Climatic and Geographic Determinants of Economic Development

Miriam Breckner

Dissertation

Munich, 2018

Tag der mündlichen Prüfung: 23. Januar 2019

Namen der Berichterstatter: Prof. Dr. Uwe Sunde, Prof. Gabriel Felbermayr, Ph.D. and Prof. Dr. Florian Englmaier

Climatic and Geographic Determinants of Economic Development

Inaugural-Dissertation
Zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig–Maximilians–Universität München

2018

vorgelegt von
Miriam Breckner

Referent: Prof. Dr. Uwe Sunde
Korreferent: Prof. Gabriel Felbermayr, Ph.D.
Promotionsabschlussberatung: 30. Januar 2019

Acknowledgements

First and foremost I would like to express my gratitude to my supervisor Uwe Sunde. I am deeply thankful for his guidance and support. I enjoyed working on joint projects and I am grateful for the enthusiasm and valuable advice which shaped many ideas and the progress of this dissertation. I would also like to thank Gabriel Felbermayr for his willingness to act as my second supervisor and for valuable comments and suggestions to my research. Further I am grateful to Florian Englmaier as co-author and mentor during my PhD. Our joint project was an important starting point of this dissertation and his advice steadily motivated me and encouraged my progress. Furthermore, I am happy to have worked with great other co-authors. I would like to thank Till Stowasser who beyond the work on our joint project was always open to help and to share his experiences. Moreover I would like to thank Marie Lechler. The joint work and the years we spent sharing an office and many interests were a huge enrichment for me. I profited a lot from the frequent exchange of thoughts and perspectives.

I would further like to thank the DFG for funding through the Research Training Group GRK1928. I would like to thank Carsten Eckel for his support as the coordinator of the program and Julia Zimmermann as well as Silke Englmaier for help in all administrative concerns. Moreover, I am extremely thankful for the great research environment that the MGSE provides and for the readiness to help that I experienced from the faculty members. My work has further profited a lot from several courses and workshops and from numerous comments and discussions at conferences. Further, I would like to express gratitude to all my colleagues from the Chair for Population Economics for a fruitful exchange on research as well as great cooperation in teaching. I would also like to thank Susan Fay for great administration at the chair. Further, I would like to thank the staff of the Center for Economic Studies and my colleagues from the 1st floor for providing me an extremely nice and helpful work environment.

I would like to thank all other friends and colleagues that have brightened everyday life. I am grateful to Ilka Gerhardts for her support and flow of inspiration, to Susanne Wildgruber, especially for our brain stormings and the discovery of “the channel” and to Lukas Buchheim for sharing his knowledge and motivation. I would further like to thank

Sebastian Horn, Mark Westcott, Florian Loipersberger, Daniel Gietl, Nadine Geiger and Zhaoxin Pu as well as Felix Klimm, Stefan Grimm, Marco Schwarz, Markus Nagler, Else Christensen, Cathrin Mohr and Lisa Simon for good times during lunch and coffee breaks, sporting activities and the persistent mutual support.

I am also thankful to my friends outside LMU for always being there for me, and happily remember all visits and joint journeys. Finally, I am more than grateful to my sister and to my parents, for their support, inspiration and belief in me. I am thankful to my father for all the guidance and advice he gave with great foresight and humor throughout the years of this dissertation.

Contents

List of Figures

List of Tables

Preface

The impact of climate and geography on economic outcomes has long been a controversial topic in the scientific debate. Initially, the debate was concerned with determining the relative importance of geography as fundamental factor of development (e.g. Acemoglu et al., 2001, Diamond, 1999, Sachs, 2003, Gallup et al., 1999, Rodrik et al., 2004). Since then the literature about the role of geography and climate for economic outcomes has moved from the cross-sectional perspective to analyses of shorter-term weather variability within countries or smaller spatial units which allowed for refined identification strategies. With this new strand of literature, which Dell et al. (2014) refer to as “The New Climate–Economy Literature”, the perception of climate as merely fixed (pre-)condition to economic development has shifted to a perception of climate as a varying factor. This perception reflects two trends that are shaping the current debate. First, the climatic and environmental parameters are in fact changing, and this change is currently taking up pace. Global average surface temperature shows a warming of 0.85° Celsius over the period 1880 to 2012 causing more frequent weather extremes and significant environmental adjustments (IPCC, 2014). Second, rapidly advancing satellite and imaging technologies improve the measurement of environmental and climatic variables and thus make its changing nature more salient.

At the same time these improvements create new opportunities to study the impact of environmental and climatic changes on economic outcomes. In recent years the literature has considerably been fueled by advances in the use of spatially disaggregated climatic data at higher frequencies and complementing geo-coded datasets on outcome variables. Central insights of existing research relate to proving the impact of climatic and environmental factors on outcomes such as economic growth (Dell et al., 2012), agricultural output (Burke and Emerick, 2016) or conflict (Burke et al., 2015a) among other outcomes. Further, the effects are shown to be quite heterogeneous across regions. On the one hand, the effects depend on the prevalent climatic conditions. For instance Burke et al. (2015b) reveal a non-linearity regarding the impact of temperature on income. On the other hand, the socio-economic environment seems to matter as well. Dell et al. (2012) show that the impact of rising temperatures on economic growth differs strongly by countries’ income levels.

The literature has recently started to disentangle the underlying factors and mechanisms of these observations in more detail, with findings relating to more specific strands of literature for the respective outcomes.

This dissertation contributes to this research in different dimensions. In particular, the spectrum of analysis covers the impact of temperature extremes, longer term warming and land degradation on conflict in chapter 1 and chapter 2, the impact of natural disasters on income levels in chapter 3 and the effect of water pollution on health in chapter 4. This dissertation aims at a precise identification of these impacts and a better understanding of the interplay between the environmental factors with economic, institutional and demographic traits in shaping the socio-economic outcomes. To achieve these goals this dissertation uses a variety of econometric approaches and presents new applications for the use of high-resolution data.

The first two chapters contribute to the debate on whether climate change and global warming cause civil conflicts. The two chapters are also methodologically linked as they both build on high-frequency high-resolution data. In chapter 1 Uwe Sunde and I analyze the effect of monthly temperature shocks and longer term changes on the risk of civil conflicts in Africa. The analysis of monthly data for 4826 grid cells of 0.75^o latitude$\times$longitude over the period 1997-2015 documents a positive effect of the occurrence of temperature extremes on conflict incidence. These effects are larger the more severe the extremes in terms of duration, and are larger in densely populated regions, in regions with lower agricultural productivity, and in regions with more pronounced land degradation. The results also point towards heterogeneity of the effect with respect to the type of violence and towards the crucial role of population dynamics. Regions experiencing a decline or outflow of population exhibit different conflict types in response to temperature extremes than regions experiencing population increase or inflows. These findings hint to different mechanisms being in place which might reconcile some of the contradictory findings documented in the literature. Considering the role of changes in the frequency of extreme events in a long-differences analysis and applying a generalized differences-in-differences strategy delivers evidence for a positive effect of a gradual increase in the frequency of extreme events on conflict. This chapter illustrates that global warming and the related increase in extremes can impact conflict, which constitutes a major barrier of development. It seems that beyond a reaction to short-term shocks, societies did not adapt yet to changes. This chapter therewith contributes to the debate by providing novel evidence on the role of extreme temperature events for armed conflict based on an analysis for the entire continent of Africa.

Chapter 2 takes a closer look at land degradation which is oftentimes discussed as a consequence of climate change and focuses on the Sahel region. In the previous chapter land degradation is shown to be an interacting factor strengthening the adverse impact that temperature extremes have on conflict. The analysis of this chapter aims at identifying the main effects of land degradation on conflict. For this purpose it builds on a time-varying measure that is constructed from satellite data. This allows for an analysis of monthly shocks and longer term changes on a 0.75 degree grid in the period 2000-2012. Moreover it allows isolating the external climate driven components of land degradation by instrumenting vegetation barrenness with monthly precipitation levels. The residual of precipitation based predictions is commonly interpreted as "human component". Thus, to shed light on the human component of land degradation I analyze the effect of the residual on conflict separately. The results of the analysis unfold a positive effect of vegetation barrenness on conflict which is driven by rainfall shortages. Beyond, the effect shows to be most strongly pronounced where societies most heavily depend on agricultural productivity. Further the findings unfold an effect heterogeneity with respect to the involved types of conflict and the seasonality of their occurrence. Taken together, the findings point towards a mechanism that works through tightened resource competition in the face of lowered agricultural productivity. Turning towards the longer term implications, the results of the long-difference analysis suggest a positive effect of land degradation on conflicts related to agriculture and pastoralism. Whereas chapter 1 focuses on temperature extremes, the findings of this chapter are more closely related to the literature that analyses the effect of precipitation shortages and drought on conflict. In this vein the findings of this chapter confirm a link between weather variability and conflict. In particular, the analysis provides new evidence for the role of vegetation dynamics as intermediary between precipitation variability and conflict.

The third chapter which is based on joint work with Florian Englmaier, Till Stowasser and Uwe Sunde turns towards countries' exposure to natural disasters and the effect on countries' income levels. Natural disasters in this analysis include hurricanes, earthquakes, floods and alike. By the design of the empirical analysis the first two chapters do not deal with the role of institutional factors as crucial determinants of how severely societies are affected by weather extremes. In contrast, the setting of this chapter uses country-year data for the period 1980-2011 and takes a global perspective with a sample of 127 countries. This enables an analysis of the mitigating role of institutions in linking natural disasters and income. In particular this chapter aims to understand the role of private insurance markets and their interplay with public institutions in shaping countries' resilience to natural disasters. Employing comprehensive data on natural disasters and related losses as well as on global insurance penetration rates the findings indicate that private insurance markets accommodate the negative effects of natural catastrophes in developed countries whereas they turn out to be ineffective in developing countries. The results further reveal that this pattern masks the role of deficient institutional quality in preventing insurance to unfold full efficiency. This implies that insurance and a stable, well-institutionalized environment complement each other in mediating the negative disaster shocks. This chapter provides new insights on the negative effect of natural catastrophes on economic development. In particular, it highlights another dimension of interacting mechanisms linking climate to economic outcomes.

Chapter 4 which is based on joint work with Marie Lechler shifts the focus to another environmental issue that affects economic outcomes as well. In particular, it turns towards the impact of industrial and urban water pollution on health of children living along the Nile in Egypt. Methodologically this chapter returns to the use of disaggregated data. Specifically, we collect geo-coded data for industrial plants, population density and urban hotspots as well as individual health outcomes of children living in households along the Nile. We find that children living in households downstream to urban areas suffer from higher risks of disease than children living upstream. Information on the opening date of industrial plants allows us to perform a difference-in-difference analysis regarding the health effect of industrial pollution on health outcomes. We also find strong negative health effects of industrial plants on children living downstream while children living upstream remain unaffected. The results further reveal that the negative health effect on the downstream population can be mitigated by access to clean drinking water. The results indicate detrimental impacts of water pollution and advise the sensible design and enforcement of environmental regulations as well as an improved provision of the access to adequate water resources to households. This chapter provides further evidence regarding the interplay between human actions and environmental conditions in shaping development outcomes.

The four essays in this thesis are self contained. Each chapter is followed by an appendix whereas a consolidated bibliography is contained at the end of the thesis.

Chapter 1

Temperature Extremes, Global Warming and Armed Conflict: New Evidence from High Resolution Data

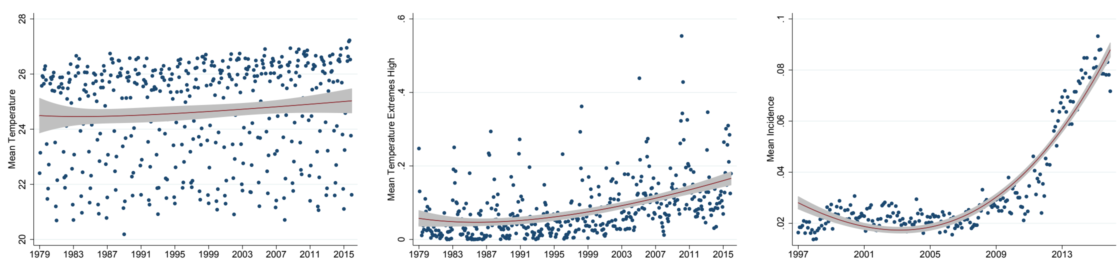
Amid the diverse social and political causes, the Darfur conflict began as an ecological crisis, arising at least in part from climate change. (...) It is no accident that the violence in Darfur erupted during the drought. (...) For the first time in memory, there was no longer enough food and water for all. Fighting broke out. (...) Any peace in Darfur must be built on solutions that go to the root causes of the conflict.
(Ban Ki Moon, 2007)

Most of today's conflicts are still essentially internal. (...) They are fuelled by competition for power and resources, inequality, marginalization and exclusion, poor governance, weak institutions, sectarian divides. They are exacerbated by climate change, population growth and the globalization of crime and terrorism.
(Antonio Guterres, 2017)

This chapter is based on joint work with Uwe Sunde.

1.1 Introduction

Climate change and civil conflict belong to the greatest challenges for developing countries, especially in Africa. Official statements, such as those by UN Secretary Generals Ban Ki Moon and Antonio Guterres, are explicit about the detrimental roles played by these phenomena, as well as their interlinkages. A look at the data confirms the increasing prevalence of both. For instance, Figure 1.1 illustrates the trend in temperature, exemplified by the average temperature per month (Panel (a)) and the prevalence of extreme temperature events in a month, measured by the incidence of deviations from calendar month specific means for a given grid cell exceeding the 95th percentile of deviations (Panel (b)). The data reveal a clear increase in both, temperature and in the frequency of extreme events over the past three decades. Over the the past two decades, conflict incidence has increased starkly (Panel (c)). Figure 1.2 depicts the dynamics in temperature extremes and conflict over the period 1997-2015 for a map of Africa in grid cells of 0.75^o latitude/longitude. Conflict seems to have increased relatively more in regions that experienced a more pronounced increase in the prevalence of temperature extremes, which is suggestive of an interrelation.[1] This paper explores the question whether weather extremes and climate change, reflected by the frequency of events of extreme temperature which is commonly viewed as a symptom of global warming, is relevant for the outbreak of violent conflict at the disaggregate level.



(a) Average Temperature (b) Temperature Extremes (c) Conflict Events

Figure 1.1: Dynamics of Temperature, Extreme Temperature and Conflict in Africa
This figure shows the monthly evolution of temperature, extremes and conflict. Additionally, it plots fractional-polynomial predictions with 95% confidence intervals. Temperature extremes are coded as 1 if the absolute temperature deviation from month-specific means for a given cell (calculated from the training period 1979-1996) exceeds the 95 percentile threshold, and 0 otherwise. See Section 3.2 for a detailed data description.

[1] Figure 1A.1 in the Appendix depicts the corresponding map for dynamics in average temperature and conflict over the period 1997-2015. See Section 3.2 and Table 1A.1 for details on data, data sources and variable construction.

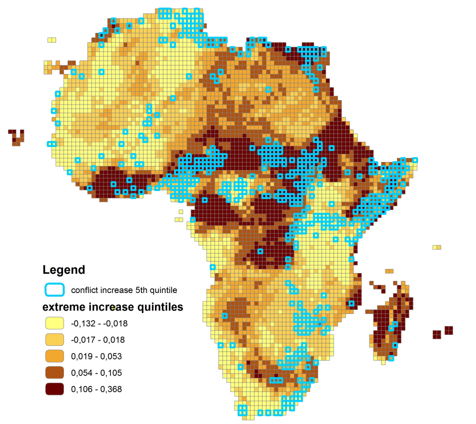


Figure 1.2: Dynamics of Temperature Extremes and Conflict Incidence
This figure plots quintiles of changes in the average occurrence of temperature extremes from the first half of the panel (1997-2006) to the second half of the panel (2006-2015) together with the cells belonging to the highest quintile of changes in average conflict incidence from the first half to the second half of the panel.

The question whether climate change and global warming cause conflicts has fueled a heated debate that has been ongoing for years. One branch of the existing academic literature has pointed at evidence that supposedly shows that weather and climate variation cause conflicts, mainly as consequence of increased resource pressure due to a deterioration in the conditions for agricultural production. Another branch of the literature has disavowed this conclusion by arguing that increased conflict was mainly a problem of institutional failure, whose consequences become aggravated in the face of increased environmental pressure as well as rapid globalization. Among the reasons for the lack of consensus in the scientific literature have been methodological issues, such as sampling bias, and the heterogeneity in the narratives about the triggers of incidences of conflict and the lack of evidence regarding the channels, which is related to the fact that most of the evidence at the core of this debate is at annual frequencies and at the country or region level. In light of this, official statements, such as those by Ban Ki Moon or Antonio Guterres, both UN Secretary General officials, usually avoid specific statements about causality.

This paper provides novel evidence about the role of extreme weather events related to climate change for conflict and contributes to the debate in several ways. In terms of data, the use of high-frequency high-resolution data for the entire continent of Africa allows isolating the role of weather extremes as triggers of conflict with greater precision than the previous literature and without restricting to particular areas that are more or less conflict prone, thus avoiding sampling bias. In particular, the analysis is based on monthly temperature and precipitation information for 4826 grid cells of 0.75^o latitude$\times$longitude over the period 1997-2015 and isolates the role of weather extremes for the incidence of violent conflict events. In contrast to previous studies on extreme weather events, we consider deviations from long-run cell-specific and calender month-specific average conditions.

We find a significantly positive effect of temperature extremes on conflict incidence for both, an increase in the average number of extreme weather events and a longer duration in terms of months. The effect on conflict of extreme events that last for two months is found to be quantitatively larger than the effect of events that last for one month only. There is also some indicative evidence that this effect is non-monotonic for events that last even longer. The quality of the data also allows for a detailed investigation of the underlying mechanisms. Among the narratives that have been mentioned in the literature, climate-related shocks to agricultural productivity and conflict for resources have been the most prominent. The analysis documents that temperature extremes have a particularly strong effect in densely populated areas and areas with low agricultural productivity.

Recent work has pointed at the role of migration in the context of climate change and conflict. When investigating the role of migration and population dynamics, our results reveal that temperature extremes have an effect on conflict mainly in areas that are loosing population or that are growing rapidly, not so much in areas with fairly stable populations. However, the types of observed conflicts turn out to differ systematically across areas that lose population and those that gain population, providing novel insights to the mechanisms linking weather extremes and conflict. Weather extremes appear to trigger mainly battles involving territorial changes and conflicts in rural and agricultural contexts in cells that experience out-migration and population loss. The results further show that land degradation is a critical factor for the link between temperature extremes and conflict with the effect of temperature extremes being much stronger in regions with greater degradation. These findings deliver novel evidence that is consistent with a mechanism working through the loss of agricultural productivity, in line with some of the conflict narratives that have been debated in the literature. At the same time, we find that temperature extremes are linked to riots and battles without territorial changes in areas that experience immigration and fast (presumably migration-related) population growth.

The analysis concludes by turning to the implications of global climate change, which is typically associated with a greater frequency and longer duration of extreme weather events. Considering the role of changes in the frequency of extreme events in a long-differences analysis and applying a generalized difference-in-differences strategy delivers evidence for a positive effect of a gradual increase in the frequency of extreme events on conflict. This sheds new light on the debate about the consequences of climate change for conflict. Our findings indicate that societies do not seem to have adjusted to an increasing frequency of extremes over the observed time period. Further, the societal vulnerability to short-run climatic shocks is mirrored in the long-run results.

Contribution to the Literature. This paper contributes to the debate about the role of weather variability and climate change for conflict, which has mainly focused on Africa. Proponents of the role of climate for conflict reported evidence pointing at a strong link between high temperatures or drought in a year and the occurrence of civil conflict (Burke, Miguel, Satyanath, Dykema, and Lobell, 2009; Burke, Dykema, Lobell, Miguel, and Satyanath, 2010; Burke, Miguel, Satyanath, Dykema, and Lobell, 2010), while critics pointed to methodological problems and structural factors being responsible for these results (Buhaug, 2010b; Buhaug, Hegre, and Strand, 2010; Buhaug, 2010a). This literature concentrated on large-scale conflicts and used annual data at the country level. More recent work used refined identification strategies and also considered small-scale conflicts, but still delivered no consensus about the role of climate, see Burke, Hsiang, and Miguel (2015a) for a recent survey.[2] Our work complements this literature by considering the role of weather extremes at a much higher frequency and with grid-level data of higher resolution, as well as by providing evidence for effect heterogeneity that allows insights regarding the underlying mechanisms.

In this dimension, our work is closely related to studies that use grid cells to analyze the effect of weather shocks like drought or temperature and precipitation extremes on conflict. There are several studies that use latitude-longitude grids (or other fine spatial units) as unit of observation and rely on annual variation for identification. Recent work by Harari and Ferrara (2018) studies the effects of drought during particularly critical phases of the crop cycle and finds evidence for a positive effect on annual conflict incidence at the 1° grid level. Using spatially defined ethnic homelands as unit of observation, recent work by von Uexkull, Croicu, Fjelde, and Buhaug (2016) analyzes the impact of growing season drought in Africa and Asia and documents a positive impact for agriculturally dependent and politically excluded groups on violence. This finding relates to Fjelde and Uexkull (2012) who document an effect of rainfall anomalies on communal conflict that is amplified in the presence of economic and political marginalization. Theisen, Holtermann,

[2]This ongoing debate includes work that finds evidence for a role of climate for conflict (Hsiang, Meng, and Cane, 2011; Hsiang, Burke, and Miguel, 2013; Hsiang and Meng, 2014), and work that questions the empirical validity of such a role (Theisen, 2012; Theisen, Holtermann, and Buhaug, 2012; Gleditsch, 2012; Hegre, Buhaug, Calvin, Nordkvelle, Waldhoff, and Gilmore, 2016).

and Buhaug (2012) study the impact of drought on conflict for the African continent at the 0.5° grid level, but attribute the cause of conflict to sociopolitical factors rather than weather shocks. Further, Theisen (2012) conducts an analysis at the 0.25° grid level for Kenya and finds the cause of conflict to be political motives rather than agricultural scarcities.[3]

The present paper is most closely related to studies that employ grid level data and rely on high frequency data for identification. These include work by O'Loughlin et al. (2012) who analyze the impact of temperature- and precipitation anomalies on a 1° degree grid in East Africa. In a follow-up study, O'Loughlin, Linke, and Witmer (2014) expand their sample to sub-Saharan Africa. This study documents a significantly positive impact of high temperature anomalies on conflicts that differs across sub-regions and types of conflict, whereas they do not find a significant effect of precipitation anomalies. Further, they assess the quantitative importance of this effect to the impact of political, economic and geographic factors. Maystadt, Calderone, and You (2015) use quarterly data at the 0.5° grid level for the case of Sudan and Maystadt and Ecker (2014) analyze the monthly impact on Somalian administrative regions.[4] Almer, Laurent-Lucchetti, and Oechslin (2017) investigate the role of monthly variation in water scarcity, using an index of evaporation and drought, for local riots. Our paper complements these works by providing an analysis that is based on detailed weather information on the grid-cell level for the entire continent of Africa at a monthly frequency, thereby mitigating concerns about selective sampling (Adams, Ide, Barnett, and Detges, 2018). Moreover, our analysis digs deeper into heterogeneity of the effects, in particular the types of conflict affected by temperature extremes, the role of population dynamics, the relevance of agricultural productivity and land degradation, and long-run patterns relating the analysis to the debate on global warming. Our work also complements recent work by Cervellati, Esposito, Sunde, and Valmori (2017) who use monthly data for a 1° grid for the entire continent of Africa to explore the role of weather fluctuations that affect the exposure of the local population to malaria as potential channel leading to increased violence. Instead of disease, our evidence focuses on mechanisms related to agricultural productivity and population dynamics.

Moreover, our analysis relates to the literature on the long-run variability in the context of climate change. In this dimension, it follows the suggestion by Burke, Hsiang, and Miguel (2015a) pointing to the need for more comprehensive evidence for the climate-conflict link from a long-run perspective. We add to that in terms of an expansion of the sample from East Africa to the entire African continent over a more recent time period, and in

[3]Hsiang, Burke, and Miguel (2013) dispute the finding that temperature has no significant effect on conflict by criticizing the methodology used by Theisen (2012).

[4] Crost et al. (2015) provide related evidence for the Philippines.

terms of methodology. Specifically, we extend the focus from considering trends of average temperature to trends of temperature extremes, documenting the driving force of conflict risk to be latter. We substantiate the long-difference results by applying a generalized difference-in-differences design which relaxes the restrictive assumption about common pre-trends. Further, our results add new insights regarding societal vulnerability towards increases in temperature (extremes) in the long run.

In line with conjectures formulated in the existing literature (see, e.g., the survey by Exenberger and Pondorfer, 2013), our findings also indicate the prevalence of coexisting mechanisms taking effect in different contexts. We find conflict risk to increase in regions with declining population and high levels of land degradation, involving conflicts related to territory, agriculture and pastoralism. Therefore, in terms of mechanisms, our study complements work that argues via the loss of agricultural productivity as potential channel, including work by Harari and Ferrara (2018) and recent evidence for ethnicity-related conflict (Sarsons, 2015; von Uexkull, Croicu, Fjelde, and Buhaug, 2016). Also, our work complements evidence by Raleigh and Urdal (2007) or Hendrix and Glaser (2007) who point out environmental degradation as a critical factor for the link between climate and conflict, although this literature relies on cross-sectional analysis or time series at an annual frequency. Using cross-sectional data only, Raleigh and Urdal (2007) emphasize the importance of analyzing the relationship between demography and environmental variables in shaping the risk of civil conflict by considering an interaction between population and land degradation. The present study broadens this focus by analyzing the interplay between climate and population dynamics, and between climate and land degradation, in shaping the risk of differential types of conflict within an extensive panel analysis. Our findings indeed suggest that climatic extremes fuel conflict risk in the face of adverse environmental conditions. Further, our finding of weather extremes leading to different types of conflict in regions with different population dynamics in terms of population growth and in-migration hints to different mechanisms being responsible for this result. The results thereby reconcile some of the contradictory findings documented in the literature, e.g., by Theisen (2012). In highlighting the role of population dynamics and migration for different conflict patterns, our work also complements recent work by Bosetti, Cattaneo, and Peri (2018) and Owain and Maslin (2018) that focuses on conflict-related migration between countries. In this respect, our work also contributes to the literature on climate migration and the links between environment, migration and conflict surveyed by Brzoska and Froehlich (2016).

The remainder of the paper is structured as follows. Section 3.2 describes the data sources and the construction of the data set used in the analysis. Section 1.3 presents the results for the short-run effects of the occurrence of temperature extremes on conflict incidence. Section 1.4 turns to a long-differences analysis of the effect of a gradual increase in the frequency of extreme events on conflicts. Section 3.4 concludes the analysis with a brief discussion of the results.

1.2 Data

To analyze the impact of temperature shocks and long-run warming on conflict risk we construct a monthly data set for Africa for the period 1997 to 2015 for 4826 grid cells of 0.75° latitude and longitude.

Monthly time series for temperature and precipitation are obtained from the Era Interim reanalysis data set provided by the European Centre for Medium Run Weather Forecast (ECMWF).[5] Reanalysis of meteorological data ensures very high data quality by combining the strengths of all available meteorological sources. Data inputs ranging from modern radiometric measurements by satellites to local weather stations, buoys or aircrafts are comprised by using a stable assimilation scheme. This guarantees temporally and spatially consistent estimates of the weather state and alleviates the concern that the extent of measurement error resulting, for instance, from unevenly distributed weather stations is correlated with omitted factors.

Our main explanatory variable is a binary measure for temperature extremes that is constructed from the monthly time series of temperature in each available grid cell. As the climate data is available from 1979 whereas the geo-coded conflict data is only available from 1997, we use the period 1979-1996 as training period for the construction of temperature extremes. Accordingly, the period 1997-2015 serves as estimation period. For the training period we calculate, for each grid cell, calendar month-specific means and define a grid-cell and calender-month specific threshold at the 95 percent percentile of absolute deviations from these means. Based on this threshold, we create a binary variable for the estimation period that takes on the value 1 if a deviation exceeds this threshold. The advantage of this approach is that the mean that serves as basis to construct the extremes stays constant over the estimation period. This avoids that the temperature extremes are by construction related to variation during the period of interest, allowing for a transparent benchmark and an investigation of long-run trends. Further, by looking at deviations from calendar month-specific means we explicitly abstract from seasonal climate effects. Otherwise, most of the extremes would be found in the hot season whereas there would not be much variation in the remaining year. This methodology is based on the

[5] See Dee et al. (2011) for details. We use data on synoptic monthly means of precipitation and average temperature at time 0:00 and time 12:00 (step 12). To obtain total precipitation for one month we sum up the values of both times and multiply that sum by the number of days in the respective month.

assumption that a deviation from the long-run normal conditions can have a disruptive impact in any month. It is therefore the variation relative to the normal conditions, and not an absolute weather event, that is modeled as extreme event and hypothesized as potential trigger of violence.

This approach constitutes a key difference to previously used measures of extreme events as most studies based their definition of temperature (precipitation) anomalies on deviations from long-run means independently of the (calendar) month. For instance, O'Loughlin, Witmer, Linke, Laing, Gettelman, and Dudhia (2012) and O'Loughlin, Linke, and Witmer (2014) define extreme events as +/- 1 or 2 SDs of the long-term means. Also see Maystadt, Calderone, and You (2015), Fjelde and Uexkull (2012), Raleigh and Kniveton (2012) for similar approaches. More complex drought indices, like the PDSI or SPEI, that have been used in the literature also reflect absolute variation and are therefore not suitable for the purposes of the present application, which tries to explicitly analyze the impact of global warming and related temperature extremes.

By construction, the frequency of extremes in the training period is 5 percent. In the estimation period this frequency is higher (11 percent), which already indicates the warming observable over the sample period that comes along with a rise in average temperature and a corresponding shift in the distribution. Further, it is notable that an increase of the intra-annual variance elevates the tails of the distribution and thereby the evolution of extremes (see Figure 1A.3). Precipitation declines on average over the sample period and therefore low precipitation extremes and drought increase in frequency while the intra-annual standard deviation declines.

The dependent variable is a binary conflict indicator that switches on if at least one conflict has taken place in a given grid and month. Geo-coded data on civil conflict is obtained from the Armed Conflict Location and Event Database (ACLED). The ACLED data set provides locations of conflicts within all African countries since 1997 (Raleigh, Linke, Hegre, and Karlsen, 2010). Events involve a range of actors, including rebels, governments, militias, armed groups, protesters and civilians. The two main categories contained in the database are battles and riots. Riots are usually (non-violent or violent) demonstrations against the government. In some cases the target might also be private entities like businesses. Battles are defined as violent events between two groups. One of the groups might be the government but it might also be that two non governmental groups fight against each other. Battles are further split into battles that result in changes of the contested territory and battles that do not affect territorial changes. The categorization of conflicts related to agriculture or pastoralism is based on key word search in informational notes that are included in the database for each incidence.[6] Besides, we generate a category

[6]Agropastoral conflicts are based on the following keywords in the contextual notes: "farm", "crop",

for rural conflicts based on the geographic location of their emergence. Conflicts are defined as rural regional conflicts if they take place outside of large agglomerations.[7]

To proxy agricultural productivity we employ the "Caloric Suitability Index" developed by Galor and Özak (2016). In contrast to previously used, weight-based measures of agricultural suitability, this index takes the caloric return of agricultural yields into account. This permits a straightforward comparison of output across regions. The Caloric Suitability Index is based on data for potential crop yields from the Global Agro-Ecological Zones project (GAEZ) by the Food and Agricultural Organization (FAO) which are transformed into caloric output using information on the caloric content of the respective crops. Estimates of potential yields are based on agro-climatic factors which ensures exogeneity with respect to to human intervention. Further, the caloric suitability measure constitutes a long-term, time invariant estimate and therefore remains unaffected by the evolution of climate.

Information on the exposure to land degradation comes from the World Atlas of Desertification (UNEP, 1992). The data on soil degradation contained in the Atlas is adopted from the Global Assessment of Human-induced Soil Degradation (GLASOD) project, funded by the United Nations Environment Program (UNEP) and coordinated at the International Soil Reference and Information Centre (ISRIC). Soil degradation is assessed and averaged over the recent past (5 to 10 years) at the time of its compilation in 1990. The measure of soil degradation is based on ratings by a large number of soil scientists that are specialists for their respective geographical regions. This expert rating classifies the extent of soil degradation into 5 categories (0-4). For instance, category 0 implies that there is no sign of present degradation whereas category 5 implies extreme degradation with the terrain being irreclaimable.

Demographic data on population density, population growth and net migration are obtained from the Gridded Population of the World Database (CIESIN, 2016). This data is provided in 5-year intervals (1995, 2000, 2005, 2010 and 2015). To be able to use these data in a time-varying specification, we linearly interpolate the population density data between these points in time on an annual basis. Population growth rates are calculated as log difference between population density levels in 1995, 2000, 2005, 2010 and 2015 and accordingly reflect 5-year growth rates. These growth rates are also interpolated at an annual level. Information on net migration takes birth rates and death rates into account by substracting the natural increase in population from the change in population density, and therefore reflects the number of people migrating into a grid. Net migration information

"cattle", "herd", "grazing", "nomad", "pasture", "water".

[7] This categorization is based on cities defined by the "World Cities Database" and considers cities with at least 100k inhabitants. The city area is approximated by a 5km buffer around the city center.

is available for the period 1990-2000.

Details about data sources and variable construction are contained in the Appendix in Table 1A.1 while Table 1A.2 presents summary statistics for the main variables.

1.3 Weather Extremes and Conflict

1.3.1 Empirical Framework

This section presents results for the short-run effects of the occurrence of weather extremes on conflict incidence. The empirical framework models conflict in grid cell i in month t as

$$c_{i,t} = \alpha + \beta TE_{i,t} + \delta \mathbf{X}_{i,t} + \nu_i + \nu_t + \nu_m + \nu_{m\cdot E} + \nu_{i\cdot T} + \epsilon_{i,t} \ , \qquad (1.1)$$

where $c_{i,t}$ is the incidence of a conflict event in grid cell i in month t, $TE_{i,t}$ is the prevalence of a temperature extreme in grid cell i in month t as described in the previous section, and $\mathbf{X}_{i,t}$ is a vector of weather controls, which include, in particular, the average temperature and precipitation in a cell during a given month. The empirical model accounts for time invariant heterogeneity across grid cells by ways of grid cell fixed effects ν_i, and for time-specific waves in conflict incidence by including fixed effects for each year of the observation period, ν_t, as well as for month-of-year (calender month) fixed effects ν_m. In addition, more extensive specifications account for country-specific time-varying factors by the inclusion of country-specific year fixed effects $\nu_{i\cdot T}$, or month effects that are allowed to vary by the location relative (in terms of North or South) to the equator, $\nu_{m\cdot E}$. The error term $\epsilon_{i,t}$ allows for clustering within cells, as well as for spatial clustering among neighboring cells (Conley robust standard errors) in some of the robustness checks.[8]

The identification of the coefficient of interest β relies on the assumption that the occurrence of a weather extreme, $TE_{i,t}$, in a cell and month is exogenous to the occurrence of a conflict event in this cell during this month. The data on weather events is from reanalysis data based on raw data from different sources, including in particular satellite data as described above. Variation in this variable relative to a threshold for extreme events that is based on the 95% interval in the pre-analysis period 1979-1997 is therefore not systematically influenced by the occurrence of (small-scale) conflict. Hence, the identifying assumption is that Variation in $TE_{i,t}$, conditional on the set of controls, is exogenous to conflict incidence is plausibly satisfied.

[8] When adjusting the standard errors for spatial dependency along the lines Conley (1999) and Conley (2008). The distance cut-off for spatial contiguity is 200km (and thus includes two neighbouring grid cells in each direction) and the lag cut-off is 20. The results are robust to alternative specifications.

1.3.2 Baseline Results

Table 3.2 displays the results for different specifications of the empirical model, with grid cell fixed effects in Column (1), grid and year fixed effects in Column (2), additional calender month effects by hemisphere in Column (3), and with country-specific year effects in Column (4). Across all specifications, the occurrence of an extreme temperature event implies a significantly higher likelihood of the incidence of a violent event. Quantitatively, the effect in Columns (2)-(4) corresponds to an increase in the likelihood of conflict incidence in a given month and cell of 0.002, or approximately 7% compared to the unconditional mean of 0.03. In line with earlier results on the role of rainfall on conflict in mainly agricultural regions (Miguel, Satyanath, and Sergenti, 2012; Harari and Ferrara, 2018), the results also indicate that a shortage of rain increases conflict incidence, presumably due to resource constraints. It is worth noting, however, that these earlier studies effectively used variation at the yearly level, or variation in rainfall during particular seasons of the year (in particular growing seasons), whereas the present application accounts for recurrent heterogeneity in conflict activity during neuralgic months by the inclusion of year and calender month fixed effects. In this sense, the results complement these earlier findings, while pointing at an independent significant effect of weather extremes. The finding that temperature extremes are associated with higher conflict risk confirms previous findings, particularly by O'Loughlin, Linke, and Witmer (2014), using a refined measure of temperature extremes and considering a sample that includes the entire African continent. Moreover, the variation in temperature extremes contains relevant information for conflict incidence in terms of predictive power.[9]

The estimates for the baseline specification correspond to the effect of an extreme weather event in a given month on conflict incidence. One might suspect that the duration of this extreme is not irrelevant for the implications for conflict. In order to explore the sensitivity of the results with respect to the length of extreme events, we estimated an extended model, where the prevalence of extreme events was decomposed into events that lasted for exactly one month, for exactly two months, or three or more months. Table 1.2 presents the corresponding results. The findings indeed suggest that the effect of extreme weather events on conflict is non-linear in the length of the extreme weather events. The effect is larger for extreme events that last for two months than for events that last only one month, presumably due to the greater impact on distress and hardship associated with extended extreme events. On the other hand, events that last for three months or longer do not seem to have an independent effect, probably because of possibilities to cope with prolonged extreme events or other reactions of adjustment, relief, or adaptation. Generally,

[9]See the ROC-curve in Figure 1A.4.

Table 1.1: Baseline Results: Extreme Temperature Events and Conflict

	(1)	(2)	(3)	(4)
Dep. var.: incidence civil conflict				
Extreme Event	0.525***	0.197***	0.191***	0.165***
	(0.0626)	(0.0588)	(0.0596)	(0.0563)
Temp (mean)	0.00650***	0.00250	0.00381	0.00284
	(0.00240)	(0.00238)	(0.00323)	(0.00325)
Prec (mean)	-0.00196***	-0.00119***	-0.00112***	-0.00110***
	(0.000287)	(0.000280)	(0.000289)	(0.000273)
Adjusted R^2	0.000	0.016	0.016	0.244
N	1100328	1100328	1100328	1100328
Grid	4826	4826	4826	4826
Grid FE	✓	✓	✓	✓
Time FE		✓	✓	✓
Month FE			✓	✓
Month×Equator FE			✓	✓
Country×Year FE				✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

the length of extreme events may trigger two opposing effects that are reflected in these results. On the one hand, the severity of distress increases with the endurance of the shock which in turn aggravates it's impact on conflict risk. On the other hand, societies may start to react to the shock and implement coping strategies. Further, it is conceivable that weather shocks do not turn into civil conflict instantaneously. Of course the time span of reaction depends on the mechanism in place; relevant factors might for instance be the speed of bio-geographic transformation processes, or, depending on the conflict type, the formation process of conflict. All variables are measured as monthly means, hence it is conceivable that the (main) reaction to the weather shock is captured if including a second month in the analysis. Figure 1A.5 shows the effect of temperature extremes that have been going on since t=1 to t=7 months with estimation results shown in Table 1A.4. It shows that the temperature extremes unfold their impact on conflict most strongly in the second month. These results also document, however, that the baseline specification delivers a conservative estimate of the effect of extreme weather events on conflict.

Instead of estimating the effect on any conflict event, one might wonder about the potential heterogeneity in the effect of weather extremes on conflicts of different types. To investigate this conjecture, we replicate the analysis while restricting attention to particular conflict types. In particular, using the categorizations and narratives supplied with the ACLED data, we consider battles resulting in a change of contested territory, battles

Table 1.2: Baseline Results: Extreme Temperature Events of Different Lengths

	(1)	(2)	(3)	(4)
Dep. var.: incidence civil conflict				
One Month Extreme	0.137*			0.148**
	(0.0703)			(0.0705)
Two Months Extreme		0.486***		0.498***
		(0.147)		(0.147)
Three or more Months Extreme			-0.0219	0.00831
			(0.117)	(0.117)
Adjusted R^2	0.016	0.016	0.016	0.016
N	1100328	1100328	1100328	1100328
Grid	4826	4826	4826	4826
Grid FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

without changes of territory, conflicts involving farmers and/or pastoralist (agropastoral conflicts), conflicts occurring in rural areas, and riots. To gain some overview, we replicate the analysis for the entire sample. The corresponding results are shown in Table 1.3. Weather extremes mainly show a positive effect on territorial conflicts and conflicts in agricultural/rural areas.

Interacting Factors: Population Density and Caloric Suitability. One problem pervading the existing literature on the effects of climate for conflict is the lack of a common and coherent narrative underlying the evidence. One reason for this lack might be the fact that climate or weather extremes affect individuals in different ways depending on the respective living environment. In the following, we explore the role of some of the interacting factors that have been mentioned in the discussion.

The existing literature suggests that extreme weather events have particularly devastating effects on health and resources in areas in which the population exhibits a high degree of vulnerability due to low resilience. Such areas are typically associated with high population density and low agricultural productivity. In order to test this conjecture, we consider an extended specification in which the effect of weather extremes on conflict is allowed to vary with population density or the suitability of the soil for food production, as measured by the potential caloric yield per unit of land. The corresponding results are

Table 1.3: Baseline Results: Extreme Temperature Events and Conflict Types

	(1) Battle (terr)	(2) Battle (non-terr)	(3) Agropastoral	(4) Rural Region	(5) Riot
Extreme Event	0.147***	0.000455	0.0508***	0.179***	0.0184
	(0.0405)	(0.0115)	(0.0167)	(0.0560)	(0.0195)
Adjusted R^2	0.003	0.000	0.003	0.013	0.002
N	1100328	1100328	1100328	1100328	1100328
Grid	4826	4826	4826	4826	4826
Grid FE	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

contained in Table 1.4. In fact, weather extremes exhibit a greater effect in areas with high population density. Likewise, the effect is amplified in areas where the caloric suitability of the soil is comparably low. Both effects are present in isolation as well as when added jointly in the same specification.[10]

The fact that permanent agricultural scarcity or high population pressure significantly weakens the resilience to climatic shocks points towards a channel that works through rising agricultural scarcity in the face of extreme climatic conditions. This hypothesis is supported by the finding that particularly conflicts in a rural and agricultural context or conflicts involving territorial changes are particularly sensitive to the occurrence of weather extremes. One facet of this mechanism might be climatic stress in the growing season of grid specific crops which translates into increasing annual conflict risk (Harari and Ferrara, 2018). However, also alternative mechanisms are conceivable such as battles over fruitful land or over water resources that may occur in any months of the year. To take a closer look at this nexus we estimated empirical models with more extensive specifications including cell-year-specific fixed effects into the regressions. The results reveal that indeed part of the effect of monthly temperature extremes on conflict risk is accounted for by grid-specific annual factors.[11] This finding is consistent with a mechanism as outlined above or might

[10]Notice that the main effect of soil suitability, which is a time-invariant variable, is absorbed by the cell fixed effects. The algorithm on which the population density data are based has changed in 2005 (from v3 to v4, see CIESIN, 2016). To account for this change in the variable construction, we replicated the estimation with a more flexible specification that additionally includes an interaction term of extremes with an indicator variable that reflects the timing of this change. The results are shown in Table 1A.5 in the Appendix and suggest that the main results remain unaffected.

[11]Table 1A.6 in the Appendix shows the corresponding results for an extended specification of Table 1.4 and a substantial increase in the variation explained by the empirical model.

indicate a grid-specific trend of climate and conflict pointing towards potential long-run consequences of global warming. This will be subject of analysis in Section 1.4. Notably, even in this highly restrictive specification that accounts for grid-year-specific fixed effects, temperature extremes retain an effect on conflict that remains statistically significant and quantitatively relevant, in particular when accounting for regional vulnerability in terms of high population pressure.

Table 1.4: Extreme Temperature Events and Conflict: The Role of Population and Productivity

	(1)	(2)	(3)
Dep. var.: incidence civil conflict			
Extreme Event	1.161***	0.400***	2.653***
	(0.222)	(0.0829)	(0.402)
ln Pop. Density	-1.109***		-1.121***
	(0.246)		(0.245)
Extreme Event×ln Pop. Density	0.191***		0.352***
	(0.0353)		(0.0535)
Extreme Event×Caloric Suitability		-0.000245***	-0.000795***
		(0.0000755)	(0.000131)
Adjusted R^2	0.016	0.016	0.016
N	1085604	1099188	1085292
Grid	4814	4821	4811
Grid FE	✓	✓	✓
Time FE	✓	✓	✓
Month FE	✓	✓	✓
Month×Equator FE	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Alternative Measures and Estimation Methods. The analysis so far has focused on extreme weather events in terms of temperature. Much of the existing literature has instead focused on rainfall and droughts. Additional results for alternative specifications of extreme events that focus on either low precipitation or droughts deliver qualitatively similar results.[12] In light of the research question underlying this paper – whether climate

[12]See Tables 1A.7 and 1A.8 for the main effects and for the results regarding population density and caloric suitability. See Table 1A.9 for results when including all weather variables in one specification. See Figure 1A.2 for the evolution of precipitation (extremes) and conflict over time. We do however find that the effect of low precipitation extremes and drought is more sensitive to the inclusion of time fixed effects, an issue that has been discussed previously by Couttenier and Soubeyran (2014).

change and global warming cause conflicts – we concentrate attention on the effects of extreme temperature events in the remainder of the analysis.

We analyze whether the relationship between temperature extremes and civil conflict is robust to different specifications. First, we investigate the effects of temperature extremes on conflict along the intensive margin. In particular, we consider the effect on the number of conflict incidences as well as the number of conflict-related fatalities. Temperature extremes have a significantly positive effect on both measures, indicating that beyond being a trigger for conflict at the extensive margin, temperature extremes also affect the severity of conflicts. This implies that ongoing conflicts can be aggravated by the appearance of climatic shocks.[13] To investigate this issue in more detail, we analyze the impact on conflict onsets separately from that on incidence by setting the conflict measure to 1 in the month of the onset of a new conflict, and 0 otherwise. While the results point to an overall significant impact, the effect is quantitatively smaller and statistically insignificant for some specifications, which suggest that part of the effect is related to the prolongation of ongoing conflict.[14] The estimation of average marginal effects using a logit model also confirms the baseline results. Finally, to correct for spatial auto-correlation of the residuals we employ Conley-robust standard errors.[15]

1.3.3 Additional Results

The Role of Population Dynamics. The previous findings suggest that temperature extremes affect individuals differently in different environments. This implies that extreme events might exhibit interactions with local conditions, reflecting different living conditions and resilience, which might be related to population dynamics and migration patterns. In particular, the occurrence of weather extremes in environments in which people are forced to move away from their homes might create different tensions than in regions that attract population inflows. Obviously, weather extremes might affect conflict in different ways and through different mechanisms, leading to potentially complicated and multi-faceted patterns of effects and narratives. To investigate this issue, we consider population dynamics in terms of population growth and migration as factors that might interact with weather extremes, thereby providing further insights regarding the vulnerability of regions towards climatic shocks.

In an attempt to address this issue, we replicate the previous analysis but group the grid cells by their population dynamics in terms of population density, population growth, or net migration. The split by quartiles of population density effectively constitutes a

[13]See Table 1A.12.
[14]See Table 1A.10.
[15]See Table 1A.12.

different way of replicating the analysis of Table 1.4. To avoid the potential concern that population dynamics are simultaneously influenced by temperature extremes in the estimation month we base the sample splits on pre-annual values of population density and population growth, while migration data comes mostly from the pre-analysis period (1990 to 2000). This avoids having to consider the potential feedbacks of conflicts on other causes of population growth and migration and allows focusing attention to the exploration of the effect of climatic distress caused by temperature extremes in the face of differential population dynamics.

The findings, shown in Panel A of Table 1.5, reveal a non-linear effect for cells characterized by different population density. In particular, we find a positive effect of weather extremes on conflict in the most densely populated quartile of cells. In contrast, the findings indicate that weather extremes have no effect on conflict in cells belonging to the three other quartiles.

Panels B and C adopt a more dynamic perspective by focusing on the heterogeneity in the effect by population dynamics. In Panel B of Table 1.5, the cells are split into quartiles by population growth. While this can include any reasons for population growth, i.e., fertility, mortality and migration, this setting delivers additional insights into the vulnerability and crowding that might be the key factor that leads to an effect of weather extremes on conflict. The results reveal a different picture than those obtained when considering heterogeneity in the level of population density. In particular, when considering population growth, the effect of weather extremes turns out to be u-shaped. The effect is effectively zero (and even negative but insignificant) in the two intermediate quartiles, whereas weather extremes appear to have a significantly positive effect on conflict in the quartiles with the lowest and the highest rates of population growth.

In order to identify the role of migration, Panel C of Table 1.5 presents results for a sample split by net migration rates. The results document a similarly u-shaped pattern, although the effect is significantly positive only for the lowest quartile, and positive but insignificant for the highest quartile. Note that the results are not affected by ongoing conflict events, as documented by results for an extended specification that includes lagged conflict incidence as control variable.[16]

These findings are consistent with completely different narratives and mechanisms for cells characterized by different population dynamics. Areas with the highest rates of population growth or net migration are likely to be destination areas for refugees, which are presumably also areas of higher density. In these areas, weather shocks might also constitute a major threat to the provision with resources, thereby triggering conflicts. Also, it is conceivable that the areas with the lowest population growth or lowest (most negative)

[16]See Table 1A.13 in the Appendix.

net migration are areas where where conflicts are most sensitive to temperature extremes. Likewise, these are the areas that individuals might see themselves forced to leave in the face of environmental shocks and emerging subsistence constraints.

Table 1.5: Extreme Temperature Events and Conflict: The Role of Population Dynamics

	Dependent Variable: Conflict Incidence			
Quartile	Q1	Q2	Q3	Q4
	Panel A: Population Density			
Extreme Event	0.00537	-0.134	0.198	0.332**
	(0.0396)	(0.0820)	(0.130)	(0.168)
Adjusted R^2	0.003	0.010	0.018	0.030
N	257616	257424	257508	257100
Grid	1326	1506	1566	1402
	Panel B: Population Growth			
Extreme Event	0.424***	0.159	-0.136	0.224**
	(0.133)	(0.124)	(0.0985)	(0.114)
Adjusted R^2	0.015	0.023	0.012	0.004
N	257088	256812	256896	257004
Grid	2415	2949	2999	2621
	Panel C: Net Migration			
Extreme Event	0.546***	-0.0879	-0.00494	0.161
	(0.157)	(0.0969)	(0.0607)	(0.131)
Adjusted R^2	0.027	0.014	0.004	0.019
N	273828	274056	273828	274056
Grid	1201	1202	1201	1202
Grid FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓

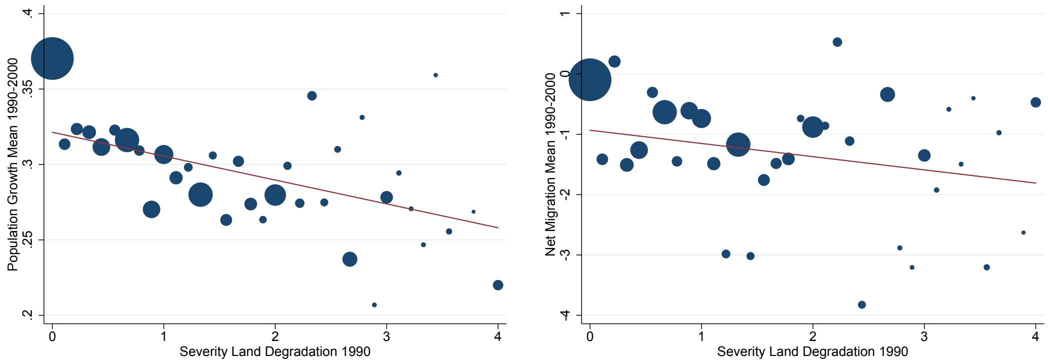
OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. The assignment to quartiles related to population density and population growth is based on values from the respective previous year. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

The previous findings indicate that the effect of weather extremes on conflict incidence might exhibit a different intensity depending on the concrete environment, as reflected by the respective population dynamics. To investigate whether there is also heterogeneity regarding the type of conflicts that are triggered by weather extremes in different environments, we replicated the analysis for different conflict types while distinguishing cells by population dynamics. The corresponding results, which are shown in Table 1A.14 in the Appendix, document that the pattern of the pooled analysis is largely confirmed when considering the quartiles of cells that grow least in terms of population or experience the greatest level of out-migration (Panels A and C). In contrast, in cells that grow in terms of population or experience in-migration, weather extremes have weaker effects on battles involving territorial changes or rural conflicts, but instead are associated with the outbreak of non-territorial battles and riots (Panels B and D). This heterogeneity in conflict types substantiates the conjecture that differential mechanisms are in place that link temperature shocks to the outbreak of conflict, depending on the local circumstances in terms of population dynamics.

Land Degradation. One factor that is often discussed as potential driver of migration and that has received revived interest in the context of climate change is environmental degradation. To investigate whether environmental stress relates to the vulnerability found for cells with the lowest population growth or migration we first look at the relationship between land degradation and population growth or net migration, respectively. Figure 1.3 illustrates this relationship. The plot suggests that indeed areas with greater land degradation in the sense of sensitivity of soil productivity experience lower population growth and more out-migration in the following decade.

To explore this aspect in more detail, we repeat the analysis by allowing for differential effects of weather extremes in cells with different degrees of land degradation in our analysis. Table 1.6 presents the corresponding estimates. The results document that weather extremes mainly affect conflict in cells with higher degrees of land degradation, supporting the vulnerability hypothesis and pointing towards a connection between the effects found in the first quartile of population growth or migration and environmental degradation. In additional results for specifications that also distinguish between different conflict types and account for for regions with low and high levels of land degradation, respectively, the pattern of the pooled analysis is largely confirmed.[17] The results also show that this is particularly the case for grid cells that experienced strong land degradation. Here the effect corresponds closely to the findings for grid cells with low population growth or net migration where battles involving changes of territory, conflicts in the context of agriculture and pastoralism and in rural regions are significantly affected by temperature shocks. Overall, the results suggest that the effect found in the lower tail of the distribution of population dynamics might be related to scarcities arising from declines in agricultural production. In contrast, the effect found in the upper tail of the distribution of population dynamics seems to unfold in a different context, potentially involving greater vulnerability arising from increased population pressure.

[17] The results are shown in Table 1A.15 in the Appendix.



(a) Population Growth (b) Net Migration

Figure 1.3: Land Degradation and Population Dynamics
This figure plots population density and net migration relative to the degree of land degradation. Severity of land degradation is rounded to the closest decimal level (0.1). Population growth (over the period 1990-2000) and net migration (over the period 1990-2000) are then averaged for each bin. The marker size corresponds to the number of grid-month observations in each bin. The figure is based on a sample that excludes cells that are outliers in terms of population growth or net migration (measured by the 95th percentile).

Table 1.6: Extreme Temperature Events and Conflict: The Role of Land Degradation

	(1) All	(2) High Degradation	(3) Low Degradation
Dep. var.: incidence civil conflict			
Extreme Event	-0.0529	0.306***	0.0458
	(0.0773)	(0.0943)	(0.0717)
Extreme Event×Land Degradation	0.249***		
	(0.0647)		
Adjusted R^2	0.016	0.020	0.011
N	1066356	558600	541728
Grid	4677	2450	2376
Grid FE	✓	✓	✓
Time FE	✓	✓	✓
Month FE	✓	✓	✓
Month×Equator FE	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. The sample split into grids with high- and low land degradation is based on the median value of land degradation in 1990. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

1.4 Long-Run Trends: Climate Change and Conflict

The results so far document that the occurrence of weather extremes is associated with the incidence of violent conflict events in high-resolution data of monthly frequency and a narrowly defined spatial environment of grid cells of 0.75^o latitude and longitude. As discussed in the Introduction, however, the frequency and severity of extreme weather events seems to have increased over the past three decades. Figure 1.4 provides additional evidence for this by plotting the average number of extreme temperature events per month at the grid cell level, weighted by the duration of the respective extreme events. The figure shows that the frequency of events as well as the duration of these events have increased. In light of the results presented in the last section, this suggests that climate change might have considerable consequences for the incidence of conflict. In contrast to being exposed to short-run fluctuations, societies may adapt to gradually moving levels of temperature or extremes. But it might also be the case that the effect exacerbates, for instance when the underlying vulnerability due to ongoing environmental degradation rises. In order to investigate this issue in more detail, this section presents the results of a long-run analysis of trends in temperature extremes and conflict incidence.

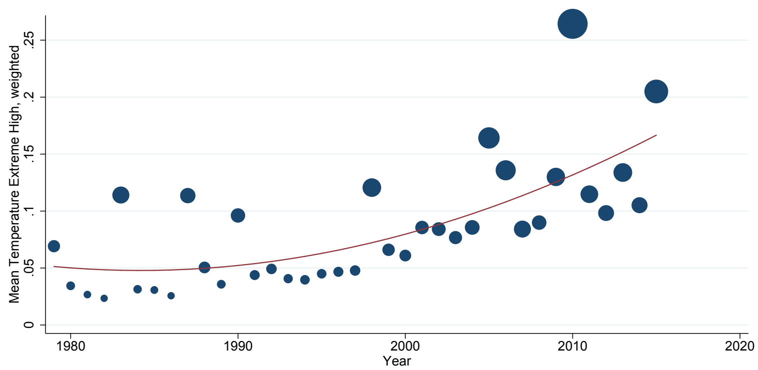


Figure 1.4: Dynamics of Extreme Temperature Events in Africa
This figure plots the average annual number of extreme temperature events across all cells by year, weighted by the average length of extreme events in terms of consecutive months. See Section 3.2 for a detailed data description.

1.4.1 Empirical Framework

The analysis is based on variation in the long-run trends of extreme temperature events and conflict incidence over the observation period 1997-2015. We conduct two sets of analyses that are based on the logic of a difference-in-differences (DiD) approach in long differences and a generalization thereof. The baseline empirical framework underlying the long-differences analysis splits the data into two equally long observation periods, 1997(1)-2006(6) and 2006(7)-2015(12) and computes the difference in the average incidence of conflict in a cell and relates it to the difference in the average incidence of extreme temperature events. In particular, with $\overline{C}_{i,1}$ denoting the average conflict incidence in grid cell i during the first half of the observation period (1997-2006), and $\overline{C}_{i,2}$ denoting the average conflict incidence in the same grid cell during the second half of the observation period (2006-2015), the (long) difference in conflict incidence in cell i is given by $\overline{DC}_i = \overline{C}_{i,2} - \overline{C}_{i,1}$.[18] This difference in conflict incidence is then related to the long difference in weather extremes, which is computed analogously as the difference in the average frequency of extreme temperature events in grid cell i during the second half of the observation period, $\overline{TE}_{i,2}$, relative to the corresponding average during the first half of the observation period, $\overline{TE}_{i,1}$, as $\overline{DTE}_i = \overline{TE}_{i,2} - \overline{TE}_{i,1}$. The long-difference estimation is then based on the empirical model

$$\overline{DC}_i = \alpha + \gamma \overline{DTE}_i + \phi_r + \phi_c + \epsilon_i \ , \tag{1.2}$$

where the empirical specification includes controls for climate zone fixed effects ϕ_r and country fixed effects ϕ_c. This setting corresponds to a difference-in-differences (DiD) estimator using long differences.[19] In essence, this specification corresponds to the graphical illustration in Figure 1.2. The identification of the coefficient of interest, γ, relies on the assumption of a common trend in conflict, α, across all cells within a given climate zone and country.[20]

To relax the common trend assumption, we also apply a generalized version of this long-differences estimator that is based on differences over three time periods. In particular, we repeat the analysis by dividing the observation period into three sub-periods (1997-

[18]Averages are calculated over the entire first half of the estimation period (9.5 years) instead of restricting to observations at arbitrary short time windows at the beginning and end of the panel. This reduces the concern of averages being driven by outlying years in terms of conflict incidence or climate. However, the results are robust to alternative specifications. See for instance table 1A.18 and 1A.20 for an alternative specification using a 7-year window or table 1A.19 and 1A.21 using a 8-year window at the beginning and end of the sample period to construct averages for the long-difference estimation.

[19]A similar approach is applied by Dell, Jones, and Olken (2012) who analyze the impact of differences in temperature on differences economic growth in the long run.

[20]Note that without controls for climate zone and country fixed effects, the estimation of (1.2) is equivalent to a regression in levels of $\overline{C}_{i,t}$ on $\overline{TE}_{i,t}$, $t = 1, 2$, with the inclusion of cell fixed effects.

2002, 2003-2008, 2009-2015), and computing the respective differences $\overline{DC}_{i,t}$ and $\overline{DTE}_{i,t}$. Since there are three sub-periods, this allows computing two differences per cell i and, consequently, estimating an extended model in differences that includes cell fixed effects μ_i and a trend component $I23$ that reflects the differences between the second and third sub-period. The generalized long difference model is then estimated as

$$\overline{DC}_{i,t} = \mu_i + \gamma \overline{DTE}_{i,t} + \psi I23 + \epsilon_{i,t} \; . \tag{1.3}$$

1.4.2 Results

Panel A of Table 1.7 presents the results for the long-differences specification (1.2). The results show separate estimates for the effect of variation in the frequency of extreme temperature events over the two sub-periods, and of the long difference in temperature, as well as both.[21] The results show that cells that experienced a stronger increase in temperature extremes, or in temperature, also experienced a more pronounced increase in conflict incidence. This finding indicates that societies do not fully adapt to slowly changing levels of mean temperature or extreme events. The results for the joint specification in Column (3) further indicate that the increase in weather extremes exhibits the stronger and more robust effect.

In order to explore the robustness of these results with respect to potentially different trends in climate and conflict at the grid-level, we also estimated the model for long differences across three sub-periods (1997-2002, 2003-2008, 2009-2015) applying the Generalized DiD model (1.3). The respective results are presented in Panel B of Table 1.7. By considering three periods, the model effectively accounts for grid-specific trends in conflict incidence and thus identifies how this trend is affected by trend changes in weather extremes (and temperature), thereby relaxing the common trend assumption underlying the DiD estimator with only two time periods in Panel A. The results indicate that there is an increase in the frequency of conflict incidence over time, as reflected by the trend coefficient ψ for $I23 = 1$. More importantly, even beyond grid-specific trends and this overall trend increase, a rise in the incidence of temperature extremes is related to rising conflict incidence in the long run. The results in Table 1.7 Column (3) indicate that an increase in the frequency of temperature extremes (reflected by an increase in differences of long-run averages by 1) leads to an increase in the frequency of conflicts by 0.084 or 0.042, respectively. For the latter, this corresponds to twice the unconditional mean of the increase in conflict incidence in terms of differences of long-run averages, which is 0.018. This implies that the coefficients of the long-run analysis are considerably larger in magnitude than the

[21] Long differences in temperature have been computed analogously to long differences in conflicts and extreme events.

coefficients found in the short-run analysis.

Additional results obtained for a Generalized DiD model that allows for time-varying trends within countries and climate zones confirm these findings.[22] In further analyses, we also investigated whether the effect of climate change in terms of an increasing frequency of weather extremes and mean temperatures becomes more pronounced over time, i.e. whether the effect of climate change on conflict becomes stronger in the later sub-periods. Although the estimates indicate a positive coefficient of the impact of extremes on conflict in later sub-periods, the effects turn out to be statistically insignificant when including region-specific time-varying trends.[23] Hence, while we find evidence for an increase in the frequency of weather extremes leading to an increase in the frequency of conflict incidence, we find no evidence that this effect becomes stronger over time.

Table 1.7: Extreme Weather Events and Conflict: Long Differences

	Dependent Variable: Diff Conflict Incidence		
	Panel A: DiD (Two Periods)		
Diff Extreme Event	0.0951***		0.0838***
	(0.0220)		(0.0272)
Diff Temp		0.0188***	0.00537
		(0.00584)	(0.00719)
r2	0.219	0.217	0.219
N	4826	4826	4826
Climate Zone Trend	✓	✓	✓
Country Trend	✓	✓	✓
	Panel B: GDD (Three Periods)		
Diff Extreme Event	0.0206		0.0416**
	(0.0147)		(0.0182)
Diff Temp		-0.00189	-0.0106
		(0.00638)	(0.00811)
I23=1	0.0312***	0.0315***	0.0306***
	(0.00157)	(0.00158)	(0.00168)
r2	0.527	0.527	0.527
N	9652	9652	9652
Grid Trend	✓	✓	✓

Panel A: OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2006 and 2006-2015, ($\overline{DC}_i$), with one observation per grid cell. Panel B: OLS fixed effects estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2002 and 2003-2008, and between 2003-2008 and 2009-2015, ($\overline{DC}_{i,t}$), with two observations per grid cell. See text for details. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.17 replicates the analysis while allowing for heterogeneity in the effect by

[22]See Panel A of Table 1A.16 in the Appendix for the results.

[23]See Panel B of Table 1A.16 in the Appendix for the results.

agricultural productivity, population density, land degradation, population growth, net migration and absolute changes in population density. By and large, the results mirror those obtained with month-by-month variation. In particular, the increase in weather extremes entails particularly strong increases in conflict incidence in areas with low agricultural productivity, high population density, population density changes, and high vulnerability of agricultural production as reflected by greater land degradation. The main effect of the increase in weather extremes remains large and significant across all specifications.

1.5 Discussion

This paper provides novel evidence for the role of climate change reflected by the frequency of temperature extremes for violent conflict in Africa. Estimations based on a fine spatial resolution of 0.75° latitude and longitude and month-by-month variation document a positive effect of the occurrence of temperature extremes on conflict incidence. These effects increase with the severity of the extreme in terms of its duration, and are larger in highly densely populated regions, in regions with lower agricultural productivity, as measured by potential caloric yield, and in regions with more pronounced land degradation. The results also point towards heterogeneity in the effect regarding the type of violence and a crucial role of population dynamics. Regions experiencing an outflow of population exhibit different types of conflict in response to weather extremes than regions experiencing population inflows.

The findings of this paper also contribute to the debate about the role of climate change for conflict by documenting a link between the increase in the frequency and severity of extreme temperature events and conflict, using a (generalized) difference-in-differences approach spanning almost two decades, from 1997-2015. The results resemble those obtained for short-run variability in the sense that regions with a higher increase in extreme temperature events are shown to have experienced a larger increase in the incidence of violence. Also the differential effects with respect to population density, population dynamics, agricultural suitability and land degradation are confirmed, as well as the robustness to differential trends in regions that experience different severity of climate change.

The results help reconciling some of the open issues in the literature. The results provide evidence in line with the introductory quotes of UN secretary generals Ban Ki Moon and Antonio Guterres. In particular, the heterogeneity of the effects is consistent with different narratives for outbreaks of violence in the context of weather shocks or climate change that have been argued to be inconsistent with monocausal views of climate change affecting conflict. Moreover, the results illustrate the central role of interacting mechanisms, in particular population dynamics, migration, and environmental degradation, in linking weather shocks or climate change to violent conflict and different types of violent events. However, by the design of the empirical analysis using high-frequency and high-resolution data, the analysis does not provide direct evidence for the role of institutions or institutional failures for the nexus between climate and conflict. Future work is needed to isolate institutional aspects and options for policy in containing and avoiding climate-driven conflict.

1.A Appendix

Table 1A.1: Data Sources and Variable Construction

	Source	*Variable*	*Specification*
		Climate	
1.	European Centre for Medium-Range Weather Forecasts (ECMWF) Era-Interim dataset Dee et al. (2011) 1979-2014; 0.75°	Temp	Monthly mean temperature in degree celsius
		Prec	Monthly mean precipitation in mm (litre per sqm)
		Extreme	Monthly binary indicator 95% pctile of deviations from month specific means in 1979-1997. All temperatures in 1997-2015 above threshold are classified as extremes.
		Extreme Prec Low	Monthly binary indicator Definition according to temp extreme for low precipitation
		Drought	Monthly binary indicator Extreme = 1 & Extreme Prec Low =1
		Conflict	
2.	ACLED (Raleigh, Linke, Hegre, and Karlsen, 2010)	Incidence	Monthly binary indicator for conflict of any type
		Type:	
		Battle (territory)	Battle-Non-state actor overtakes territory Battle-Government regains territory
		Battle (non-territory)	Battle-No change of territory
		Riot and Protest	Riots/Protests
		Agropastoral	Keywords in contextual notes: “farm” “crop” “cattle” “herd” “grazing” “nomad” “pasture” “water”
		Location:	
		Rural	Outside of city area (5km buffer). City is defined by "World Cities Database" (Includes cities above 100k inhabitants like national capitals, provincial capitals, major population centers, and landmark cities)
		Cell-specific Characteristics	
3.	(Galor and Özak, 2016)	Caloric Suitability	Post 1500 Caloric yield in 1000 million kilo calories per hectar per year
4.	World ATLAS of Desertification (UNEP, 1992)	Severity Land Degradation	Grid mean of severity index [0,4]
5.	Climate Zones		Dry: Arid, semi-arid, Tropical: Tropical dry, tropical wet, Temperate: Humid subtropical, mediterranean, highlands, marine
		Demographic Characteristics	
6.	Gridded Population of the World (GPW), v3 (1995, 2000) and v4 (2005, 2010, 2015), (CIESIN, 2016)	Population Density	Annual interpolation from data in 1995, 2000, 2005, 2010, 2015. 1000 per square kilometer
		Population Growth	Annual interpolation from 5-year growth rates in 1995, 2000, 2005, 2010, 2015. Growth rates are calculated from population data (e.g. popgrowth_2000 = (pop_2000-pop_1995)/pop_1995
		Quartiles	Interpolated value of population density/growth in L12 falls into quartile threshold in L12.
7.	Global Estimated Net Migration Grids by Decade, v1, (CIESIN, 2016)	Net Migration 1990-2000	Population in time period 2 is substracted from the population in time period 1, and then the natural increase (births minus deaths) is substracted.

Table 1A.2: Summary Statistics

	Mean	SD	Min	P95	Max	Obs
Incidence Conflict	0.03	0.18	0.00	0.00	1.00	1,100,328
Temperature (degree Celsius)	24.80	5.50	2.80	34.32	41.97	1,100,328
Temperature Extreme	0.11	0.32	0.00	1.00	1.00	1,100,328
Pop. Density (in 1000 per sqkm)	0.04	0.13	0.00	0.15	4.97	1,087,428
Pop. Growth (5-year, in 100 percent)	0.34	8.67	-1.00	0.60	944.06	1,085,076
Net Migration (per sqm, 1990-2000)	-0.77	5.63	-97.45	2.40	47.67	1,095,768
Caloric Suitability (calories/hectar, year)	901.73	810.94	0.00	2140.42	2784.73	1,099,188
Severity Land Degradation Index [0,4]	1.00	0.98	0.00	3.00	4.00	1,066,356

Table 1A.3: Lagged Extremes

	(1)	(2)	(3)	(4)
Dep. var.: incidence civil conflict				
L.Extreme Event	0.517***	0.195***	0.177***	0.160***
	(0.0642)	(0.0604)	(0.0607)	(0.0597)
Temp (mean)	0.00836***	0.00402*	0.00597*	0.00419
	(0.00245)	(0.00242)	(0.00325)	(0.00334)
Prec (mean)	-0.00222***	-0.00130***	-0.00124***	-0.00135***
	(0.000289)	(0.000282)	(0.000291)	(0.000289)
Adjusted R^2	0.000	0.016	0.016	0.065
N	1095502	1095502	1095502	1027856
Grid	4826	4826	4826	4528
Grid FE	✓	✓	✓	✓
Time FE		✓	✓	✓
Country×Year FE				✓
Month FE			✓	✓
Month×Equator FE			✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

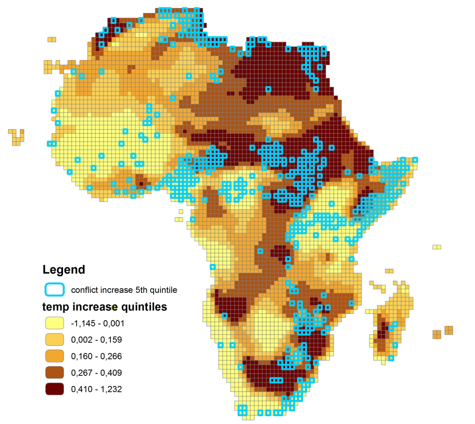
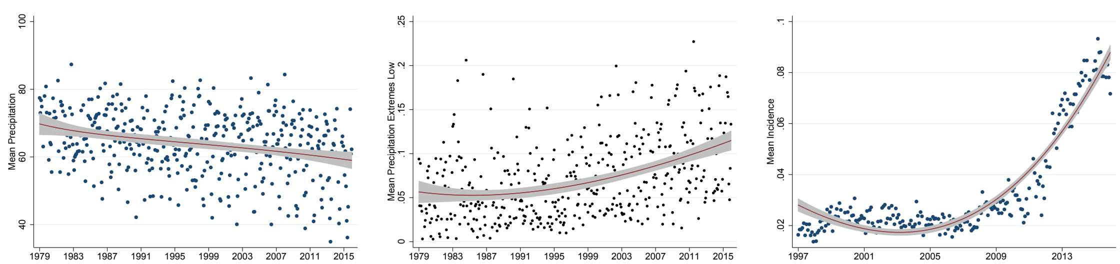


Figure 1A.1: Dynamics of Average Temperature and Conflict Incidence
This figure plots quintiles of changes in the average temperature from the first half of the panel (1997-2006) to the second half of the panel (2006-2015) together with the cells belonging to the highest quintile of changes in average conflict incidence from the first half to the second half of the panel.



(a) Average Precipitation (b) Precipitation Extremes (c) Conflict Events

Figure 1A.2: Dynamics of Precipitation, Extreme Precipitation Events and Conflict
This figure shows the monthly evolution of precipitation, low precipitation extremes and conflict. Additionally, it plots fractional-polynomial predictions with 95% confidence intervals. Precipitation extremes are coded as 1 if the absolute precipitation deviation from month-specific means for a given cell (calculated from the training period 1979-1996) falls below the 5 percentile threshold, and 0 otherwise. See Section 3.2 for a detailed data description.

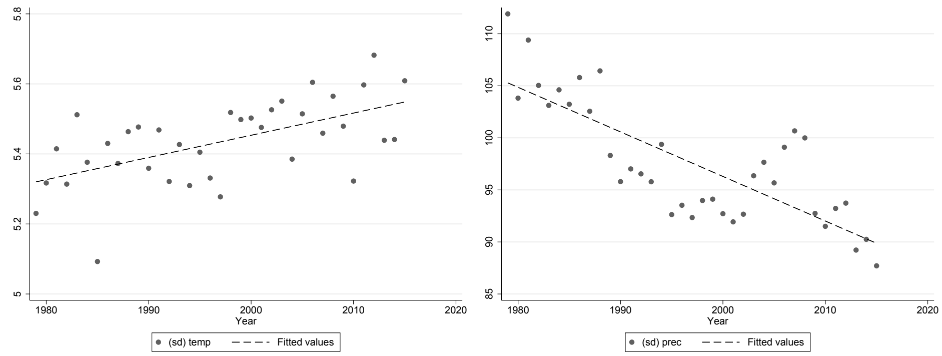


Figure 1A.3: Trend of Intra-Annual SD of Temperature and Precipitation

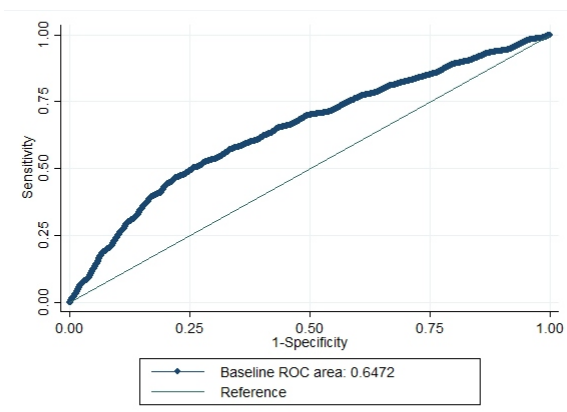


Figure 1A.4: ROC Curve Baseline Model

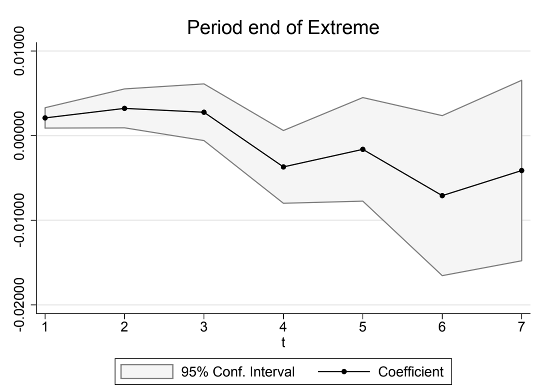


Figure 1A.5: Length of Extreme Events
This figure plots the effect of extreme events of various duration on conflicts as shown in Table 1A.4. Extremes may be ongoing after the respective period.

Table 1A.4: Length of Extreme Periods

	(1)
Dep. var.: incidence civil conflict	
t_1	0.209***
	(0.0622)
t_2	0.308***
	(0.118)
t_3	0.272
	(0.171)
t_4	-0.420*
	(0.220)
t_5	-0.168
	(0.313)
t_6	-0.766
	(0.483)
t_7	-0.437
	(0.544)
Temp (mean)	0.00341
	(0.00323)
Prec (mean)	-0.00112***
	(0.000289)
Adjusted R^2	0.016
N	1100328
Grid	4826
Grid FE	✓
Time FE	✓
Country×Year FE	
Month FE	✓
Month*times*Equator FE	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All coefficients are multiplied by 100. Length of extreme periods in months. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.5: Gridded Population of the World (GPW), v4, 2005-2015

	(1)
Dep. var.: incidence civil conflict	
Extreme Event	0.717***
	(0.222)
ln Pop. Density	-1.110***
	(0.246)
Extreme Event×ln Pop. Density	0.182***
	(0.0350)
v42005	6.748***
	(0.293)
Extreme Event×v42005	0.580***
	(0.124)
Adjusted R^2	0.016
N	1085604
Grid	4814
Grid FE	✓
Time FE	✓
Month FE	✓
Month×Equator FE	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. v42005 corresponds to a binary indicator that takes value 0 for the period pre 2005 (for v3 of the GPW data), and 1 thereafter (for v4). Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.6: Grid × Year Fixed Effects

	(1)	(2)	(3)	(4)
Dep. var.: incidence civil conflict				
Extreme Event	0.0508	0.298*	0.0761	0.584**
	(0.0518)	(0.175)	(0.0570)	(0.256)
Extreme Event×ln Pop Density		0.0478*		0.0790**
		(0.0263)		(0.0342)
Extreme Event×Caloric Suitability			-0.0000300	-0.000152*
			(0.0000635)	(0.0000852)
Adjusted R^2	0.368	0.371	0.371	0.368
N	1100328	1085604	1099188	1085292
Grid	4826	4814	4821	4811
Grid FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓
Grid*times*Year FE	✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. The main effect of population density is omitted because population density varies at yearly level. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.7: Extreme Events and Conflict: Precipitation

	(1)	(2)	(3)	(4)	(5)
Dep. var.: incidence civil conflict					
Extreme Prec Low	0.311***	0.0195	0.728***	-0.0416	1.296***
	(0.0664)	(0.0650)	(0.225)	(0.0681)	(0.370)
Temp (mean)	0.00964***	0.00568*	0.00746**	0.00582*	0.00749**
	(0.00246)	(0.00326)	(0.00329)	(0.00327)	(0.00329)
Prec (mean)	-0.00227***	-0.00123***	-0.00128***	-0.00123***	-0.00128***
	(0.000290)	(0.000290)	(0.000292)	(0.000291)	(0.000292)
ln Pop. Density			-1.100***		-1.103***
			(0.246)		(0.246)
Extreme Prec Low×ln Pop. Density			0.141***		0.203***
			(0.0344)		(0.0496)
Extreme Prec Low×Caloric Suitability				0.0000681	-0.000285**
				(0.0000760)	(0.000118)
Adjusted R^2	0.000	0.016	0.016	0.016	0.016
N	1100328	1100328	1085604	1099188	1085292
Grid	4826	4826	4814	4821	4811
Grid FE	✓	✓	✓	✓	✓
Time FE		✓	✓	✓	✓
Month FE		✓	✓	✓	✓
Month×Equator FE		✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.8: Extreme Events and Conflict: Droughts

	(1)	(2)	(3)	(4)	(5)
Dep. var.: incidence civil conflict					
Drought	0.449***	-0.0738	0.966***	0.192	2.346***
	(0.105)	(0.103)	(0.350)	(0.157)	(0.569)
Temp (mean)	0.00929***	0.00587*	0.00673**	0.00568*	0.00667**
	(0.00245)	(0.00325)	(0.00326)	(0.00325)	(0.00326)
Prec (mean)	-0.00220***	-0.00123***	-0.00126***	-0.00125***	-0.00130***
	(0.000288)	(0.000289)	(0.000290)	(0.000289)	(0.000291)
ln Pop. Density			-1.095***		-1.097***
			(0.246)		(0.246)
Drought×ln Pop. Density			0.220***		0.353***
			(0.0591)		(0.0773)
Drought×Caloric Suitability				-0.000258*	-0.000736***
				(0.000133)	(0.000186)
Adjusted R^2	0.000	0.016	0.016	0.016	0.016
N	1100328	1100328	1085604	1099188	1085292
Grid	4826	4826	4814	4821	4811
Grid FE	✓	✓	✓	✓	✓
Time FE		✓	✓	✓	✓
Month FE		✓	✓	✓	✓
Month×Equator FE		✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.9: Extreme Events and Conflict: Extended Specification

	(1)	(2)	(3)	(4)
Extreme Event	0.545***	0.290***	0.289***	0.205***
	(0.0686)	(0.0656)	(0.0665)	(0.0631)
Extreme Prec Low	0.271***	0.0963	0.0950	0.0519
	(0.0757)	(0.0738)	(0.0747)	(0.0706)
Drought	-0.270**	-0.396***	-0.408***	-0.180
	(0.130)	(0.130)	(0.130)	(0.128)
Temp (mean)	0.00630***	0.00264	0.00376	0.00284
	(0.00240)	(0.00238)	(0.00323)	(0.00325)
Prec (mean)	-0.00199***	-0.00119***	-0.00112***	-0.00110***
	(0.000290)	(0.000283)	(0.000291)	(0.000275)
Adjusted R^2	0.000	0.016	0.016	0.052
N	1100328	1100328	1100328	1100328
Grid	4826	4826	4826	4826
Grid FE	✓	✓	✓	✓
Time FE		✓	✓	✓
Country×Year FE				✓
Month FE			✓	✓
Month×Equator FE			✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.10: Onset of Conflicts

	(1)	(2)	(3)	(4)
Extreme Event	0.0657	0.388***	0.150***	0.924***
	(0.0439)	(0.141)	(0.0511)	(0.212)
ln Pop. Density		-0.326***		-0.330***
		(0.0819)		(0.0818)
Extreme Event×ln Pop. Density		0.0647***		0.123***
		(0.0213)		(0.0280)
Extreme Event×Caloric Suitability			-0.0000986*	-0.000285***
			(0.0000531)	(0.0000721)
Adjusted R^2	0.004	0.004	0.004	0.004
N	1100328	1085604	1099188	1085292
Grid	4826	4814	4821	4811
Grid FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.11: Alternative Specification 1

	(1)	(2)	(3)	(4)
Extreme Event pct95	0.337***	0.0327	0.0417	0.0822
	(0.0692)	(0.0677)	(0.0694)	(0.0689)
Temp (mean)	0.00456*	0.00335	0.00486	0.00271
	(0.00258)	(0.00254)	(0.00347)	(0.00346)
Prec (mean)	-0.00214***	-0.00129***	-0.00121***	-0.00115***
	(0.000289)	(0.000281)	(0.000291)	(0.000274)
Adjusted R^2	0.000	0.016	0.016	0.052
N	1100328	1100328	1100328	1100328
Grid	4826	4826	4826	4826
Grid FE	✓	✓	✓	✓
Time FE		✓	✓	✓
Country×Year FE				✓
Month FE			✓	✓
Month×Equator FE			✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All coefficients are multiplied by 100. In this specification, extreme events are coded as 1 if monthly temperature exceeds the 95 percentile threshold of the grid- (but not calendar month) specific temperature distribution. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.12: Alternative Specification 2

	(1) Number Conflicts	(2) Number Fatalities	(3) Logit	(4) Conley SE
Extreme Event	1.089***	29.47*	0.890**	0.191***
	(0.358)	(17.60)	(0.437)	(0.0680)
N	1100328	1100328	580716	1100328
Grid	4826	4826	.	.
Grid FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓
Country×Year FE				

The dependent variable is the number of conflicts in a given cell and month in Column (1), the number of fatalities in a given cell and month in Column (2), the incidence of a conflict event in a given cell and month in Columns (3) and (4). OLS (linear probability model) estimation results in Columns (1), (2) and (4), marginal effects of Logit estimates in Column (3). All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses in Columns (1)-(3). ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.13: Lagged Conflict as Control

	Dependent Variable: Diff Conflict Incidence			
Quartile	Q1	Q2	Q3	Q4
	Panel A: Population Density			
L.Incidence	20.73***	21.04***	25.08***	23.02***
	(2.982)	(1.201)	(1.068)	(0.784)
Extreme Event	-0.000311	-0.121	0.111	0.240
	(0.0370)	(0.0755)	(0.117)	(0.153)
Adjusted R^2	0.046	0.053	0.079	0.082
N	256274	255858	255887	255684
Grid	1326	1506	1566	1402
	Panel B: Population Growth			
L.Incidence	21.06***	24.43***	18.28***	18.80***
	(1.126)	(1.118)	(0.908)	(1.217)
Extreme Event	0.327***	0.0472	-0.154	0.170
	(0.123)	(0.110)	(0.0942)	(0.105)
Adjusted R^2	0.059	0.081	0.045	0.039
N	253903	252613	252414	253397
Grid	2415	2949	2999	2621
	Panel C: Net Migration			
L.Incidence	25.07***	23.84***	22.21***	25.16***
	(0.926)	(1.094)	(2.105)	(1.083)
Extreme Event	0.433***	-0.103	-0.0249	0.0882
	(0.140)	(0.0881)	(0.0553)	(0.116)
Adjusted R^2	0.088	0.069	0.052	0.081
N	272627	272854	272627	272854
Grid	1201	1202	1201	1202
Grid FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.14: Extreme Temperature Events and Conflict Types: Sample Splits

	Dependent Variable: Conflict Incidence				
Conflict Type	Battle (terr)	Battle (non-terr)	Agropastoral	Rural Region	Riot
	(1)	(2)	(3)	(4)	(5)
	Panel A: Population Growth 1st Quartile				
Extreme Event	0.190**	0.0324*	0.0907**	0.309***	-0.0117
	(0.0846)	(0.0186)	(0.0375)	(0.117)	(0.0361)
Adjusted R^2	0.002	0.000	0.002	0.013	0.001
N	257088	257088	257088	257088	257088
Grid	2415	2415	2415	2415	2415
	Panel B: Population Growth 4th Quartile				
Extreme Event	0.148*	-0.00709	0.0633**	0.210*	0.0884**
	(0.0807)	(0.0257)	(0.0282)	(0.111)	(0.0406)
Adjusted R^2	0.001	0.001	0.000	0.003	0.002
N	257004	257004	257004	257004	257004
Grid	2621	2621	2621	2621	2621
	Panel C: Net Migration 1st Quartile				
Extreme Event	0.386***	-0.0264	0.0648	0.452***	-0.0480
	(0.111)	(0.0286)	(0.0408)	(0.148)	(0.0470)
Adjusted R^2	0.006	0.001	0.005	0.024	0.005
N	273828	273828	273828	273828	273828
Grid	1201	1201	1201	1201	1201
	Panel D: Net Migration 4th Quartile				
Extreme Event	0.135*	0.0557**	0.0840**	0.183	0.0850*
	(0.0792)	(0.0280)	(0.0332)	(0.117)	(0.0509)
Adjusted R^2	0.002	0.000	0.002	0.013	0.003
N	273828	273828	273828	273828	273828
Grid	1201	1201	1201	1201	1201
Grid FE	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.15: Extreme Temperature Events and Conflict Types: The Role of Land Degradation

	Dependent Variable: Conflict Incidence				
Conflict Type	Battle (terr)	Battle (non-terr)	Agropastoral	Rural Region	Riot
	(1)	(2)	(3)	(4)	(5)
	Panel A: Low Land Degradation				
Extreme Event	0.0495	-0.0167	0.0438**	0.0446	0.0206
	(0.0511)	(0.0157)	(0.0191)	(0.0674)	(0.0229)
Adjusted R^2	0.003	0.000	0.001	0.009	0.002
N	541728	541728	541728	541728	541728
Grid	2376	2376	2376	2376	2376
	Panel B: High Land Degradation				
Extreme Event	0.224***	0.0179	0.0521*	0.280***	0.00898
	(0.0632)	(0.0172)	(0.0273)	(0.0884)	(0.0316)
Adjusted R^2	0.004	0.001	0.004	0.017	0.003
N	558600	558600	558600	558600	558600
Grid	2450	2450	2450	2450	2450
Grid FE	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. All regressions include average temperature and precipitation levels as controls. All coefficients are multiplied by 100. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.16: Timing of Long-Run Effect

	Dependent Variable: Diff Conflict Incidence		
	Panel A: GDD (Three Periods)		
Diff Extreme Event	0.0852***		0.0962***
	(0.0154)		(0.0184)
Diff Temp		0.0109**	-0.00557
		(0.00436)	(0.00519)
r2	0.224	0.221	0.224
N	9650	9650	9650
Climate Zone Trend	✓	✓	✓
Country Trend	✓	✓	✓
	Panel B: GDD (Three Periods) Time Effects		
Diff Extreme Event	0.0723***		0.0871***
	(0.0196)		(0.0196)
I23=1*Diff Extreme Events	0.0232		0.0137
	(0.0301)		(0.0363)
Diff Temp		0.00622	-0.00638
		(0.00515)	(0.00525)
I23=1*Diff Temp		0.0124	0.00325
		(0.00935)	(0.0117)
r2	0.224	0.222	0.224
N	9650	9650	9650
Climate Zone Trend	✓	✓	✓
Country Trend	✓	✓	✓

OLS fixed effects estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2002 and 2003-2008, and between 2003-2008 and 2009-2015, ($\overline{DC}_{it}$), with two observations per grid cell. See text for details. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.17: Extreme Weather Events and Conflict: Heterogeneity in Long Differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var.: diff civil conflict incidence							
Diff Extreme	0.0951***	0.220***	0.466***	0.0501*	0.160***	0.0938***	0.0997***
	(0.0220)	(0.0422)	(0.0694)	(0.0257)	(0.0456)	(0.0233)	(0.0223)
Diff Extreme×Caloric Suitability		-0.000118***					
		(0.0000299)					
Diff Extreme×Population Density			0.0837***				
			(0.0129)				
Diff Extreme×Land Degradation				0.0533***			
				(0.0193)			
Diff Extreme×Population Growth					-0.222*		
					(0.124)		
Diff Extreme×Net Migration						-0.00104	
						(0.00492)	
Diff Extreme×Diff Pop.							0.0218***
							(0.00434)
Diff Pop.							-0.0135***
							(0.00383)
r2	0.219	0.222	0.234	0.231	0.220	0.219	0.229
N	4826	4821	4808	4677	4808	4806	4749
Climate Zone Trend	✓	✓	✓	✓	✓	✓	✓
Country Trend	✓	✓	✓	✓	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2006 and 2006-2015, ($\overline{DC_i}$), with one observation per grid cell. See text for details. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.18: Long Run: Robustness 7-Year Window

	(1)	(2)	(3)
Dep. var.: diff civil conflict incidence			
Diff Extreme	0.0934***		0.0659**
	(0.0205)		(0.0262)
Diff Temp		0.0232***	0.0120*
		(0.00558)	(0.00714)
r2	0.224	0.223	0.224
N	4826	4826	4826
Climate Zone Trend	✓	✓	✓
Country Trend	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2004 and 2008-2015, ($\overline{DC}_i$), with one observation per grid cell. See text for details. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.19: Long Run: Robustness 8-Year Window

	(1)	(2)	(3)
Dep. var.: diff civil conflict incidence			
Diff Extreme	0.0774***		0.0597**
	(0.0202)		(0.0257)
Diff Temp		0.0179***	0.00782
		(0.00552)	(0.00703)
r2	0.219	0.218	0.219
N	4826	4826	4826
Climate Zone Trend	✓	✓	✓
Country Trend	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2005 and 2007-2015, ($\overline{DC}_i$), with one observation per grid cell. See text for details. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.20: Long Run: Interactions, Robustness 7-Year Window

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var.: diff civil conflict incidence							
Diff Extreme	0.0934***	0.238***	0.494***	0.0442*	0.141***	0.0923***	0.0973***
	(0.0205)	(0.0366)	(0.0422)	(0.0264)	(0.0307)	(0.0207)	(0.0206)
Diff Extreme×Caloric Suitability		-0.000128***					
		(0.0000268)					
Diff Extreme×Population Density			0.0910***				
			(0.00842)				
Diff Extreme×Land Degradation				0.0536***			
				(0.0146)			
Diff Extreme×Population Growth					-0.159**		
					(0.0760)		
Diff Extreme×Net Migration						-0.000782	
						(0.00175)	
Diff Extreme×Diff Pop.							0.0249***
							(0.00595)
Diff Pop.							-0.0134***
							(0.00259)
r2	0.224	0.227	0.242	0.236	0.225	0.223	0.234
N	4826	4821	4808	4677	4808	4806	4747
Climate Zone Trend	✓	✓	✓	✓	✓	✓	✓
Country Trend	✓	✓	✓	✓	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2004 and 2008-2015, ($\overline{DC_i}$), with one observation per grid cell. See text for details. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 1A.21: Long Run: Interactions, Robustness 8-Year Window

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var.: diff civil conflict incidence							
Diff Extreme	0.0774***	0.223***	0.480***	0.0358	0.136***	0.0761***	0.0829***
	(0.0202)	(0.0367)	(0.0419)	(0.0263)	(0.0327)	(0.0203)	(0.0203)
Diff Extreme×Caloric Suitability		-0.000128***					
		(0.0000268)					
Diff Extreme×Population Density			0.0911***				
			(0.00833)				
Diff Extreme×Land Degradation				0.0458***			
				(0.0147)			
Diff Extreme×Population Growth					-0.202**		
					(0.0874)		
Diff Extreme×Net Migration						-0.000972	
						(0.00170)	
Diff Extreme×Diff Pop.							0.0210***
							(0.00539)
Diff Pop.							-0.0134***
							(0.00256)
r2	0.219	0.223	0.238	0.230	0.220	0.219	0.228
N	4826	4821	4808	4677	4808	4806	4747
Climate Zone Trend	✓	✓	✓	✓	✓	✓	✓
Country Trend	✓	✓	✓	✓	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 1997-2005 and 2007-2015, ($\overline{DC}_i$), with one observation per grid cell. See text for details. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Chapter 2

Land Degradation and Armed Conflict: Evidence from the Sahel

2.1 Introduction

According to the Intergovernmental Panel of Climate Change (IPCC) global warming contributes to the degradation of the environment and thus to a sustained loss of the productive capacity of soils (IPCC, 2017). This evolution constitutes a threat to global food supply and is therefore of major relevance when discussing the implications of climate change for society. What is more, the issue of land degradation and the resulting scarcities are discussed as crucial drivers of conflict in the public debate.

The debate focuses on the Sahel region where the climate is exceptionally dynamic and unpredictable. The region experienced several severe droughts in the 70s and 80s that were largely unexpected given relatively abundant rainfall in the 50s and 60s (Giannini, Biasutti, and Verstraete, 2008). The chain of droughts jointly with excessive land use steadily dried up the landscape and caused a narrative of ongoing land degradation. The United Nations Convention to Combat Desertification (UNCCD) was founded in 1994 to counteract the initial adverse development. In 2007 it launched the "Great Green Wall Initiative" - a project that aims to restore Africa's degraded landscapes and to avoid a further spread of desert by planting a 8000km long vegetation belt from the east- to the west coast. Although recent evidence points towards an average "regreening" of the Sahel land degradation still constitutes a major concern in many regions. At the same time the Sahel suffers from frequent and disruptive conflicts with many of them being observed in times of climatic stress.

In this study I analyze the effect of vegetation dynamics on civil conflict in the Sahel. The main explanatory variable, vegetation barrenness, is constructed as a reversed scale from the Net Difference Vegetation Index (NDVI) which commonly serves as proxy for the greenness or productivity of vegetation. The NDVI is retrieved from satellite spectral reflectance measurements and therefore available at extremely high resolution and frequency which makes it unique for the purpose of this study. For the short-run analysis I estimate the monthly impact of vegetation barrenness on conflict incidence on a 0.75° grid within the period 2000-2012. Further, to isolate external climate driven components of land degradation I instrument vegetation barrenness with monthly precipitation levels. The residual of precipitation based predictions is commonly interpreted as "human component" in the literature that discusses the causes of land degradation (Herrmann, Anyamba, and Tucker, 2005). The productivity of vegetation may also be altered by anthropogenic impacts such as forestation, irrigation systems, the use of pesticides, cropland expansion etc.. Thus, to shed light on the human component of land degradation I analyze the effect of the residual on conflict incidence separately. The results unfold a positive effect of vegetation barrenness on conflict which is driven by rainfall shortages.

Further, I analyze regional and temporal characteristics that shape the vulnerability towards environmental shocks. I find that regions with extremely high densities of cultivation and pastoralism, scarce water resources or ethnic division react most strongly to vegetation barrenness shocks. Further the findings indicate that especially conflicts related to agriculture, pastoralism or territorial changes are involved. Looking at effect heterogeneity across seasons further shows that the effect mainly unfolds in the rainy season and during months of seasonal pastoralist movement.

Then I turn to the long-run implications of these findings in a long-difference analysis. The results suggest that regions facing more severe land degradation become more prone to conflicts related to agriculture or pastoralism in the long run. Reflecting results of the short-run analysis, the overall positive effect of land degradation on conflict is driven by declining rainfall. To shed more light on environmental factors that may impact conflict risk in the long run I further consider the effect of three different types of land cover changes on conflict incidence (forestation, grassland- and cropland expansion). I find indications that cropland expansions are related to rising conflict risk.

There are several contributions of this study to the literature. First, this study contributes to the literature that analyzes the effect of land degradation on conflict risk. It adds to this literature by employing a time-varying measure of land degradation, allowing to analyze the impact at monthly frequency as well as in a long-difference framework. Further, in contrast to previous studies the vegetation barrenness index employed in this analysis is based on satellite data which provides an external and very fine scaled measure. Second, the findings add to previous research that analyzes the impact of drought or precipitation shortages on conflict risk. In contrast to common drought indices vegetation barrenness is a more direct way to describe the productivity of land, which in turn may be disentangled into its climate and human-induced components. Therewith the present study emphasizes the role of vegetation as intermediary between climate and conflict and documents a positive effect of low plant productivity on conflict at the monthly level. Finally, the fine scale of the data allows for an analysis of regional and temporal factors shaping the vulnerability towards the environmental shocks. This adds new insights regarding a mechanism linking weather variability to conflict which is triggered by the loss of agricultural productivity.

This study relates to the concept of supply-induced scarcity which is in this case induced by changing supply of fruitful vegetation.[1] According to Collier and Hoeffler (1998, 2002, 2004) rising scarcity translates into conflict as follows. When individuals encounter lower economic returns from their activity, the opportunity costs of engaging into conflict decline, i.e. lower income thus increases the incidence of civil conflict. In the present study vegetation barrenness serves as a proxy for agricultural productivity of land. Accordingly, a decline in the fruitfulness of soil results in lower economic returns from production. A related line of arguments states that higher resource scarcity intensifies competition that can eventually lead to the outbreak of violent conflict. Further, several studies argue, among them Homer-Dixon (1999), that climate-induced scarcity puts additional stress on societies that interacts with societal traits to trigger the outbreak of civil conflicts.

The literature regarding the effect of environmental degradation on conflict risk reveals mixed evidence. Hendrix and Glaser (2007) analyse the impact of environmental factors such as freshwater resources, types of agriculture and land degradation on conflict in Sub-Saharan Africa at the country-year cell level. They do not find a significant impact of land degradation on conflict. Hauge and Ellingsen (1998) despite employing a similar approach, document a significantly positive impact of land degradation on conflict. Raleigh and Urdal (2007) analyze the impact of land degradation on conflict on a more disaggregated scale and perform a global cross-section analysis on a 100km$\times$100km grid. They show that higher levels of land degradation are related to an increased conflict risk. The studies cited above employ expert rating data on land degradation from the Global Assessment of Human-induced Soil Degradation (GLASOD) project. Soil degradation is measured statically and provided on a discrete scale from 0 to 4. Thus, the present study adds to this literature by providing new evidence from a time-varying, external and very fine scaled measure. The results obtained from a month-based analysis confirm a significantly positive impact of vegetation barrenness on the incidence of conflicts. Further, the analysis adds new insights on the long-run perspective, suggesting a positive effect of climate-induced land degradation on conflicts related to agriculture and pastoralism.

[1] Homer-Dixon (1999) distinguish supply-induced scarcity (e.g. land degradation), demand-induced scarcity (e.g. population growth) and structural scarcity (e.g. unequal distribution).

There are several studies that analyze the impact of drought or precipitation anomalies on civil conflict pointing towards the loss of agricultural productivity as a trigger of conflict. These studies relate to a more general controversy debating the causal pathway from different climatic variables to conflict.[2] By its econometric approach the present study is more deeply rooted in this strand of literature. Harari and Ferrara (2018) use a 1°x1° grid for Africa and find that a drought occuring during the growing season of a location specific crop increases the risk of conflict. Almer, Laurent-Lucchetti, and Oechslin (2017) build on the approach by Harari and Ferrara (2018) using a 0.5°x0.5° grid to analyze the impact of drought on smaller scale conflicts. They find that unusually dry months increase the incidence of riots in Sub-Saharan Africa. Further, they find that the effect is particularly pronounced where water resources are limited. Von Uexcull et al. (2016) focus on locations of politically relevant ethnic groups in Asia and Africa as unit of observation and approve that growing season droughts have a positive impact on conflict. The documented effect unfolds in the presence of poverty, political exclusion or high agricultural dependence. Maystadt, Calderone, and You (2015) analyze the impact of quarterly temperature and precipitation anomalies on civil conflict in North- and South Sudan on a 0.5°x0.5° grid. They find the positive effect on conflict to be most pronounced in areas with high livestock density or the presence of pastoral and agro-pastoral ethnic groups, referring to a history of conflicts between groups of herders and farmers over fruitful land. Mitigating factors are found to be availability of water resources and alluvial soil. Hendrix and Salehyan (2012) find that extreme deviations in rainfall increase the risk of political conflict and argue via an increased pressure on agricultural production. The present study documents a positive impact of precipitation driven vegetation barrenness on conflict risk. Therewith it confirms previous findings linking precipitation shortages to conflict and at the same time points towards the role of vegetation as intermediate factor. Further, the findings complement previous research by showing up regional and temporal factors shaping the vulnerability towards the environmental shocks. In this dimension, the findings reemphasize a mechanism that works through declining agricultural productivity.

The remainder of this paper is structured as follows. Section 2.2 describes the compilation of the data set as well as specifications of the main variables of interest. Section 2.3 presents results of the month-based analysis. Section 2.4 presents results from the long-differences analysis. Section 2.5 concludes the analysis with a discussion.

[2]For a more detailed description of this literature see Section 1.1.

2.2 Data

In order to analyze the impact of vegetation barrenness on conflict I construct a data set for 156 months from 2000-2012 for 558 grids of size 0.75° within the Sahel.

The main explanatory variable is constructed from the Net Difference Vegetation Index (NDVI), which is the most widely used index in the remote sensing literature to proxy vegetation greenness and the overall performance of the ecosystem (Herrmann, Anyamba, and Tucker, 2005). I employ NDVI data from NASAs Vegetation Index Products (NASA, 2015). The NDVI is defined as the ratio of visible light to near infrared light that is reflected by vegetation and collected by satellites.[3] The idea behind this ratio is that healthy vegetation absorbs most of the visible light that reaches it in the process of photosynthesis and emits near infrared light. In contrast, sparse vegetation reflects more visible light and less near infrared light. Accordingly, no greenness at all gives a value of 0 and values close to 1 indicate a very high density of green leaves. In this sample the NDVI ranges from 0.14 to 0.90. For ease of interpretation and a better semantic fit to the debate regarding land degradation and conflict I reverse the scale by substracting it from one. The resulting variable is from now on called "vegetation barrenness". The vegetation barrenness index also lies between 0 and 1, with 0 indicating no barrenness and 1 indicating complete barrenness.[4] Further, in this study I employ the term "land degradation" for a positive trend of vegetation barrenness, i.e. a negative trend of the NDVI.[5]

Figure 2.1 shows deciles of average barrenness over the sample period for all grids in the study area. The area spreads between 11.625° and 20.625° latitude and -16.875° and 39.375 ° longitude. Barrenness exhibits a relatively high latitudinal spatial correlation, i.e. grids further northern towards the Sahara yield more barren soil on average. The Sahel countries involve (from west to east) Mauritania (8.21% of the sample), Mali (18.21%), Niger (23.21%), Chad (13.57%) and the Republic of Sudan (24.30%). At the northern border Algeria intersects partly and to the southern border Burkina Faso, Eritrea, Senegal, Nigeria, Ethiopia intersect with less than 5% of the sample.

[3]Specifically, the NDVI is calculated by the following formula: $NDVI = (NIR - VIS)/(NIR + VIS)$ where NIR stands for near infrared light (0.7 to 1.1 μm) and VIS stands for visible light (0.4 to 0.7 μm).

[4]An increase of the NDVI within one grid may arise through the extensive margin (for instance expansion of cropland or forest) or through the intensive margin (for instance more productive crops through the use of fertilizers). For both cases it holds that an increase in the NDVI approximates the overall productive capacity within the grid.

[5]The literature provides different definitions of the term "land degradation". For instance, Stocking (2001) define land degradation as "...the temporary or permanent decline in the productive capacity of the land, and the diminution of the productive potential, including its major land uses (e.g. rainfed arable, irrigation, forests), its farming systems (e.g. smallholder subsistence), and its value as an economic resource."

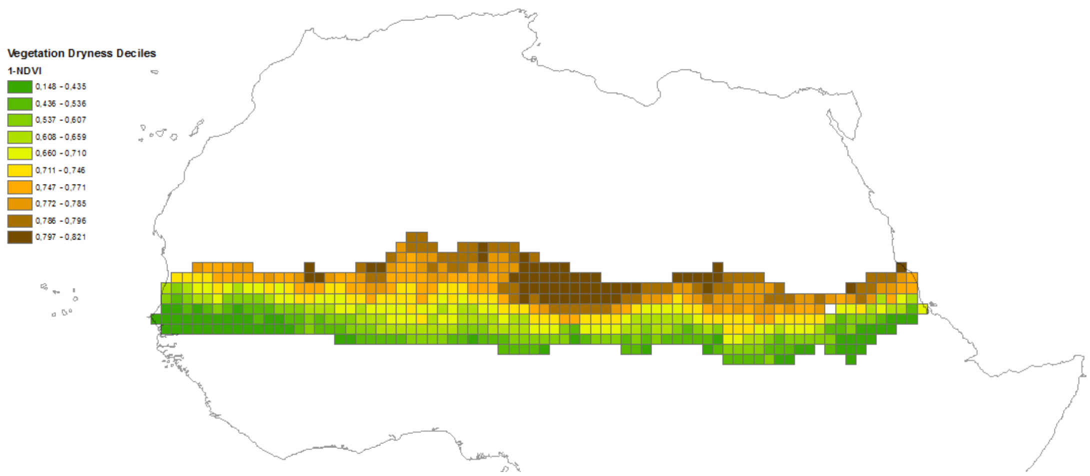


Figure 2.1: Average of Vegetation Barrenness 2000-2012
This figure plots deciles of the average vegetation barrenness index across cells over the sample period 2000-2012. The vegetation barrenness index is constructed as 1-NDVI. See text of Section 2.2 for details.

The soil properties which determine how efficiently rain is translated into fruitful vegetation vary widely across regions and depend on a range of climatic and anthropogenic impact factors. Therefore, vegetation barrenness should not be considered equal to precipitation or drought. For instance, consider a situation where precipitation turns sandy soil into mud. Whereas the vegetation barrenness index remains equal to 1 and the land is indicated to be unproductive, the implications from precipitation levels or common drought indices are the opposite. Thus, looking at (deviations from) average rainfall patterns and looking at vegetation barrenness does not necessarily catch the same phenomenon. Several recent studies in the context of civil conflicts have used the Standardized Precipitation-Evapotranspiration Index (SPEI). The SPEI reflects the difference between monthly precipitation and potential evapotranspiration (PET), where the calculation of the PET only requires monthly mean temperature as input (Vicente-Serrano, Beguería, and López-Moreno, 2010). This clarifies that vegetation barrenness does not describe the same phenomenon as common drought indices that are based on a set of climatic variables. Vegetation barrenness directly describes the productivity of land, which may be disentangled into climatic and human-induced components.

Climatic data on monthly precipitation is obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) Era-Interim data set (Dee et al., 2011). The data is based on reanalysis which implies an unchanging assimilation scheme at a regular time scale that comprises information from all available data sources. Reanalysis data is considered to be superior to weather station data but also to satellite data alone because it exploits the strengths of all different sources. For instance, in the Sahel the density of weather stations and rain gauges is rather sparse and the resulting measurement error through spatial interpolation might be correlated to unobserved factors.

Data on civil conflict is obtained from The Armed Conflict Location and Event Data Project (ACLED) (Raleigh, Linke, Hegre, and Karlsen, 2010). The dependent variable is a binary variable that indicates the incidence of a civil conflict in a given grid-month cell. The database classifies conflict incidences into different types. One major conflict type is a "battle" which constitutes an armed dispute between two groups and might involve militia groups or the government. For instance, this category might capture a battle between two groups of pastoralists over contested grazing land. I further distinguish between battles leading to a change of contested territory and those that do not. Another major type of conflict are "riots" which are usually demonstrations against the government. I create further conflict types of interest like "Agropastoral" and "Political" based on key word search in the informational notes contained in the database.[6]

Data on ethnic diversity comes from "Geo-referencing of Ethnic Groups" (GREG) project which provides a global set of polygons for ethnic territories (Weidmann, Rød, and Cederman, 2010).[7] Within the sample of this study there are 35 different ethnic groups coded and there are on average 2.2 ethnicities present in one grid. To proxy for ethnic diversity I create an indicator that switches on if there are at least two ethnicities present within the grid.

To distinguish the different components of the land cover changes (forestation, grassland- and cropland expansion) I employ satellite data on land cover. This data is obtained from the Climate Change Initiative (CCI) of the European Space Agency (ESA, 2017) and is extracted from 300mx300m resolution annual data. In 2017, the European Space Agency released their new (and in terms of frequency and disaggregation unique) annual global land cover time series from 1992 to 2015. The data provides 38 types of land cover that are classified via remote sensing techniques. I subsume the specific categories to the superordinate types "cropland", "forest" and "grassland" and calculate the share of the respective type of land within the 0.75° grid.[8] For instance, grids in the sample are on average covered

[6]In Section 1.3 we employ a similar approach to classify the respective conflict types. For the category "Agropastoral" I search for the keywords "farm", "crop", "cattle", "herd", "grazing", "nomad", "pasture", "water" within the informational notes. For the category "Political" I search for the keywords "president" and "government".

[7]The GREG data is based on ethnographic and geographic maps contained in the Atlas Narodov Mira (ANM, Bruk and Apenchenko, 1964). Much of the contained information is assembled by the Institute of Ethnography at the USSR Academy of Sciences. Further, population census data and ethnographic publications of government agencies are utilized for the construction of the maps.

[8]"Cropland" includes the original categories "rainfed cropland", "irrigated or post-flooded cropland"

to 18% by cropland, to 3% by forest and to 21% by grassland.

To measure the intensity of pastoralist agriculture I employ data for cattle density from the Gridded Livestock of the World (GLW) database (FAO, 2007). The data is based on predictions from sub-national livestock census data and predictions from environmental variables that determine the suitability for certain types of livestock holding. The data is available on a 3 arc-minute grid resolution and measured time constantly with 2005 serving as base year. To measure the density of sedentary agriculture I use the average share of cropland in each grid, generated from ESA's land cover time series.

Data on the availability of natural water resources comes from the "World Waterbodies" database (ESRI Garmin International, 2016). This data provides a map of the lakes, seas and rivers of the world and is compiled in the year 2016. As an alternative specification I employ the average share of water over the sample period within each grid cell. Annual data on the share of water is obtained from ESA's land cover time series. The employed measures do not capture devices creating additional water access such as wells or storage systems, which constitutes a relevant target for future research. Table 2.1 shows the summary statistics for the main variables.[9]

Table 2.1: Summary Statistics of Main Variables

	Mean	SD	Min	Max	Obs
Vegetation Barrenness	0.719	0.098	0	1	87,360
Precipitation (in mm)	13.447	31.598	0	486	87,360
Conflict Incidence	0.017	0.130	0	1	87,360
Battle Incidence	0.008	0.091	0	1	87,360
Riot Incidence	0.003	0.051	0	1	87,360

2.3 Vegetation Barrenness and Conflict

2.3.1 Empirical Framework

This section presents the analysis of the short-run effects of vegetation barrenness on conflict incidence. The methodology is related to the methodology as outlined in Section 1.3. Specifically, to analyze the impact of vegetation barrenness on conflict I estimate the

and "mosaic cropland". The category "forest" includes "broad leaved tree cover" (evergreen and deciduous), "needle leaved tree cover" (evergreen and deciduous), "mixed leaf type", and "mosaic forest cover". Grassland constitutes a separate category.

[9]Table 2A.2 reports summary statistics for all other variables.

following model:

$$c_{i,t} = \alpha + \beta VB_{i,t} + \gamma c_{i,t-1} + \nu_i + \nu_t + \nu_m + \nu_{m \cdot C} + \epsilon_{i,t}, \tag{2.1}$$

where $c_{i,t}$ is a binary indicator for the incidence of a civil conflict in grid i and month t and $VB_{i,t}$ reflects the vegetation barrenness index in grid i and month t as defined in the previous section. Accordingly, β is the coefficient of main interest as it represents the effect that vegetation barrenness has on conflict incidence. The model further accounts for lagged conflict incidence by the inclusion of $c_{i,t-1}$.[10] In the context of this study the lagged dependent variable gains relevance beyond its auto-regressive impact as lagged conflict incidence may be related to current vegetation outcomes. Further, the inclusion of grid fixed effects, ν_i, accounts for any grid-specific factors such as the prevalent climate or cultural and institutional traits. Common temporal impacts like global climate change or political shocks are accounted for by year fixed effects ν_t. To account for seasonality, i.e. systematic occurrences of vegetation barrenness and conflict in particular months of the year, the model further includes month of year (calendar month) fixed effects ν_m. Further, the model accounts for climate zone-specific calendar month effects, $\nu_{m \cdot C}$, which allow seasonality to differ by climate zone.[11] The standard errors $\epsilon_{i,t}$ are clustered at the grid cell level but I show results for spatial clustering or bootstrapping of standard errors in further robustness checks.

In the specification as outlined above the remaining variation comes from climate zone-specific seasonal deviations of vegetation barrenness within the grid, beyond common time effects. Thus, it abstracts from a spectrum of factors that might impact both, vegetation barrenness and civil conflict. However, I do run more extensive specifications to check on the robustness of estimates. In particular, to capture a common evolution in vegetation barrenness and conflict, I include country-, climate zone- or grid-specific year fixed effects into the regression and to account for seasonality at the finest possible spatial unit I include grid-specific calendar month fixed effects.

The identification of the coefficient β relies on the assumption that monthly vegetation barrenness in a given cell is exogenous to conflict occurrence of a conflict incidence in this month. One remaining concern may be reverse causation in the impact month. Importantly, instrumenting vegetation barrenness by precipitation levels does not only serve for decomposing land degradation into a climatic- and a human component. It also accounts

[10] The fixed effects estimator in combination with a lagged dependent variable leads to a mechanical correlation between the transformed error term and the lagged dependent variable (Nickell, 1981). However, Judson and Owen (1999) show that the asymptotic order of the resulting bias is 1/T which implies a negligible magnitude of bias for T=156.

[11] The sample region entails three climate zones as defined by the Köppen climate classification scheme: Arid, semi-arid and tropical dry.

for potential reverse causation as monthly rainfall levels are plausibly exogenous to monthly conflict occurrence.

2.3.2 Results

Baseline Results Table 2.2 presents the baseline results from the monthly panel regressions. Column (1) shows the effect of vegetation barrenness on civil conflict of any type whereas Column (2) and Column (3) show results for battles and riots, respectively.[12] The incidence of battles increases significantly in more barren months whereas the effect on overall conflict incidence or riots turns out to be insignificant. Quantitatively, if the barrenness index changes from 0 to 1, i.e. from entirely fertile- to entirely barren soil, the likelihood of battles rises by 0,01 or approximately 80% relative to the unconditional mean of 0,017.

Table 2.2: Baseline Results: Vegetation Barrenness and Conflict

	(1) All	(2) Battle	(3) Riot
Vegetation Barrenness	0.0134	0.0128**	-0.00861
	(0.00950)	(0.00579)	(0.00557)
Adjusted R^2	0.044	0.023	0.062
N	87360	87360	87360
Grid	560	560	560
Time FE	✓	✓	✓
Grid FE	✓	✓	✓
Month FE	✓	✓	✓
MonthXClimate FE	✓	✓	✓

OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Decomposition of Vegetation Barrenness In the next step, I decompose the vegetation barrenness index into a climate- and a human-induced component. The decomposition of the NDVI into its components constitutes a key tool in the ongoing debate on the causes of land degradation.[13] There are two camps in this debate: The first one considers land degradation as mainly climate-induced and therefore as a temporary phenomenon. The other one sees the causes of land degradation in human impacts such as overuse- and mismanagement of soil (Herrmann, Anyamba, and Tucker, 2005). The statistic most widely used to draw conclusions on human impacts is the residual trend of the NDVI, called RE-

[12] I subsume the category "Riots/Protests" as it is originally named in the ACLED database under the term "Riots".

[13] The literature including Herrmann, Anyamba, and Tucker (2005) refers to trends in greening.

STREND (Ibrahim, Balzter, Kaduk, and Tucker, 2015).[14] I follow the methodology as outlined in Herrmann, Anyamba, and Tucker (2005) and regress the vegetation barrenness index on a 3-month average of precipitation (current and last two months). This method reflects the first stage of an instrumental variable (IV) approach using precipitation as instrument for vegetation barrenness.[15] Accordingly, the prediction picks up the precipitation induced component of land degradation. Precipitation is by far the best climatic predictor of vegetation barrenness, therefore, the literature encourages this simple version of prediction. For completeness I discuss results including temperature into the prediction in Section 2.3.3.

Moreover, it is plausible that climatic- and human impacts interact to some extent. For instance, a human intervention to soil through the implementation of irrigation systems may affect both, the residual and the efficiency by which rain is translated into greenness. This might lead to an overestimation of the human impact, while the estimation for precipitation constitutes a lower bound. For comparability I stick to the methodology provided by the literature, although it would be conceivable in further analysis to consider interactions between the components.

To analyze the impact of the precipitation driven part of barrenness on conflict, I employ precipitation as IV for vegetation barrenness and regress civil conflict on the linear prediction from the 3-month average of precipitation. To analyze the impact of the human driven part of barrenness on conflict, I regress civil conflict on the monthly residual. To be able to single out the residual I estimate the IV stages "manually". Using the generated prediction and residual in the second stage may mislead the degrees of freedom available. Therefore I report checks on the computation method of standard errors in Section 2.3.3.[16]

Table 2.3 shows the estimated effects of both components on conflict risk. The effect of the prediction from precipitation is significantly positive for the combined conflict category and for battles. The effect of the residual on the other hand is insignificant for all conflict types. This shows that the positive short-run effect observed in the combined vegetation barrenness index is driven by the part that is predicted by rainfall. For all conflicts taken together (Table 2.2, Column (1)) the positive effect of rainfall and the negative effect of the residual seem to cancel out each other leading to an overall small and insignificant effect.

[14] Another concept is the Rain Use Efficiency (RUE), the ratio of the NDVI to precipitation. However, this concept is less feasible when comparing values across space as the average product declines with increasing precipitation and the RUE approaches infinity when precipitation is close to zero.

[15] The F-test statistic of this prediction is $F = 257.97$ with $\text{Prob} > F = 0.0000$. All control variables and fixed effects of the second stage are included to the prediction.

[16] I employ bootstrapped standard errors upon estimating the second stage which rely on computational resampling and therefore abstract from the degrees of freedom. Further, I run regressions for the prediction from precipitation with Stata's implemented 2SLS command.

The finding that a loss in the productivity of vegetation (predicted by rainfall) increases conflict is well in line with previous research that points out the impact of drought or precipitation shortages on conflict risk. Harari and Ferrara (2018) show at the grid-year level that drought unfolds a positive impact on conflict when it occurs during the growing season of a grid-specific crop. This finding implies a mechanism that works through the loss of agricultural yields. Thus, the estimated effects as presented in Table 2.3 potentially entail a mechanism as outlined above. But the effect of vegetation barrenness on conflict may also occur in contexts and times during the year other than the growing season, for instance during seasonal movement of livestock when tensions between pastoralists and sedentary farmers arise. In more extensive specifications as outlined in Section 2.3.3 I include cell-year-specific fixed effects into the regression and find a significant effect of vegetation barrenness on conflict, suggesting also other mechanisms being at work.

Table 2.3: Weather- and Human-Induced Components of Vegetation Barrenness

	(1) All	(2) Battle	(3) Riot	(4) All	(5) Battle	(6) Riot
Veg. Barrenness (Prediction)	0.0545***	0.0357***	-0.0103			
	(0.0176)	(0.0134)	(0.0112)			
Veg. Barrenness (Residual)				-0.00722	0.00267	-0.00715
				(0.0118)	(0.00672)	(0.00782)
Adjusted R^2	0.044	0.023	0.062	0.044	0.023	0.062
N	87360	87360	87360	87360	87360	87360
Grid	560	560	560	560	560	560
Time FE	✓	✓	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓	✓	✓

Column (1) through Column (3) show IV results with the 3-month (t, t-1 and t-2) average of precipitation as instrument for vegetation barrenness. Column (4) through Column (6) show results for the effect of the residual. The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

2.3.3 Robustness Analysis

Before turning to the mechanisms in more detail, this section shows that the findings remain valid for an extensive set of robustness checks. Some of the variations presented already narrow the range of plausible explanations for the observed effects, leading up to to Section 2.3.4.

To start with, data on conflict incidence as well as on precipitation levels is available for a longer time frame than the NDVI monthly time series from which the vegetation barrenness index is constructed (1997 to 2015 instead of 2000 to 2012). Therefore, estimations involving predictions from rainfall can be conducted for an extended time frame. To analyze whether the results remain valid for the period 1997-2015, I replicate the estimations of the prediction from precipitation on conflict for this period. The results indicate an even stronger effect when considering this longer time horizon.[17]

Next, I adjust the prediction of vegetation barrenness by also including temperature. The findings correspond quantitatively and qualitatively very closely to the baseline results.[18] In fact, temperature adds positively but negligibly to the prediction of vegetation barrenness. The within R-squared of the first stage regression rises by approximately 1.5 percent (from 0.72 to 0.73). Next, I predict vegetation barrenness with the 3-month average of precipitation of the previous year (t-12, t-13 and t-14). Due to the storage capacity of soils, previous seasons' rainfall still affects the current state of vegetation. At the same time previous seasons' rainfall cannot prompt current reactions related to conflict such as for instance heavy rainfall preventing a riot. This rules out that this kind of mechanism is driving the results. Of course, this approach does not account for other potential channels that evenly work via impacts of past rainfall. The results unfold that coefficients remain statistically significant and even slightly increase in magnitude.[19]

One main feature of the analysis is the use of monthly shocks instead of annual aggregates. So far, to account for seasonality the estimation includes calendar month fixed effects and climate zone-specific calendar month fixed effects. Identification thus relies on the assumption that seasonal patterns such as particular crop-, rainfall and conflict cycles are homogeneous within climate zones. To relax this assumption and to allow for the finest possible spatial heterogeneity in this context, I include grid-specific calendar month fixed effects. The results remain statistically significant and even increase in magnitude for the prediction from precipitation. In contrast, the effect of overall vegetation barrenness becomes statistically insignificant and the residual effect becomes more negative.[20]

Conflict risk is closely related to the institutional strength and the political situation of a country. At the same time, environmental policies are oftentimes implemented at the country level. An example is the engagement in the Great Green Wall Initiative (GGWI) which is based on federal decisions. To account for country-specific time-varying factors, I add country-specific year fixed effects to the regression. The positive effect of vegetation barrenness on battles remains significantly positive and is comparable in magnitude to the baseline result. Also the effect of the prediction from precipitation remains statistically significant but slightly declines in magnitude. The impact of the

[17]See Table 2A.4. This also holds for a comparison of the reduced form estimates.

[18]See Table 2A.3.

[19]I employ the longer time period in this exercise as this analysis is exclusively performed for the prediction. See Table 2A.5.

[20]See Table 2A.6, Column (4).

residual remains insignificant.[21]

It might be even more plausible that common trends are attached to the climate zone, especially in the case of the prediction from precipitation. To capture climate zone-specific time-varying factors, I include climate zone-specific year fixed effect. The baseline effect and the effect of the prediction from precipitation stay significantly positive and slightly increase in magnitude. The impact of the residual remains insignificant.[22] Overall, the findings hint to some common evolution of vegetation barrenness and conflict within countries and within climate zones.

The common evolution in vegetation barrenness and conflict risk may be attached to even finer spatial levels. For instance, different crop types may deliver different reflection levels and at the same time relate to differential conflict risk. If bio-geographic factors (that are constant over time) determine the crop type, the inclusion of grid fixed effects accounts for this concern. It might however be conceivable that there is a time-varying component, i.e. a common cause (e.g. migration) that alters the main crop type and conflict risk at the same time. Note that for a systematic bias the direction of change in reflection levels would systematically have to be the same as the direction of change in conflict levels which seems to be unlikely. Nevertheless, I include grid-specific year fixed effects to check for robustness. Given this very restrictive specification involving a vast number of fixed effects which absorb much of the variation, particularly the coefficients from the prediction from precipitation decline and standard errors increase. Notwithstanding, the baseline results and the effect of the prediction from precipitation (long panel) remain significantly positive.[23]

It might be the case that conflict itself affects the state of vegetation. For instance if the vegetation deteriorates by enduring battles. This concern is however only relevant for the baseline estimation and the residual, not for the prediction from precipitation. I replicate the main results of these specifications and include lagged instead of current levels of vegetation barrenness, controlling for lagged conflict at the same time. The effect of the lagged variables turns out to be insignificant.[24] However, it remains unclear whether this is due to a bias in the baseline result or because the environmental shock unfolds an immediate impact which is not visible in the following month.

Next I turn to a robustness analysis regarding the estimation of standard errors. One concern is the spatial correlation of variables between grids which might lead to an underestimation of standard errors. I include Conley robust standard errors that account for cross-sectional spatial correlation and also location-specific serial correlation in the esti-

[21]See Table 2A.6, Column (2).
[22]See Table 2A.6, Column (3).
[23]See Table 2A.6, Column (5).
[24]See Table 2A.7.

mation of standard errors (Conley, 1999, 2008).[25] The results indicate that the standard errors do indeed increase slightly. However, the effect of the prediction from precipitation remains statistically significant. In contrast, the baseline effect falls slightly below conventional significance levels and the effect from the residual remains insignificant.[26] Further, I report results for estimations employing bootstrapped standard errors and regressions applying Stata's implemented 2SLS command to account for potential generated regressor issues that might arise from the manually conducted instrumental variable calculation. It shows that standard errors only vary slightly in these cases and the significance levels are not affected.[27]

In the following, I explore the sensitivity of the results with respect to alternative measures and the use of a non-linear estimator. As the effect of the residual has overall turned out to be insignificant, I restrict this analysis to estimates involving the prediction from precipitation. First I analyze the role of vegetation barrenness as trigger civil conflict *onset*. In this specification only the starting month of a conflict episode is coded as 1 whereas all other months are coded as 0. The results indicate a significantly positive impact on conflict onset, although the effect slightly declines compared as to the effect on conflict incidence.[28] This implies that the overall impact is to some extent, but not exclusively, driven by prolonging ongoing conflicts. Further, I explore the intensive margin of conflict incidence. The results reveal mixed evidence. The number of conflicts in a grid cell in a given month turns out to be significantly affected by increasing vegetation barrenness, while the number of related fatalities as another facet of the intensity of conflict does not.[29] I close the analysis by presenting results from the estimation of a Logit model. The results confirm the baseline results.[30]

2.3.4 Vulnerability towards Vegetation Barrenness

In this section I am going to explore regional and temporal factors that shape the vulnerability towards the environmental shocks. For this analysis, I focus on the part of barrenness that is predicted by precipitation as this has turned out to be of major relevance, whereas the residual remained without significant consequence for conflict risk.

The Sahel has a long history of herder-farmer conflicts or conflicts between different groups of herders. These clashes have become more frequent in recent years (Moritz,

[25] I allow for a spatial correlation that decreases up to a cutoff of 200km and thus includes two neighbouring grid cells in each direction. Serial correlation is accounted for across 20 time periods.
[26] See Table 2A.8.
[27] See Table 2A.9.
[28] See Table 2A.10.
[29] See Table 2A.10.
[30] see Table 2A.10.

2010). One reason might be that herds are more often owned by urban investors with insufficient knowledge on cultural conventions that traditionally organize property rights (Oyama, 2014). Further, tensions might have intensified due to policies that favor settled agriculture and expanding croplands (Oyama, 2014). At the core of these conflicts lies the competition for resources. There are several case-studies documenting the prevalance of land-use conflicts and studies that revise the causes for these conflicts in specific regional contexts (e.g. Benjaminsen et al., 2012 or Benjaminsen and Ba, 2009). Conflicts may for instance involve battles over grazing land and livestock or disputes on crop damage by passing cattle in the course of seasonal herder movements. To explore the hypothesis that agricultural resources are indeed involved in the emergence of conflict I first look at the conflict type in more detail. I create the category "Agropastoral" that is based on key word search in the informational notes for each incidence contained in the ACLED database. The key words are extracted from articles on land degradation and conflict, particularly Oyama (2014). A resulting note is for example:

"Farmers and herders fought each other. Farmers had seized cows grazing on their lands. 6 people died in the clashes."

Further, I create a "placebo" category for political conflicts. Beyond, within the category of battles I look at battles that lead to a change in the contested territory and battles that do not lead to such a change. Panel A of Table 2.4 shows the estimation results for the different types of conflict. It shows that the effect is most strongly pronounced for agropastoral conflicts and battles involving changes of territory. The other categories, which are mainly unrelated to agricultural production, do not react significantly to the vegetation barrenness shock.

To further substantiate the argument that competition for agricultural resources relates the barrenness shock to conflict I split the sample by the intensity of agricultural production. Agricultural production may involve sedentary crop farming or pastoralism. Thus, I look at the intensity of both categories separately and combined. The results are shown in Panel B of Table 2.4. I proxy the intensity of pastoralism by the density of cattle and proxy the intensity of farming with the share of cropland within the grid. Then I split the sample into grids with high intensity of farming and pastoralism and those with low intensity respectively.[31]

Further, I analyze whether the presence of large water resources moderates the impact

[31]The correlation between the two types of agriculture is very high within the sample (0.55). Therefore there is only a very small number of grids that fall exclusively into the low/high category of one of the two types. Table 2A.11 shows results for combined categories; i.e. both high, both low, one high and one low. The effect is strongly pronounced and statistically significant where densities of both types of agriculture are high.

of monthly vegetation barrenness. I split the sample into grids containing freshwater resources (like a river or a lake) and grids without freshwater resources. I also take into account information on the share of water within the grid, obtained from satellite land cover data, and split the sample at the median of this measure. The results in Panel C of Table 2.4 unfold that the effect stems from regions with insufficient access to water resources.

Table 2.4: Vulnerability towards Vegetation Barrenness:
Conflict Types, Land Use and Water Resources

	Dependent Variable: Conflict Incidence			
	Panel A: Conflict Types			
	Battle (terr)	Battle (non-terr)	Agropastoral	Political
Veg. Barrenness (Prediction)	0.0404***	0.00333	0.0150***	0.00890
	(0.0155)	(0.00468)	(0.00541)	(0.00690)
Adjusted R^2	0.022	-0.000	0.002	0.005
N	87360	87360	87360	87360
Grid	560	560	560	560
	Panel B: Livestock- and Farmland Density			
	Cattle Dense	Cattle Sparse	Agr. Dense	Agr. Sparse
Veg. Barrenness (Prediction)	0.0459**	-0.0499	0.0517**	0.00848
	(0.0196)	(0.0425)	(0.0200)	(0.0493)
Adjusted R^2	0.047	0.042	0.044	0.050
N	43524	43524	41832	45216
Grid	279	279	269	290
	Panel C: Water Scarcity			
	Source	No Source	Water>Med.	Water<Med.
Veg. Barrenness (Prediction)	0.0362	0.0632***	0.0257	0.0649***
	(0.0312)	(0.0184)	(0.0223)	(0.0242)
Adjusted R^2	0.049	0.044	0.041	0.053
N	25116	61932	30444	56604
Grid	161	397	196	363
Grid FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓

IV results using 3-month (t, t-1 and t-2) average precipitation as instrument for vegetation barrenness. In Panel B, the sample is split into grids with sparse and dense presence of cattle- and farmland is based on the median value of cattle density (measured in 2005) in Column (1) and (2) and the average value of the share of cropland over the sample period 2000-2012 in Columns (3) and (4). The dependent variable is the incidence of a conflict event in a given cell and month. In Panel C the sample is split into grids with access to a major natural water source in Column (1) and no such source in Column (2) and with a share of water above the median in Column (3) and below the median in Column (4). The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Conflicts in the Sahel oftentimes take place along ethnic lines that traditionally also determine which type of agriculture is practiced. Therefore ethnic diversity in this context may have two implications. First, differential and potentially competing systems of agriculture being in place. Second, the existence of boundaries along which property rights are contested. To examine whether the effect of barrenness is amplified in ethnically diverse regions I split the sample into grids that are ethnically diverse, i.e. in grids where more than one ethnicity is present, and grids where only one ethnicity is present. The results are shown in Table 2.5. I find that the effect in ethnically divided regions is twice as large and statistically highly significant in contrast to grids with only one ethnicity.[32]

Table 2.5: Vulnerability towards Vegetation Barrennness: Ethnic Diversity

	(1) Ethnic Diversity	(2) No Ethnic Diversity
Dep. var.: incidence		
Veg. Barrenness (Prediction)	0.0722***	0.0324
	(0.0228)	(0.0246)
Adjusted R^2	0.051	0.032
N	49452	37908
Grid	317	243
Time FE	✓	✓
Grid FE	✓	✓
Month FE	✓	✓
MonthXClimate FE	✓	✓

IV results using 3-month (t, t-1 and t-2) average precipitation as instrument for vegetation barrenness. The dependent variable is the incidence of a conflict event in a given cell and month. The sample split is based on the number of ethnic homelands being located present within the grid. At least two ethnicities within a grid constitute "Ethnic Diversity". Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

During the rainy season farmers grow and harvest their crops. The land is exclusively used by them and any invasion is considered as offense. In contrast, in the dry season the land may even be open to public use. This exemplifies that benefits are merely competed for in the rainy season. Figure 2.2 shows the distribution of barrenness across months. Table 2.6, Column (1) and Column (2), shows estimation results for the rain and the dry season separately. It shows that in the rainy season increasing barrenness raises conflict risk significantly. In contrast, in the dry season the effect turns out to be statistically insignificant.

Seasonal movement of pastoralists optimizes livestock grazing over the year. Pastoralists leave their homes at the end of the rainy season and come back at the start of the rainy season (Thebaud, 2017). In the transition months there is an increased risk of conflict emergence as passing herds potentially feed on or destroy croplands, especially in the

[32]The results are replicated in table 2A.12 using alternative data on ethnicity from Murdock (1967).

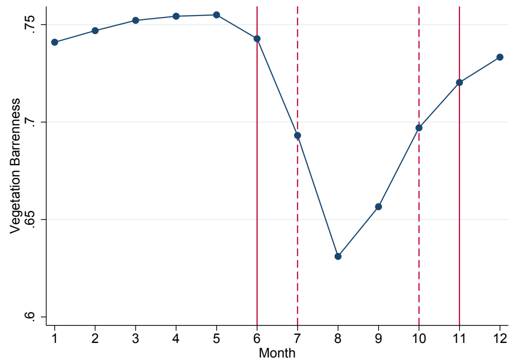


Figure 2.2: Vegetation Barrenness by Month

The figure shows average vegetation barrenness by calendar month. The rainy season is framed in solid red lines. Transition months are defined as the first and the last month of the rainy season and are thus graphically located between the solid and dashed red lines.

face of an increasing number of transhumance corridors being closed for other usage. I additionally split the sample into transition months and other (dispersion) months. The transition months are defined as the first and the last month of the rainy season as depicted in Figure 2.2. The estimation results are presented in Table 2.6, Column (3) and Column (4). The effect appears to be relevant in both sub-periods. However, the effect of barrenness during the transition months is larger than in dispersion months and even larger than considering the whole rainy season. This unfolds an interesting non-linearity with respect to the amount of rain as the transition months are located at the edges of the rainy season where the amount of rain is relatively low.

The results presented in this section set a frame for potential mechanisms explaining the observed effects. Given that vegetation barrenness affects conflict risk most severely where societies rely on the productivity of soil, carry out sedentary- or pastoralist agriculture and have insufficient access to other water resources when rain stays out, points towards rising scarcity being at the core of dispute. Thinking about how this rising scarcity translates into conflict, the results suggest that competition for available resources is involved, potentially between groups defined by ethnicity or the type of agriculture.

2.4 Long-Run Trends: Land Degradation and Conflict

In this section I turn towards the long-run implications of the documented short-run results. In particular, I address the question whether long-term changes in vegetation barrenness are related to long-term changes in conflict risk. Figure 2.3 depicts the monthly evolution of the vegetation barrenness index over the estimation period 2000-2012. Contrary to common

Table 2.6: Vulnerability towards Vegetation Barrenness: Seasons

	(1) Rain	(2) Dry	(3) Transition	(4) Dispersion
Dep. var.: incidence				
Veg. Barrenness (Prediction)	0.0572**	-0.109	0.219**	0.0449**
	(0.0274)	(0.156)	(0.0876)	(0.0178)
Adjusted R^2	0.050	0.038	0.061	0.043
N	43680	43680	14560	72800
Grid	560	560	560	560
Time FE	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓

IV results using 3-month (t, t-1 and t-2) average precipitation as instrument for vegetation barrenness. According to seasons as depicted in Figure 2.2 estimation in Column (1) includes calendar months 6, 7, 8, 9, and 10, in Column (2) calendar months 1, 2, 3, 4, 5, 11, and 12, in Column (3) calendar months 6 and 10 and in Column (4) 1, 2, 3, 4, 5, 7, 8, 9, 11 and 12, respectively. The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

perceptions of an ongoing desertification in the Sahel, average vegetation barrenness slightly declines over the sample period.[33] This observation is in line with recent research that refutes an the prevalence of an irreversible and ongoing land degradation and rather finds a "regreening of the Sahel" in terms of an overall positive trend of the NDVI (e.g. Tong et al., 2017, , Hermann et al., 2005). Hermann et al. (2005) further discuss the spatial heterogeneity in the extent of land degradation in the Sahel. The authors show that there is a strong heterogeneity in trends within the Sahel, with some regions not experiencing any considerable changes and some regions being exposed to strong positive or negative trends in the NDVI. In essence, I exploit this spatial heterogeneity to estimate the effect of land degradation on conflict risk in the long run.

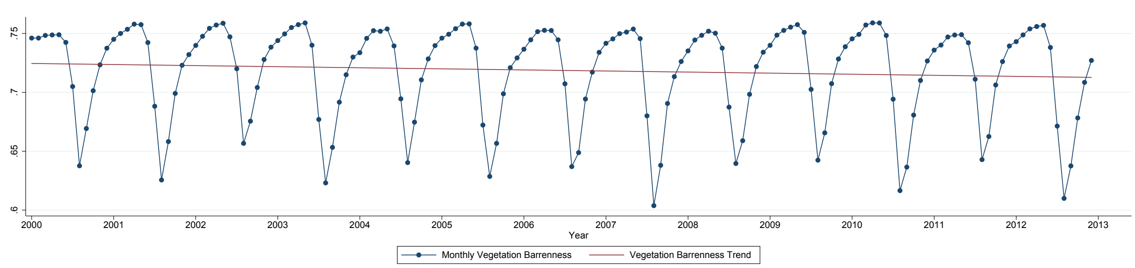


Figure 2.3: Evolution of Monthly Vegetation Barrenness 2000-2012

[33]This is mainly driven by a negative trend in the residual. See Figure 2A.1 for a graph depicting the trend of the residual and the prediction from precipitation separately.

2.4.1 Empirical Framework

The methodology of this section closely follows the strategy as outlined in Section 1.4. It is based on the logic of a difference-in-differences (DiD) approach in long differences.[34] The idea of this approach is to regress grid-specific differences in average conflict incidence between the early and the late part of the panel on the respective differences in vegetation barrenness. Specifically, I compute averages of conflict incidence in a cell for the first three years of the panel (2000-2002) as denoted by $\overline{C}_{i,1}$ and in the last three years of the panel (2010-2012) as denoted by $\overline{C}_{i,2}$.[35] The long difference in conflict incidence in cell i is then given by $\overline{DC}_i = \overline{C}_{i,2} - \overline{C}_{i,1}$. The same procedure is applied to the vegetation barrenness index which yields the long difference of average vegetation barrenness between the last years of the panel and the first years of the panel ($\overline{DVB}_i = \overline{VB}_{i,2} - \overline{VB}_{i,1}$). Then I estimate the following model:

$$\overline{DC}_i = \alpha + \mu \overline{DVB}_i + \phi_r + \epsilon_i \; , \tag{2.2}$$

where ϕ_r controls for climate zone-specific time trends and captures a common evolution of land degradation and conflict within climate zones. ϵ_i constitutes the error term. The coefficient of interest, μ, captures the effect that changes in vegetation barrenness yield on changes in conflict incidence and relies on the common trend assumption across cells within the given climate zones.

2.4.2 Results

Table 2.7 shows the long-run results for different types of conflicts. The results indicate that cells experiencing a stronger increase in vegetation barrenness also experience a stronger increase in the incidence of conflicts related to agriculture or pastoralism. The coefficients estimating the impact on the other conflict types turn out to be statistically insignificant. In particular, the result in Column (5) indicates that an increase in the extent of vegetation barrenness in long differences by 1, i.e. an increase from zero to full vegetation barrenness,

[34]Given the availability of NDVI data I have only access to a shorter time period, so I refrain from a generalization of the difference-in-differences (DiD) to a generalized difference-in-differences (GDD) framework as outlined in Section 1.4.

[35]I vary this time window in further robustness checks. See Table 2A.15 for an alternative specification using a 4-year window at the beginning and end of the sample period. However, when expanding the windows used for the long-difference calculation by more and more years, the standard errors increase as the available variation in differences declines. When employing the full halfs of the sample even the effect on agropastoral conflicts looses its statistical significance (see Table 2A.16). A longer time horizon would be helpful to detect the medium- or long-run consequences of land degradation on conflict.

leads to an increase in the frequency of conflicts by 0.063. This corresponds to a threefold of the unconditional mean of the increment over time in conflict which is 0.0165.

Table 2.7: Land Degradation and Conflict: Long Differences

	(1) All	(2) Riot	(3) Battle (terr)	(4) Battle (no terr)	(5) Agropastoral
Diff Veg. Barrenness	0.0859	0.108	0.0240	0.00664	0.0637**
	(0.268)	(0.0753)	(0.165)	(0.0204)	(0.0263)
Adjusted R^2	0.041	0.016	0.028	0.005	0.016
N	560	560	560	560	560
Climate Zone Trend	✓	✓	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 2000-2002 and 2010-2012, $\overline{DC_i}$, with one observation per cell. Robust standard errors in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Next, I consider the effect of long differences in the prediction from precipitation and long differences in the residual separately.[36] The results reveal that the overall impact of land degradation on conflict, corresponding to Column (1) of Table 2.7, turns out to be insignificant as it consists of a positive impact of predicted vegetation barrenness and a negative impact of the residual. However, when including temperature into the prediction, the negative effect of the residual becomes statistically insignificant.[37] Thus, in contrast to land degradation determined by climatic variables, the "human component" as defined via the RESTREND methodology does not seem to unfold a significant impact on conflict risk in the long run.

Besides the impact of precipitation, the inter-annual variation in vegetation dynamics is mainly determined by land cover changes. So, to shed more light on a potential relation between land degradation and conflict in the long run, I analyze the effect of changes in land cover on changes in conflict risk.[38] Among the different land cover types as described in Section 2.2, the share of cropland yields the strongest negative correlation with vegetation barrenness (-0.727), followed by forest cover (-0.279) and the share of grassland (-0.185). Based on the methodology as outlined above, I calculate long differences for the respective land cover types. Then I regress the changes in conflicts on land cover changes.[39] The results unfold that an increasing share of cropland seems to raise conflict risk in the long run, which is in line with abundant narratives on disputes arising from closed transhumance

[36]See Table 2A.13.

[37]See Table 2A.14

[38]Begue et al. (2011) study the case of Mali and observe a strong positive trend of the NDVI for the period 1982-2006 and that cropping contributed to the greening. Several other studies document an increase in cropland in the Sahel in recent years and sketch a link to increasing greenness values (e.g. Nutini et al., 2013).

[39]See Table 2A.17.

corridors by cropland expansions or policies favoring sedentary over pastoralist agriculture in the region. However, the positive impact of cropland is significant only at the 10 percent level and should therefore be considered with some caution. However, this finding points towards an underestimation of the effect of the human-induced component, as cropland expansion has a positive impact on conflict risk and a negative impact on the vegetation barrenness index. When including the difference of the residual and the difference of the share of cropland jointly into the regression the effect of the residual change indeed increases in magnitude and looses its statistical significance.[40]

2.5 Discussion

Many studies have tackled the question of whether climate change impacts the risk of civil conflict. One phenomenon that is oftentimes related to changing climate is land degradation. In the Sahel between 2000 and 2012 there is strong heterogeneity in the extent of land degradation, although on average the region seems to have become slightly greener. In this study I analyze the impact of monthly vegetation barrenness shocks and medium-term land degradation on conflict risk on a 0.75° resolution grid in a month-to-month and a long-difference framework between 2000 and 2012.

The findings indicate that in the short-run vegetation barrenness has a significantly positive impact on conflict. Employing monthly precipitation levels as instrument for vegetation barrenness unfolds that the short-run effect is merely driven by monthly rainfall variability. The results further point towards stronger effects in regions where societies strongly rely on the capacity of soil. This is the case for regions with high densities of settled agriculture or pastoralism or those with scarce freshwater resources. Also the effects are more pronounced in regions that are ethnically fractionalized which may imply competing systems of agriculture being in place or the existence boundaries along which property rights are contested. Further, the effects unfold in particular for types of conflict that may plausibly arise through tightened competition for land resources - involving conflicts related to agriculture or pastoralism or battles inducing territorial changes. Looking at the heterogeneity across seasons reveals that conflict risk responds most strongly in the rainy season and in months of seasonal pastoralist movement. In the long-run analysis I find suggestive evidence that changes in barrenness increase the risk of conflicts related to agriculture or pastoralism and that cropland expansions are positively related to conflict risk. The analysis of impacts of the "human-induced component" of land degradation, defined as the residual after the prediction from precipitation, does not deliver robust results in the short- and in the long run.

[40]See Table 2A.17.

This study sets different starting points to think about policy implications. Anthropogenic global warming and its related impacts on local rainfall patterns are to a large extent a consequence of global human activity such as aggregate CO2 emissions. Beyond the difficulty to achieve and enforce global policy, the scientific community is still clarifying the relationships and spatial dynamics of climatic and environmental variables. Taking environmental shocks and changes as matter of fact suggests a reduction of societal vulnerability and the implementation of locally targeted adaption mechanisms. The human impact on land degradation unfolds locally and in shorter time horizons which yields a more tangible target for policy implications. But further research is needed to address the interactions between climatic and human impacts in shaping the risk for conflict in more detail. Moreover, the analysis of environmental dynamics over longer time horizons could add valuable insights complementing the findings of the present study.

2.A Appendix

Table 2A.1: Data Sources and Variable Construction

	Source	*Variable*	*Specification*
		Net Difference Vegetation Index (NDVI)	
1.	National Aeronautics and Space Administration (NASA, 2015), Vegetation Index Products	Vegetation Barrenness	Monthly index [0,1], calculated as 1-NDVI.
		Land Cover	
2.	European Space Agency (ESA, 2017), CCI Land Cover Products	Share of Cropland Share of Grassland Share of Forest Share of Water	Annual share of Land cover type within grid calculated from 300m×300m raster.
		Cell-specific Characteristics	
2.	Geo-referencing of Ethnic Groups (GREG) (Weidmann, Rød, and Cederman, 2010)	Ethnic Diversity	Binary indicator, equals 1 if there is more than 1 ethnic territory within the grid, 0 otherwise.
	World Waterbodies (ESRI Garmin International, 2016) Share of Water (ESA, 2017)	Water Resources	Binary indicator, equals 1 if water source available, 0 otherwise. Indicator created from waterbody polygons.
3.	Gridded Livestock of the World (GLW) (FAO, 2007)	Pastoralism	Estimated number of cattle per sqkm.
		Conflict	
2.	ACLED (Raleigh, Linke, Hegre, and Karlsen, 2010)	Incidence	Monthly binary indicator for conflict of any type.
		Battle (territory)	Monthly binary indicator for Battle-Non-state actor overtakes territory Battle-Government regains territory.
		Battle (non-territory)	Monthly binary indicator for battle-No change of territory.
		Riot and Protest	Monthly binary indicator for riots/protests.
		Agropastoral	Monthly binary indicator for conflict based on keyword match in contextual notes: “farm”, “crop”, “cattle”, “herd”, “grazing”, “nomad”, “pasture”, “water”.
		Political	Monthly binary indicator for conflict based on keyword match in contextual notes: “government”, “president”.

Table 2A.2: Summary Statistics of Additional Variables

	Mean	SD	Min	Max	Obs
Battle (Territorial Change)	0.008	0.090	0.000	1.000	87,360
Battle (no Territorial Change)	0.001	0.024	0.000	1.000	87,360
Agropastoral Conflict	0.001	0.029	0.000	1.000	87,360
Political Conflict	0.002	0.049	0.000	1.000	87,360
Ethnic Diversity (GREG)	0.566	0.496	0.000	1.000	87,360
Cattle Density (Number per sqkm)	0.000	0.015	0.000	1.000	87,373
Share of Cropland	0.183	0.271	0.000	0.986	87,373
Share of Grassland	0.207	0.249	0.000	0.925	87,373
Share of Forest	0.024	0.076	0.000	0.767	87,373
Share of Water	0.004	0.025	0.000	0.276	87,373
Water Source	0.291	0.454	0.000	1.000	87,373

Table 2A.3: Prediction accounting for Temperature

	(1) All	(2) Battle	(3) Riot	(4) All	(5) Battle	(6) Riot
Veg. Barrenness (Prediction)	0.0573***	0.0344**	-0.00835			
	(0.0175)	(0.0135)	(0.0101)			
Residuals				-0.0106	0.00230	-0.00823
				(0.0121)	(0.00715)	(0.00775)
Adjusted R^2	0.044	0.023	0.062	0.044	0.023	0.062
N	87360	87360	87360	87360	87360	87360
Grid	560	560	560	560	560	560
Time FE	✓	✓	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓	✓	✓

Column (1) through Column (3) show IV results with the 3-month (t, t-1 and t-2) average of precipitation and temperature as instruments for vegetation barrenness. Column (4) through Column (6) show results for the effect of the residual. The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.4: Extended Time Horizon

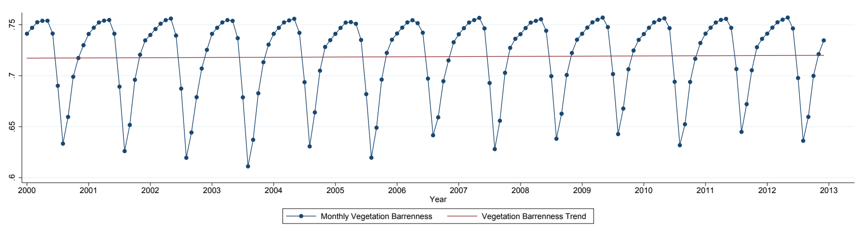
	(1) 2000-2012	(2) 1997-2015
Veg. Barrenness (Prediction)	0.0545***	0.0861***
	(0.0176)	(0.0203)
Adjusted R^2	0.044	0.092
N	87360	127120
Grid	560	560
Time FE	✓	✓
Grid FE	✓	✓
Month FE	✓	✓
MonthXClimate FE	✓	✓

IV results using the 3-month (t, t-1 and t-2) average of precipitation as instrument for vegetation barrenness. The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

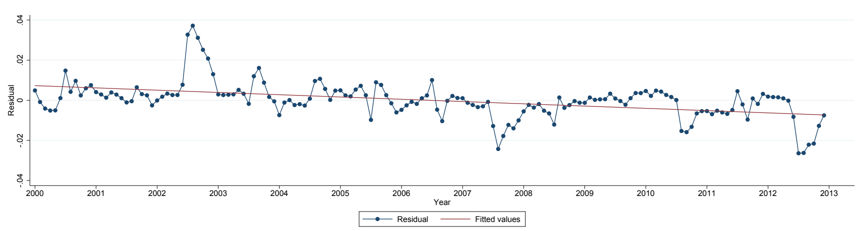
Table 2A.5: Prediction from past Precipitation (t-12) to (t-14)

	(1) All	(2) Battle	(3) Riot
Veg. Barrenness (Prediction)	0.0775**	0.0722***	-0.00977
	(0.0347)	(0.0260)	(0.0154)
Adjusted R^2	0.021	0.006	0.004
N	127120	127120	127120
Grid	560	560	560
Time FE	✓	✓	✓
Grid FE	✓	✓	✓
Month FE	✓	✓	✓
MonthXClimate FE	✓	✓	✓

IV results using 3-month average of precipitation in the previous year (t-12, t-13 and t-14) as instrument for vegetation barrenness. The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.



(a) Monthly Barrenness and Prediction from Rainfall



(b) Monthly Barrenness Residuals

Figure 2A.1: Evolution of Weather- and Human-Induced Components of Vegetation Barrenness 2000-2012

This figure shows the monthly evolution of the predicted vegetation barrenness by precipitation in Panel (a), which excludes year fixed effects for visualization. Panel (b) shows the monthly evolution of the residual obtained from the prediction, i.e. the "human component" of land degradation.

Table 2A.6: Alternative Specifications

FE Model	Baseline (1)	Country-Year (2)	Climate-Year (3)	Grid-Month (4)	Grid-Year (5)
	Panel A: Baseline - Battles				
Vegetation Barrenness	0.0128**	0.0138**	0.0146**	-0.00222	0.0156**
	(0.00579)	(0.00585)	(0.00591)	(0.0153)	(0.00652)
Adjusted R^2	0.132	0.135	0.134	0.123	0.193
N	87360	87360	87360	87360	87360
Grid	560	560	560	560	560
	Panel B: Prediction from Rainfall - Incidence (1997-2015)				
Veg. Barrenness (Prediction)	0.0861***	0.0529***	0.0861***	0.140***	0.0427**
	(0.0203)	(0.0190)	(0.0205)	(0.0367)	(0.0205)
Adjusted R^2	0.238	0.247	0.243	0.231	0.349
N	127120	127120	127120	127120	127120
Grid	560	560	560	560	560
	Panel C: Prediction from Rainfall - Incidence (2000-2012)				
Veg. Barrenness (Prediction)	0.0545***	0.0375**	0.0660*	0.0660*	0.0394*
	(0.0176)	(0.0184)	(0.0365)	(0.0365)	(0.0211)
Adjusted R^2	0.208	0.213	0.196	0.196	0.283
N	87360	87360	87360	87360	87360
Grid	560	560	560	560	560
	Panel D: Residual - Incidence				
Veg. Barrenness (Residual)	-0.00722	0.00247	-0.00722	-0.0479**	0.00737
	(0.0118)	(0.0116)	(0.0118)	(0.0227)	(0.0131)
Adjusted R^2	0.208	0.213	0.203	0.196	0.283
N	87360	87360	87360	87360	87360
Grid	560	560	560	560	560
Grid FE	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Month×Equator FE	✓	✓	✓	✓	✓

Column (2) adds country-specific year fixed effects to the regression, Column (3) climate zone-specific year fixed effects, Column (4) grid-specific (calendar) month fixed effects and Column (5) grid-specific year fixed effects. Panel A: OLS (linear probability model) fixed effects estimation results. The dependent variable is the incidence of a battle event in a given cell and month. Panel B and C: IV results using the 3-month (t, t-1 and t-2) average of precipitation as instrument for vegetation barrenness. The dependent variable is the incidence of any conflict in a given cell and month. Panel D: Results for the effect of the residual. The dependent variable is the incidence of any conflict in a given cell and month. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.7: Effect of Lagged Vegetation Barrenness

	(1) All	(2) Battle	(3) Riot	(4) All	(5) Battle	(6) Riot
L.Vegetation Barrenness	0.0134	0.00893	-0.00910			
	(0.0108)	(0.00672)	(0.00606)			
L.Incidence	0.200***			0.200***		
	(0.0190)			(0.0190)		
L.Battle		0.145***			0.145***	
		(0.0195)			(0.0195)	
L.Riot			0.246***			0.246***
			(0.0397)			(0.0397)
L.Veg. Barrenness (Residual)				-0.0000928	0.00594	-0.0129
				(0.0135)	(0.00917)	(0.00845)
Adjusted R^2	0.046	0.023	0.061	0.046	0.023	0.061
N	87360	87360	87360	87360	87360	87360
Grid	560	560	560	560	560	560
Time FE	✓	✓	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓	✓	✓

Results of OLS (linear probability model) estimates. The dependent variable is the incidence of a conflict event in a given cell and month. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.8: Conley Standard Errors

	(1) Baseline	(2) Conley SE	(3) Baseline	(4) Conley SE	(5) Baseline	(6) Conley SE
Vegetation Barrenness	0.0128**	0.0128*				
	(0.00579)	(0.00737)				
Veg. Barrenness (Prediction)			0.0545***	0.0545**		
			(0.0176)	(0.0246)		
Veg. Barrenness (Residual)					-0.00722	-0.00722
					(0.0118)	(0.0149)
Adjusted R^2	0.023	0.021	0.044	0.039	0.044	0.039
N	87360	87360	87360	87360	87360	87360
Grid	560	.	560	.	560	.
Time FE	✓	✓	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓	✓	✓

Columns (1), (3) and (5) employ standard errors clustered at the grid level. Columns (2), (4) and (6) show the respective results employing Conley robust standard errors. Spatial correlation is accounted for up to a cutoff of 200km. Serial correlation is accounted for across 20 time periods. The dependent variable is the incidence of a battle in Columns (1) and (2) and the incidence of any conflict in Columns (3) through Column (6). ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.9: 2SLS and Bootstrapped Standard Errors

	(1) BS	(2) HW	(3) BS (2SLS)	(4) HW (2SLS)	(5) BS Res	(6) HW Res
Veg. Barrenness (Prediction)	0.0545***	0.0545***				
	(0.0204)	(0.0176)				
Veg. Barrenness (2SLS)			0.0545***	0.0545***		
			(0.0186)	(0.0178)		
Veg. Barrenness (Residual)					-0.00722	-0.00722
					(0.0119)	(0.0118)
Adjusted R^2	0.038	0.044			0.038	0.044
N	87360	87360	87360	87360	87360	87360
Grid	560	560	560	560	560	560
Time FE	✓	✓	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓	✓	✓

"HW" indicates Huber-White standard errors clustered at the grid level, as employed in the baseline regressions. "BS" indicates bootstrapped standard errors. In Columns (1) and (2) the instrumental variable approach is conducted manually, Columns (3) and (4) employ Stata's implemented 2SLS command. Columns (5) and (6) document results for the residual. The dependent variable is the incidence of a conflict event in a given cell and month. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.10: Conflict Onset, Intensive Margin and Logit

	(1) Conflict Onset	(2) Number Conflicts	(3) Number Fatalities	(4) Logit
Veg. Barrenness (Prediction)	0.0419**	0.121**	0.0906	0.222***
	(0.0173)	(0.0584)	(3.151)	(0.0523)
Adjusted R^2	0.017	0.059	0.007	
N	87360	87360	87360	35412
Grid	560	560	560	.
Time FE	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓

IV results with the 3-month (t, t-1 and t-2) average of precipitation as instrument for vegetation barrenness in Column (1), (2) and (3), and marginal effects of Logit estimates in Column (4). In Column (1) the dependent variable is the onset of a conflict in cell in a given month. In Column (2) the dependent variable is the number of conflicts within a grid in a given month. In Column (3) the dependent variable is the number of fatalities arising from conflict in a given cell and month. In Column (4) the dependent variable is conflict incidence. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.11: Cattle- and Cropland Density Matrix

	(1) cuH caH	(2) cuL caL	(3) cuL caH	(4) cuH caL
Dep. var.: incidence civil conflict				
Veg. Barrenness (Prediction)	0.0625***	-0.0526	-0.0983	0.0922
	(0.0223)	(0.0509)	(0.168)	(0.141)
Adjusted R^2	0.046	0.041	0.068	0.053
N	39852	39900	3828	3780
Grid	257	256	25	26
Time FE	✓	✓	✓	✓
Grid FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
MonthXClimate FE	✓	✓	✓	✓

IV results using 3-month (t, t-1 and t-2) average precipitation as instrument for vegetation barrenness. The dependent variable is the incidence of a conflict event in a given cell and month. "cuH caH" indicates high density of cultivation and cattle, "cuL caL" low density of cultivation and cattle, "cuL caH" low density of cultivation and high density of cattle and "cuH caL" high density of cultivation and low density of cattle. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.12: Alternative Ethnicity Measure

	(1) Ethnic Diversity	(2) No Ethnic Diversity
Dep. var.: incidence		
Veg. Barrenness (Prediction)	0.0568***	0.0269
	(0.0200)	(0.0391)
Adjusted R^2	0.052	0.021
N	56940	30420
Grid	365	195
Time FE	✓	✓
Grid FE	✓	✓
Month FE	✓	✓
MonthXClimate FE	✓	✓

IV results using 3-month (t, t-1 and t-2) average precipitation as instrument for vegetation barrenness. The dependent variable is the incidence of a conflict event in a given cell and month. Data is obtained from Murdock's Ethnographic Atlas. "Ethnic Diversity" applies when at least two ethnic homelands are present within the grid. Clustered standard errors at grid level in parentheses. ***/**/* indicate significance at 1%/5%/10%, respectively. All regressions include lagged conflict as control.

Table 2A.13: Long Run: Weather- and Human-Induced Components of Land Degradation

	(1)	(2)	(3)
Diff Veg. Barrenness	0.0859		
	(0.268)		
Diff Veg. Barrenness (Prediction)		0.315*	
		(0.189)	
Diff Veg. Barrenness (Residual)			-0.191
			(0.157)
Adjusted R^2	0.041	0.045	0.043
N	560	560	560
Climate Zone Trend	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 2000-2002 and 2010-2012, $\overline{DC}_i$, with one observation per cell. Robust standard errors in parantheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 2A.14: Long Run: Weather- and Human-Induced Components of Land Degradation accounting for Temperature

	(1)	(2)	(3)
Diff Veg. Barrenness	0.0859		
	(0.268)		
Diff Veg. Barrenness (Prediction with Temp)		0.333	
		(0.212)	
Diff Veg. Barrenness (Residual)			-0.186
			(0.172)
Adjusted R^2	0.041	0.045	0.042
N	560	560	560
Climate Zone Trend	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 2000-2002 and 2010-2012, $\overline{DC}_i$, with one observation per cell. Robust standard errors in parantheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 2A.15: Long Run: 4-Year Window

	(1)	(2)	(3)	(4)	(5)
	All	Riot	Battle (terr)	Battle (no terr)	Agropastoral
Diff Veg. Barrenness	-0.0190	0.0432	-0.102	0.00301	0.0598***
	(0.245)	(0.0845)	(0.147)	(0.0171)	(0.0215)
Adjusted R^2	0.035	0.013	0.023	0.006	0.021
N	560	560	560	560	560
Climate Zone Trend	✓	✓	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 2000-2003 and 2009-2012, $\overline{DC}_i$, with one observation per cell. Robust standard errors in parantheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 2A.16: Long Run: 7-Year Window

	(1) All	(2) Riot	(3) Battle (terr)	(4) Battle (no terr)	(5) Agropastoral
Diff Veg. Barrenness	-0.0508	0.0816	-0.0538	-0.0237	-0.0208
	(0.175)	(0.0808)	(0.129)	(0.0215)	(0.0277)
Adjusted R^2	0.011	0.009	0.002	0.000	0.001
N	560	560	560	560	560
Climate Zone Trend	✓	✓	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 2000-2006 and 2007-2012, $\overline{DC}_i$, with one observation per cell. Robust standard errors in parantheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Table 2A.17: Long Run: Land Cover Changes

	(1)	(2)	(3)	(4)
Diff Share Cropland	0.750*	0.557		
	(0.436)	(0.508)		
Diff Share Forest	-0.0333		0.0190	
	(1.053)		(1.009)	
Diff Share Grassland	0.128			0.0443
	(0.104)			(0.102)
Diff Veg. Barrenness (Residual)		-0.0611	-0.191	-0.193
		(0.201)	(0.158)	(0.157)
Adjusted R^2	0.045	0.045	0.041	0.041
N	560	560	560	560
Climate Zone Trend	✓	✓	✓	✓

OLS estimation results. The dependent variable is the difference in the average incidence of conflict events in a given cell and month between 2000-2002 and 2010-2012, $\overline{DC}_i$, with one observation per cell. Robust standard errors in parantheses. ***/**/* indicate significance at 1%/5%/10%, respectively.

Chapter 3

Resilience to Natural Disasters: Insurance and Institutions

3.1 Introduction

The common perception regarding the key determinants of economic development is that good institutions foster development while natural catastrophes constitute one of the key impediments to development. However, a glance at the empirical literature reveals an unresolved controversy about whether natural catastrophes indeed have significant and persistent negative or positive effects on income, and under which circumstances these effects unfold. As is discussed in more detail below, the existing evidence reveals a surprisingly heterogeneous picture of the development consequences of natural catastrophes, with institutions being one of the main determinants of the sign of the effect. Most of the existing literature presents reduced form effects, with little evidence for the channels and mechanisms that influence the effect of natural catastrophes on economic development.

This paper contributes to the debate by providing new evidence on the effect of natural catastrophes on economic development, and in particular on the determinants of the sign of this effect. The analysis uses a comprehensive data set of natural catastrophes as well as a measure of the damages caused by the catastrophes. This data has global coverage on all natural disasters and related losses, thus allowing for an estimation of the economic consequences of natural catastrophes by distinguishing the extensive and intensive margin. The results suggest that the effect of natural catastrophes depends on the

---

This chapter is based on joint work with Florian Englmaier, Till Stowasser and Uwe Sunde. A brief version of this Chapter is published as Breckner, Englmaier, Stowasser, and Sunde (2016), see `https://doi.org/10.1016/j.econlet.2016.09.023`. We are grateful to Munich Re's NatCat Service for granting access to comprehensive data on natural disasters and to Munich Re's Economic Research Unit for the provision of comprehensive data on insurance market penetration rates.

access to insurance in the form of private insurance markets in combination with public disaster relief. The findings demonstrate that private insurance markets and a stable, well-institutionalized environment complement each other in accommodating the negative effects of natural catastrophes. This implies that market forces and public institutional infrastructure are both essential in providing economies with resilience against natural catastrophes.

This paper makes several contributions to the existing literature, which has shown that the impact of natural disasters on income depends on the type and severity of natural disasters, as well as on the economic and institutional environment. In particular, almost all studies using cross-country panel data find negative effects of natural disasters on income in the short-run, in particular in developing countries and for severe disasters (Noy, 2009; Hochrainer, 2009; Raddatz, 2009; Loayza, Olaberría, Rigolini, and Christiaensen, 2012; Fomby, Ikedab, and Loayza, 2013), whereas there is some evidence that suggests a positive effect on income in developed economies, see, e.g. Noy (2009). While the literature lacks a coherent explanation for this finding, some suggest that this effect is mechanical as reconstruction investment is part of GDP while the loss due to destruction of capital is not (e.g. Tol and Leek, 1999). Some recent studies provide evidence that international openness and access to finance can raise a country's resilience to natural hazards, with higher openness to trade, higher financial openness and more advanced financial markets being attenuating factors that operate towards economic recovery in the aftermath of a natural disaster (Noy, 2009; McDermott, Barry, and Tol, 2013; Felbermayr and Gröschl, 2014). Noy (2009) also suggests that higher levels of government spending belong to the list, whereas higher foreign exchange reserves appear to worsen the disaster impact. Our study adds to this a novel measure of insurance market development. This measure reflects insurance market penetration based on micro level data and allows for a precise measurement of the role of insurance for resilience to natural catastrophes. Von Peter, von Dahlen, and Saxena (2012) present the first evidence that links the effect of natural disasters to insurance markets and show that, when treating uninsured and insured losses separately, uninsured disaster-related losses lead to income declines whereas there is no negative effect for insured losses. While we have access to the same data, this paper broadens the focus by considering insurance market penetration as control and as a further mitigating factor.

A distinct strand of the literature suggests that particular institutional attributes are relevant for mitigating the economic consequences of natural disasters, with countries with more stable and more democratic regimes appearing to be more capable to withstand the disaster shock (Noy, 2009; Cavallo, Galiani, Noy, and Pantano, 2013; Felbermayr and Gröschl, 2014). Our paper provides an important link between the functioning of insurance markets in attenuating the effects of catastrophes and the institutional environment.

Most other published studies employ the Em-Dat disaster database with one exception being Felbermayr and Gröschl (2014) who introduce the Geological and Meteorological Events Database (GAME) to the literature. This data is based on measures of the physical attributes of disasters. We employ the NatCat database provided by MunichRe, which constitutes the most comprehensive data set for disaster-related losses, including information on whether the losses were insured or not. Only few other studies have used these data in the context of macroeconomic resilience to natural disasters (von Peter, von Dahlen, and Saxena, 2012; Felbermayr and Gröschl, 2014).

The remainder of the paper is structured as follows. Section 3.2 describes the data and the empirical framework. Section 3.3 presents the main results and robustness analysis. Section 3.4 concludes the analysis.

3.2 Data and Empirical Framework

3.2.1 Data

We construct a panel data set with yearly data for 129 countries for the period 1980 to 2011.[1] Data on natural catastrophes is provided by the NatCat Service of the global insurance- and reinsurance group MunichRe. The data set contains information on the incidences of natural catastrophes on a global scale.[2] The data further reports the disaster types and also includes measures of the intensity of these catastrophes in terms of direct monetary losses and the number of fatalities, and provides information on different kinds of infrastructure assets affected. Of particular relevance for our analysis is the classification of disasters into severity categories 0-4, which are defined according to fatality- and monetary loss thresholds.[3]

The main advantage of the NatCat data over alternative data sets on natural catastrophes such as the publicly available Em-Dat data set which is employed in almost all published studies, is their comprehensiveness as well as the assessment of losses caused by the catastrophe. These loss data are of very high quality as they are essential for the estimation of reinsurance liabilities and the adequate risk pricing of contracts by MunichRe,

[1]Due to missing observations for some countries, the panel is not balanced.

[2]Wirtz, Kron, Loew, and Steuer (2014) provide an extensive description of data bases on natural disasters with a special focus on NatCat data.

[3]For instance, in order to be classified into category 4 in a high-income economy, a disaster must have caused either 2.5 billion$ or 1000 fatalities. See Table 3A.1 for all threshold definitions. For classification, losses are normalized by a normalization factor (current income to income in the respective year) which accounts for inflation and the increase in values.

which is the largest reinsurance company worldwide.[4] The calculation of disaster-related losses is mainly based on replacement and repair costs and draws on various sources, including the insurance industry, scientific reports, weather services, news agencies, NGOs and GOs. According to their own assessment, NatCat Service provides the most comprehensive natural catastrophe loss database in the world (NatCatService, 2014). The loss data distinguishes between insured losses and economic (overall) losses. The accuracy of loss data and the distinction between overall and insured losses makes the NatCat data unique for the purpose of this study. For instance, the smallest loss registered in the NatCat database amounts to 4450 US$, while disasters need to meet specific severity criteria before they are entered into alternative data bases, such as the Em-Dat database.[5]

In the empirical analysis, we employ two different specifications to capture natural catastrophes. First, we code a binary measure for natural disasters which is 1 if a severe disaster (category 4) occurred in country i, year t, and 0 otherwise. This measure only exploits the extensive margin of a natural catastrophe occurrence. Because the majority of events is related to relatively small losses that have ambiguous effects on income, we code catastrophes to be severe events (category 4). Second, as a measure of catastrophe intensity, we use the sum of direct losses caused by natural disasters in country i, year t, normalized by the level of GDP (of the preceding year). This measure exploits the intensive margin of disaster occurrence and makes use of the availability of high accuracy loss data. Losses are normalized by GDP to set the catastrophe intensity in relation to the country size. The weighted loss measure (losses per GDP) exhibits an outlier problem, where in some rare cases losses can amount to twice the level of GDP in extremely small countries. In particular, some small island states are affected in this respect. In order to accommodate this problem, and to allow for a straightforward interpretation, the log of the weighted loss is taken, which yields a rather normal distribution (see Figure 3.1). We also only consider shares of losses exceeding 0.1 percent in the baseline analysis to rule out that extremely small losses influence the estimated coefficients that can clearly not affect aggregate income.[6]

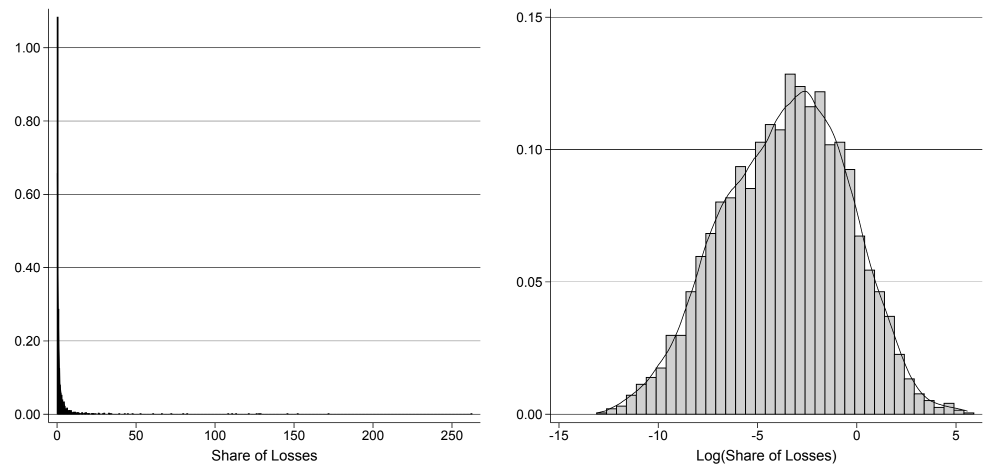
Another innovation in this paper concerns the availability of data on the development of insurance markets. In particular, the Economic Research Department of MunichRe provided us with data on national insurance market penetration for a worldwide panel. The availability of this data allows us to investigate whether access to insurance markets can help mitigating the consequences of natural catastrophes on economic development.

[4] Source: Standard & Poors, see `http://de.statista.com/statistik/daten/studie/188545`.

[5] For instance, for a disaster to be entered into the Em-Dat database at least one of the following criteria must be fulfilled: Ten or more people reported killed, hundred or more people reported affected, declaration of a state of emergency or a call for international assistance.

[6] Figure 3A.1 depicts the evolution of the two measures for natural disasters over time.

Figure 3.1: Histogram and Kernel Density Plot of Loss Measures

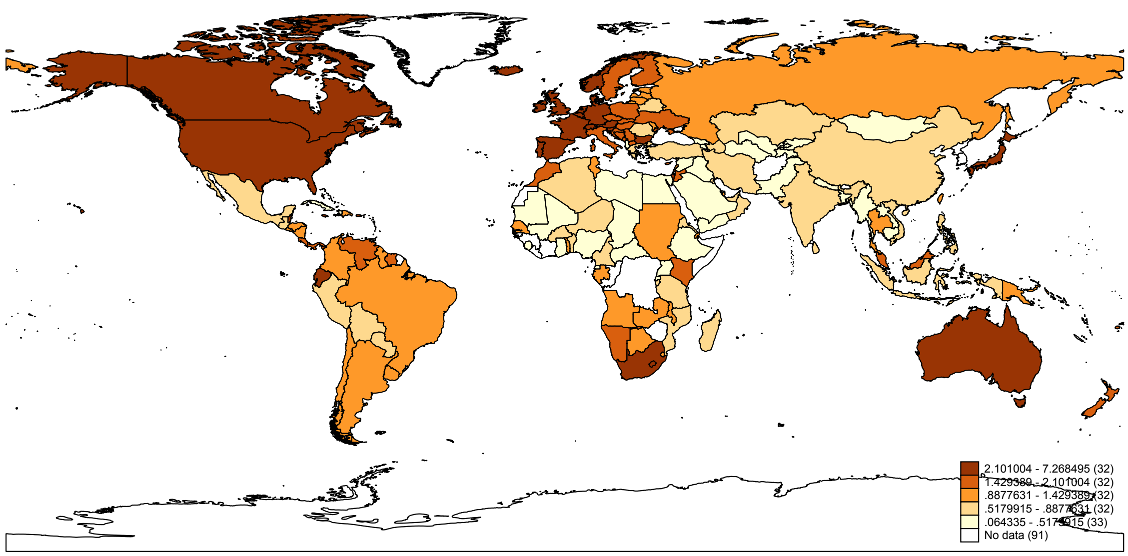


The left panel depicts losses per GDP in absolute terms whereas the right panel depicts losses per GDP in logs.

The main measure we employ is the insurance penetration rate, which is defined as the annual sum of insurance premia paid in a country divided by the country's GDP. We focus on insurance premia excluding health- and life insurance. This leaves us with property- and casualty insurance, which we consider as best proxy for insurance coverage in the case of natural catastrophes. The availability of other insurance measures is used in further robustness checks. In addition to the investigation of insurance penetration as a mitigating factor, the availability of this measure enables us to account for the concern that the effect of natural disasters on income might be upward biased if better developed insurance markets correlate with both, the measurement (or selection) of disasters and the level of development, as suggested by Felbermayr and Gröschl (2014), by including the insurance penetration rate as control variable.

Figure 3.2 shows a map of the average insurance penetration rate across countries over the observation period 1980-2014. Data on aggregate and per capita GDP, as well as on population is obtained from the Worldbank's Development Indicators (WDI). Data on the capital stock and human capital in terms of a human capital index is taken from the Penn World Tables 8.0 (Feenstra, Inklaar, and Timmer, 2015). The human capital index draws on the database of Barro and Lee (2013) and reflects a function of the average years of schooling for the population aged 15 or older. Data on institutions draws on several sources. The two main indicators for institutional quality, civil liberties and political rights, are obtained from Freedom House. Civil liberties involve feedom of expression and belief, associational and organizaitonal rights, rule of law and personal autonomy without inference from the state. Political rights involve the quality of the electoral process, political pluralism and participation as well as the functioning of the government.[7] For further robustness analysis, we employ a new measure of the quality of political institutions provided by Kunčič (2014). This measure comprises different concepts of measuring the well-functioning of political institutions to a new aggregate index. This provides us with a measure of the higher order attributes which is the latent quality of political institutions. Further, we employ the polity2 index from the polity4 database for robustness which focuses on institutionalized democracy (Marshall, Gurr, and Jaggers, 2016). Codings of the competitiveness of political participation, the openness and competitiveness of executive recruitment, and constraints on the chief executive yield a scale moving from complete autocracy to full democracy (-10,-10). Table 3.1 contains summary statistics of the main variables used in the empirical analysis.[8]

Figure 3.2: Quintiles of Countries' Average Insurance Penetration Rates, 1980-2014



3.2.2 Empirical Strategy

To investigate the effect of natural disasters on income we estimate the following empirical model:

$$lnY_{i,t} = \alpha + \beta lnY_{i,t-1} + \gamma DIS_{i,t} + \mu X_{i,t-1} + \nu_i + \nu_t + \nu_{i \cdot T} + \epsilon_{i,t}, \tag{3.1}$$

[7] See `https://freedomhouse.org/` for details. We use version 20 Dec 13, see (Teorell, Charron, Dahlberg, Holmberg, Rothstein, Sundin, and Svensson, 2013).

[8] See Table 3A.1 for details on data, data sources and variable construction.

Table 3.1: Summary Statistics - Estimation Sample

	N	Min	Max	Mean	SD
Disaster(Cat4)	2,572	0.00	1.00	0.13	0.34
Disaster(Log Loss)	2,562	-2.30	4.41	-0.20	0.70
Insurance Penetration Rate	2,519	1.00	7.00	3.09	1.69
Civil Liberties	2,519	1.00	7.00	3.00	2.02
Political Rights	2,572	0.00	32.05	1.49	1.33
log GDP per Capita	2,572	-1.94	4.46	1.40	1.60
log Capital Stock	2,572	7.03	17.61	12.30	1.98
log Population	2,572	12.20	21.02	16.30	1.65
log Human Capital Index	2,572	0.12	1.29	0.87	0.25
Disaster(Cat 4) Events only	341	1.00	1.00	1.00	0.00
Disaster(Log Loss) Events only	721	-2.30	4.41	-0.72	1.18

where the dependent variable $lnY_{i,t}$ is the log of per capita income in country i and year t. One lag of the dependent variable is included to capture convergence effects. The coefficient of primary interest is γ, which captures the impact of natural disasters on income. The variable $DIS_{i,t}$ represents the incidence of a natural catastrophe in country i and year t. Catastrophes are measured in two ways as described in more detail in the previous section. The first specification considers a binary indicator that takes value 1 in a year of a severe (cat 4) disaster, and 0 otherwise. The second specification considers the occurence of any disaster together with the log of the weighted sum of overall (monetary) disaster-related losses that occurred within the disaster year to analyze the impact of the intensive margin of disaster occurrence. The vector X denotes a set of control variables and contains the capital stock, total population and human capital.[9] All control variables enter in lags to avoid endogeneity due to a simultaneous impact of a disaster on dependent- and explanatory variables. The specification includes country fixed effects, ν_i, to account for time-invariant country characteristics and a full set of time (year) dummies, ν_t, to capture common time trends. In addition, the specification includes country-specific linear time trends, $\nu_{i \cdot T}$, to account for unobserved country-specific factors that are varying systematically over time. The inclusion of country specific linear time trends captures the diverse evolution of incomes over time and facilitate an accurate estimation of disaster shocks to differential income paths.[10] Standard errors are clustered at the country level and are robust to heteroskedasticity. This is necessary as for instance the measurement precision might be

[9]The specification thus reflects the factors of production in a human capital augmented Solow growth model (Mankiw, Romer, and Weil, 1992).

[10]In particular, the inclusion of country-specific linear time trend ensures that no unobserved country-specific trends drive the results. For instance one might think of improvements in disaster data quality or reporting that have been especially strong in transition economies.

correlated with the amount of losses. Combining a fixed effects estimator with a lag of the endogenous variable on the right hand side of the equation leads to biased estimates (Nickell, 1981). However, the asymptotic order of this bias is shown to decline with the length of the panel (Judson and Owen, 1999). In the above estimation framework with T=34 the Nickell bias therefore does not constitute a major concern.[11]

Another potential concern is that the measure of natural catastrophes is endogenous to economic development and insurance market development (Felbermayr and Gröschl, 2014). The first reason is that the amount of monetary losses caused by a natural disaster might correlate with the (insurance market-) development status of a country. Moreover, (insurance market-) development might correlate with the selection of disasters into the data set, if the insurance industry is a major source of information for compilation.[12] According to McDermott, Barry, and Tol (2013) the first concern is addressed by employing a dichotomous measure as is done in our first specification. The second concern regarding a potential selection bias is accounted for by the inclusion of country fixed effects and country-specific trends into the regression. Further, it is alleviated by employing a specification that only considers the most striking natural disasters based on a classification scheme that moreover accounts for differential income levels across countries. However, to fully address these issues it is necessary to include an interaction term between disasters and insurance market development into the regression to absorb the omitted effect at the moment that the disaster strikes. In the course of investigating the mediating effect of insurance markets by including an interaction term into the regression we therefore implicitly alleviate this concern.[13] To investigate the mediating effect of insurance markets we estimate the following empirical model:

$$lnY_{i,t} = \alpha + \beta lnY_{i,t-1} + \gamma DIS_{i,t} + \delta DIS_{i,t} * INS_{i,t-1} + \mu X_{i,t-1} + \nu_i + \nu_t + \nu_{i \cdot T} + \epsilon_{i,t}, \quad (3.2)$$

where an interaction term between the natural disaster and the insurance market penetration rate $DIS_{i,t} * INS_{i,t-1}$ is added to equation (3.1). The insurance penetration rate enters as lag such that it is not affected by the disaster shock and is included in the vector X. The coefficient of interest is δ, which measures the mediating effect of insurance markets on the effects that natural disasters have on income.

[11] Judson and Owen (1999) show that the asymptotic order of bias is 1/T and that for an average number of T=30 the bias is already moderate.

[12] On the other hand one might argue that disasters in poorer countries cause more fatalities (Kahn, 2005) and therefore will be more extensively covered in the databases, particularly the Em-Dat database which sets fatality thresholds to select disasters into the database. A robustness check to this concern is thus to validate the results using the publicly available Em-Dat data, which uses different selection criteria to sort events into their data set than the NatCat data.

[13] The robustness section further provides estimates including an interaction term with income.

3.3 Main Results

3.3.1 Baseline Effect of Natural Disasters

The main results regarding the effect of natural catastrophes on economic development are presented in Table 3.2. Columns (1)-(3) show the results when focusing attention on the extensive margin in terms of the incidence of a category 4 natural catastrophe in a given year. Two findings are relevant. First, on average the incidence of a natural catastrophe appears to be detrimental for development by reducing GDP per capita by more than half a percent, as indicated by the results in Column (1). Second, there appears to be pronounced heterogeneity in the effect, depending on the level of development. In particular, the effect is negative but not statistically significant in OECD countries as shown in Column (2), whereas the effect is larger in size and statistically significant negative in non-OECD countries, displayed in Column (3). This replicates the broad picture revealed by the existing literature, but it leaves open whether the negative effect is affected by the size of disaster-related losses (the intensive margin). Moreover, it leaves open the reasons for why developed countries are apparently more resilient to the occurrence of natural catastrophes than less developed countries.

Columns (4)-(6) address the question regarding the intensive margin by presenting results for an extended specification that includes both measures, the measure for disaster incidence and the disaster-related losses. The results of this specification show that the severity of the natural catastrophe, rather than the mere occurrence, matters for the economic consequences. Regarding the sub-samples, the extended specification yields qualitatively very similar results to the baseline specification with the disaster indicator.

Overall, these results suggest a negative effect of natural disasters on GDP per capita based on different disaster specifications, in line with Noy (2009) and Felbermayr and Gröschl (2014). Existing research has pointed to the fact that the impact of natural disasters on income depends on different features of the socio-economic environment (e.g., trade openness, financial openness, share of insured losses), as well as the quality of institutions (e.g., democratic institutions, political stability), see Noy (2009), McDermott, Barry, and Tol (2013), Felbermayr and Gröschl (2014), Fomby, Ikedab, and Loayza (2013), von Peter, von Dahlen, and Saxena (2012) and Loayza, Olaberría, Rigolini, and Christiaensen (2012).

3.3.2 Effects of Insurance Markets

In order to investigate in more detail why developing countries suffer more from natural disasters, we first explore whether better developed insurance markets in terms of higher insurance penetration help to mitigate the negative effect of natural disasters in OECD

Table 3.2: The Effect of Natural Catastrophes on Development

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
Disaster Cat4(t)	-0.619**	-0.277	-0.776**			
	(0.267)	(0.206)	(0.363)			
Disaster(t)				-0.553**	-0.552**	-0.533*
				(0.241)	(0.232)	(0.288)
Losses(t)				-0.394***	-0.245	-0.424**
				(0.141)	(0.148)	(0.169)
N	3844	868	2976	3822	868	2954
Countries	129	33	104	129	33	104
R-squared	0.961	0.992	0.956	0.963	0.992	0.958
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

countries. Insurance penetration has not played a great role as one of the potential reasons for the apparent heterogeneity in the effects of natural catastrophes in the literature so far. Thus, as a first step, we investigate the development of the insurance market as potential reason for the heterogeneity of different effects in the different samples.

Table 3.3 presents the results from estimating an empirical specification that includes an interaction term between the insurance penetration rate and the respective disaster measure (incidence, loss). The table follows the same structure as Table 3.2. Columns (1)-(3) show the results for specification using disaster incidence, while columns (4)-(6) show the results for the specification that also accounts for the intensive margin in terms of overall losses. By itself, insurance penetration does not appear to be related to economic development above and beyond the lagged controls from a standard development accounting framework. Regarding the effect of natural catastrophes, the negative coefficient for the entire sample is slightly larger than in the baseline specification, and significant. This is true for the full sample as well as the two sub-samples. In OECD countries, the coefficient of the main effect of natural disasters is significant and even larger than for the sample of non-OECD countries. At the same time, the results provide evidence for a significant positive interaction between insurance penetration and disasters in the full sample. This effect is mainly driven by the OECD sample and not significant in the non-OECD countries. Thus, at least in the OECD sample, the negative effect of the occurrence of a natural catastrophe is mitigated by higher insurance penetration.

The results are similar when considering the extended specification that also includes

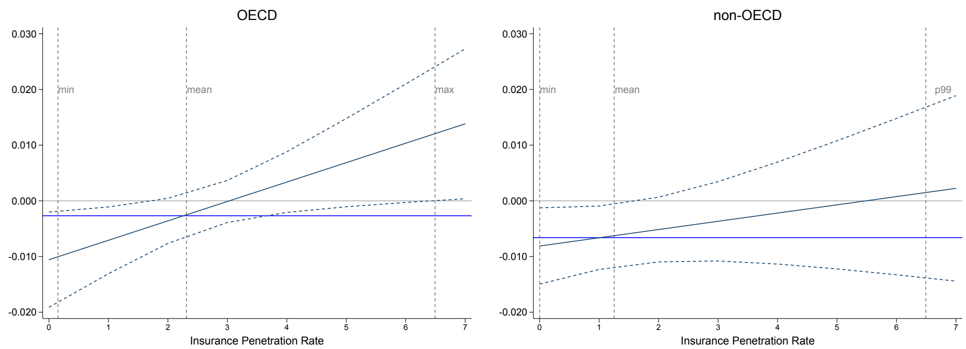
Table 3.3: The Mitigating Effect of Insurance

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
Disaster Cat4(t)*INS(t-1)	0.290**	0.349**	0.148			
	(0.124)	(0.151)	(0.143)			
Disaster Cat4(t)	-0.870***	-1.057**	-0.810**			
	(0.326)	(0.436)	(0.350)			
INS(t-1)	0.179	-0.498	0.141	0.0853	-0.415	0.0512
	(0.123)	(0.405)	(0.138)	(0.164)	(0.338)	(0.181)
Disaster*INS(t-1)				0.104	0.103	0.103
				(0.109)	(0.0956)	(0.131)
Disaster(t)				-0.351**	-0.421	-0.335
				(0.172)	(0.258)	(0.206)
Losses(t)*INS(t-1)				0.0800*	0.446***	0.0349
				(0.0469)	(0.152)	(0.0350)
Losses(t)				-0.400***	-1.218***	-0.363**
				(0.152)	(0.430)	(0.158)
N	2572	677	1895	2562	677	1885
Countries	126	33	101	126	33	101
R-squared	0.977	0.991	0.975	0.980	0.991	0.979
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

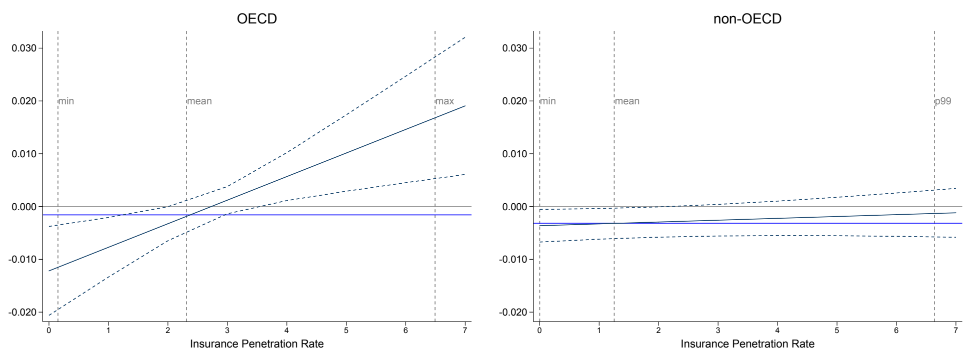
Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

the measure of catastrophe severity in terms of losses. Again, higher losses imply more negative development effects, but insurance penetration dampens this effect significantly, at least in developed economies. Hence, ignoring the role of insurance markets appears to confound negative effects of natural catastrophes with the mitigation due to higher insurance penetration, which leads estimates of the average effect to be insignificant in the OECD sample. Overall, the findings suggest that natural disasters have a negative effect on income in both samples, but that insurance markets attenuate the income decline in OECD countries.

Figure 3.3 depicts the total effect of the occurrence of a natural catastrophe in terms of the occurrence of severe natural disasters in Panel (a) and the log share of losses in Panel (b) on GDP per capita. As long as insurance penetration is below a certain threshold, natural catastrophes have an unambiguously negative effect on income per capita. In OECD countries, this effect is mitigated with increasing access to insurance, in terms of higher insurance penetration, and, with average penetration the effect is already insignificant. The same is true when considering losses. In non-OECD countries, however, the mitigating effect of insurance penetration is substantially weaker. In particular, even at average insurance penetration levels (or at average levels exhibited by OECD countries) the effect of the occurrence of a natural catastrophe is negative. This raises the question why insurance markets appear not to abate the consequences of natural catastrophes in non-OECD countries.



(a) Effect of Natural Disasters (Cat 4) on GDP per capita ($\gamma + \delta * InsurancePenetration(\%)$)



(b) Effect of Natural Disasters (Losses) on GDP per capita ($\gamma + \delta * InsurancePenetration(\%)$)

Figure 3.3: The Mitigating Effect of Insurance
Graphs include 95% confidence interval calculated via the delta method. The horizontal light blue lines indicate point estimates of table 3.2, i.e. ignoring insurance.

3.3.3 Robustness

In the following, we will report on the robustness of these findings to the use of alternative measures and estimation approaches, before investigating in more depth the mechanisms behind the results. The tables with the respective results are contained in the Appendix.

The first step of the robustness analysis explores the sensitivity of the results with respect to alternative measures. The findings are robust to the use of alternative measures of losses and a restriction to large disasters.[14] Likewise, the results also hold when accounting for the ratio of insured losses over all losses as an alternative measure for the insurance penetration rate.[15] This measure might even be a more accurate measure for insurance coverage regarding the destructed assets. The results are not confined to the use of the Nat-Cat data on natural catastrophes and also replicate when using the Em-Dat data.[16] Finally, the results for the OECD sample hold when using data on insurance penetration provided by the OECD.[17]

In a second step, we investigate the robustness of the results when adding additional interaction terms with natural catastrophe occurrence to rule out that insurance penetration picks up other factors, such as the level of development. When including an interaction term between natural catastrophes with income together with an interaction term with insurance penetration, the effect is essentially as in the baseline specification.[18] It turns out that the interaction of natural catastrophes with income remains statistically insignificant even without the inclusion of insurance penetration.[19] Adding additional controls, such as institutional quality, domestic credit, trade openness or government expenditures leaves the results unchanged.[20] At the same time, the positive interaction between natural catastrophes and insurance penetration remains unaffected by adding interaction terms of catastrophes with these additional controls.[21]

3.3.4 The Role of Institutions

Having documented a significant role of insurance in moderating the adverse effects of natural catastrophes on economic development, at least in the developed countries, we now turn to the question about the reasons for the apparent heterogeneity in this mitigation. In particular, the previous results suggest that it is not merely the level of development per se that is responsible for the finding that the negative effects of natural catastrophes are diminished by access to insurance in developed countries, but not in less developed countries. In particular, the effect of insurance remains when adding an interaction term with income. This suggests that it might be another factor that is related to the level of

[14]See Table 3A.2 in the Appendix.
[15]See Table 3A.3 in the Appendix.
[16]See Table 3A.4 in the Appendix.
[17]See Table 3A.5 in the Appendix.
[18]See Table 3A.6 in the Appendix.
[19]See Table 3A.7 in the Appendix.
[20]See Table 3A.8 in the Appendix.
[21]See Tables 3A.9, 3A.10, 3A.11, 3A.12, and 3A.13 in the Appendix.

development.

Felbermayr and Gröschl (2014) and Noy (2009) find that a prime candidate is higher quality of institutions by showing that higher institutional quality mediates the negative consequences of natural disasters. In the following we show that institutional quality also unfolds an indirect effectiveness through the channel of functioning insurance markets. To investigate this hypothesis, we estimate the model for additional sample splits by institutional quality. Table 3.4 reports the corresponding estimation results when splitting the full sample by institutional quality. Panel A of Table 3.4 reports results employing the indicator specification, while Panel B of table 3.4 reports respective results considering the extensive as well as the intensive margin. Columns (1)-(3) show results for a sample split according to a measure of civil liberties. Columns (1) and (2) split the sample at the median of country averages of this measure. Column (3) contains countries that constitute high institutional quality with respect to civil liberties, but do belong to the non-OECD sample. Column (4)-(6) report results for sample splits according to a measure of political rights. Column (4) and (5) split the sample at the median of country averages of this measure. Column (6) contains countries that constitute high quality of political institutions, but do belong to the OECD sample.[22]

[22] Institutions are measured via an ordinal index without straightforward cardinal interpretation, therefore we prefer the analysis via sample splits rather than using institutions in a linear regression as an interacting variable.

Table 3.4: The Role of Institutions

	Dependent Variable: log GDP p.c.(t)					
Institutions	Civil Liberties			Political Rights		
	high	low	high	high	low	high
Sample	All	All	non-OECD	All	All	non-OECD
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Extensive Margin					
Disaster Cat4(t)	-1.350**	-0.572	-1.633**	-1.292**	-0.807**	-0.913**
	(0.569)	(0.528)	(0.718)	(0.716)	(0.345)	(0.448)
INS(t-1)	0.259**	1.647**	0.247**	0.0781	0.311**	0.208**
	(0.0937)	(0.840)	(0.116)	(0.148)	(0.176)	(0.117)
Disaster Cat4(t)*INS(t-1)	0.442**	-0.0194	0.402**	0.456**	0.212	0.198**
	(0.154)	(0.533)	(0.180)	(0.260)	(0.131)	(0.118)
N	1271	1283	931	1260	1294	918
Countries	60	65	50	59	66	49
R-squared	0.985	0.974	0.981	0.986	0.972	0.980
	Panel B: Extensive and Intensive Margin					
Disaster(t)	-0.290	-0.552**	-0.332	-0.239	-0.486**	-0.193
	(0.266)	(0.198)	(0.308)	(0.241)	(0.216)	(0.269)
Losses(t)	-0.399**	-0.198	-0.592**	-0.650**	-0.389**	-0.414**
	(0.228)	(0.263)	(0.269)	(0.228)	(0.179)	(0.186)
INS(t-1)	0.267**	0.979	0.272**	0.0872	0.170	0.250**
	(0.109)	(0.890)	(0.123)	(0.184)	(0.126)	(0.124)
Disaster(t)*INS(t-1)	-0.0266	0.601	-0.0380	0.0212	0.134	-0.0523
	(0.0748)	(0.361)	(0.0910)	(0.120)	(0.174)	(0.0902)
Losses(t) * INS(t-1)	0.0834	-0.203	0.0859**	0.224**	0.00802	0.0654**
	(0.0519)	(0.242)	(0.0497)	(0.0730)	(0.0297)	(0.0320)
N	1269	1275	929	1258	1286	916
Countries	60	65	50	59	66	49
R-squared	0.985	0.980	0.981	0.986	0.978	0.981
Controls	yes	yes	yes	yes	yes	yes
Country FE	yes	yes	yes	yes	yes	yes
Country Trends	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes

Notes: Controls are GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Results in Columns (1) and (3) are based on the sub-sample with country averages of the civil liberties measure better than the median of the entire sample. Column (3) refers to countries with a civil liberties measure better than the median and that do not belong to the OECD sample. Results in Columns (4) and (6) are based on the sub-sample with country averages of the political rights measure better than the median of the entire sample. Column (6) refers countries with a political rights measure better than the median and that do not belong to the OECD sample. Accordingly, Columns (2) and (5) show results for samples with institutional quality measures worse than the median. Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Two findings are relevant here. First, insurance markets appear to have a mediating effect in countries with high quality of institutions. The coefficient on the interaction term is more pronounced and significant in column (1) and (4) than in Column (2) and (5). Second, this complementarity also unfolds detached from the development status. In Columns (3) and (6) we observe that access to insurance markets helps to mitigate the disaster shock in countries that have good institutions but are part of the non-OECD sample. The findings indicate that insurance penetration indeed only works as a mitigating factor for the adverse effects of natural catastrophes on economic development in environments with institutional quality above the median. This suggests an additional subtlety related to the earlier results, namely that the failure of finding the mitigating effect of insurance in non-OECD countries might be related to the lower institutional quality in that sub-sample. The results are supported by alternative indicators for institutional quality.[23]

3.4 Concluding Remarks

A number of studies have tackled the macroeconomic consequences of natural disasters. While the main part of the literature finds that natural disasters are harmful for income per capita in the short-run, some studies suggest that natural disasters may improve the macroeconomic performance. This paper contributes to the debate by providing new evidence on the economic effects of natural catastrophes, and in particular on how insurance markets influence the effects of natural disasters on income. We show that insurance markets mitigate the negative disaster shock in developed economies. Neglecting insurance markets may result in an insignificant negative or even positive effect of disasters on income per capita. However, when adding insurance, the results show a strong negative baseline effect of natural disasters on income, which is mitigated by insurance markets. Further, this paper provides new evidence for an interaction between access to insurance and institutional quality in mitigating the adverse economic effects of natural catastrophes. The results show that the failure of finding the mitigating effect of insurance in developing countries might be related to the lower institutional quality in that sub-sample. This finding implies that insurance and a stable, well-institutionalized environment complement each other in mediating the negative disaster shock.

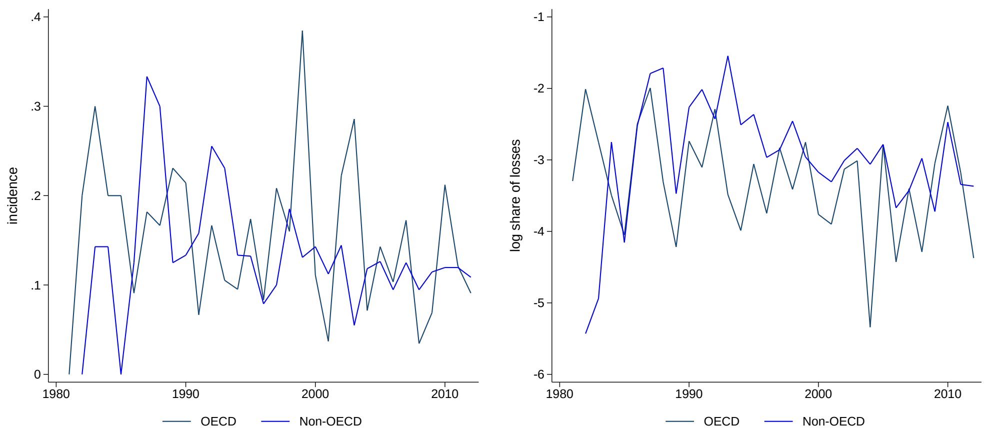
[23]Table 3A.14 reports the results for alternative measures of institutional quality. First, it shows results based on an index on the overall institutional performance of countries constructed by Kunčič (2014). Second, it shows results based on the polity2 index.

3.A Appendix

Table 3A.1: Data Sources and Variable Construction

	Source	*Variable*	*Specification*
		Natural Disasters	
1.	Natcat Service (MunichRe) (NatCatService, 2014) Robustness analysis based on EM-DAT (Guha-Sapir, Below, and Hoyois, 2015)	Disaster Cat4	Classification by income quartile: High: 2.5 billion$ | 1000 fatalities Upper-middle: 830 million$ | 1000 f Low-middle: 280 million$ | 1000 f Low: 90 million$ | 1000 f
		Disaster	Any disaster.
		Losses	Log sum of direct losses (t), normalized by GDP (t-1).
		Insurance Penetration	
2.	MunichRe Robustness analysis based on OECD database, see `https://data.oecd.org/`	Insurance Penetration Rate	Sum of insurance premia (non-life, non-health), normalized by GDP (in percent).
		GDP and Control Variables	
3.	World Bank Development Indicators (WDI) (World Bank, 2014)	GDP	Log GDP p.c. at constant 2005 national prices in million 2005 US$.
		Population	Log total population.
		Domestic Credit	Sum of domestic credit to private sector, normalized by GDP (in percent).
		Trade Openness	Sum of exports and imports of goods and services, normalized by GDP (in percent).
4.	Penn World Tables 8.0 (PWT) (Feenstra, Inklaar, and Timmer, 2015)	Physical Capital	Log capital stock at constant 2005 national prices in million 2005 US$.
		Human Capital Index	Quality adjusted measure of average years of schooling for population aged 15 or older.
		Government Expenditure	Sum of all government current expenditures for purchases of goods and services, normalized by GDP (in percent).
		Institutions	
5.	Freedom House (Teorell, Charron, Dahlberg, Holmberg, Rothstein, Sundin, and Svensson, 2013)	Civil Liberties Political Rights	Index [1,7], based on expert rating.
6.	Kuncic Political Institutions (Kunčič, 2014)	Institutional Quality	Index [0,1], based on computation of the latent quality of political institutions.
7.	Polity2 (Marshall, Gurr, and Jaggers, 2016)	Polity2	Index [-10,10].

Figure 3A.1: Yearly Average of Disasters by Specification for OECD and non-OECD



This figure depicts the annual evolution of the average of severe (Cat 4) disasters in the left panel and the average of the log loss measure in the right panel. The sample is split between developed (OECD) and developing (non-OECD) countries.

Table 3A.2: Alternative Measure of Losses

	(1) All	(2) OECD	(3) non-OECD
DV: log GDP p.c.			
INS(t-1)	0.00164	-0.00375	0.00138
	(1.34)	(-1.06)	(1.01)
Disaster>15(t)	0.104	-0.888***	0.401***
	(1.43)	(-3.38)	(13.05)
Losses(t)	-0.00218**	-0.0198***	-0.00171*
	(-2.23)	(-5.62)	(-1.82)
Disaster>15(t)*Losses(t)	-0.000683	0.0601***	-0.00218**
	(-0.42)	(3.59)	(-2.52)
Disaster>15(t)*INS(t-1)	-0.0846*		-0.0931***
	(-1.83)		(-5.89)
Losses(t)*INS(t-1)	0.000262	0.00813***	0.0000102
	(0.62)	(5.34)	(0.02)
Disaster>15(t)*Losses(t)*INS(t-1)	0.00294**		
	(2.29)		
Disaster(t)	-0.000537	0.0158**	-0.00168
	(-0.24)	(2.24)	(-0.73)
Disaster(t)*INS(t-1)	0.000382	-0.00733**	0.00153
	(0.45)	(-2.68)	(1.43)
Observations	2562	677	1885
N	2562	677	1885
Countries	126	33	101
R-squared	0.980	0.991	0.979
Controls	✓	✓	✓
Year FE	✓	✓	✓
Country FE	✓	✓	✓
Country Trends	✓	✓	✓

Notes: Share of losses, including dummy for very large disasters (>15% of GDP). There are only 3 observation within the non-OECD sample (Belize 2000, Belize 2001, Honduras 1998) and 2 observations in the OECD sample (Chile 2010, New Zealand 2011) for which the 15 percent criterion holds. As the interaction effect between these disasters and insurance yields more variation (degrees of freedom) than the indicator itself, the coefficient on the interaction between these large disasters can be estimated while the baseline effect is omitted. Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.3: Alternative Insurance Penetration Measure

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
Disaster Cat4(t)	-0.00842**	-0.00923**	-0.00684			
	(-2.49)	(-2.47)	(-1.56)			
Disaster Cat4(t)*INS (insured/overall loss)	0.0279***	0.0247**	-0.0590			
	(2.62)	(2.46)	(-0.59)			
Disaster(t)				-0.00604**	-0.0111***	-0.00575*
				(-2.14)	(-3.02)	(-1.66)
Disaster(t)*INS (insured/overall loss)				0.0149	0.0277**	-0.00983
				(1.06)	(2.49)	(-0.11)
Losses(t)				-0.00430**	-0.00862**	-0.00299
				(-2.57)	(-2.65)	(-1.52)
Losses(t)*INS (insured/overall loss)				0.00906	0.0257**	-0.0794
				(1.14)	(2.72)	(-1.33)
Observations	3844	868	2976	3822	868	2954
N	3844	868	2976	3822	868	2954
Countries	129	33	104	129	33	104
R-squared	0.961	0.992	0.956	0.963	0.992	0.958
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: INS is defined as country-specific average of the ratio insured losses/overall losses of all disasters that occurred 1980-2014. Since INS is time-invariant, the main effect is absorbed by country fixed-effects. Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.4: Equivalent to Table 7 using Em-Dat Data.

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
INS(t-1)	0.00177	-0.00511	0.00132	0.00215	-0.00761	0.00168
	(1.48)	(-1.26)	(0.99)	(1.34)	(-1.56)	(0.97)
Disaster Cat4(t) EmDat*INS(t-1)	0.00500*	0.00665***	0.00675			
	(1.97)	(3.41)	(0.88)			
Disaster dummy (cat4) EMDAT	-0.0128**	-0.0194***	-0.0131*			
	(-2.51)	(-4.26)	(-1.75)			
Disaster EmDat*INS(t-1)				-0.000414	0.00815**	-0.000421
				(-0.43)	(2.09)	(-0.47)
Disaster Em-Dat(t)				0.00150	-0.0192*	0.00154
				(0.52)	(-1.88)	(0.49)
Losses EmDat*INS(t-1)				0.000363	0.00489*	0.000274
				(0.68)	(1.88)	(0.52)
Losses Em-Dat(t)				-0.00232	-0.0142**	-0.000775
				(-1.50)	(-2.11)	(-0.49)
Observations	2627	677	1950	2617	677	1940
N	2627	677	1950	2617	677	1940
Countries	129	33	104	129	33	104
R-squared	0.977	0.991	0.975	0.980	0.991	0.979
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.5: Equivalent to Table 7, Column (2) and (5), using OECD Data.

DV: log GDP p.c.	(1) OECD	(2) OECD
Non-Life Ins. Penetration (t-1) OECD	-0.00509*	-0.00354
	(-1.84)	(-1.32)
Disaster Cat4(t)	-0.00761***	
	(-2.82)	
Disaster Cat4(t)*Non-Life Ins. Penetration (t-1) OECD	0.00236***	
	(2.78)	
Disaster(t)		-0.00938*
		(-1.77)
Disaster(t)*Non-Life Ins. Penetration (t-1) OECD		0.00185
		(1.27)
Losses(t)		-0.0108***
		(-3.00)
Losses(t)*Non-Life Ins. Penetration (t-1) OECD		0.00303***
		(3.26)
Observations	769	769
N	769	769
Countries	33	33
R-squared	0.991	0.992
Controls	✓	✓
Year FE	✓	✓
Country FE	✓	✓
Country Trends	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Only the non-life insurance penetration rate is available from the OECD data source, therewith this insurance measure includes health insurance in contrast to the baseline results. Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.6: Including two Interaction Terms, Income and Insurance

DV: log GDP p.c.	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
INS(t-1)	0.00176 (1.42)	-0.00496 (-1.23)	0.00141 (1.02)	0.00171 (1.00)	-0.00377 (-1.03)	0.00102 (0.55)
Disaster Cat4(t)	-0.00871*** (-2.67)	-0.00746 (-0.66)	-0.00812** (-2.40)			
Disaster Cat4(t)*INS(t-1)	0.00209 (1.60)	0.00432** (2.22)	0.00154 (1.10)			
Disaster cat4(t)*log GDP p.c.(t-1)	0.000974 (0.57)	-0.00159 (-0.36)	-0.000177 (-0.05)			
Disaster(t)				-0.00347** (-2.06)	-0.00419 (-1.66)	-0.00353* (-1.75)
Disaster*INS(t-1)				0.0000159 (0.01)	0.000483 (0.22)	0.000448 (0.34)
Disaster(t)*log GDP p.c.(t-1)				0.00167 (1.08)	0.000420 (0.27)	0.00153 (0.62)
Losses(t)				-0.00391** (-2.36)	-0.0142 (-1.42)	-0.00350** (-2.09)
Losses(t)*INS(t-1)				0.000773 (1.39)	0.00409*** (3.68)	0.000525 (1.36)
Losses(t)*log GDP p.c.(t-1)				0.00000418 (0.01)	0.000885 (0.31)	-0.000730 (-0.53)
Observations	2572	677	1895	2562	677	1885
N	2572	677	1895	2562	677	1885
Countries	126	33	101	126	33	101
R-squared	0.977	0.991	0.975	0.980	0.991	0.979
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.7: Including an Interaction Term with Income

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
INS(t-1)	0.00165	-0.00499	0.00133	0.00165	-0.00446	0.00135
	(1.40)	(-1.22)	(1.00)	(1.43)	(-1.12)	(1.04)
Disaster cat4(t)*log GDP p.c.(t-1)	0.00192	0.00311	0.000424			
	(1.34)	(0.97)	(0.12)			
Disaster Cat4(t)	-0.00688**	-0.0124	-0.00673**			
	(-2.10)	(-1.19)	(-2.06)			
Disaster(t)				-0.00341**	-0.00423	-0.00337*
				(-2.01)	(-1.63)	(-1.66)
Disaster(t)*log GDP p.c.(t-1)				0.00167	0.000855	0.00177
				(1.41)	(1.29)	(0.80)
Losses(t)*log GDP p.c.(t-1)				0.000366	0.00461	-0.000442
				(0.62)	(1.48)	(-0.35)
Losses(t)				-0.00313*	-0.0166	-0.00294*
				(-1.93)	(-1.53)	(-1.70)
Observations	2572	677	1895	2562	677	1885
N	2572	677	1895	2562	677	1885
Countries	126	33	101	126	33	101
R-squared	0.977	0.991	0.975	0.980	0.991	0.979
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.8: Adding further Controls

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
INS(t-1)	0.00224***	-0.00629	0.00210*	0.000702	-0.00490	0.000435
	(2.63)	(-1.64)	(1.92)	(0.71)	(-1.50)	(0.38)
Disaster Cat4(t)	-0.00820**	-0.0117**	-0.00806**			
	(-2.40)	(-2.71)	(-2.20)			
Disaster Cat4(t)*INS(t-1)	0.00246*	0.00371**	0.00155			
	(1.92)	(2.43)	(1.21)			
Disaster*INS(t-1)				0.00165	0.000461	0.00181
				(1.25)	(0.61)	(1.11)
Disaster(t)				-0.00390**	-0.00397	-0.00386*
				(-2.29)	(-1.34)	(-1.89)
Losses(t)*INS(t-1)				0.000719	0.00424***	0.000346
				(1.65)	(2.91)	(1.11)
Losses(t)				-0.00372**	-0.0119***	-0.00344**
				(-2.41)	(-2.80)	(-2.15)
Observations	2249	600	1649	2247	600	1647
N	2249	600	1649	2247	600	1647
Countries	120	32	96	120	32	96
R-squared	0.982	0.991	0.980	0.982	0.991	0.980
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1), log human capital (t-1), polity2 (t-1), domestic credit (t-1), trade openness (t-1) and government expenditure (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.9: Including two Interaction Terms, Polity2 and Insurance

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
INS(t-1)	0.00151	-0.00473	0.000980	0.00000842	-0.00243	-0.000649
	(1.19)	(-1.19)	(0.70)	(0.00)	(-0.64)	(-0.28)
Polity (t-1)	0.000629	0.000951	0.00108*	0.000498	0.00140	0.000881
	(1.19)	(0.13)	(1.69)	(0.83)	(0.20)	(1.20)
Disaster Cat4(t)	-0.00907**	-0.0150	-0.00792**			
	(-2.36)	(-0.90)	(-2.02)			
Disaster Cat4(t)*INS(t-1)	0.00243**	0.00326*	0.00181			
	(2.06)	(1.75)	(1.49)			
Disaster Cat4(t)*Polity(t-1)	0.000260	0.000460	0.0000415			
	(0.64)	(0.22)	(0.10)			
Disaster(t)				-0.00326**	-0.00413	-0.00318*
				(-2.02)	(-1.53)	(-1.70)
Disaster*INS(t-1)				0.00173	-0.000983	0.00191
				(1.03)	(-0.51)	(1.04)
Disaster(t)*Polity(t-1)				-0.0000998	0.000382	-0.000117
				(-0.23)	(0.62)	(-0.22)
Losses(t)				-0.00362*	-0.0285*	-0.00327*
				(-1.95)	(-1.96)	(-1.76)
Losses(t)*INS(t-1)				0.000805	0.00353**	0.000313
				(1.65)	(2.56)	(0.87)
Losses(t)*Polity(t-1)				-0.0000696	0.00191	-0.000113
				(-0.34)	(1.38)	(-0.51)
Observations	2440	643	1797	2433	643	1790
N	2440	643	1797	2433	643	1790
Countries	121	32	97	121	32	97
R-squared	0.979	0.990	0.977	0.981	0.991	0.980
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.10: Including two Interaction Terms, Trade Openness and Insurance

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
INS(t-1)	0.00134	-0.00471	0.00104	0.00210	-0.00449	0.00164
	(0.95)	(-1.17)	(0.68)	(1.44)	(-0.90)	(1.02)
Trade openness (t-1)	0.000529***	0.000370*	0.000470***	0.000414***	0.000409**	0.000362**
	(3.26)	(1.77)	(2.68)	(2.98)	(2.31)	(2.39)
Disaster Cat4(t)	-0.00924*	-0.0118**	-0.00922			
	(-1.81)	(-2.07)	(-1.53)			
Disaster Cat4(t)*INS(t-1)	0.00261*	0.00351**	0.00136			
	(1.96)	(2.22)	(0.92)			
Disaster cat4(t)*Trade Openness(t-1)	0.0000272	0.0000183	0.0000331			
	(0.53)	(0.16)	(0.56)			
Disaster(t)				-0.00363**	-0.00434	-0.00369*
				(-2.13)	(-1.54)	(-1.77)
Disaster*INS(t-1)				-0.00104	0.00182	-0.000878
				(-1.38)	(0.53)	(-0.97)
Disaster(t)*Trade Openness(t-1)				0.0000649**	-0.0000156	0.0000632**
				(2.56)	(-0.18)	(2.31)
Losses(t)				-0.00519**	-0.0132**	-0.00626*
				(-2.06)	(-2.55)	(-1.93)
Losses(t)*INS(t-1)				0.000668	0.00463***	0.0000823
				(1.57)	(3.00)	(0.19)
Losses(t)*Trade Openness(t-1)				0.0000173	0.00000981	0.0000365
				(0.76)	(0.41)	(1.11)
Observations	2555	677	1878	2547	677	1870
N	2555	677	1878	2547	677	1870
Countries	126	33	101	126	33	101
R-squared	0.978	0.991	0.977	0.981	0.991	0.980
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.11: Including two Interaction Terms, Government Expenditure and Insurance

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
INS(t-1)	0.00220**	-0.00807*	0.00193*	0.00244**	-0.00833**	0.00227*
	(2.38)	(-1.93)	(1.77)	(2.24)	(-2.29)	(1.90)
Government expenditure (t-1)	0.000736	0.00660**	0.000541	0.000882	0.00704**	0.000644
	(0.75)	(2.19)	(0.63)	(1.14)	(2.56)	(0.90)
Disaster Cat4(t)	-0.0153**	-0.00625	-0.0173**			
	(-2.54)	(-1.14)	(-2.42)			
Disaster Cat4(t)*INS(t-1)	0.00348***	0.00448**	0.00233*			
	(2.63)	(2.38)	(1.74)			
Disaster cat4(t)*Government Expenditure(t-1)	0.000776	-0.000957	0.00101*			
	(1.45)	(-0.97)	(1.67)			
Disaster(t)				-0.00520***	-0.00347	-0.00533**
				(-2.96)	(-1.27)	(-2.62)
Disaster*INS(t-1)				-0.000621	0.00237	-0.000718
				(-0.77)	(1.06)	(-0.80)
Disaster(t)*Government Expenditure(t-1)				0.000635**	-0.000531	0.000647**
				(2.28)	(-0.62)	(2.23)
Losses(t)				-0.0107***	-0.0119	-0.0120***
				(-2.99)	(-1.33)	(-2.94)
Losses(t)*INS(t-1)				0.000996*	0.00411**	0.000482
				(1.74)	(2.61)	(1.40)
Losses(t)*Government Expenditure(t-1)				0.000702**	0.0000629	0.000839**
				(2.44)	(0.08)	(2.53)
Observations	2407	634	1773	2399	634	1765
N	2407	634	1773	2399	634	1765
Countries	125	33	100	125	33	100
R-squared	0.976	0.991	0.974	0.980	0.991	0.978
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.12: Including two Interaction Terms, Domestic Credit and Insurance

	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
DV: log GDP p.c.						
INS(t-1)	0.00225**	-0.00483	0.00224*	0.00223*	-0.00501	0.00198
	(2.29)	(-1.17)	(1.79)	(1.75)	(-1.35)	(1.34)
Domestic credit (t-1)	-0.000418***	-0.0000305	-0.000547***	-0.000443***	-0.0000294	-0.000556***
	(-3.51)	(-0.27)	(-4.71)	(-3.92)	(-0.22)	(-4.82)
Disaster Cat4(t)	-0.0109***	-0.00736	-0.0166**			
	(-2.76)	(-1.64)	(-2.61)			
Disaster Cat4(t)*INS(t-1)	0.00178	0.00479***	0.000793			
	(1.42)	(2.78)	(0.42)			
Disaster cat4(t)*Domestic Credit(t-1)	0.0000452	-0.0000529**	0.000184*			
	(1.22)	(-2.61)	(1.86)			
Disaster(t)				-0.00392**	-0.00425	-0.00389*
				(-2.33)	(-1.65)	(-1.94)
Disaster*INS(t-1)				-0.0000603	0.00132	0.000229
				(-0.06)	(0.60)	(0.21)
Disaster(t)*Domestic Credit(t-1)				0.0000470	-0.00000956	0.0000455
				(1.27)	(-0.20)	(0.93)
Losses(t)				-0.00449**	-0.0141***	-0.00415
				(-2.32)	(-3.17)	(-1.66)
Losses(t)*INS(t-1)				0.000594	0.00437***	0.000353
				(1.61)	(3.05)	(1.15)
Losses(t)*Domestic Credit(t-1)				0.0000147	0.0000183	0.0000214
				(1.26)	(1.59)	(0.79)
Observations	2509	663	1846	2507	663	1844
N	2509	663	1846	2507	663	1844
Countries	126	33	101	126	33	101
R-squared	0.981	0.991	0.980	0.981	0.991	0.980
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.13: Including two Interaction Terms, Financial Openness and Insurance

DV: log GDP p.c.	(1) All	(2) OECD	(3) non-OECD	(4) All	(5) OECD	(6) non-OECD
INS(t-1)	0.00110	-0.00221	0.000793	0.0000755	-0.00103	-0.000275
	(0.82)	(-0.43)	(0.54)	(0.04)	(-0.21)	(-0.13)
Financial openness (t-1)	0.00144	0.00870***	-0.00166	0.000548	0.00719**	-0.00189
	(0.75)	(3.06)	(-0.80)	(0.22)	(2.73)	(-0.66)
Disaster Cat4(t)	-0.00808**	-0.0102**	-0.00752**			
	(-2.56)	(-2.64)	(-2.18)			
Disaster Cat4(t)*INS(t-1)	0.00263**	0.00367***	0.00181			
	(2.31)	(2.78)	(1.33)			
Disaster cat4(t)*Financial Openness(t-1)	0.000891	-0.000132	0.000840			
	(0.72)	(-0.07)	(0.44)			
Disaster(t)				-0.00317*	-0.00370	-0.00298
				(-1.89)	(-1.46)	(-1.51)
Disaster*INS(t-1)				0.00104	-0.0000164	0.00107
				(0.88)	(-0.01)	(0.80)
Disaster(t)*Financial Openness(t-1)				0.00123	0.00225	0.000662
				(0.64)	(1.18)	(0.28)
Losses(t)				-0.00330**	-0.0126***	-0.00309*
				(-2.12)	(-2.82)	(-1.95)
Losses(t)*INS(t-1)				0.000834	0.00352***	0.000347
				(1.53)	(3.12)	(0.98)
Losses(t)*Financial Openness(t-1)				-0.000498	0.00151	-0.00106
				(-0.80)	(1.25)	(-1.04)
Observations	2479	665	1814	2473	665	1808
N	2479	665	1814	2473	665	1808
Countries	122	32	97	122	32	97
R-squared	0.978	0.991	0.976	0.981	0.991	0.980
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Huber-White robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Table 3A.14: The Mitigating Effect of Insurance by Institutional Quality: Alternative Measures

	Institutional Quality			Polity2		
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Low	High	High non-OECD	Low	High	High non-OECD
	Panel B: Extensive and Intensive Margin					
Disaster Cat4(t)*INS(t-1)	0.363	0.367***	0.285***	0.224*	0.481	0.219*
	(0.484)	(0.104)	(0.0986)	(0.134)	(0.308)	(0.113)
Disaster Cat4(t)	-0.804	-1.107***	-1.388***	-0.801**	-1.442	-1.067**
	(0.579)	(0.369)	(0.397)	(0.325)	(0.863)	(0.468)
INS(t-1)	0.581	0.243***	0.276***	0.266*	0.0880	0.200
	(1.196)	(0.0883)	(0.0934)	(0.152)	(0.139)	(0.127)
N	1245	1214	879	1293	1251	917
Countries	65	56	48	68	57	49
R-squared	0.980	0.985	0.978	0.978	0.986	0.982
	Panel B: Extensive and Intensive Margin					
Disaster*INS(t-1)	0.733**	-0.0290	0.0188	0.109	0.0583	-0.0546
	(0.298)	(0.0917)	(0.115)	(0.158)	(0.123)	(0.0937)
Disaster(t)	-0.588***	-0.343	-0.498**	-0.595***	-0.207	-0.104
	(0.211)	(0.227)	(0.240)	(0.184)	(0.277)	(0.306)
Losses(t)*INS(t-1)	-0.104	0.0838*	0.0576*	0.0230	0.321***	0.0604*
	(0.276)	(0.0469)	(0.0295)	(0.0279)	(0.0772)	(0.0304)
Losses(t)	-0.295	-0.446**	-0.496***	-0.331*	-0.922***	-0.384*
	(0.295)	(0.177)	(0.174)	(0.168)	(0.266)	(0.217)
INS(t-1)	0.0462	0.251**	0.247**	0.152	0.0614	0.244*
	(1.110)	(0.124)	(0.111)	(0.126)	(0.181)	(0.139)
N	1245	1214	879	1293	1251	917
Countries	65	56	48	68	57	49
R-squared	0.981	0.985	0.978	0.978	0.986	0.982
Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Country Trends	✓	✓	✓	✓	✓	✓

Notes: Controls are log GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Notes: Controls are GDP p.c. (t-1), log population (t-1), log capital stock (t-1) and log human capital (t-1). Results in Columns (1) and (3) are based on the sub-sample with country averages of the institutional quality measure better than the median of the entire sample. Column (3) refers to countries with an institutional quality measure better than the median and that do not belong to the OECD sample. Results in Columns (4) and (6) are based on the sub-sample with country averages of the polity2 measure better than the median of the entire sample. Column (6) refers countries with a polity2 measure better than the median and that do not belong to the OECD sample. Accordingly Columns (2) and (5) show results for samples with institutional quality measures worse than the median. Huber-White Robust standard errors clustered at country-level are reported in brackets. ***, **, * indicate significance at 1-, 5-, and 10-% level, respectively. Coefficients and standard errors are multiplied by 100.

Chapter 4

Health on the Nile: The Curse of Living Downstream

4.1 Introduction

Urbanization and industrialization characterize economic development worldwide. Standards of living seem to be higher in urban and industrial areas, which promise job opportunities, better infrastructure and public good provision. Therefore many people are attracted to these areas. But one important downside of urbanization and industrialization is pollution. This is particularly problematic in developing countries as the institutional framework is usually weak and environmental regulation is not enforced. Industrial wastewater and sewage containing toxic substances are often discharged into rivers. Hence, water pollution is a particular concern associated with high population density and industrial activity. The WWF calls it "one of the most serious ecological threats we face today" (WWF, 2017) and the WHO estimated that more than 360,000 children under 5 years die due to diarrhea, as a result of poor access to clean water, sanitation, and hygiene (WHO, 2017c).

Water pollutants are spread by moving water, and so the flow of rivers through cities and past industrial plants determines which households are affected by polluted water. We use this quasi-experimental variation in exposure to pollutants and examine the effect of pollution on children's health. We construct a novel panel data set based on geo-coded DHS survey data between 1992 and 2014, geo-coded factory locations and finely gridded population density in Egypt. We aggregate this data for 78 segments along the Nile. We then estimate the effect of population density and the existence of a factory on the health of children living up- or downstream of the pollutant. Further, we use information on the

This chapter is based on joint work with Marie Lechler.

opening year of factories. We are thus able to exploit temporal and spatial variation in water pollution and to hold region and time specific factors constant.

We find a negative effect of urbanization (measured by population density) and factory presence on health for the downstream population whereas we do not find a negative health effect on the upstream population. This differential health effect is most likely caused by water pollution and suggests that factories and agglomerations emit pollutants into the Nile, which worsen health outcomes. The effect does not seem to be driven by sorting as we do not find differential wealth effects. Further results show that the negative health impact of urban and industrial areas can be mitigated by access to clean drinking water. Moreover, we find that in the case of urbanization the negative health effect is localized and vanishes with increased distance between the pollutants and the population. In contrast, we find evidence that cumulative industrial pollution over several upstream grids still has detrimental health effects. This is due to the different nature of pollutants. While industrial pollution is to a large extent persistent due to its chemical composition, pollution caused by agglomerations degrades more rapidly.

With this study we contribute to a large body of literature, which links water pollution to poor health outcomes in developing countries. First, we add methodologically to the existing literature. Based on our unique data set we exploit both spatial and temporal variation in pollution and are therefore able to deal with potential endogeneity issues, as for instance systematic differences between upstream and downstream location and time-specific shocks. Second, we directly examine the health effects of the two most important pollutants, urban areas and industrial plants. Third, our study adds a highly relevant case, namely the Nile in Egypt, to the body of literature. Moreover, while most other studies linking water pollution to poor health outcomes only focus on specific cases we provide an extensive analysis of the effects along the entire course of the Nile river in Egypt.

There are a few papers that are methodologically related to our approach analyzing differential effects on the population depending on the relative location to the pollutant. Duflo and Pande (2007) compare agricultural productivity and vulnerability to rainfall shocks in Indian districts downstream of a dam with other districts. They thereby assume that people living downstream of a dam tend to benefit while those living in the vicinity or upstream do not. Garg, Hamilton, Hochard, Plous, and Talbot (2016) show for Indonesia that human bathing in upstream villages increases diarrheal incidence, while bathing of downstream villages has no effect. Romero and Saavedra (2016) examine health effects of mines in Columbia and find that while mothers living in the vicinity of a mine are positively affected, mothers living downstream from a mine are negatively affected.

Methodologically less related but relevant in the context of water pollution is a study by Brainerd and Menon (2014), who show that water quality has an effect on infant and child health in India, exploiting seasonal and geographic variation in the use of fertilizers. Greenstone and Hana (2014) also focus on India and study the impact of environmental regulations on infant mortality. They find however no significant effect of water regulations. Ebenstein (2012) uses variation in water pollution across river basins in China and shows that lower water quality is associated with a higher digestive cancer death rate. Galiani, Gertler, and Schargrodsky (2005) find that water privatization in Argentina decreases child mortality using variation in ownership of water provision across time and space. Another strand of literature deals with the effects of water pollution on other outcomes such as labor productivity. A recent study by Zhang and Xu (2016) finds for example a positive effect of a water treatment program in China on education. Zivin and Neidell (2013) provide an overview over quasi-experimental evidence on the negative effects of pollution on individual well-being in general and Currie, Zivin, Mullins, and Neidell (2014) provide an overview over the literature about early-childhood exposure to pollution and health and human capital outcomes later in life. This literature predominantly finds significantly negative effects of water pollution on health and human capital. While our study confirms this relationship for the case of Egypt it goes beyond the existing literature by studying two of the most hazardous pollutants, agglomerations and industrial plants.

Egypt provides an ideal setting to study health effects of water pollution. The Nile river is the country's only major river and around 90% of Egyptians live in the Nile valley and are thus directly or indirectly affected by polluted water. The Nile is the 'life artery' of Egypt and constitutes the most important freshwater resource for almost all water demands. The Nile water can thus reach human organisms through fishing, irrigation, the groundwater (which is for instance used for washing) and even as drinking water. The Nile's water quality has been deteriorating over several decades due to the disposal of industrial effluents and human sewage (Wahaab and Badawy, 2004; El-Ayouti and Abou-Ali, 2013; Ali, Shabaan-dessouki, Soliman, and Shenawy, 2014; Abdel-Satar, Ali, and Goher, 2017). Abdel-Satar, Ali, and Goher (2017) document spatial differences in the measured water quality of the Nile, which reflect "combinations of natural and human activities". In our study we focus on industrial activities and urbanization, which generate industrial wastewater and human sewage that are often disposed into the Nile. Water pollution is a serious concern as around 40% of the Egyptian population does not have access to 'safely managed'[1] sanitation (WashWatch, 2017), which fosters the transmission of diarrhoeal diseases. These diseases are particularly dangerous for children, who are extremely sensitive to dehydration and the related loss of electrolytes (WHO, 2017a). According to UNICEF, diarrhea is the second leading cause of death among under 5 year old children in Egypt

[1] 'Safely managed' sanitation refers to improved sanitation facilities that are not shared with other households and where excreta are safely disposed of in site or transported and treated off site.

(3,500 to 4,000 under 5 year old children die of diarrhea every year (UNICEF, 2017)). Egypt thus constitutes a compelling case to study the hazardous effects of water pollution on children's health.

This chapter is structured as follows. Section 4.2 gives a brief overview over the types of industrial and urban pollutants and their potential effects on health outcomes. Section 4.3 describes the compilation of the data set. Section 4.4 presents the analysis of urbanization whereas Section 4.5 presents the analysis of industrialization as source of pollution. Section 4.6 shows an analysis of accumulated water pollution. Section 4.7 concludes with a discussion.

4.2 Health Impact of Industrial and Urban Water Pollution

Water pollutants can be broadly classified into biodegradable and non-biodegradable pollutants. Biodegradable pollutants consist of organic matter that is broken down into simple organic molecules by natural agents like water, oxygen and micro-organisms. These molecules eventually return into the environment.[2] However, the speed rate of the degradation process differs strongly by material (e.g. paper towels naturally take approximately 1-2 weeks whereas a plastic bottle takes 100 years to biodegrade). At the extreme, non-biodegradable substances are entirely resistant to natural degradation processes - they are environmentally persistent and bioaccumulate.

So called Persistent Organic Pollutants (POPs) constitute a great environmental concern. These materials are widely resistant to natural degradation processes and are particularly toxic to living organisms (Schwarzenbach, Egli, Hofstetter, von Gunten, and Wehrli, 2010).[3] POPs are used in agriculture, manufacturing and industrial processes (e.g. fertilizers) but can also emerge unintentionally as by-products of industrial production (e.g. dioxins in textile production) (Križanec and Majcen Le Marechal, 2006).

Urban sewage mainly consists of "raw sewage" containing excrement and debris (e.g. sanitary towels or plastic). However, in developing and emerging economies the organic part of urban sewage like bacteria, parasites as well as viruses constitute the major health concern (Schwarzenbach, Egli, Hofstetter, von Gunten, and Wehrli, 2010). Diseases caused by the respective bacteria and viruses can involve gastro-enteritis, diarrhea, typhoid, cholera, but also respiratory diseases like the Acute Severe Respiratory Syndrome (SARS)

[2]The level of organic pollution is measured by the biochemical oxygen demand (BOD). Two counteracting effects determine the BOD: The organic pollution load and natural cleaning.

[3]The POP 'Dirty Dozen' are aldrin, chlordane, DDT, dieldrin, endrin, heptachlor, hexachlorobenzene, mirex, toxaphene, polychlorinated biphenyls (PCBs), dioxins and furans. POPs have been subject of two international environmental treaties, the Stockholm Convention on Persistent Organic Pollutants (2004) and the Aarhus Protocol (1998).

involving heavy cough as major symptom (Feachem, R. G.; Bradley, D. J.; Garelick, H.; Mara, 1983). An important cause of waterborne illness is the Escherichia coli bacterium that is commonly found in intestines of humans and animals. Five groups of pathogenic excreted viruses are particularly important: adenoviruses, enteroviruses (including poliovirus), hepatitis A virus, reoviruses and diarrhea causing viruses (especially rota virus).

In contrast, the dominant pollutants in industrial sewage are non-degradable pollutants such as POPs. They constitute around 95% of industrial effluents in Egypt (Dahshan, Megahed, Abd-Elall, Abd-El-Kader, Nabawy, and Elbana, 2016) and have hazardous effects on human health (WHO, 2017b). The most toxic are so called dioxins and dioxin-like compounds (Križanec and Majcen Le Marechal, 2006).[4] Dioxins are considered to have detrimental effects on the immune system making people more vulnerable to acute infections. Further, they can damage the gastrointestinal tract, organs and the reproductive system. In industrialized countries the use of many of these substances in production processes is forbidden and particular technologies are in place to destroy material containing POPs. Developing and emerging economies however often lack regulations and funds to pursue consequent environmental strategies to combat POPs. Industrial waste water may also contain harmful components of inorganic pollutants including heavy metals like lead, cadmium, mercury and arsenic. The related health effects are however expected to unfold after long-term exposure and often only emerge later in life. Given our research design we are not able to detect these effects.

4.3 Data

Our analysis is based on a panel data set of 78 river segments and 7 time periods between 1992 and 2014. The spatial dimension of the data set is based on equally spaced 10x10 square kilometer grids along the course of the Nile. We include grids, whose center is located within a radius of 30km from the river line. This captures the entire Nile valley and thereby the vast majority of Egypt's population as only few people live in the desert regions beyond the Nile valley. In addition we create a subsample, which only includes grids that intersect with the Nile. In Egypt the Nile runs relatively straight from its southern border with Sudan northwards into the Mediterranean Sea. Accordingly, we group grids by latitude. Employing this strategy, we obtain 78 horizontal river segments (see Figure 1). We exclude the Nile delta, as the river disperses into multiple arms so that the population may be affected by polluted water from several sources and a clear assignment is not

[4]The main groups of dioxins are polychlorinated dibenzo-p-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs). Polychlorinated biphenyl (PCB) are also chlorinated hydrocarbons with a similar structure as dioxin (Umweltbundesamt, 2018).

possible anymore. The temporal dimension of the data set is determined by the availability of DHS survey waves for Egypt. All, factory-, population density and household DHS data contain geographic coordinates. We use geospatial software to aggregate all data for each segment.[5]

To assess health outcomes of children we use data from the Demographic and Health Survey. The survey is conducted by USAID and collects detailed health and demographic data for a wide range of developing countries. For Egypt there are seven survey rounds available that contain geographic coordinates of the surveyed households for the years 1992, 1995, 2000, 2003, 2005, 2008 and 2014.[6] We employ data from the survey's 'Individual Recode' asking women about their own and their children's health. Accordingly, for data on children's health we use questions asking whether any child had diarrhea (H11), fever (H22) or cough (H31) during the two-week period before the survey. We construct binary variables (0 indicating that no child had the respective illness and 1 indicating that at least one child suffered from the illness) and combine them to a one-dimensional health index by taking the average of the three measures.[7]

We also construct a wealth index for each of these households (DHS only provides a wealth index from 2003 onwards). The DHS index is based on households' ownership of selected assets. We follow the same methodology and take the first principal component of the following survey items, which are covered in all surveys since 1992: type of toilet facility (dummy variables), possession of radio, TV, refrigerator, bicycle and electricity and type of floor material (dummy variables). Our newly constructed index highly correlates with the index provided by DHS for rounds 2003-2014 (correlation coefficient: 0.8). For ease of interpretation we normalize the measure to scale the range in [0, 1].

In addition we examine sub-samples with respect to the source of drinking water. The indicator variable turns 1 if drinking water is piped into dwelling and remains 0 if public taps, the Nile, wells and springs are the household's source of drinking water.

We thus obtain information on children's health and household characteristics for 2500

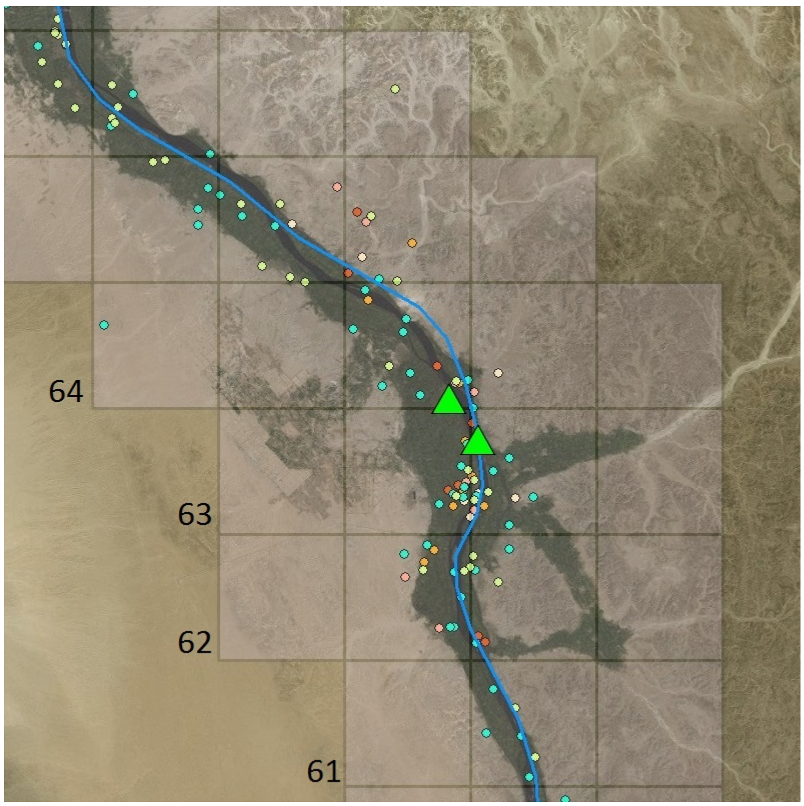
[5]Following, we use the terms segment and grid interchangeably, referring to the 78 horizontal segments along the Nile.

[6]In order to ensure respondent confidentiality, the longitude/latitude information are randomly displaced. "Urban clusters contain a minimum of 0 and a maximum of 2 kilometers of error. Rural clusters contain a minimum of 0 and a maximum of 5 kilometers of positional error with a further 1% of the rural clusters displaced a minimum of 0 and a maximum of 10 kilometers" (http://dhsprogram.com/What-We-Do/GPS-Data-Collection.cfm). Given that our segments span around 40 x 10 kilometers there should only be few cases in which a DHS cluster is mistakenly allocated to the previous or subsequent segment. In these cases we only introduce random noise and measurement error, which would bias our results towards zero.

[7]We also construct an alternative health index based on the first principal component. The correlation is 0.999 and we therefore use the former index, which is easier to interpret.

households in 1992 and up to 4400 households in 2014 along the Nile river grid, which amounts to 23,700 households in total. These households are grouped into 3200 clusters , which are georeferenced (the points in Figure 4.1 represent the clusters).[8] Finally, we aggregate the health data to the river segment level, which leaves us with data on children's health for 78 segments at 7 points in time. We observe on average 22 DHS clusters per year and river segment.[9]

Figure 4.1: Nile River Grid and DHS Cluster



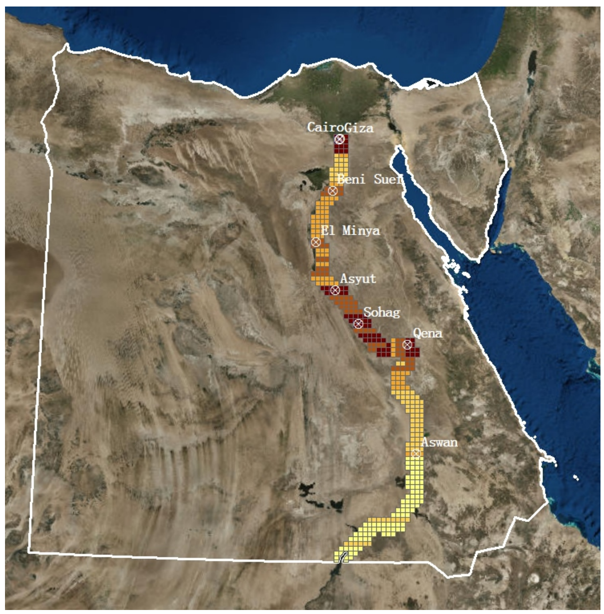
In our main specification we use population density data as a continuous measure for the level of urbanization in each grid. Data on population density on a 2.5 arc-minutes grid is obtained from the Gridded Population of the World Database (v3) provided by the Socioeconomic Data and Applications Center (SEDAC). The data provides estimates of population density based on counts consistent with national censuses and population registers in 5 year intervals. To match the years to our DHS waves, we take the value

[8]242 clusters in 1992, 433 clusters in 1995, 395 clusters in 2000, 391 clusters in 2003, 614 clusters in 2005, 530 clusters in 2008 and 724 clusters in 2014.

[9]The median number of grids per year and river segment is 9.

closest to the respective DHS year. We then also aggregate them to the river segment level. As the population is concentrated around the Nile and the fringes of the Nile valley are sparsely inhabited, the highest rather than average population density in each segment is the relevant measure to determine the extent of pollution. In an additional specification we use cities as binary measure for urbanization. Cities are defined by the "World Cities Database" and we focus on Egyptian cities with at least one million inhabitants along the Nile (ordered by total population: Cairo/Giza, Asyut, Aswan, Minya, Beni Suef, Quena, Sohag).

Figure 4.2: Nile River Grid



Finally, we use data on plant establishment to measure industrialization. Data on industrial plants comes from the "Plants Database" of Industrial Info Resources, a provider of global market intelligence. The database tracks the 368 most important industrial facilities in Egypt.[10] The data provides geographic coordinates as well as opening and closing dates, which enables us to conduct comparisons of health outcomes across both, space and time. To the best of our knowledge this is the first time this data is used for scientific purposes in the context of analyzing health effects of industrialization. We select all 65 plants, which are located within our grid. The industry types represented in our sample involve power plants, chemical processing, metals and minerals, and pulp, paper and wood production. The Nile provides approximately 65 percent of the industrial needs

[10]Plants are considered 'important' if they qualify for industry-specific criteria, e.g. mines with capacity of 250.000 tons per anum and greater.

Table 4.1: Summary Statistics

	Mean	SD	Min	Max	Obs
Health Index	0.72	0.12	0	1	459
Wealth Index	0.67	0.14	0	1	459
Population Density (in 10K per sqkm)	0.25	0.45	0	4	539
Factory (binary)	0.23	0.42	0	1	539
Cumulative Population Density (weighted)	0.06	0.05	0	0	539
Cumulative Factory Presence (weighted)	0.06	0.06	0	1	539
Population Share with Access to piped Water	0.73	0.24	0	1	459
Distance along Nile (from South to North)	481.03	286.03	0	921	539

for fresh water used in the production process and receives approximately 57 percent of industrial effluents. It has been shown that industries in our sample produce toxic wastewater with detrimental effects on human health (Megahed, Dahshan, Abd-El-Kader, Abd-Elall, Elbana, Nabawy, and Mahmoud, 2015; Balabanič, Filipič, Krivograd Klemenčič, and Žegura, 2017). An indicator variable denotes whether there existed a plant in a river segment at a certain point in time. Summary statistics on all variables of interest are provided in Table 4.1.

Figure 4.3: Factories Along the Nile in Egypt

4.4 Urbanization

4.4.1 Empirical Framework

To estimate the effect of urbanization on children's health we exploit variation in population density across space. Although we use our panel data set spanning seven time periods, the variation in population density mainly stems from spatial differences whereas there is relatively little idiosyncratic variation in population density over time. We estimate the effect of population density upstream of grid i on the health index in grid i. To control for neighboring population density with potential spill-over effects we also control for population density downstream of the respective grid. At the same time this serves as an important placebo check. According to our assumption that water pollution affects upstream and downstream population differentially we do not expect strong health effects for population living upstream of the pollutant. Population density in the same grid is an important control variable as it is correlated with population density in the previous and following grids and may affect health in grid i. We thus estimate the following model in order to identify the effect of upstream population density on health:

$$Health_{i,t} = \alpha + \beta Pop\ Density^{U}_{i,t} + \delta Pop\ Density^{D}_{i,t} + \sigma_r + \sigma_t + \epsilon_{i,t} \tag{4.1}$$

$Health_{i,t}$ denotes the health index in grid i at time t, $Pop\ Density^{U}_{i,t}$ is the population density in the grid directly upstream of grid i at time t and β therefore constitutes the coefficient of interest. $Pop\ Density^{D}_{i,t}$ is the population density downstream of grid i at time t. By including $Pop\ Density^{D}_{i,t}$ we are able to take advantage of the quasi-experimental setting that only upstream pollution affects health in grid i while downstream pollution should have no negative effect. Further, we use DHS wave fixed effects, σ_t, to capture common time trends in urbanization and health. We also include 10 region fixed effects, σ_r, to compare households within the same subnational administrative unit, which are exposed to the same institutional, economic, and cultural environment. In additional specifications we include distance along the Nile and population density in grid i as control variables. Distance along the Nile is measured from the southern border of Egypt and constitutes an important control variable that captures linear trends along the Nile such as downstream increases in cumulative pollution or factors related to the distance from Cairo. Finally, we also show that our results are robust to including region-year fixed effects to control for all region-specific changes over time such as changes in health legislation or enforcement of regulations. This specification uses the variation in population density as pollutant most efficiently as it captures the localized effects. Section 4.6 presents results for a cumulative measure of population density.

The correlation between population density in two contiguous grids is however rather high (0.72) and therefore multicollinearity may be an issue when estimating OLS. The coefficients would not be biased but standard errors tend to be very high in the presence of multicollinearity. Estimates of variance inflation factors for the estimates for $Pop\ Density^{D}_{i,t}$ and $Pop\ Density^{U}_{i,t}$ are however below 5 and thus do not reveal evidence for excessive multicollinearity. In the baseline specification we cluster standard errors at the grid level. To more precisely account for both, cross-sectional spatial correlation and location-specific serial correlation we conduct robustness checks using Conley standard errors with a spatial HAC correction (Conley, 1999, 2008).[11]

Our main identifying assumption is that upstream population density is uncorrelated with unobserved factors that are correlated with health and that affect the down- and upstream population differentially. We use data on upstream polluting behavior rather than data on local pollution levels and are therefore not relying on correlating local pollution with local health outcomes - in fact we even control for local population density and thereby capture all factors associated with both local level of urbanization and local health.

The inclusion of downstream population density as a placebo check rules out all concerns related to urbanization effects besides water pollution (such as working conditions in cities, health care provision, air pollution, risk of contagion). If water pollution is a relevant channel downstream population density should not affect health. A remaining concern threatening our identifying assumption is sorting. Particularly, poorer people may sort into more polluted areas and accordingly settle downstream of high population densities. To address this concern we estimate the effect of both, upstream and downstream population density on wealth in order to check for systematic differences.

4.4.2 Results

We find that upstream population density has a significantly negative effect on health. An increase in population density by 10,000 inhabitants per sqkm, which corresponds to 2 standard deviations, decreases the health index by roughly 5 percentage points, which corresponds to half a standard deviation of the health index (see Table 4.2). This increase corresponds to one additional disease in every 6th household.[12] The effect remains statis-

[11]Spatial autocorrelation is assumed to linearly decrease with distance up to a cutoff of 100 km and we account for serial correlation across all seven time periods.

[12]The health index is an average over the number of diseases (diarrhea, cough and fever) per household, which is then averaged across all households in the grid. Accordingly, if the health index increases from 0 to 1 all three diseases will occur in all households within the grid. If the health index increases by 1/3 there will be on average one additional disease in every household. As the health index increases by 0.05, there will be on average 1/6 additional disease in every household or in other words: one more disease in

Table 4.2: Effects of Urbanization on Health

	(1) Health	(2) Health	(3) Health
Upstream Pop Density	-0.0428***	-0.0455***	-0.0506***
	(0.0130)	(0.0116)	(0.0124)
Downstream Pop Density	0.0260	0.0272*	0.0296
	(0.0174)	(0.0154)	(0.0180)
Year FE	✓	✓	✓
Region FE	✓	✓	✓
Controls		✓	✓
Year x Region FE			✓
N	450	450	450
Cluster	74	74	74
Mean DV	0.72	0.72	0.72

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Pop Density is population density in the grid upstream of grid i. Downstream Pop Density is population density in the grid downstream of grid i. Controls include distance along Nile and population density in grid i. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

tically significant when including distance along the Nile and population density in grid i as control variables (Column (2)) and also when adding year-specific region fixed effects (Column (3)). These results are also robust to using Conley spatial HAC standard errors, which account for both, spatial- and temporal correlation of the standard errors (see Table 4.A1). The negative effect of downstream population density on health is significantly different from the (positive) effect of upstream population density (p-value of 0.00). Hence, the negative health effects are unique to the downstream population. This indicates that population density affects health through water pollution as it is tied to the direction of the river flow.

In order to further substantiate water pollution as a driver of the observed negative health effects on the downstream population, we take into account the households' source of drinking water. For this analysis we calculate the fraction of households within a grid that have access to piped water and split the sample at the mean of this variable (75%). We find that the effect of upstream population density is almost twice as large in grids where drinking water of less than 75% of households comes from clean sources (these households use public taps, the Nile, wells and springs) as opposed to grids where more than 75% of households have access to clean water (Table 4.3). The negative health effect

every 6th household.

Table 4.3: Effect Heterogeneity: The Role of Piped Water

	Health	
	(1) Piped Water<0.75	(2) Piped Water>0.75
Upstream Pop Density	-0.0796***	-0.0416**
	(0.0155)	(0.0159)
Year FE	✓	✓
Region FE	✓	✓
Controls	✓	✓
N	177	270
Cluster	58	73
Mean DV	0.71	0.73

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Pop Density is population density in the grid upstream of grid i. Controls include distance along Nile and population density. The sample is split according to the fraction of households with access to piped water. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

is considerably mitigated by the access to clean drinking water. This again points to water being the crucial link between pollutants and diseases.

4.4.3 Robustness

To address sorting as a potential concern we analyze whether there are differential effects of urbanization on wealth given the relative location of the household to the pollutant. In fact we replicate our analysis from the previous section using wealth as outcome. We find that both, downstream and upstream population density have positive, albeit not statistically significant effects on wealth (see Table 4.4). The upstream and downstream coefficients in Table 4.4, do not differ significantly from each other (p-value of 0.84). This shows that there is no heterogeneity in wealth outcomes depending on the geographic location of households indicating that sorting between up- and downstream locations based on wealth is unlikely. Hence, more polluted water downstream of agglomerations does not prevent wealthy people from living there (unlike air pollution caused by industrial activity in 19th century Britain (Heblich, Trew, and Zylberberg, 2016).
We also include wealth as control variable in an additional specification (Table 4.A2 in Appendix) because it may be an omitted variable in the main specification. The results are robust to the inclusion of this additional control variable. This however introduces bad control bias (Angrist J. D. and J. S. Pischke, 2008) as wealth itself is an outcome of the treatment.

Table 4.4: Effects of Urbanization on Wealth

	(1) Wealth	(2) Wealth	(3) Wealth
Upstream Pop Density	0.0295	0.0346*	0.0255
	(0.0449)	(0.0178)	(0.0206)
Downstream Pop Density	0.0468	0.0256	0.0350
	(0.0444)	(0.0242)	(0.0280)
Year FE	✓	✓	✓
Region FE	✓	✓	✓
Controls		✓	✓
Year x Region FE			✓
N	450	450	450
Cluster	74	74	74
Mean DV	0.72	0.72	0.72

Wealth is a wealth index, which lies between 0 and 1 and is averaged over all households living in grid i. It is based on households' ownership of selected assets. Upstream Pop Density is population density in the grid upstream of grid i. Downstream Pop Density is population density in the grid downstream of grid i. Controls include distance along Nile and population density in grid i. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

As additional robustness check we estimate the same model using only grids that intersect with the Nile river rather than grids within a 30km radius.[13] We expect the effect to be stronger for people living extremely close to the river as opposed to people who live on the fringe of the Nile valley, close to the desert. The disadvantage of using this subsample is however that we observe fewer DHS clusters in each segment (though the most affected ones).[14] Table 4.A3 in the Appendix shows that the negative effect of population density on downstream health is indeed stronger than for the main sample (the effect size is 30% - 40% larger). In contrast, the effect on the health of the upstream population is significantly positive in all specifications estimated for this subsample. These findings stress the asymmetric health effects of population density.

Finally, we focus on large and densely populated cities as the most severe cases of concentrated pollution. To do so, we restrict our sample to grids surrounding large cities, excluding all other observations. We thus only compare grids located downstream of a city to grids located upstream of the same city rather than comparing all grids in the region.. The downside of this approach is the limited number of observations given that there are

[13]The number of river segments (grids grouped by latitude) is still similar. The sample size only reduces by 26 segment-year cells due to missing DHS observations).

[14]We here observe on average 16 DHS cluster per segment-year cell as opposed to 20 in the main sample.

Table 4.5: Case Study: Effect of Cities on Health

	(1) Health
City Upstream	-0.0413*
	(0.0182)
Year FE	✓
City FE	✓
Controls	✓
N	112
SE	6 City Cluster

The sample is restricted to grids that are located directly upstream or downstream of a city. Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. City Upstream is a binary variable, which indicates whether a city is located upstream of grid i (=1) or whether a city is located downstream of grid (=0). Controls include distance along Nile and population density. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

only 6 major cities along the Nile.[15]

We estimate the following model:

$$Health_{i,t} = \alpha + \beta City_i^U + \sigma_c + \sigma_t + \mathbf{X}_{i,t} + \epsilon_{i,t} \tag{4.2}$$

where $Health_{i,t}$ denotes health in grid i at time t. $City_{i,t}^U$ indicates whether the grid is "treated", this indicator variable turns 1 if the city is located upstream of grid i and zero if the city is located downstream of gird i. σ_c denotes city cluster fixed effects. Each of these clusters contains three grids: the grid where the city is located, the upstream (control) and the downstream (treatment) grid. The city cluster fixed effects thus ensure that we only compare grids around the same city. σ_t denotes time fixed effects and $X_{i,t}$ is a vector of control variables for grid i, including distance along the Nile and population density.

[15] We exclude Cairo from the regression because we cannot specify a unique downstream grid. The city is located at the end of the river stream and spreads across the Nile Delta, which we have excluded from the analysis.

We find that children living downstream of a city are significantly sicker than children living upstream of the same city. Specifically, their health index is 4 percentage points lower, which corresponds to 0.4 standard deviations of the health index for this sample. This provides further evidence for asymmetric health effects of agglomerations. Children living downstream of these pollutants exhibit significantly worse health outcomes than children living upstream.

4.5 Industrialization

4.5.1 Empirical Framework

In this section we analyze the impact of water pollution caused by industrial plant openings close to the Nile on children's health. Information on the opening date of industrial plants allows us to exploit temporal variation in addition to spatial variation to estimate the health effect of industrial water pollution. Applying a difference-in-differences strategy we compare health changes of children living downstream of a factory with health changes of children in unaffected grids.[16] We estimate the following model:

$$Health_{i,t} = \alpha + \beta Factory^{U}_{i,t} + \sigma_i + \sigma_t + \epsilon_{i,t} \tag{4.3}$$

$Health_{i,t}$ denotes health in grid i at time t. $Factory^{U}_{i,t}$ indicates whether a factory is located in the upstream grid at time t and thus constitutes a binary treatment variable. σ_i are grid fixed effects, which account for grid specific factors, as for instance potential systematic differences between northern and southern areas.[17] σ_t denotes time fixed effects, which account for health trends over time. $\epsilon_{i,t}$ is the error term and standard errors are clustered on a grid level.

In an additional specification we include population density as a control variable in order to account for changes in population density caused by the opening of a factory, which at the same time may influence health.[18] We also apply a placebo check by estimating the effect of a downstream factory on health.

Since we control for all grid-specific and time-specific factors the only threat to identification would be an event that occurred simultaneously with the factory opening and

[16] The variation in this set-up stems from 31 factories (out of the 65), which opened between 1992 and 2014.

[17] Here we can exploit temporal variation within the grid and are therefore able to include gird fixed effects as opposed to our specification in Section 4.4.1.

[18] It becomes obsolete to control for distance along the Nile as this variable is time-invariant and is captured by grid FE.

Table 4.6: Effects of Industrialization on Health

	(1) Health	(2) Health
Upstream Factory	-0.0763*	-0.0751*
	(0.0429)	(0.0421)
Downstream Factory	0.00857	0.0113
	(0.0430)	(0.0423)
Year FE	✓	✓
Grid FE	✓	✓
Pop Density		✓
N	450	450
Cluster	74	74
Mean DV	0.72	0.72

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Factory indicates whether a factory is located in the grid upstream of grid i. Downstream Factory indicates whether a factory is located in the grid downstream of grid i. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Control for population density is included.

affects downstream and upstream health differentially. As this seems rather unlikely, we are confident that water pollution caused by the factory contributes to poor health outcomes.

4.5.2 Results

We find that a factory located upstream affects children's health negatively. The effect is however only marginally significant (see Table 4.6, Columns (1) and (2)). The effect appears to be quantitatively sizable - the opening of an upstream factory decreases the health index by roughly 8 percentage points, which corresponds to roughly 0.7 standard deviations of the dependent variable. In other words, the factory opening leads to an additional child disease in every 4th household. Contrary, we find that a downstream factory has no effect on health. These results are also robust to using Conley spatial HAC standard errors, which account for both spatial and temporal correlation of the standard errors (see Table 4.A4).

In order to further examine water pollution as driver of the negative health effects we again split the sample into grids with less than 75% of households with access to piped water and grids where more than 75% of households have access to piped water. We find

Table 4.7: Effect Heterogeneity: The Role of Piped Water

	Health	
	(1) Piped Water<0.75	(2) Piped Water>0.75
Upstream Factory	-0.101***	-0.0450
	(0.0257)	(0.0574)
Year FE	✓	✓
Grid FE	✓	✓
N	177	277
Cluster	58	74
Mean DV	0.71	0.73

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Factory indicates whether a factory is located in the grid upstream of grid i. Controls include distance along Nile and population density. The sample is split according to the fraction of households with access to piped water within each grid. Controls include distance along Nile and population density. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

that the effect of an upstream factory is highly significant and twice as large for the group with less than 75% of households with access to piped water and turns insignificant for grids where more than 75% of households use piped water as drinking water. This demonstrates that children who have contact with industrial wastewater are more likely to fall sick than children who have access to clean drinking water.

4.5.3 Robustness

To tackle the potential concern of sorting due to polluting factories we examine the effects on wealth. Table 4.8 shows that there are no statistically significant effects of factories (both upstream and downstream) on wealth. The coefficients on upstream and downstream factories are also not statistically different from each other (p-value: 0.49), which suggests that there is no sorting based on wealth around factories.

One might argue that control grids may in fact be "treated" by a factory two or three grid further upstream. Even though this would only bias our coefficients towards zero we address this issue twofold. First, to account for accumulating pollution we construct a cumulative measure (see Section 4.6). Second, we construct factory clusters to better distinguish between treated and control grids. We therefore generate a new sample where geographically close factory grids are defined as factory clusters. We define grids that are located downstream of the entire factory cluster as treated grids and grids, which are

Table 4.8: Effects of Industrialization on Wealth

	(1) Wealth	(2) Wealth
Upstream Factory	-0.0374	-0.0399
	(0.0279)	(0.0263)
Downstream Factory	-0.0119	-0.0171
	(0.0160)	(0.0150)
Year FE	✓	✓
Grid FE	✓	✓
Pop Density		✓
N	450	450
Cluster	74	74
Mean DV	0.72	0.72

Wealth is a wealth index, which lies between 0 and 1 and is averaged over all households living in grid i. It is based on households' ownership of selected assets. Upstream Factory indicates whether a factory is located in the grid upstream of grid i. Downstream Factory indicates whether a factory is located in the grid downstream of grid i. Control for population density is included. Standard errors (clustered by grids) in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.9: Case Study: Effect of Factory Clusters on Health

	(1) Health
Factory Upstream	-0.0926* (0.0552)
Year FE	✓
Factory Cluster	✓
Controls	✓
N	80
Mean DV	0.73

Sample consists only of grids, which are directly located upstream or downstream of an industry cluster (consecutive grids that are characterized as industrial area). Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Factory Upstream indicates whether a factory cluster is located upstream of grid i (=1) or whether the factory cluster is located downstream of grid i (=0). Controls include population density and distance along Nile. Standard errors (in parentheses) are bootstrapped, *** p<0.01, ** p<0.05, * p<0.1.

located upstream of the cluster as control grids. In 2014, we for example observe six factory clusters consisting of 3 to 16 individual grids (as opposed to 22 individual factory-grids). Instead of grid fixed effects we include factory cluster fixed effects and thereby directly compare children living upstream to children living downstream of the same factory cluster.[19] Due to the small number of observations we bootstrap the standard errors.

We again find that children living downstream of a factory are significantly sicker than children living upstream. Specifically, their health index is around 9 percentage points lower, which corresponds to roughly 0.8 standard deviations of the dependent variable

[19]For each factory cluster we observe health in two grids (upstream and downstream) over 8 years. As these clusters change over time (due to factory openings and closings) we only observe 5 of these clusters over the entire sample period.

(Table 4.9, Column (1)). This provides further evidence for adverse health effects of the presence of industrial plants.

4.6 Cumulative Pollution

Finally, we shift the focus from localized pollution to cumulative pollution along the Nile. Here we analyze the aggregate effects of population density and factory presence over all grids located upstream. To do so, we weight population density and factory presence respectively by the inverse of the distance to grid i. This weighting accounts for cumulative pollution as people living downstream are not only exposed to pollution from the previous grid but also to water pollution originating in grids further upstream.

$$Weighted\ Population\ Density_i = \sum_{j=1}^{78} \frac{1}{Distance_{i,j}} * Population\ Density_j * I(j < i)$$

The mean of the weighted population density is 0.06 and the standard deviation 0.05. We applied the same formula to calculate the weighted factory presence, which has a mean of 0.07 and a standard deviation of 0.06. We assume linear degradation of pollutants in this setting. In an alternative specification we weight by the inverse of the quadratic distance, which gives closer grids a higher weight as compared to linear weighting. We include this measure as regressors in specification 4.1 (see Section 4.4.1).

We do not find an effect of the cumulative population density measure on health, irrespective of the weighting function (Table 4.10, Columns (1) and (2)). While we detect negative health effects of population density in the adjacent upstream grid (Section 4.4.2) we do not identify a statistically significant effect of cumulative population density. This finding suggests that the negative health effect of population density is localized. The effects of the cumulative factory measure are in turn highly statistically significant (Table 4.10, Columns (3) and (4)). If a factory opens 10km upstream (increase of weighted factory measure by 0.1) the health index decreases by 0.02 points which corresponds roughly to 0.2 standard deviations of the health index. The results reflect a key difference between the two types of pollution in terms of degradability. While organic pollutants in urban waste are subject to natural cleaning and dilution, large fractions of industrial pollution are not. In line with this pattern we observe that the effects of urbanization decline more rapidly with distance to the pollutant than the effect of industrial pollution on human health.[20] Finally, we also include both sources of pollution in the specification at the same time and

[20]Depending on industry type the scale of biodegradability can vary, e.g. slaughter houses contain higher extend of biodegradable components whereas chemical plants or pulp and paper mills contain higher content of persistent organic pollutants.

find that the effect of cumulative factory presence remains statistically significant negative while the effect of cumulative population density remains insignificant.

4.7 Conclusion

Using a newly constructed, geo-coded data set for Egypt we find detrimental effects of urban and industrial water pollution on health. Our research design exploits the direction of the river flow as natural experiment, where population density and industrial plant location constitute the sources of urban and industrial pollution.

We find a strong negative impact on the health of children who live downstream of an agglomeration. Children living upstream are not negatively affected. This heterogeneity of health outcomes suggests that urbanization affects health through water pollution. The opening date of industrial plants allows us to employ a difference-in-differences strategy to analyze the health effect of industrialization. Here we also find a significantly negative effect on the health of children living downstream of a factory while children living upstream are again not affected.

The health effects of urbanization and industrialization are strongest for grids where a significant proportion of households uses untreated water as source of drinking water. This finding substantiates the argument that the water transmits diseases caused by the pollutants. Finally we show that the cumulative effect of factories along the course of the Nile is significant whereas we do not find a cumulative effect of population density. This finding reflects the difference between these two different sources of pollution in terms of degradability. While organic pollutants in urban waste dilute, large fractions of the industrial pollutants are persistent. The health effects of water pollution caused by urbanization are therefore expected to decline more rapidly with distance to the pollutant than the effect of water pollution caused by factories.

The findings of this study have important implications for policy in developing countries. In order to improve health outcomes it is important to put environmental regulations in place and to enforce them. This is particularly important for countries that are industrializing rapidly and are positioned on the polluting trajectory of the Environmental Kuznet Curve[21]. Weak political and legal institutions are however a stumbling block to regulation. It is thus important to fight corruption, promote law enforcement and increase awareness for sustainable growth. In addition the international community has to support developing countries in these efforts as they often do not have the resources and technologies needed for environmental protection.

In Egypt, various authorities are in charge of water management and quality control,

[21]It illustrates that over the course of development pollution first increases and then decreases again.

Table 4.10: Distance Weighted Treatments

	(1) Health	(2) Health	(3) Health	(4) Health	(5) Health	(6) Health
Pop Dens (weighted by distance)	-0.190				-0.0994	
	(0.153)				(0.149)	
Pop Dens (weighted by distance^2)		-1.542				-0.904
		(2.160)				(2.412)
Factory Presence (weighted by distance)			-0.231***		-0.222***	
			(0.0712)		(0.0714)	
Factory Presence (weighted by distance^2)				-0.804**		-0.745***
				(0.309)		(0.267)
Year FE	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
R^2	0.29	0.29	0.29	0.29	0.29	0.29
N	459	459	459	459	459	459
Cluster	76	76	76	76	76	76
Mean DV	0.72	0.72	0.72	0.72	0.72	0.72

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. The weights on population density and factory presence are constructed as the inverse of the linear distance and quadratic distance respectively. Pop Dens (weighted by distance or distance^2) and Factory Presence (weighted by distance or distance^2) constitute cumulative measures of pollution, which are based on all grids that are located upstream of grid i. Control for distance along Nile is included. Standard errors (clustered by grids) in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

which leads to ambiguous responsibilities.[22] Water quality thus has to become a priority for the Egyptian government and it has to assign responsibilities clearly.

Moreover, adverse health effects of urbanization have to be taken into account for city planning in developing countries. The direction of the river flow determines who is affected most by water pollution and thereby leaves people living downstream worse off. As long as there is no environmental protection in place governments have to work on improved sanitation in these areas. Institutional changes are thus key to reduce the detrimental health effects of industrialization and urbanization in emerging economies.

[22]"The absence of a single administrative body in charge of water management and quality improvement from the High Dam to the riverbed, and up to the point where it [water] is delivered to people's homes, is the reason behind water pollution in Egypt." (Kareem Khaled, 2015).

4.A Appendix

Table 4.A1: Effect of Urbanization on Health using Conley spatial HAC Standard Errors

	(1) Health	(2) Health	(3) Health
Upstream Pop Density	-0.0428	-0.0455*	-0.0457*
	(0.0275)	(0.0275)	(0.0275)
Downstream Pop Density	0.0260	0.0272	0.0274
	(0.0275)	(0.0274)	(0.0276)
Year FE	✓	✓	✓
Region FE	✓	✓	✓
Controls		✓	✓
Region x Year FE			✓
N	450	450	450

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Pop Density is population density in the grid upstream of grid i. Downstream Pop Density is population density in the grid downstream of grid i. Controls include distance along Nile and population density. Conley spatial HAC standard errors estimated, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.A2: Including Wealth as Control

	(1) Health	(2) Health	(3) Health
Upstream Pop Density	-0.0412*** (0.0118)	-0.0437*** (0.0116)	-0.0499*** (0.0122)
Downstream Pop Density	0.0286* (0.0160)	0.0285* (0.0149)	0.0307* (0.0179)
Year FE	✓	✓	✓
Region FE	✓	✓	✓
Controls		✓	✓
Year x Region FE			✓
N	450	450	450
Cluster	74	74	74
Mean DV	0.72	0.72	0.72

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Pop Density is population density in the grid upstream of grid i. Downstream Pop Density is population density in the grid downstream of grid i. All specifications include wealth in grid i as control. Other controls include distance along Nile and population density. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 4.A3: Subsample: Only Grids intersecting with Nile

	(1) Health	(2) Health	(3) Health
Upstream Pop Density	-0.0599***	-0.0571***	-0.0649***
	(0.0121)	(0.0129)	(0.0126)
Downstream Pop Density	0.0384**	0.0387***	0.0412***
	(0.0156)	(0.0133)	(0.0147)
Year FE	✓	✓	✓
Region FE	✓	✓	✓
Controls		✓	✓
Year x Region FE			✓
N	424	424	424
Cluster	71	71	71
Mean DV	0.72	0.72	0.72

Data for segments along the Nile is only based on grids that intersect with NIle river (information from grids that are located close to the desert is excluded). Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Pop Density is population density in the grid upstream of grid i. Downstream Pop Density is population density in the grid downstream of grid i. Controls include distance along Nile and population density. Standard errors (clustered by grids) in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 4.A4: Effect of Industrialization on Health using Conley spatial HAC Standard Errors

	(1) Health	(2) Health
Upstream Factory	-0.0763*	-0.0751*
	(0.0410)	(0.0405)
Downstream Factory	0.00857	0.0113
	(0.0517)	(0.0510)
Year FE	✓	✓
Grid FE	✓	✓
Pop Density		✓
N	450	450

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Factory indicates whether a factory is located in the grid upstream of grid i. Downstream Factory indicates whether a factory is located in the grid downstream of grid i. Population density in grid i is included as control variable in column (2). Conley spatial HAC standard errors estimated, *** p<0.01, ** p<0.05, * p<0.1.

Table 4.A5: Including Wealth as Control

	(1) Health	(2) Health
Upstream Factory	-0.0759* (0.0410)	-0.0752* (0.0425)
Downstream Factory		0.0112 (0.0424)
Year FE	✓	✓
Grid FE	✓	✓
Pop Density	✓	✓
N	457	450
Cluster	75	74
Mean DV	0.72	0.72

Health is the share of households in each grid where no child suffered from diarrhea, fever or cough in the past 14 days. It lies between 0 and 1. Upstream Factory indicates whether a factory is located in the grid upstream of grid i. Downstream Factory indicates whether a factory is located in the grid downstream of grid i. All specifications include wealth and population density in grid i as controls. Standard errors (clustered by grids) in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Bibliography

Abdel-Satar, A. M., M. H. Ali, and M. E. Goher (2017): "Indices of Water Quality and Metal Pollution of Nile River, Egypt," *Egyptian Journal of Aquatic Research*, 43(1), 21–29.

Acemoglu, D., S. Johnson, and J. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91(5), 1369–1401.

Adams, C., T. Ide, J. Barnett, and A. Detges (2018): "Sampling Bias in Climate-Conflict Research," *Nature Climate Change*, 8, 200–203.

Ali, E. M., S. A. Shabaan-dessouki, A. R. I. Soliman, and S. E. Shenawy (2014): "Characterization of Chemical Water Quality in the Nile River, Egypt," *International Journal of Pure & Applied Bioscience*, 2(3), 35–53.

Almer, C., J. Laurent-Lucchetti, and M. Oechslin (2017): "Water Scarcity and Rioting: Disaggregated Evidence from Sub-Saharan Africa," *Journal of Environmental Economics and Management*, 86, 193–209.

Angrist J. D. and J. S. Pischke (2008): *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton University Press, Princeton.

Balabanič, D., M. Filipič, A. Krivograd Klemenčič, and B. Žegura (2017): "Raw and Biologically Treated Paper Mill Wastewater Effluents and the Recipient Surface Waters: Cytotoxic and Genotoxic Activity and the Presence of Endocrine Disrupting Compounds," *Science of the Total Environment*, 574, 78–89.

Barro, R., and J.-W. Lee (2013): "A New Data Set of Educational Attainment in the World, 1950-2010," *Journal of Development Economics*, 104, 184–198.

Bégué, A., E. Vintrou, D. Ruelland, M. Claden, and N. Dessay (2011): "Can a 25-year Trend in Soudano-Sahelian Vegetation Dynamics be Interpreted in Terms of

Land Use Change? A Remote Sensing approach," *Global Environmental Change*, 21(2), 413–420.

Benjaminsen, T. A., K. Alinon, H. Buhaug, and J. T. Buseth (2012): "Does Climate Change Drive Land-Use Conflicts in the Sahel?," *Journal of Peace Research*, 49(1), 97–111.

Benjaminsen, T. A., and B. Ba (2009): "Farmer-Herder Conflicts, Pastoral Marginalisation and Corruption: A Case Study from the Inland Niger Delta of Mali," *Geographical Journal*, 175(1), 71–81.

Bosetti, V., C. Cattaneo, and G. Peri (2018): "Should They Stay or Should They Go? Climate Migrants and Local Conflicts," *National Bureau of Economic Research Working Paper*, 24447.

Brainerd, E., and N. Menon (2014): "Seasonal Effects of Water Quality: The Hidden Costs of the Green Revolution to Infant and Child Health in India," *Journal of Development Economics*, 107, 49–64.

Breckner, M., F. Englmaier, T. Stowasser, and U. Sunde (2016): "Resilience to natural disasters — Insurance penetration, institutions, and disaster types," *Economics Letters*, 148, 106–110.

Brzoska, M., and C. Froehlich (2016): "Climate Change, Migration and Violent Conflict: Vulnerabilities, Pathways and Adaptation Strategies," *Migration and Development*, 5(2), 190–210.

Buhaug, H. (2010a): "Climate Not to Blame for African Civil Wars," *Proceedings of the National Academy of Sciences*, 107(38), 16477–16482.

——— (2010b): "Reply to Burke et al.: Bias and Climate War Research," *Proceedings of the National Academy of Sciences*, 107(51), E186–E187.

Buhaug, H., H. Hegre, and H. Strand (2010): "Sensitivity Analysis of Climate Variability and Civil War," *PRIO Paper*.

Burke, Miguel, Satyanath, Dykema, and Lobell (2009): "Warming Increases the Risk of Civil War in Africa," *Proceedings of the National Academy of Sciences*, 106(49), 20670–20674.

Burke, M., J. Dykema, D. Lobell, E. Miguel, and S. Satyanath (2010): "Climate and Civil War: Is this Relationship Robust?," *National Bureau of Economic Research Working Paper*, 16440.

Burke, M., and K. Emerick (2016): "Adaptation to Climate Change: Evidence from US Agriculture," *American Economic Journal: Economic Policy*, 8(3), 106–140.

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

Burke, M., S. M. Hsiang, and E. Miguel (2015b): "Global Non-Linear Effect of Temperature on Economic Production," *Nature*, 527(7577), 235–239.

Burke, M. B., E. Miguel, S. Satyanath, J. A. Dykema, and D. B. Lobell (2010): "Climate Robustly Linked to African Civil War," *Proceedings of the National Academy of Sciences*, 107(51), E185–E185.

Cavallo, E., S. Galiani, I. Noy, and J. Pantano (2013): "Catastrophic Natural Disasters and Economic Growth," *Review of Economics and Statistics*, 95(5), 1549–1561.

Cervellati, M., E. Esposito, U. Sunde, and S. Valmori (2017): "Malaria Risk and Civil Violence," *CESifo Working Paper Series*, 6413.

CIESIN (2016): "Gridded Population of the World, GPWv3 and GPWv4: Population Density Grid," *Center for International Earth Science Information Network (CIESIN) - Columbia University. NY: NASA Socioeconomic Data and Applications Center (SEDAC)*.

Collier, P., and A. Hoeffler (1998): "On Economic Causes of Civil War," *Oxford Economic Papers*, 50(4), 563–573.

——— (2002): "On the Incidence of Civil Wars in Africa," *Journal of Conflict Resolution*, 46(1), 13–28.

——— (2004): "Greed and Grievance in Civil War," *Oxford Economic Papers*, 56, 563–595.

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

——— (2008): "Spatial Econometrics," in *The New Palgrave Dictionary of Economics*, ed. by S. Durlauf, and L. Blume. Palgrave Macmillan, London, 2 edn.

Couttenier, M., and R. Soubeyran (2014): "Drought and Civil War in Sub-Saharan Africa," *Economic Journal*, 124(575), 201–244.

Crosby, A. (1986): *Ecological Imperialism: The Biological Expansion of Europe, 900-1900*. Cambridge University Press, New York.

Crost, B., J. Felter, D. Rees, B. Crost, C. Duquennois, J. Felter, and D. Rees (2015): "Climate Change, Agricultural Production and Civil Conflict: Evidence from the Philippines," *IZA Discussion Papers*, 8965.

Currie, J., J. G. Zivin, J. Mullins, and M. Neidell (2014): "What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?," *Annual Review of Resource Economics*, 6(1), 217–247.

Dahshan, H., A. M. Megahed, A. M. M. Abd-Elall, M. A.-G. Abd-El-Kader, E. Nabawy, and M. H. Elbana (2016): "Monitoring of Pesticides Water Pollution. The Egyptian River Nile.," *Journal of Environmental Health Science and Engineering*, 14(1), 15.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. Mcnally, B. M. Monge-Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. N. Thépaut, and F. Vitart (2011): "The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System," *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

Dell, M., B. Jones, and B. Olken (2012): "Climate Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, 4(3), 66–95.

Dell, M., B. F. Jones, and B. A. Olken (2014): "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, 52(3), 740–798.

Diamond, J. (1999): *Guns, Germs, and Steel: The Fates of Human Societies*. W.W. Norton & Co, New York.

Duflo, E., and R. Pande (2007): "Dams," *The Quarterly Journal of Economics*, 122(2), 601–645.

Ebenstein, A. (2012): "The Consequences of Industrialization : Evidence From Water Pollution and Digestive Cancers in China," *The Review of Economics and Statistics*, 94(1), 186–201.

El-Ayouti, A., and H. Abou-Ali (2013): "Spatial Heterogeneity of the Nile Water Quality in Egypt," *Journal of Environmental Statistics*, 4(8), 1–12.

ESA (2017): "ESA annual LC maps v2.0.7," http://maps.elie.ucl.ac.be/CCI/viewer/ (Accessed March 2016).

ESRI Garmin International (2016): "World Waterbodies," https://www.arcgis.com/home/item.html?id=e750071279bf450cbd510454a80f2e63 (Accessed December 2017).

Exenberger, A., and A. Pondorfer (2013): "Climate Change and the Risk of Mass Violence: Africa in the 21st Century," *Peace Economics, Peace Science and Public Policy*, 19(3), 381–392.

FAO (2007): "Gridded Livestock of the World, by G. Wint and T.Robinson," Rome, 131.

Feachem, R. G.; Bradley, D. J.; Garelick, H.; Mara, D. D. (1983): *Sanitation and Disease: Health Aspects of Excreta and Wastewater Management.* John Wiley and Sons, Chichester.

Feenstra, R., R. Inklaar, and M. Timmer (2015): "The Next Generation of the Penn World Table," *American Economic Review*, 105(10), 3150–3182.

Felbermayr, G., and J. Gröschl (2014): "Naturally Negative: The Growth Effects of Natural Disasters," *Journal of Development Economics*, 111, 92–106.

Fjelde, H., and N. Uexkull (2012): "Climate Triggers: Rainfall Anomalies, Vulnerability and Communal Conflict in Sub-Saharan Africa," *Political Geography*, 31, 444–453.

Fomby, T., Y. Ikedab, and N. Loayza (2013): "The Growth Aftermath of Natural Disasters," *Journal of Applied Econometrics*, 28, 412–434.

Galiani, S., P. Gertler, and E. Schargrodsky (2005): "Water for Life: The Impact of the Privatization of Water Services on Child Mortality," *Journal of Political Economy*, 113(1), 83–120.

Gallup, J. L., J. D. Sachs, and A. D. Mellinger (1999): "Geography and Economic Development," *International Regional Science Review*, 22(2), 179–232.

Galor, O., and Ö. Özak (2016): "The Agricultural Origins of Time Preference," *American Economic Review*, 106(10), 3064–3103.

Garg, T., S. E. Hamilton, J. P. Hochard, E. M. Plous, and J. Talbot (2016): "(Not So) Gently Down The Stream: River Pollution and Health in Indonesia," *GRI Working Paper Series*.

Giannini, A., M. Biasutti, and M. Verstraete (2008): "A Climate Model-Based Review of Drought in the Sahel: Desertification, the Re-Greening and Climate Change," *Global and Planetary Change*, 64(3), 119–128.

Gleditsch, N. P. (2012): "Whither the Weather? Climate Change and Conflict," *Journal of Peace Research*, 49(1), 3–9.

Greenstone, M., and R. Hana (2014): "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India," *American Economic Review*, 104(10), 3038–3072.

Guha-Sapir, D., R. Below, and P. Hoyois (2015): *EM-DAT: International Disaster Database - www.emdat.be - Université Catholique de Louvain - Brussels - Belgium.*

Guterres, A. (2017): "Maintenance of International Peace and Security: Conflict Prevention and Sustaining Peace," https://www.un.org/sg/en/content/sg/statement/2017-01-10/secretary-generals-remarks-security-council-open-debate-maintenance (Accessed May 2018).

Harari, M., and E. L. Ferrara (2018): "Conflict, Climate and Cells: A Disaggregated Analysis," *The Review of Economics and Statistics*, (forthcoming).

Hauge, W., and T. Ellingsen (1998): "Beyond Environmental Scarcity: Causal Pathways to Conflict," *Journal of Peace Research*, 35(3), 299–317.

Heblich, S., A. Trew, and Y. Zylberberg (2016): "East Side Story: Historical Pollution and Persistent Neighborhood Sorting," *SERC Discussion Paper*, 208.

Hegre, H., H. Buhaug, K. Calvin, J. Nordkvelle, S. Waldhoff, and E. Gilmore (2016): "Forecasting Civil Conflict Along the Shared Socioeconomic Pathways," *Environmental Research Letters*, 11(5), 1–8.

Hendrix, C., and S. Glaser (2007): "Trends and Triggers: Climate, Climate Change and Civil Conflict in Sub-Saharan Africa," *Political Geography*, 26, 695–715.

Hendrix, C., and I. Salehyan (2012): "Climate Change, Rainfall, and Social Conflict in Africa," *Journal of Peace Research*, 49(1), 35–50.

HERRMANN, S., A. ANYAMBA, AND C. TUCKER (2005): "Recent Trends in Vegetation Dynamics in the African Sahel and their Relationship to Climate," *Global Environmental Change*, 15(4), 394–404.

HOCHRAINER, S. (2009): "Assessing the Macroeconomic Impacts of Natural Disasters. Are there Any?," *World Bank Policy Research Working Paper*, 4968.

HOMER-DIXON, T. (1999): *Environment, Scarcity, and Violence*. Princeton University Press, Princeton.

HSIANG, S., M. BURKE, AND E. MIGUEL (2013): "Reconciling Temperature-Conflict Results in Kenya," *CEGA Working Papers*, 032.

HSIANG, S., AND K. MENG (2014): "Reconciling disagreement over climate-conflict results in Africa," *Proceedings of the National Academy of Sciences*, 111(6), 2100–2103.

HSIANG, S., K. MENG, AND M. CANE (2011): "Civil Conflicts are Associated with the Global Climate," *Nature*, 476(7361), 438–441.

IBRAHIM, Y., H. BALZTER, J. KADUK, AND C. TUCKER (2015): "Land Degradation Assessment Using Residual Trend Analysis of GIMMS NDVI3g, Soil Moisture and Rainfall in Sub-Saharan West Africa from 1982 to 2012," *Remote Sensing*, 7(5), 5471–5494.

IPCC (2014): "Climate Change 2014 Synthesis Report Summary Chapter for Policymakers," .

——— (2017): "Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems (SR2)," .

JUDSON, R., AND A. OWEN (1999): "Estimating dynamic panel data models: a guide for macroeconomists," *Economic Letters*, 65(1), 9–15.

KAHN, M. (2005): "The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions," *The Review of Economics and Statistics*, 87(2), 271–284.

KRIŽANEC, B., AND A. MAJCEN LE MARECHAL (2006): "Dioxins and Dioxin-Like Persistent Organic Pollutants in Textiles and Chemicals in the Textile Sector," *Croatica Chemica Acta*, 79(2), 177–186.

KUNČIČ, A. (2014): "Institutional quality dataset," *Journal of Institutional Economics*, 10(1), 135–161.

LOAYZA, N., E. OLABERRÍA, J. RIGOLINI, AND L. CHRISTIAENSEN (2012): "Natural Disasters and Growth: Going beyond the Averages," *World Development*, 40(7), 1317–1336.

MANKIW, N. G., D. ROMER, AND D. WEIL (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107(2), 407–437.

MARSHALL, M. G., T. R. GURR, AND K. JAGGERS (2016): "Polity IV Project Dataset Users' Manual, v.2015," *Polity IV Project*, pp. 1–86.

MAYSTADT, J.-F., M. CALDERONE, AND L. YOU (2015): "Local Warming and Violent Conflict in North and South Sudan," *Journal of Economic Geography*, 15, 649–671.

MAYSTADT, J. F., AND O. ECKER (2014): "Extreme Weather and Civil War: Does Drought fuel Conflict in Somalia through Livestock Price Shocks?," *American Journal of Agricultural Economics*, 96(4), 1157–1182.

MCDERMOTT, T., F. BARRY, AND R. TOL (2013): "Disasters and Development: Natural Disasters, Credit Constraints, and Economic Growth," *Oxford Economic Papers*, November, 1–24.

MEGAHED, A. M., H. DAHSHAN, M. A. ABD-EL-KADER, A. M. M. ABD-ELALL, M. H. ELBANA, E. NABAWY, AND H. A. MAHMOUD (2015): "Polychlorinated Biphenyls Water Pollution along the River Nile, Egypt," *Scientific World Journal*, 2015.

MIGUEL, E., S. SATYANATH, AND E. SERGENTI (2012): "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," 112(4), 725–753.

MOON, B. K. (2007): "A Climate Culprit in Darfur," *Washington Post, http://www.washingtonpost.com/wp-dyn/content/article/2007/06/15.html (Accessed May 2018).*

MORITZ, M. (2010): "Understanding Herder-Farmer Conflicts in West Africa: Outline of a Processual Approach," *Human Organization*, 69(2), 138–148.

MURDOCK, G. P. (1967): "Ethnographic Atlas: A Summary," *Ethnology*, pp. 109–236.

NASA (2015): "MOD13C1 MODIS/Terra Vegetation Indices 1 mo L3 Global 0.25Deg CMG V006," https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD_NDVI_M (Accessed December 2017).

NATCATSERVICE (2014): "NatCat Data," .

Nickell, S. (1981): “Biases in Dynamic Models with Fixed Effects,” *Econometrica*, 49(6), 1417–1426.

Noy, I. (2009): “The Macroeconomic Consequences of Disasters,” *Journal of Development Economics*, 88, 221–231.

Nutini, F., M. Boschetti, P. A. Brivio, S. Bocchi, and M. Antoninetti (2013): “Land-Use and Land-Cover Change Detection in a Semi-Arid Area of Niger Using Multi-Temporal Analysis of Landsat Images,” *International Journal of Remote Sensing*, 34(13), 4769–4790.

O’Loughlin, J., A. Linke, and F. Witmer (2014): “Effects of Temperature and Precipitation Variability on the Risk of Violence in Sub-Saharan Africa, 1980-2012,” *Proceedings of the National Academy of Sciences*, 111(47), 16712–16717.

O’Loughlin, J., F. Witmer, A. Linke, A. Laing, A. Gettelman, and J. Dudhia (2012): “Climate Variability and Conflict Risk in East Africa, 1990-2009,” *Proceedings of the National Academy of Sciences*, 109(45), 18344–18349.

Owain, E. L., and M. Maslin (2018): “Assessing the Relative Contribution of Economic, Political and Environmental Factors on Past Conflict and the Displacement of People in East Africa,” *Palgrave Communications*, (2018).

Oyama, S. (2014): “Farmer-Herder Conflict, Land Rehabilitation, and Conflict Prevention in the Sahel Region of West Africa,” *African Study Monographs*, 50, 103–122.

Raddatz, C. (2009): “The Wrath of God: Macroeconomic Costs of Natural Disasters,” *World Bank Policy Research Working Paper Series*, 5039.

Raleigh, C., and D. Kniveton (2012): “Come Rain or Shine: An Snalysis of Conflict and Climate Variability in East Africa,” *Journal of Peace Research*, 49(1), 51–64.

Raleigh, C., A. Linke, H. Hegre, and J. Karlsen (2010): “Introducing ACLED: An Armed Conflict Location and Event Dataset,” *Journal of Peace Research*, 47(5), 651–660.

Raleigh, C., and H. Urdal (2007): “Climate Change, Environmental Degradation and Armed Conflict,” *Political Geography*, 26, 674–694.

Rodrik, D., A. Subramanian, and F. Trebbi (2004): “Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development,” *Journal of Economic Growth*, 9(2), 131–165.

ROMERO, M., AND S. SAAVEDRA (2016): "The Effects of Gold Mining on Newborns' Health," .

SACHS, J. D. (2003): "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," *National Bureau of Economic Research Working Paper*, 9490.

SARSONS, H. (2015): "Rainfall and Conflict: A Cautionary Tale," *Journal of Development Economics*, 115, 62–72.

SCHWARZENBACH, R. P., T. EGLI, T. B. HOFSTETTER, U. VON GUNTEN, AND B. WEHRLI (2010): "Global Water Pollution and Human Health," *Annual Review of Environment and Resources*, 35(1), 109–136.

STOCKING, M. (2001): "Land Degradation," in *International Encyclopedia of the Social & Behavioral Sciences*, ed. by N. Smelser, and P. Baltes, pp. 8242–8247. Pergamon, Oxford.

TEORELL, J., N. CHARRON, S. DAHLBERG, S. HOLMBERG, B. ROTHSTEIN, P. SUNDIN, AND R. SVENSSON (2013): "The Quality of Government Dataset, version 20Dec13.," *University of Gothenburg: The Quality of Government Institute.*

THEBAUD, B. (2017): "Pastoral and Agropastoral Resilience in the Sahel: Portrait of the 2014-2015 and 2015-2016 Transhumance," *AFL-NCG.*

THEISEN, O. M. (2012): "Climate Clashes? Weather Variability, Land Pressure, and Organized Violence in Kenya, 1989-2004," *Journal of Peace Research*, 49(1), 81–96.

THEISEN, O. M., H. HOLTERMANN, AND H. BUHAUG (2012): "Climate Wars? Assessing the Claim That Drought Breeds Conflict," *International Security*, 36(3), 79–106.

TOL, R., AND F. LEEK (1999): "Economic Analysis of Natural Disasters," in *Climate, Change and Risk*, ed. by T. Downing, A. Olsthoorn, and F. Leek, chap. 12. Routledge, London.

TONG, X., M. BRANDT, P. HIERNAUX, S. M. HERRMANN, F. TIAN, A. V. PRISHCHEPOV, AND R. FENSHOLT (2017): "Revisiting the Coupling between NDVI Trends and Cropland Changes in the Sahel Drylands: A Case Study in Western Niger," *Remote Sensing of Environment*, 191, 286–296.

UMWELTBUNDESAMT (2018): *Dioxins.*

UNEP (1992): *World Atlas of Desertification.* Edward Arnold, London Baltimore.

UNICEF (2017): "Water, Sanitation, Hygiene," https://www.unicef.org/egypt/wes.html (accessed november 2017).

Vicente-Serrano, S., S. Beguería, and J. López-Moreno (2010): "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index," *Journal of Climate*, 23(7), 1696–1718.

von Peter, G., S. von Dahlen, and S. Saxena (2012): "Unmitigated Disasters? New Evidence on the Macroeconomic Cost of Natural Catastrophes," *BIS Working Papers*, 394(394), 1–38.

von Uexkull, N., M. Croicu, H. Fjelde, and H. Buhaug (2016): "Civil Conflict Sensitivity to Growing-Season Drought," *Proceedings of the National Academy of Sciences*, 113(44), 12391–12396.

Wahaab, R., and M. Badawy (2004): "Water Quality Assessment of the River Nile System : An Overview," *Biomedical and Environmental Sciences*, 17, 87–100.

WashWatch (2017): "WashWatch Egypt Statistics, accessed 2017," https://www.washwatch.org (Accessed November 2014).

Weidmann, N. B., J. K. Rød, and L.-E. Cederman (2010): "Representing Ethnic Groups in Space: A New Dataset," *Journal of Peace Research*, 47(4), 491–499.

WHO (2017a): "Diarrhoeal Disease," http://www.who.int/mediacentre/factsheets/fs330/en/ (Accessed November 2014).

——— (2017b): "Persistant Organic Pollutants (POPs)," http://www.who.int/foodsafety/areas_work/chemical-risks/pops/en/ (Accessed November 2014).

——— (2017c): "The Cost of a Polluted Environment," http://www.who.int/mediacentre/news/releases/2017/pollution-child-death/en/ (Accessed November 2014).

Wirtz, A., W. Kron, P. Loew, and M. Steuer (2014): "The Need for Data: Natural Disasters and the Challenges of Database Management," *Natural Hazards*, 70(1), 135–157.

World Bank (2014): *World Development Indicators*. Washington, DC.

WWF (2017): "Water Pollution," http://wwf.panda.org/about_our_earth/.

Zhang, J., and L. C. Xu (2016): "The Long-Run Effects of Treated Water on Education: The Rural Drinking Water Program in China," *Journal of Development Economics*, 122, 1–15.

Zivin, J. G., and M. Neidell (2013): "Environment, Health, and Human Capital," *Journal of Economic Literature*, 51(3), 689–730.