions) | (1) | (2) |
|---|---|---|
| Any swarm in year | 93.4 | |
| | (145.8) | |
| Any swarm previous year | 185.0$^{*}$ | |
| | (92.7) | |
| Any swarm in year during upsurge | | 159.7 |
| | | (171.5) |
| Any swarm previous year during upsurge | | 312.9$^{**}$ |
| | | (141.8) |
| Observations | 936 | 936 |
| Outcome mean, no swarms | 932.8 | 965.8 |
| Country FE | Yes | Yes |
| Year FE | Yes | Yes |

Note: This table shows estimates of the correlation between any locust swarm being recorded in a country in a given year and total international development flows to agriculture received (FAOSTAT 2023b). The first two rows consider locust swarms recorded at any point in the sample period (1997-2018). The second two rows consider only locust swarms recorded during the major 2003-2005 locust upsurge which received significant international attention. Regressions are at the country-year level and include country and year fixed effects. SEs clustered at the country level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.1.6: Average impacts of exposure to agricultural shocks on violent conflict risk

| Outcome: Any violent conflict event | (1) Swarm | (2) Swarm | (3) Swarm | (4) Drought | (5) Drought | (6) Drought |
|---|---|---|---|---|---|---|
| Exposed to shock | 0.008* | -0.002 | 0.002 | 0.005** | -0.003 | 0.001 |
| | (0.004) | (0.003) | (0.002) | (0.002) | (0.002) | (0.001) |
| Exposed to shock × Any cropland or pasture in cell | | 0.017** | | | 0.015*** | |
| | | (0.007) | | | (0.004) | |
| Any violent conflict elsewhere in 1 degree cell | | | 0.140 | | | 0.053* |
| | | | (0.095) | | | (0.031) |
| Exposed to shock × Any violent conflict elsewhere in 1 degree cell=1 | | | 0.051*** | | | 0.028*** |
| | | | (0.015) | | | (0.007) |
| Observations | 473754 | 473754 | 473754 | 432843 | 429524 | 432843 |
| Cell FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Weather and population controls FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Inverse propensity weights FE | Yes | Yes | Yes | No | No | No |

Note: The dependent variable is a dummy for any violent conflict event observed. The columns indicate which agricultural shock is analyzed. Cells are regarded as 'affected' by a shock for all years starting from the first year in which the shock is recorded in the cell. Inverse propensity weights for swarm exposure are calculated based on estimates of the propensities of cells to have been exposed to any swarm. Columns with interactions also include interactions with other right-hand side variables. Observations are grid cells approximately 28×28km by year for 1997-2018 for swarms and 1997-2014 for drought. SEs clustered at the sub-national region level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.1.7: Impacts of agricultural shocks on the conflict risk, treating shocks as temporary vs. persistent

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Locust swarms | | Drought | |
| | Temporary effects | Persistent effects | Temporary effects | Persistent effects |
| Any swarm in cell | -0.015*** | | | |
| | (0.004) | | | |
| Any swarm in cell previous year | -0.006 | | | |
| | (0.004) | | | |
| Affected by any locust swarm | | 0.021*** | | |
| | | (0.005) | | |
| Any drought in year | | | 0.004 | |
| | | | (0.003) | |
| Any drought previous year | | | 0.002 | |
| | | | (0.002) | |
| Affected by any drought | | | | 0.005** |
| | | | | (0.002) |
| Year 0 event study estimate | -0.002 | | 0.009*** | |
| | (0.003) | | (0.003) | |
| $p$-value, equality of estimates | 0.009 | | 0.198 | |
| Observations | 508269 | 482691 | 408744 | 408752 |
| Country-Year FE | Yes | Yes | Yes | Yes |
| Cell FE | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes |

Note: The dependent variable is a dummy for any violent conflict event observed. Locust swarm and drought exposure in columns (1) and (3) is based only on whether the shock was observed in a particular year. Cells are considered 'affected' by any locust swarm or any drought in columns (2) and (4) for all years after these shocks are first observed in the cell. At the bottom of columns (1) and (3) I show coefficients for the treatment effect in the year of exposure from a persistent effects event study specification along with the $p$-value for the test of equality of the coefficients. Controls in all regressions include total cell population and current and prior year measures of total rainfall and maximum annual temperature. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

## A.2 Desert locusts background

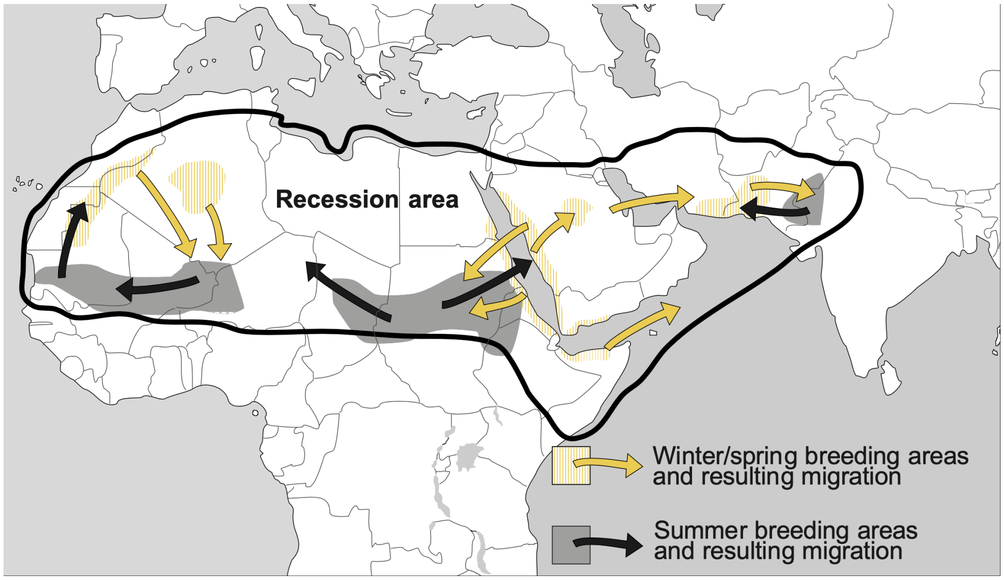
The desert locust is considered the world's most dangerous and destructive migratory pest (Cressman et al., 2016; Lazar et al., 2016). Locusts consume any available vegetation, and swarms frequently lead to the total destruction of local agricultural output (Showler, 2019). Damages from locust shocks can be extreme, with a small swarm covering one square kilometer can consume as much food in one day as 35,000 people. During the last locust upsurge in 2003-2005 in North and West Africa, 100, 90, and 85% losses on cereals, legumes, and pastures respectively were recorded, affecting more than 8 million people (Renier et al., 2015; Brader et al., 2006). Damages to crops alone were estimated at \$2.5 billion USD and \$450 million USD was required to bring an end to the upsurge (ASU, 2020).

In the most recent upsurge from 2019-2021 in East Africa and the Arabian Peninsula, over 40 million people in 10 countries faced severe food insecurity due to crop destruction. Locust control operations undertaken by the United Nations Food and Agriculture Organization (FAO) and its partners, primarily via ground and aerial spraying of pesticides, and global food aid efforts helped reduce the damages (Food and Agriculture Organization of the United Nations (FAO), 2022a). The FAO estimates that 3.5 million people were affected by locust destruction, but that control efforts saved agricultural production worth \$1.7 billion USD.

Small numbers of locusts are always present in desert 'recession' areas from Mauritania to India (Figure A.2.1). The population can grow exponentially under favorable climate conditions: periods of repeated rainfall and vegetation growth overlapping with the breeding cycle. The 2019-2021 upsurge persisted in large part because of repeated heavy precipitation out of season due to cyclones, prompting explosive reproduction (Cressman and Ferrand, 2021). The 2003-2005 upsurge was initiated by good rainfall over the summer of 2003 across four separate breeding areas. This was followed by two days of unusally heavy rains in October 2003 from Senegal to Morocco, after which environmental conditions were favorable for reproduction over the following 6 months (FAO and WMO, 2016).

Unique among grasshopper species, after reaching a particular population density desert locusts undergo a process of 'gregarization' wherein they mature physically and form large bands or swarms which move as a cohesive unit (Symmons and Cressman, 2001). Locust bands may extend over several kilometers and alternate between roosting and marching, typically downwind (FAO and WMO 2016). Locust swarms form when bands of locusts remain highly concentrated when they reach the adult stage and become able to fly. This formation of swarms can lead to 'outbreaks,' where locusts spread out from their largely desert initial breeding areas. Locusts in swarms have increased appetites and accelerated reproductive cycles, and are thus particularly threatening to agriculture. The FAO distinguishes different levels of locust swarm activity (Symmons and Cressman, 2001). I use the terms 'outbreak'

Figure A.2.1: Desert locust recession and breeding areas




Source: Symmons and Cressman (2001).

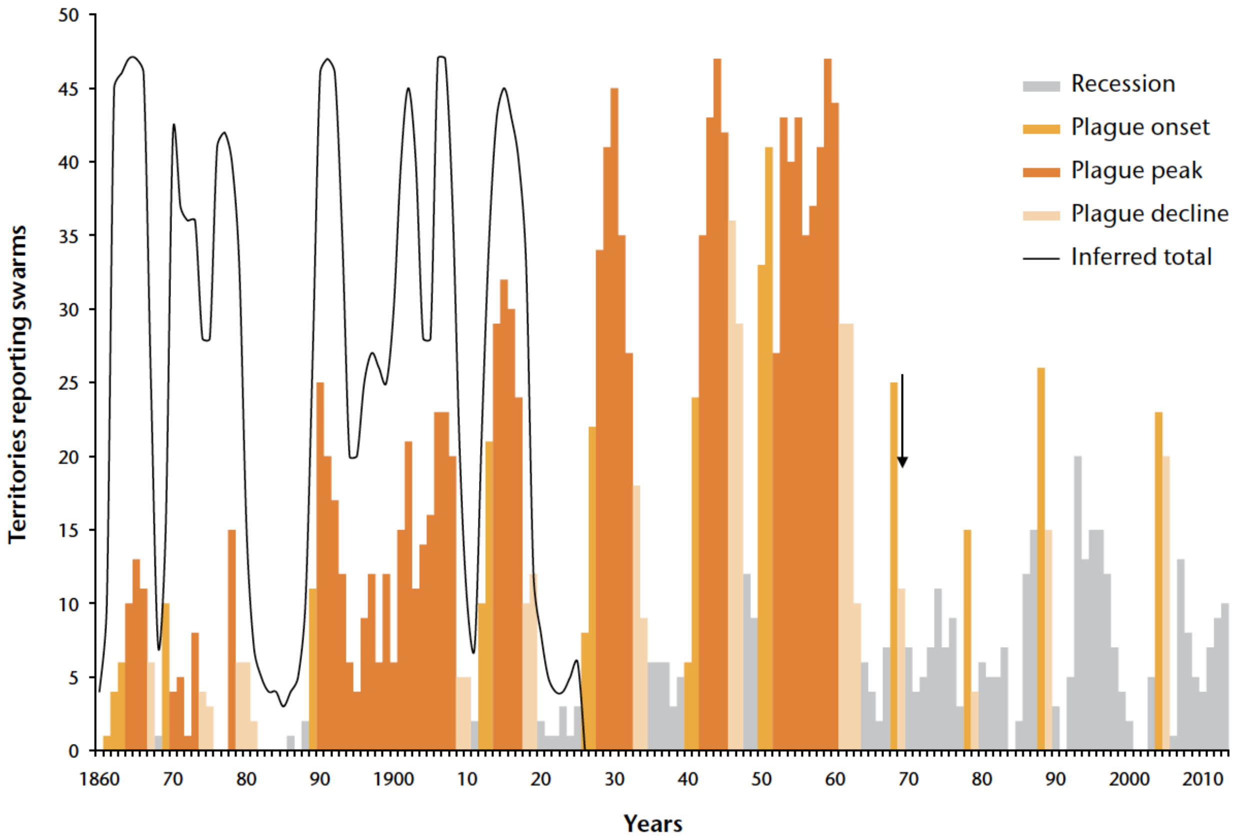
and 'upsurge' interchangeably to refer to any locust swarm activity. By the FAO definition 'outbreaks' refer to localized increases in locust numbers while 'upsurges' refer to broader and more sustained locust activities. A third level, 'plagues,' is characterized by larger and more widespread locust infestations. Few locust swarms are observed outside of major outbreaks, as conditions favoring swarm formation tend to produce large swarms which reproduce and spread rapidly and are very difficult to control.

As illustrated by Figure 1.1, locust swarms are not observed with any regularity over time or space. Desert locusts are migratory, moving on after consuming all available vegetation, and outside of outbreak periods are ultimately restricted to desert 'recession' areas. Unlike many other insect species, therefore, the arrival of a desert locust swarm does not signal a permanent change in local agricultural pest risk. Instead, the arrival of a swarm can be considered a locally and temporally concentrated natural disaster where all crops and pastureland are at risk (Hardeweg, 2001).

The frequency of large-scale outbreaks has fallen since around the 1980s (Figure A.2.2), in large part due to increases in coordinated preventive operations (Cressman and Stefanski, 2016), as shown by the figure below. Given their tolerance for extreme heat and responsiveness to periods of heavy precipitation, however, climate change might create conditions conducive to more frequent desert locust outbreaks.

Farmers have no proven effective recourse when faced with the arrival of a locust swarm,

Figure A.2.2: Desert locust observations by year




Source: Cressman and Stefanski (2016), Figure 6.

though activities such as setting fires, placing nets on crops, and making noise are commonly attempted. While these may slow damage they have little effect on locust population or total damages (Dobson, 2001; Hardeweg, 2001; Thomson and Miers, 2002). Locust outbreaks end due to a combination of migration to unfavorable habitats, failure of seasonal rains in breeding areas, and control operations (Symmons and Cressman, 2001). The only current viable method of swarm control is direct air or ground spraying with pesticides (Cressman and Ferrand, 2021). These control operations do not prevent immediate agricultural destruction as they take some time to kill the targeted locusts, but will limit their spread. The 2003-2005 locust upsurge ended due to lack of rain and colder temperatures which slowed down the breeding cycle, combined with intensive ground and aerial spraying operations which treated over 130,000km$^2$ at a cost of over US$400 million (FAO and WMO, 2016).

Desert locust control is most effective before locust populations surge, and the FAO manages an international network of early monitoring, warning, and prevention systems in support of this goal (Zhang et al., 2019). While improvements in desert locust management have been largely effective in reducing the frequency of outbreaks (as seen in Figure A.2.2), many challenges remain. Desert locust breeding areas are widespread and often in remote or insecure areas. Small breeding groups are easy to miss by monitors, and swarms can migrate quickly. In addition, control operations are slow and costly, resources for monitoring and control are limited outside of upsurges, and the cross-country nature of the thread creates coordination issues. Insecurity may also limit locust control activities (Showler and Lecoq, 2021).

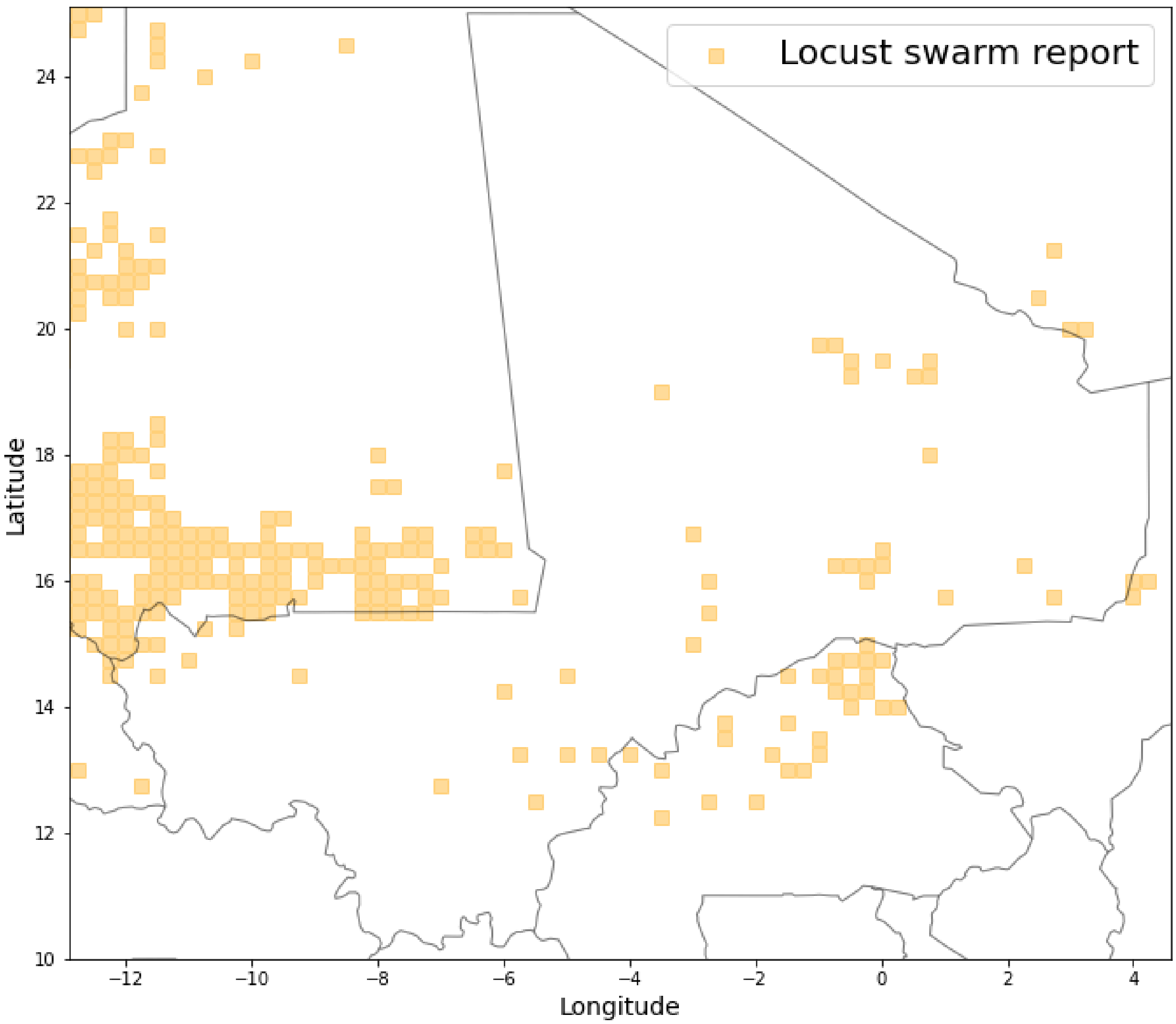
Locust swarms vary in their density and extent (Symmons and Cressman, 2001). The average swarm includes around 50 locusts per $m^2$ with a range from 20-150, and can cover under 1 square kilometers to several hundred (Symmons and Cressman, 2001). About half of swarms exceed 50km$^2$ in size (FAO and WMO 2016, meaning swarms typically include over a billion individuals. Swarms fly downwind from a few hours after sunrise to an hour or so before sunset when they land and feed. Swarms do not always fly with prevailing winds and may wait for warmer winds. Small deviations in the positions of individual locusts in the swarm can also lead to changes in swarm flight trajectory, making their movements difficult to predict. Seasonal changes in these winds tend to bring locusts to seasonal breeding areas at times when rain and the presence of vegetation is most likely, allowing them to continue breeding (FAO and WMO, 2016).

These movement characteristics inform efforts to predict locust swarm movements, but these remain highly imprecise. The desert locust bulletins produced monthly by the FAO include forecasts of areas at risk of desert locust activity, but the areas described are quite large, often encompassing several countries in periods with increased swarms. While breeding regions and the broad areas at risk over different time periods can generally be predicted with some accuracy (Latchininsky, 2013; Samil et al., 2020; Zhang et al., 2019), predicting specific local variation in swarm presence remains a challenge due to the multiple factors influencing specific flight patterns (FAO and WMO 2016).

Patterns in swarm movements lead to local variation in locust swarm exposure. After taking off, swarms fly for 9-10 hours rather than landing as soon as they encounter new vegetation. A swarm can easily move 100km or more in a day even with minimal wind (Symmons and Cressman, 2001). Consequently, the flight path of a locust swarm will include both affected and unaffected areas, with the affected areas determined by largely by patterns of wind direction and speed over time from the initial swarm formation in breeding areas. An important result of the local variation in locust swarm damages during outbreaks is that macro level impacts may be muted, since outbreaks occur in periods of positive rainfall shocks which tend to increase agricultural production in unaffected areas. Several studies find that impacts of locust outbreaks on national agricultural output and on prices are minimal, despite devastating losses in affected areas (Joffe, 2001; Krall and Herok, 1997; Showler, 2019; Thomson and Miers, 2002; Zhang et al., 2019).

Figure A.2.3 illustrates the variation in areas affected by locust swarms over space, showing reports of locust swarms in 2004—the peak of the 2003-2005 upsurge—around Mali. Swarm reports are densely clustered in the breeding areas in southern Mauritania where locust swarms reproduced in summer 2004. Outside of this area there is considerable variation in where swarms were reported, with distances between reported swarm over time consistent with typical flight distances. I leverage the quasi-random variation over both time and space in the areas affected by swarms to identify their impact on conflict.

Figure A.2.3: Reports of locust swarms around Mali in 2004




Note: The figure illustrates the grid cells in which locusts swarms were reported for the area around the country of Mali in 2004, the peak year of the 2003-2005 locust upsurge. Locust swarm reports are from the FAO Locust Watch database.

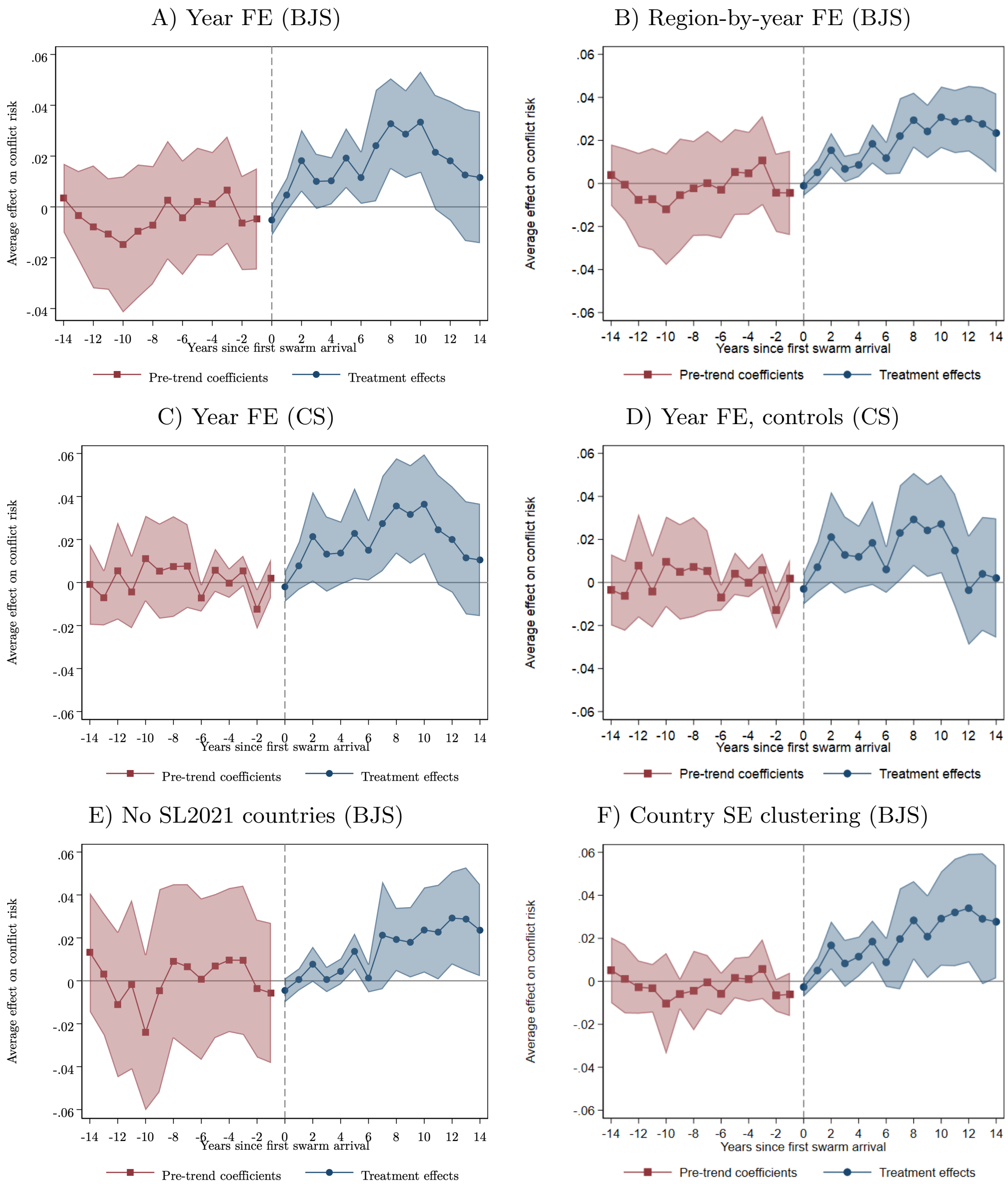
## A.3 Robustness

Table A.3.1: Balance by exposure to a locust swarm, including inverse propensity weights

| | N | Never swarm Mean (SD) | Any swarm treatment diff (SE) | N | Never swarm Mean (SD) | 2003-2005 swarm treatment diff (SE) |
|---|---|---|---|---|---|---|
| Population in 2000 | 21536 | 1.26 | 0.27 | 21147 | 1.26 | -0.36 |
| (10,000s) | | (6.16) | (0.35) | | (6.15) | (0.35) |
| Gross cell product | 21536 | 0.24 | -0.00 | 21147 | 0.24 | -0.07 |
| in 2005 (USD PPP) | | (0.92) | (0.07) | | (0.92) | (0.07) |
| Mean of cell | 21536 | 0.05 | 0.00 | 21147 | 0.05 | -0.00 |
| nightlights 1996-2012 (0-1) | | (0.04) | (0.00) | | (0.04) | (0.00) |
| First month rainy | 21536 | 6.21 | 0.13 | 21147 | 6.22 | 0.16 |
| season | | (3.34) | (0.21) | | (3.35) | (0.28) |
| Percent of cell | 21536 | 4.97 | -0.32 | 21147 | 4.97 | -1.56 |
| covered by crop land in 2000 | | (13.63) | (1.04) | | (13.62) | (1.31) |
| Percent of cell | 21536 | 16.42 | -2.46 | 21147 | 16.40 | -6.56* |
| covered by pasture land in 2000 | | (25.95) | (2.37) | | (25.94) | (3.52) |
| Mean annual rainfall | 21536 | 2.50 | 0.06 | 21147 | 2.50 | -0.65 |
| 1997-2018 (dm) | | (3.90) | (0.45) | | (3.90) | (0.56) |
| Mean annual max | 21536 | 37.88 | -0.27 | 21147 | 37.89 | 0.52 |
| temperature 1997-2018 (deg C) | | (5.12) | (0.61) | | (5.13) | (0.79) |
| | | | $F = 0.99$ | | | $F = 0.95$ |
| Joint significance | | | $p = 0.450$ | | | $p = 0.477$ |

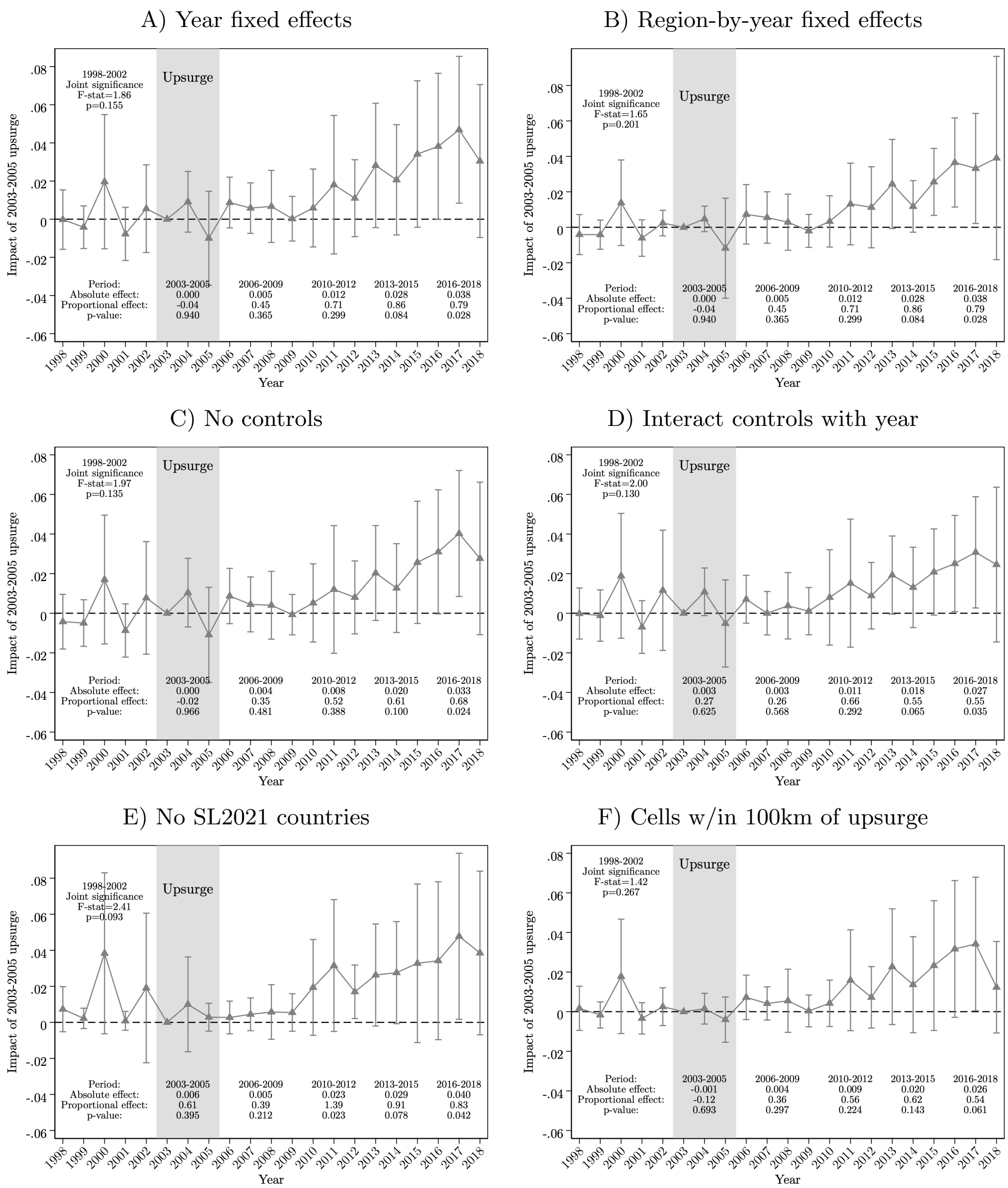
Note: The table shows results from separate bivariate regressions of baseline or mean cell outcomes on locust swarm exposure, including inverse propensity weights. The rows indicate which dependent variable is used. Inverse propensity weights are calculated based on estimates of the propensities of cells to have been exposed to any swarm during the study period and to have been exposed during the 2003-2005 upsurge. The first set of columns compares cells where a swarm was first observed between 1998-2018 ('Any sample period exposure') to cells where no swarm was observed from 1990-2018. The second set of columns compares cells where a swarm was first observed during the major 2003-2005 locust upsurge ('2003-2005 swarm exposure') to the same control cells. At the bottom are results of joint tests that there is no relationship between any of the characteristics and swam exposure. Observations are grid cells approximately 28×28km by year. SEs clustered at the sub-national region level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A.3.1: Sensitivity of locust swarm exposure event study results



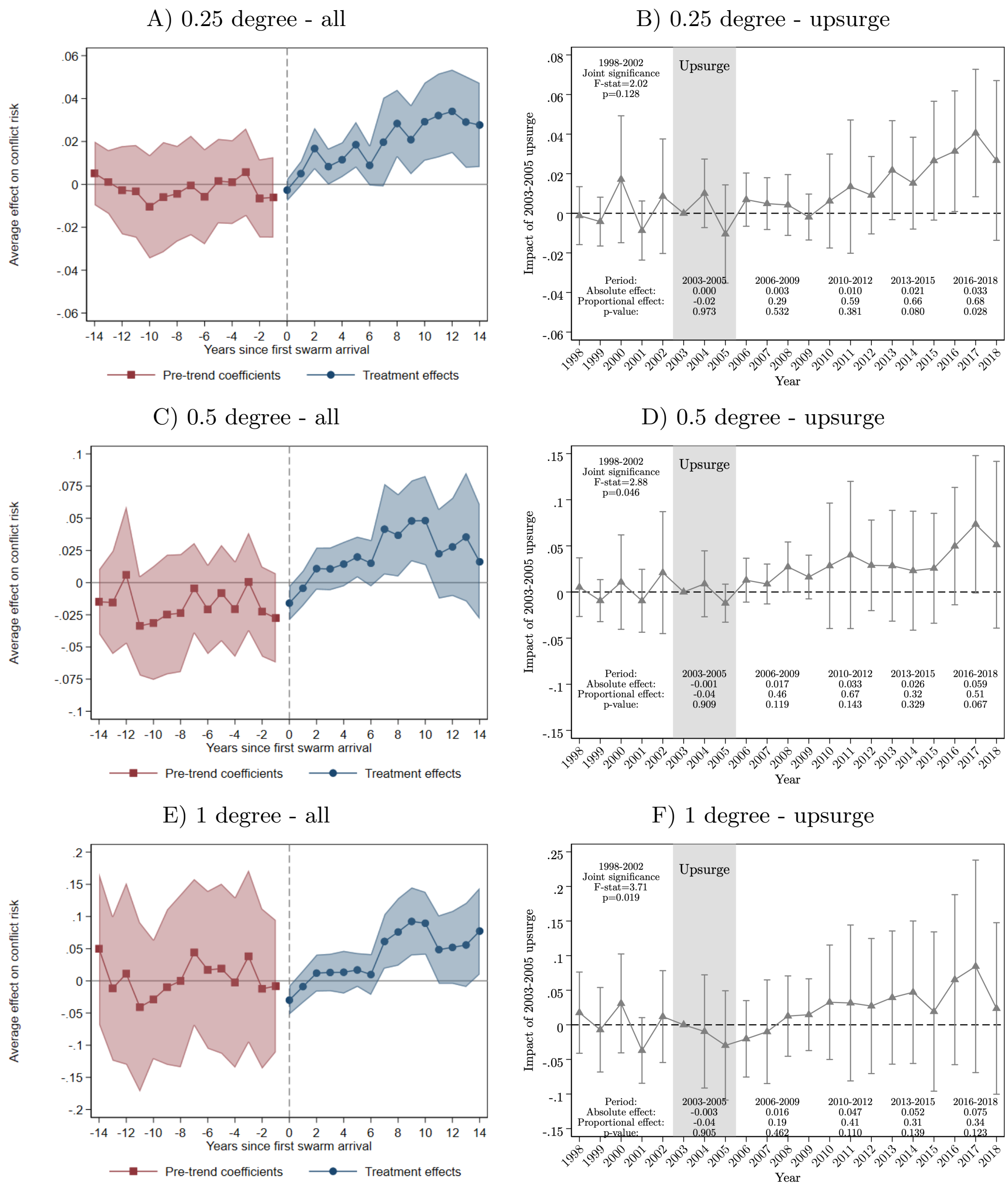

Each panel replicates Figure 1.3 Panel A but changes some aspect of the specification as indicated in the panel title. ‘BJS’ refers to the Borusyak et al. (2021) method. ‘CS’ refers to the Callaway and Sant’Anna (2021) method. The BJS method calculates all effects relative to the full pre-exposure period. The CS method calculates pre-trend effects using adjacent years and treatment effects relative to the year prior to exposure. ‘SL2021’ countries in panel E refers to the countries listed in Showler and Lecoq (2021) as areas where insecurity has limited desert locust control operations during the sample period. See the figure note for Figure 1.3 for more detail.

Figure A.3.2: Sensitivity of upsurge swarm exposure event study results



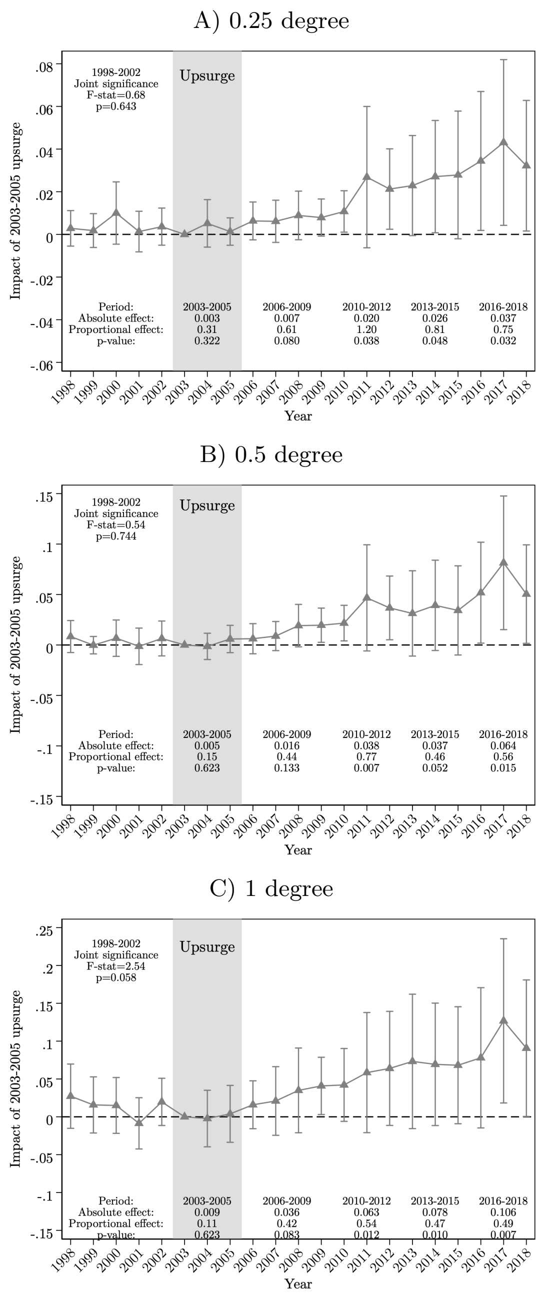

Note: Each panel replicates Figure 1.3 Panel B but changes some aspect of the specification as indicated in the panel title. 'SL2021' countries in panel E refers to the countries listed in Showler and Lecoq (2021) as areas where insecurity has limited desert locust control operations during the sample period. See the figure note for Figure 1.3 for more detail.

Figure A.3.3: Locust swarm exposure event study results by cell size



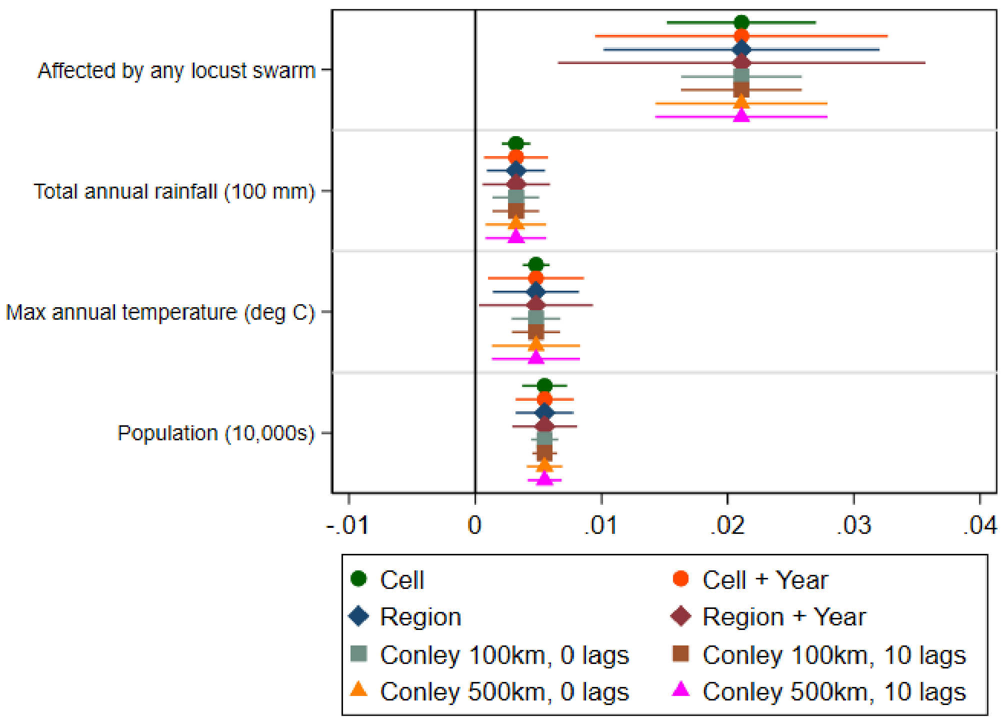

Note: Panels A, C, and E replicate Figure 1.3 Panel A with different cell sizes. Panels B, D, and F replicate Figure 1.3 Panel B with different cell sizes. See the figure notes for Figure 1.3 for more detail. When collapsing to larger cells I take the maximum of the swarm exposure and violent conflict variables and the mean of control variables across smaller cells within the aggregate cell. Shading and bars represent 95% confidence intervals using SEs clustered at the sub-national region level.

Figure A.3.4: Upsurge swarm exposure event study by cell size, no inverse propensity weights



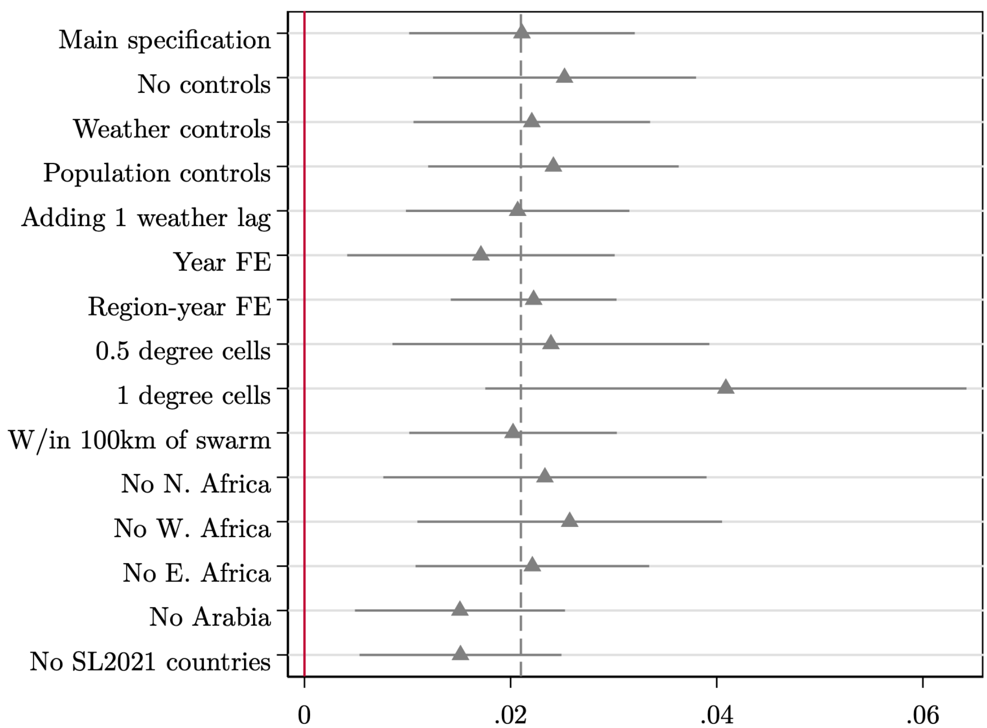

Note: The figure replicates Figure 1.3 Panel B but without inverse propensity weights and with different cell sizes. See the figure note for Figure 1.3 for more detail. When collapsing to larger cells I take the maximum of the swarm exposure and violent conflict variables and the mean of control variables across smaller cells within the aggregate cell.

Figure A.3.5: Estimated coefficients from Equation 1.1 with different SEs



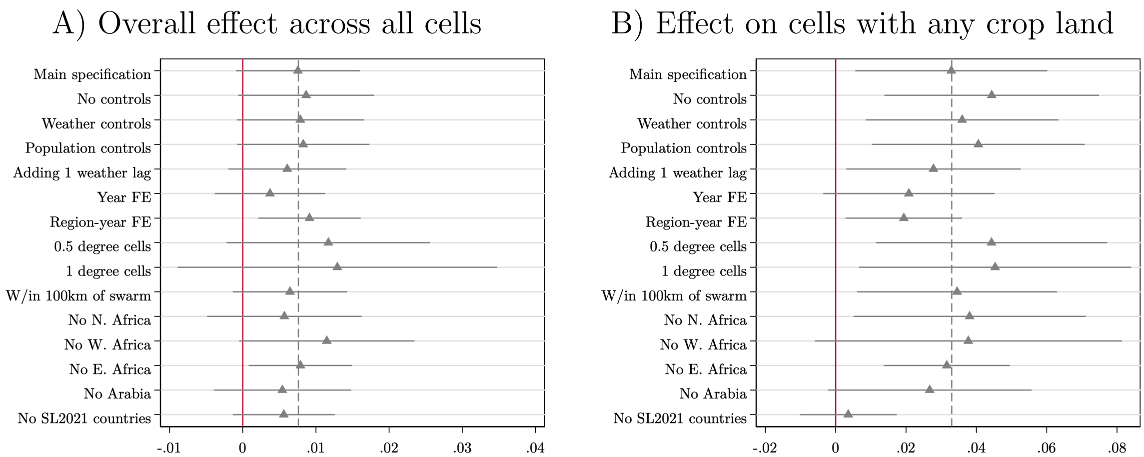

Note: The outcome variable is a dummy for any violent conflict observed. The figure shows 95% confidence intervals for estimates from Table A.1.3 column (1) applying different clustering for the SEs. Observations are grid cells approximately 28×28km by year. Regressions also include country-by-year and cell FE.

Figure A.3.6: Average impacts of locust swarm exposure on violent conflict risk, varying specifications and samples




Note: The outcome variable is a dummy for any violent conflict observed. The figure shows 95% confidence intervals for estimates from Table A.1.3 column (1) varying controls, fixed effects, cell size, and countries included in the sample. 'SL2021' countries are those listed in Showler and Lecoq (2021) as having areas where insecurity limited desert locust control operations during the sample period, and include Chad, Mali, Somalia, Sudan, Western Sahara, and Yemen. Observations are grid cells approximately 28×28km by year. Regressions all include cell FE as well as country-by-year FE unless otherwise stated. SEs are clustered at the sub-national region level.

Figure A.3.7: Average impacts of locust swarm exposure on violent conflict risk with inverse propensity weights, varying specifications and samples



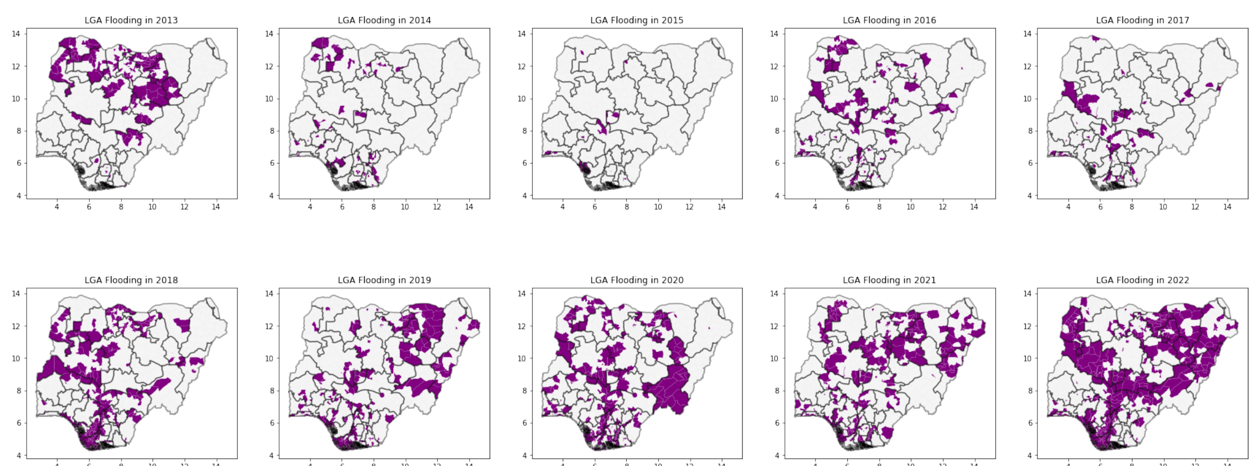

Note: The outcome variable is a dummy for any violent conflict observed. Panel A shows 95% confidence intervals for estimates from Table A.1.3 column (2) varying controls, fixed effects, cell size, and countries included in the sample. Panel B shows the same for the total effect in crop cells from Table A.1.3 column (9). 'SL2021' countries are those listed in Showler and Lecoq (2021) as having areas where insecurity limited desert locust control operations during the sample period, and include Chad, Mali, Somalia, Sudan, Western Sahara, and Yemen. Inverse propensity weights are applied in all regressions and are calculated separately based on estimates of the propensities of cells to have been exposed to any swarm and to the 2003-2005 upsurge. Observations are grid cells approximately 28×28km by year. Regressions all include cell FE as well as country-by-year FE unless otherwise stated. SEs are clustered at the sub-national region level. All estimates in Panel B except 'No SL2021 countries' are significant at a 10% confidence level or less.

# Appendix B

# Flooding and livelihood diversification in Nigeria: the view from the sky and the view from the ground

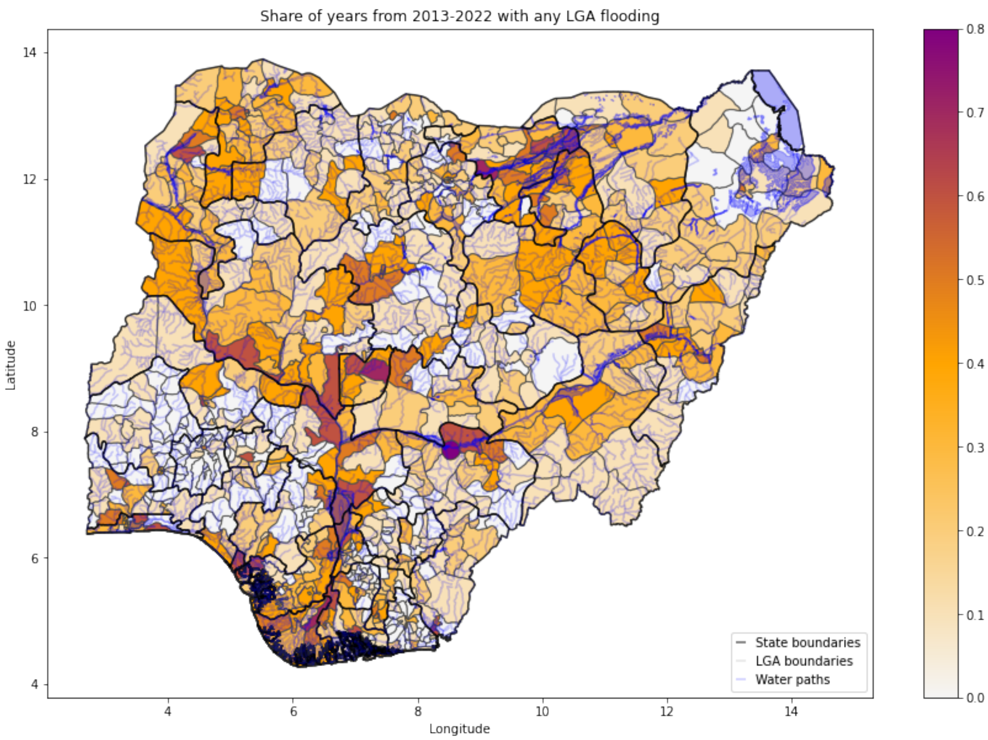
## B.1 Additional figures

Figure B.1.1: Local Government Areas reporting flood events by year



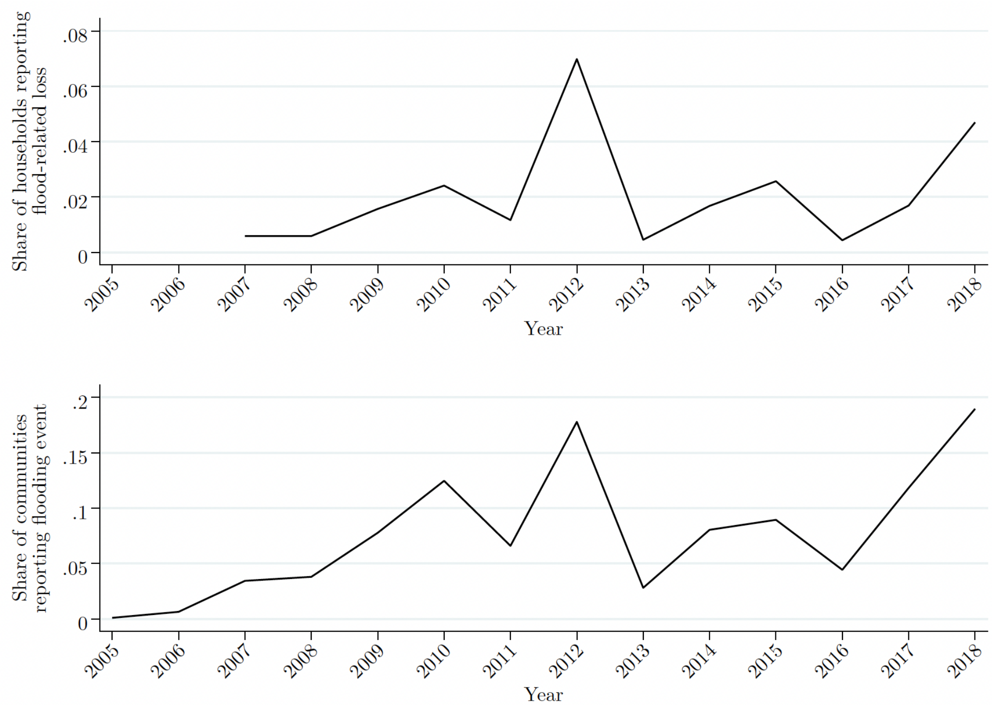

Note: Data on LGA-level flooding obtained by direct communication with the Nigeria Hydrological Services Agency.

Figure B.1.2: Share of years with any reported flood events from 2013-2022 by Local Government Areas




Note: Data on LGA-level flooding obtained by direct communication with the Nigeria Hydrological Services Agency.

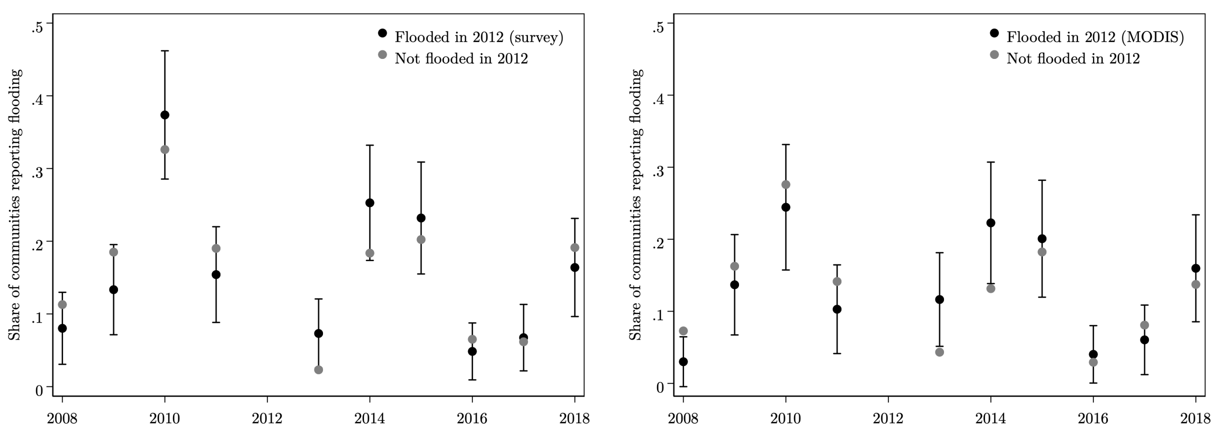
Figure B.1.3: GHSP household and community flooding reports by year




Note: Flood reports are from the Nigeria GHSP. In each survey round, households are asked to recall whether they experienced specific economic shocks over the last few years, including flooding that caused harvest failure and loss of property due to flood. Community informants are asked to recall major community events over the last few years, including floods. These data were collected in the post-harvest survey waves in 2009, 2013, 2016, and 2019.

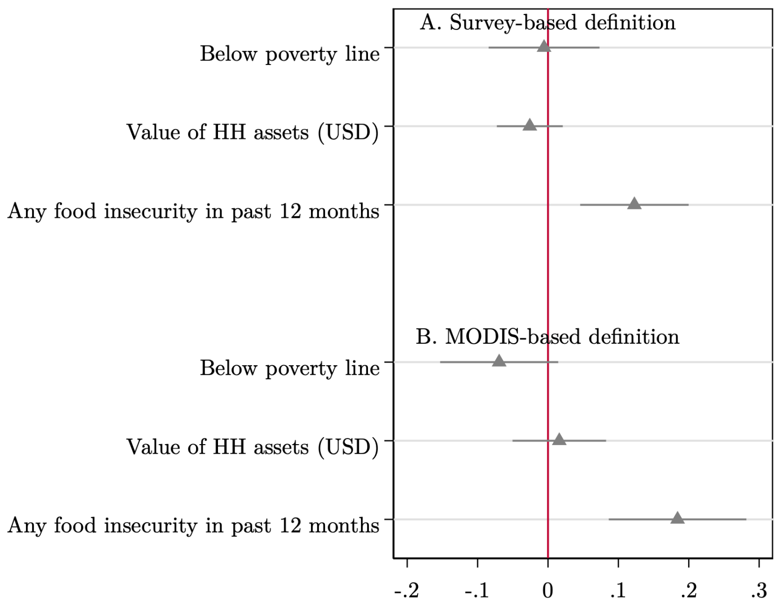
Figure B.1.4: Trends in community survey reports of flooding events over time by exposure to 2012 floods, including inverse propensity weights

A) Trends by survey-reported flooding B) Trends by proximity to MODIS flooding



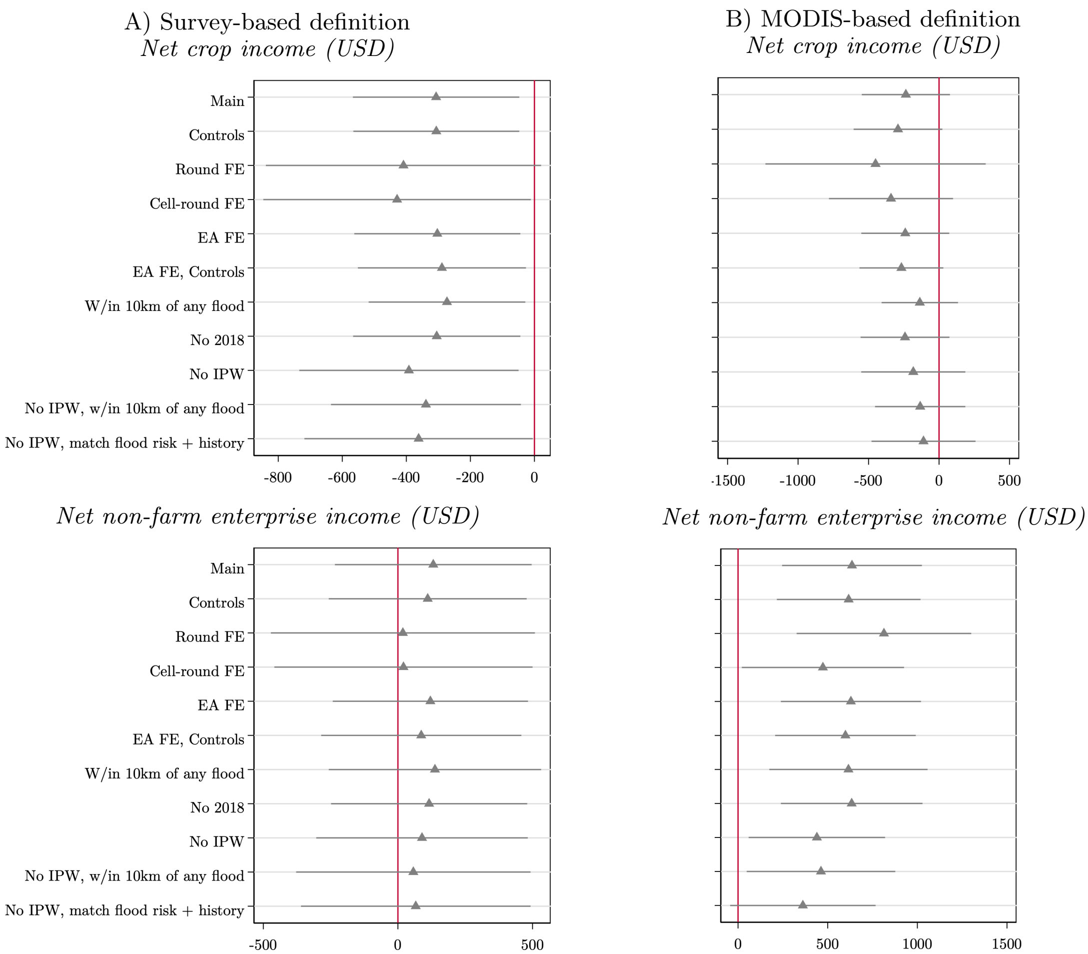

Note: The figures shows the share of communities reporting any flooding each year, weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. In Panel A, flood exposure and the sample weights are defined based on survey reports for the community of residence in 2012. In Panel B, they are defined based on the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data.

Figure B.1.5: Average impacts of flood exposure in 2012 on household well-being outcomes, by flooding definition



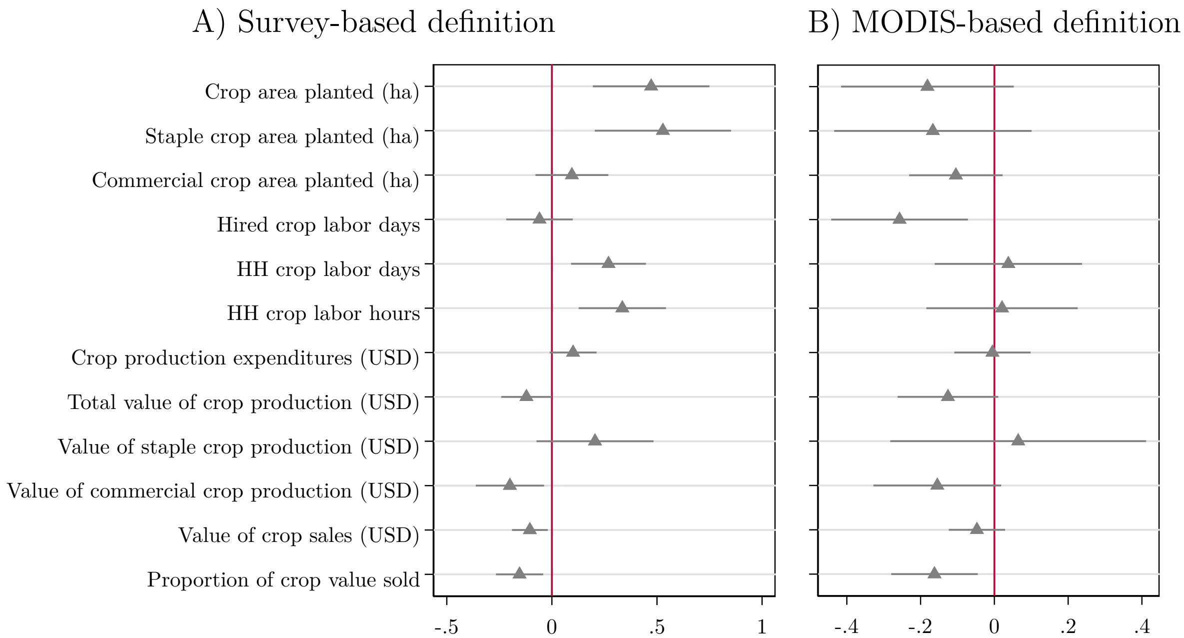

Note: The figure shows the coefficient and 95% confidence interval from separate regressions of household well-being outcomes on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. All outcome variables have been normalized so that the units are standard deviations away from the 2010-11 mean in non-flooded communities. The poverty line is defined as total household consumption being less than USD 1.90 PPP per capita per day. All regressions include household and state-by-year fixed effects. Standard errors are clustered at the level of the community of residence in 2012. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. In Panel A, flood exposure and the sample weights are defined based on survey reports for the community of residence in 2012. In Panel B, they are defined based on the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data.

Figure B.1.6: Robustness of impacts of flood exposure in 2012 on household income measures



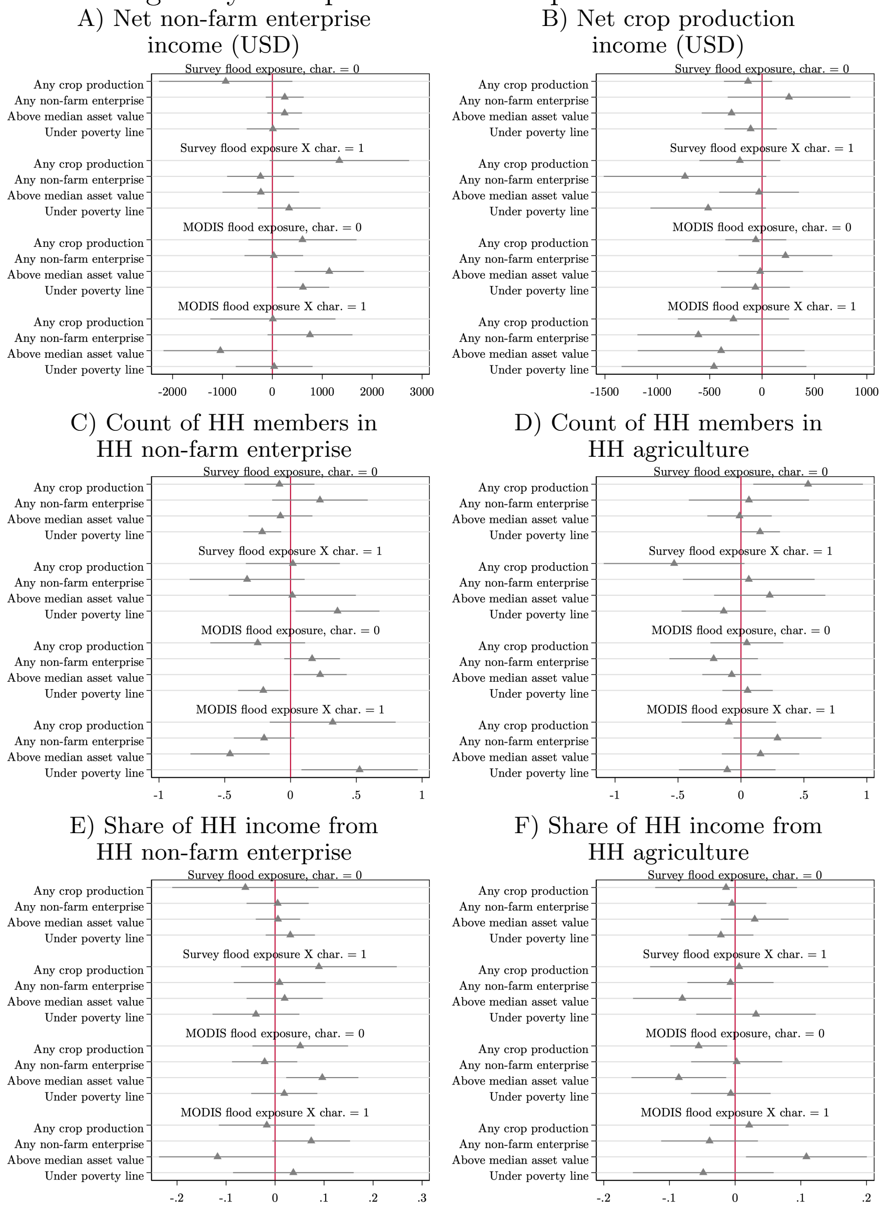

Note: The figure shows the coefficient and 95% confidence interval from separate regressions of household outcomes on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. Standard errors are clustered at the level of the community of residence in 2012. In Panel A, flood exposure and the sample weights use the survey-based flooding definition and in Panel B they use the MODIS-based definition. Net crop income is calculated by taking the value of all crop production (including output produced for own consumption) and subtracting expenditures over the prior agricultural year. Net non-farm enterprise income is calculated as total revenues minus total expenses over the prior 12 months. All regressions include household and state-by-year fixed effects, unless otherwise specified. 'Cells' are the 1 degree grid cell containing a given community. Controls are time-varying community variables and baseline household controls. The specification 'w/in 10km of any flood' excludes control communities farther than 10km from any community that ever flooded. The specification 'no 2018' drops the 2018-19 survey round when the sample was partially refreshed. Observations are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012 (IPW) unless otherwise specified. The row 'match flood risk + history' drops observations within strata of 2005-2011 flooding histories that are outside the range of common support for estimated flood risk in 2012, and includes strata by round fixed effects.

Figure B.1.7: Average impacts of flood exposure in 2012 on crop production outcomes, by flooding definition



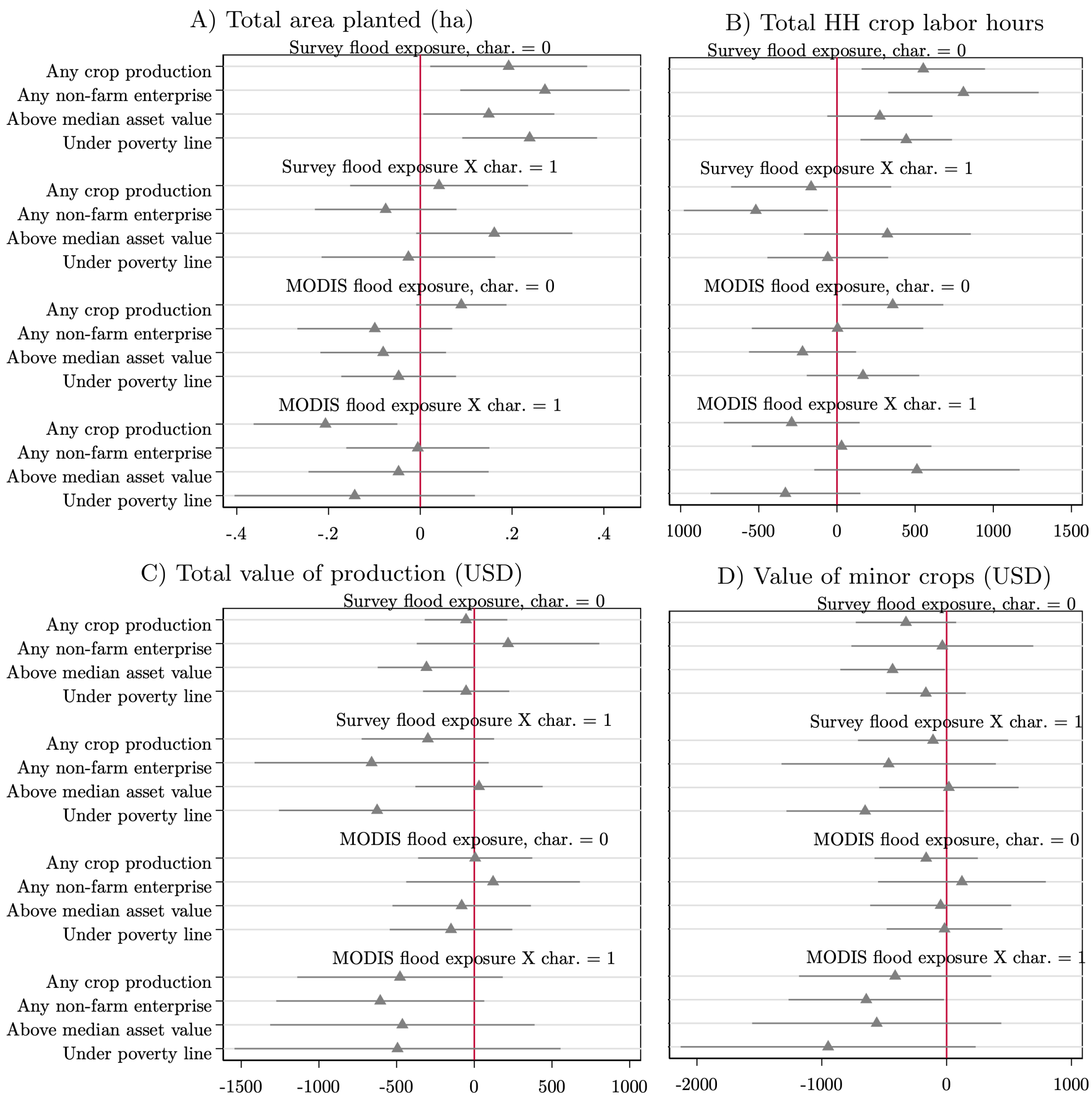

Note: The figure shows the coefficient and 95% confidence interval from separate regressions of household crop production outcomes on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. All outcome variables have been normalized so that the units are standard deviations away from the 2010-11 mean in non-flooded communities. Staple crops include maize, millet, sorghum, rice, cassava, yam, cowpea, and groundnut. 'Commercial' crops include all other crops. The value of crop production includes sold crops valued at their sales prices and unsold crops valued at the median price for sales of that crop within the community (or state, if there are fewer than 3 observations of sales for that crop in the community). All regressions include household and state-by-year fixed effects. Standard errors are clustered at the level of the community of residence in 2012. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. In Panel A, flood exposure and the sample weights are defined based on survey reports for the community of residence in 2012. In Panel B, they are defined based on the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data. Results and FDR-adjusted q-values are shown in Table B.2.6.

Figure B.1.8: Heterogeneity in impacts of flood exposure in 2012 on income activities




Note: The figure shows the coefficient and 95% confidence interval from separate regressions of household outcomes on the interaction between 2012 flood exposure and a baseline household characteristic from the 2010-11 survey round. The headings for each set of 4 coefficients indicate the flooding definition and whether the coefficient is for the base effect of flood exposure when the characteristic has a value of 0 or for the interaction of flood exposure with that characteristic having a value of 1 (the differential effect). All regressions include household and state-by-year fixed effects. Standard errors are clustered at the level of the community of residence in 2012. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012.

Figure B.1.9: Heterogeneity in impacts of survey-based flood exposure in 2012 on crop production outcomes




Note: The figure shows the coefficient and 95% confidence interval from separate regressions of household outcomes on the interaction between 2012 flood exposure and a baseline household characteristic. Baseline characteristics are from the 2010-11 survey round. The headings for each set of 4 coefficients indicate the flooding definition and whether the coefficient is for the base effect of flood exposure when the characteristic has a value of 0 or for the interaction of flood exposure with that characteristic having a value of 1. Minor crops include all but the 7 most commonly cultivated crops for Nigeria. All regressions include household and state-by-year fixed effects. Standard errors are clustered at the level of the community of residence in 2012. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012.

## B.2 Additional tables

Table B.2.1: Alignment of household survey- vs. community survey-based definitions of 2012 GHSP community flood exposure

| | | Community survey flooding | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | Row total | Share agreed |
| HH survey flood definition | | 331 | 134 | 465 | |
| Any household with | 0 | 275 | 62 | 337 | 0.82 |
| shocks module flooding | 1 | 56 | 72 | 128 | 0.56 |
| Any household with | 0 | 233 | 46 | 279 | 0.84 |
| any reported flooding | 1 | 98 | 88 | 186 | 0.47 |
| At least 2 households with | 0 | 302 | 88 | 390 | 0.77 |
| shocks module flooding | 1 | 29 | 46 | 75 | 0.61 |
| At least 2 households with | 0 | 281 | 63 | 344 | 0.82 |
| any reported flooding | 1 | 50 | 71 | 121 | 0.59 |

Table B.2.2: Balance in baseline characteristics by survey-based flood exposure

| | A. Without inverse propensity weights | | B. With inverse propensity weights | |
|---|---|---|---|---|
| | Not flooded (n=2570) Mean (SD) | Flooded (n=1171) Treatment diff (SE) | Not flooded (n=2168) Mean (SD) | Flooded (n=1082) Treatment diff (SE) |
| *Community characteristics* | | | | |
| Rural | 0.71 | 0.14*** | 0.76 | -0.02 |
| | (0.45) | (0.04) | (0.43) | (0.09) |
| Distance to nearest | 65.67 | 10.26** | 64.27 | 6.02 |
| market (km) | (41.12) | (4.39) | (39.44) | (9.33) |
| Annual mean | 26.33 | 0.11 | 26.37 | 0.24 |
| temperature (deg C) | (0.90) | (0.11) | (0.94) | (0.17) |
| Annual precipitation | 1491.23 | -277.58*** | 1475.68 | -38.03 |
| (mm) | (680.37) | (68.49) | (718.64) | (138.94) |
| Percent ag land w/in | 32.53 | 12.72*** | 34.52 | -4.91 |
| approx 1km | (25.35) | (2.85) | (25.99) | (4.62) |
| Slope (percent) | 3.28 | -0.63** | 3.09 | -0.20 |
| | (2.79) | (0.26) | (2.84) | (0.35) |
| Elevation (m) | 275.03 | 49.41** | 281.07 | -31.81 |
| | (203.03) | (24.79) | (211.47) | (36.53) |
| Years community | 0.62 | 0.75*** | 0.68 | -0.14 |
| reported flooding from 2007-2011 | (0.93) | (0.13) | (0.96) | (0.20) |
| Mean of cell annual | 0.09 | -0.00 | 0.09 | -0.00 |
| share of months with drought 1998-2014 | (0.02) | (0.00) | (0.02) | (0.00) |
| *Household characteristics* | | | | |
| Female-headed HH | 0.15 | -0.08*** | 0.14 | -0.03 |
| | (0.36) | (0.02) | (0.35) | (0.02) |
| Count of HH members | 5.40 | 1.39*** | 5.58 | 0.43* |
| | (2.92) | (0.17) | (2.95) | (0.23) |
| Livestock holdings | 0.87 | 0.63*** | 1.00 | -0.11 |
| (TLU) | (2.64) | (0.22) | (2.79) | (0.23) |
| Value of HH assets | 1520.42 | -445.61*** | 1378.15 | 233.60 |
| (USD) | (3237.70) | (143.60) | (2963.32) | (395.29) |
| Count of months in | 0.42 | -0.07 | 0.44 | -0.13* |
| last 12 that HH experienced food insecurity | (1.07) | (0.05) | (1.09) | (0.07) |
| HH under 1.90 USD | 0.31 | 0.17*** | 0.33 | 0.03 |
| PPP per capita daily poverty line | (0.46) | (0.03) | (0.47) | (0.06) |
| Household uses | 0.32 | -0.13*** | 0.29 | 0.01 |
| formal financial services | (0.47) | (0.03) | (0.45) | (0.07) |
| Years HH reported | 0.04 | 0.12*** | 0.05 | -0.00 |
| flood from 2007-2011 (0-5) | (0.24) | (0.02) | (0.26) | (0.03) |
| *Household productive activities* | | | | |
| Any HH crop | 0.71 | 0.18*** | 0.75 | -0.00 |
| production activity | (0.45) | (0.03) | (0.43) | (0.07) |
| Any HH livestock | 0.46 | 0.20*** | 0.50 | 0.00 |
| activity | (0.50) | (0.03) | (0.50) | (0.06) |
| Any HH non-farm | 0.72 | 0.04* | 0.72 | 0.04 |
| enterprise activity | (0.45) | (0.02) | (0.45) | (0.03) |
| Any HH wage work | 0.48 | -0.04 | 0.49 | 0.01 |
| activity | (0.50) | (0.03) | (0.50) | (0.04) |
| Total crop and | 1.11 | 1.51*** | 1.18 | 0.25 |
| pasture holdings (ha) | (2.42) | (0.23) | (2.51) | (0.28) |
| Household uses any | 0.42 | 0.08* | 0.45 | -0.00 |
| inorganic fertilizer | (0.49) | (0.04) | (0.50) | (0.07) |
| Value of crop sales | 1143.81 | 981.24*** | 1193.90 | 495.42 |
| (USD) | (2424.93) | (237.20) | (2447.43) | (321.07) |
| Proportion of value | 0.23 | -0.02 | 0.20 | 0.03 |
| of crop production sold | (0.30) | (0.02) | (0.28) | (0.04) |
| Test of joint significance | $F = 6.45$ | $p < 0.001$ | $F = 2.48$ | $p < 0.001$ |

Note: This table shows the baseline (2010-11) control mean and difference by 2012 flood exposure status for selected community and household characteristics. The results in Panel A do not include any weights, while those in Panel B include inverse propensity weights based on the estimated probability that a community reported flooding in 2012. Standard errors are clustered at the 2012 community level. The bottom row shows results for the test of the hypothesis that the relationship between 2012 flooding and all variables is jointly 0. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2.3: Balance in baseline characteristics by MODIS-based flood exposure

| | A. Without inverse propensity weights | | B. With inverse propensity weights | |
|---|---|---|---|---|
| | Not flooded (n=3009) | Flooded (n=760) | Not flooded (n=2734) | Flooded (n=696) |
| | Mean (SD) | Treatment diff (SE) | Mean (SD) | Treatment diff (SE) |
| *Community characteristics* | | | | |
| Rural | 0.78 | -0.12** | 0.78 | 0.06 |
| | (0.42) | (0.06) | (0.41) | (0.06) |
| Distance to nearest | 69.72 | -3.38 | 67.68 | -1.46 |
| market (km) | (41.98) | (4.73) | (40.10) | (5.42) |
| Annual mean | 26.31 | 0.26* | 26.38 | -0.32 |
| temperature (deg C) | (0.91) | (0.14) | (0.86) | (0.41) |
| Annual precipitation | 1361.19 | 229.15** | 1374.01 | -55.89 |
| (mm) | (614.47) | (98.82) | (635.02) | (97.31) |
| Percent ag land w/in | 38.02 | -8.21** | 38.54 | -2.17 |
| approx 1km | (26.36) | (3.21) | (26.57) | (6.56) |
| Slope (percent) | 3.27 | -0.86*** | 2.81 | -0.41** |
| | (2.88) | (0.22) | (1.90) | (0.21) |
| Elevation (m) | 307.14 | -87.26*** | 290.49 | 82.90 |
| | (200.29) | (31.12) | (192.03) | (95.28) |
| Years community | 0.82 | 0.17 | 0.79 | -0.15 |
| reported flooding from 2007-2011 | (1.14) | (0.13) | (1.12) | (0.14) |
| Mean of cell annual | 0.09 | -0.00 | 0.09 | -0.00 |
| share of months with drought 1998-2014 | (0.02) | (0.00) | (0.02) | (0.00) |
| *Household characteristics* | | | | |
| Female-headed HH | 0.12 | 0.03 | 0.12 | 0.03 |
| | (0.33) | (0.02) | (0.33) | (0.03) |
| Count of HH members | 5.87 | -0.23 | 5.87 | 0.28 |
| | (3.20) | (0.20) | (3.20) | (0.45) |
| Livestock holdings | 1.13 | -0.34* | 1.06 | -0.11 |
| (TLU) | (3.01) | (0.19) | (2.83) | (0.23) |
| Value of HH assets | 1331.63 | 229.50 | 1311.90 | -23.46 |
| (USD) | (2922.78) | (212.09) | (2811.99) | (201.13) |
| Count of months in | 0.40 | 0.02 | 0.40 | -0.10 |
| last 12 that HH experienced food insecurity | (1.02) | (0.07) | (1.03) | (0.07) |
| HH under 1.90 USD | 0.37 | -0.07** | 0.37 | 0.04 |
| PPP per capita daily poverty line | (0.48) | (0.04) | (0.48) | (0.07) |
| Household uses | 0.26 | 0.13*** | 0.26 | 0.06 |
| formal financial services | (0.44) | (0.04) | (0.44) | (0.07) |
| Years HH reported | 0.07 | 0.04 | 0.07 | 0.00 |
| flood from 2007-2011 (0-5) | (0.33) | (0.03) | (0.33) | (0.02) |
| *Household productive activities* | | | | |
| Any HH crop | 0.79 | -0.16*** | 0.79 | -0.02 |
| production activity | (0.40) | (0.04) | (0.41) | (0.05) |
| Any HH livestock | 0.55 | -0.15*** | 0.54 | -0.02 |
| activity | (0.50) | (0.04) | (0.50) | (0.07) |
| Any HH non-farm | 0.73 | 0.02 | 0.72 | 0.04 |
| enterprise activity | (0.45) | (0.02) | (0.45) | (0.03) |
| Any HH wage work | 0.47 | 0.03 | 0.46 | 0.06 |
| activity | (0.50) | (0.03) | (0.50) | (0.06) |
| Total crop and | 1.62 | -0.22 | 1.59 | 0.24 |
| pasture holdings (ha) | (2.92) | (0.24) | (2.92) | (0.28) |
| Household uses any | 0.45 | 0.01 | 0.45 | 0.12 |
| inorganic fertilizer | (0.50) | (0.05) | (0.50) | (0.10) |
| Value of crop sales | 1451.69 | -51.87 | 1393.15 | 856.73 |
| (USD) | (2675.41) | (320.65) | (2627.09) | (622.32) |
| Proportion of value | 0.23 | -0.03 | 0.23 | -0.03 |
| of crop production sold | (0.29) | (0.02) | (0.30) | (0.02) |
| Test of joint significance | $F = 4.76$ | $p < 0.001$ | $F = 1.61$ | $p = 0.037$ |

Note: This table shows the baseline (2010-11) control mean and difference by 2012 flood exposure status for selected community and household characteristics. Values are in constant 2016 USD. Flood exposure is defined based on the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data. The results in Panel A do not include any weights, while those in Panel B include inverse propensity weights based on the estimated probability that a community was within 5km of a flooded pixel in 2012. Standard errors are clustered at the 2012 community level. The bottom row shows results for the test of the hypothesis that the relationship between 2012 flooding and all variables is jointly 0. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2.4: Impacts of 2012 flood exposure on engagement in different income activities, by flooding measure

| | N | Control Mean (SD) | Flood × Post (SE) | FDR q-value |
|---|---|---|---|---|
| *A. Survey-based measure* | | | | |
| Any HH farm activity | 10266 | 0.79 | 0.02* | 0.191 |
| | | [0.41] | (0.01) | |
| Any HH crop production activity | 10256 | 0.76 | 0.03** | 0.108 |
| | | [0.43] | (0.01) | |
| Any HH livestock activity | 10256 | 0.51 | 0.05** | 0.062 |
| | | [0.50] | (0.02) | |
| Any HH non-farm work activity | 10266 | 0.81 | 0.01 | 0.563 |
| | | [0.39] | (0.02) | |
| Any HH non-farm enterprise activity | 10259 | 0.72 | -0.01 | 0.583 |
| | | [0.45] | (0.02) | |
| Any HH wage work activity | 10266 | 0.39 | 0.03 | 0.234 |
| | | [0.49] | (0.02) | |
| Any other HH income source | 10256 | 0.09 | 0.01 | 0.630 |
| *B. MODIS-based measure* | | | | |
| Any HH farm activity | 10974 | 0.79 | -0.03* | 0.242 |
| | | [0.40] | (0.01) | |
| Any HH crop production activity | 10960 | 0.77 | -0.03* | 0.242 |
| | | [0.42] | (0.01) | |
| Any HH livestock activity | 10960 | 0.51 | -0.02 | 0.391 |
| | | [0.50] | (0.02) | |
| Any HH non-farm work activity | 10974 | 0.81 | -0.00 | 0.956 |
| | | [0.40] | (0.02) | |
| Any HH non-farm enterprise activity | 10963 | 0.71 | 0.00 | 0.939 |
| | | [0.45] | (0.02) | |
| Any HH wage work activity | 10974 | 0.37 | -0.00 | 0.984 |
| | | [0.48] | (0.02) | |
| Any other HH income source | 10960 | 0.09 | -0.02 | 0.242 |
| | | [0.28] | (0.01) | |

Note: This table shows the results from separate regressions of household activities on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. In Panel A flood exposure is defined based on survey reports, while in Panel B it is based on proximity to MODIS-identified flooding. Each row represents an outcome. Household engagement each survey round is based on any household or household member activity in a given area in either the post-planting or post-harvest survey. All regressions include household and state-by-year fixed effects. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. Standard errors are clustered at the level of the community of residence in 2012. FDR $q-values$ adjusting for multiple hypothesis testing are included in the last column. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2.5: Impacts of 2012 flood exposure on household income from different activities, by flooding measure

| | N | Control Mean (SD) | Flood × Post (SE) | FDR q-value |
|---|---|---|---|---|
| *A. Survey-based measure* Total net income (USD) | 10365 | 3550.71 | -179.57 | 0.594 |
| | | [6104.10] | (281.33) | |
| Net income from farm activities (USD) | 10365 | 1127.38 | -284.01** | 0.104 |
| | | [2289.27] | (133.10) | |
| Net income from crop production (USD) | 10365 | 1069.45 | -307.05** | 0.071 |
| | | [2257.92] | (132.11) | |
| Net income from livestock (USD) | 10365 | 57.92 | 23.04 | 0.251 |
| | | [310.12] | (16.16) | |
| Net income from non-farm activities (USD) | 10365 | 2423.34 | 104.44 | 0.704 |
| | | [5827.69] | (238.46) | |
| Net income from HH non-farm enterprise (USD) | 10365 | 862.92 | 131.37 | 0.583 |
| | | [3439.77] | (185.86) | |
| Income from wage labor (USD) | 10365 | 1485.49 | -17.03 | 0.897 |
| | | [4641.48] | (131.00) | |
| Income from all other sources (USD) | 10365 | 74.92 | -9.90 | 0.594 |
| | | [391.77] | (15.25) | |
| HH income Herfindahl-Hirschman | 8686 | 0.84 | -0.01 | 0.251 |
| Index | | [0.23] | (0.01) | |
| *B. MODIS-based measure* | | | | |
| Total net income (USD) | 11054 | 3565.00 | 428.48 | 0.288 |
| | | [6018.65] | (304.98) | |
| Net income from farm activities (USD) | 11054 | 1124.60 | -217.10 | 0.292 |
| | | [2243.67] | (158.54) | |
| Net income from crop production (USD) | 11054 | 1069.96 | -235.29 | 0.265 |
| | | [2218.07] | (159.16) | |
| Net income from livestock (USD) | 11054 | 54.63 | 18.18 | 0.265 |
| | | [287.19] | (12.28) | |
| Net income from non-farm activities (USD) | 11054 | 2440.41 | 645.58** | 0.174 |
| | | [5748.45] | (288.04) | |
| Net income from HH non-farm enterprise (USD) | 11054 | 862.15 | 635.50*** | 0.039 |
| (USD) | | [3273.35] | (198.39) | |
| Income from wage labor (USD) | 11054 | 1510.13 | 30.51 | 0.953 |
| | | [4749.28] | (183.55) | |
| Income from all other sources (USD) | 11054 | 68.12 | -20.43 | 0.391 |
| | | [365.22] | (18.73) | |
| HH income Herfindahl-Hirschman | 9156 | 0.83 | 0.01 | 0.592 |
| Index | | [0.23] | (0.01) | |

Note: This table shows the results from separate regressions of household incomes on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. In Panel A flood exposure is defined based on survey reports, while in Panel B it is based on proximity to MODIS-identified flooding. Each row represents an outcome. All incomes are in constant 2016 USD. Net income from household farm activities is calculated by taking the value of all household farm output (including output produced for own consumption) and subtracting expenditures on all purchased inputs over the prior agricultural year. Net income from household non-farm enterprise is calculating as total revenues minus total expenses over the prior 12 months. Income from wage work is calculated as total earnings over the prior 12 months. The Herfindahl-Hirschman Index (HHI) is the sum of squared income shares for household farm income (crop and livestock), non-farm household enterprise income, and wage income and can take a value between 0 and 1, where 1 indicates that all household income comes from a single activity. All regressions include household and state-by-year fixed effects. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. Standard errors are clustered at the level of the community of residence in 2012. FDR $q - values$ adjusting for multiple hypothesis testing are included in the last column. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2.6: Impacts of 2012 flood exposure on crop production, by flooding measure

| | N | Control Mean (SD) | Flood × Post (SE) | FDR q-value |
|---|---|---|---|---|
| *Survey-based measure* | | | | |
| Total area planted (ha) | 10365 | 0.34 | 0.23*** | 0.014 |
| | | [0.81] | (0.07) | |
| Total staple crop area planted (ha) | 10375 | 0.27 | 0.21*** | 0.014 |
| | | [0.67] | (0.06) | |
| Total commercial crop area planted (ha) | 10365 | 0.07 | 0.02 | 0.398 |
| | | [0.38] | (0.02) | |
| Hired crop labor days in past year | 10365 | 13.80 | -1.44 | 0.583 |
| | | [32.00] | (1.97) | |
| HH crop labor days in past year | 10365 | 195.67 | 65.91*** | 0.022 |
| | | [395.28] | (22.19) | |
| HH crop labor hours in past year | 10375 | 1121.32 | 437.29*** | 0.014 |
| | | [2397.46] | (137.83) | |
| Total crop production expenditures (USD) | 10365 | 314.91 | 57.74* | 0.172 |
| | | [760.76] | (32.44) | |
| Value of crop production (USD) | 10365 | 1397.75 | -293.56** | 0.116 |
| | | [2603.55] | (147.71) | |
| Value of staple crop production (USD) | 10375 | 661.18 | 157.95 | 0.251 |
| | | [1521.38] | (108.93) | |
| Value of commercial crop production (USD) | 10365 | 735.92 | -451.95** | 0.062 |
| | | [2135.97] | (187.34) | |
| Value of crop sales (USD) | 10365 | 315.79 | -97.50** | 0.062 |
| | | [883.68] | (40.35) | |
| Proportion of value of crop production sold | 6786 | 0.21 | -0.05*** | 0.042 |
| | | [0.28] | (0.02) | |
| *MODIS-based measure* | | | | |
| Total area planted (ha) | 11054 | 0.46 | -0.10 | 0.265 |
| | | [0.94] | (0.07) | |
| Total staple crop area planted (ha) | 11068 | 0.37 | -0.08 | 0.347 |
| | | [0.83] | (0.06) | |
| Total commercial crop area planted (ha) | 11054 | 0.08 | -0.03 | 0.242 |
| | | [0.39] | (0.02) | |
| Hired crop labor days in past year | 11054 | 14.81 | -7.75*** | 0.058 |
| | | [32.87] | (2.84) | |
| HH crop labor days in past year | 11054 | 229.58 | 10.45 | 0.868 |
| | | [431.82] | (27.98) | |
| HH crop labor hours in past year | 11068 | 1341.85 | 30.86 | 0.953 |
| | | [2612.08] | (154.43) | |
| Total crop production expenditures (USD) | 11054 | 355.53 | -3.42 | 0.956 |
| | | [811.02] | (34.79) | |
| Value of crop production (USD) | 11054 | 1437.37 | -334.96* | 0.242 |
| | | [2547.79] | (184.98) | |
| Value of staple crop production (USD) | 11068 | 783.81 | 51.79 | 0.868 |
| | | [1674.49] | (141.47) | |
| Value of commercial crop production (USD) | 11054 | 653.13 | -391.15* | 0.242 |
| | | [1976.07] | (222.73) | |
| Value of crop sales (USD) | 11054 | 351.81 | -46.39 | 0.347 |
| | | [937.11] | (38.12) | |
| Proportion of value of crop production sold | 6853 | 0.23 | -0.05*** | 0.058 |
| | | [0.29] | (0.02) | |

Note: This table shows the results from separate regressions of household crop production outcomes on a dummy for having resided in a community in 2012 that was exposed to flooding and being observed after 2012. In Panel A flood exposure is defined based on survey reports, while in Panel B it is based on proximity to MODIS-identified flooding. Each row represents an outcome. Staple crops include maize, millet, sorghum, rice, cassava, yam, cowpea, and groundnut. 'Commercial' crops include all other crops. All monetary values are in constant 2016 USD. The value of crop production includes sold crops valued at their sales prices and unsold crops valued at the median price for sales of that crop within the community (or state, if there are fewer than 3 observations of sales for that crop in the community). All regressions include household and state-by-year fixed effects. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012. Standard errors are clustered at the level of the community of residence in 2012. FDR $q - values$ adjusting for multiple hypothesis testing are included in the last column. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2.7: IV Impacts of 2012 flood exposure measures on household outcomes

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Net non-farm enterprise income (USD) | | Net crop production income (USD) | | Any food insecurity last 12 months | |
| MODIS flood × Post | 365.0 | | -1437.1* | | 0.2** | |
| | (875.8) | | (823.5) | | (0.1) | |
| Survey flood × Post | | 1687.2* | | -313.6 | | 0.3*** |
| | | (948.7) | | (663.9) | | (0.1) |
| Observations | 10971 | 10365 | 10971 | 10365 | 10985 | 10375 |
| First stage F-stat | 10.8 | 12.3 | 10.8 | 12.3 | 10.8 | 12.3 |
| State-Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Household FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Inv. Prop. Weight | MODIS | Survey | MODIS | Survey | MODIS | Survey |

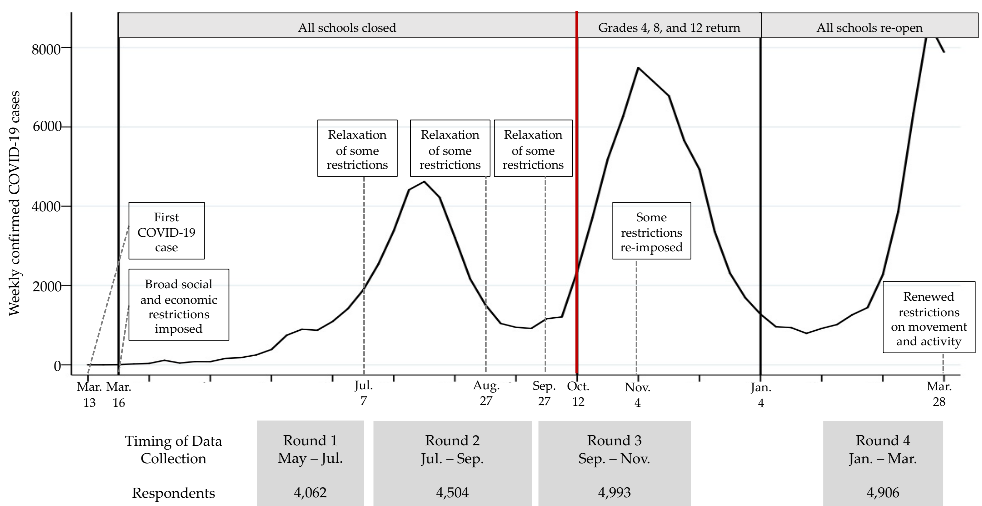
Note: This table shows the results from separate IV regressions of household outcomes on 2012 flooding × Post dummies for a given flooding definition instrumented using the other flooding definition. Survey-based flood exposure is defined based on survey reports for the community of residence in 2012, while MODIS-based exposure is defined as the community of residence in 2012 being within 5km of a pixel identified as flooded using MODIS data. I include first stage Kleibergen-Paap Wald $F$-statistics for tests of weak identification. All incomes are in constant 2016 USD. Net income from household farm activities is calculated by taking the value of all household farm output (including output produced for own consumption) and subtracting expenditures on all purchased inputs over the prior agricultural year. Net income from household non-farm enterprise is calculating as total revenues minus total expenses over the prior 12 months. All regressions include household and state-by-year fixed effects. Observations in all regressions are weighted by the inverse of the estimated probability of having been exposed to flooding in 2012 using either the survey-based or MODIS-based definition. Standard errors are clustered at the level of the community of residence in 2012. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

# Appendix C

# Balancing work and childcare: Evidence from COVID-19 school closures and reopenings in Kenya – Appendix

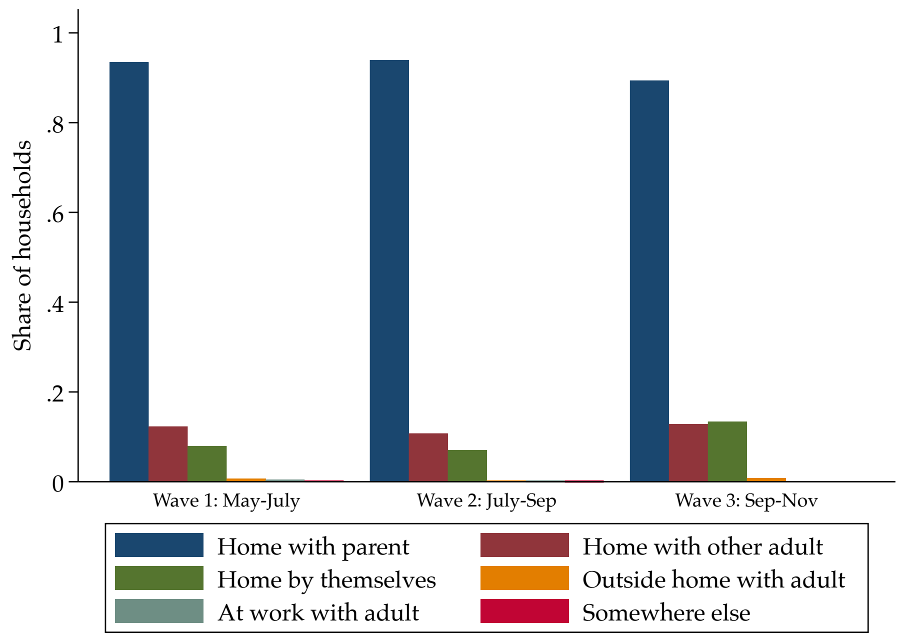
## C.1 Additional figures and tables

Figure C.1.1: Kenya COVID-19 cases, pandemic policy, and data collection timeline



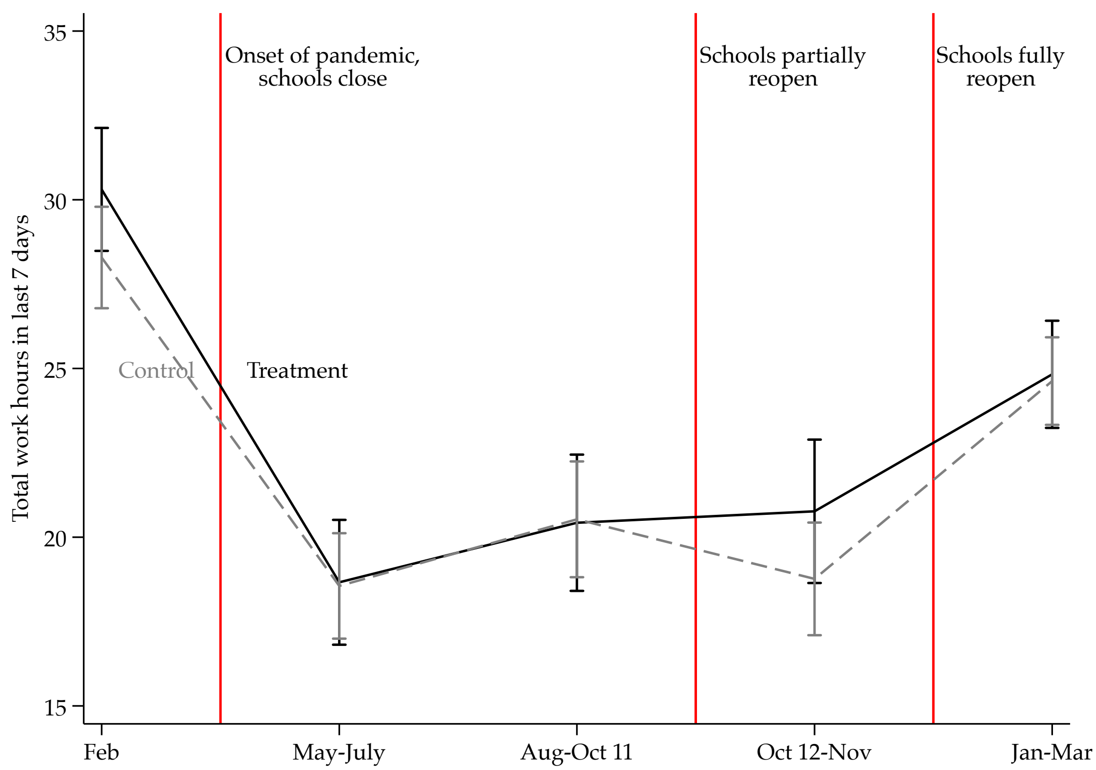

Note: The figure shows the evolution of weekly confirmed COVID-19 cases in Kenya over time, along with the timing of key pandemic policy changes. The red bar indicates the partial school reopening on 12 October, the focus of the analysis. 'Relaxation of some restrictions' indicates that one or more of the initial pandemic constraints were at least partially reduced. Specific policy changes are outlined in Appendix B.2.
Sources: COVID-19 government response timeline for Kenya; Kenya COVID Tracker; Presidency of Kenya; Kenya Ministry of Education Twitter feed; COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University

Figure C.1.2: Childcare arrangements when children are out of school




Note: Respondents are asked to specify all of the situations where a randomly selected child spent at least some time when out of school in the past week. ‘Somewhere else’ combines ’daycare/other childcare’ and ’at home with a maid/domestic helper.’ The figure uses information on childcare arrangements for analysis sample households with at least one child in any grade from 3 to 9, but the distribution is nearly identical when considering all households with children.

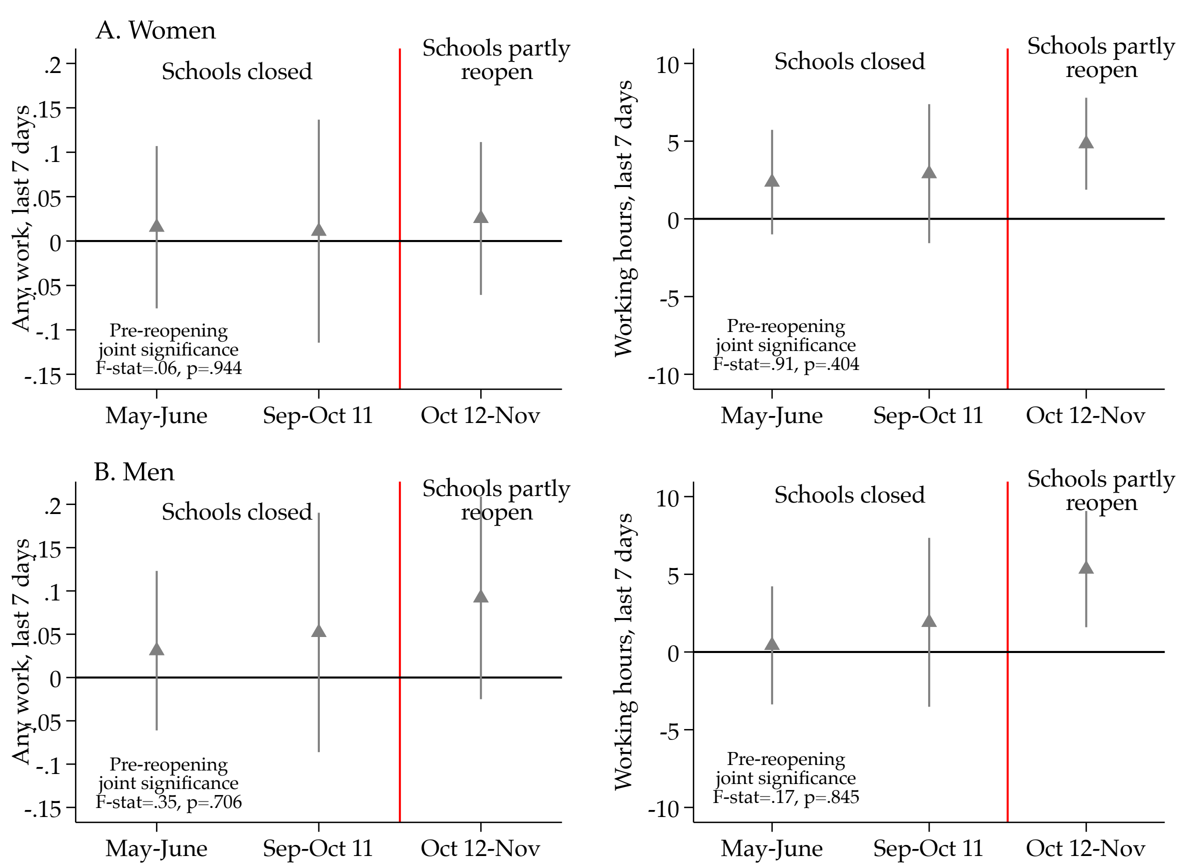
Figure C.1.3: Respondent work hours in the last 7 days by survey round and treatment status




Note: The figure shows raw means and 95% confidence intervals for household respondents' total work hours in the last 7 days by treatment status in each time period. Means are shown for the respondent only due to missing data on pre-pandemic working hours for other household adults. Treatment households have a child enrolled in grades 4 or 8, and control households have a child enrolled in grades 3, 5, 6, 7, or 9. We do not show means for mixed households with children in both grade groups.

Data for February are based on recall from the first time a respondent is surveyed. We combine observations from the first two weeks of survey round 3, before the partial school reopening, with data from round 2. The red bars indicate changes in Kenya's school closures policy. The fall in hours after the partial reopening for control households reflects the end of main harvest period in Kenya, as 64% of households are engaged in agriculture.

Figure C.1.4: Impact of treatment on labor participation in the last 7 days, by time period and sex




Note: The figure replicates Figure 3.3 but estimating the regressions separately for women and men. It shows estimated coefficients and 95% confidence intervals for the interaction between *Treat* and time period from Equation 3.1, where *Post* is replaced with time period dummies. Time periods prior to the partial school reopening are constructed to have roughly equal sample sizes. The reference period is July-August 15, while schools were closed and before the partial reopening was announced. The red bar indicates the timing of Kenya's partial school reopening. Outcomes are any work engagement (left) and total work hours (right) in the 7 days prior to the interview. Treatment households have a child enrolled in grades 4 or 8, and control households have a child enrolled in grades 3, 5, 6, 7, or 9. We do not show coefficients for mixed households with children in both grade groups. $p$-values for the test of joint significance of the pre-reopening coefficients and for tests of equality between the pre-reopening coefficient and the post-reopening coefficient are shown.

Table C.1.1: Baseline balance by treatment status

| | Control HH Mean | N | Mixed HH Mean | N | Treatment HH Mean | N | C-T Diff. | C-T p-value | C-M Diff. | C-M p-value |
|---|---|---|---|---|---|---|---|---|---|---|
| *Respondent characteristics* | | | | | | | | | | |
| Age | 40.12 | 934 | 41.22 | 362 | 41.36 | 327 | -1.24 | 0.108 | -1.10 | 0.119 |
| Female | 0.58 | 934 | 0.57 | 362 | 0.57 | 327 | 0.01 | 0.669 | 0.01 | 0.798 |
| Completed primary school | 0.88 | 934 | 0.84 | 362 | 0.86 | 327 | 0.02 | 0.477 | 0.04 | 0.066 |
| Completed secondary school | 0.48 | 934 | 0.41 | 362 | 0.49 | 327 | -0.01 | 0.739 | 0.07 | 0.029 |
| Completed school beyond secondary | 0.15 | 934 | 0.13 | 362 | 0.17 | 327 | -0.02 | 0.415 | 0.02 | 0.302 |
| Married | 0.75 | 927 | 0.80 | 360 | 0.73 | 323 | 0.02 | 0.461 | -0.05 | 0.058 |
| Is the household head | 0.64 | 934 | 0.62 | 362 | 0.64 | 327 | -0.00 | 0.950 | 0.02 | 0.594 |
| *Household characteristics* | | | | | | | | | | |
| Female household head | 0.29 | 934 | 0.26 | 362 | 0.30 | 327 | -0.01 | 0.746 | 0.03 | 0.314 |
| Age of household head | 44.49 | 934 | 45.30 | 362 | 46.11 | 327 | -1.62 | 0.036 | -0.81 | 0.266 |
| Count adults | 2.54 | 934 | 2.76 | 362 | 2.61 | 327 | -0.07 | 0.352 | -0.22 | 0.007 |
| More than 2 household adults | 0.40 | 934 | 0.49 | 362 | 0.40 | 327 | -0.00 | 0.937 | -0.10 | 0.002 |
| Age of youngest household child | 6.75 | 934 | 6.10 | 362 | 6.89 | 327 | -0.13 | 0.620 | 0.66 | 0.005 |
| Any below-school-age (0-4) children | 0.42 | 934 | 0.43 | 362 | 0.37 | 327 | 0.05 | 0.098 | -0.01 | 0.688 |
| Any young (0-8) children | 0.66 | 934 | 0.72 | 362 | 0.67 | 327 | -0.02 | 0.587 | -0.06 | 0.038 |
| Count young (0-4) children | 0.56 | 934 | 0.62 | 362 | 0.48 | 327 | 0.08 | 0.101 | -0.06 | 0.240 |
| Count school (5-17) children | 2.47 | 934 | 3.28 | 362 | 2.31 | 327 | 0.15 | 0.059 | -0.81 | 0.000 |
| Count adolescent (10-17) children | 0.29 | 934 | 0.45 | 362 | 0.29 | 327 | -0.00 | 0.952 | -0.16 | 0.007 |
| Household wealth index | -0.06 | 934 | -0.15 | 362 | 0.03 | 327 | -0.09 | 0.137 | 0.10 | 0.099 |
| Connected to electricity grid | 0.45 | 934 | 0.41 | 362 | 0.51 | 327 | -0.05 | 0.095 | 0.05 | 0.118 |
| Urban household | 0.46 | 934 | 0.46 | 362 | 0.47 | 327 | -0.01 | 0.647 | -0.00 | 0.877 |
| Household engaged in agriculture | 0.64 | 934 | 0.67 | 362 | 0.63 | 327 | 0.01 | 0.771 | -0.03 | 0.312 |
| Any child engaged in household farm labor | 0.26 | 934 | 0.31 | 362 | 0.26 | 327 | -0.00 | 0.976 | -0.06 | 0.049 |
| Household engaged in enterprise | 0.15 | 934 | 0.17 | 362 | 0.19 | 327 | -0.04 | 0.118 | -0.01 | 0.517 |
| *Respondent labor participation* | | | | | | | | | | |
| Engaged in any work in last 7 days | 0.70 | 934 | 0.69 | 362 | 0.71 | 327 | -0.01 | 0.670 | 0.01 | 0.823 |
| Engaged in wage employment in last 7 days | 0.10 | 934 | 0.08 | 362 | 0.13 | 327 | -0.03 | 0.204 | 0.02 | 0.214 |
| Engaged in HH agriculture in last 7 days | 0.57 | 934 | 0.60 | 362 | 0.54 | 327 | 0.03 | 0.374 | -0.03 | 0.290 |
| Engaged in HH non-ag enterprise in last 7 days | 0.09 | 934 | 0.11 | 362 | 0.13 | 327 | -0.03 | 0.119 | -0.01 | 0.440 |
| Engaged in any work in February 2020 | 0.76 | 934 | 0.76 | 362 | 0.75 | 327 | 0.00 | 0.866 | -0.00 | 0.919 |

Note: The table presents means for control households (C) with a child in grade 3, 5, 6, 7, or 9, treatment households (T) with a child in grade 4 or 8, and mixed households (M) with a child in both grade groups. Data are from the first time a household is observed, typically in survey round 1 (May-early July) while schools were fully closed. Individual-level data are for the survey respondent.
Columns on the right present differences and means and $p$-values for tests of equality for control households compared to treatment and mixed households, separately. The joint F-stat for differences across control and treatment households is 0.90, with p-value 0.594. It is 9.14 (p¡0.001) for differences across control and mixed households.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.1.2: Impacts of partial school reopening on household income in the last 14 days

| | KSH | | | | Log(KSH+1) | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Source of income | Total | Wage work | HH ag. | HH ent. | Total | Wage work | HH ag. | HH ent. |
| Post x Treat | -18.103 | 188.933 | -240.140 | 33.104 | -0.517 | -0.208 | -0.247 | -0.152 |
| | (656.326) | (449.291) | (168.963) | (448.344) | (0.359) | (0.233) | (0.248) | (0.221) |
| Post x Mixed | -736.476 | -151.989 | -313.544** | -270.943 | -0.581 | -0.200 | -0.381* | -0.117 |
| | (690.646) | (395.970) | (136.337) | (544.078) | (0.359) | (0.230) | (0.211) | (0.233) |
| Observations | 3103 | 3103 | 3103 | 3103 | 3103 | 3103 | 3103 | 3103 |
| Mean, pre-reopen control | 2360.892 | 814.774 | 298.003 | 1248.115 | 2.165 | 0.703 | 0.739 | 0.908 |

Note: This table presents estimates of Equation 3.1 for measures of household income over the last 14 days. Columns 1-4 report estimates for income in Kenyan Shillings (KSH; USD 1 $\approx$ 107 KSH) and columns 5-8 report estimates for the log of income + 1. Observations include data from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat'), control households with children in an adjacent grade, and 'mixed' households with both types of children. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.1.3: Heterogeneity in impacts of partial school reopening on adult working hours

| Interaction term $Z$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Female | Ag HH | Urban | Large Urban | Above Mean Wealth | Any Kids Aged 0-4 | ¿2 HH Adults |
| Post | -2.271 | -1.069 | 2.121 | 2.121 | -1.246 | -0.527 | -4.488 |
| | (2.541) | (2.835) | (3.727) | (3.733) | (3.069) | (3.056) | (3.037) |
| Post × Treat | 3.684* | 2.924 | 3.063 | 3.063 | 6.523*** | 5.029** | 1.567 |
| | (1.961) | (2.211) | (2.410) | (2.414) | (2.524) | (2.209) | (2.287) |
| Post × Z=1 | 2.791 | 0.786 | -2.650 | 0.296 | 2.182 | -1.919 | 6.322 |
| | (1.865) | (4.132) | (4.523) | (5.421) | (4.248) | (4.075) | (4.144) |
| Post × Treat × Z=1 | -0.453 | 3.210 | -0.868 | -1.107 | -6.612** | -3.500 | 3.724 |
| | (1.751) | (3.416) | (3.176) | (4.245) | (3.205) | (3.261) | (3.108) |
| Observations | 8717 | 8717 | 8717 | 5300 | 8717 | 8717 | 8717 |
| Mean, pre-reopen control | 16.578 | 16.578 | 16.578 | 16.578 | 16.578 | 16.578 | 16.578 |

Note: This table presents estimates of Equation 3.1 but interacting a characteristic $Z$ with all right-hand side variables except the household fixed effects. The column label indicates which characteristic $Z$ is being used. Values for household characteristics are from the first time they are observed in the data. 'Closures' work participation is based on any participation in a given sector from May-October 2020. 'Large Urban' is a dummy for location in one of Kenya's largest urban areas (Nairobi, Mombasa, Nakuru, Kisumu, Kiambu) relative to any rural area, while 'Urban' is a dummy for location in any urban area. 'Above Mean Wealth' is a dummy for whether an index of household wealth, based on housing and asset ownership, is above the sample mean. 'Girl Returns' is a dummy for any girl in grade 3-9.
The dependent variable is total working hours over the last 7 days, with individuals not working coded as working 0 hours. Observations include data from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat'), control households with children in an adjacent grade, and 'Mixed' households with both. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Coefficients with $Z = 1$ represent the sum of the $Post \times Treat$ and $Post \times Treat \times Z$ terms, and analogously for Mixed households. We include $p$-values for tests of whether the interaction term is equal to 0. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.1.4: Heterogeneity in impacts of partial reopening by child grade

| Adult hours in last 7 days in | All work activities | | | HH agriculture work | | | HH enterprise or wage work | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post x Treat | 3.554 | 5.083** | 0.605 | 2.942 | 4.927** | 0.612 | 0.612 | 0.156 | -0.008 |
| | (2.190) | (2.329) | (5.567) | (1.795) | (1.982) | (5.289) | (1.387) | (0.974) | (1.350) |
| Post x Mixed | 0.016 | 1.253 | -3.577 | 0.300 | 1.671 | -0.414 | -0.284 | -0.418 | -3.162** |
| | (1.827) | (1.951) | (2.698) | (1.543) | (1.635) | (2.327) | (0.950) | (0.970) | (1.553) |
| Observations | 6724 | 6284 | 2551 | 6724 | 6284 | 2551 | 6724 | 6284 | 2551 |
| Mean, pre-reopen control | 16.693 | 16.702 | 15.958 | 12.073 | 12.634 | 11.700 | 4.619 | 4.068 | 4.259 |
| Households with children in grades: | 2-6 | 6-10 | 10-12 | 2-6 | 6-10 | 10-12 | 2-6 | 6-10 | 10-12 |

Note: This table presents estimates of Equation 3.1 for different sub-samples, in comparison to Table 3.1 which includes all households in the main analysis sample (with children in grades 3-9). For households with a child in grades 2-6, treatment means having a child in grade 4 eligible to return to school, while treatment is driven by grade 8 children for households with grade 6-10 children and by grade 12 children for households with grade 10-12 children. In all cases the set of comparison households includes those with children within 2 grades of the 'treated' grade. Households with a child in another treated grade outside the focus range are categorized as 'mixed.'
Dependent variables are defined over the last 7 days, and take a value of 0 for individuals not working in a particular activity. Observations include data from May to November 2020. All regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.1.5: Impacts of partial school reopening by baseline wealth and respondent sex

| Respondent hours in last 7 days in | All work activities | | | All childcare activities | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Post × Treat, Below | 11.603*** | 5.858 | 21.997*** | 1.367 | 12.524 | -19.797 |
| mean baseline wealth | (3.465) | (4.169) | (5.997) | (8.136) | (10.656) | (12.667) |
| Post × Treat, Above | 2.322 | 4.848 | -0.809 | 2.623 | -2.406 | 3.265 |
| mean baseline wealth | (3.072) | (4.156) | (4.665) | (6.329) | (8.948) | (8.792) |
| Observations | 2288 | 1331 | 902 | 2284 | 1328 | 901 |
| Mean, pre-reopen control | 19.272 | 17.211 | 22.478 | 52.747 | 60.059 | 41.175 |
| Sample | All | Women | Men | All | Women | Men |
| p-value, diff. by baseline wealth | 0.045 | 0.864 | 0.003 | 0.903 | 0.284 | 0.136 |

Note: This table estimates impacts of the partial school reopening on respondent hours spent in the last 7 days on work and on childcare. Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample, though results are similar when they are included. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. We interact treatment with baseline wealth before the pandemic, defined using an index of household wealth and housing characteristics. Coefficients for 'Above mean baseline wealth' represent the sum of the $Post \times Treat$ and $Post \times Treat \times Above\ mean\ baseline\ wealth$ terms. We include $p$-values for tests of whether the interaction terms are equal to 0. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.1.6: Impacts of partial school reopening on respondent total work hours by sex and selected characteristics

| Interaction term Z | None | | | Any child age 0-4 | | | Any child ag labor before reopening | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post × Treat | 5.894** | 5.215* | 7.613** | | | | | | |
| | (2.330) | (2.987) | (3.814) | | | | | | |
| Post × Treat, Z=0 | | | | 8.326*** | 8.094** | 8.837* | 3.557 | 1.951 | 6.892 |
| | | | | (3.156) | (4.092) | (5.313) | (2.769) | (3.599) | (4.398) |
| Post × Treat, Z=1 | | | | 2.964 | 1.606 | 6.679 | 10.849** | 13.607*** | 8.534 |
| | | | | (3.453) | (4.188) | (5.689) | (4.211) | (5.071) | (6.456) |
| Observations | 2288 | 1331 | 902 | 2288 | 1331 | 902 | 2288 | 1331 | 902 |
| Mean, pre-reopen control | 19.272 | 17.211 | 22.478 | 19.272 | 17.211 | 22.478 | 19.272 | 17.211 | 22.478 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men |
| p-value, diff. by Any child age 0-4 | | | | 0.252 | 0.268 | 0.782 | 0.148 | 0.061 | 0.834 |

Note: This table estimates impacts of the partial school reopening on respondent total work hours in the last 7 days. Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. The first three columns show average estimated impacts, and the following show estimated impacts from regressions fully interacted with a specified household characteristics. Coefficients for $Z = 1$ represent the sum of the $Post \times Treat$ and $Post \times Treat \times Z$ terms. We include $p$-values for tests of whether the interaction term is equal to 0. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

## C.2 Major pandemic policy changes in Kenya

The following outline summarizes when major nation-wide pandemic-related policies were implemented and relaxed over the course of 2020 after the first COVID-19 cases in Kenya on March 13. The dates for the announcements of new restrictive policies are in *italics* and the dates when these policies were relaxed or ended are in **bold**. We also include announcements related to school closures, even though policies did not necessarily change with these announcements. Most policies were extended multiple times after first being imposed; we do not list the dates of policy extensions, except for school closures.

- *March 13-20*
  - Suspend all public gatherings, meetings, games, events
  - Ban on gatherings of more than 10 people
  - All schools closed
  - Recommend working from home where possible
  - Ban on foreigner entry; quarantine requirements for entry of nationals and visa holders
  - Public transport asked to reduce to 60% of capacity
- *March 24-27*
  - Ban on national and international flights
  - Closure of bars and restaurants for in-person service
  - Direct cash payments implemented for vulnerable citizens
  - Stay at home requirements imposed, except for 'essential' trips
  - Curfew imposed from 1700 to 0500 hours
  - Public transit closed between 'infected' and 'not infected' areas
- April 26: School closures extended to June 4
- **April 27**: Partial reopening of restaurants for take-out service
- June 6: School closures extended until further guidance from the Ministry of Health
- **June 7**: Nightly curfew revised to between 2100 and 0400 hours

- June 24: Announcement that school might reopen on September 1
- **July 7**
  - Phased reopening of religious gatherings
  - Up to 100 people permitted to attend weddings and funerals
  - Local air travel within Kenya to resume July 15
  - International air travel to resume August 1
- July 7: Announcement that schools will remain closed until January 2021, final exams are cancelled, and students would repeat the year; colleges and universities following strict guidelines might reopen in September
- ***July 27***
  - Restaurants reopened, must close by 1900 hours
  - Ban on sale of alcoholic drinks and beverages in eateries and restaurants
- **August 27**
  - Restaurants may remain open until 2000 hours
  - Ban on sale of secondhand clothing lifted
  - Licensed hotels may sell alcohol
- September 15: Ministry of Education releases guidelines for safe reopening of schools
- September 21: Ministry of Education calls all teachers to report back to schools by September 28
- **September 27**
  - Nightly curfew revised to between 2300 and 0400 hours
  - Bars may reopen; restaurants and eateries may sell alcohol; bars, restaurants, and eateries may remain open until 2200 hours
  - Religious gatherings may open for up to 1/3 of capacity
  - Up to 200 people may attend funerals and weddings
- October 6: Ministry of Education announces that students in examination grades (4, 8, and 12) shall return to classes on October 12

- **October 12**: Students in examination grades (4, 8, and 12) to return to classes
- *November 4*
  - Requests for government work to be done remotely when possible
  - Political gatherings suspended
  - Nightly curfew revised to between 2200 and 0400 hours
  - Bars, restaurants, and eateries must close by 2100 hours
- November 4: Announcement that schools to fully reopen in January 2021
- **January 4**: Schools fully reopen

Other policies were implemented that specifically affected certain parts of the country. For example, on April 6 the government instituted a 21 day movement ban/lockdown for Nairobi, Kilifi, Kwale, and Mombasa, and Mandera was added soon after. This lockdown was extended multiple times. These were the only counties affected. The lockdowns for Kilifi and Kwale ended on June 7 and those for Nairobi, Mombasa, and Mandera ended on July 8.

Sources: COVID-19 government response timeline for Kenya; Kenya COVID Tracker; Presidency of Kenya; Kenya Ministry of Education Twitter feed

## C.3 Data

Data come from the Kenya COVID-19 Rapid Response Phone Surveys (RRPS), collected by the Kenya National Bureau of Statistics with support from the World Bank and the University of California at Berkeley. Pape et al. (2021) describe the survey methodology and implementation in detail.

The main RRPS sample is drawn from the nationally representative Kenya Integrated Household Budget Survey (KIHBS) conducted in 2015-2016: 9,009 households that were interviewed and provided a phone number served as the primary sampling frame for the RRPS. All households in the sample were targeted in each round regardless of whether they were reached in a previous round. By the fourth round of the RRPS, 5,499 KIHBS households had been successfully surveyed at least once. The KIHBS sample is supplemented by random digit dialing (RDD). From a sampling frame of 5,000 randomly selected numbers, of which 4,075 were active, 1,554 households had completed at least one survey by round four.

The sample is intended to be representative of the population of Kenya using cell phones. In the 2019 Kenya Continuous Household Survey 80% of households nationally report owning a mobile phone, though certain counties—notably in the northeast—have much lower mobile phone penetration. Pape et al. (2021) report that KIHBS households that provided a phone number and those that were successfully surveyed in the RRPS have better socioeconomic conditions—measured by housing materials and asset ownership—than households that did not provide a phone number or that did but were not reached for the RRPS.

The RRPS data include household survey weights adjusting for selection and differential response rates across counties and rural/urban strata, attempting to recover national representativeness. We do not apply these household weights for our individual-level regression analyses, but do apply them for population-level inference based on our results.

The surveys include information on household composition, assets and housing, labor outcomes for household adults, and child schooling and care, as well as more general household information and COVID-specific modules. Detailed questions on child care, schooling, labor, and other outcomes are included for a randomly selected child in each round. We use data from the first four rounds of the RRPS, covering May 2020-March 2021, and also construct measures for February 2020, before the first COVID-19 cases in Kenya using recall questions from the first time a household was surveyed. Each round lasted approximately 2.5 months and covered a representative cross-section of households each week within each wave.

Data on childcare arrangements for a randomly selected child include questions on which household member has primary responsibility for the child's care, which household member was with the child in the last 15 minutes, and where and in whose company the child stayed during the day when out of school (from a set of general categories).[1] The surveys also ask respondents for their hours spent on childcare in the last 7 days.[2] Childcare hours from other household adults are measured starting in round 3, while childcare hours from all household children combined and all non-household members combined are included in round 4 only, after schools fully reopened.

Labor supply is captured using modules on household agricultural production, household enterprise, and wage employment. For both household agricultural production and for each household enterprise, respondents report all household adults engaged in those activities over the last 7 days and their hours of work. Wage employment is reported for each household adult. An individual not working in a given activity is coded as working 0 hours. Recall on labor supply for February 2020 is only available for survey respondents.

Respondents report estimates of total household earnings from agricultural activities and from household enterprises over the last 14 days. For the few households with multiple enterprises, we sum earnings and profits across enterprises. Wage earnings in the last 14 days are reported for individual wage workers. For comparability with the measures of household agriculture and enterprise earnings, we aggregate wage earnings to the household level. Earnings data are limited—for all activities the 90th percentile of household earnings in the analysis sample is 0—in part due to a focus on the last 14 days, which does not accommodate seasonality or other variability in earnings.

We winsorize reported individual hours of work and household earnings across activities at the 99th percentile. We winsorize reported childcare hours at 140 per week.

[1] Respondents are instructed to select all childcare arrangements used. Nevertheless, respondents might omit types of childcare that are used less frequently or that are seen as less socially acceptable (e.g., leaving a child at home by themselves).

[2] The survey asks "In the last 7 days, how many hours did you spend doing childcare?" and does not distinguish between time actively spent caring for a child and time spent on other activities while responsible for a child. We topcode reported childcare hours at 140, or 20 hours a day. Over 15% of respondents in our analysis sample indicate spending at least this many hours on childcare.

## C.4 A simple static partial equilibrium model of childcare and labor supply

Here, we develop a simple model of adult labor supply and childcare decisions to generate predictions to take to the data. The model considers a static problem for adult household members with children. For simplicity, we assume that household adults take decisions jointly, and thus model the decision as that of a single person. We focus on a static, partial equilibrium labor supply decision, and set aside possible impacts of shocks to labor demand to focus on shocks to the childcare burden. Key aspects of the context that we aim to reflect in the model are the availability of child labor in household agriculture, as well as childcare of younger children by older children. To reflect women's larger role in childcare in this setting, we assume that female adult household members have either a comparative advantage in childcare of younger children or, similarly, that social norms are such that the costs of refraining from childcare for women (or the cost of engaging in childcare for men) are exogenously larger. Moreover, in line with our data, we assume economies of scale in childcare provision, as well as the ability of the caregiver to combine some types of childcare with household production in agriculture.

Household adults get utility $U = U(C, L, \{Q_k\}_{k=1}^N)$ from consumption, leisure, and the well-being of $k$ household children. They can spend time on leisure, wage work for a fixed wage, home production[3], and childcare and face a time constraint $T = L + t_w + t_h + t_c$. Wage work earns a wage $w$. Home production $H = H(t_h, X, S)$ is a concave function satisfying the Inada conditions, which depends on the adult time input and other household characteristics $X$ such as the availability of household agricultural land or existence of a household enterprise, as well as the availability of child labor (which is a function of the age distribution of children and school closure policy $S$). We normalize the price of consumption to 1 and assume that household production can be sold at this same price.

Adults provide childcare $CCM = \psi_g\big(t_c + \theta_h t_h\big)$, which includes "active" childcare time $t_c$ focused on children as well as some portion $\theta$ of home production time that simultaneously provides passive childcare. $\psi_g$ is a cost-shifter for childcare provision that takes on takes on smaller values for men than for women.[4] Adult childcare is a public good that all household children can access, reflecting the economies of scale we observe in this context.

[3]The key distinction this model makes is between a work sector which accommodates both simultaneous childcare and child labor contribution, and another work sector which does not. We therefore primarily think of the home production activity as being household agriculture, with household enterprise activities being a form of home production which has some of the same characteristics but less so.

[4]This cost-shifter can be rationalized in multiple ways. It could be that women's childcare hours count for more relative to men's due to a social expectation of women to provide childcare, or a social stigma of childcare for men. This shifter is also isomorphic to a model where women are relatively more productive in

Total childcare for child $k$ is given by $CC_k = CCM + I_k$, adults' childcare provision plus any childcare provided by older children.[5] All children receive the same amount of care from household adults, but childcare from other children varies based on the age distribution of siblings. Total informal childcare available to the household $I = I(X, S)$ is a function of household characteristics (notably the presence of older siblings) and of school closures; older siblings provide more informal childcare when schools are closed. Child well-being $Q_k = Q_k(CC_k, \bar{CC}_k(age_k, S))$ is a concave function of childcare provided to the child and their minimum required care.[6] Minimum required care $\bar{CC}_k(age_k, S)$ decreases with age, and for school-age children it increases when schools are closed.

Adults take as given household characteristics $X$ such as presence of other adults and child composition, school closure policy $S$, and non-labor income $Y$. We model $S$ as a binary variable taking a value of 1 if schools are closed and 0 otherwise.[7]

Adults' static optimization problem is

$$\max_{t_w, t_h, t_c} U(C, L, \{Q_k\}_{k=1}^{N}) \tag{C.1}$$

Subject to

$$C = wt_w + H(t_h, X, S) + Y \tag{C.2}$$

$$T = L + t_w + t_h + t_c \tag{C.3}$$

$$L \geq 0;\ t_w \geq 0;\ t_h \geq 0;\ t_c \geq 0 \tag{C.4}$$

$$CC_k = \psi_g\big(t_c + \theta_h t_h\big) + I_k \tag{C.5}$$

$$\sum_{k=1}^{N} I_k \leq I(X, S) \tag{C.6}$$

$$Q_k = Q_k(CC_k, \bar{CC}_k(age_k, S)) \tag{C.7}$$

Adults maximize their utility over choices of time use, subject to the following constraints: 1) the household budget constraint, 2) their time endowment, 3) non-negativity constraints on time use, 4) the childcare provision function, 5) availability of childcare from older siblings, and 6) the child well-being function. The budget constraint states that spending on consumption must equal the sum of wage income, the value of home production, and non-labor income.

childcare, and require fewer hours to achieve the same increase in child welfare.

[5]Very few households in the data use non-household sources of childcare, so we abstract away from this possibility.

[6]We can think of this as a Stone-Geary type of function.

[7]$S$ may vary by child, as in the case of the partial school reopening in Kenya, but we abstract from this point.

We are interested in the impacts of changes in school closure policy $S$ on adult labor supply $t_w$ and $t_h$. $S$ enters the model through childcare needs, the availability of sibling childcare, and household child labor availability. We expect $\bar{CC}_k(age_k, 1) > \bar{CC}_k(age_k, 0)$ for children enrolled in school, meaning schools being open decreases household childcare needs. On the other hand, $I(X, 1) > I(X, 0)$ for households with older children, as those children can provide more informal childcare when schools are closed. Moreover, $H(t_h, X, 1) > H(t_h, X, 0)$ as children may contribute more to home production when they are home from school.

Adults thus trade off their time among wage work, home production, childcare, and leisure, at an interior solution setting the marginal returns to each equal to each other:

$$u'_C w = u'_C H'_t(X, S) + \phi(\{age_k\}_{k=1}^N, S, \psi_g)\theta_h = \phi(\{age_k\}_{k=1}^N, S, \psi_g) = u'_L \qquad \text{(C.8)}$$

where these terms are, respectively, the marginal utility of working one more hour for a wage $w$, the marginal utility of an additional hour in home production (providing both consumption and some child well-being value due to joint work and childcare), the marginal value to adults of an additional hour of childcare, and the marginal utility of an additional hour of leisure, and where we define

$$\phi = \sum_{k=1}^{N} u'_{Q_k} Q'_k(age_k, S)\psi_g \qquad \text{(C.9)}$$

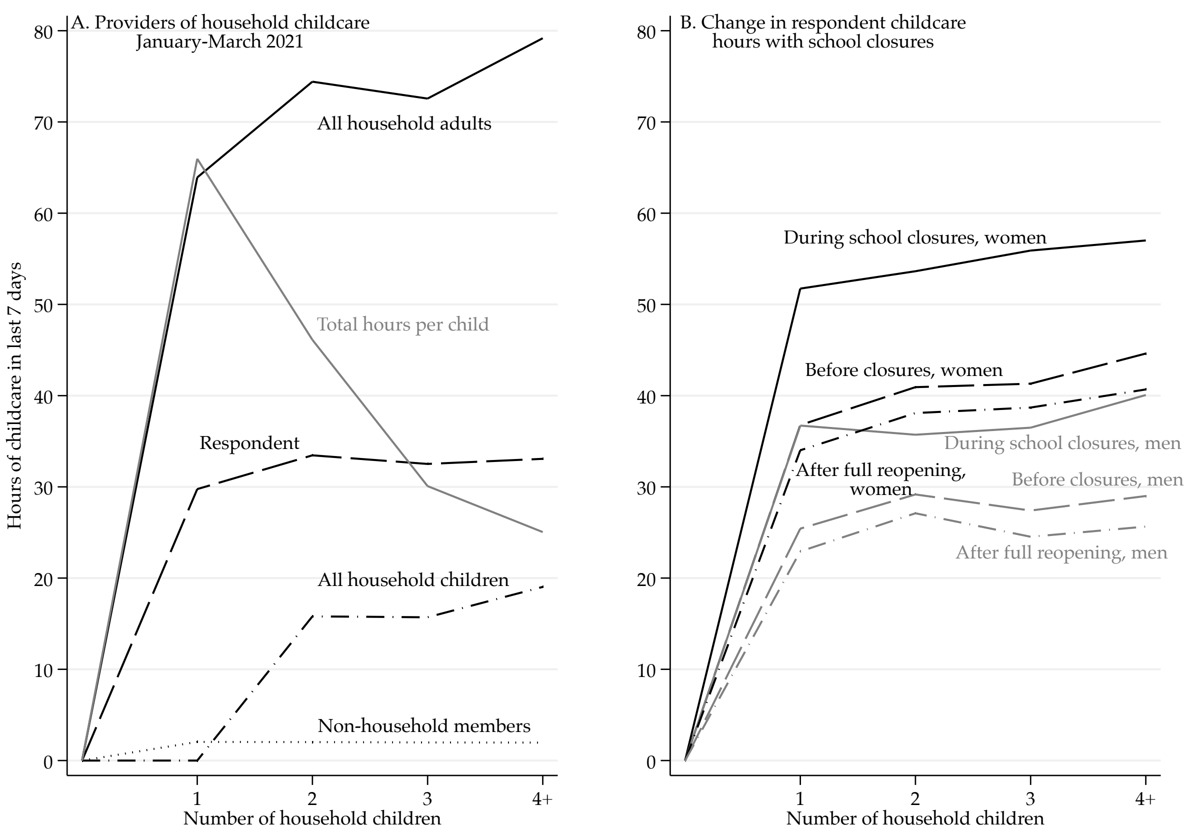
School reopening (moving from $S = 1$ to $S = 0$) affects the solution to adults' problem through two channels. First, it lowers child labor in home production, thus likely raising the marginal return of adults working in home production after school reopenings ($H'_t(X, 0) > H'_t(X, 1)$. Second, it changes the net demand (and thus net return) to childcare for the remaining children. When there are no other children present, the utility return to adult childcare hours decreases for sure ($\phi(K = 1, S = 1) < \phi(K = 1, S = 0)$, which also reduces the return to work in home production. But when there are other children present, the sign of this effect is ambiguous and depends on the age distribution of remaining children. When remaining children are very young, for instance, the marginal utility from adult childcare may actually *increase*, since the school reopening decreases sibling childcare ($I(X, 1) > I(X, 0)$) and changes the household childcare constraint.

The simple model thus allows us to generate several hypotheses for the effect of schools reopening to take to the data. First, in households where *all* children return to school, hours spent in childcare should decline while labor supply overall should increase. Second, in agricultural households which used child labor while schools were closed, adult labor supply after school reopening is furthermore expected to shift relatively towards agriculture to make up for lost child labor. Third, adults with young children will likely be providing a higher level of childcare on average when schools are open: sibling childcare falls as older siblings return to school, but childcare needs remain high since these are driven relatively more by younger children. This matters given mothers' ability to care for multiple children simultaneously, suggested by the economies of scale in childcare hours we observe in the data. Labor supply may therefore decrease in households with young children given the increased marginal utility from adult childcare.

Finally, if women play a larger role in childcare (as is the case in our setting), either because children benefit more from female care, or because social norms and economic circumstances are such that men's social cost to childcare are larger, then $\psi_{female} > \psi_{male}$. This suggests that women should engage in more childcare overall, and that school closures should impact their childcare hours more. They will also be more likely to supply relatively more labor to home production, as this can be combined with childcare whereas wage work cannot. If, in addition, we suppose that childcare needs and norms for infants and young children are particularly gendered, then labor supply responses to school reopening when young children remain present in the household should be particularly muted (or even reversed) for women relative to men.

## C.5 Robustness

Figure C.5.1: Childcare hours in the last 7 days among all households with school-age children, by provider of care and school closure status




Note: The figures match Figure 3.1 but include all households with school-age children. The figures show mean childcare hours in the last 7 days among all households with at least one school-age child by number of household children and by time period.
Panel A presents data from RRPS round 4 (January-March 2021) which asks about childcare hours for each household adult, for all children in total, and for all non-household members in total. Previous rounds only ask about childcare hours for the respondent. The hours for 'all household adults' include the respondent's hours.
Panel B presents data for female (black) and male (gray) *respondent* childcare hours over time as school closure policies changed. Data on childcare hours before and during the school closures period for other care providers are not available.

Table C.5.1: Impacts of partial school reopening on respondent labor supply

| | N | Control Mean (SD) | Post x Treat (SE) | Post x Mixed (SE) |
|---|---|---|---|---|
| Engaged in any work in last 7 days | 2939 | 0.621 | 0.072 | 0.047 |
| | | (0.485) | (0.050) | (0.050) |
| Engaged in wage employment in last 7 days | 2939 | 0.084 | 0.006 | -0.003 |
| | | (0.277) | (0.024) | (0.022) |
| Engaged in HH agriculture in last 7 days | 2939 | 0.518 | 0.075 | 0.030 |
| | | (0.500) | (0.052) | (0.048) |
| Engaged in HH non-ag enterprise in last 7 days | 2939 | 0.109 | 0.039 | 0.011 |
| | | (0.312) | (0.026) | (0.027) |
| Total work hours, last 7 days | 2939 | 18.756 | 5.812** | 2.182 |
| | | (21.509) | (2.336) | (2.192) |
| Wage hours, last 7 days | 2939 | 2.606 | 0.589 | 0.699 |
| | | (10.514) | (1.083) | (0.979) |
| Ag hours, last 7 days | 2939 | 12.353 | 4.184** | 1.451 |
| | | (15.707) | (1.642) | (1.554) |
| Enterprise hours, last 7 days | 2939 | 3.925 | 1.230 | -0.124 |
| | | (13.127) | (1.412) | (1.248) |

Note: This table presents estimates of Equation 3.1 for survey respondent labor supply. Individuals not working in a given sector are coded as working 0 hours. From left to right, the columns show the dependent variable, number of observations, the control mean prior to the partial reopening, and the impacts of being in the partial reopening period for treatment households (Post x Treat) and mixed households (Post x Mixed). Impacts for control households are absorbed by month fixed effects. Control households have a child in grades 3, 5, 6, 7, or 9, treatment households have a child in grades 4 or 8, and mixed households have both. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. Standard errors are clustered at the household level. Data include observations for adults age 18-64 from May to November 2020. Significant treatment impacts on total and agricultural work hours are robust to multiple testing adjustment using FDR q-values.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.2: Impacts of partial school reopening on total working hours, varying controls

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Post | -2.251*** | -3.489*** | | | | | |
| | (0.831) | (0.960) | | | | | |
| Post x Treat | 1.048 | 4.836*** | 4.749*** | 3.243* | 4.865*** | 4.737*** | 3.954*** |
| | (1.688) | (1.839) | (1.825) | (1.884) | (1.832) | (1.821) | (1.376) |
| Post x Mixed | 0.613 | 0.595 | 0.594 | -1.126 | 0.538 | 0.610 | -0.011 |
| | (1.601) | (1.824) | (1.799) | (1.821) | (1.805) | (1.804) | (1.362) |
| Observations | 8587 | 8538 | 8538 | 8538 | 7765 | 8538 | 8538 |
| Mean, pre-reopen control | 16.486 | 16.494 | 16.494 | 16.494 | 16.383 | 16.494 | 16.494 |
| Household FE | N | Y | Y | Y | N | Y | Y |
| Individual FE | N | N | N | N | Y | N | N |
| Month FE | N | N | Y | N | Y | Y | Y |
| County-month FE | N | N | N | Y | N | N | N |
| Individual controls | N | N | N | N | N | Y | Y |
| Household controls | N | N | N | N | N | N | Y |

Note: This table presents estimates of Equation 3.1 with varying controls. The dependent variable is total working hours over the last 7 days, taking a value of 0 for individuals not working. Observations include data from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat'), control households with children in an adjacent grade, and 'Mixed' households with both. 'Post' is a dummy for being observed on or after the partial school reopening on October 12.
Column 3 is the primary specification. Individual controls include sex, age, and household head status. Household controls include number of adults, young children (age 0-4), and school-age children (5-17) in the household, and dummies for engagement in agriculture and in enterprise. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.3: Impacts of partial school reopening on total working hours, varying sample

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Post x Treat | 4.577** | 4.271** | 4.085** | 4.739*** | 4.060** | 4.924** | 4.552** | 4.357** |
| | (2.065) | (1.823) | (1.975) | (1.825) | (1.754) | (1.992) | (1.988) | (2.004) |
| Post x Mixed | -0.166 | 0.799 | -0.021 | | 0.422 | 1.491 | 0.378 | -0.335 |
| | (1.982) | (1.755) | (1.995) | | (2.044) | (1.659) | (1.661) | (1.622) |
| Observations | 5369 | 8998 | 5929 | 6575 | 7641 | 9252 | 8824 | 9407 |
| Mean, pre-reopen control | 17.482 | 16.390 | 17.608 | 16.494 | 16.660 | 16.594 | 16.433 | 16.043 |
| Sample | Prime age adults age 25-50 | All adults age 18+ | Potential primary care-givers | No mixed households | Grade 6 out of ctl | Grade 2 in ctl | Grade 2 in, Grade 9 out of ctl | Grade 12 in trt, Grades 10-11 in ctl |

Note: This table presents estimates of Equation 3.1 with varying samples. The main sample includes adults ages 18-64. We test robustness to focusing on adults age 25-50 (the most likely to be parent caregivers and engaged in work), to including all adults aged 18 and up, and to including only adults identified as potential parents—between 14 and 55 years older than the oldest household child—or sole caregivers (the only household adult). We also test the robustness to varying the grades included in the definition of control households. Omitting grade 6 students from the control definition focuses just on students in grades immediately adjacent to those treated. Including grade 2 students in the control group mirrors the inclusion of grade 6 relative to grade 8. Omitting grade 9 students prevents comparing primary to secondary school students. Adding grade 12 students to the treatment definition and grade 10 and 11 students to the control definition expands the definition to include all grades eligible for the partial reopening. The dependent variable is total working hours over the last 7 days, taking a value of 0 for individuals not working. Observations include data from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat'), control households with children in an adjacent grade, and 'Mixed' households with both, unless otherwise indicated. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.4: Robustness to defining post by timing of reopening announcement, 21 Sept 2020

| | (1) Any work | (2) Wage work | (3) HH ag. | (4) HH ent. | (5) Total hrs | (6) Wage hrs | (7) HH ag hrs | (8) HH ent hrs |
|---|---|---|---|---|---|---|---|---|
| Post × Treat | 0.028 | 0.005 | 0.031 | 0.011 | 3.269** | 0.054 | 2.586** | 0.632 |
| | (0.032) | (0.011) | (0.032) | (0.015) | (1.354) | (0.511) | (1.088) | (0.598) |
| Observations | 8717 | 8717 | 8717 | 8717 | 8717 | 8717 | 8717 | 8717 |
| Mean, pre-reopen control | 0.604 | 0.062 | 0.530 | 0.065 | 16.784 | 2.139 | 12.455 | 2.306 |

Note: This table presents estimates of variations of Equation 3.1, but *Post* is defined not by the date schools reopened on 12 October 2020 but by the timing it was announced, 27 September.
Dependent variables are defined over the last 7 days, and take a value of 0 for individuals not working in a particular activity. Observations include data from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat'), control households with children in an adjacent grade, and 'Mixed' households with both. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.5: Impacts of partial school reopening on work and childcare hours by returning child grade and presence of young children

| Respondent hours in last 7 days in | All work activities | | | | All childcare activities | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Post x Grade 4 | 4.959 | 6.307 | 8.651 | 11.339 | 1.186 | -9.339 | -16.489 | 11.870 |
| | (4.629) | (5.299) | (5.916) | (9.341) | (10.545) | (11.502) | (10.970) | (23.900) |
| Post x Grade 4 × Any child age 0-4 | | | -10.620 | -6.956 | | | 47.030* | -34.685 |
| | | | (9.200) | (11.338) | | | (24.417) | (27.313) |
| Post x Grade 8 | 7.902** | 7.481 | 10.343** | 7.714 | 8.544 | -7.235 | 0.652 | -4.433 |
| | (3.220) | (5.614) | (4.981) | (6.441) | (9.455) | (8.512) | (14.344) | (9.527) |
| Post x Grade 8 × Any child age 0-4 | | | -3.973 | -0.272 | | | 19.202 | -3.250 |
| | | | (6.522) | (13.070) | | | (19.071) | (19.710) |
| Observations | 1317 | 892 | 1317 | 892 | 1314 | 891 | 1314 | 891 |
| Mean, pre-reopen control | 17.211 | 22.478 | 17.211 | 22.478 | 60.059 | 41.175 | 60.059 | 41.175 |
| Sample | Women | Men | Women | Men | Women | Men | Women | Men |

Note: This table estimates impacts of the partial school reopening on respondent hours spent in the last 7 days on all work activities (columns 1-4), and on childcare (columns 5-8). Observations include data for survey respondents from May to November 2020, and include households with at least one child in grades 3-9. We separately estimate the impact of the partial reopening for grades 4 and 8. 'Mixed' households with children in both a treated grade and an adjacent grade and households with children in both treated grades are dropped from the sample, though results are similar when they are included. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.6: Robustness of analysis of childcare mechanism to including mixed households

| Respondent last 7 days hours in | HH agriculture work | | | HH enterprise or wage work | | | Childcare | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post x Treat | 4.595** | 3.915 | 4.844 | 3.706* | 4.332 | 3.880 | -5.703 | -10.049 | 0.395 |
| | (2.190) | (2.568) | (4.150) | (2.101) | (2.846) | (3.150) | (6.222) | (8.365) | (9.465) |
| Post x Treat × Any child age 0-4 | -0.837 | 0.498 | -1.192 | -4.779 | -7.246* | -1.912 | 17.242* | 32.167** | -10.193 |
| | (3.300) | (4.043) | (5.715) | (3.596) | (4.261) | (5.448) | (10.242) | (14.449) | (14.828) |
| Post x Mixed | 1.930 | 1.254 | 2.537 | -1.289 | -1.777 | 1.050 | -0.601 | -6.662 | 10.419 |
| | (2.338) | (2.917) | (4.041) | (1.958) | (2.365) | (3.460) | (7.429) | (9.725) | (11.357) |
| Post x Mixed × Any child age 0-4 | -0.622 | 1.948 | -4.096 | 3.311 | 4.683 | 1.361 | 1.974 | 19.163 | -22.313 |
| | (3.134) | (4.123) | (5.095) | (2.965) | (3.877) | (4.775) | (9.919) | (13.435) | (14.765) |
| Observations | 2939 | 1710 | 1163 | 2939 | 1710 | 1163 | 2935 | 1707 | 1162 |
| Mean, pre-reopen control | 12.947 | 11.702 | 14.827 | 6.325 | 5.509 | 7.651 | 52.747 | 60.059 | 41.175 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men |

Note: This table estimates impacts of the partial school reopening on respondent hours spent in the last 7 days on household agricultural labor (columns 1-3), wage employment and household non-agricultural enterprise labor (columns 4-6), and childcare (columns 7-9). Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households include both types of children. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.7: Robustness of analysis of child agriculture mechanism to including mixed households

| Last 7 days hours in | HH agriculture work (Respondent) | | | HH enterprise or wage work (Respondent) | | | Household agriculture (All household children) | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post x Treat | 4.184** | 5.262** | 1.731 | 1.628 | 3.282* | 5.435** | -0.673 | -1.006 | -0.554 |
| | (1.642) | (2.162) | (2.804) | (1.744) | (1.927) | (2.594) | (1.001) | (1.317) | (0.960) |
| Post x Treat × Any child ag labor before reopening | | | 8.738** | | | -5.036 | | | -0.078 |
| | | | (4.358) | | | (3.856) | | | (2.479) |
| Post x Mixed | 1.451 | 2.019 | 2.478 | 0.731 | 1.016 | 3.112 | -1.725 | -2.000 | 0.735 |
| | (1.554) | (1.891) | (2.689) | (1.505) | (1.493) | (2.083) | (1.175) | (1.424) | (1.100) |
| Post x Mixed × Any child ag labor before reopening | | | 1.019 | | | -4.269 | | | -2.239 |
| | | | (3.776) | | | (2.945) | | | (2.558) |
| Observations | 2939 | 2300 | 2300 | 2939 | 2300 | 2300 | 2939 | 2300 | 2300 |
| Mean, pre-reopen control | 12.947 | 16.806 | 16.806 | 6.325 | 5.325 | 5.325 | 4.167 | 5.409 | 5.409 |
| Sample of households | All | Ag. HH | Ag. HH | All | Ag. HH | Ag. HH | All | Ag. HH | Ag. HH |

Note: This table estimates impacts of the partial school reopening on respondent hours of work in the last 7 days for household agriculture (columns 1-3) and other activities (wage employment and household non-agricultural enterprise, columns 4-6), and on total child agricultural labor hours in the last 7 days (columns 7-9). Columns with the 'Ag. HH' sample include households engaged in agricultural production at any point in 2020. Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households include both types of children. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.8: Baseline balance by survey timing in round 3

| | Post Oct 12 Mean | N | Pre Oct 12 Mean | N | Difference | p-value |
|---|---|---|---|---|---|---|
| *Household characteristics* | | | | | | |
| Female household head | 0.28 | 1436 | 0.30 | 208 | -0.02 | 0.582 |
| Age of household head | 45.05 | 1436 | 44.23 | 208 | 0.81 | 0.360 |
| Count adults | 2.61 | 1436 | 2.63 | 208 | -0.02 | 0.832 |
| Count of all kids age 0-17 | 3.14 | 1436 | 3.27 | 208 | -0.13 | 0.262 |
| Any child age 0-4 | 0.36 | 1436 | 0.46 | 208 | -0.11 | 0.004 |
| Household wealth index | -0.05 | 1436 | -0.16 | 208 | 0.11 | 0.140 |
| Connected to electricity grid | 0.46 | 1436 | 0.45 | 208 | 0.00 | 0.895 |
| Urban/Rural | 0.45 | 1436 | 0.57 | 208 | -0.12 | 0.002 |
| Household engaged in agriculture | 0.61 | 1436 | 0.60 | 208 | 0.01 | 0.702 |
| Any child engaged in household farm labor | 0.27 | 1436 | 0.27 | 208 | -0.00 | 0.924 |
| Household engaged in enterprise | 0.15 | 1436 | 0.20 | 208 | -0.04 | 0.128 |
| *Respondent characteristics* | | | | | | |
| Age | 40.82 | 1436 | 38.65 | 208 | 2.17 | 0.012 |
| Female | 0.58 | 1436 | 0.61 | 208 | -0.03 | 0.402 |
| Completed primary school | 0.87 | 1434 | 0.87 | 207 | -0.01 | 0.760 |
| Completed secondary school | 0.47 | 1434 | 0.47 | 207 | -0.00 | 0.897 |
| Completed school beyond secondary | 0.14 | 1434 | 0.16 | 207 | -0.02 | 0.469 |
| Married | 0.76 | 1414 | 0.72 | 207 | 0.03 | 0.345 |
| Is the household head | 0.64 | 1436 | 0.59 | 208 | 0.06 | 0.120 |
| *Respondent labor participation* | | | | | | |
| Engaged in any work in last 7 days | 0.68 | 1436 | 0.65 | 208 | 0.03 | 0.396 |
| Engaged in wage employment in last 7 days | 0.11 | 1436 | 0.08 | 208 | 0.02 | 0.259 |
| Engaged in HH agriculture in last 7 days | 0.56 | 1436 | 0.56 | 208 | -0.00 | 0.899 |
| Engaged in HH non-ag enterprise in last 7 days | 0.10 | 1436 | 0.10 | 208 | 0.01 | 0.802 |
| Engaged in any work in February 2020 | 0.83 | 1436 | 0.84 | 208 | -0.01 | 0.795 |

Note: The table presents sample sizes and means for households in the analysis sample (with a child in grade 3-9) whose round 3 survey date fell before or after the partial reopening of schools on October 12. Data are from the first time a household is observed, typically in survey round 1 (May-early July) while schools were fully closed. Individual-level data are for the survey respondent.
Columns present differences and means and $p$-values for tests of equality for households by round 3 survey timing. The joint F-stat for differences across the two groups is 1.87, with p-value 0.008. It is 2.01 ($p = 0.003$) for control households and 0.78 ($p = 0.762$) for treatment households.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.9: Impacts of partial school reopening on labor supply, dropping households surveyed before the partial reopening in round 3

| | N | Control Mean (SD) | Post x Treat (SE) | Post x Mixed (SE) |
|---|---|---|---|---|
| Engaged in any work in last 7 days | 7262 | 0.566 | 0.085* | 0.048 |
| | | (0.496) | (0.051) | (0.051) |
| Engaged in wage employment in last 7 days | 7262 | 0.058 | 0.011 | -0.006 |
| | | (0.233) | (0.012) | (0.012) |
| Engaged in HH agriculture in last 7 days | 7262 | 0.495 | 0.080 | 0.039 |
| | | (0.500) | (0.053) | (0.050) |
| Engaged in HH non-ag enterprise in last 7 days | 7262 | 0.069 | 0.014 | 0.013 |
| | | (0.253) | (0.023) | (0.020) |
| Total work hours, last 7 days | 7262 | 16.195 | 4.784*** | 0.594 |
| | | (19.862) | (1.824) | (1.799) |
| Wage hours, last 7 days | 7262 | 1.888 | 0.443 | -0.148 |
| | | (9.085) | (0.524) | (0.522) |
| Ag hours, last 7 days | 7262 | 11.836 | 4.110*** | 0.876 |
| | | (15.423) | (1.513) | (1.516) |
| Enterprise hours, last 7 days | 7262 | 2.591 | 0.315 | -0.264 |
| | | (11.155) | (0.823) | (0.790) |

Note: This table presents estimates of Equation 3.1 for adult labor supply dropping households surveyed before the partial reopening in round 3 which do not contribute to identifying the Post × Treat effect. Survey timing was randomized in round 1 and maintained in subsequent rounds. Table C.5.8 tests for balance on baseline characteristics of omitted households against the rest of the analysis sample. Individuals not working in a given sector are coded as working 0 hours. From left to right, the columns show the dependent variable, number of observations, the control mean prior to the partial reopening, and the impacts of being in the partial reopening period for treatment households (Post x Treat) and mixed households (Post x Mixed). Impacts for control households are absorbed by month fixed effects. Control households have a child in grades 3, 5, 6, 7, or 9, treatment households have a child in grades 4 or 8, and mixed households have both. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. Standard errors are clustered at the household level. Data include observations for adults age 18-64 from May to November 2020.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.10: Impacts of partial school reopening by presence of children age 0-4 and respondent sex, dropping households surveyed before the partial reopening in round 3

| Respondent last 7 days hours in | HH agriculture work | | | HH enterprise or wage work | | | Childcare | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post × Treat, No child age 0-4 | 4.579** | 3.833 | 4.774 | 3.722* | 4.098 | 4.165 | -5.817 | -11.325 | 0.127 |
| | (2.195) | (2.569) | (4.158) | (2.089) | (2.840) | (3.077) | (6.168) | (8.165) | (9.579) |
| Post × Treat, Any child age 0-4 | 3.905 | 4.620 | 4.208 | -1.035 | -3.131 | 2.236 | 11.883 | 23.786** | -11.065 |
| | (2.450) | (3.081) | (3.890) | (2.912) | (3.163) | (4.448) | (8.208) | (11.903) | (11.749) |
| Observations | 1909 | 1113 | 756 | 1909 | 1113 | 756 | 1905 | 1110 | 755 |
| Mean, pre-reopen control | 13.671 | 12.358 | 15.552 | 6.405 | 5.891 | 7.313 | 52.901 | 60.297 | 41.718 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men |
| p-value, diff. by any child age 0-4 | 0.838 | 0.844 | 0.921 | 0.185 | 0.090 | 0.722 | 0.085 | 0.015 | 0.461 |

Note: This table estimates impacts of the partial school reopening on respondent work hours dropping households surveyed before the partial reopening in round 3 which do not contribute to identifying the Post × Treat effect. Survey timing was randomized in round 1 and maintained in subsequent rounds. Table C.5.8 tests for balance on baseline characteristics of omitted households against the rest of the analysis sample. Outcomes include hours spent in the last 7 days on household agricultural labor (columns 1-3), wage employment and household non-agricultural enterprise labor (columns 4-6), and childcare (columns 7-9). Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. The first panel shows average estimated impacts, and the second panel shows estimated impacts from regressions fully interacted with a dummy for the presence of children age 0-4. Coefficients for 'Any child age 0-4' represent the sum of the $Post \times Treat$ and $Post \times Treat \times Any\ child\ age\ 0\text{-}4$ terms. We include $p$-values for tests of whether the interaction term is equal to 0. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.11: Impacts of partial school reopening by household characteristics, all adults

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $Z$ = Any child age 0-4 | | | | | | $Z$ = Any baseline child ag labor | | | | | |
| Last 7 days hours in | HH agriculture | | | HH enterprise or wage work | | | HH agriculture | | | HH enterprise or wage work | | |
| Post × Treat, $Z = 0$ | 3.352 | 2.953 | 3.967 | 0.867 | 0.869 | 0.700 | 2.442 | 2.032 | 3.171 | 0.774 | 1.066 | 0.316 |
| | (2.094) | (2.063) | (2.544) | (1.193) | (1.614) | (1.499) | (1.693) | (1.691) | (2.105) | (1.211) | (1.601) | (1.563) |
| Post × Treat, $Z = 1$ | 5.015** | 4.660** | 5.837** | 0.302 | -0.244 | 0.791 | 7.378** | 7.146** | 8.026** | 0.708 | -0.733 | 2.119 |
| | (2.177) | (2.278) | (2.368) | (1.613) | (1.837) | (2.254) | (2.929) | (2.964) | (3.156) | (1.631) | (1.722) | (2.229) |
| Observations | 6575 | 3414 | 2880 | 6575 | 3414 | 2880 | 6575 | 3414 | 2880 | 6575 | 3414 | 2880 |
| Mean, pre-reopen control | 12.167 | 11.688 | 12.623 | 4.327 | 3.579 | 4.965 | 12.167 | 11.688 | 12.623 | 4.327 | 3.579 | 4.965 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men | All | Women | Men |
| p-value, diff. | 0.582 | 0.579 | 0.591 | 0.778 | 0.649 | 0.973 | 0.145 | 0.134 | 0.201 | 0.974 | 0.445 | 0.508 |

Note: This table estimates impacts of the partial school reopening on adults hours of work in the last 7 days in household agriculture and in other activities (wage employment and household non-agricultural enterprise), by sex of the adult and household characteristics. Regressions are fully interacted with a dummy variable $Z$. Coefficients for $Z = 1$ represent the sum of the $Post \times Treat$ and $Post \times Treat \times Z$ terms. We include $p$-values for tests of whether the interaction term is equal to 0. Observations include data for adults age 18-65 from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.12: Impacts of partial school reopening by household characteristics, respondents

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $Z$ = Any child age 0-4 | | | | | | $Z$ = Any baseline child ag labor | | | | | |
| Last 7 days hours in | HH agriculture | | | HH enterprise or wage work | | | HH agriculture | | | HH enterprise or wage work | | |
| Post × Treat, Z=0 | 4.568** | 3.831 | 4.769 | 3.758* | 4.263 | 4.068 | 1.273 | -0.453 | 2.541 | 2.284 | 2.404 | 4.350 |
| | (2.195) | (2.570) | (4.158) | (2.094) | (2.848) | (3.073) | (1.807) | (2.167) | (3.368) | (2.187) | (2.759) | (3.275) |
| Post × Treat, Z=1 | 3.905 | 4.614 | 4.270 | -0.941 | -3.008 | 2.409 | 10.342*** | 16.068*** | 6.854 | 0.508 | -2.462 | 1.680 |
| | (2.450) | (3.097) | (3.901) | (2.916) | (3.178) | (4.410) | (3.322) | (4.430) | (4.787) | (2.826) | (2.165) | (4.050) |
| Observations | 2288 | 1331 | 902 | 2288 | 1331 | 902 | 2288 | 1331 | 902 | 2288 | 1331 | 902 |
| Mean, pre-reopen control | 12.947 | 11.702 | 14.827 | 6.325 | 5.509 | 7.651 | 12.947 | 11.702 | 14.827 | 6.325 | 5.509 | 7.651 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men | All | Women | Men |
| p-value, diff. | 0.840 | 0.846 | 0.930 | 0.191 | 0.089 | 0.758 | 0.017 | 0.001 | 0.462 | 0.619 | 0.166 | 0.609 |

Note: This table estimates impacts of the partial school reopening on respondent hours of work in the last 7 days in household agriculture and in other activities (wage employment and household non-agricultural enterprise), by sex of the adult and household characteristics. Regressions are fully interacted with a dummy variable $Z$. Coefficients for $Z = 1$ represent the sum of the $Post \times Treat$ and $Post \times Treat \times Z$ terms. We include $p$-values for tests of whether the interaction term is equal to 0. Observations include data for survey respondents from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.13: Impacts of partial school reopening by household characteristics, all adults in HHs with at least one male and female

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $Z$ = Any child age 0-4 | | | | | | $Z$ = Any baseline child ag labor | | | | | |
| Last 7 days hours in | HH agriculture | | | HH enterprise or wage work | | | HH agriculture | | | HH enterprise or wage work | | |
| Post × Treat, $Z = 0$ | 3.668* | 4.080* | 3.308 | 0.404 | 0.031 | 0.640 | 2.805 | 3.283* | 2.443 | 0.416 | 0.516 | 0.272 |
| | (2.216) | (2.230) | (2.473) | (1.262) | (1.731) | (1.541) | (1.822) | (1.886) | (2.046) | (1.204) | (1.500) | (1.605) |
| Post × Treat, $Z = 1$ | 5.314** | 5.106** | 5.795** | 0.648 | 0.096 | 1.106 | 7.363** | 6.856** | 8.238*** | 1.069 | -0.512 | 2.465 |
| | (2.343) | (2.584) | (2.377) | (1.554) | (1.518) | (2.237) | (3.050) | (3.155) | (3.164) | (1.691) | (1.875) | (2.200) |
| Observations | 5952 | 2874 | 2792 | 5952 | 2874 | 2792 | 5952 | 2874 | 2792 | 5952 | 2874 | 2792 |
| Mean, pre-reopen control | 12.551 | 12.304 | 12.702 | 4.004 | 2.861 | 4.912 | 12.551 | 12.304 | 12.702 | 4.004 | 2.861 | 4.912 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men | All | Women | Men |
| p-value, diff. | 0.610 | 0.764 | 0.469 | 0.903 | 0.978 | 0.864 | 0.200 | 0.331 | 0.124 | 0.753 | 0.669 | 0.421 |

Note: This table estimates impacts of the partial school reopening on adults hours of work in the last 7 days in household agriculture and in other activities (wage employment and household non-agricultural enterprise), by sex of the adult and household characteristics. Regressions are fully interacted with a dummy variable $Z$. Coefficients for $Z = 1$ represent the sum of the $Post \times Treat$ and $Post \times Treat \times Z$ terms. We include $p$-values for tests of whether the interaction term is equal to 0. Observations include data for adults age 18-65 from May to November 2020 in households with at least one adult female and at least one adult male, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5.14: Impacts of partial school reopening by household characteristics, controlling for respondent gender

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $Z$ = Any child age 0-4 | | | | | | $Z$ = Any baseline child ag labor | | | | | |
| Last 7 days hours in | HH agriculture | | | HH enterprise or wage work | | | HH agriculture | | | HH enterprise or wage work | | |
| Post × Treat, $Z = 0$ | 3.427 | 3.087 | 3.991 | 0.845 | 0.852 | 0.679 | 2.536 | 2.152 | 3.218 | 0.793 | 1.096 | 0.324 |
| | (2.100) | (2.069) | (2.550) | (1.195) | (1.621) | (1.501) | (1.694) | (1.695) | (2.101) | (1.209) | (1.600) | (1.562) |
| Post × Treat, $Z = 1$ | 5.247** | 4.774** | 6.137*** | 0.440 | -0.155 | 0.953 | 7.427** | 7.171** | 8.071** | 0.786 | -0.742 | 2.278 |
| | (2.182) | (2.287) | (2.344) | (1.613) | (1.843) | (2.242) | (2.948) | (2.986) | (3.165) | (1.653) | (1.750) | (2.216) |
| Observations | 6575 | 3414 | 2880 | 6575 | 3414 | 2880 | 6575 | 3414 | 2880 | 6575 | 3414 | 2880 |
| Mean, pre-reopen control | 12.167 | 11.688 | 12.623 | 4.327 | 3.579 | 4.965 | 12.167 | 11.688 | 12.623 | 4.327 | 3.579 | 4.965 |
| Sample | All | Women | Men | All | Women | Men | All | Women | Men | All | Women | Men |
| p-value, diff. | 0.548 | 0.584 | 0.536 | 0.840 | 0.682 | 0.919 | 0.151 | 0.144 | 0.202 | 0.997 | 0.438 | 0.471 |

Note: This table estimates impacts of the partial school reopening on adults hours of work in the last 7 days in household agriculture and in other activities (wage employment and household non-agricultural enterprise), by sex of the adult and household characteristics. Regressions are fully interacted with a dummy variable $Z$ and control for respondent gender. Coefficients for $Z = 1$ represent the sum of the $Post \times Treat$ and $Post \times Treat \times Z$ terms. We include $p$-values for tests of whether the interaction term is equal to 0. Observations include data for adults age 18-65 from May to November 2020, and include treatment households with children in grades 4 or 8 (indicated by 'Treat') and control households with children in an adjacent grade. 'Mixed' households with both types of children are dropped from the sample. 'Post' is a dummy for being observed on or after the partial school reopening on October 12. Regressions include household and month fixed effects. SEs clustered at household level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$