From Swarms to Strife: the Desert Locust and Conflict in Africa

Elias Farnleitner*

September 2023

Abstract

This paper represents the first empirical study on the role of a migratory pest in civil violence in Africa. Based on a disaggregated analysis, we exploit plausible exogenous within cell × year variation in the occurrence of desert locust swarms using panel data from 2000 to 2021 at a resolution of 1° latitude × 1° longitude. We estimate conflict outcomes for hexagonal cells as a function of swarms, a set of controls and both temporal and spatial dependencies. While the results of our baseline models suggest that swarms occurring during the growing season are negatively correlated with conflict incidence, we also find that conflict intensity tends to increase following a locust shock. The main transmission channel appears to be farm income.

JEL codes: C33, D74, O13, Q12, Q54

Keywords: Africa, agricultural shocks, civil violence, climate change, desert locust, disaggregated analysis, migratory pest

*Paris School of Economics, EHESS (farnleitner.elias@psestudent.eu).

I am sincerely grateful to Oliver Vanden Eynde for his supervision, expertise, and accessibility throughout the process of completing this Master's thesis. His valuable guidance played a crucial role in shaping the outcome of this work. All the remaining mistakes are my own.

I also would like to thank Jérémie Gignoux, Liam Wren Lewis, Keith Cressman, Milos Popovic, François Libois and participants at the PSE Development research seminar (especially Hannes Tepper) for their constructive comments and engaging discussions, which greatly enriched this research. Also, without the support of my family, this would have been much more difficult than it already was - a thousand thanks for that. And finally, I am infinitely grateful to Marita, my better half, for her unconditional support during our time in Paris. Without you this would not have been possible.

Contents

1 Introduction **1**

1.1 Related literature 2

2 Conceptual Framework **5**

2.1 The desert locust 5

2.2 Sample and cell specification 7

3 Empirical Framework **9**

3.1 Identification strategy 9

3.2 Models 13

4 Data **16**

5 Results & Discussion **18**

6 Mechanisms **22**

6.1 Swarms in growing season drive intensity rather than incidence 22

6.2 Heterogeneity among crop types and dependence on crop 23

6.3 Heterogeneity among conflict actors 25

7 Robustness **26**

7.1 IV with hoppers as instrument 26

7.2 Adding luminosity data 27

7.3 Alternative cell shape and size 28

7.4 Alternative sample criterion 28

7.5 Excluding conflict outliers 29

7.6 Alternative conflict variable 29

7.7 Alternative spatial weight matrix, temporal and spatial lags 29

7.8 Anticipation of swarms 30

8 Conclusion **30**

Table 1: List of Abbreviations.

Abbreviation	Definition
ACLED	Armed Conflict Location and Event Data Project
AR_LOC	A numeric variable that captures the number of desert locust swarms within the growing season of the cell's main crop in a given cell and year ("agriculturally relevant swarms")
CROP_AREA	A numeric variable that captures the cultivated area in hectares of the cell's main crop in a given year
GDP	Gross Domestic Product
FAO	Food and Agriculture Organization of the United Nations
FAO_AREA	A numeric variable that captures the treatment area in hectares of FAO-monitored ground operations to control the desert locust population in a given cell and year
FATAL	A dummy variable that indicates if a cell has experienced non-zero fatalities due to conflicts in a given cell and year
FE	Fixed Effects
HOA	Horn of Africa
HOP	A numeric variable that captures the number of clustered occurrences of desert locust hoppers in a given cell and year
IV	Instrumental Variable
LOC	A numeric variable that captures the number of desert locust swarms in a given cell and year
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation
NA	Non-available data
OLS	Ordinary Least Squares
SD	Standard Deviation
SE	Standard Errors
SPEI	Standardized Precipitation Evapotranspiration Index
TWFE	Two-Way Fixed Effects
WMO	World Meteorological Organization
WWF	World Wild Fund

1 Introduction

Conflict is a major obstacle to economic development in many regions of the world. In recent years, considerable research efforts have been made to investigate the causes of civil violence.[1] The focus has often been on the role of negative economic shocks, with particular attention paid to commodity price fluctuations and anomalies in temperatures or rainfall as triggers of conflict. However, in the face of climate change, other climate-related shocks may play a more important role in the future - such as migratory pests, especially in the form of locust swarms.[2]. To the author's knowledge, this paper represents the first empirical study about the role of desert locust swarms in civil violence in Africa.

Locust swarms are a fascinating and at the same time frightening phenomenon. Their ability to move in perfect harmony, as if controlled by an invisible force, provides an impressive spectacle. First and foremost, however, locust swarms pose a significant risk to agriculture and have devastating consequences far beyond. As one of the oldest migratory pests, they have had a lasting impact on human societies, both demographically and culturally. Ancient written records from the time of the Assyrians around 3200 BC document devastating locust plagues that were often interpreted as a form of divine punishment. Even in modern times, locust swarms are still feared, wreaking widespread havoc and posing a significant threat to regional food security.

The last major locust event, an upsurge[3], occurred from 2019 to 2021. It was triggered by a combination of factors, including unusually heavy rains and cyclones that created ideal breeding conditions for locusts.[4]. The locusts multiplied rapidly, forming large swarms

[1] In this study, the terms *conflict* and *civil violence* are used interchangeably unless otherwise noted. We define them as a form of collective violence between a government and an organized non-state actor *or* between organized non-state actors, that arises because of disagreements related to governance, resources and territory. Notably, civil violence excludes interstate conflict and encompasses a wide range of events, from peaceful protests to more violent forms of insurgency - see section 4.

[2] Migratory pests, typically insects, regularly move between locations in search of food, habitat, or breeding sites, causing crop damage and harming other resources along their routes (IPPC (2021)).

[3] The Food and Agriculture Organization of the United Nations (FAO) introduced several terms to better describe locust population dynamics: *Recessions* are periods without widespread locust infestations. Swarming populations are minimal, and mainly involve small groups or hopper infestations. *Outbreaks* happen when locust numbers multiply, leading to gregarious behavior and the formation of hopper bands and swarms. *Upsurges* involve a widespread increase in locust numbers, triggering simultaneous outbreaks and multiple breeding seasons. *Plagues* are severe infestations affecting extensive areas for consecutive years, with large numbers of swarms and hopper bands causing significant damage (Van Huis et al. (2007)).

[4] In particular, tropical cyclones Mekunu and Luban in May and October 2018 caused intense precipitation events in the Arabian Peninsula. Throughout 2019, locust swarms gradually spread to Yemen, Saudi Arabia, and Iran. Eventually, the locusts crossed the Red Sea, reaching East Africa (Retkute et al. (2021)).

that devastated crops and vegetation as they moved across several countries. The Greater Horn of Africa region and Yemen were particularly hard hit. Additionally, the COVID-19 pandemic further complicated the response efforts and hindered the movement of personnel and resources. Although the assessment of the damage has not yet been fully completed, the FAO estimates that in 2020, the peak of the upsurge, up to 2.5 million people suffered a significant income shock (FAO (2021)). And according to the World Bank, the resulting costs could amount to almost $10 billion in East Africa and Yemen alone (World Bank (2021)). All this exacerbates living conditions in a region where many people are already struggling to make ends meet.

1.1 Related literature

While research on the adverse impacts of locust swarms on agricultural production has made considerable progress, our understanding of their broader impacts on economic development, migration, education, etc. is still at an early stage. The primary transmission of the locust shock, however, is often similar: locusts lead to a reduction in output from agricultural production, which in turn decreases income. The absence of a functioning crop failure insurance market further exacerbates this situation, so that alternative adjustment mechanisms are subsequently activated. Studies have shown that locust shocks can lead to an increase in school dropouts (De Vreyer et al. (2015)), contribute to the deterioration of children's health (Le and Nguyen (2022)) or that it might incentivize farmers to shift their production towards more robust but less valuable crops (Marending and Tripodi (2023)).

In the context of conflicts, there is only anecdotal evidence (The Guardian (2020)). To the author's knowledge, no study to date has comprehensively attempted to measure the effects of locust swarms on conflict. Based on evidence from other areas of conflict research, however, two main mechanisms are conceivable following an economic shock. On the one hand, scholars supporting the *opportunity cost channel* argue that conflicts are more likely to increase when agriculture becomes less profitable (eg. Fearon and Laitin (2003), Miguel et al. (2004)). Lower returns to agriculture mean that the opportunity costs of participating in violence or joining armed groups have become lower for people engaged in agriculture. Consequently, individuals may find it more appealing to invest their time and resources in conflict-related activities, which may offer alternative economic or social opportunities. Thus, the decline in agricultural profitability is seen as a contributing factor

to the escalation of conflicts. On the other hand, the *rapacity channel* suggests that conflict decreases when income goes down (Fearon (2005)). The idea behind is that conflicts arise when individuals or groups have a desire to capture and control resources, often driven by greed or economic motivations. When overall income levels decline, the potential gains from engaging in violence or resource capture diminish, resulting in a decrease in conflicts.

After all, empirical research has highlighted that the context is crucial and both channels can dominate. In particularly, Dube and Vargas (2013) find that the effect of an income shock actually depends on the returns to labor, ie., shocks in areas that are relatively more dependent on the production of labor-intensive goods increase civil violence through the opportunity cost mechanism, while shocks in areas that are relatively more dependent on the production of labor-unintensive goods such as natural resources increase civil violence through the rapacity mechanism. With this in mind, one would expect the opportunity cost channel to predominate in the case studied here, as African agriculture is typically labor intensive. Nevertheless, it is difficult to assess a priori the impact of a locust-induced income shock. This is because other mechanisms could also play an important role. For instance, locust swarms may create a common threat perception among affected communities. When faced with a shared crisis, communities often unite to confront the threat and cooperate in mitigating its impact. In this case, the presence of a common enemy in the form of locust swarms could potentially foster a sense of cohesion and collaboration among communities. This potential mechanism suggests that locust swarms may reduce conflict by promoting collective action and cooperation. Support for this *common enemy channel* idea, albeit from a different perspective, comes from Depetris-Chauvin et al. (2020), who finds a conflict-reducing and nation-building effect of victories by a national football team.

In any case, exploring the connection between locust swarms and conflicts adds to the understanding of conflicts dynamics. In particular, this paper contributes to several parts of the literature. First and foremost, to the author's knowledge, this is the first work ever to quantify the link between a migratory pest and civil violence. The paper establishes a new, small strand in the literature on nature-induced conflict. Thematically, it is most comparable to the strand of conflicts triggered by vector-borne diseases (Cervellati et al. (2017), Gonzalez-Torres and Esposito (2016), Cervellati et al. (2022)). However, there are notable differences, eg. in terms of their transmission dynamics and control strategies. Filling this gap is therefore a valuable extension to existing knowledge. In a broader sense, this paper is also a contribution to the ongoing debate on the link between climate and

conflict (eg. Burke et al. (2009), Buhaug (2010), Burke et al. (2015)).

This paper also contributes to the studies that conduct disaggregated analyses to study drivers of civil violence (eg. Besley and Reynal-Querol (2014), Michalopoulos and Papaioannou (2016), Almer et al. (2017), Cervellati et al. (2022)). By breaking down the data into smaller units, researchers gain a more precise understanding of the conflict patterns. This fine granularity can uncover variations that might be masked in aggregated analyses. Moreover, the use of hexagons as a cell shape is a feature that has rarely been used in this literature. The advantage over square cells is that hexagons always have borders of equal length to neighboring cells. Assuming that the locust shock is transmitted in space circularly rather than rectangularly, this then allows for more accurate and effective modeling of spatial spillover effects.

Finally, this paper also contributes to improve the temporal coverage of a climate-related shock using a technique Harari and La Ferrara (2018). Because the main link between locust shocks and conflict is believed to be through farm income, we attempt to isolate the component of locust swarm fluctuations that is more relevant to agriculture. That is, we isolate the locust swarms that directly affect agriculture. Rather than using annual averages, we analyze the occurrence of locust swarms during the growing season when crops are most sensitive to adverse conditions. This has been firstly done by Harari and La Ferrara as they where trying to better capture droughts that are more relevant to farm income. As they noted, this approach is data intensive, requiring both within-year variation in the timing of shocks and spatial variation in plant cover to identify the main crop of a cell. To the author's knowledge, there is only one study (Cervellati et al. (2022)) that has replicated this procedure.

The remainder of this paper is structured as follows. Section 2 presents the conceptual framework that forms the basis for the analysis. In section 3, we present the empirical framework outlining the identification strategy and models employed. Section 4 is devoted to the data sources and presenting descriptive statistics of our main variables. In section 5, we present and discuss the main results and in section 6, we investigate mechanisms and heterogeneous effects. Section 7 provides robustness checks, and section 8 concludes with a summary of the most important findings.

2 Conceptual Framework

2.1 The desert locust

The desert locust, *Schistocerca gregaria*, is often considered the most destructive locust type in the world because it can reproduce rapidly and travel long distances. Therefore, it is also the best documented and most monitored migratory pest. In general, locusts belong to the grasshopper family *Acridadae*, but they substantially differ from common grasshoppers. In particular, they differ in their adaptability to respond to new environmental stimuli, which allows them to change their physiology and behavior (FAO and WMO (2016)).

The lifespan of locusts can vary widely from three to five months due to various factors such as food competition, weather conditions, and other environmental factors. The life cycle consists of three stages: the egg, the hopper, and the adult stage. During the egg stage, it takes about two to three weeks for the eggs to hatch. After hatching, locusts are initially restricted to crawling on the ground, hence the name "hopper", which lasts about 30 to 40 days. During this time, locusts go through several sub-stages and increase in size and weight. Upon reaching maturity, locusts enter the adult stage, which is characterized by their ability to fly. This marks an important moment in their life cycle, enabling them to make longer journeys in search of food and suitable breeding sites. What sets them apart from most other species of the grasshopper family is their ability to travel long distances using wind as a vehicle, which makes their flight paths highly unpredictable. This is even enhanced when they transition from the solitary to the gregarious phase, forming swarms that can be tens of square kilometers in size.[5] To make these numbers more concrete, a swarm 1 km^2 in size can consume the same amount of food as 35,000 humans in a single day (FAO and WMO (2016)).

Moreover, swarms can fly along wind directions for up to 10 hours per day, covering distances of 90 km or more. During recessions, desert locusts are mostly restricted to the semi-arid and arid deserts of Africa, the Middle East, and southwest Asia, where annual precipitation is less than 200 mm. This area covers 16 million km^2 and includes more than 30 countries. When extreme locust events happen, however, desert locusts can spread

[5]Desert locusts usually behave as individuals when their numbers are low. When certain conditions are met, such as increased population density and limited food resources, the transition to swarm behavior occurs. However, the reasons for this are not yet fully understood. Researcher have recently identified a pheromone molecule that is believed to trigger the process of swarm formation (Vosshall (2020).

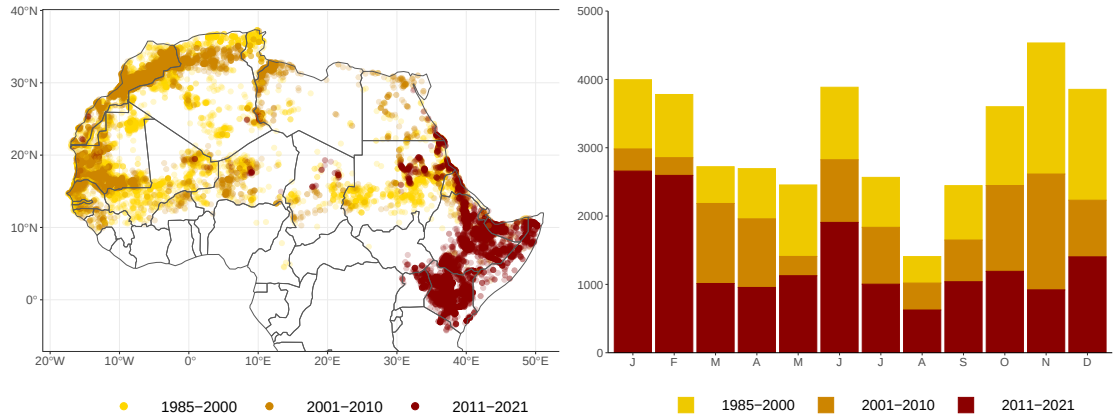

Figure 1: Spatial distribution of desert locust swarms by period (left) and desert locust swarms occurrence by month and period (right). Data Source: FAO.

to an area of up to 30 million km^2 which has the potential to affect 10% of the world's population livelihood. Also, their actual migration pattern has changed substantially in the last 20 years. As a result, many new ares, in particular in the Greater Horn of Africa, have been recently affected. This trend in the desert locusts' migration routes is captured in left panel of Figure 1 (FAO and WMO (2016)).

Since the late 1970s, the FAO has been actively engaged in coordinating and monitoring global initiatives to curb the spread of locusts. With a focus on early detection and rapid response, FAO works closely with member countries, regional partners and other international organizations to improve coordination and knowledge sharing. Various technologies such as remote sensing, drones, pesticides, etc. are used. Given the decline in locust plagues in recent decades, FAO's efforts can certainly be considered a success.[6] However, changing climate patterns due to global warming are making it more and more difficult to predict the locust population. Regions that have not previously been affected by locust swarms and therefore have no experience of controlling them are increasingly exposed to this threat, as was the case in Kenya 2019. When the first swarms arrived from the Arabian Peninsula, the country was caught off guard, having "no equipment, no expertise, no pesticides, no aircraft", according to a FAO senior (National Geographic (2021)). Therefore, one would expect locust shocks to be generally harsher in newly affected areas due to lack of expertise and resources.

[6]Figure C16 in the appendix documents this decline.

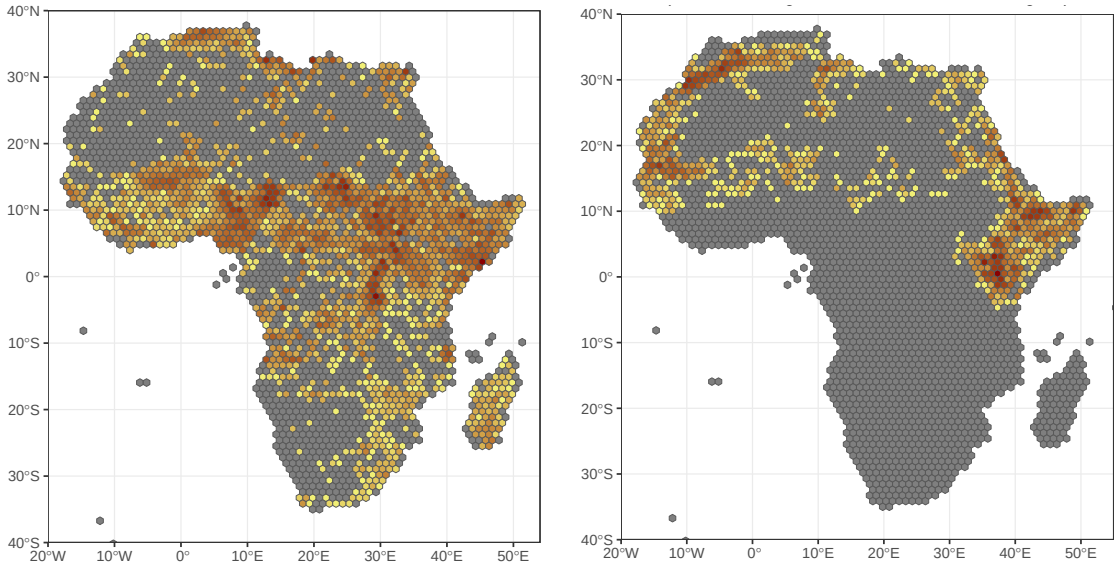

Figure 2: Spatial distribution of reported conflict fatalities (left) and desert locust swarms (right). 1.0x1.0 hexagons, aggregated data from 2000 - 2021, log-transformed. Intensity of red color corresponds to higher occurrences of reported fatalities and swarms, while grey cells indicate NA. Data Source: ACLED, FAO.

2.2 Sample and cell specification

As shown in the left panel of Figure 1, not all of Africa is affected by desert locust swarms. Limiting the analysis to their habitat is not only intuitive, but also increases the precision of the estimation by avoiding unnecessary noise. However, finding the right sample is not straightforward. Simply excluding cells that have never experienced a locust shock could lead to a bias, as the locust swarm status of an area could be influenced by an unknown third factor that may also affect conflicts. This dilemma demonstrates the need for a well-defined criterion that determines the sample - especially to restrict it in the south, since everywhere else the sample is naturally bounded by the ocean. Therefore, we propose the equatorial climate zone as the main criterion for defining our sample. More specifically, we include any cell in mainland Africa that lies above the *equatorial winter dry climate zone*, which is a tropical climate sub-type that corresponds to the Köppen-climate classification categories "Aw" (Kottek et al. (2006), see Figure B5 in appendix). This criterion seems to limit the southward migration of locusts, at least there is a strong negative correlation between its climatic conditions and the appearance of desert locust swarms. Yet, this criterion is not flawless, which makes it necessary to make an exception: the Horn of

Africa.[7] Most of this region does not meet the criterion as defined above but has recently been one of the hotspots for desert locusts. If the Horn of Africa were excluded from the analysis, important information would probably be lost - the exemption thus seems crucial. However, it should be noted that we perform several checks to ensure robustness, such as excluding the Horn of Africa, or another sample that we obtain with a different criterion, ie., based on the tropical forest frontier as defined in the World Wildlife Fund's list of terrestrial ecoregions in 2000 (Olson et al. (2001)) - see section 7.4. A graph of the sample can be found in the appendix (Figure B7).

For the analysis, we construct a panel at the cell $\times$ year level. The period chosen is 2000 to 2021, which is primarily due to data availability constraints. Before 2000 and after 2021, certain data that are crucial for our analysis do not exist. As for the shape of the cells, we are among the first in the conflict literature not to favor the squares as the main specification. To ensure more accurate modeling of spatial spillovers without complex weight matrices, we use hexagons in our baseline sample, as they have been used in environmental science for a long time (see Birch et al. (2007) for an overview). Hexagonal cells have the advantage that each neighboring cell has a boundary of equal length. This property makes the use of a binary contiguity matrix, where only direct neighbors are included, more plausible. As cell size we have chosen $1°$ latitude $\times$ $1°$ longitude, which corresponds to an area of about 10,530 km^2 per cell.[8] This eventually leads to a sample size of 35,706 (ie., 1623 cells in each of the 22 years). The choice of grid cell resolution is not trivial, as there is no theoretical guidance. The degree of localization of locust shocks and the extent of spatial conflict transmission are therefore empirical questions. To validate the choice of the cell specification, we conduct the analysis with different cell shapes and at both higher and lower spatial scales - see section 7.3.

[7] This includes Djibouti, Eritrea, Ethiopia, Kenya and Somalia. Djibouti, Eritrea and most of Somalia would have been included in the sample even without the exemption.

[8] Since the cells are hexagonal, the correct notation for the size is actually different, namely $1° \times 1° \times 1°$. For the sake of simplicity, we stick with $1° \times 1°$.

3 Empirical Framework

3.1 Identification strategy

To evaluate the impact of desert locust swarms on conflict, we rely on plausible exogenous variation in their occurrence and use a two-way fixed effects (TWFE) estimator to quantify the parameters of interest. In the last two decades, desert locust migration patterns, which were previously somewhat consistent, have shown considerable variation, presumably due to the effects of climate change (Salih et al. (2020), Guan et al. (2021)). Migration routes have become increasingly unpredictable, which makes it more reasonable to assume that locust swarms occur unanticipated. Moreover, the temporal distribution within the year has also changed considerably, so that their occurrence is increasingly shifting towards the beginning of the year (ie., January) - see the right panel of Figure 1. Further support for this claim is derived from the regression results presented in Table 2. Specification (1) clearly shows a significant relationship between a one-year increase in swarm occurrence and a shift in occurrence toward the beginning of the year, implying that swarming tends to occur earlier in the year as time progresses. In specification (2), we also include a squared term and the findings reveal a positive correlation in the initial period, suggesting that swarms occurred later in the year. However, this effect diminishes over time, and from 2001 onwards[9], we observe a reversion to a negative correlation. The findings of this nonlinear model provide even stronger evidence for unexpected occurrence of desert locust swarms.

Table 2: Relationship between desert locust swarm occurrence by start month and year (1985 - 2021).

	Month of swarm occurrence	
	(1)	**(2)**
Year of swarm occurrence	-0.069^{***}	0.048^{***}
	(0.001)	(0.007)
Year of swarm occurrence^2		-0.003^{***}
		(0.000)
Constant	8.138^{***}	7.354^{***}
	(0.038)	(0.060)
Observations	37,935	37,935

Estimated by OLS • Heteroscedasticity-corrected SE in parenthesis • Year of occurrence is standardized with 1985 ≡ 1, 1986 ≡ 2, etc. • Month of occurrence is a numeric variable, with Jan. ≡ 1, Feb. ≡ 2, etc. • Data source: FAO • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

[9] $1985 + 0.048/0.003 = 2001$

Yet, there are a number of factors that challenge our empirical strategy. The first one is the potential issue of reverse causality. Areas that have ongoing conflicts are limited in their capabilities of monitoring and fighting locust populations. Yemen, for example, has been affected by conflict for several years now. As the World Banks notes, it used to be a reliable partner on the Arabian Peninsula in curbing the spread of locust swarms before the war broke out (World Bank (2021)). In the 2019-2021 upsurge, Yemen was the epicenter for many months. Situations like this raise concerns that the presence of conflict could lead to more locust swarms, rather than the other way around. However, we show that this issue is arguably not that large. That is, we set up a simple model for the relationship between FAO-monitored control measures and conflicts. To this end, we use two specifications for the control measures: one is a binary variable that is equal to 1 if there is at least one control measure in cell i and year t, and the other is the treated area (in hectares) of the corresponding control measure in cell i and year t ("FAO_AREA"). The models also include cell and year fixed effects and, for some specifications, lags of the conflict variable. The coefficient of interest is that of "FATAL"[10], which gives a somewhat ambiguous picture - as can be seen in Table 3. While specifications (1) and (2) suggest a negative correlation between conflicts and ground operations, specifications (3) and (4) show the opposite. The intuition behind these results may be that the dummy dependent variable in the first two models captures the extensive margin of FAO-monitored ground operations, while the dependent variable, "FAO_AREA", in the other models captures the intensive margin. This suggests that conflict may reduce the number of ground operations, but this effect could be counteracted by a more concentrated and intense treatment (ie., larger treated areas). One effect may be offset by the other - but at the very least, there is no clear evidence that conflict unambiguously affects the occurrence of swarms. Nevertheless, it is essential to acknowledge that the evidence presented does not definitively rule out the possibility of reverse causality. Therefore, to address any remaining concerns, we also propose an instrumental variable strategy, for which we exploit a feature of desert locusts that is exogenous to civil violence - see section 7.1.

Another challenge arises from the correlation between locust swarms and crop yields. As mentioned above, the dynamics of locust migration are complicated and depend on various environmental factors, including temperature and precipitation. These very factors also play an important role in agricultural production, especially in rainfed regions like most

[10]"FATAL" takes the value 1 if the cell had at least one conflict-related death in a given year and 0 otherwise.

Table 3: Relationship between FAO-monitored ground operations and conflict.

	Y = 1 if FAO_AREA > 0		Y = FAO_AREA	
	(1)	**(2)**	**(3)**	**(4)**
FATAL	-0.010^{**}	-0.008^{**}	253.023^{**}	258.423^{**}
	(0.004)	(0.004)	(103.723)	(101.303)
FATAL_{t-1}	0.006		25.224	
	(0.004)		(103.711)	
LOC	0.007^{***}	0.007^{***}	456.226^{***}	456.241^{***}
	(0.000)	(0.000)	(3.359)	(3.358)
Year FE	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓
Observations	35,705	35,706	35,705	35,706

Estimated by MLE • Cell-level clustered SE in in parentheses • Data source: FAO
• * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

cells in our sample. It has been observed that a year with a higher occurrence of locust swarms can coincide with better agricultural conditions due to increased rainfall. Not accounting for this correlation could lead to bias in the estimates. Therefore, we include a drought index in our models that captures the variation in climatic conditions.

An additional aspect that needs attention stems from recent findings in the literature on TWFE regressions. Researchers have shown (De Chaisemartin and d'Haultfoeuille (2020), Goodman-Bacon (2021), Sun and Abraham (2021)) that when using TWFE and deviating from the canonical case (ie., two periods, two groups, one of which is treated in the second period and the other is never treated), an additional assumption is required for a well-identified treatment effect. That is, they must be homogeneous across groups and periods, meaning that there must be no dynamic changes in treatment effects. But it is unlikely that this is the case in the setting we are studying (eg. affected cells may have different adaptive capacities and locust swarms represent a continuous rather than binary treatment). As these findings are still quite recent, there is not yet an estimator that fits the framework of our work, ie., that takes into account the different timing of (continuous) treatments, spatial and temporal spillover effects etc. This is a reason why our result should be treated with caution. Nevertheless, there is a bright spot: when the never-treated group increases, the bias of TWFE decreases. In our sample, the fraction of never-treated cells is relatively high, ie., 59%, which mitigates the identification concerns raised by these recent econometric developments.

And finally, the estimates could also be biased if we do not take general economic activity into account. This poses two major challenges. First, obtaining reliable disaggregated data on economic activity is proving to be a difficult task, so researchers often resort to alternatives. One common approach is to use light emission during the night as a proxy. However, it is important to acknowledge that this method has its limitations. Secondly, using nighttime light data tends to overlook economic activities that are less dependent on light/electricity, such as agricultural production in impoverished regions. Therefore, we prefer not to rely on luminosity in the baseline models. For robustness, we still estimate models using data on nighttime lights - see section 7.2.

Even after addressing all these issues, the estimation additionally needs to satisfy the conditional parallel trends assumption, ie., in the absence of locust swarms, conflict incidence for cells exposed to swarms and cells not exposed would have followed parallel paths over time conditional on a set of covariates. Under this assumption, any deviation in the mean of the outcome variable observed between exposed and unexposed cells can be attributed to the effect of locust swarms. Typically, evidence in favor of this assumption is presented through a plot of the outcome for both exposed and non-exposed cells over a sufficiently long period before the swarms occurred. In our case, this is not trivial as locust swarms have existed long before the dataset's first documented occurrence. As a result, it is impossible to demonstrate parallel trends in pre-treatment data for the entire sample. Nevertheless, we can analyze a selected subsample in which swarms never occurred, or at least did not occur for a long period of time. Specifically, we can compare the average of "FATAL" for cells in the Horn of Africa region hit by the 2019-2021 upsurge with cells that have not been affected in that region. This is because most of the Horn of Africa was not affected by desert locust swarms between 1985 and 2019 (see Figure 1). In the appendix, we included two charts (C13, C14) showing a comparison of both pooled cells and cells grouped by country, where a high degree of parallelism can be observed for the pre-treatment (ie., before 2019) conflict variable. One should note, however, that this piece of evidence is neither necessary nor sufficient to ensure the method's validity.

3.2 Models

It is often observed that data gathered for points in space are not independent, but rather spatially dependent. Observations from one location tend to have similar values to those from nearby locations. As some studies show (eg. Murdoch and Sandler (2002)), this is also true for civil violence.[11] Therefore, to study the effects of desert locust swarms on conflicts, it is important to use models that account for this spatial interdependence. We chose to estimate four baseline models that are adapted from Harari and La Ferrara (2018), each of which provides increasingly sophisticated insights. The first - the most naïve - includes the number of swarms and SPEI as regressors and their lags for two years each; the second adds the number of swarms that occur in the growing season with two lags each. Building on this, the third includes a one year lag of the dependent conflict variable. While model 1 - 3 account for the spatial dimension in the error term, the fourth and most conservative model includes direct spatial dependencies between cell neighbors. To account for differences in time-invariant factors across cells (such as shared border, distance to river, road density, etc.), we include cell fixed effects in all our models. To account for country-specific time trends, we add country × year fixed effects in models 1 - 3; for model 4, we use common year fixed effects.

Before going into more details, two points need to be clarified. First, we chose to use two time lags for all our baseline models because - as we show in Table 8 the temporal spillover effect seems to fade out after two years at the latest, and also because this is a typical specification in the literature (eg. Harari and La Ferrara (2018), Cervellati et al. (2022)). Nevertheless, we have also estimated models with 1 and 3 temporal lags in the appendix - see section 7.7. Second, for the sake of easier interpretation and implementation, we chose to use a dummy variable instead of a count variable as the dependent variable in all our baseline models. However, in section 6.1, I also estimate models with a count variable, revealing an interesting mechanism.

Model 1:

$$FATAL_{c,i,t} = \beta_0 + \sum_{k=0}^{2} \beta_{1,k} LOC_{i,t-k} + \sum_{k=0}^{2} \beta_{3,k} SPEI_{i,t-k} + \mu_c + \gamma_{i,t} + \epsilon_{c,i,t} \tag{1}$$

[11] Based on a Moran's I test, we find evidence that this is the same in our sample, ie., we can reject the H_0 that there is no spatial autocorrelation in "FATAL".

Model 1 aims to show the relationship between conflicts and locust swarms in an immediate, direct way. Here c represents cell, i represents country, and t represents year. μ_c captures the cell fixed effects and $\gamma_{i,t}$ the country-specific time trends. As it can be easily integrated with spatial econometrics techniques, we use a linear probability model estimated by OLS. To account for spatial dependence in the error term, we apply a method by Hsiang (2010) based on the approach of Conley (1999). This method can be used to correct standard errors for both spatial and serial correlation. We set the spatial threshold to 180 km[12] and the temporal one to two years. Cells beyond these cut-offs are therefore considered irrelevant. In section 7.7, we estimate models with different lag specifications.

Model 2:

$$\begin{aligned} FATAL_{c,i,t} = \beta_0 + \sum_{k=0}^{2} \beta_{1,k} LOC_{i,t-k} + \sum_{k=0}^{2} \beta_{2,k} AR_LOC_{i,t-k} \\ + \sum_{k=0}^{2} \beta_{3,k} SPEI_{i,t-k} + \mu_c + \gamma_{i,t} + \epsilon_{c,i,t} \end{aligned} \tag{2}$$

Model 2 builds on model 1 by adding locust swarms in the growing season. By doing so, we can better capture the variation in swarms that is more relevant to agricultural. The corresponding coefficient, β_2, is therefore the one of most interest for us. Standard errors and fixed effects are specified and computed as before.

Model 3:

$$\begin{aligned} FATAL_{c,i,t} = \beta_0 + \phi FATAL_{c,i,t-1} + \sum_{k=0}^{2} \beta_{1,k} LOC_{i,t-k} + \sum_{k=0}^{2} \beta_{2,k} AR_LOC_{i,t-k} \\ + \sum_{k=0}^{2} \beta_{3,k} SPEI_{i,t-k} + \mu_c + \gamma_{i,t} + \epsilon_{c,i,t} \end{aligned} \tag{3}$$

Some of the observed spatial correlation in conflict is due to correlation between conflict determinants, while some is due to direct contagion. By including a one-year lag of the dependent variable, model 3 aims to separate these two effects. Standard errors and fixed effects are specified and computed as before.

[12] This corresponds to the inclusion of the six direct neighbor cells.

Model 4:

$$
\begin{aligned}
FATAL_{c,i,t} = \beta_0 + \phi FATAL_{c,i,t-1} + \rho W \times FATAL_{c,i,t} + \sum_{k=0}^{2} \beta_{1,k} LOC_{i,t-k} \\
+ \sum_{k=0}^{2} \beta_{2,k} AR_LOC_{i,t-k} + \sum_{k=0}^{2} \beta_{3,k} SPEI_{i,t-k} + \beta_4 W \times LOC_{i,t} \\
+ \beta_5 W \times AR_LOC_{i,t} + \beta_6 W \times SPEI_{i,t} + \mu_c + \gamma_t + \epsilon_{c,i,t}
\end{aligned}
\tag{4}
$$

While the method we use for handling spatial dependencies in models 1 to 3 is appropriate when spatial correlation exists within the error term, it does not account for spatial interdependence within the process itself. However, considering the likely spatial correlation among covariates and conflicts due to direct effects spilling over neighboring cells, a more sophisticated model is needed. Model 4 allows us to account for just that by considering climatic conditions, swarms, and agriculturally relevant swarms of neighboring cells and a temporal lag of the dependent variable, which therefore becomes a dynamic, spatial Durbin model that is estimated by maximum likelihood. Standard errors are corrected for clustering at the cell level.

W is a 1623×1623 row-standardized weight matrix that quantifies the connection between the cells in our sample. This is necessary to impose some structure on the spatial dependence when relaxing the conventional assumption of independent observations in a cross-section. For simplicity, we are considering only the connections between direct neighbor cells for the process we are trying to model.[13] More specifically, we assume that locust swarms beyond the six direct neighbor cells have no effect on the base cell's conflict incidence unless they trigger further local conflict, which then spills over into space. The use of hexagons instead of squares makes this assumption more reasonable. Despite this, there are feedback effects arising from the fact that marginal changes in variables within the cell lead to a series of marginal effects in other cells due to spillover effects, which can then affect the initial cell again. As a consequence, there are direct, indirect and total marginal effects. These are used to assess the signs and magnitudes of impacts arising from changes in the explanatory variables, since estimates cannot be interpreted as partial derivatives in the sense of a linear regression model. For inference, we compute standard errors based on Markov Chain Monte Carlo simulations following the approach proposed by LeSage and Pace (2009).

[13] W is therefore also referred to as a contiguity matrix of order 1.

4 Data

Since we are conducting a disaggregated analysis, it is necessary that the data have sufficient spatial granularity. Although this limits the scope - especially in the temporal dimension, we were still able to compile a high-quality cell $\times$ year panel dataset for which we mainly used four different sources. Firstly, the data on civil violence are obtained from the Armed Conflict Location & Event Data Project (ACLED, Raleigh et al. (2010)), which is widely recognized as one of the most comprehensive sources for disaggregated analyses of internal political violence. ACLED provides extensive information on various conflict indicators, including actors, event types, and the number of reported fatalities.[14] The dataset covers all of Africa from 1997 onwards. ACLED data are compiled from various sources, with a focus on reports from war zones, humanitarian agencies, and research publications. For geographical precision, all conflict events with the lowest precision level are removed from the sample.[15] To measure the overall conflict incidence, we construct a variable that sums up the reported fatalities across different conflict event types for each cell c and year t and transform it to a dummy ("FATAL"). We focus on a measure based on reported fatalities because it is a more tangible measure and easier to aggregate, which allows for direct comparisons across regions and time periods. This is particularly useful when examining spatial and temporal spillover effects. For robustness, we also use an alternative specification that is an aggregate of conflict event types ("ALLEVENT") - see section 7.6.

Secondly, information on locust events comes from the FAO Locust Hub, which provides reports on the exact location, date and various characteristics of locust assemblages from 1985 onwards.[16] This includes details of locust in both adult and hopper stages, as well as FAO-monitored ground operations, including area treated ("FAO_AREA") and pesticide use. For our baseline specification, we use the number of swarms ("LOC") from this source, which we also use to construct the number of swarms relevant to agriculture ("AR_LOC", more details below). Besides the many great possibilities offered by this dataset, there are

[14] ACLED distinguishes between the following actors: state forces/governments, rebels, political militias, identity militias, civilians, rioters, protesters and external forces. For our baseline analysis, we consider only reported fatalities from events initiated by the first four categories. Events are classified as follows: battles, explosions/remote violence, violence against civilians, protests, riots, and strategic developments - the latter is excluded. While there are other disaggregated data bases for conflict events (eg. UCDP Geolocated Event Dataset, Presidential Instability Task Force Worldwide Atrocities Dataset) the advantage of ACLED is that it also includes events that do not involve casualties. See Figure C15 for more details.

[15] That is, ACLED's "GEO_PRECIS" = 3.

[16] The data can be found here: `https://locust-hub-hqfao.hub.arcgis.com`.

also some disadvantages. For example, there is no way to distinguish locust swarms by size - complete data on swarm size have only been available for a few years. Therefore, all swarms are treated with the same weight. Furthermore, the data collection method is not able to reflect the mobility of a swarm over time. Thus, it may be that swarms are counted more than once if they have moved to a different location over time. This is not problematic as long as the swarms within a cell are not counted multiple times in a given year. The use of smaller cells, as done in section 7.3, makes this requirement more plausible and therefore also serves as a robustness check for this issue.

Thirdly, for assessing climate conditions, we rely on the Standardized Precipitation Evapotranspiration Index (SPEI) data from the Climate Research Unit Timeseries (CRU TS, Harris et al. (2020)). SPEI is a drought indicator that considers the difference between precipitation and potential evapotranspiration, incorporating temperature and rainfall. This accounts for the fact that a plant's ability to absorb water depends not only on rainfall but also on the soil's capacity to retain water. As the focus of this analysis is on seasonal droughts, we decided to take the SPEI measure at the one-month time scale. Moreover, SPEI is expressed in units of standard deviation from the historical average (ie., 1955-2021), making it comparable across time and space, where positive values indicate surplus water supply and negative values imply water scarcity.

Finally, we use data of monthly irrigated and rainfed crop areas around the year 2000 (MIRCA2000, Portmann et al. (2010)) to get information on crop coverage. This dataset contains information on the cultivated area of the 26 most widely grown crops, such as wheat, millet or rice. The indicators in the dataset represent the mean values for the period 1998-2002 and have a spatial resolution of 5 arc minutes. Conveniently, the data also include the dates of the harvest calendar, indicating the beginning and end of harvest for each crop, which may vary a lot depending on the location. As mentioned above, we use the data to construct "AR_LOC" by identifying the cell's main crop, ie., the crop with the largest acreage within the cell per year.[17] After identifying them, we use the crop calendar data to aggregate only the locust swarms that occurred within the growing season. For example, the main crop for the cell with ID 291 is barley. The month in which the growing season for barley begins at the cell location (-6.67573° longitude and 31.69215° latitude) is December, and the end is June. For the year 2004, this results in a number of 34 swarms ("LOC"), but only 17 of them fell in the main growing season of barley, which ultimately

[17]See Figure B8 for the spatial distribution of the main crops and Figure B9 for the main crop's acreage.

results in "AR_LOC". Another variable that comes from this dataset is "CROP_AREA", which captures the acreage of the main crop. It is therefore a proxy for the dependence of each cell on its main crop, which we later use to study the transmission channel.

Importantly, all of the data sources used in this study are publicly available. The summary statistics are reported in Table 4. Key findings include that the mean value of "FATAL" indicates that cells in the sample experienced an average of three years[18] with reported deaths due to conflict. The average cell also experienced desert locust swarms for 12.5 years and agriculturally relevant swarms for 3.7 years. From the negative mean of the drought index, it appears that the climate in our sample has become drier on average than in the 1955-2021 historical sample used to calculate the SPEI. To get an idea of the spatial distribution, we show the total number of reported fatalities and swarms at the cell level in Figure 2.[19]

Table 4: Summary statistic.

	N	Mean	SD
FATAL	35,706	0.14	0.34
ALLEVENT	35,706	0.17	0.38
LOC	35,706	0.57	7.98
AR_LOC	35,706	0.17	4.42
HOP	35,706	0.86	9.26
SPEI	34,700	−0.40	0.43
CROP_AREA (in ha)	35,706	36,902	56,499
FAO_AREA (in ha)	35,706	389	6,117

For some cells SPEI data are not available. This can be explained by the fact that the cells of these missings are all located on the coast and most of their area is at sea - see Figure B1 in the appendix. Also for other variables data are missing, but they can be reasonably set to 0. Note, however, that the mean is affected by this procedure. • As a reminder, the area of an 1×1 hexagonal cell is 1,053,000 ha.

5 Results & Discussion

Table 5 shows the main results of our analysis. As a reminder, of particular interest is the coefficient associated with "AR_LOC". Throughout models (1) - (4), "AR_LOC" exhibits a negative correlation, similar to "LOC" and "$W \times$AR_LOC", its spatial lag. All this suggests that swarms are associated with a decrease in conflict incidence, which in turn weakens

[18] 22 years × 0.14 = 3.08 years

[19] In the appendix, we also included one for the SPEI (Figure B1).

the opportunity cost channel and points more to the *rapacity* or *common enemy* theory as the primary mechanism(s). However, it should be noted that most of the coefficients are not significant. The coefficients of the drought index "SPEI" tell a more coherent story, as they are almost all significantly negative. This means that more favorable agricultural conditions, characterized by more rainfall and/or lower temperatures, are associated with a decrease in conflict incidence, which provides further evidence for scholars' claims that less rain and/or higher temperatures increase the risk of civil violence (eg. Burke et al. (2009), Harari and La Ferrara (2018)). Comparing the relative magnitude of "SPEI", we notice that the contemporaneous version is considerably larger than "SPEI$_{t-1}$" and "SPEI$_{t-2}$", suggesting that seasonal weather patterns translate directly into an economic shock and then slowly subside over the years. While this is in contrast to the majority of studies (eg. Miguel et al. (2004), Harari and La Ferrara (2018)), where the one-year lag is associated with the strongest effect, few others (eg. O'Loughlin et al. (2012)) observe a similar pattern. What is also striking is that conflict spillovers are significant in both time and space, with the spatial lag proving to be stronger than the temporal one. This is also somewhat unexpected, as others studies have come to opposite results (eg. Michalopoulos and Papaioannou (2016), Harari and La Ferrara (2018)).

To meaningfully interpret the coefficients of model 4, our preferred specification, we need summary measures. This is necessary because of feedback effects, i.e., marginal changes in variables within the cell lead to a series of marginal effects in other cells due to spillover effects, which can then affect the initial cell again. LeSage and Pace (2009) introduced a procedure that combines all marginal effects with three key aggregates. First, the "direct effect" measures the impact of increasing the regressor x_k by one unit within a given cell on y within the same cell, which is then averaged over all cells. Second, the "indirect effect" evaluates how increasing the regressor x_k by one unit in a particular cell jointly affects y in all other cells - also averaged over cells. And finally, the "total effect" examines how increasing the variable x_k by one unit within a particular cell on average affects y in all cells - averaged over cells where the impulse can potentially occur.

In Table 6, we present the average marginal effects of (4) from Table 4 according to LeSage and Pace (2009). In general, it is noticeable that for most coefficients the indirect effects outweigh the direct ones, indicating complex shock dynamics. Furthermore, if we compare the total average marginal effects of "LOC" and "AR_LOC", we observe that their signs seem to differ now. Specifically, a one-standard-deviation increase of swarms occurring in

Table 5: Locust swarms and conflict, baseline results.

	(1)	**(2)**	**(3)**	**(4)**
	Model 1	Model 2	Model 3	Model 4
LOC	-0.0005*	-0.0004	-0.0002	-0.0002
	(0.0003)	(0.0004)	(0.0004)	(0.0005)
LOC_{t-1}	-0.0001	-0.0003	-0.0006	0.0002
	(0.0003)	(0.0003)	(0.0004)	(0.0003)
LOC_{t-2}	0.0000	-0.0003	-0.0004	0.0004*
	(0.0002)	(0.0003)	(0.0003)	(0.0002)
SPEI	-0.0211**	-0.0210**	-0.0156*	-0.0401
	(0.0073)	(0.0073)	(0.0071)	(0.0265)
SPEI_{t-1}	-0.0080	-0.0081	-0.0141*	-0.0005
	(0.0072)	(0.0072)	(0.0073)	(0.0048)
SPEI_{t-2}	0.0025	0.0025	0.0022	0.0001
	(0.0070)	(0.0070)	(0.0068)	(0.0044)
AR_LOC		-0.0003	-0.0005	-0.0004
		(0.0007)	(0.0006)	(0.0007)
AR_LOC_{t-1}		0.0003	0.0008	0.0003
		(0.0009)	(0.0009)	(0.0007)
AR_LOC_{t-2}		0.0009	0.0005	-0.0004
		(0.0006)	(0.0004)	(0.0003)
FATAL_{t-1}			0.1873***	0.2415***
			(0.0113)	(0.0118)
$W\times$LOC				0.0008
				(0.0007)
$W\times$AR_LOC				-0.0006
				(0.0013)
$W\times$SPEI				0.0436
				(0.0278)
$W\times$FATAL				0.3459***
				(0.0100)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	
Year FE				✓
Observations	34,700	34,700	34,700	34,700

(1) - (3) are estimated by OLS, SE in parenthesis are adjusted for both spatial and serial correlation à la Hsiang (2010) • (4) is estimated by MLE, cell-level clustered SE in parenthesis • The constant is not shown • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

the growing season within a particular cell is on average associated with a contemporary decrease in conflict incidence by 1.1 percentage points in all cells. To put this in perspective, Harari and La Ferrara (2018) find that a one-standard-deviation decrease in SPEI (ie. more droughts) during the growing season is significantly associated with a 1.3 percentage point increase in conflict incidence. Note, however, that this refers to the one-year lagged variable - their main variable of interest.[20]

Table 6: Average marginal effects of the baseline model 4.

	(1)	**(2)**	**(3)**
	Total	Indirect	Direct
LOC	0.0014	0.0016	-0.0002
	(0.0013)	(0.0014)	(0.0007)
LOC_{t-1}	0.0004	0.0002	0.0002
	(0.0007)	(0.0003)	(0.0004)
LOC_{t-2}	0.0011**	0.0005**	0.0006**
	(0.0004)	(0.0002)	(0.0002)
AR_LOC	-0.0025	-0.0018	-0.0007
	(0.0031)	(0.0029)	(0.0009)
AR_LOC_{t-1}	0.0008	0.0004	0.0005
	(0.0017)	(0.0007)	(0.0010)
AR_LOC_{t-2}	-0.0010	-0.0004	-0.0006
	(0.0008)	(0.0003)	(0.0005)
SPEI	0.0088	0.0625*	-0.0537
	(0.0119)	(0.0348)	(0.0318)
$SPEI_{t-1}$	-0.0011	-0.0005	-0.0006
	(0.0117)	(0.0050)	(0.0066)
$SPEI_{t-2}$	0.0003	0.0001	0.0002
	(0.0103)	(0.0045)	(0.0059)

Computed for Model 4, Table 5 à la LeSage and Pace (2009), where SE (in parenthesis) are constructed by MCMC simulations • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Considering the existing findings, our study's results might raise a few eyebrows. Could it really be that swarms that occur during the growing season have no different effect on conflicts than swarms that occur in general? Is the economic impact of locust swarms an immediate consequence, devoid of any time lag? As it turns out, the crux of the matter lies in how we define conflict. In the next section, we show that our results are different if we move away from conflict incidence and instead use a measure that captures conflict intensity. We also examine some other mechanisms that may influence the results.

[20]The contemporary effects suggests that one-standard-deviation decrease in SPEI (ie. more droughts) during the growing season is associated with a 0.8 percentage point *decrease* in conflict incidence ($0.0207 \times 0.365 \times 100 \approx 0.8$). However, the estimate is not significant. An additional comparison with a study of vector-borne diseases shows that a one-standard-deviation increase in Ebola incidence leads to an increase by 10-17% of a standard deviation in conflict incidence (Gonzalez-Torres and Esposito (2016)). Due to the lack of certain descriptive statistics, we cannot calculate a ppt equivalent for this paper.

6 Mechanisms

6.1 Swarms in growing season drive intensity rather than incidence

Table 7: Mechanisms I.

	(1)	**(2)**	**(3)**	**(4)**
	Model 1	Model 2	Model 3	Model 4
LOC	-0.1861**	-0.2756*	-0.2467*	-0.1405
	(0.0939)	(0.1514)	(0.1452)	(0.1240)
LOC_{t-1}	0.2870	-0.3371	-0.4655*	-0.1787
	(0.3188)	(0.1801)	(0.2458)	(0.1641)
LOC_{t-2}	-0.0109	-0.0351	-0.0879	-0.0892
	(0.0329)	(0.0474)	(0.0729)	(0.1133)
AR_LOC		0.1873	0.1783	0.0477
		(0.1587)	(0.1548)	(0.0903)
AR_LOC_{t-1}		1.4987	1.6651	1.2228
		(1.1334)	(1.1900)	(1.0038)
AR_LOC_{t-2}		0.2231	0.7053	0.6402
		(0.2283)	(0.6467)	(0.6595)
$W\times$LOC				-0.1309
				(0.1174)
$W\times$AR_LOC				0.4050
				(0.3268)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	
Year FE				✓
Observations	34,700	34,700	34,700	34,700

(1) - (3) are estimated by OLS and SE (in parenthesis) are adjusted for both spatial and serial correlation à la Hsiang (2010) • (4) is estimated by MLE and SE are clustered at the cell level • Variables that are not shown include the constant, "SPEI" - "SPEI_2", "FATAL.count_1" and the corresponding spatial lags • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

When the dependent variable is defined as a count variable instead of a dummy, ie., all reported deaths related to civil violence in a cell in a given year are summed, conflict intensity rather than conflict incidence is captured. It is a proxy for how severe the conflict has become in terms of its impact on human lives. Table 7 shows the regression results of model 1 - 4 with the adjusted dependent variable. What stands out at first glance is not only the larger coefficients due to the extended interval of Y, but also the reversed signs for "LOC" and "AR_LOC". This suggests that generic swarms are linked to lower conflict intensity, while the presence of swarms during the growing season is associated with higher intensity, which also provides evidence that the main channel is through farm income. In addition, there has also been a shift in relative magnitudes, which now shows the pronounced effects of "AR_LOC_{t-1}" on conflict intensity. This supports the thesis that the effects of locust swarms translate into an economic shock only after a certain

period of time (probably due to savings). Overall, the results presented in Table 7 suggest that conflict incidence is mostly influenced by factors other than locust swarms, as their influence appears to be relatively small. Instead, the results suggest that locust swarms may play an important role in exacerbating already existing conflicts.

6.2 Heterogeneity among crop types and dependence on crop

If the assumption that the propagation of the locust shock is primarily through farm income is correct, then one should observe stronger effect in cells that are more dependent on agriculture as a source of income. Unfortunately, we do not have data from which to infer dependence directly. Instead, we have information on the proportion of each cell devoted to the main crop, referred to as "CROP_AREA". The assumption is that the higher the proportion, the more the cell depends on its main crop. Using this information, we create a variable to divide the cells into different groups (quintiles) based on this proportion, indicating the degree of their dependence on the main crop. Then we combine this variable with "AR_LOC" to see how they interact. If the presumed relationship does exist, one would expect the effects to be stronger for higher quintiles. In Table 8, we show the results in the last two columns. In (3), we investigate the mechanism with the dependent variable "FATAL" and in (4), we incorporate the findings from the previous section and use "FATAL.count" as the dependent variable. For both specifications, we can observe an initial increase in the effects before they start to decrease from the 3rd and 4th quintile, respectively. While this gives (partial) support for the farm income mechanism, it also suggests that there is a point beyond which the reliance on the main crop becomes less crucial. How can we make sense of it? One explanation could be that in cells that are heavily dependent on a single crop, crop failures can seriously threaten the existence of local farmers more seriously compared to cells that have a more balanced agriculture. Therefore, it is all the more important that they provide particular protection against external risks such as desert locust swarms. Thus, it may well be that highly dependent cells are, on average, better protected, ultimately mitigating the impact of locust shock. All this could explain why the upper edges of the distribution show lower interaction effects. Yet, to get to the bottom of this mechanism, the necessary data are not available.

Table 8: Mechanisms II.

	(1)	(2)	(3)	(4)
	FATAL	FATAL.count	FATAL	FATAL.count
LOC	0.0000	-0.3464	0.0002	-0.3184
	(0.0005)	(0.2356)	(0.0005)	(0.2259)
AR_LOC	-0.0003	0.1474	-0.0019*	0.0558
	(0.0010)	(0.4018)	(0.0011)	(0.3001)
$CA_{2^{nd}quintile}$×AR_LOC			0.0007	0.2842
			(0.0011)	(0.1956)
$CA_{3^{rd}quintile}$×AR_LOC			0.0018	0.3145
			(0.0016)	(0.2241)
$CA_{4^{th}quintile}$×AR_LOC			0.0011	0.3955
			(0.0017)	(0.3863)
$CA_{5^{th}quintile}$×AR_LOC			0.0010	0.0578
			(0.0010)	(0.3780)
Barley×AR_LOC	0.0003	0.6498		
	(0.0012)	(0.4010)		
Coffee×AR_LOC	-0.0089***	-0.3226		
	(0.0019)	(0.4477)		
Fodder_grasses×AR_LOC	-0.0001	0.1747		
	(0.0010)	(0.3054)		
Groundnuts×AR_LOC	0.0027	-1.1154		
	(0.0078)	(2.2037)		
Maize×AR_LOC	-0.0006	0.2763		
	(0.0009)	(0.2887)		
Others_an.×AR_LOC	0.0004	-0.5450		
	(0.0019)	(0.6127)		
Others_pe.×AR_LOC	0.0524*	1.2550		
	(0.0299)	(0.8449)		
Pulses×AR_LOC	0.0046	0.0079		
	(0.0058)	(1.0473)		
Rice×AR_LOC	-0.0532***	-0.1590		
	(0.0169)	(1.0413)		
Sorghum×AR_LOC	-0.0010	-0.1251		
	(0.0013)	(0.3701)		
Wheat×AR_LOC	0.0002	0.4609		
	(0.0009)	(0.3146)		
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	✓
Observations	14,066	14,066	14,066	14,066

(1) - (4) specified as Model 3, which is estimated by OLS and SE (in parenthesis) are adjusted for both spatial and serial correlation à la Hsiang (2010) • Variables that are not shown include the constant, "LOC_1", "LOC_2", "AR_LOC_1", "AR_LOC_2", "SPEI" - "SPEI_2" and "Y_1" • Only observations with "CROP_AREA" > 0 are included • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

In this context, it is also interesting to know whether there are heterogeneous effects with respect to crop types. In other words, do locust shocks spread differently among certain cultures? The first two columns of table 7 provide information on this. The procedure here is the same as before, ie., we interact the crop types with "AR_LOC", where millet,

the most used crop in our sample, is the base category. While most coefficients are not estimated precisely enough, we find that coffee and rice tend to attenuate the effects on conflict incidence and that "Others_perenial"[21] increase it compared to millet.

6.3 Heterogeneity among conflict actors

Table 9: Mechanisms III.

	(1) FATAL_idmil	(2) FATAL_polmil	(3) FATAL_state	(4) FATAL_rebel
LOC	-0.0001	-0.0002	0.0000	-0.0003
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
LOC_{t-1}	-0.0002	-0.0001	0.0004	0.0003
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
LOC_{t-2}	0.0001	0.0000	0.0003**	0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
AR_LOC	-0.0004	0.0007	-0.0008	0.0001
	(0.0006)	(0.0005)	(0.0005)	(0.0004)
AR_LOC_{t-1}	0.0002	0.0004	0.0004	0.0004
	(0.0005)	(0.0004)	(0.0007)	(0.0007)
AR_LOC_{t-2}	0.0001	0.0001	-0.0000	-0.0005
	(0.0002)	(0.0002)	(0.0002)	(0.0003)
$W \times$LOC	-0.0003	-0.0002	0.0014*	0.0009
	(0.0005)	(0.0004)	(0.0007)	(0.0006)
$W \times$AR_LOC	0.0022*	-0.0010	-0.0021*	-0.0013*
	(0.0011)	(0.0008)	(0.0011)	(0.0007)
Cell FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	34,700	34,700	34,700	34,700

(1) - (4) specified as Model 4, which is estimated by MLE and SE (in parenthesis) are clustered at the cell level • Variables that are note shown include "SPEI" - "$SPEI_{t-2}$", "$FATAL.count_{t-1}$" and the corresponding spatial lags • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Similarly, it would be interesting to know whether there is heterogeneity with respect to the perpetrating actor. To find out, we break "FATAL" into four subcategories: "FATAL_polmil" (reported fatalities from conflict events initiated by political militias), "FATAL_idmil" (reported fatalities from conflict events initiated by identity militias), "FATAL_rebel" (reported fatalities from conflict events initiated by rebel groups) and "FATAL_state" (reported fatalities from conflict events initiated by state forces).[22] This time we do not use interactions, but run the regression directly on the respective subcategory - the results are presented in Table 9. Among other things, there are considerable

[21] In the MIRCA2020 data, there is insufficient statistical information available for some perennial crops to be recorded individually, which are placed in this category.

[22] For a detailed description of the various conflict actors, please refer to: this ACLED document.

differences in the coefficients for "AR_LOC" between the models. While we can observe a positive correlation with respect to political militias, it is interesting to observe the negative estimates of "FATAL_state" and "FATAL_idmil". This suggests that the impact of locust swarms may differ across different conflict types. In the appendix, Table A13, we repeat the exercise with a count variable version as outcome. We find a similar pattern as in Table 7, but what is striking is the relationship between droughts and conflict events initiated by political militias in (2). That is, we find a strongly positive and significant coefficient for "SPEI" and a strongly negative and significant coefficient for "$W\times$SPEI", which indicates that droughts may affect the intensity of conflicts initiated by political militias differently depending on whether the shock occurs inside or outside the cell.

7 Robustness

In this section, we perform a number of robustness checks to strengthen confidence in our baseline results. The order of the checks is chosen arbitrarily.

7.1 IV with hoppers as instrument

We first address the endogeneity concerns that may result from reverse causality as outlined in section 3. As a reminder, there may be concerns that conflicts are affecting the occurrence of swarms due to ground operations not being conducted in conflict zones. To overcome this, we propose to use an instrumental variable (IV) strategy using hoppers as instrument. As described in section 2, hoppers are young and wingless locusts. They are limited in their mobility, which is why they do not move far from their breeding ground. The location where locusts lay their eggs usually depends on factors beyond human control, making it difficult to determine them in advance and even more difficult for ground operations to take countermeasures prior to hatching. As a result, efforts to control locust populations often focus on the hopper stage, when they are more developed and easier to locate as in their egg stage. Once they are detected, hoppers are documented and eventually treated by field officers.[23] Therefore, we expect the occurrence of hoppers at the time of data collection to be less influenced by conflict and thus that they are less likely to be subject to reverse causality than common desert locust swarms. The idea, then, is to use the

[23]Subsequently, this and a range of other information is then transmitted to both the National Locust Control Centre and the FAO.

presence of hoppers in a given cell and year to filter out the variation in the occurrence of swarms that is exogenous, or at least has a higher degree of exogeneity. Data on hoppers also come from the FAO. This being said, hoppers must meet some criteria to qualify as a valid instrument. First, hoppers must affect the conflict outcome only through swarms (ie., exclusion restriction). While this is not testable, the only threats we can think of could come from the adverse health effects of pesticide use on humans, which could eventually translate into conflict events (Pinstrup-Andersen and Shimokawa (2008)) as well as from the adverse effects on the environment (eg. killing other insects with locust pesticides (National Geographic (2021))), which also has the potential to affect farm income and eventually conflict outcomes. Nonetheless, given the less-documented adverse health effects on humans and on the environment, we find it reasonably to assume exclusivity. Second, the identification also requires that the average causal effect of hoppers on swarms is non-zero (ie., relevance restriction), which is by definition the case.[24] It is thus argued that the presence of hoppers is a valid instrument for swarms in a given cell and year.[25]

In Table A1, specification (1) and (2), we present the results of the second-stage for models 2 and 3[26]. While the significance of the coefficient of the instrumented variable "AR_LOC" has not changed much, the magnitude has decreased by fourfold, which suggests a positive bias in the baseline results shown in Table 5. We repeat this exercise with conflict intensity as outcome (see (3) and (4)), which gives a similar picture. The magnitude of the coefficient has increased substantially indicating a negative bias in our baseline results. This ultimately confirms the pattern that conflict incidence is less sensitive to locust shock than conflict intensity. However, these results must be taken with a grain of salt again, as they are estimated very imprecisely.

7.2 Adding luminosity data

As an additional robustness check, we add nighttime light data ("NL") as a proxy for GDP. The data come from the AfroGrid database (Schon and Koren (2022)), which only cover the period from 2000 to 2018 and thus reduce the sample size. In Figure B2, we present the

[24]For completeness, the assumptions of monotonicity and independence must also be satisfied.

[25]See Figure B3 in the appendix for their spatial distribution in Africa.

[26]The first-stage and reduced-form results are presented in Tables A2 and A3 respectively. Note also that the application of an IV for model 4 was technically not possible, as to the author's knowledge no package existed at the time of the analyses (neither in Stata nor R) that is capable of running IVs with dynamic spatial Durbin models.

spatial distribution of the nighttime light data averaged over all available years, which shows that there is very little variation. We therefore chose to use a logarithmic transformation.[27] Table A12 reveals that the results only change marginally, which increases confidence in our baseline results. Interestingly, the coefficient of log(NL+1) indicates a negative correlation while $\log(\text{NL}+1)_{t-2}$ shows a positive correlation. This could suggest that there is no rigid income-conflict mechanism that is constantly dominant over the years. That is, initially the opportunity cost channel might predominate, while the greed channel becomes dominant only over time.

7.3 Alternative cell shape and size

Next, to make sure that our results are not driven by our cell modification, we rerun our baseline models using a different cell shape and size. In Table A4, specification (1), we estimate model 4 with a new sample based on squares instead of hexagons. In (2) and (3), we perform the same estimation with a sample based on larger (2.0×2.0) and smaller hexagonal cells (0.5×0.5). For all specifications, we find no meaningful deviation from the baseline results.

7.4 Alternative sample criterion

Additionally, to verify whether the criterion we use for our baseline sample is affecting the results, we come up with an alternative criterion. That is, we include any cell in mainland Africa that lies north of the tropical and subtropical moist broadleaf forest biome as defined in the World Wildlife Fund's list of terrestrial ecoregions in 2000 (Olson et al. (2001), see figure B6), which shows a high negative correlation with locust swarm migration. The dense forest cover of the tropical rainforest appears to be a physical barrier to locust swarm migration, or at least plays a role in their dispersal. While there is also a high correlation between the climate zone criterion and rainforest frontier criterion, there are some differences in cell cover due to factors such as deforestation. It should be noted that cells in the Horn of Africa region are again excluded from the criterion. In Table A5, we present the results from model 1 - 4 with the rainforest frontier sample, which show no substantial qualitative or quantitative differences from our baseline findings.

[27] That is, log("NL"+1). This transformation also ensures that we do not lose the observations that have zero nighttime light emissions.

7.5 Excluding conflict outliers

To understand the extent to which the results are shaped by particular events of concentrated violence, we repeat the analysis with a sample that excludes conflict outliers. Compared to the original sample, cells in the Horn of Africa region show significantly higher conflict rates than the rest. This is documented in Table A6, where we also show the results of a Student's t-test suggesting that the H_0 of true difference in means of "FATAL" is equal to 0 can be rejected. In Table A7, we present the results from our models excluding cells in the Horn of Africa, ie., either in Djibouti, Eritrea, Ethiopia, Kenya or Somalia. Across all specifications, we get virtually identical results.

7.6 Alternative conflict variable

We also examine whether our approach to quantifying conflict affects the results. This is due to the concern that fatality data may be biased, since they are particularly prone to manipulation for political purposes (ACLED (2021)). Hence, instead of using a dummy that indicates whether a cell has experienced one or more reported fatalities in a given year, the dependent dummy now indicates whether a cell has experienced one or more conflict events in a given year, which is denoted with "ALLEVENT". Aggregating over conflict events is the measure preferred by many papers in the literature, eg. Buhaug (2010), Dell et al. (2012), Harari and La Ferrara (2018), Cervellati et al. (2022). Table A8 documents the results of this modification. Again, no meaningful deviation from the baseline findings can be observed.

7.7 Alternative spatial weight matrix, temporal and spatial lags

Moreover, we run several versions of our baseline models with different weight matrices and lag structure to see if our results are sensitive towards this dimension. First, in Table A9, specification (1) - (3), we account for spatial dependence on first- and second-order neighbors in the error term. In (4), we do the same for model 4 but we directly model the spatial dependence through the weight matrix W. In (5) - (6), we play around with the temporal lag structure using a one-year and three-year lag of our regressors. For all specifications, we find no meaningful deviation from the baseline results.

7.8 Anticipation of swarms

Finally, we also conduct a placebo test to make sure that there has not been any anticipation effects related to future locust shocks. In the specifications presented in Table A11, we add two leads of the main explanatory variable "AR_LOC" and find no evidence that leads are significant conflict predictors.

8 Conclusion

This paper represents a pioneering investigation of the impact of desert locust swarms on civil violence in Africa from 2000 to 2021. We assemble a disaggregated panel dataset at the cell $\times$ year level and estimate models that account for spatial and temporal dependencies. We use a number of novel features in our analysis, such as hexagons as a cell shape and the replication of the technique à la Harari and La Ferrara (2018), to better link the impact of locust swarms to farm income. We also provide a comprehensive set of robustness checks to ensure that our results are not affected by confounding factors.

We find that an increase in desert locust swarms during the growing season of the cell's main crop is associated with a decrease in conflict incidence. However, a notable discrepancy emerges when we consider conflict intensity, which tends to increase with locust swarms. Our research also confirms the existing literature by showing that droughts contribute to increases in conflict incidence and by highlighting the persistent nature of conflict across time and space. In addition, we provide evidence that locust shocks are transmitted through farm income. Specifically, we find a nonlinear relationship between main crop dependence and swarms. That is, a cell's dependence on its main crop is positively related to the occurrence of conflict, but only up to a certain threshold, before it declines again. While this puzzles us, it could be the starting point for future work to better understand the factors that contribute to mitigating locust shocks.

References

ACLED (2021). Armed Conflict Location & Event Data Project (ACLED) Codebook. https://acleddata.com/acleddatanew/wp-content/uploads/2021/11/ACLED_Codebook_v1_January-2021.pdf (accessed: 29.08.2023).

Almer, C., Laurent-Lucchetti, J., and Oechslin, M. (2017). Water scarcity and rioting: Disaggregated evidence from Sub-Saharan Africa. *Journal of Environmental Economics and Management*, 86:193–209. Special issue on environmental economics in developing countries.

BBC (2020). How a single locust becomes a plague. https://www.bbc.co.uk/news/resources/idt-84994842-8967-4dfd-9490-10f805de9f68 (accessed: 29.07.2023).

Besley, T. and Reynal-Querol, M. (2014). The legacy of historical conflict: Evidence from Africa. *The American Political Science Review*, 108(2):319–336.

Birch, C. P., Oom, S. P., and Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling*, 206(3-4):347–359.

Buhaug, H. (2010). Climate not to blame for African civil wars. *Proceedings of the National Academy of Sciences*, 107(38):16477–16482.

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

Burke, M., Miguel, E., Satyanath, S., Dykema, J. A., and Lobell, D. B. (2009). Warming increases the risk of civil war in Africa. *Proceedings of the National Academy of Sciences*, 106(49):20670–20674.

Cervellati, M., Esposito, E., and Sunde, U. (2022). Epidemic shocks and civil violence: Evidence from malaria outbreaks in Africa. *Review of Economics and Statistics*, 104(4):780–796.

Cervellati, M., Sunde, U., and Valmori, S. (2017). Pathogens, weather shocks and civil conflicts. *The Economic Journal*, 127(607):2581–2616.

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

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

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.

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.

Depetris-Chauvin, E., Durante, R., and Campante, F. (2020). Building nations through shared experiences: Evidence from African football. *American Economic Review*, 110(5):1572–1602.

Dube, O. and Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *The Review of Economic Studies*, 80(4 (285)):1384–1421.

FAO (2021). Desert locust upsurge – progress report on the response in the Greater Horn of Africa and Yemen (January–April 2021).

FAO and WMO (2016). Weather and desert locusts. *WMO-No. 1175*.

Fearon, J. D. (2005). Primary commodity exports and civil war. *Journal of Conflict Resolution*, 49(4):483–507.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1):75–90.

Gonzalez-Torres, A. and Esposito, E. (2016). Epidemics and conflict: Evidence from the ebola outbreak in Western Africa. *Available at SSRN 3544606*.

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

Guan, J., Li, M., Ju, X., Lin, J., Wu, J., and Zheng, J. (2021). The potential habitat of desert locusts is contracting: predictions under climate change scenarios. *PeerJ*, 9:e12311.

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

Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7(1):109.

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).

IPPC (2021). Scientific review of the impact of climate change on plant pests – a global challenge to prevent and mitigate plant pest risks in agriculture, forestry and ecosystems.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15:259–263.

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

LeSage, J. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Chapman & Hall/CRC.

Marending, M. and Tripodi, S. (2023). Gone with the wind: Productive response to agricultural shock. *Working Paper*.

Michalopoulos, S. and Papaioannou, E. (2016). The long-run effects of the scramble for Africa. *American Economic Review*, 106(7):1802–1848.

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.

Murdoch, J. C. and Sandler, T. (2002). Economic growth, civil wars, and spatial spillovers. *Journal of Conflict Resolution*, 46(1):91–110.

National Geographic (2021). A locust plague hit East Africa. the pesticide solution may have dire consequences. https://www.nationalgeographic.com/environment/article/locust-plague-hit-east-africa-pesticide-solution-may-have-dire-consequences (accessed: 15.07.2023).

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem, K. R. (2001). Terrestrial ecoregions of the world: a new map of life on earth. *Bioscience*, 51(11):933–938.

O'Loughlin, J., Witmer, F. D. W., Linke, A. M., Laing, A., Gettelman, A., and Dudhia, J. (2012). Climate variability and conflict risk in East Africa, 1990–2009. *Proceedings of the National Academy of Sciences*, 109(45):18344–18349.

Pinstrup-Andersen, P. and Shimokawa, S. (2008). Do poverty and poor health and nutrition increase the risk of armed conflict onset? *Food Policy*, 33(6):513–520.

Portmann, F. T., Siebert, S., and Döll, P. (2010). MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, 24(1).

Raleigh, C., Linke, R., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: An armed conflict location and event dataset. *Journal of Peace Research*, 47(5):651–660.

Retkute, R., Hinton, R. G. K., Cressman, K., and Gilligan, C. A. (2021). Regional differences in control operations during the 2019-2021 desert locust upsurge. *Agronomy*, 11(12).

Salih, A. A., Baraibar, M., Mwangi, K. K., and Artan, G. (2020). Climate change and locust outbreak in east africa. *Nature Climate Change*, 10(7):584–585.

Schon, J. and Koren, O. (2022). Introducing AfroGrid, a unified framework for environmental conflict research in Africa. *Scientific Data*, 9(1):116.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

The Guardian (2020). Kenya's pastoralists face hunger and conflict as locust plague continues. https://www.theguardian.com/global-development/2020/may/15/kenyas-pastoralists-face-hunger-and-conflict-as-locust-plague-continues (accessed: 15.07.2023).

Van Huis, A., Cressman, K., and Magor, J. I. (2007). Preventing desert locust plagues: optimizing management interventions. *Entomologia Experimentalis et Applicata*, 122(3):191–214.

Vosshall, L. (2020). Catching plague locusts with their own scent. *Nature*, 584.

World Bank (2021). The World Bank Group and the locust crisis. https://www.worldbank.org/en/topic/the-world-bank-group-and-the-desert-locust-outbreak (accessed: 15.07.2023).

Appendix A

Table A1: Robustness: IV using hoppers as instrument, second-stage.

	(1)	(2)	(3)	(4)
	FATAL	FATAL	FATAL.count	FATAL.count
LOC	0.0003	0.0003	-0.4870	-0.4756
	(0.0016)	(0.0016)	(0.4203)	(0.4114)
LOC_{t-1}	-0.0003	-0.0002	-0.3199**	-0.2430
	(0.0003)	(0.0003)	(0.1689)	(0.1780)
LOC_{t-2}	-0.0003	-0.0003	-0.0411	-0.0030
	(0.0003)	(0.0003)	(0.0497)	(0.0555)
AR_LOC (HOP)	-0.0020	-0.0020	0.6949	0.6875
	(0.0036)	(0.0036)	(0.8356)	(0.8142)
AR_LOC_{t-1}	0.0002	0.0002	1.5162	1.4453
	(0.0009)	(0.0009)	(1.1405)	(1.1500)
AR_LOC_{t-2}	0.0009	0.0005	0.2261	-0.0437
	(0.0006)	(0.0006)	(0.2278)	(0.3186)
SPEI	-0.0208***	-0.0201***	1.3490	1.9022
	(0.0073)	(0.0072)	(1.6351)	(1.5688)
$SPEI_{t-1}$	-0.0080	-0.0083	1.2605	0.9914
	(0.0072)	(0.0072)	(1.3992)	(1.4093)
$SPEI_{t-2}$	0.0026	0.0017	1.1234	0.4404
	(0.0070)	(0.0069)	(1.2298)	(1.1972)
Y_{t-1}		0.0002***		0.1954***
		(0.0000)		(0.0343)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	✓
Observations	34,700	34,700	34,700	34,700

(1) & (3) are specified as Model 2 and (2) & (4) are specified as Model 3. They are all estimated by OLS and the standard errors are adjusted for both spatial and serial correlation à la Hsiang (2010)

• * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A2: Robustness: IV using hoppers as instrument, first-stage.

	(1)	(2)	(3)	(4)
	AR_LOC	AR_LOC	AR_LOC	AR_LOC
LOC	0.3878***	0.3879***	0.3878***	0.3891***
	(0.0797)	(0.0797)	(0.0797)	(0.0793)
LOC_{t-1}	-0.0315	-0.0315	-0.0315	-0.0526
	(0.0274)	(0.0256)	(0.0256)	(0.0434)
LOC_{t-2}	0.0112**	0.0111**	0.0112**	0.0104*
	(0.0049)	(0.0049)	(0.0049)	(0.0056)
HOP	0.0509	0.0509	0.0509	0.0510
	(0.0319)	(0.0319)	(0.0319)	(0.0319)
AR_LOC_{t-1}	-0.0315	-0.0314	-0.0315	-0.0099
	(0.0396)	(0.0396)	(0.0396)	(0.0488)
AR_LOC_{t-2}	-0.0054	-0.0056	-0.0054	-0.0312*
	(0.0115)	(0.0114)	(0.0115)	(0.0184)
SPEI	0.0726	0.0799	0.0726	0.1020
	(0.0788)	(0.0793)	(0.0788)	(0.0840)
$SPEI_{t-1}$	0.0178	0.0191	0.0178	-0.0085
	(0.0826)	(0.0825)	(0.0826)	(0.0885)
$SPEI_{t-2}$	0.0915	0.0906	0.0915	0.1051
	(0.0776)	(0.0775)	(0.0776)	(0.0827)
Y_{t-1}		0.1952**		0.0001
		(0.0966)		(0.0002)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	✓
Observations	34,700	34,700	34,700	34,700

(1) & (3) are specified as Model 2 and (2) & (4) are specified as Model 3. They are all estimated by OLS and the standard errors are adjusted for both spatial and serial correlation à la Hsiang (2010) • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A3: Robustness: IV using hoppers as instrument, reduced-form.

	(1)	(2)	(3)	(4)
	FATAL	FATAL	FATAL.count	FATAL.count
LOC	-0.0005	-0.0004	-0.2175*	-0.1920
	(0.0003)	(0.0003)	(0.1257)	(0.1197)
LOC_{t-1}	-0.0002	-0.0003	-0.3418*	-0.4727*
	(0.0003)	(0.0003)	(0.1827)	(0.2500)
LOC_{t-2}	-0.0003	-0.0004	-0.0334	-0.0864
	(0.0003)	(0.0003)	(0.0473)	(0.0727)
HOP	-0.0001	-0.0001	0.0354	0.0350
	(0.0002)	(0.0002)	(0.0340)	(0.0329)
AR_LOC_{t-1}	0.0003	0.0004	1.4943	1.6640
	(0.0009)	(0.0008)	(1.1341)	(1.1919)
AR_LOC_{t-2}	0.0009*	0.0007	0.2223	0.7008
	(0.0005)	(0.0005)	(0.2287)	(0.6464)
SPEI	-0.0209***	-0.0144**	1.3994	1.6264
	(0.0073)	(0.0070)	(1.6525)	(1.6278)
$SPEI_{t-1}$	-0.0080	-0.0069	1.2729	0.8127
	(0.0072)	(0.0070)	(1.3956)	(1.4605)
$SPEI_{t-2}$	0.0024	0.0016	1.1870	0.5372
	(0.0070)	(0.0066)	(1.2319)	(1.2202)
Y_{t-1}		0.1714***		0.2591***
		(0.0104)		(0.0412)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	✓
Observations	34,700	34,700	34,700	34,700

(1) & (3) are specified as Model 2 and (2) & (4) are specified as Model 3. They are all estimated by OLS and the standard errors are adjusted for both spatial and serial correlation à la Hsiang (2010)
• * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A4: Robustness: Cell size and shape.

	(1)	**(2)**	**(3)**
	1.0 × 1.0 SQ	2.0 × 2.0 HEX	0.5 × 0.5 HEX
LOC	-0.0003	-0.0000	-0.0003
	(0.0004)	(0.0002)	(0.0003)
LOC_{t-1}	-0.0000	-0.0002	0.0007
	(0.0003)	(0.0004)	(0.0006)
LOC_{t-2}	0.0005**	0.0002	0.0005**
	(0.0002)	(0.0003)	(0.0002)
SPEI	0.0057	-0.0174	0.0032
	(0.0290)	(0.0439)	(0.0080)
SPEI_{t-1}	-0.0007	-0.0038	-0.0021
	(0.0054)	(0.0131)	(0.0021)
SPEI_{t-2}	-0.0003	0.0017	0.0026
	(0.0052)	(0.0120)	(0.0019)
AR_LOC	-0.0004	-0.0000	0.0009
	(0.0008)	(0.0004)	(0.0009)
AR_LOC_{t-1}	0.0008	0.0006	-0.0002
	(0.0006)	(0.0005)	(0.0009)
AR_LOC_{t-2}	-0.0005	-0.0001	0.0006
	(0.0003)	(0.0004)	(0.0007)
FATAL_{t-1}	0.2501***	0.2358***	0.2544***
	(0.0123)	(0.0187)	(0.0092)
$W\times$LOC	0.0012*	0.0000	0.0001
	(0.0006)	(0.0003)	(0.0005)
$W\times$AR_LOC	-0.0012	-0.0000	-0.0010
	(0.0013)	(0.0006)	(0.0010)
$W\times$SPEI	-0.0050	0.0174	-0.0014
	(0.0308)	(0.0501)	(0.0083)
$W\times$FATAL	0.2873***	0.2825***	0.2605***
	(0.0108)	(0.0169)	(0.0078)
Cell FE	✓	✓	✓
Year FE	✓	✓	✓
Observations	31,680	8,414	137,412

(1) - (3) are specified as Model 4, which is estimated by MLE and SE are clustered at the cell level • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A5: Robustness: WWF rainforest frontier sample.

	(1)	**(2)**	**(3)**	**(4)**
	Model 1	Model 2	Model 3	Model 4
LOC	-0.0005*	-0.0004	-0.0003	-0.0002
	(0.0003)	(0.0004)	(0.0004)	(0.0005)
LOC_{t-1}	-0.0001	-0.0003	-0.0006	0.0002
	(0.0003)	(0.0003)	(0.0004)	(0.0003)
LOC_{t-2}	0.0000	-0.0003	-0.0004	0.0006**
	(0.0002)	(0.0003)	(0.0003)	(0.0002)
SPEI	-0.0171**	-0.0169**	-0.0133*	-0.0394
	(0.0071)	(0.0071)	(0.0070)	(0.0274)
SPEI_{t-1}	-0.0021	-0.0022	-0.0080	0.0009
	(0.0070)	(0.0070)	(0.0072)	(0.0048)
SPEI_{t-2}	0.0038	0.0037	0.0028	-0.0001
	(0.0068)	(0.0068)	(0.0066)	(0.0044)
AR_LOC		-0.0003	-0.0005	-0.0004
		(0.0007)	(0.0006)	(0.0007)
AR_LOC_{t-1}		0.0003	0.0008	0.0002
		(0.0009)	(0.0009)	(0.0007)
AR_LOC_{t-2}		0.0008	0.0005	-0.0005
		(0.0006)	(0.0004)	(0.0003)
FATAL_{t-1}			0.1707***	0.2302***
			(0.0097)	(0.0099)
$W\times$LOC				0.0008
				(0.0007)
$W\times$AR_LOC				-0.0007
				(0.0014)
$W\times$SPEI				0.0417
				(0.0288)
$W\times$FATAL				0.3418***
				(0.0090)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	
Year FE				✓
Observations	40,100	40,100	40,100	40,100

(1) - (3) are estimated by OLS and the standard errors (SE) are adjusted for both spatial and serial correlation à la Hsiang (2010), (4) is estimated by MLE and SE are clustered at the cell level. • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A6: Summary statistic for cells in Horn of Africa.

	N	Mean	SD	P-value of t-test
FATAL	4,900	0.38	0.49	0.00
ALLEVENT	4,900	0.45	0.49	0.00
LOC	4,900	1.94	15.51	0.00
AR_LOC	4,900	0.64	9.63	0.00
HOP	4,900	1.18	12.14	0.06
SPEI	4,619	−0.09	0.35	0.00
CROP_AREA (in ha)	4,900	19,673	21,704	0.00
FAO_AREA (in ha)	4,900	576	6,273	0.03

For some cells SPEI data are not available. This can be explained by the fact that the cells of these missings are all located on the coast and most of their area is at sea - see Figure B1 in the appendix. As a reminder, the area of an 1×1 hexagonal cell is 1,053,000 ha. Also for other variables data are missing, but they can be reasonably set to 0. Note, however, that the mean is affected by this procedure.

Table A7: Robustness: HOA excluded sample.

	(1)	**(2)**	**(3)**	**(4)**
	Model 1	Model 2	Model 3	Model 4
LOC	-0.0003*	-0.0004*	-0.0002	-0.0003
	(0.0001)	(0.0002)	(0.0001)	(0.0002)
LOC_{t-1}	-0.0001	-0.0003	-0.0002	0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
LOC_{t-2}	0.0000	0.0000	0.0001	0.0002
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
SPEI	-0.0271***	-0.0270***	-0.0214**	-0.0385
	(0.0071)	(0.0071)	(0.0068)	(0.0255)
SPEI_{t-1}	-0.0038	-0.0038	-0.0084	0.0059
	(0.0070)	(0.0070)	(0.0069)	(0.0047)
SPEI_{t-2}	0.0014	0.0014	0.0008	0.0002
	(0.0070)	(0.0070)	(0.0065)	(0.0043)
AR_LOC		0.0002	0.0000	0.0001
		(0.0003)	(0.0003)	(0.0003)
AR_LOC_1		0.0004	0.0004	0.0002
		(0.0004)	(0.0004)	(0.0004)
AR_LOC_{t-2}		0.0001	-0.0004	-0.0005
		(0.0003)	(0.0003)	(0.0003)
FATAL_{t-1}			0.2251***	0.2637***
			(0.0141)	(0.0146)
$W\times$LOC				0.0006
				(0.0005)
$W\times$AR_LOC				-0.0006
				(0.0009)
$W\times$SPEI				0.0371
				(0.0268)
$W\times$FATAL				0.3652***
				(0.0125)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	
Year FE				✓
Observations	29,800	29,800	29,800	29,800

(1) - (3) are estimated by OLS and the standard errors (SE) are adjusted for both spatial and serial correlation à la Hsiang (2010), (4) is estimated by MLE and SE are clustered at the cell level. • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A8: Robustness: ALLEVENT.

	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 3	Model 4
LOC	-0.0005**	-0.0002	-0.0001	-0.0000
	(0.0002)	(0.0004)	(0.0003)	(0.0005)
LOC_{t-1}	-0.0001	-0.0003	-0.0005	0.0002
	(0.0003)	(0.0003)	(0.0004)	(0.0003)
LOC_{t-2}	0.0002	0.0001	0.0001	0.0002
	(0.0002)	(0.0003)	(0.0003)	(0.0002)
SPEI	-0.0270***	-0.0268***	-0.0215**	0.0101
	(0.0076)	(0.0076)	(0.0074)	(0.0287)
SPEI_{t-1}	-0.0137	-0.0138*	-0.0182**	-0.0052
	(0.0075)	(0.0075)	(0.0075)	(0.0048)
SPEI_{t-2}	0.0058	0.0058	0.0072	0.0072
	(0.0073)	(0.0073)	(0.0070)	(0.0049)
AR_LOC		-0.0006	-0.0008	-0.0008
		(0.0007)	(0.0006)	(0.0008)
AR_LOC_{t-1}		0.0004	0.0007	0.0002
		(0.0009)	(0.0009)	(0.0008)
AR_LOC_{t-2}		0.0005	-0.0004	-0.0005
		(0.0005)	(0.0005)	(0.0004)
ALLEVENT_{t-1}			0.1971***	0.2472***
			(0.0105)	(0.0109)
$W\times$LOC				0.0004
				(0.0008)
$W\times$AR_LOC				0.0000
				(0.0016)
$W\times$SPEI				-0.0142
				(0.0302)
$W\times$ALLEVENT				0.3291***
				(0.0096)
Cell FE	✓	✓	✓	✓
Country×Year FE	✓	✓	✓	
Year FE				✓
Observations	34,700	34,700	34,700	34,700

Models 1 - 3 are estimated with OLS and the SE are adjusted for both spatial and serial correlation (Hsiang 2010), Model 4 is estimated with MLE and SE are clustered at the cell level. • W denotes the binary, contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A9: Robustness: Different spatial and temporal specifications.

	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 4	Model 3	Model 4	Model 4
LOC	-0.0003	-0.0004*	-0.0002	-0.0001	-0.0002	-0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0004)	(0.0005)	(0.0005)
LOC_{t-1}	-0.0001	-0.0003	-0.0002	0.0001	0.0002	0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
LOC_{t-2}	0.0000	0.0000	0.0001	0.0003	0.0005**	
	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	
SPEI	-0.0271***	-0.0270***	-0.0214**	-0.0039	-0.0403	-0.0291
	(0.0089)	(0.0089)	(0.0081)	(0.0192)	(0.0272)	(0.0254)
SPEI_{t-1}	-0.0038	-0.0038	-0.0084	0.0002	-0.0026	-0.0003
	(0.0090)	(0.0090)	(0.0083)	(0.0047)	(0.0049)	(0.0046)
SPEI_{t-2}	0.0014	0.0014	0.0008	0.0007	0.0016	
	(0.0083)	(0.0083)	(0.0074)	(0.0044)	(0.0047)	
AR_LOC		0.0002	0.0000	-0.0004	-0.0005	-0.0005
		(0.0003)	(0.0003)	(0.0007)	(0.0007)	(0.0007)
AR_LOC_{t-1}		0.0004	0.0004	0.0002	0.0002	0.0002
		(0.0004)	(0.0004)	(0.0007)	(0.0007)	(0.0007)
AR_LOC_{t-2}		0.0001	-0.0004	-0.0004	-0.0005	
		(0.0004)	(0.0004)	(0.0003)	(0.0003)	
FATAL_{t-1}			0.2251***	0.2137***	0.2405***	0.2423***
			(0.0162)	(0.0118)	(0.0122)	(0.0118)
LOC_{t-3}					0.0006	
					(0.0005)	
AR_LOC_{t-3}					-0.0008	
					(0.0008)	
SPEI_{t-3}					0.0041	
					(0.0048)	
$W\times$LOC				0.0005	0.0007	0.0007
				(0.0008)	(0.0007)	(0.0007)
$W\times$AR_LOC				-0.0006	-0.0005	-0.0006
				(0.0017)	(0.0014)	(0.0014)
$W\times$SPEI				0.0066	0.0420	0.0317
				(0.0217)	(0.0285)	(0.0267)
$W\times$FATAL				0.5891***	0.3419***	0.3512***
				(0.0130)	(0.0102)	(0.0097)
Cell FE	✓	✓	✓	✓	✓	✓
Country×Year FE	✓	✓	✓			
Year FE				✓	✓	✓
Observations	34,700	34,700	34,700	34,700	34,700	34,700

(1) - (3) are estimated by OLS and the standard errors (SE) are adjusted for both spatial and serial correlation à la Hsiang (2010), where the spatial cut-off is set at 360 km (4) - (6) are estimated by MLE and SE are clustered at the cell level • W denotes the contiguity matrix of order 1 except in (4), which uses a matrix of order 2 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A10: Robustness: Different spatial lags, Model 4.

	(1)	**(2)**	**(3)**
	Model 4	Model 4	Model 4
LOC	-0.0002	-0.0002	-0.0002
	(0.0005)	(0.0005)	(0.0005)
LOC_{t-1}	0.0002	0.0002	0.0002
	(0.0003)	(0.0004)	(0.0004)
LOC_{t-2}	0.0005*	0.0005*	0.0002
	(0.0002)	(0.0002)	(0.0002)
SPEI	-0.0405	-0.0438	-0.0436
	(0.0272)	(0.0271)	(0.0270)
SPEI_{t-1}	-0.0019	0.0264	0.0276
	(0.0049)	(0.0273)	(0.0274)
SPEI_{t-2}	0.0020	0.0018	-0.0081
	(0.0046)	(0.0046)	(0.0253)
AR_LOC	-0.0005	-0.0005	-0.0005
	(0.0007)	(0.0007)	(0.0007)
AR_LOC_{t-1}	0.0002	0.0001	0.0002
	(0.0007)	(0.0007)	(0.0007)
AR_LOC_{t-2}	-0.0005	-0.0005	-0.0005
	(0.0003)	(0.0003)	(0.0003)
FATAL_{t-1}	0.2405***	0.2406***	0.2405***
	(0.0122)	(0.0122)	(0.0122)
$W\times$LOC	0.0007	0.0007	0.0007
	(0.0007)	(0.0008)	(0.0008)
$W\times$AR_LOC	-0.0005	-0.0004	-0.0004
	(0.0014)	(0.0014)	(0.0014)
$W\times$SPEI	0.0421	0.0455	0.0453
	(0.0285)	(0.0284)	(0.0283)
$W\times\text{LOC}_{t-1}$		-0.0000	-0.0000
		(0.0006)	(0.0006)
$W\times\text{AR_LOC}_{t-1}$		0.0005	0.0004
		(0.0015)	(0.0015)
$W\times\text{SPEI}_{t-1}$		-0.0301	-0.0313
		(0.0287)	(0.0288)
$W\times\text{LOC}_{t-2}$			0.0003
			(0.0003)
$W\times\text{AR_LOC}_{t-2}$			0.0006
			(0.0008)
$W\times\text{SPEI}_{t-2}$			0.0105
			(0.0269)
$W\times$FATAL	0.3422***	0.3420***	0.3420***
	(0.0102)	(0.0102)	(0.0102)
Cell FE	✓	✓	✓
Year FE	✓	✓	✓
Observations	34,700	34,700	34,700

(1) - (3) are estimated by MLE and SE are clustered at the cell level. • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A11: Robustness: Leads.

	(1)	(2)	(3)
	Model 2	Model 3	Model 4
LOC	-0.0005*	-0.0004*	-0.0005*
	(0.0002)	(0.0002)	(0.0002)
LOC_{t-1}	0.0001	-0.0002	0.0003*
	(0.0002)	(0.0002)	(0.0002)
LOC_{t-2}	-0.0003	-0.0004	0.0005**
	(0.0003)	(0.0003)	(0.0002)
AR_LOC	0.0005	0.0004	0.0003
	(0.0003)	(0.0003)	(0.0003)
AR_LOC_{t-1}	-0.0000	0.0003	0.0001
	(0.0004)	(0.0004)	(0.0004)
AR_LOC_{t-2}	0.0008	0.0004	-0.0005
	(0.0005)	(0.0004)	(0.0003)
AR_LOC_{t+1}	0.0004	0.0006	0.0008
	(0.0005)	(0.0004)	(0.0005)
AR_LOC_{t+2}	-0.0008	-0.0007	-0.0003
	(0.0005)	(0.0005)	(0.0006)
SPEI	-0.0206**	-0.0152*	-0.0303
	(0.0074)	(0.0073)	(0.0278)
SPEI_{t-1}	-0.0089	-0.0159*	-0.0052
	(0.0072)	(0.0074)	(0.0050)
SPEI_{t-2}	0.0027	0.0022	-0.0005
	(0.0075)	(0.0073)	(0.0048)
FATAL_{t-1}		0.1641***	0.2266***
		(0.0119)	(0.0123)
$W\times$LOC			0.0006
			(0.0005)
$W\times$AR_LOC			-0.0000
			(0.0009)
$W\times$SPEI			0.0317
			(0.0292)
$W\times$FATAL			0.3344***
			(0.0106)
Cell FE	✓	✓	✓
Country×Year FE	✓	✓	
Year FE			✓
Observations	34,700	34,700	34,700

(1) - (2) are estimated by OLS and the standard errors (SE) are adjusted for both spatial and serial correlation à la Hsiang (2010), (3) is estimated by MLE and SE are clustered at the cell level. • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A12: Robustness: Luminosity.

	(1)	(2)	(3)	(4)
	FATAL	FATAL	FATAL	FATAL
LOC	-0.0002*	-0.0003	-0.0003	-0.0004*
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
LOC_{t-1}	-0.0001	-0.0002	-0.0001	0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
LOC_{t-2}	0.0001	0.0002	0.0003	0.0005**
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
SPEI	-0.0126	-0.0125	-0.0099	-0.0308
	(0.0143)	(0.0144)	(0.0140)	(0.0289)
SPEI_{t-1}	-0.0272*	-0.0272*	-0.0258*	-0.0038
	(0.0142)	(0.0142)	(0.0141)	(0.0055)
SPEI_{t-2}	-0.0026	-0.0026	-0.0002	-0.0003
	(0.0142)	(0.0142)	(0.0138)	(0.0051)
log(NL+1)	-0.0036	-0.0036	-0.0038	-0.0037**
	(0.0056)	(0.0055)	(0.0054)	(0.0018)
$\log(\text{NL+1})_{t-1}$	0.0023	0.0023	0.0032	0.0018
	(0.0059)	(0.0059)	(0.0060)	(0.0017)
$\log(\text{NL+1})_{t-2}$	-0.0007	-0.0007	-0.0013	0.0035*
	(0.0056)	(0.0056)	(0.0054)	(0.0020)
AR_LOC		0.0003	0.0003	0.0001
		(0.0003)	(0.0003)	(0.0003)
AR_LOC_{t-1}		0.0004	0.0004	0.0000
		(0.0005)	(0.0005)	(0.0004)
AR_LOC_{t-2}		-0.0003	-0.0003	-0.0005
		(0.0004)	(0.0004)	(0.0003)
FATAL_{t-1}			0.1048***	0.2175***
			(0.0141)	(0.0129)
$W\times$LOC				0.0006
				(0.0005)
$W\times$AR_LOC				-0.0002
				(0.0009)
$W\times$SPEI				0.0313
				(0.0304)
$W\times$log(NL+1)				0.0107***
				(0.0029)
$W\times$FATAL				0.3222***
				(0.0106)
Observations	31,546	31,546	31,546	31,546

(1) - (3) are estimated by OLS and the standard errors (SE) are adjusted for both spatial and serial correlation à la Hsiang (2010), (4) is estimated by MLE and SE are clustered at the cell level. • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Table A13: Extension of Mechanism III (count variable as outcome).

	(1)	(2)	(3)	(4)
	FATAL_idmil.c	FATAL_polmil.c	FATAL_state.c	FATAL_rebel.c
LOC	-0.0315	-0.0589	-0.0394	-0.0141
	(0.0384)	(0.0811)	(0.0249)	(0.0094)
LOC_{t-1}	-0.0691	-0.1310	0.0208	-0.0034
	(0.0481)	(0.1135)	(0.0136)	(0.0113)
LOC_{t-2}	-0.0190	-0.0831	0.0018	0.0149^{*}
	(0.0304)	(0.0816)	(0.0068)	(0.0090)
AR_LOC	0.0046	0.0193	0.0175	0.0125
	(0.0348)	(0.0556)	(0.0175)	(0.0109)
AR_LOC_{t-1}	0.4014	0.7616	-0.0050	0.0943
	(0.2863)	(0.7029)	(0.0407)	(0.0549)
AR_LOC_{t-2}	0.1606	0.4945	0.0078	-0.0096
	(0.1704)	(0.4820)	(0.0162)	(0.0216)
SPEI	0.2747	4.3428^{**}	-0.2802	4.0460
	(1.6028)	(2.1966)	(2.8664)	(2.6924)
SPEI_{t-1}	0.1660	-0.0309	0.1761	-0.4195^{*}
	(0.2274)	(0.2208)	(0.3496)	(0.2315)
SPEI_{t-2}	0.3112	0.4354	-0.2749	0.0934
	(0.2913)	(0.2895)	(0.3810)	(0.2989)
Y_{t-1}	0.2254^{***}	0.1175^{**}	0.2263^{***}	0.1434^{***}
	(0.0515)	(0.0600)	(0.0460)	(0.0505)
$W\times$LOC	-0.0659^{*}	-0.1070	0.0446	-0.0131
	(0.0351)	(0.0772)	(0.0350)	(0.0169)
$W\times$AR_LOC	0.1531^{*}	0.2677	-0.0313	0.0131
	(0.0926)	(0.2258)	(0.0382)	(0.0247)
$W\times$SPEI	0.2524	-4.3239^{**}	1.1244	-3.1946
	(1.6677)	(2.1127)	(3.0474)	(2.7569)
$W\times$Y	0.4057^{***}	0.4068^{***}	0.2638^{***}	0.4233^{***}
	(0.0507)	(0.0397)	(0.0439)	(0.0679)
Cell FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	34,700	34,700	34,700	34,700

(1) - (4) are specified as Model 4, which is estimated by MLE and the standard errors are clustered at the cell level • W denotes the contiguity matrix of order 1 • * denotes significant at the 10% level, ** at the 5% level, *** at the 1% level

Appendix B

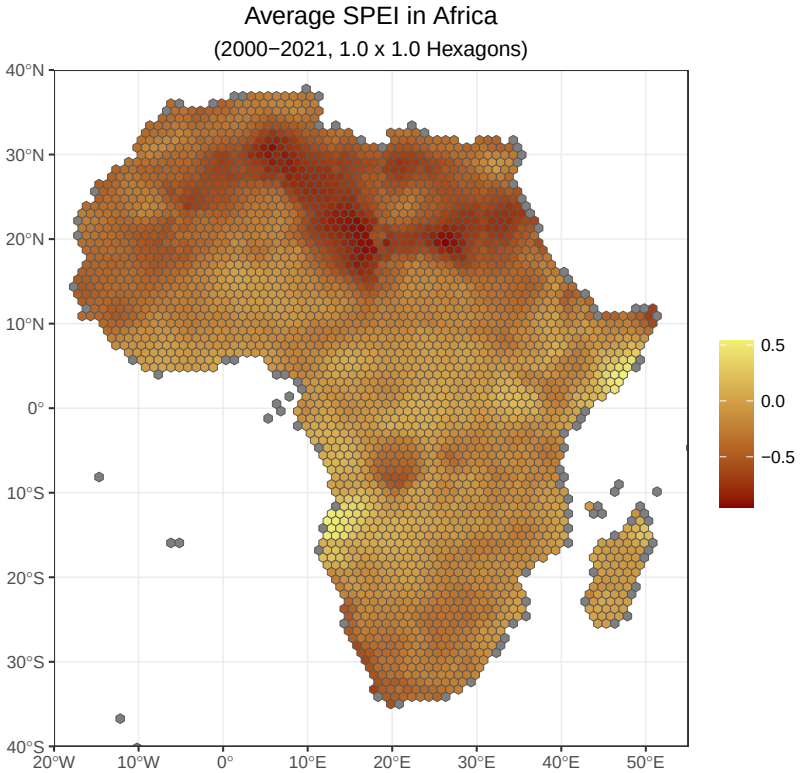

Figure B1: Spatial distribution of SPEI. 1.0 x 1.0 hexagons, aggregated data from 2000 - 2021. Intensity of red color indicates drier conditions, while grey cells indicate NAs. Data Source: CRU TS.

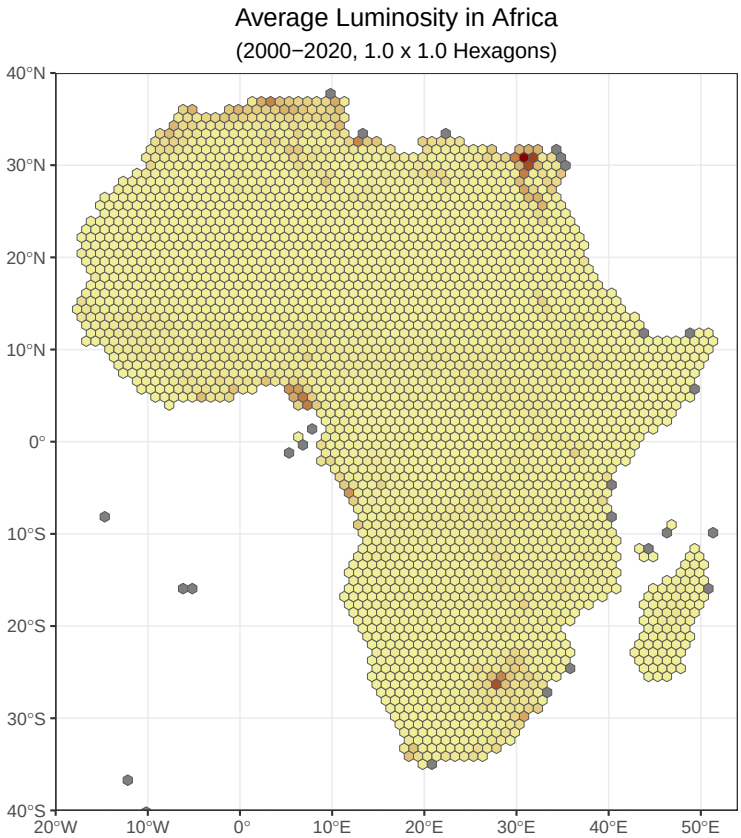

Figure B2: Spatial distribution of nighttime lights ("luminosity") data. 1.0 x 1.0 hexagons, mean over 2000 - 2020. Intensity of red color indicates more light, while grey cells indicate NAs. Data Source: AfroGrid.

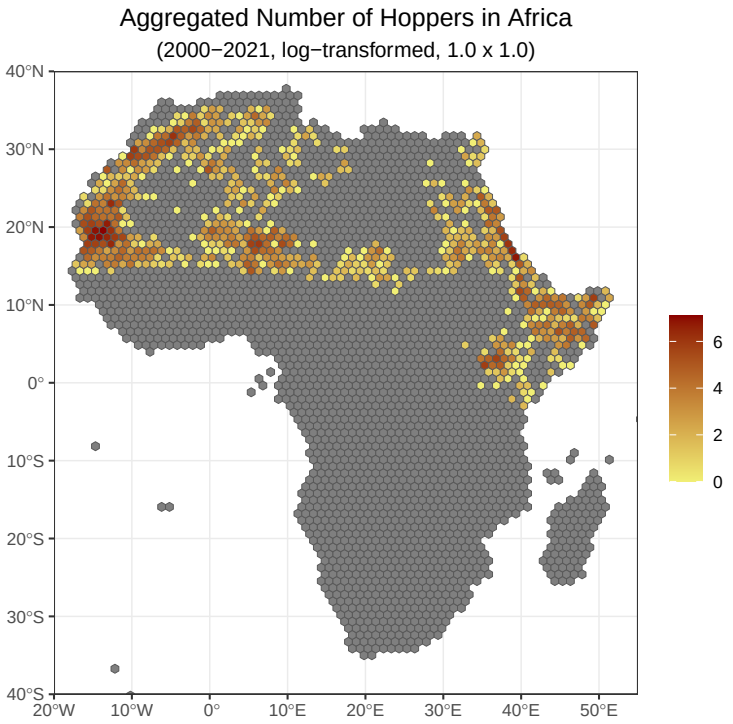

Figure B3: Spatial distribution of hoppers. 1.0 x 1.0 hexagons, aggregated data from 2000 - 2021. Intensity of red color indicates more hoppers, while grey cells indicate NAs. Data Source: FAO.

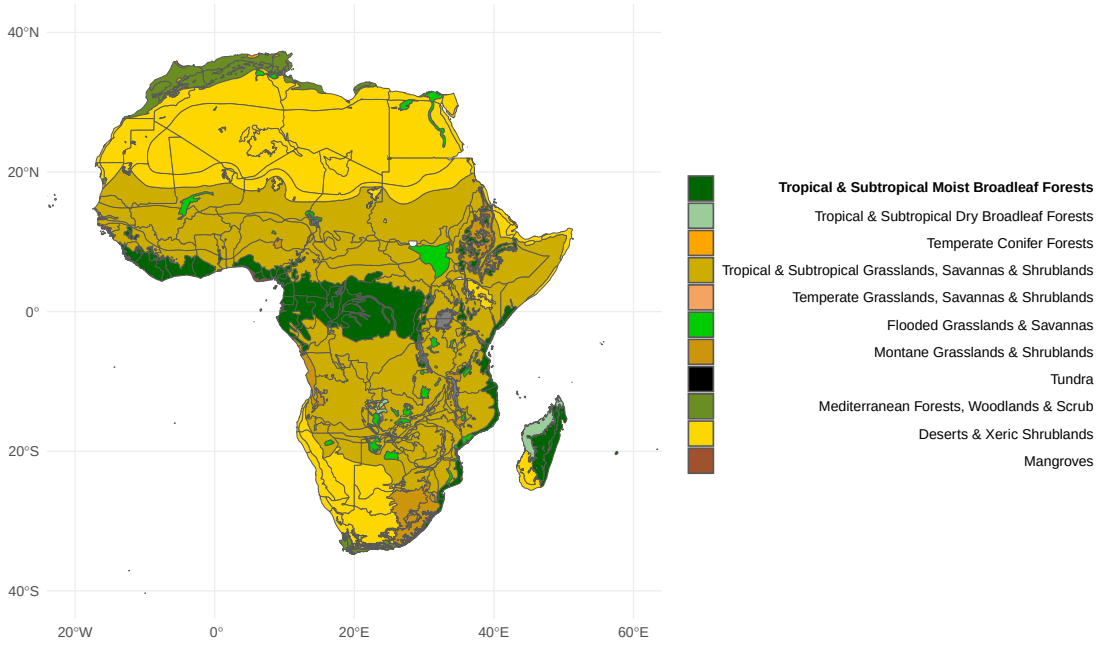

Figure B4: Global biomes as defined by the WWF in their 2004 Terrestrial Ecoregions of the World. Data Source: Olson et al. (2001).

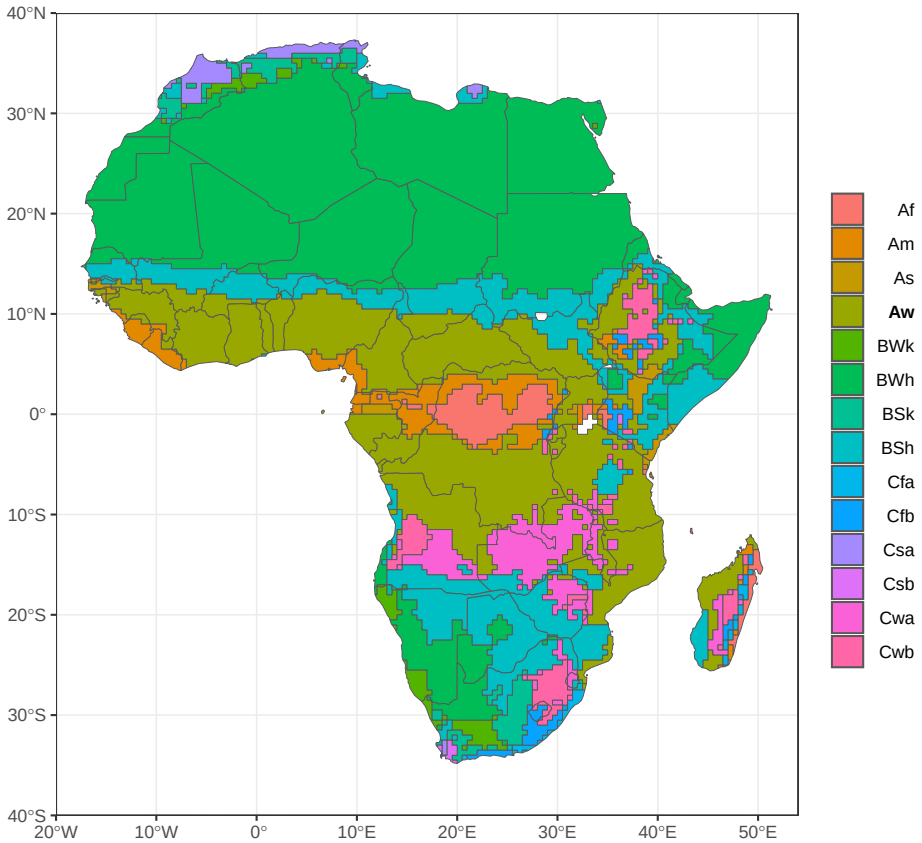

Figure B5: Köppen-Geiger climate classification of Africa using CRU TS 2.1 temperature and GPCC Full v4 precipitation data, period 1976-2000. Data Source: Kottek et al. (2006).

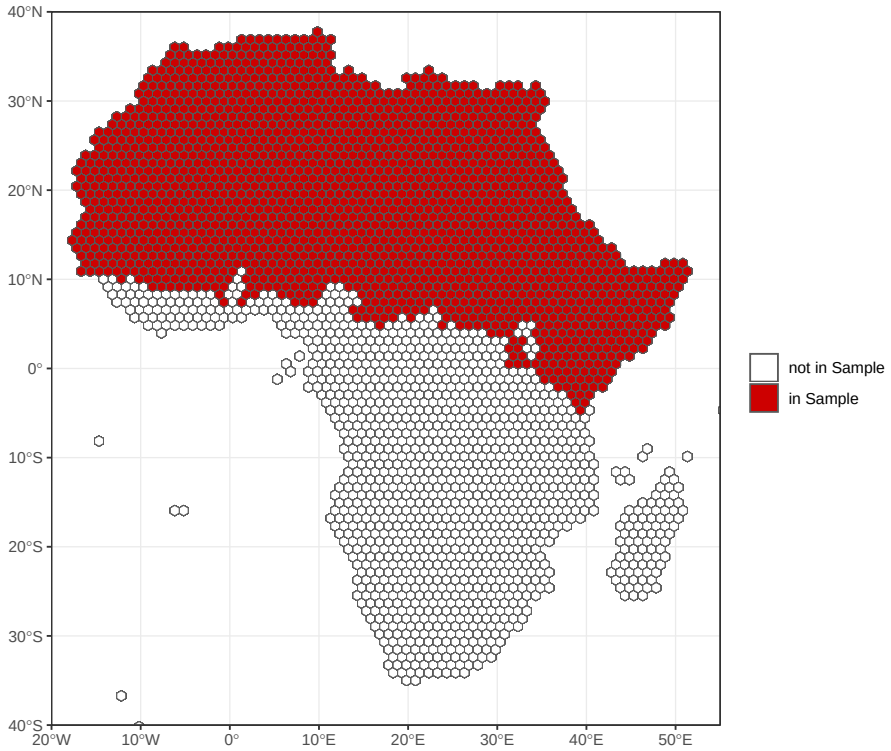

Figure B6: Forest frontier sample (based on the WWF forest frontier and Horn of Africa).

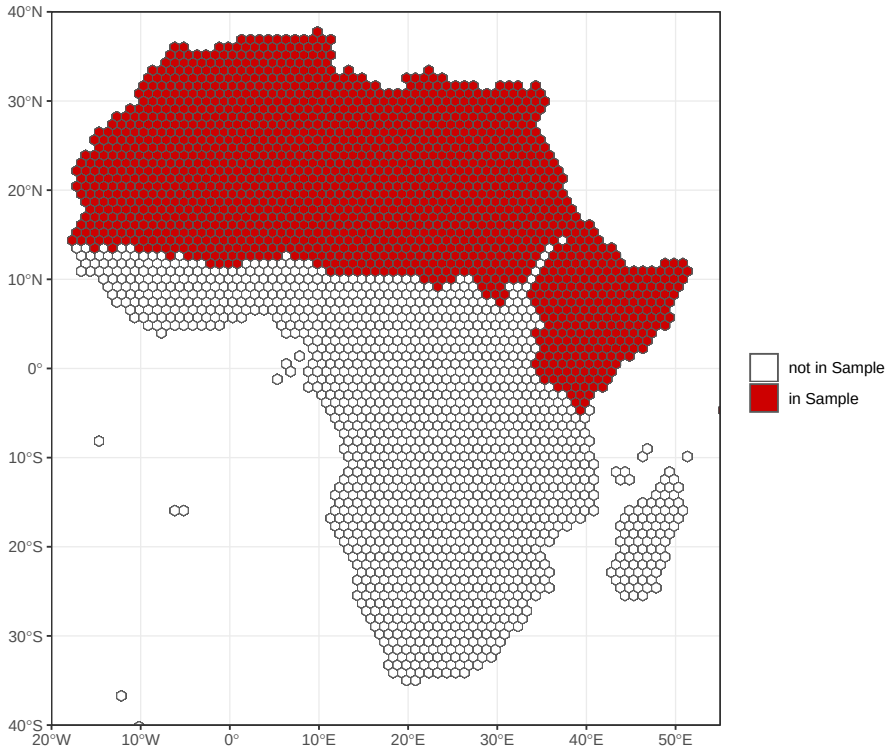

Figure B7: Climate zones sample (based on the Köppen-Geiger climate zones and Horn of Africa).

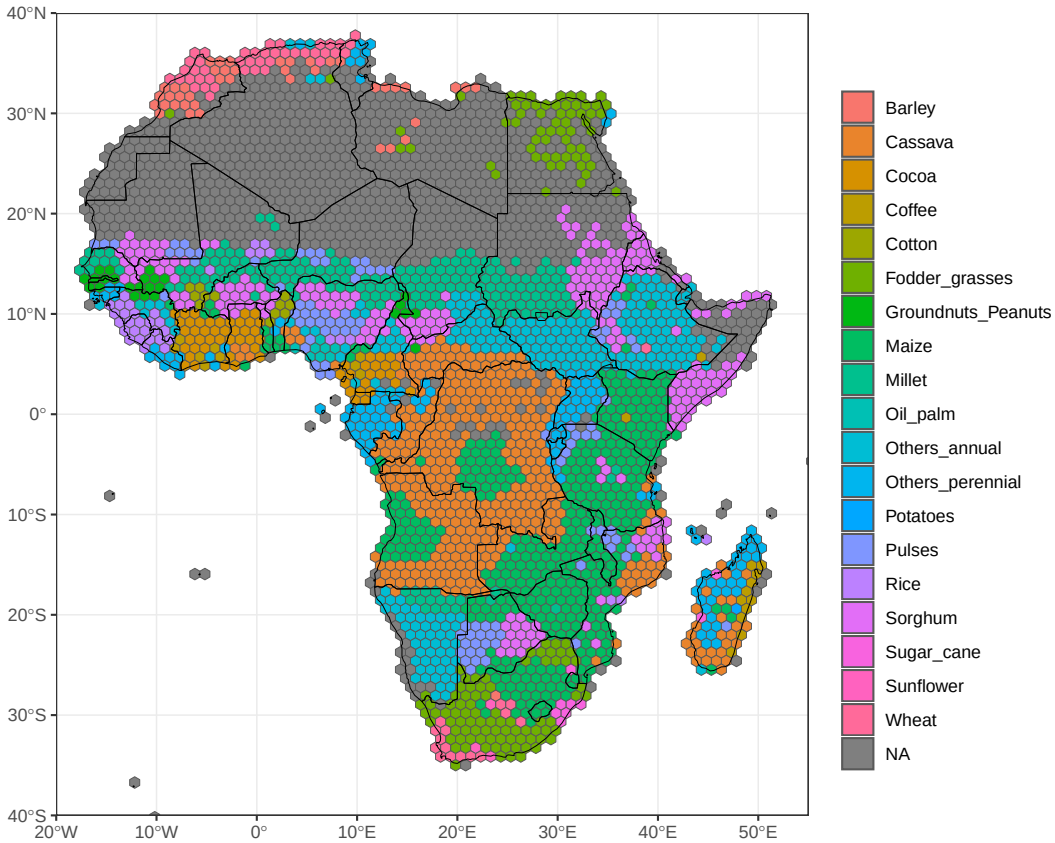

Figure B8: Main crop of cell (by area). Average from 1998 - 2002. Data source: MIRCA2000

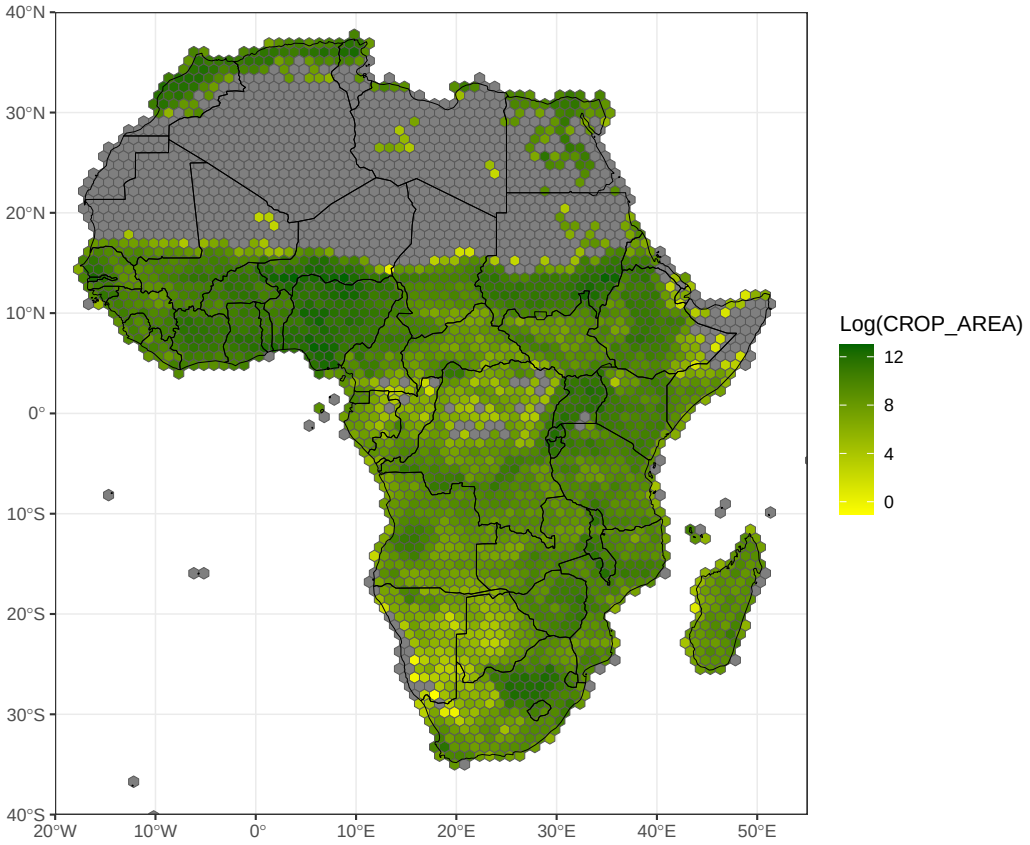

Figure B9: Main crop's acreage. Log-transformed. Average from 1998 - 2002. Data source: MIRCA2000

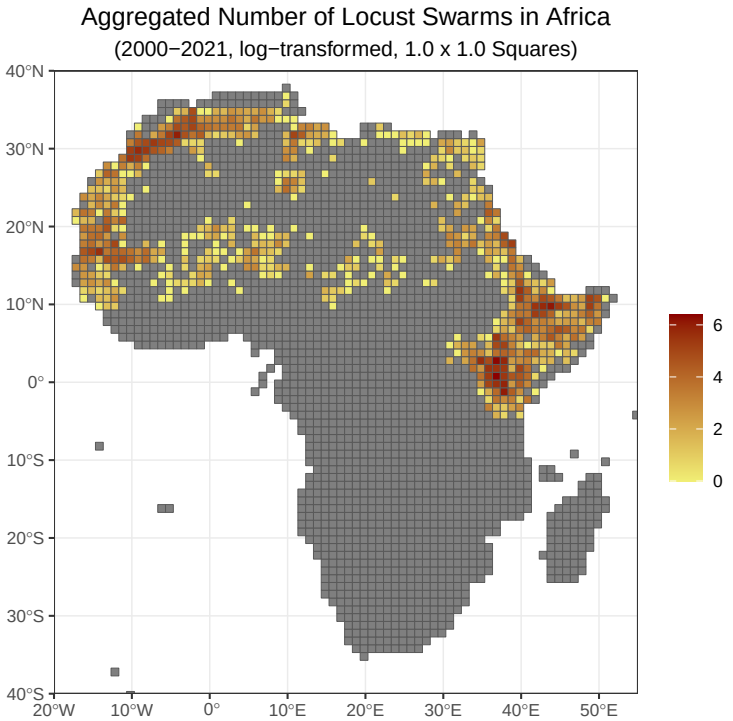

Figure B10: Spatial distribution of desert locust swarms. 1.0 x 1.0 squares, aggregated data from 2000 - 2021, log-transformed. Intensity of red color corresponds to higher occurrences of reported fatalities, while grey cells indicate NAs. Data Source: FAO.

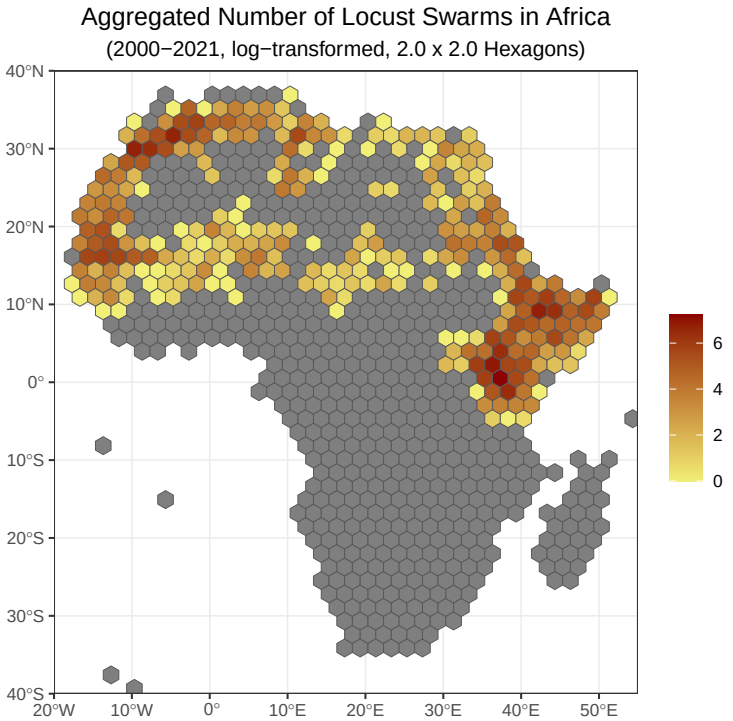

Figure B11: Spatial distribution of desert locust swarms. 2.0 x 2.0 hexagons, aggregated data from 2000 - 2021, log-transformed. Intensity of red color corresponds to higher occurrences of reported fatalities, while grey cells indicate NAs. Data Source: FAO.

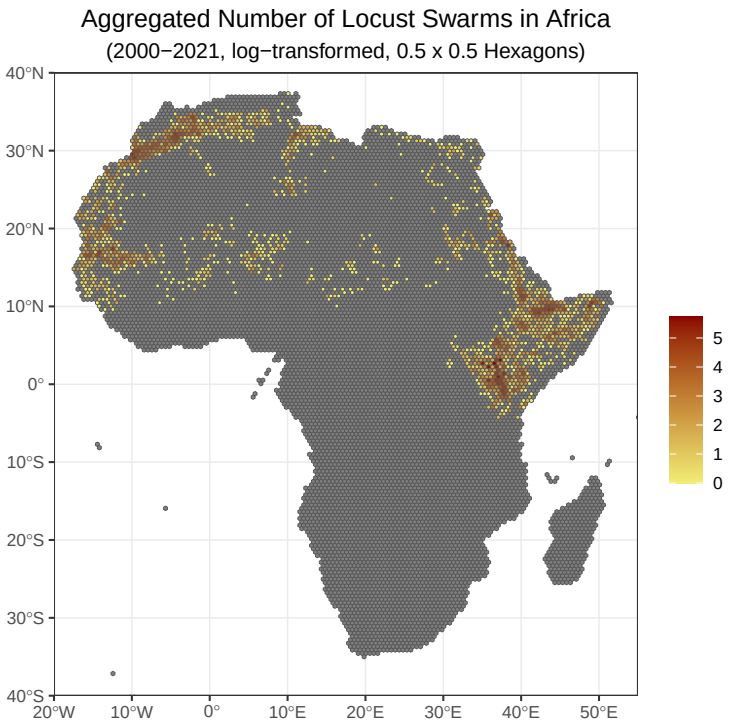

Figure B12: Spatial distribution of desert locust swarms. 0.5 x 0.5 hexagons, aggregated data from 2000 - 2021, log-transformed. Intensity of red color corresponds to higher occurrences of reported fatalities, while grey cells indicate NAs. Data Source: FAO.

Appendix C

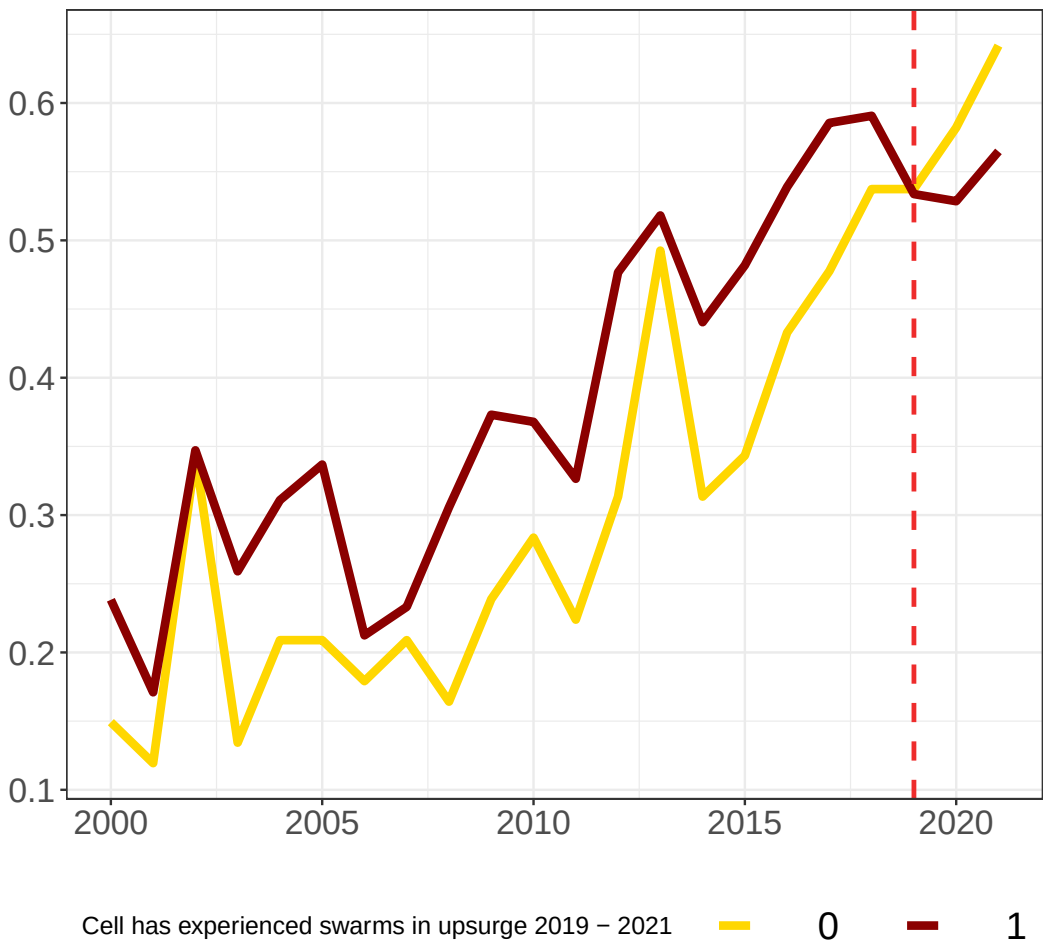

Figure C13: Parallel trends grouped for cells exposed and not exposed to swarms in the 2019-2021 upsurge. Average of "FATAL" on y-axis. Horn of Africa cells only. Data source: ACLED and FAO.

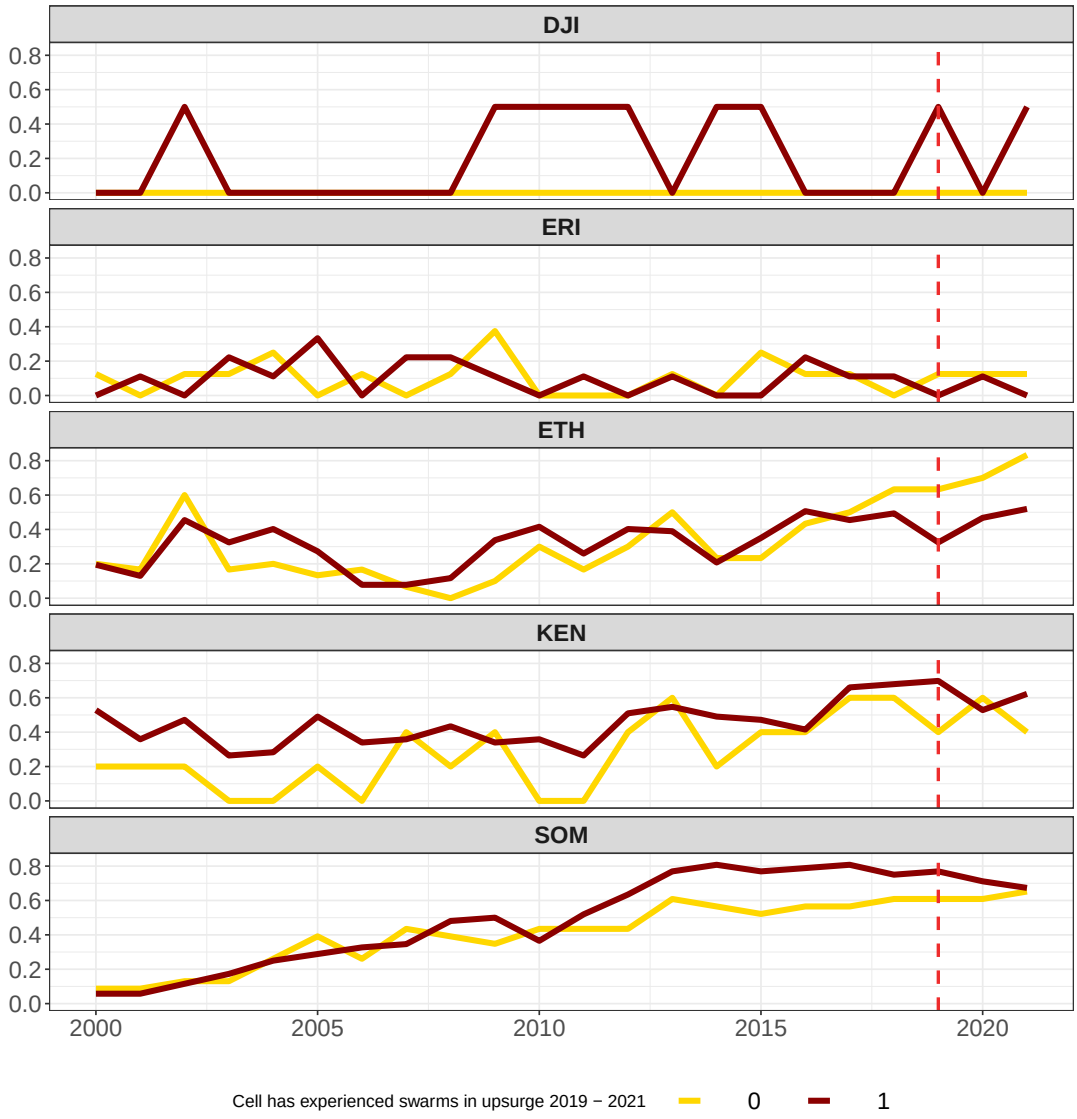

Figure C14: Parallel trends grouped by country for cells exposed and not exposed to swarms in the 2019-2021 upsurge. Average of "FATAL" on y-axis. Cells in Horn of Africa only. Data source: ACLED and FAO.

General	Event Type	Sub-Event Type
Violent events	Battles	*Armed clash*
		Government regains territory
		Non-state actor overtakes territory
	Explosions/Remote violence	*Chemical weapon*
		Air/drone strike
		Suicide bomb
		Shelling/artillery/missile attack
		Remote explosive/landmine/IED
		Grenade
	Violence against civilians	*Sexual violence*
		Attack
		Abduction/forced disappearance
Demonstrations	Protests	*Peaceful protest*
		Protest with intervention
		Excessive force against protesters
	Riots	*Violent demonstration*
		Mob violence
Non-violent actions	Strategic developments	*Agreement*
		Arrests
		Change to group/activity
		Disrupted weapons use
		Headquarters or base established
		Looting/property destruction
		Non-violent transfer of territory
		Other

Figure C15: ACLED event types. Source: Trivedi et al. (2021).

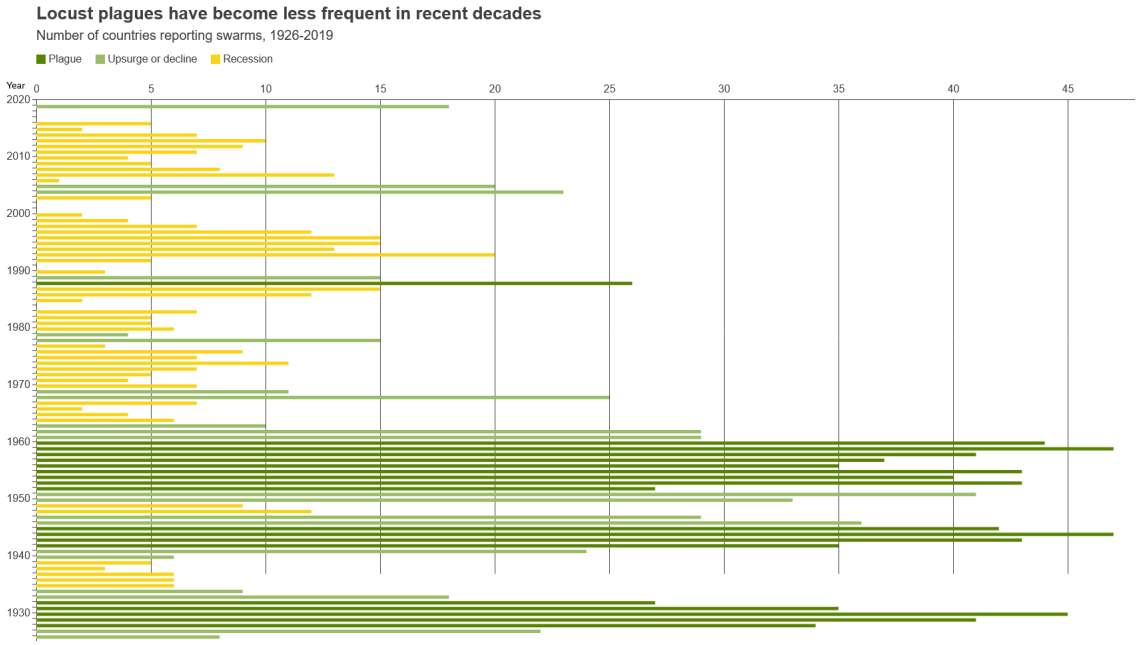

Figure C16: Locust events in the long-rung. Source: FAO (via BBC (2020)).