WORKING PAPER JUNE 2018

WEATHER AND CONFLICTS IN AFGHANISTAN

Monir Elias Bounadi

Weather and Conflicts in Afghanistan

EBA Working paper

Monir Elias Bounadi
June 2018

Underlagsrapport 2018 till

Expertgruppen för Biståndsanalys (EBA)

In collaboration with the Master's Programme at the Department of Economics, Stockholm University, the EBA regularly engages as Master's thesis co-supervisors. A successfully graded thesis may be published as an EBA Working Paper.

Please refer to the present paper as:

Bounadi, Monir Elias (2018), Weather and Conflicts in Afghanistan, EBA Working Paper, June, 2018, Expert Group for Aid Studies, Sweden.

The Expert Group for Aid Studies – EBA – is a Government committee analysing and evaluating Swedish international development aid. This report can be downloaded free of charge at www.eba.se

Cover design by Julia Demenchenko

Weather and Conflicts in Afghanistan

By MONIR ELIAS BOUNADI*

Abstract

I combine high-resolution data on temperature and precipitation with georeferenced data on conflict events to explore the link between local weather variations and conflict incidence for all districts of Afghanistan between July 2005 and December 2016. By utilizing exogenous interannual variation in daily temperature and precipitation within district-months, I find that exchanging colder for warmer days tends to increase the likelihood of conflict and that precipitation does not drive the occurrence of conflict. I provide suggestive evidence that temperature shocks to opium production do not explain the observed temperature-conflict link.

Keywords: weather, conflicts, Afghanistan

Department of Economics
Stockholm University

EC9901 Master Thesis in Economics
June 18, 2018

*Acknowledgments: I thank Tessa Bold for the encouragement and valuable advice. I am also indebted to Jossy Purayidathil Jacob at the National Aeronautics and Space Administration Goddard Space Flight Center for access to unique temperature data.

1 Introduction

The weather of Afghanistan has grown harsher since the 20$^{\text{th}}$ century. The steady increase in temperature and intensification of local droughts and floods—changes predicted to continue under global climate change—have shaken the region, degrading infrastructure and causing adaptive challenges for farmers (NEPA and UNEP, 2015). Meanwhile conflicts plague Afghanistan. Recent data from the Uppsala Conflict Data Program (UCDP) indicate that the number of conflict events rose from about 800 to 1,900 between 2006 and 2016. The trend in harsher weather conditions combined with the escalation of conflicts raises the fundamental question of a causal relationship.

Much of the violence in Afghanistan is likely driven by ethnolinguistic intolerance and competition for scarce resources.[2] As such, the weather is not the sole conflict driver. Still, the weather acts as an ever-present support factor that under certain circumstances raises the likelihood of a violent incident. A case of relevance is the Afghan 1969-1972 drought crisis. The failure of the last king to respond to the crisis weakened the support for the monarch and opened up for the successful July 1973 coup d'etat (Ruttig, 2013). In the absence of droughts, there would be no room for such a failure, and the conflict history of Afghanistan would perhaps have looked different.

The present thesis explores the impact of weather variations on conflicts for all districts of Afghanistan (i.e., second administrative-level regions) between July 2005 and December 2016. The empirical analysis is based on a com-

[2] In a series of interviews in Kabul in 2017 by the Afghanistan Research and Evaluation Unit (AREU, 2017), interviewees were found to blame the current Kuchi-Hazara conflict on the blocking by local Hazara militias of Kuchi migration routes. At the same time, Hazara farmers seem to view these measures as justified by recent experiences of crop destruction connected to Kuchi migration efforts.

bination of monthly location-specific information on conflict events from the UCDP with an original weather dataset. The latter include high-resolution daily temperature and precipitation data from the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) and the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS).

My use of a conservative panel data fixed effects model enables causal identification. I argue that by including district-month fixed effects in all specifications, I can utilize exogenous interannual variation in daily temperature and precipitation within district-months. I find that exchanging colder for warmer days tends to increase the likelihood of a conflict and that precipitation does not drive the occurrence of conflict. The observed temperature-conflict link describes a net effect of temperature variations on conflict incidence and is consistent with several theories. Returning to a dominant topic in empirical studies on conflicts in Afghanistan, I provide suggestive evidence that temperature shocks to opium production do not explain the observed temperature-conflict link.

To the best of my knowledge, this study adds to the literature in three ways. First, it is the first subnational fixed effects study on the weather-conflict relationship in Afghanistan.[3] Second, it is first to employ such a high spatial resolution to study the impact of high-frequency monthly variation in weather on intergroup conflicts. Third, it provides suggestive evidence on the role of opium production in explaining the connection between variation in tempera-

[3]There is one study on the weather-conflict link in Afghanistan by Carter and Veale (2013). They use an event count model. Their exclusion of fixed effects (including unobservable characteristics such as farmland values) makes it doubtful that they are exploiting exogenous variation in their weather variables. Furthermore, they include a host of controls that are potentially endogenous to weather variation (e.g., opium cultivation), so that the coefficients on their weather variables are potentially biased.

ture and conflicts.

The remainder of this thesis is organized as follows. Section 2 identifies main findings and econometric insights from the existing literature. It further contains a summary of mechanisms that make the weather-conflict link ex-ante probable and a sketch of the weather risk profile of Afghanistan. Section 3 presents the data. Section 4 outlines the empirical strategy. Then Section 5 displays the baseline results, various sensitivity analyses and an analysis of the role of opium production in explaining the observed temperature-conflict link. Section 6 concludes with a discussion.

2 Existing Literature and Conceptual Framework

I here summarize the main findings and econometric insights from recent studies. For a comprehensive review of the general effects of weather variations and their impact on conflicts, the reader is referred to Dell, Jones and Olken (2014) and Burke, Hsiang and Miguel (2015).

Many studies on the weather-conflict relationship have used weather variation as an instrument for non-climatic causal factors of conflicts. For example, Miguel, Satyanath and Sergenti (2004) use rainfall variation as an instrument for economic growth and Maystadt and Ecker (2014) use a temperature-based drought indicator as an instrument for livestock prices. However, the key identifying assumption that the instrumental variable only affects conflicts through a particular intermediate variable (i.e., the exclusion restriction) is unlikely to hold in general as climatic events affect other likely causes of conflicts. These causes include human health, agricultural income, demographics and psychological attitudes towards violence (Baysan et al., 2015; Carleton and Hsiang, 2016; Dell, Jones and Olken, 2014).

That the effect of climate on conflicts tends to operate through a plethora of channels has encouraged a shift towards reduced-form analyses. On the one hand, reduced-form estimates seldom pin down any specific mechanism that explains the weather-conflict link. On the other hand, no exclusion restriction need to be satisfied, and we identify the net effect of weather variations per se on conflicts. This net effect is causal under the weaker assumption that weather variations are (unconfounded) random draws from a climate distribution (after controlling for a broad set of covariates).[4]

Most reduced-form estimates from recent studies rely on panel data fixed effects approaches. These approaches tend to have robust identification properties as the inclusion of panel-specific and time-varying fixed effects allow one to absorb unobserved fixed spatial characteristics and time-varying regional shocks. Although the inclusion of fixed effects eliminates the need to explicitly control for all confounders (as in a cross-sectional regression), the use of time series variation in the panel data means that we only identify causal effects at specific frequencies. Thus, though short-run time series variation may identify effects of high-frequency weather variations this does not necessarily inform the debate on the impacts of climate change on conflicts.[5]

Given a panel data fixed effects model additional methodological challenges abound. One concerns the time dependence of conflicts on the weather. Climatic events can displace events in time such as delaying the ending of a conflict event. They can also have persistent effects. For example, Carleton (2017) find that the impact of growing season temperature on suicide rates last for about five years. Similar dynamics are present spatially. For just as local con-

[4]Note that I implicitly think of weather variations as short-run realizations from a long-run climate distribution.

[5]Hsiang (2016) derives sufficient (but not necessary) econometric conditions on when measured weather variations identify causal effects of stochastic perturbations to the climate.

flict events spill over temporally they may spill over to neighboring locations (Harari and La Ferrara, 2017).

Also, there is the task to specify the correct dose-response relationship between weather and conflicts. Early analyses used affine transformations of weather variables or climate-based indicators. However, these seem insufficient characterization of the weather-conflict link. Because at least across some regions, the relationship appears nonlinear (Hsiang, Burke and Miguel, 2013).

Of importance has also been to identify the correct set of fixed effects. Failure to include a fixed effect can generate significant effects that are the results of simple coincident time series variation. This failure is empirically illustrated by Couttenier and Soubeyran (2014) who find that adding year to country fixed effects (that, among other things, control for worldwide climate changes) reduces the effect of temperature on civil war by about 2/3. However, this should not encourage us to include fixed effects whenever possible. Inevitable measurement error in weather variables causes attenuation bias of our estimators, and this attenuation bias can be amplified when adding fixed effects (Fisher et al., 2012; Wansbeek and Koning, 1991).

These methodological concerns explain the broad set of different approaches in the weather-conflict literature. Adding the use of different datasets to these concerns make it hard to summarize the evidence on the weather-conflict link. To examine the weather-conflict link more systematically, Burke, Hsiang and Miguel (2015) conduct a meta-analysis of 55 fixed effects studies on the weather-conflict relationship. These studies cover a broad spectrum of violence including both interpersonal and intergroup violence (e.g., violent crime and civil conflicts). Meta-analysing these studies in a hierarchical Bayesian framework, they find that contemporaneous temperature has the most substantial mean effect on conflict. Specifically, they find that each within-location standard deviation increase in contemporaneous temperature induces a 2.4 percent increase in interpersonal conflicts and an 11.3 percent increase in intergroup conflicts, relative to the baseline average probability of conflict (posterior p-value < 0.001). The contemporaneous effect of precipitation on interpersonal and intergroup conflict is only marginally significant (posterior p-value < 0.100).

As discussed in Burke, Hsiang and Miguel (2015), the rich weather-conflict literature suggests several channels that could explain this empirically robust weather-conflict link. First, there is the *productivity* channel. Extreme weather events such as long periods of droughts and heavy rainfall can lower productivity and wages within the agricultural sector, causing a worsening of current living standards. These adverse shocks can spur conflicts by reducing the opportunity cost of conflict by more than it alters the value of peace, as in a model by Chassang and Padro-i Miquel (2009). Second, there is the *migration* channel. If, e.g., urban labor markets cannot absorb rural climate immigrants, per capita income may decline and, in turn, induce a rise in conflicts. Also, weather-induced migration can strengthen ethnolinguistic fragmentation and increase ethnic violence. Third, there is the *physical geography* channel. This channel is active when weather variations cause changes in the physical geography that raise or lower the probability of a successful attack. These changes include generated constraints on ground operations that depend upon logistics and intelligence gathering. Finally, there is the *psychological* channel. Events such as heat waves can alter the physical and psychological stress and, as a result, affect the psychological cost of acting on intents to act violently.[6]

[6] For *productivity*, see Maystadt and Ecker (2014); for *migration* and the comment regarding ethnolinguistic fragmentation, see Bohra-Mishra, Oppenheimer and Hsiang (2014) and Ray and Esteban (2017); for *physical geography*, see the discussions on flood destruction of road networks in Miguel, Satyanath and Sergenti (2004) and on the role of physical geography in Afghanistan in Carter and Veale (2013); and for *psychological*, see Baysan et al. (2015).

It is possible that all these channels are present in Afghanistan. Nevertheless, it is valuable to delineate the weather-conflict risk profile of Afghanistan (NEPA and UNEP, 2015; NEPA, UNEP and WFP, 2016). The most important characteristic of this profile seems to be long periods of droughts and increased temperature. Such events can increase the internal displacement in the region and, in turn, exacerbate existing tensions between ethnolinguistic groups in a country with more than a dozen different major ethnolinguistic groups (e.g., ethnic violence between the Shia Muslim group Hazaras and the militantly Sunni Pashtun Taleban). High regional tensions that result from extreme weather events can further deepen the competition for the use of scarce productive rangelands and exacerbate existing nomad-farmer conflicts, especially as about 45 percent of the total land mass is under permanent pasture (NEPA, UNEP and WFP, 2016).

The risk factor most studied in empirical conflict studies on Afghanistan concerns how changes in the Afghan opium economy drive conflicts (Bove and Elia, 2013; Gehring, Langlotz and Kienberger, 2017; Lind, Moene and Willumsen, 2014). For suppose long periods of droughts induce farmers to grow more of the opium poppy, a drought-resistant crop. Then if the Taleban use revenues from opium production to finance insurgencies, we expect increased opium production to raise conflict levels.[7] Gehring, Langlotz and Kienberger (2017) provide some evidence on this. Instrumenting indicative district-level opium cultivation with a drought index they find an adverse effect of opium cultivation last year on battle-related deaths this year. However, if droughts change other aspects associated with conflicts, their estimates are biased (as the exclusion restriction does not hold).

[7] According to the UNODC World Drug Report 2017, up to 85 percent of opium poppy cultivation in Afghanistan was in Taleban territory during 2016 (UNODC, 2017).

3 Data

3.1 Data Description

The structure of the dataset is a georeferenced balanced panel across 398 districts in 34 provinces of Afghanistan from July 2005 to December 2016. My unit of analysis is a district-year-month. The administrative boundaries are fixed to those recognized by the Afghan Ministry of the Interior in June 2005. This fixation ensures that my unit of analysis is not endogenous to conflict events during the sample period.[8]

In the next section, I study how weather variations affects the probability that a conflict occurs in a given district during a given year-month. This section provides information about the main datasets and variables. Additional data details appear in Appendix A.

Conflict Events.—Monthly information on conflict events at the district-level comes from the UCDP Georeferenced Event Data (GED). The primary variable

[8]The number of unofficial and temporary districts is continually changing in Afghanistan. To date, the number of districts is above 400. If political violence drives administrative boundary changes, or vice verse (Bazzi and Gudgeon, 2016), fixing the administrative boundaries to a year-month contained in the sample period would make my unit of analysis endogenous to conflict events. Compare with Berman et al. (2017) who use the PRIO-GRID dataset. This dataset defines a spatiotemporal grid structure of $0.5° \times 0.5°$ cells that are by construction unrelated to administrative boundaries and hence not endogenous to conflict events. However, a similar setup is unfeasible in this case. The reason is that a grid structure that covers all districts of Afghanistan would roughly be a grid of $0.05° \times 0.05°$ cells. It is impossible to match these grid cells with the UCDP measure of conflict events as these are not coded at such a fine-grained level. Furthermore, for district assignment of conflict events, I had to ensure that the UCDP acknowledge the particular division of administrative regions I use, and this seems to be the case from July 2005 to December 2016 (see Appendix A.2).

is (the best estimate of) battle-related deaths resulting from an event.[9] An event is an incident that meets the following four main criteria (Croicu and Sundberg, 2017, pp. 9–10, 15). First, the incident must involve armed force by "an organi[z]ed actor against another organized actor, or against civilians". Second, the incident must result in at least one death "relating to either combat between warring parties or violence against civilians". Third, it must be possible to represent the incident as involving "two conflicting primary parties or party killing unarmed civilians". Four, the incident must pass the threshold of 25 annual deaths, counting from 1989 to 2016.

The consequences of these restrictions for my results are unknown but is not a concern as there is no comparable dataset of similar quality.[10] Sundberg and Melander (2013) discuss the limitations of the threshold of 25 annual deaths for inclusion in the UCDP GED. The take-home message is that the strict adherence to the 25 annual deaths threshold obscures minor conflicts but captures

[9] The other measures of the number of battle-related deaths in the UCDP GED are the lowest and highest reliable estimates of ditto. The lowest and highest reliable estimate coincides with the best for around 95.1 and 84.3 percent of the events, respectively. I note that the baseline results in Subsection 5.1 are barely affected by the chosen measure (not shown).

[10] That there is no comparable dataset of similar quality comes from comparing the UCDP GED to two other comparable datasets. First, conflict data from the Global Database of Events, Language, and Tone (GDELT) Project at https://www.gdeltproject.org/data.html/. The GDELT Project data contain information on conflict events from world local media in 100 different languages. In contrast, the UCDP GED is primarily derived from a large number of sources from Factiva and consequently contain almost exclusively English material. However, the GDELT Project dataset does not record events at the district-level. The second comparable dataset is version 1 of the Armed Conflict Location & Event Data Project (ACLED) dataset at https://www.acleddata.com/data/. The ACLED dataset is as precise as the UCDP GED and further contain conflict events that do not meet the threshold of 25 annual deaths (i.e., minor conflicts). However, version 1 of the ACLED dataset is by construction an incomplete pilot dataset and contain no information on conflict events after 2010.

sensible definitions of concepts of major (i.e., high intensity) armed conflicts. That is, the UCDP GED is an unrepresentative sample of all conflict events in Afghanistan but is more likely a representative sample of all major conflict events in Afghanistan.

The UCDP GED was spatially joined to the 398 Afghan districts and restricted to the period from July 2005 to December 2016. The number of conflict events is 19,846. For the empirical analysis, I restrict the dataset to events known to occur at the district-level for no more than 30 days. This restriction drops 5,378 events. I also drop five events that occurred at the border between two districts and that has no district name assigned. Out of 14,468 remaining events (about 72.9 percent of the full sample), 110 spans two months. Battle-related deaths related to these 110 events are not assigned to both months but are instead assigned to the second month when the event ends.[11]

Temperature and Precipitation.—Daily temperature data come from the Asia Land Information System (LIS) Framework developed at the Hydrological Sciences Laboratory at the NASA GSFC (Kumar et al., 2006). The core of the NASA LIS Framework consists of a land surface model and tools for high-performance computing. Input data include satellite and ground-based observational data (e.g., data on topography, vegetation, snowpack and soil moisture). The final product is a series of variables on land surface states and fluxes. Among these, the sole variable of interest is the daily average land surface temperature in Kelvin over all of Central Asia at 0.01° resolution (approximately 1.11 km at the equator) for each day between July 2005 to December 2016. For each day and district of Afghanistan, I compute an area-weighted mea-

[11] The baseline results in Subsection 5.1 are essentially unchanged if these events are dropped or if battle-related deaths are assigned to the first month when the event starts (not shown).

sure of land surface temperature, namely the mean daily land surface temperature converted into °C across pixels whose centroids falls within the district's boundary.[12]

Daily precipitation data is the CHIRPS. The dataset contains information on daily precipitation over a quasi-global grid at 0.05° resolution (approximately 5.55 km at the equator) from 1981 to near-present. Data was initially generated using interpolation techniques incorporating satellite imagery with in-situ station data (Funk et al., 2015). The dataset contains information on precipitation in millimeters (mm) over $0.05° \times 0.05°$ longitude and latitude cells across all 398 districts of Afghanistan for each day between July 2005 and December 2016. For each day and district of Afghanistan, I compute the mean daily (mm) precipitation across pixels whose centroids falls within the district's boundary.

Among the principal types of weather data, ground station data is believed to most reliably measure weather for the areas where stations are located (Dell, Jones and Olken, 2014). However, entry and exit of weather observations in conflict-ridden countries such as Afghanistan could cause the measured quantities to be endogenous to conflict events (Auffhammer et al., 2013). I believe that the interpolation and reanalysis data on temperature and precipitation are exempt from this endogeneity issue. The reason is that these are primarily products of interpolation methods and climate data models. By construction, these methods are unrelated to conflicts in Afghanistan. Therefore, as these methods dominate data generation, the endogeneity of the input data (e.g., quantities measured at the weather stations) unlikely translate into endogeneity of the generated temperature and precipitation data. Nonetheless, the methods and models used simplify the physical relationship between climatic elements and introduce measurement error into my regression estimates. Though

[12] Area-weighting is discussed in Appendix A.3.

these measurement errors decrease the efficiency of the regression estimators, they can be treated as statistical noise orthogonal to conflict events.

Opium Data.—I construct an environmental opium suitability index based on a method by Kienberger et al. (2017). Each input variable used to construct the index is a characteristic of the environment and consequently exogenous to conflicts. The index varies from 0 (not suitable) to 1 (very suitable) and is a district-level measure of the ability to meet the abiotic environmental requirements of the opium poppy (*Papaver somniferum*). I further collect yearly indicative district-level data on opium cultivation from 2005 to 2016, and information on the period for opium planting across the 34 provinces of Afghanistan, from the United Nations Office on Drugs and Crime (UNODC).

Historically opium has been a dominant source of income for Afghan farmers. During 2006 to 2008, the size of the Afgan opium economy accounted for about 40 percent of licit GDP (UNODC, 2008). The summary statistic has since then fallen almost linearly to around 5 percent in 2016 (UNODC, 2016).

3.2 Descriptive Statistics

Table 1 presents descriptive statistics on weather and conflicts for all 398 districts of Afghanistan from July 2005 to December 2016. Battle-related deaths are measured at the district-year-month-level. I observe an average number of 1.41 battle-related deaths. Conditioning on the presence of a battle-related death the average is about 9.57. Thus, during the sample period around 77,500 have died in (major) conflict events. The average probability of conflict measured as the presence of a battle-related death is 15 percent. Investigating the characteristics of the conflict events that make up the previous figures, I find that almost all involve the Government of Afghanistan and the Taleban as the two primary conflicting parties. Remaining conflict events involve the Taleban killing unarmed civilians and, since January 2016, conflicts between the Government of Afghanistan and the Islamic State (IS) (of Iraq and Syria).

TABLE 1—DESCRIPTIVE STATISTICS: WEATHER AND CONFLICTS

			SD	
	Observations	Mean	Overall	Within
Battle-related deaths†				
All	54,924	1.41	8.29	7.55
If > 0	8,096	9.57	19.71	18.05
$\mathbf{1}$(Battle-related deaths $> 0)^{\dagger}$	54,924	0.15	0.35	0.30
Conflict Events†				
All	14,468			
Government of Afghanistan vs. Taleban	13,246			
Taleban vs. Civilians	646			
Government of Afghanistan vs. IS	263			
Daily Temperature (°C)‡	1,817,666	13.28	13.24	3.64
Daily Precipitation (mm)‡	1,817,666	0.96	3.20	3.04

Notes: The summary statistic Overall SD stands for the overall standard deviation of the corresponding variable. The summary statistic Within SD stands for the overall standard deviation of the corresponding variable after removing district-month fixed effects. The variable $\mathbf{1}$(Battle-related deaths > 0) is 1 if there is at least one battle-related death, and 0 otherwise. The acronym IS stands for Islamic State (of Iraq and Syria). The sample period is July 2005 to December 2016. All 398 districts are included in the sample. Numbers are correct to two decimal places. †Measured at district-year-month-level. ‡Measured at district-year-month-day-level.

Source: Author's calculations based on data from the CHIRPS, NASA GSFC and UCDP.

Temperature and precipitation are measured at the district-year-month-day-level. The total number of observations is about 1.8 million. The average daily temperature is 13.28 °C and the average daily precipitation is 0.96 mm. Thus, on average during the sample period, a day in Afghanistan is cool and dry.

I also tabulate the overall and within standard deviation. The within standard deviation represents the standard deviation within district-months and is a summary statistic of interannual within district-month variation of a variable. Interannual within district-month variation is the variation I use as part of my identification strategy that I present in the next section. Consider the standard deviations of the weather variables. I find that there is no substantial difference in overall and within standard deviation in daily precipitation. However, the within standard deviation in temperature is about 1/4 of the overall standard deviation. Thus, while there is substantial daily variation in temperature across Afghanistan over time, interannual within district-month variation roughly occur in a small ± 3.64 °C band.

Figure 1 shows that the overall and within-province variation in the opium suitability index is noticeable. The index is 0 (1) in the district in which it is least (most) suitable to grow opium poppies, and between 0 and 1 for all other districts. I expect the density of opium production to correlate positively with the index. For 134 out of 398 districts that have never cultivated opium during the sample period, I expect planned opium cultivation to increase with the index. Regarding the temporal variation, 10 out of 34 provinces cultivate opium during spring (most in the Northern and Central regions), 14 during Autumn (most in the Eastern and Southern regions), and the remaining 10 during spring and autumn (not concentrated to any region).

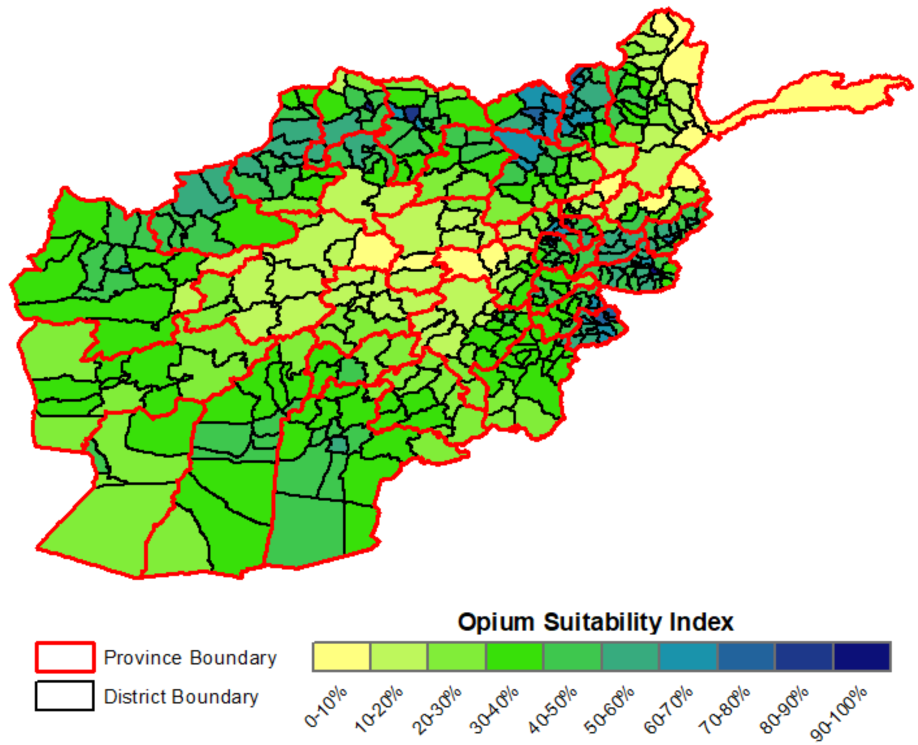


FIGURE 1. ENVIRONMENTAL SUITABILITY FOR OPIUM POPPY CULTIVATION IN AFGHANISTAN

Notes: The figure shows the distribution of the constructed environmental opium suitability index. The index measures the ability of a district's location to meet the abiotic environmental requirements of the opium poppy (*Papaver somniferum*). See the text for more details.

Source: Author's presentation based on calculations using data from Globcover 2009, USGS HydroSHEDS and WorldClim. See Appendix A for more details.

4 Empirical Strategy

I begin by describing the baseline specification that I use to estimate the weather-conflict relationship. Then I outline my baseline standard error correction method used for statistical inference.

Baseline Specification.—My baseline specification utilizes plausibly random interannual weather variation in district-month-specific weather distributions. Mathematically I employ ordinary least squares (OLS) to estimate the following linear probability model:

$$C_{dt} = \sum_{l=0}^{1} \Big(\sum_{i} \beta_i^l T_{d,t-l}^i + \sum_{j} \gamma_j^l P_{d,t-l}^j \Big) + \delta_{dm} + \pi_{pt} + q_{dm}(y) + \epsilon_{dt}, \qquad (1)$$

where d denotes the district, p the province, m the month (January to December), y the year and t the year-month. The dependent variable C_{dt}—conflict incidence for short—is my measure of the presence of a conflict. Specifically, C_{dt} is a binary variable equal to 1 if there is at least one battle-related death in district d year-month t, and 0 otherwise.

The variables T_{dt}^i and P_{dt}^j are measures of (land surface) temperature and precipitation. Specifically, T_{dt}^i and P_{dt}^j denotes the number of days temperature and precipitation falls in bin i and j in district d year-month t, respectively. These temperature- and precipitation-day bins are my explanatory variables of interest and were constructed to approximate the potentially nonlinear relationship between weather and conflicts.[13]

[13] If the weather-conflict relationship is nonlinear, a linear parametric function of weather bins as in (1) approximate the nonlinear relationship. See Appendix B for a detailed discussion. Also, see Barreca et al. (2016) for a similar methodology applied to temperature and mortality in the United States of America.

To capture the exposure to the full distribution of temperature and precipitation I define the temperature- and precipitation-day bins as follows. For temperature I use 2 extreme bins from the minimum value (about -41 °C) to -20 °C and from 40 °C to the maximum value (about 45 °C). Then 12 bins of length 5 °C from -20 °C to 40 °C are defined. For precipitation, I follow literature in hydrometeorology and classify bins into dry-day bins that cover days with less than 1 mm precipitation and wet-day bins that cover days with at least 1 mm precipitation. I define five dry-day bins of length 0.2 mm, seven wet-day bins of length 2 mm from 1 mm to 15 mm and one bin from 15 mm to the maximum value (about 117 mm). To avoid perfect multicollinearity one bin has to be omitted for both temperature and precipitation.[14] I omit a bin if it contains the mean temperature or precipitation across all observations. Equivalently, I omit the $[10, 15)$ °C and $[0.8, 1)$ mm bin. Figure 2 illustrates the distribution of temperature and precipitation across these baseline bins.

To control for any possible direct effects that weather variations in prior year-months might have on conflict incidence in the current year-month, I include the first order lags $T^i_{d,t-1}$ and $P^j_{d,t-1}$. These lags are interesting in themselves, but also control for potential serial correlation in weather and possible delayed effects of weather shocks on current conflict incidence that if ignored could make my estimators inconsistent (Burke, Hsiang and Miguel, 2015).

The parameters of interest are the coefficients β^l_i and γ^l_j on $T^i_{d,t-l}$ and $P^j_{d,t-l}$. These are to be interpreted as the contemporaneous or delayed effect ($l = 0$ or 1, respectively) of exchanging one day in the omitted bin for a day in the bin specified by the sub-index. For example, $\beta^0_{[25,30)\,°C}$ is the contemporaneous effect on conflict incidence from increasing the monthly count of days in

[14] All temperature- or precipitation-day bins in a given month sum to the number of days in that month.

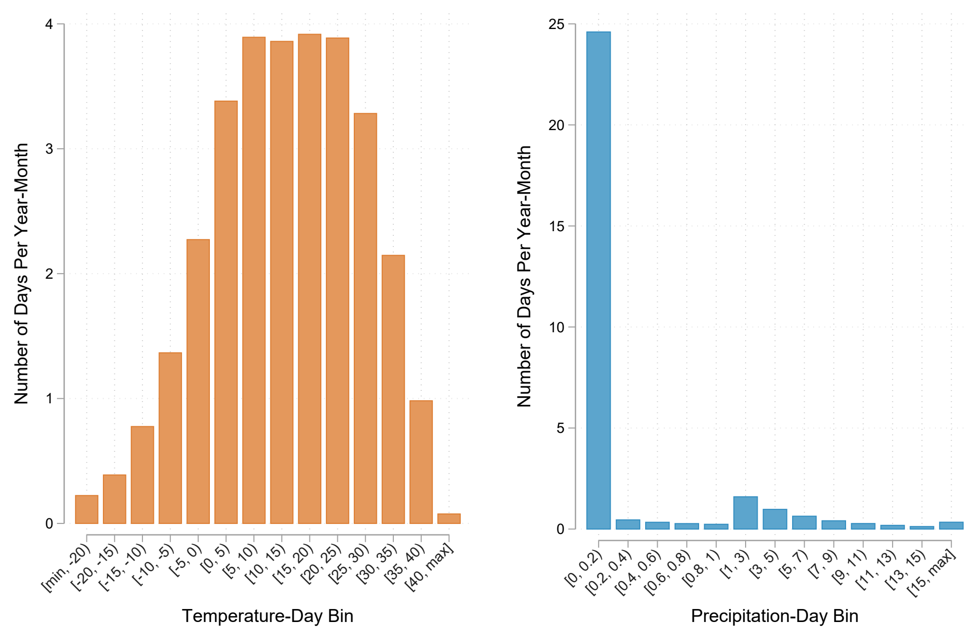


FIGURE 2. BASELINE TEMPERATURE- AND PRECIPITATION-DAY BINS

Notes: The figure shows the average distribution of daily average temperature and precipitation across 14 temperature-day bins (left panel) and 13 precipitation-day bins (right panel). Each bar represents the average number of days per year-month in each temperature or precipitation category across all 398 districts of Afghanistan over the sample period July 2005 to December 2016. Minimum daily temperature is about −41 °C, and maximum daily temperature is about 45 °C. Maximum daily precipitation is about 117 mm. See the text for more details.

Source: Author's calculations based on data from the CHIRPS and NASA GSFC.

the $[25, 30)$ °C bin by one, which implicitly requires removing a day from the omitted bin $[10, 15)$ °C.[15]

My focus on two rather than one weather variable accounts for the correlation of conflict-inducing weather variables. For example, Auffhammer et al. (2013) find a negative correlation between annual temperature and precipitation in hot areas where more precipitation and the associated evaporation cools the temperature level.[16] Since it is ex-ante unclear if temperature, precipitation or both affect conflict incidence, I include both.[17]

The baseline specification includes a full set of district-month fixed effects δ_{dm} and province-year-month fixed effects π_{pt}. The district-month fixed effects ensure that my parameters of interest are identified from interannual variation in district-month-specific weather distributions. The province-year-month fixed effects nonparametrically filter out shocks at the provincial level across all time periods. To exemplify the resulting variation I am exploiting, consider the Kabul District of the Kabul Province. For this district, the specification exploits interannual variation in a specific month (e.g., January) after controlling for shocks to weather and conflict incidence at the province-year-month-level

[15] To find all relevant comparisons from an interpretative point of view let $\Xi^l(q, r; \xi) \equiv \xi_q^l - \xi_r^l$, where ξ is β or γ. Then, e.g., $\Xi^0(q, r; \beta)$ is the contemporaneous effect on conflict incidence from exchanging one day in temperature-day bin q with a day in temperature-day bin r. Thus, for, e.g., 14 temperature-day bins there are $\binom{14}{2} = 91$ relevant comparisons for each lag order l.

[16] The sample Pearson's correlation coefficient of daily temperature and precipitation is about -0.15 in my sample and is significant at the 0.1 percent α-level.

[17] There may be other weather variables that affect conflicts such as humidity, evapotranspiration and snow cover. However, since most of these are response variables that react to changes in temperature and precipitation (e.g., water tend to vaporize, and snow melt, at higher temperatures) these are so-called bad controls. I, therefore, accept the working assumption that temperature and precipitation are most probably sufficient statistics for capturing the weather-conflict relationship and view them as broad climatic driver variables.

(e.g., shocks occurring in the Kabul Province in January 2006 to January 2016).[18]

Econometrically the importance of the fixed effects also stems from their ability to nonparametrically control for omitted determinants of conflict incidence that covaries with weather variations (i.e., omitted variables). First, the district-month fixed effects account for intraannual district-specific variation in farmland values (driven by, e.g., irrigation seasonality) that could covary with weather outcomes and the opportunity cost of engaging in violence (Jia, 2014).[19] Second, the province-year-month fixed effects nonparametrically control for regional shocks such as climate change-induced shifts in the distribution of extreme weather events that may drive conflict events through, e.g., regional price shocks (IPCC, 2014; Maystadt and Ecker, 2014).[20] The province-

[18] The inclusion of the district-month and province-year-month fixed effects implies that the estimation of (1)—described as a least squares dummy variable model—using OLS yields multidimensional within estimators of all parameters. To derive the specific within transformation, view each variable z_{dt} as z_{dpmy}. Then the within transformation $\tilde{z}_{dpmy}$ of z_{dpmy} is (Balazsi, Matyas and Wansbeek, 2018)

$$\tilde{z}_{dpmy} = z_{dpmy} - \overline{z}_{dpm\cdot} - \overline{z}_{\cdot pmy} + \overline{\overline{z}}_{\cdot pm\cdot}.$$

This equality algebraically illustrates how (1) utilize district-month specific interannual variation (first minus second term) after accounting for time-varying provincial shocks (third term). The last term ensures that transformations of fixed effects are not subtracted twice.

[19] For the same reason the district-month fixed effects also exclude temperature- and precipitation-day bins that are never realized within the specific district-month (i.e., zero for all years).

[20] Note that I implicitly assume that temperature and precipitation do not entirely explain climate change-induced shifts. Besides it is important to note that since the province subgroup is foremost a political rather than an agroclimatic zone I implicitly assume that differing weather distributions across provinces capture relevant differences in agroclimatic zones (NEPA, UNEP and WFP, 2016). However, if climate change induces district-specific stochastic perturbations of the number of extreme events, this will not be picked up by the province-year-month fixed effects. This is partially controlled for by the district-month-specific yearly trends $q_{dm}(y)$.

year-month fixed effects further control for conflict events spilling over to one district from nearby districts, given that they are located in the same province.[21]

Finally, I control for district-month-specific smooth quadratic yearly trends $q_{dm}(y)$ that semi-parametrically control for differential trends in conflict incidence driven by time-varying unobservables (e.g., district-level climate change-induced trends in conflict incidence and extreme weather events). The trend component also removes potential spurious regression phenomena generated by interannual coincidental time series variation (e.g., increasing media coverage of conflict events combined with trends in extreme weather events). Though there are potential determinants of conflict incidence that I could control for, these are themselves outcomes of weather variations and would, if included, bias the estimated coefficients on the temperature- and precipitation-day bins. These potentially endogenous controls were intentionally omitted.[22]

Few potential confounders seem to remain after including such a rich set of fixed effects. I, therefore, believe that my identifying assumption is satisfied. In other words, that the variation in daily temperature and precipitation is as if interannually randomly assigned across district-months.[23]

[21]More subtle is that the province-year-month fixed effects also filter out trends in the reporting of violent conflicts, or the presence of some systematically biased over-reporting of conflict events, across provinces. However, I think such trends are unrelated to my weather variables.

[22]Examples of potentially endogenous controls are several in the Afghan context: The market price for wheat or opium (Gehring, Langlotz and Kienberger, 2017); the local snow depth water equivalent that determines seasonal variation in irrigation capacity (NEPA, UNEP and WFP, 2016); and aid inflows induced by extreme weather events (Zürcher, 2017). If any of these are outcomes of weather variations and included in the baseline specification, the estimated coefficients on the temperature- and precipitation-day bins would be conditional on a given level of a control variable. Including these so-called bad controls would, however, invalidate a causal interpretation of the estimators (Angrist and Pischke, 2008, Subsection 2.2.3).

[23]Though I argue that my fixed effects are needed to ensure exogenous weather variation, they can induce overfitting and exacerbate attenuation bias by reducing the signal-to-noise ra-

There is one additional issue worth mentioning. It concerns the fact that since my baseline specification is a linear probability model, it can predict values of conflict incidence outside the strict unit interval of a probability. If there is a nonzero probability of predicting outside the unit interval, then the estimated parameters are not realizations from consistent estimators (Horrace and Oaxaca, 2006). Since there are predicted values outside the unit interval for all of my estimated linear probability models (not shown), my estimated coefficients are biased.[24]

However, this problem is unsolvable as my multidimensional fixed effects make standard solutions infeasible. Specifically, estimating a logit or probit model lead to an incidental parameter problem so that parameters become inestimable.[25] Though this removes my ability to predict more than marginal changes in temperature and precipitation, the linear probability model for binary responses is a convenient approximation of the underlying response probability and a minimum mean squared error linear approximation of the underlying conditional expectation function (Angrist and Pischke, 2008; Wooldridge, 2010). Furthermore, even if a logit or probit model would be estimable, there is no a priori reason to assume that the error term in (1) is well-modeled by the

tio of the weather variables. The extent of overfitting in my baseline specification is illustrated by the R^2-values from regressing each baseline bin on the baseline fixed effects. I find that for the temperature-day bins, the mean R^2 is about 85.26 percent, and for the precipitation-day bins, the mean R^2 is about 53.44 percent.

[24]For the estimated baseline specification in Section 5.1, about 67.68 percent of the predicted values are inside the unit interval.

[25]The incidental parameters problem was highlighted by Neyman and Scott (1948). Here the problem is that when the number of district-months goes off to infinity in the asymptotic analysis of the consistency of the estimators, incidental parameters (i.e., district-months) are inconsistently estimated and further contaminate the common parameters (i.e., the coefficients on the temperature- and precipitation-day bins). Model-specific solutions exist, but there is no unified solution (Charbonneau, 2017; Lancaster, 2000).

Gaussian or logistic distribution.

Baseline Statistical Inference.—For my baseline specification, I employ a two-way clustering design and allow the transitory shocks ϵ_{dt} to be serially correlated of unspecified form within districts and spatially correlated of unspecified form within year-months, as modeled by sandwich estimators of one-way clustered variance-covariance matrices.[26] There are two critical aspects of my two-way clustering design. First, it accounts for potential serial and spatial correlation in conflict events and my measures of temperature and precipitation.[27] This is important since failure to account for serial and spatial correlation of the error terms can overestimate precision and thereby cause erroneous statistical inference (Moulton, 1986).

The other aspect is that the size of each cluster dimension (i.e., district and year-month) must go off to infinity for the cluster-robust variance-covariance estimator to be consistent. Thus, the suitability of the design depends on the minimum of the number of districts (398 in the baseline) and the number of

[26]The key assumption is $\mathbb{E}(\epsilon_i \epsilon_j | \boldsymbol{x}_i, \boldsymbol{x}_j) = 0$ if observation i and j does *not* lie in the same district or the same year-month, where $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote all covariates for observation i and j, respectively. The two-way cluster-robust variance-covariance matrix estimator (CRVE) $\hat{V}$ that hold under this assumption is the sum of the one-way CRVEs for the first and second cluster dimension, minus the one-way CRVE for the intersection of the two dimensions (Cameron and Miller, 2015). If $\hat{V}$ is not positive-semidefinite, I follow Cameron, Gelbach and Miller (2011) and use the Eigen Decomposition Theorem to construct an alternative CRVE $\hat{V}^+$. The alternative $\hat{V}^+$ tend to be positive-semidefinite and a suitable alternative to $\hat{V}$.

[27]For example, Harari and La Ferrara (2017) use a spatial regression model and finds both significant serial and spatial correlation in their conflict variable. Furthermore, the precipitation interpolation data may mechanically introduce spatial correlation with measured precipitation even if none exist (Dell, Jones and Olken, 2014). Note though that the precipitation data generated by CHIRPS attempt to reduce this bias by estimating a set of local decorrelation structures that limit the extent of the spatial correlation in precipitation (Funk et al., 2015).

year-months (137 in the baseline) (Cameron and Miller, 2015). There is no clear-cut rule on the exact number of clusters needed. However, current consensus seems to suggest 50 (Cameron and Miller, 2015), and my minimum baseline cluster size 137 passes this threshold by almost two-and-a-half.[28]

5 Results

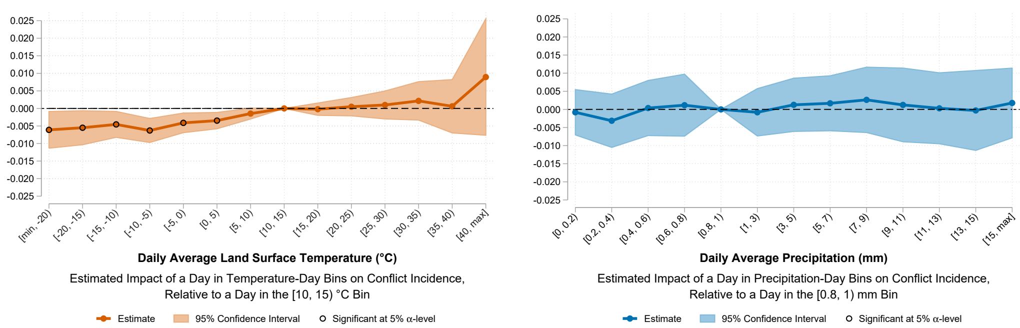
I now turn to the results. First, I report the baseline results. Second, I check if the baseline results are sensitive to alternative specifications. Third, I present suggestive evidence that temperature shocks to opium production do not explain the observed link between temperature and conflict incidence.

5.1 Baseline Results

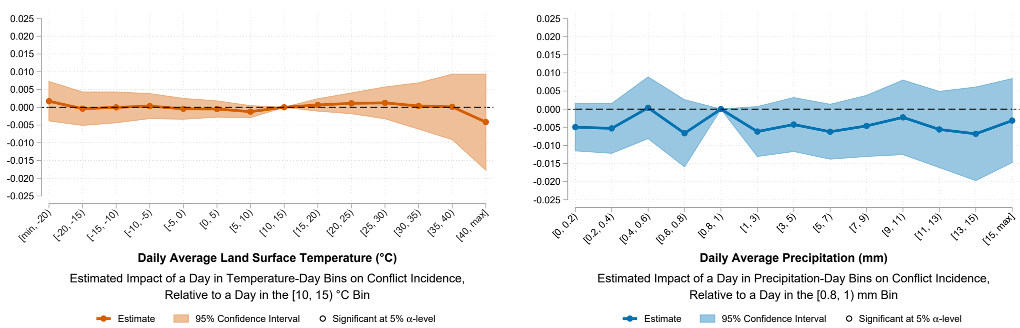
Figure 3 displays the weather-conflict relationship obtained by estimating the baseline specification (1). The red (blue) thick lines approximate the continuous dose-response relationship between conflict incidence and temperature (precipitation). To be more specific, the thick red (blue) lines mark out the estimated impacts on current conflict incidence from exchanging days with temperature (precipitation) levels in the omitted bin to a day with temperature

[28]There is no theorem to the effect that 50 clusters are enough. However, in the context of a difference-in-differences specification, Bertrand, Duflo and Mullainathan (2004) find in simulations that Wald tests based on the cluster-robust variance-covariance matrix estimator with critical value 1.96 had rejection rate 0.063 (i.e., close to 0.05). Based on another data generating process Cameron, Gelbach and Miller (2008) find that a cluster size of 30 gives the same rejection rate. Optimally, to find a suitable minimum cluster size for my two-way CRVE, I would do a Monte Carlo experiment to compute and compare several similarly defined two-way CRVEs with my baseline multidimensional fixed effects based on pseudorandom data from simulating a data generating process imitating the one observed.

(precipitation) levels in a bin indexed by the horizontal axis. The exchange is made either in the current or prior year-month, but the impact on conflict incidence is always an effect in the current year-month. For statistical inference, I add light red (blue) regions to depict 95 percent confidence bands of the temperature (precipitation) response functions.



A. CURRENT YEAR-MONTH



B. PRIOR YEAR-MONTH

FIGURE 3. CONFLICT INCIDENCE RESPONSE FUNCTIONS

Notes: The left and right figure of Panel A (B) plots estimates and 95 percent confidence bands of the contemporaneous (one year-month lagged) conflict incidence temperature and precipitation response functions—i.e., $\{\beta_i^l\}_i$ and $\{\gamma_j^l\}_j$ for $l = 0$ (1)—obtained by estimating (1). The omitted temperature- and precipitation-day bins are $[10, 15)$ °C and $[0.8, 1)$ mm. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

I find that the effect on conflict incidence from contemporaneously exchanging a day with temperature between 10-15 °C to a warmer day is insignificant, but that there is a significantly negative effect if the exchange is made for a cooler day. Using an F-test I reject the null hypothesis that all coefficients on the current temperature-day bins are zero (p-value < 0.05). This finding suggests that there is a significant link between temperature and the likelihood of a conflict. The link is also substantial. For example, exchanging a day with temperature between 10-15 °C to a day with temperature below 10 °C decrease conflict incidence by about 0.5 percent, which is 1/30 of the average probability of conflict (15 percent). Hence, the predicted decrease in conflict incidence from, e.g., five such exchanges is 1/6 of the average.

The remaining response functions are however insignificant. Indeed, using an F-test, I do not reject the null hypothesis that all coefficients on the one year-month lagged temperature-day bins, and the current and one year-month lagged precipitation-day bins, are zero (p-value ≈ 0.7).[29] This result suggests that precipitation does not drive the occurrence of conflict and that current weather variations do not delay conflicts that will eventually occur.

That contemporaneously exchanging days in the omitted temperature-day bin to colder, but not warmer, days significantly affect conflict incidence, is counterintuitive. To get a better sense of the result I vary the omitted bin in Figure 4 (cf. footnote 15). I find that exchanging a day with temperature below 5 °C for a 5-20 °C warmer day significantly raise the conflict risk. Thus, for the six figures that represent this pattern, I get the result that temperature increases above the omitted temperature-day bin lead to significant increases in the conflict risk. However, the remaining figures show that temperature in-

[29] The p-value associated with this F-test is virtually unaffected by the choice of omitted bins (not shown).

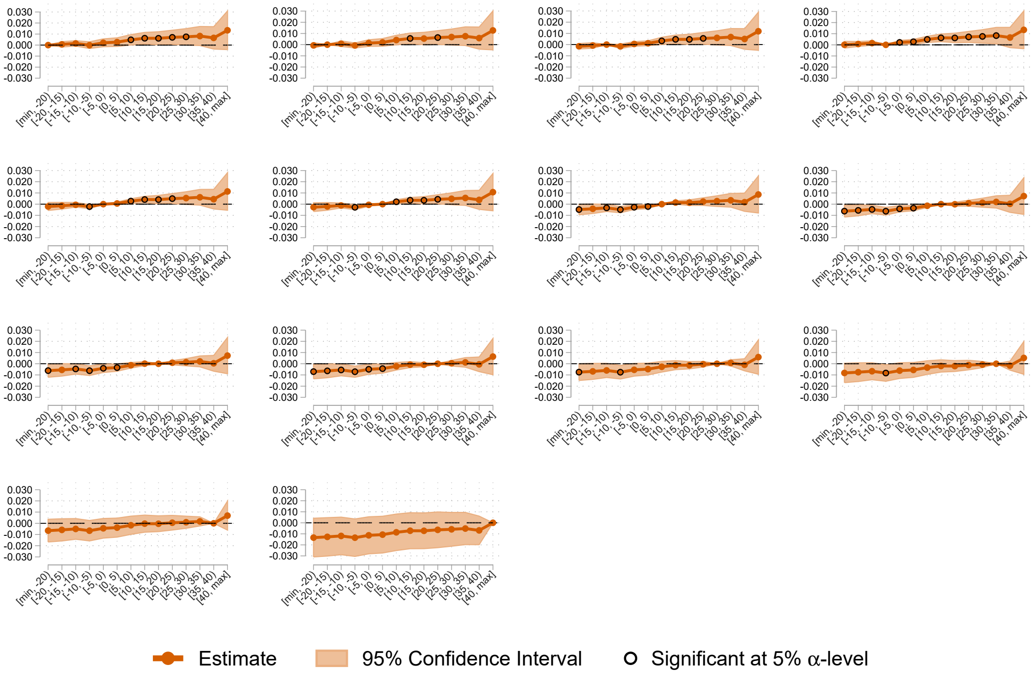


FIGURE 4. COMPLETE CONFLICT INCIDENCE TEMPERATURE RESPONSE

Notes: These figures plots estimates and 95 percent confidence intervals of $\{\beta_i^0\}_i$ obtained by estimating (1). For each temperature-day bin i there is a corresponding figure such that bin i is the omitted bin when estimating (1). The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

creases above the omitted bin do not increase the conflict risk.

To reconcile these seemingly contradictory results, note that the OLS estimates of the regression coefficients $\{\beta_i^0\}_i$ in (1) capture changes in the influence of temperature on conflict incidence relative to a baseline influence.[30] Thus, Figure 4 is to be interpreted as follows. First, that there is no significant *change* in the influence of diurnal temperature variation on the average risk of conflict from 5 °C and above. Second, that this does not imply that there is no influence of diurnal temperature variation on the average risk of conflict from 5 °C and above. The reason is that we observe a temperature treatment effect at levels between 5-35 °C (but not above 35 °C) in the sense that the influence of temperature on conflict incidence between 5-35 °C is significantly higher than the influence of ditto at levels below 5 °C.[31] Consequently, there is a significant effect of diurnal temperature variations on conflict incidence. Furthermore, the magnitude of this effect is in general nondecreasing in temperature and tends to increase with higher temperature levels.

5.2 Sensitivity Analysis

In this subsection, I provide additional justification of my baseline model and insight into my baseline temperature-conflict link. I also underline that base-

[30]Formally, $\sum_i \beta_i^0 T_{dt}^i = k_{m(t)}\beta_{i_0}^0 + \sum_{i \neq i_0}(\beta_i^0 - \beta_{i_0}^0)T_{dt}^i$, where $k_{m(t)}$ is the number of days in month m during year-month t. Hence, since $k_{m(t)}\beta_{i_0}^0$ is picked up by the (province-)year-month fixed effects, estimating (1) by OLS gives estimates of the influence of temperature on conflict incidence relative to the influence present when temperature varies in the omitted bin i_0 (i.e., $\beta_i^0 - \beta_{i_0}^0$). Also, see Appendix B where I mathematically motivate why the $\{\beta_i^0\}_i$ capture the average diurnal influence of temperature on conflict incidence.

[31]In other words, for each fixed bin $q \in \{[5,10), \ldots, [30,35)\}$ °C, I cannot, at the 5 percent α-level, reject the null hypothesis H_0 that $\beta_q^0 - \beta_r^0 = 0$ for any bin $r \in \{[5,10), \ldots, [30,35)\}$ °C, but I can reject H_0 for multiple bins r with maximum temperature levels below 5 °C, and the OLS estimates of these significant $\beta_q^0 - \beta_r^0$ are positive.

line results are robust to a battery of alternative specifications.

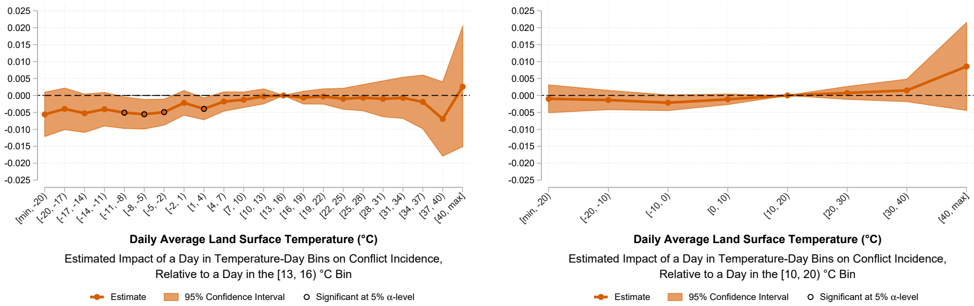
Alternative Bin Construction.—The primary functional form assumption of my baseline specification (1) is that the weather-conflict relationship is constant within the temperature- and precipitation-day bins. If this modeling assumption is false, the estimated weather-conflict relationship may mask important nonlinearities. Regarding temperature the assumption is investigated in Figure 5 where the baseline specification is re-estimated with alternative temperature-day bins. For temperature-day bins of width 3 °C, I observe no additional nonlinearity (Figure 5a). However, for temperature-day bins of width 10 °C, nonlinearities previously observed at the lower and upper end of the temperature distribution are masked (Figure 5b).[32]

The choice of temperature-day bins therefore stand between bins of width 5 or smaller (e.g., 3 °C). To make a selection, note that the constructed temperature-day bins are better measured for temperature-day bins of width 5 °C than of smaller widths.[33] Thus, since temperature-day bins of width 5 °C do not seem to mask any important additional nonlinearity when compared to temperature-day bins of width 3 °C, and are less susceptible to measurement error than bins of smaller widths, the former is a more sensible choice.

Precipitation over Afghanistan is non-normally distributed and positively

[32]Figure C.1 in Appendix C illustrates the distribution of temperature and precipitation over the alternative temperature- and precipitation-day bins discussed in this subsection. Note that the distribution of precipitation across the bins in Figure 2 and C.1 are similar but not identical.

[33]Consider a bin $b = [\underline{b}, \overline{b}]$ and let $\tilde{w}$ denote a measure of a weather variable w. Further suppose that $\tilde{w}$ is measured with additive error u such that $\tilde{w} = w + u$. Suppose $\tilde{w} \in b$ and set $l = \min(\tilde{w} - \underline{b}, \overline{b} - \tilde{w})$. Then $w = \tilde{w} - u \in b$ if and only if $|u| \leq l$. Since l is nondecreasing in the expansion of b (i.e., decreasing $\underline{b}$ or increasing $\overline{b}$) it follows that the likelihood that $|u| \leq l$ is nondecreasing in the length of the bin. That is, the measurement error of the bins tend to decrease with their width.



A. TWENTY-TWO TEMPERATURE-DAY BINS

B. EIGHT TEMPERATURE-DAY BINS

FIGURE 5. ALTERNATIVE CONFLICT INCIDENCE TEMPERATURE RESPONSE FUNCTIONS

Notes: Panel A (B) plots estimates and 95 percent confidence intervals of the effects from exchanging a day with temperature levels between 13 and 16 °C (10 and 20 °C) to a day in another temperature-day bin obtained by estimating (1) with alternative temperature-day bins. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

skewed. To account for this, I apply an alternative definition of precipitation-day bins where I first order precipitation (plus 0.001) on a log scale. Then I define precipitation-day bins as unit intervals on this scale. Using these bins I re-estimate (1). I find in Figure 6 that precipitation still plays no significant role and that the predicted pattern is similar to the baseline pattern.[34]

[34]The choice of temperature- and precipitation-day bins is somewhat arbitrary. Though diagnostic tools for assessing model fit (e.g., information criteria and coefficients of determination) could be applied to choose that set of bins which maximized fit, it is hard to correct for the degree of measurement error associated with each set. So I still prefer the analysis presented here. Furthermore, there is a curse of dimensionality problem in that there is an uncountably infinite number of different sets of bins. Optimally, I would in a first stage estimate a nonparametric model controlling for my set of fixed effects to identify important nonlinearities before continuing with my parametric analysis.

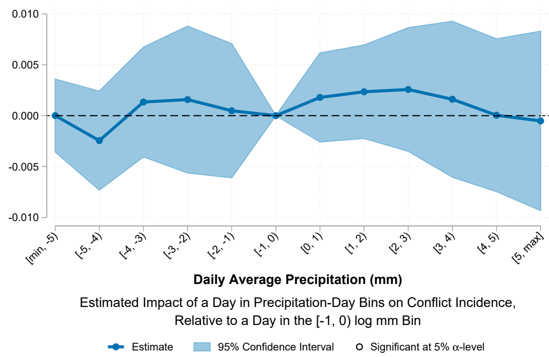


FIGURE 6. ALTERNATIVE CONFLICT INCIDENCE PRECIPITATION RESPONSE FUNCTION

Notes: This figure plots estimates and 95 percent confidence intervals of the effects from exchanging a day with log precipitation (plus 0.001) levels between -1 and 0 to another log precipitation (plus 0.001)-day bin obtained by estimating (1) with alternative precipitation-day bins. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

Onset and Ending.—About 33 percent of all conflict onsets continue with conflicts next year-month. My baseline specification ignores such persistence. Instead of including a lagged dependent variable,[35] I separately model conflict onset and ending. Conflict onset is defined as $Onset_{dt} = \mathbf{1}(C_{dt} = 1 | C_{d,t-1} = 0)$

[35] Nickell (1981) show that including a lagged dependent variable in a balanced panel make regression estimators inconsistent of $O(1/T)$, where T denote the number of observations within a group. Specifically, as the number of district-months goes off to infinity for fixed T, under homoskedasticity and no serial correlation of unobservables, the asymptotic bias on all coefficients scales like $-\frac{\mathbb{V}(\epsilon_{dt})}{T^2}\frac{(T-1)-T\rho+\rho^T}{(1-\rho)^2}$, where ρ is the autoregressive coefficient. Thus, even for $T = 10$ this analytically derived biasing factor is around -0.16 if $\mathbb{V}(\epsilon_{dt}) = 1$ and $\rho = 0.5$. Monte Carlo experiments by Judson and Owen (1999) further suggest that the bias is nonnegligible for $T = 15$. Here the groups are district-months with $T = 11$ or 12 years each, and thus including a lagged dependent variable would lead to sizable bias. Further note that the inclusion of a lagged dependent variable would be a bad control for the lagged temperature- and precipitation-day bins by construction.

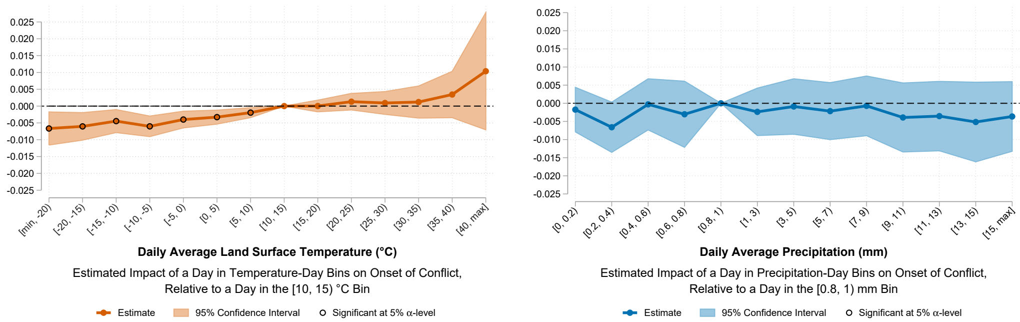
and ending as $Ending_{dt} = \mathbf{1}(C_{dt} = 0|C_{d,t-1} = 1)$. That is, the onset of a new conflict is modeled as the occurrence of a battle-related death in the current year-month given no battle-related death last year-month, and 0 otherwise; and conflict ending is similarly interpreted. The model for conflict onset (ending) is then given by replacing C_{dt} with $Onset_{dt}$ ($Ending_{dt}$) in (1). Following Bazzi and Blattman (2014), the regression for onset (ending) excludes year-months of continuing conflict (peace). If this sample selection is not made, we will constrain weather variations to have the same effect in year-months of peace as in year-months of conflict.[36]

Figure 7 and 8 show the response functions for onset and ending.[37] I find that precipitation neither plays a significant role in driving onsets nor endings. More importantly, I find that contemporaneous but not lagged temperature variations drive conflict onsets. Contemporaneous temperature does not significantly drive conflict endings, and by and large lagged temperature does not either. However, the estimated lagged conflict ending temperature response function seem to suggest that past temperature variations delay the onset of peace, and weakly significant coefficients near the $[-10, -5)$ °C bin gives weak credence to this statement. That the share of observations with continuing peace is about 77.1 percent explains the fact that estimates for endings are

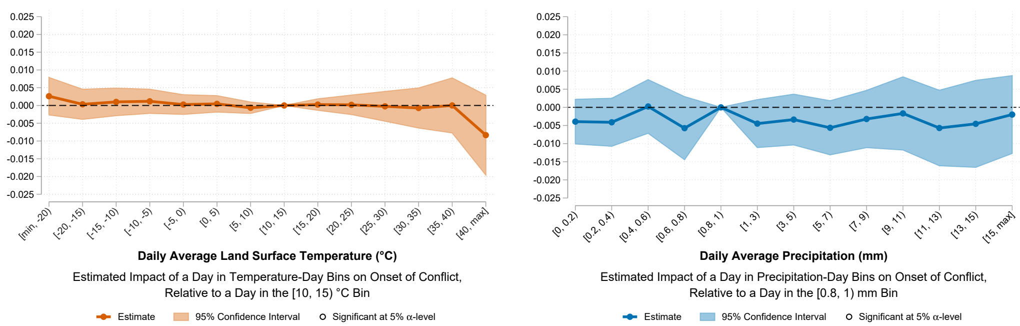
[36]It is theoretically possible to treat all cases (onset, ending, the continuation of peace and conflict) in an ordinal regression model (e.g., ordered probit or logit). However, in that case, due to the incidental parameter problem, I cannot include the multidimensional fixed effects necessary for credible causal identification.

[37]Note that the number of observations is less than for the baseline results as I for conflict onset (ending) exclude district-year-months of continuing conflict (peace), where continuing conflict (peace) is coded as $\mathbf{1}(C_{dt} = 1|C_{d,t-1} = 1) = 1$ ($\mathbf{1}(C_{dt} = 0|C_{d,t-1} = 0) = 1$). This exclusion also results in singletons (i.e., fixed effects groups with one observation) that I drop to avoid overstating statistical significance. See Appendix E for details on not maintaining singletons.

much less precise (see footnote 37). I conclude that there is still evidence for a temperature-conflict link and that the baseline contemporaneous temperature response function seems to capture an effect of temperature variations on the onsets of conflicts.



A. CURRENT YEAR-MONTH

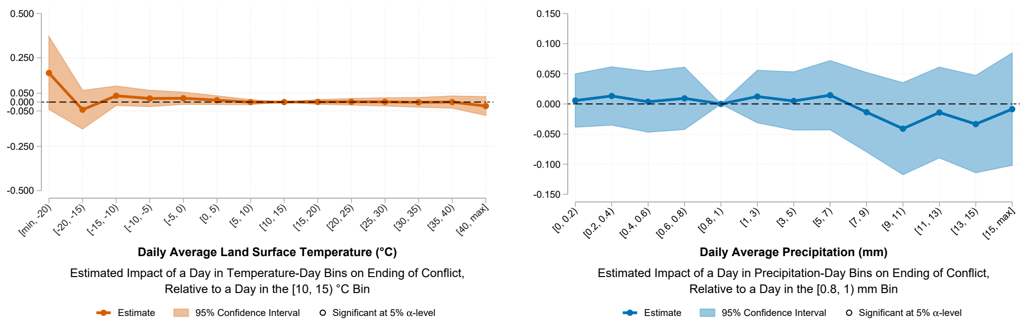


B. PRIOR YEAR-MONTH

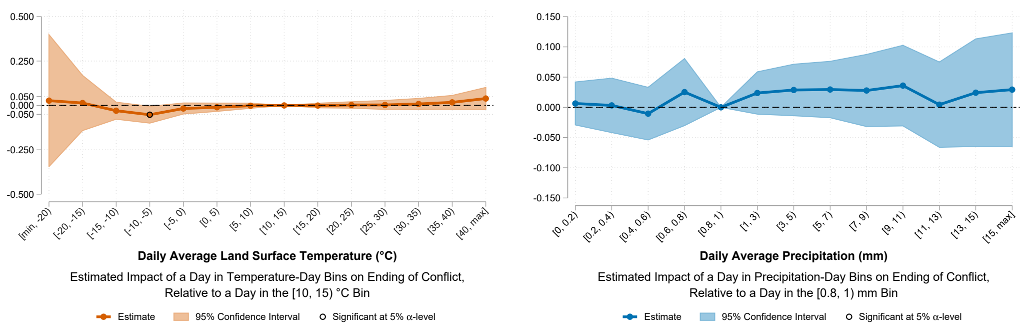
FIGURE 7. ONSET RESPONSE FUNCTIONS

Notes: The left and right figure of Panel A (B) plots estimates and 95 percent confidence bands of the contemporaneous (one year-month lagged) temperature and precipitation response functions obtained by estimating (1) after replacing the dependent variable with conflict onset and excluding year-months of continuing conflicts. The number of observations is 50,984 from an unbalanced panel of 398 districts and 137 year-months. Mean conflict onset is about 9 percent. Standard errors two-way clustered at the district- and year-month-level.

Adaption.—My baseline specification (1) does not account for adaptation as



A. CURRENT YEAR-MONTH



B. PRIOR YEAR-MONTH

FIGURE 8. ENDING RESPONSE FUNCTIONS

Notes: The left and right figure of Panel A (B) plots estimates and 95 percent confidence bands of the contemporaneous (one year-month lagged) temperature and precipitation response functions obtained by estimating (1) after replacing the dependent variable with conflict ending and excluding year-months of continuing peace. The number of observations is 10,860 from an unbalanced panel of 284 districts and 137 year-months. Mean conflict ending is about 36 percent. Standard errors two-way clustered at the district- and year-month-level.

the marginal effect of an additional hot day is by construction constant. This assumption is violated if, e.g., district-specific populations adapt to a high number of hot days within a month. Ex ante it may seem implausible that any substantial adaptation occurs within months. In any case, in an attempt to account for within district-month adaptation, I replace $\sum_{l=0}^{1}\sum_i \beta_i^l T_{d,t-l}$ with $\sum_{m=1}^{3}\sum_{l=0}^{1}\sum_i \beta_i^{lm}(T_{d,t-l}^i)^m$ in (1). That is, I include a polynomial of order 3 in each temperature- and precipitation-day bin.[38] For the resulting specification the contemporaneous total effect on conflict incidence from exchanging T^i days with temperature levels in the omitted bin to T^i days in bin i is $\sum_{m=1}^{3}\beta_i^{0m}(T^i)^m$. Estimating this specification might suggest a convex temperature-conflict link such that, e.g., within district-months populations adapt to a steady increase in the number of warm days.

Figure 9 plots estimates and 95 percent confidence intervals of $\sum_{m=1}^{3}\beta_i^{0m}(T^i)^m$ for various T^i. I find no significant evidence of adaptation within district-months. Estimated relationships are either approximately linear or, if not, insignificant in the area in which the non-linearity takes off. Thus, there is no evidence of within district-month adaptation, and the temperature-conflict link seems well-described by the baseline constant marginal effects model (1).

Seasonal Heterogeneity.—Figure 10 plots seasonal conflict incidence temperature response functions. Specifically, the temperature- and precipitation-day bins of (1) interacted with a vector of four seasonal dummies covering winter, spring, summer and autumn.[39] This adjustment allows me to test if the temperature-conflict link is heterogeneous across the seasons. One reason for expecting seasonal heterogeneity is the seasonal variation in planting seasons

[38] Deryugina and Hsiang (2017) use this type of specification in a similar context.

[39] Winter covers December to February; spring covers March to May; summer covers June to August; and autumn covers September to November.

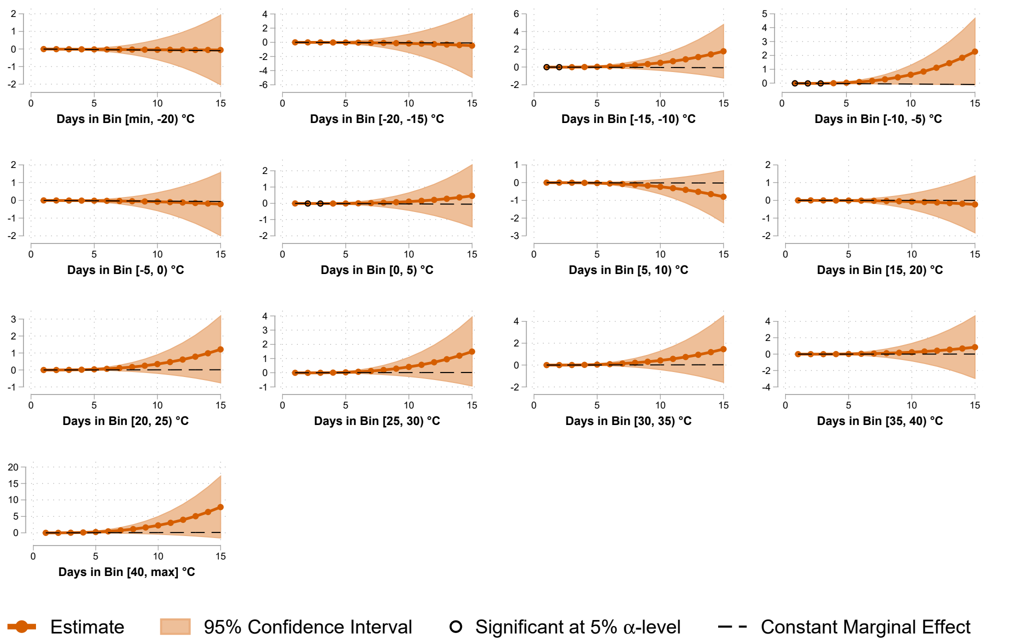


FIGURE 9. CUBIC ADAPTATION MODEL

Notes: These figures plots estimates and 95 percent confidence intervals of $\{\sum_{m=1}^{3} \beta_i^{0m}(T^i)^m\}_i$ for all $T^i \in \{1, 2, \ldots, 15\}$ obtained by estimating (1) after replacing $\sum_{l=0}^{1} \sum_i \beta_i^l T_{d,t-l}$ with $\sum_{m=1}^{3} \sum_{l=0}^{1} \sum_i \beta_i^{lm}(T^i_{d,t-l})^m$. The omitted temperature-day bin is $[10, 15)$ °C. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

(e.g., for the highly prioritized common wheat, planting tends to occur during winter and spring). Hence, if, e.g., variation in temperature affects conflicts through the agricultural sectors, we expect seasonal heterogeneity.

I find that the temperature response functions are insignificant during the summer and marginally significant during winter and autumn.[40] During spring, the temperature-conflict link is significant and alike the full sample temperature response function except at the positive end of the temperature distribution at which the estimated impact is negative (but insignificant). Further note that during winter and summer the predicted impact of high temperature levels is positive, but negative during spring and autumn. This finding suggests that there is seasonal heterogeneity in the temperature response functions.[41]

Additional Robustness Checks.—I perform various robustness checks in Appendix D. For brevity I list these here: (i) performing a falsification test where one year-month leads of all temperature- and precipitation-day bins are included (Figure D.1); (ii) testing for robustness to different standard error correction methods (Figure D.2); (iii) testing the relevance of higher order lags (two to five) of the temperature- and precipitation-day bins (Figure D.3); (iv) including spatial lags of all temperature- and precipitation-day bins (Figure D.4, D.5 and D.6); (v) replacing province-year-month fixed effects with fixed effects based on longitude and latitude (Figure D.7); (vi) controlling for district-year fixed effects (Figure D.8); and (vii) investigating the relationship between weather variations and conflict intensity (Figure D.9). Baseline results are by and large

[40] The large estimated coefficient on the $[25, 30)$ °C bin during winter seems to be an artifact of the fact that there are only seven observations that have ever occurred in this bin during winter for the full sample period.

[41] Formally, using an F-test, I reject the null hypothesis that the contemporaneous conflict incidence temperature response functions for the four seasons are equal (p-value < 0.001).

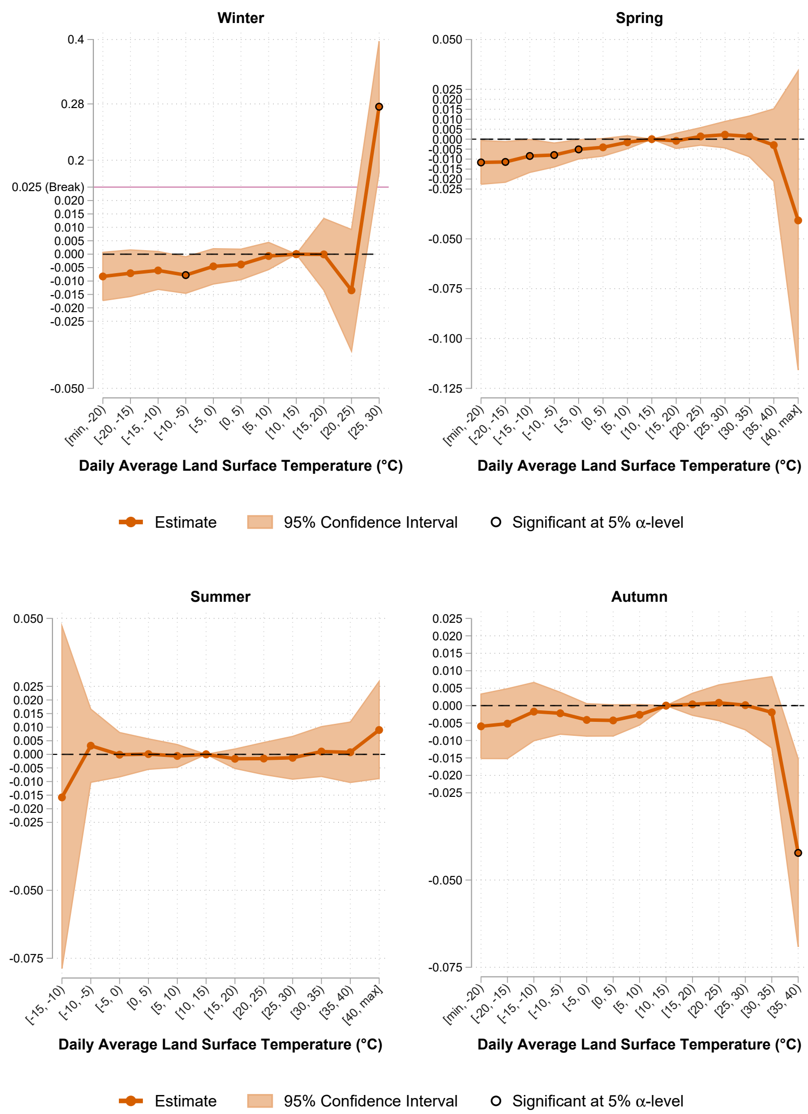


FIGURE 10. SEASONAL CONFLICT INCIDENCE TEMPERATURE RESPONSE FUNCTIONS

Notes: These figures plots estimates and 95 percent confidence intervals of seasonal conflict incidence temperature response functions. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

robust to these checks.

5.3 On the Role of Opium Production

The absence of district-level data and district-year-month-level data, in particular, make it hard to pin down what mechanism generates the baseline temperature-conflict link. Empirical studies on conflicts in Afghanistan suggests that there is a connection between opium and conflict (Bove and Elia, 2013; Gehring, Langlotz and Kienberger, 2017; Lind, Moene and Willumsen, 2014). These studies motivate a discussion of whether temperature shocks to opium production partly explain the baseline temperature-conflict link.

In this section, I provide suggestive evidence on the role of opium production in explaining the baseline temperature-conflict link. To this end, I augment my baseline specification (1) by adding

$$\sum_{l=0}^{1} \sum_{i} \theta_i^l \left(T_{d,t-l}^i \times OpiumPlantingSeason_{pm} \times \Pi_d \right), \tag{2}$$

where $OpiumPlantingSeason_{pm}$ is 1 if the historical opium planting season period in province p covers month m, and 0 otherwise; and Π_d is a district-level proxy for planned opium production and density of opium poppy sites.[42]

I first set Π_d to a dummy variable that is 1 if opium cultivation has occurred in district d over the sample period, and 0 otherwise. The use of this proxy allows me to capture the additional effect of temperature on conflict incidence during province-specific opium planting seasons for districts where opium cultivation has taken place during the sample period. Though I have annual information on opium cultivation, the use of this long-term indicator

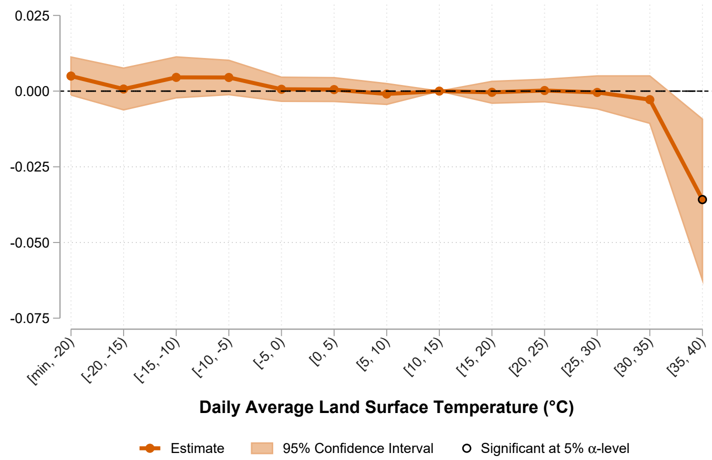
[42]Note that the term $OpiumPlantingSeason_{pm} \times \Pi_d$ is controlled for by the district-month fixed effects δ_{dm}.

represents an attempt to avoid reverse causality bias from conflict incidence to opium cultivation (Lind, Moene and Willumsen, 2014).[43]

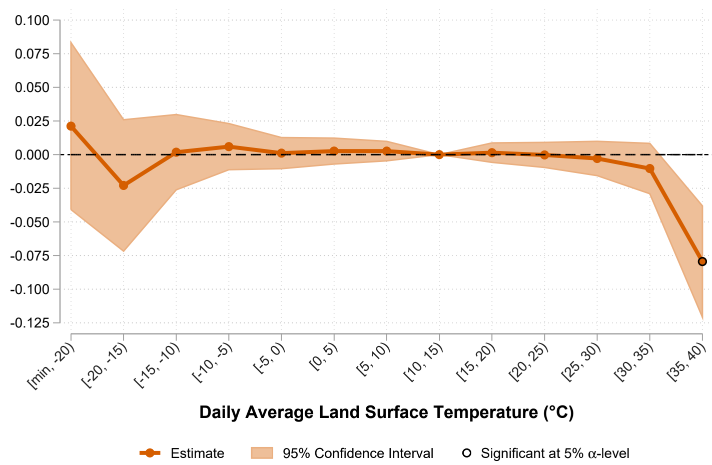
However, if one-time conflict shocks affect the likelihood of ever cultivating opium, OLS estimators of θ_i^l will be inconsistent. Also, the district-level measure of opium cultivation is indicative only and inferred from province-level statistics of opium cultivation (UNODC, 2016). These concerns motivate the use of a proxy exogenous to conflicts, namely my constructed environmental opium suitability index. Specifically, my second choice of Π_d goes from 0 to 1, where Π_d is 0 and 1 for the districts for which it is least and most suitable to grow opium poppies, respectively. The idea is that land suitable for growing opium poppies is a valuable resource for opium production and is inelastically supplied by Nature. Thus, all other things being equal, I expect planned opium cultivation to correlate positively with environmental opium suitability and production output to be dense in districts with high environmental opium suitability. The use of this proxy, therefore, allows me to capture the additional effect of temperature on conflict incidence during opium planting seasons in districts where it is relatively suitable to grow opium poppies.

I find in Figure 11a that there is no additional impact during the opium planting seasons for districts that have been growing opium poppies over the sample period for all temperature-day bins, except that in which temperature is above 35 °C. The effect is quantitatively meaningful as exchanging a day with mean temperature to a day with temperature above 35 °C during the opium planting season decrease the likelihood of a conflict incident by about 3.6 percent (s.e. about 1.3 percent). Figure 11b compares districts suitable for growing opium poppies with that in which it is least suitable. Again, I only find an additional effect at temperature levels above 35 °C, with a decrease in conflict

[43] I provide simple descriptive statistics relating to these proxies in Table C.1 in Appendix C.



A. CULTIVATION-BASED PROXY



B. ENVIRONMENTAL SUITABILITY-BASED PROXY

FIGURE 11. OPIUM SHOCK CONFLICT INCIDENCE RESPONSE FUNCTIONS

Notes: These figures plots estimates and 95 percent confidence intervals of $\{\theta_i^l\}_i$ obtained by estimating (1) after adding (2). The proxy for planned opium production and density of opium poppy sites Π used in Panel A is an indicator that is 1 if the district has ever produced opium over the sample period, and 0 otherwise. The proxy Π used in Panel B is my district-level environmental opium suitability index. The omitted temperature-day bin is $[10, 15)$ °C. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors are two-way clustered at the district- and year-month-level.

incidence by about 7.9Π percent (s.e. about 2.1Π percent), where Π is the value of the opium suitability index. The median value of Π is about 0.37, so for the 50 percent most opium suitable districts the effect is at least about 2.9 percent (s.e. about 0.78 percent).

These findings suggest that exogenous temperature shocks to opium production lower the risk of a conflict event occurring since the maximum temperature for germination of the opium poppy is about 36 °C (Kamkar et al., 2012). Now, in Subsection 5.1 we found that the influence of temperature variations in $[35, 40)$ °C was null. However, as we here see, adverse shocks to opium production induced by rising temperature seem to dampen the conflict risk. Consequently, this suggests that it cannot be temperature shocks to opium production that explains the baseline temperature-conflict link.

6 Conclusion and Discussion

In this thesis, I construct a novel panel dataset on weather and conflicts across all 398 districts of Afghanistan from July 2005 to December 2016. By fitting this dataset to fixed effects models that, I argue, allow me to utilize exogenous interannual variation in daily temperature and precipitation within district-months, I make three robust findings. First, exchanging colder for warmer days tends to significantly increase the likelihood of a conflict, and this link is quantitatively meaningful. Second, precipitation does not drive the occurrence of conflict. Third, there are no delayed effects of either variation in temperature or precipitation on conflict incidence.

I emphasize that exchanging colder for warmer days tends to, but do not always, significantly increase the likelihood of a conflict. According to my baseline results, the influence of temperature on conflict incidence stops changing from 5 °C and above. I hypothesize that this result is explained by qualitative differences in conditions present at certain temperature levels. For example, most prioritized fruits, nuts, and field crops in Afghanistan (e.g., raisins, almonds, wheat and opium poppies) hardly grow at temperature levels below 5 °C.[44] In contrast, 5-35 °C indicate suitable growing conditions. Hence, comparing days with temperature levels above 5 °C to days with temperature levels below 5 °C represent important differences in growing conditions that are unessential when comparing days with temperature levels between 5-35 °C. This argument is sound under the premise that temperature mainly affects conflict incidence via the agricultural channel. However, there are also important non-agricultural differences between days with positive and negative temperature levels. For example, precipitation falls as snow when the air temperature is below 0 °C. Hence, if the quality of roads in Afghanistan is susceptible to excessive snow covers, comparing days with air temperature below 0 °C to warmer days may represent changes in the relative likelihood for different groups to win a battle. For example, the Government of Afghanistan may have vehicles that are well-functioning in snow-covered terrain, while the Taleban do not. These changes may not occur when air temperature varies from 0 °C and above.

Thus, we see that the identified reduced-form effect of temperature on conflict incidence may operate through numerous causal pathways. The primary drawback of this thesis is that I do not pin down any specific path. For example, temperature shocks may affect economic productivity, the composition of ethnic groups, the likelihood to win battles or attitudes towards violence. However, lack of monthly information on district characteristics hinders me

[44] See the Food and Agriculture Organization of the United Nations Ecocrop database at http://ecocrop.fao.org/.

from assessing the role of these factors in driving conflicts. Nevertheless, I contribute to the literature in three ways.

First, I provide suggestive evidence that shocks to opium production induced by temperature levels above 35 °C reduce the likelihood of a conflict. Since the baseline results indicates that the net effect of temperature levels above 35 °C on conflict incidence is null, this result suggests that it is not temperature shocks to opium production that explains the observed temperature-conflict link.

Second, I present the first subnational fixed effects study on the weather-conflict relationship in Afghanistan. My finding of a significant temperature-conflict link motivates further examination of questions related to mechanisms and heterogeneity. The result that precipitation plays no significant role is also of interest. The climate in Afghanistan is dry, and about 80 percent of all cereals come from irrigated areas (NEPA, UNEP and WFP, 2016). Hence, if water supply shocks drive violence in Afghanistan, availability of irrigation water may be a more critical causal factor behind conflicts than precipitation. Irrigation-fed crops in Afghanistan are often heavily dependent on snowmelt from the Hindu Kush mountain range. It is, therefore, a natural and exciting extension of this project to use snow-related data from NASA to identify how, e.g., snowmelt-related droughts in the mountains affect district-level river flows, irrigation capacity and, in turn, conflicts.

Third, I am first to employ such a high spatial resolution to study the impact of high-frequency monthly variation in weather on intergroup conflicts. Though the temporal frequency represented by year-months, and the spatial resolution served by districts (i.e., second administrative-level regions), have been previously used (Fetzer, 2014; Maystadt and Ecker, 2014), no study utilizes both of these levels of disaggregation. Consequently, I am first to find a link between monthly variation in temperature and intergroup conflicts at the district-level.

References

Afghanistan Research and Evaluation Unit (AREU). 2017. "Mapping Nomad-Farmer Conflict in Afghanistan." Kabul: Afghanistan Research and Evaluation Unit.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Auffhammer, Maximilian, Solomon M. Hsiang, Wolfram Schlenker, and Adam Sobel. 2013. "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change." *Review of Environmental Economics and Policy*, 7(2): 181–198.

Balazsi, Laszlo, Laszlo Matyas, and Tom Wansbeek. 2018. "The Estimation of Multidimensional Fixed Effects Panel Data Models." *Econometric Reviews*, 37(3): 212–227.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S. Shapiro. 2016. "Adapting to Climate Change: The Remarkable Decline in the Us Temperature-Mortality Relationship over the Twentieth Century." *Journal of Political Economy*, 124(1): 105–159.

Baysan, Ceren, Marshall Burke, Felipe Gonzalez, Solomon M. Hsiang, and Edward Miguel. 2015. "Economic and Non-Economic Factors in Violence: Evidence from Organized Crime, Suicides and Climate in Mexico." *Unpublished manuscript, University of California, Berkeley, California, USA*.

Bazzi, Samuel, and Christopher Blattman. 2014. "Economic Shocks and Conflict: Evidence from Commodity Prices." *American Economic Journal: Macroeconomics*, 6(4): 1–38.

Bazzi, Samuel, and Matthew Gudgeon. 2016. "Local government proliferation, diversity, and conflict." *Unpublished manuscript, Boston University, Massachusetts, USA*.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig. 2017. "This Mine is Mine! How Minerals Fuel Conflicts in Africa." *American Economic Review*, 107(6): 1564–1610.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Bohra-Mishra, Pratikshya, Michael Oppenheimer, and Solomon M. Hsiang. 2014. "Nonlinear Permanent Migration Response to Climatic Variations but Minimal Response to Disasters." *Proceedings of the National Academy of Sciences*, 111(27): 9780–9785.

Bove, Vincenzo, and Leandro Elia. 2013. "Drugs and Violence in Afghanistan: A Panel VAR with Unobserved Common Factor Analysis." *Defence and Peace Economics*, 24(6): 535–554.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel. 2015. "Climate and Conflict." *Annual Review of Economics*, 7(1): 577–617.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge University Press.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3): 414–427.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Cameron, Colin A., and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources*, 50(2): 317–372.

Carleton, Tamma A. 2017. "Crop-Damaging Temperatures Increase Suicide Rates in India." *Proceedings of the National Academy of Sciences*, 114(33): 8746–8751.

Carleton, Tamma A., and Solomon M. Hsiang. 2016. "Social and Economic Impacts of Climate." *Science*, 353(6304): aad9837.

Carter, Timothy Allen, and Daniel Jay Veale. 2013. "Weather, Terrain and Warfare: Coalition Fatalities in Afghanistan." *Conflict Management and Peace Science*, 30(3): 220–239.

Charbonneau, Karyne B. 2017. "Multiple Fixed Effects in Binary Response Panel Data Models." *The Econometrics Journal*, 20(3): S1–S13.

Chassang, Sylvain, and Gerard Padro-i Miquel. 2009. "Economic Shocks and Civil War." *Quarterly Journal of Political Science*, 4(3): 211–228.

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

Croicu, Mihai, and Ralph Sundberg. 2017. "UCDP Georeferenced Event Dataset Codebook Version 17.1." *Journal of Peace Research*, 50(4): 523–532.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. "What Do We Learn from the Weather? The New Climate-Economy Literature." *Journal of Economic Literature*, 52(3): 740–98.

Deryugina, Tatyana, and Solomon M. Hsiang. 2017. "The Marginal Product of Climate." National Bureau of Economic Research.

Fetzer, Thiemo. 2014. "Social Insurance and Conflict: Evidence from India." *Unpublished manuscript*.

Fisher, Anthony C., Michael W. Hanemann, Michael J. Roberts, and Wolfram Schlenker. 2012. "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Comment." *American Economic Review*, 102(7): 3749–60.

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell, et al. 2015. "The Climate Hazards Infrared Precipitation with Stations—A New Environmental Record for Monitoring Extremes." *Scientific Data*, 2: 150066.

Gehring, Kai, Sarah Langlotz, and Stefan Kienberger. 2017. "The (Illegal) Resource Curse? Opium and Conflict in Afghanistan." *Unpublished manuscript*.

Harari, Mariaflavia, and Eliana La Ferrara. 2017. "Conflict, Climate and Cells: A Disaggregated Analysis." *Unpublished manuscript*.

Horrace, William C., and Ronald L. Oaxaca. 2006. "Results on the Bias and Inconsistency of Ordinary Least Squares for the Linear Probability Model." *Economics Letters*, 90(3): 321–327.

Hsiang, Solomon M. 2016. "Climate Econometrics." *Annual Review of Resource Economics*, 8: 43–75.

Hsiang, Solomon M., Marshall Burke, and Edward Miguel. 2013. "Quantifying the Influence of Climate on Human Conflict." *Science*, 341(6151): 1235367.

Intergovernmental Panel on Climate Change (IPCC). 2014. *Climate Change 2014 – Impacts, Adaptation and Vulnerability: Part B: Regional Aspects: Working Group II Contribution to the IPCC Fifth Assessment Report.* Vol. 2, Cambridge University Press.

Jia, Ruixue. 2014. "Weather Shocks, Sweet Potatoes and Peasant Revolts in Historical China." *The Economic Journal*, 124(575): 92–118.

Judson, Ruth A., and Ann L. Owen. 1999. "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists." *Economics Letters*, 65(1): 9–15.

Kamkar, Behnam, Majid Jami Al-Alahmadi, Abdolmajid Mahdavi-Damghani, and Francisco J. Villalobos. 2012. "Quantification of the Cardinal Temperatures and Thermal Time Requirement of Opium Poppy (*Papaver Somniferum L.*) Seeds to Germinate Using Non-Linear Regression Models." *Industrial Crops and Products*, 35(1): 192–198.

Kienberger, Stefan, Raphael Spiekermann, Dirk Tiede, Irmgard Zeiler, and Coen Bussink. 2017. "Spatial Risk Assessment of Opium Poppy Cultivation in Afghanistan: Integrating Environmental and Socio-Economic Drivers." *International Journal of Digital Earth*, 10(7): 719–736.

Kumar, Sujay V., Christa D. Peters-Lidard, Yudong Tian, Paul R. Houser, James Geiger, S. Olden, L. Lighty, Joseph L. Eastman, B. Doty, P. Dirmeyer, et al. 2006. "Land Information System: An Interoperable Framework for High Resolution Land Surface Modeling." *Environmental Modelling & Software*, 21(10): 1402–1415.

Lancaster, Tony. 2000. "The Incidental Parameter Problem Since 1948." *Journal of Econometrics*, 95(2): 391–413.

Lind, Jo Thori, Karl Ove Moene, and Fredrik Willumsen. 2014. "Opium for the Masses? Conflict-Induced Narcotics Production in Afghanistan." *Review of Economics and Statistics*, 96(5): 949–966.

Maystadt, Jean-François, and Olivier 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.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy*, 112(4): 725–753.

Moulton, Brent R. 1986. "Random Group Effects and the Precision of Regression Estimates." *Journal of Econometrics*, 32(3): 385–397.

National Environmental Protection Agency of Afghanistan (NEPA) and United Nations Environment Programme (UNEP). 2015. "Climate Change and Governance in Afghanistan." Kabul: National Environmental Protection Agency and United Nations Environment Programme.

National Environmental Protection Agency of Afghanistan (NEPA), United Nations Environment Programme (UNEP) and World Food Programme (WFP). 2016. "Climate Change in Afghanistan: What Does It Mean for Rural Livelihoods and Food Security?"

Neyman, Jerzy, and Elizabeth L. Scott. 1948. "Consistent Estimates Based on Partially Consistent Observations." *Econometrica: Journal of the Econometric Society*, 16(1): 1–32.

Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica: Journal of the Econometric Society*, 49(6): 1417–1426.

Ray, Debraj, and Joan Esteban. 2017. "Conflict and Development." *Annual Review of Economics*, 9: 263–293.

Ruttig, Thomas. 2013. "How It All Began: A Short Look at the Pre-1979 Origins of Afghanistan's Conflicts." Kabul: Afghanistan Analysts Network.

Sundberg, Ralph, and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset." *Journal of Peace Research*, 50(4): 523–532.

United Nations Office on Drugs and Crime (UNODC). 2008. "Afghanistan Opium Survey 2008." Vienna: United Nations, Office on Drugs and Crime.

United Nations Office on Drugs and Crime (UNODC). 2016. "Afghanistan Opium Survey 2016." Vienna: United Nations, Office on Drugs and Crime.

United Nations Office on Drugs and Crime (UNODC). 2017. "World Drug Report 2017." United Nations publication.

Wansbeek, Tom J., and Ruud H. Koning. 1991. "Measurement Error and Panel Data." *Statistica Neerlandica*, 45(2): 85–92.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* MIT Press.

Zürcher, Christoph. 2017. "What Do We (Not) Know About Development Aid and Violence? A Systematic Review." *World Development*, 98: 506–522.

Contents of Appendix

A Additional Data Details

A.1 Data Sources

A shapefile from the Empirical Studies of Conflict Project (ESOC) define the division of Afghanistan into administrative units and their respective boundaries. The shapefile is accessible at https://esoc.princeton.edu/files/administrative-boundaries-398-districts.

The dataset on conflict events is the UCDP GED Global version 17.2. It is the most recent version of the UCDP GED. The dataset is accessible as a shapefile at http://ucdp.uu.se/downloads/.

Raster data on temperature comes from a product developed by the LIS team at NASA GSFC. The raster data is the NASA LIS version 7 Noah-36. The data is not publicly available, but NASA GSFC provides subsets of the data upon permission. Contact information is available at https://www.nasa.gov/content/contact-goddard.

Raster data on daily precipitation comes from the CHIRPS v2.0 Global Daily netCDF products. The raster data is accessible at the FTP server ftp://ftp.chg.ucsb.edu/pub/org/chg/products/CHIRPS-2.0/global_daily/netcdf/p05/.

To construct my opium suitability index, I used several sources. First, global land cover maps from GlobCover 2009 at http://due.esrin.esa.int/page_globcover.php (Arino et al., 2012). Second, DIVA-GIS-specific world climate data between 1950 and 2000 at 2.5 arcminutes resolution from WorldClim version 1.3 at http://www.diva-gis.org/climate. Third, two river networks over parts of Asia and Southwest Asia from the United States Geological Survey (USGS) mapping product Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales (HydroSHEDS) at https:

//hydrosheds.cr.usgs.gov/datadownload.php?reqdata=15rivs.

The UNODC provide indicative district-level data on opium cultivation. Contact information is available at https://www.unodc.org/unodc/en/about-unodc/contact-us.html. The data is also accessible from the UNODC Afghanistan Opium Surveys 2005-2016 available at https://www.unodc.org/unodc/en/crop-monitoring/index.html?tag=Afghanistan. Information on the typical period for opium planting at the provincial level comes from the UNODC Afghanistan opium surveys (UNODC, 2008; UNODC, 2013) available at https://www.unodc.org/unodc/en/crop-monitoring/index.html?tag=Afghanistan.

A.2 District Assignment of UCDP GED Events

Here I highlight a caveat related to the UCDP GED Global version 17.2. It concerns the fact that I cannot confirm that the district polygons in the ESOC shapefile contain the correct coordinates used in the UCDP assignment of UCDP GED events to districts. Because though the UCDP register the longitude and latitude of a UCDP GED event to the centroid of a district if the event is known to occur at the district-level, but not at a more fine-grained level, they use no, and cannot provide any, shapefile with information on the particular district centroid used. Instead, the UCDP use gazetteers. The gazetteers used for all UCDP GED events is unknown. However, most of them are in the GEOnet Names Server at https://www.nga.mil/ProductsServices/GeographicNames/Pages/default.aspx (Petterson, Therese, and Stina Högblad at Uppsala University, personal communication, March 13 and 20, 2018).

To investigate this further, I noted that the UCDP internally track administrative changes in a data structure called a geotree. Upon contact, the UCDP provided the geotree for Afghanistan in the form of a table. The table confirms that compared to the ESOC shapefile, from July 2005 and onwards, the UCDP has used a more fine-grained subdivision of Afghanistan for all UCDP GED events and that the geotree register all districts in the ESOC shapefile. Thus, there is suggestive, but not conclusive, evidence that the number of false assignments is low when I assign UCDP GED events to the districts of Afghanistan as defined by the ESOC shapefile.

A.3 Area-Weighting

The area-weighted averages of temperature and precipitation were computed as unweighted arithmetic averages across degree grid cells in the World Geodetic System 1984 (WGS84) Geographic Coordinate System. Formally, let w_{gdk} denote temperature or precipitation in grid cell g in district d day k. Then the area-weighted average is

$$w_{dk} \equiv \frac{\sum_{g \in \mathcal{G}_d} w_{gdk}}{|\mathcal{G}_d|}, \tag{A.1}$$

where $\mathcal{G}_d$ is the set of grid cells whose centroids fall inside that of district d's boundary as defined by a polygon in the ESOC shapefile.

There are two objections to this area-weighting procedure. The first is the following. Since an oblate spheroid can approximate the figure of the Earth, the area of each grid cell varies in longitude and latitude. Therefore it would be sensible to first project the georeferenced weather data to a suitable coordinate system with units in meters.[45] Then one average across meter grid cells of

[45]For Afghanistan—as an Atlas of Earth shows—it would be sensible to project the raster data on the Universal Transverse Mercator (UTM) 41N and 42N zones. This projection is believed to introduce little distortion along both dimensions as it projects small chunks of the surface of the Earth onto a flat surface (Dell, 2009).

equal area. My response is that the method proposed by the objection may introduce more error into the measured weather variables as it involves the use of a resampling algorithm when projecting the data. Furthermore, since districts are tiny, grid cell areas are approximately constant within districts of Afghanistan.

The other objection against my weighting scheme emphasizes that I do not weight by, e.g., population or the area of agricultural land. My response is that such a weighting scheme is unnecessary in this case as I focus on the reduced-form total effect of weather variations on conflict incidence. Other weighting schemes would emphasize some particular aspect. For example, weighting by population size is sensible if we, e.g., have the hypothesis that rising temperature levels can affect people's propensity to do violence against others.[46] Another example is to weight by the agricultural land area, which is more sensible if we focus on some agricultural mechanism that is supposed to explain the weather-conflict relationship. However, since I focus on the overall, rather than any particular, average treatment effect of weather variations on conflict incidence, neither of these weighting schemes is appropriate for my research question.

A.4 Data Processing

Data processing was first carried out using ArcPy. ArcPy is a module in Python 2.7 included in ArcGIS 10. I wrote Python scripts that imported the ArcPy geoprocessing tools to process the georeferenced datasets. The scripts accomplished the following tasks.

- To match UCDP GED events to districts of Afghanistan as defined by the

[46]See Baysan et al. (2015) for a recent discussion on the role of noneconomic psychological and physiological factors in driving the temperature-conflict relationship.

ESOC shapefile of Afghanistan. Specifically, to intersect the UCDP GED and ESOC shapefile.

- To compute the area-weighted average of precipitation for each day and district of Afghanistan. Specifically, to for loop over a folder of the CHIRPS v2.0 daily precipitation datasets and compute the area-weighted average for all districts within each loop.

- To compute the area-weighted average of temperature for each day and district of Afghanistan. Specifically, to iteratively spawn subprocesses, each of which processes four days of NASA LIS version 7 Noah-36 temperature data at a time.[47]

Together with the ArcPy module and DIVA-GIS (Hijmans et al., 2001) I also accomplished the following task:

- To compute the opium suitability index. First, normalize the GlobCover 2009 global land cover map with the normalized values in Table 2 in Kienberger et al. (2017). Second, with DIVA-GIS, compute a normalized climatic suitability index of opium poppy (*Papaver somniferum*) based on tolerable climatic conditions specified in Table 3 in Kienberger et al. (2017). Third, construct a proxy for water availability by computing and normalizing the river density across Afghanistan based on two river networks over parts of Asia and Southwest Asia from the USGS HydroSHEDS. Fourth, compute a weighted arithmetic average based on (Analytical Hierarchy Process) weights in Kienberger et al. (2017) that are suggested by

[47] Due to the high resolution of the temperature datasets, the size of a folder containing all daily temperature datasets is about 3.5 terabytes as for each day there is a corresponding dataset of size about 500 megabytes. Since the ArcPy module (ultimately ArcGIS 10) reserves memory space for each computation within a for loop, I had to circumvent the finite barrier of the operating system by processing the data using a set of subprocesses.

a sample of six expert consultants.[48] Finally, aggregate to the district-level and normalize to a value between 0 and 1.[49]

The output data that resulted from the completion of the above tasks were imported to Stata 14 for cleaning and analysis. Replication files are available upon request (including results not shown).

[48] In contrast to Kienberger et al. (2017) I do not account for varying soil suitability across Afghanistan. Two reasons explain this. One, the six experts consulted in Kienberger et al. (2017) assigned a small weight (11 percent) to its importance as a measure of environmental poppy suitability. Two, the exact weighting scheme of the FAO-74 soil classification system was not disclosed by Kienberger et al. (2017).

[49] Normalizing is given by the transform $x_i \mapsto \frac{x_i - \min_i(x_i)}{\max_i(x_i) - \min_i(x_i)}$ for all districts i.

B Mathematical Appendix

Daily weather data can be used to approximate nonlinear weather-conflict relationships.[50] To see this, suppose that the weather-conflict relationship over some fixed sample period $\tau = [\underline{t}, \overline{t}]$ is

$$C_{dt} = \sum_{k \in \mathcal{K}_t} f(w_{dtk}) + \eta_{dt}, \quad t \in \tau, \tag{B.1}$$

where f is a continuous function; C_{dt} is some measure of conflicts (e.g., conflict incidence) in district d time period t (e.g., year-month); w_{dtk} is a weather event of district d day k in time period t that is a realization from a climatic element whose empirical distribution function across the days has a compact support $\mathcal{S}_{dt}$ that is specific for each district d and time period t;[51] $\mathcal{K}_t$ is the set of days in time period t; and η_{dt} is a nuisance parameter containing controls, determinants, fixed effects and error terms in a regression model with C_{dt} as the dependent variable (e.g., my baseline specification (1)). In other words, for each district, the assumption is that effect of weather on conflicts during a given year-month is additively separable into a sum of continuous diurnal weather effects.

The importance of this assumption is its implication. For if given, my baseline specification retrieves inherently nonlinear effects at the daily level. Indeed, as Proposition 1 below shows, Assumption B.1 implies that a linear parametric function with parameters describing averages of nonlinear weather effects approximates the effect of the particular weather phenomena on conflict.

Proposition 1 (Binned Weather-Conflict Dose-Response Function)**.** *Assumption*

[50] See Subsection 4.1 in Hsiang (2016) for an overview on how to identify nonlinear effects of weather variations.

[51] The support define the physical limits of weather variations. For a formalization, see Hsiang (2016).

B.1 justifies the approximation

$$C_{dt} \approx \sum_b \beta_b w_{dt}^b + \eta_{dt}, \quad t \in \tau, \tag{B.2}$$

where for a bin b of $\mathcal{S}_{dt}$, $w_{dt}^b = \sum_{k \in \mathcal{K}_t} \mathbf{1}_b(w_{dtk})$ with $\mathbf{1}_b(w_{dtk}) = 1$ if $w_{dtk} \in b$ and 0 otherwise.

Proof. Let $t \in \tau$ and partition $\mathcal{S}_{dt}$ into nontrivial and nonoverlapping bins.[52] Then, since f is continuous, by the Mean Value Theorem for Integrals, for each bin $b = [\underline{w}_b, \overline{w}_b]$ there is a real $\xi_b \in b$ such that

$$f(\xi_b) = \frac{1}{\overline{w}_b - \underline{w}_b} \int_{\underline{w}_b}^{\overline{w}_b} f(w)\, dw. \tag{B.3}$$

The above equation justifies the approximation

$$f(w_{dtk}) \approx \sum_b f(\xi_b) \mathbf{1}_b(w_{dtk}). \tag{B.4}$$

Equation B.3 and Assumption B.1 then implies

$$C_{dt} = \sum_{k \in \mathcal{K}_t} f(w_{dtk}) + \eta_{dt} \tag{B.5}$$

$$\approx \sum_{k \in \mathcal{K}_t} \sum_b f(\xi_b) \mathbf{1}_b(w_{dtk}) + \eta_{dt} \tag{B.6}$$

$$= \sum_b \underbrace{f(\xi_b)}_{\beta_b} \underbrace{\sum_{k \in \mathcal{K}_t} \mathbf{1}_b(w_{dtk})}_{w_{dt}^b} + \eta_{dt}. \tag{B.7}$$

■

From Proposition 1, two important corollaries follows. First, (B.2) shows that the parameters $\{\beta_b\}_b$ can be retrieved by regressing C_{dt} on $\{w_{dt}^b\}_b$, after accounting for terms in η_{dt}.

[52] The partitioning is possible as the empirical support is countable.

Second, each β_b identifies an average effect of potentially nonlinear effects of weather on conflict, as seen in (B.3). It is therefore beneficial concerning functional form specification to tighten the length of each bin as it makes the functional form restriction described above more plausible. Though decreasing the length of each bin reduces the approximation error theoretically, in practice, it increases the measurement error (cf. footnote 33) and also the number of parameters to be estimated so that the degrees-of-freedom decreases. In the limit as the length of each bin approaches zero, the number of in-sample perfectly multicollinear w_{dt}^b increases and estimation cannot be carried out using OLS. This trade-off cannot be parametrically resolved using known statistical inferential tools. Instead, the trade-off was experimentally addressed by cutting the support of the empirical weather distribution at end points and then choosing bins of equal lengths within the remaining support such that the nonlinear weather-conflict relationship was retrieved, as discussed in Section 4 and 5.

Augmenting Assumption B.1 with the following assumption yields my baseline specification (1):

$$C_{dt} = \sum_{k \in \mathcal{K}_t} \boldsymbol{f}(\boldsymbol{w}_{dtk}) + \eta_{dt}, \tag{B.8}$$

where $\boldsymbol{f}$ is a multivariate function and $\boldsymbol{w}_{dtk}$ a vector of weather variables. Under the assumption that temperature and precipitation are sufficient statistics for capturing the weather-conflict relationship I let $\boldsymbol{f}$ be bivariate and $\boldsymbol{w}_{dtk} = (T_{dtk}, P_{dtk})$. To acquire (1) I assume that the effect of temperature and precipitation is additively separable; i.e., that

$$\boldsymbol{f}(\boldsymbol{w}_{dtk}) = f_1(T_{dtk}) + f_2(P_{dtk}), \tag{B.9}$$

where f_1 and f_2 are two univariate continuous functions. Now, by analogy, following the steps of Proposition 1 we acquire (1).

To conclude: I have argued that my baseline specification (1) approximate any continuous weather-conflict relationship (linear or nonlinear). However, I do not derive the statistical approximation error.[53] Lastly, I note that one can relax the assumption that the functions are invariant across space and time by interacting weather variables with, e.g., regional or seasonal dummies. It is also simple to include temporally and spatially lagged functions (see, e.g., Appendix D).

[53]It would be of interest to statistically assess the error made. One idea is the following. First, view the empirical support of temperature and precipitation as a set of 2-cells. Second, with, e.g., 14 temperature-day, and 13 precipitation-day, bins as in my baseline specification, define $14 \cdot 13 = 182$ 2-cells. Then, adjust (1) by excluding the temperature- and precipitation-day bins, and including all these 2-cells. Lastly, test if the temperature response functions vary by precipitation 1-cells. This test was carried out on my dataset. However, because of the fixed effects, many 2-cells were omitted due to multicollinearity, making the empirical content of the test uncertain.

C Additional Descriptive Statistics

This section provides additional descriptive statistics. Figure C.1 illustrates the alternative temperature- and precipitation-day bins used in Section 5.2, and Table C.1 provide descriptive statistics relating to the analysis in Section 5.3.

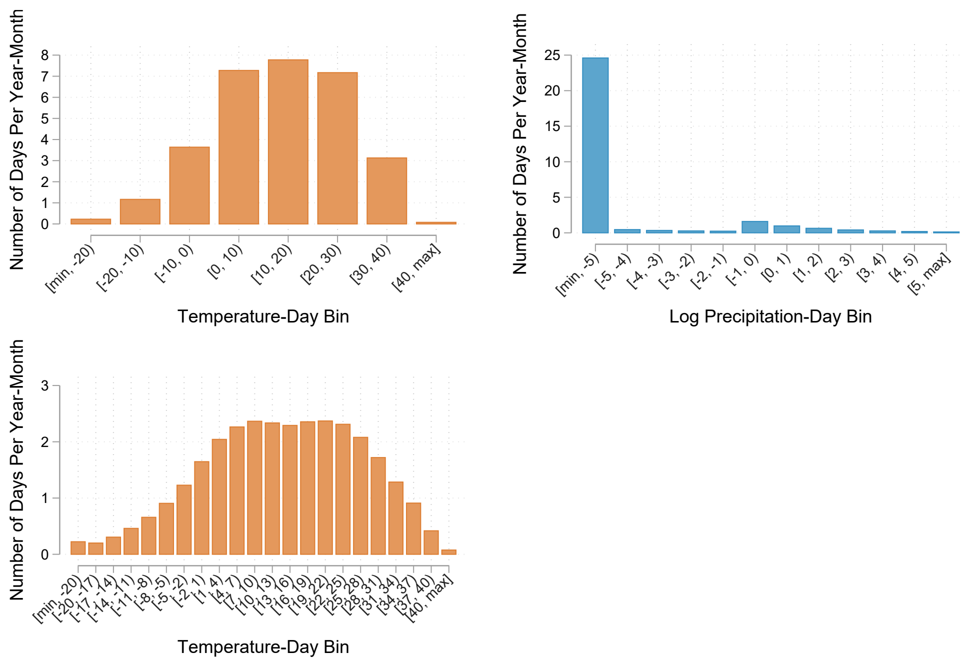


FIGURE C.1. ALTERNATIVE TEMPERATURE- AND PRECIPITATION-DAY BINS

Notes: The figure shows the average distribution of daily average temperature and precipitation across 8 and 22 temperature-day bins (upper and lower left panel) and 13 log precipitation-day bins (upper right panel). Each bar represents the average number of days per year-month in each temperature or precipitation category across all 398 districts of Afghanistan over the sample period July 2005 to December 2016. Minimum daily temperature is about −41 °C, and maximum daily temperature is about 45 °C. Maximum daily precipitation is about 117 mm. See the text for more details.

Source: Author's calculations based on data from the CHIRPS and NASA GSFC.

TABLE C.1—DESCRIPTIVE STATISTICS BY OPIUM GROWING STATUS: WEATHER AND CONFLICTS

			SD	
	Observations	Mean	Overall	Within
Panel A: Opium Growing Districts				
Battle-related deaths†				
All	36,432	1.59	8.98	8.17
If > 0	5,636	10.28	20.78	17.53
$\mathbf{1}$(Battle-related deaths > 0)†	36,432	0.15	0.36	0.30
Conflict Events†				
All	10,453			
Government of Afghanistan vs. Taleban	9,537			
Taleban vs. Civilians	450			
Government of Afghanistan vs. IS	253			
Daily Temperature (°C)‡	1,205,688	14.50	13.26	3.66
Daily Precipitation (mm)‡	1,205,688	0.92	3.10	2.94
Opium Suitability Index$^{\mathbb{P}}$	264	0.38	0.16	
Panel B: Non-Opium Growing Districts				
Battle-related deaths†				
All	18,492	1.06	6.73	6.73
If > 0	2,460	7.96	16.89	13.30
$\mathbf{1}$(Battle-related deaths > 0)†	18,492	0.13	0.34	0.29
Conflict Events†				
All	4,015			
Government of Afghanistan vs. Taleban	13,246			
Taleban vs. Civilians	196			
Government of Afghanistan vs. IS	10			
Daily Temperature (°C)‡	611,978	10.88	12.88	3.61
Daily Precipitation (mm)‡	611,978	1.02	3.39	3.23
Opium Suitability Index$^{\mathbb{P}}$	134	0.38	0.19	

Notes: This table reproduces the information in Table 1 and information on opium suitability by opium growing status. A district is opium growing if it has in any year between 2006 to 2016 cultivated opium poppies. The summary statistic Overall SD stands for the overall standard deviation of the corresponding variable. The summary statistic Within SD stands for the overall standard deviation of the corresponding variable after removing district-month fixed effects. The variable $\mathbf{1}$(Battle-related deaths > 0) is 1 if there is at least one battle-related death, and 0 otherwise. The variable Opium Suitability Index is the constructed district-level index for environmental suitability for opium poppy cultivation. The acronym IS stands for Islamic State (of Iraq and Syria). The sample period is July 2005 to December 2016. All 398 districts are included in the sample. Numbers are correct to two decimal places. †Measured at district-year-month-level. ‡Measured at district-year-month-day-level. $^{\mathbb{P}}$Measured at district-level.

Source: Author's calculations based on data from the CHIRPS, NASA GSFC and UCDP.

D Additional Robustness Checks

I now show that my baseline results in Section 5.1 are robust to a series of tests.

Falsification Test.—I plot estimates of $\{(\beta_i^{-1}, \gamma_j^{-1})\}_{(i,j)}$ from

$$C_{dt} = \sum_{l=-1}^{1} \Big(\sum_i \beta_i^l T_{d,t-l}^i + \sum_j \gamma_j^l P_{d,t-l}^j \Big) + \delta_{dm} + \pi_{pt} + q_{dm}(y) + \epsilon_{dt}. \qquad \text{(D.1)}$$

That is, I include one year-month leads of the temperature- and precipitation-day bins to my baseline specification. Since I do not expect future diurnal variation in temperature and precipitation to affect current conflict incidence, I hypothesize that the coefficients on the leads are all equal to zero. However, using an F-test, I reject this hypothesis at the 5, but not 1, percent α-level. Still, as made clear by Figure D.1, this seems to be due to a seemingly spurious coefficient on the $[0.2, 0.4)$ mm bin, as all others are insignificant. Indeed, if I ignore the $[0.2, 0.4)$ mm bin, I cannot reject the hypothesis at the 10 percent α-level. This suggests that, overall, forward temperature and precipitation play no significant role in driving conflict incidence.

Different Standard Error Corrections.—In my baseline specification, I two-way cluster at the district- and year-month-level. This clustering design allows for serial correlation within districts and spatial correlation within year-months. This section tests for robustness to three alternative designs.

The baseline design does not account for temporally lagged spatial dependence (e.g., the dependence of observations in two districts separated in time). I, therefore, check for robustness to two other two-way clustering designs. First I cluster by district and region-year. This design allows for arbitrary correlation within region-years (minimum cluster size is 72). That is, all district-months

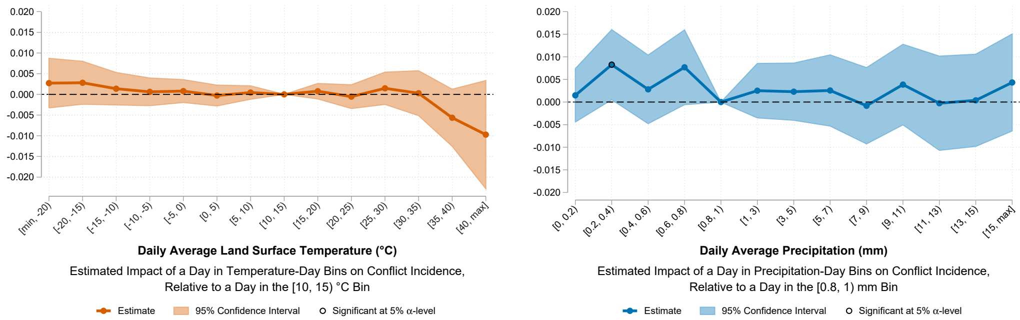


FIGURE D.1. CONFLICT INCIDENCE RESPONSE FUNCTIONS NEXT YEAR-MONTH

Notes: These figures plots estimates and 95 percent confidence bands of the one year-month lead conflict incidence temperature and precipitation response functions obtained by estimating (D.1). The omitted temperature- and precipitation-day bins are $[10, 15)$ °C and $[0.8, 1)$ mm. The number of observations is 54,128 from a balanced panel of 398 districts and 136 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

that lie in the same region-year (e.g., north of Afghanistan in 2006) are allowed to depend on each other both spatially and temporally. Second, since the former assumes unobservables in district-months in different regions are unrelated, I also cluster by district and season-year, which allows for correlation between all district-months within a season-year (e.g., spring 2006) (minimum cluster size is 47).

I also consider robustness to Conley (1999) spatial-HAC standard errors. The spatial-HAC correction account for heteroskedasticity and within-time spatial and within-location serial correlation of unobservables.[54] For the spatial correction, I use a uniform kernel that is assumed to discontinuously fall from 1 to 0 at some spatial cutoff distance (Conley, 2010). I let the cutoff distance be so large that arbitrary spatial correlation is allowed for (i.e., vanishes at 1,000 kilometers). For the HAC correction, I use the Newey-West (Bartlett) kernel

[54]Implementation is based on Fetzer (2014) and Hsiang, Meng and Cane (2011).

that weight pairs of observations within districts such that the weights decay linearly across time periods. I assume that there is no bound to the serial correlation (i.e., vanishes at 100,000 year-months). The Conley (1999) spatial-HAC standard error correction is conceptually similar to my baseline two-way clustering design but is not computationally equivalent.

Figure D.2 show the result from re-estimating the baseline specification and employing the three above adjustments. Note that I also include the baseline standard error correction and one-way clustered standard errors for completeness. Since I am unable, for technical reasons, to include the trend component $q_{dm}(y)$ when using the Conley (1999) spatial-HAC standard error correction, I further exclude $q_{dm}(y)$ for comparability. The figure shows that the two-way clustering designs yield nearly identical confidence bands, but the baseline two-way clustering design produces a wider confidence band at the highest temperature-day bin. Confidence bands based on the Conley (1999) spatial-HAC correction are less conservative. The one-way clustered confidence bands are only marginally smaller than the two-way equivalents. In conclusion, in comparison to these adjustments, my baseline standard error correction method does not overestimate standard errors and is in a worst case scenario too conservative.

Temporal Lag Length.—I consider the significance of delayed effects of higher order by adding two to five year-month lags to my baseline specification (1). I find in Figure D.3 that none of these delayed effects are significant for either temperature or precipitation.

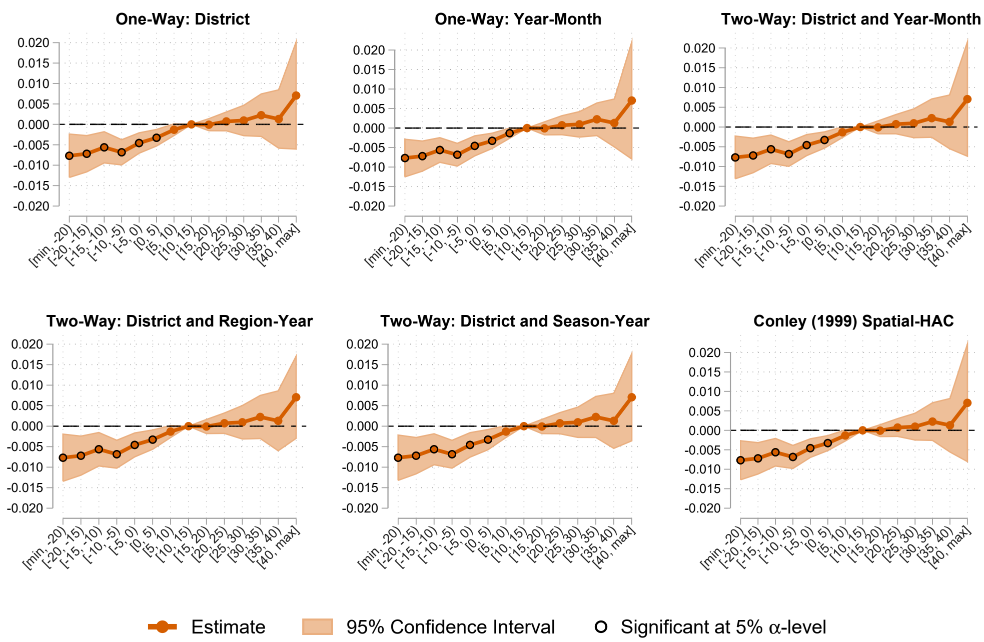
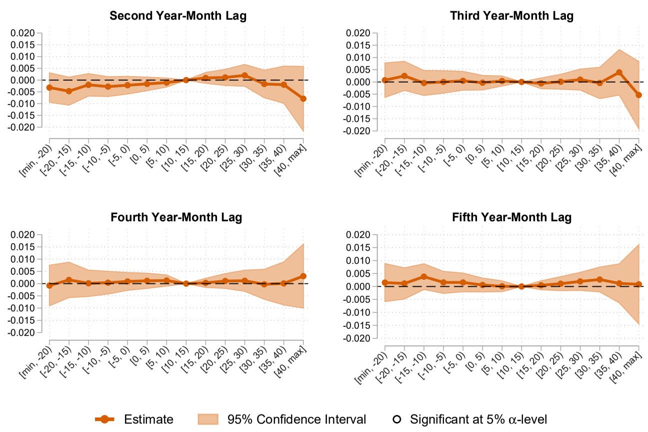
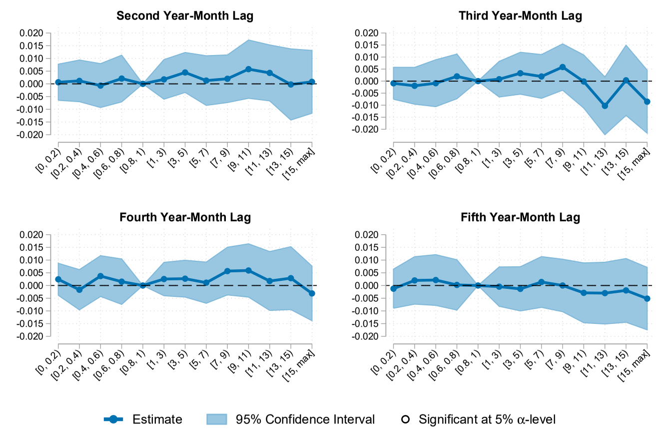


FIGURE D.2. CONFLICT INCIDENCE TEMPERATURE RESPONSE FUNCTIONS—CONFIDENCE BANDS BY STANDARD ERROR CORRECTION METHOD

Notes: These figures plot estimates and 95 percent confidence intervals of $\{\beta_i^0\}_i$ obtained by estimating (1) after excluding the trend component $q_{dm}(y)$. Standard errors are corrected using six different standard error correction methods. Upper left panel: One-way clustering at the district-level (cluster size is 398). Upper middle panel: Clustering at the year-month level (cluster size is 137). Upper right panel: Two-way clustering at district- and year-month-level (minimum cluster size is 137). Lower left panel: Two-way clustering at the district and region-year level (minimum cluster size is 72). Lower middle panel: Two-way clustering at the district and season-year level (minimum cluster size is 47). Lower right panel: Conley (1999) spatial-HAC correction with 1,000 km as the spatial distance cutoff in a uniform kernel weight function, and Newey-West (Bartlett) kernel with 100,000 year-months as the temporal distance cutoff. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent.



A. Lagged Temperature response functions



B. Lagged Precipitation response functions

Figure D.3. Lagged conflict incidence response functions

Notes: These figures plots estimates and 95 percent confidence intervals of two to five year-month lagged conflict incidence temperature and precipitation response functions obtained by estimating (1) with the second to fifth temporal lag order of all temperature- and precipitation-day bins added. The number of observations is 52,934 from a balanced panel of 398 districts and 133 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

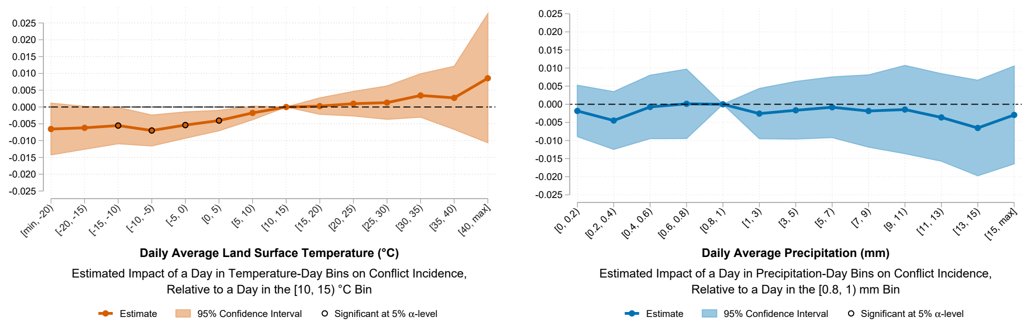
Spatial Spillovers.—Climatic events can cause outcomes to be displaced across space and spill over to neighboring districts. Spillovers at the provincial-level have been accounted for by the province-year-month fixed effects of my baseline specification. However, I do not consider spillovers across provincial boundaries. Spillovers across provincial boundaries due to weather variations can be partly accounted for by the following model:

$$C_{dt} = \sum_{l=0}^{1} \sum_{\pi=0}^{3} \Big(\sum_{i} \beta_i^{\pi l} T^i_{\pi d,t-l} + \sum_{j} \gamma_j^{\pi l} P^j_{\pi d,t-l} \Big) + \delta_{dm} + \pi_{pt} + q_{dm}(y) + \epsilon_{dt}, \quad \text{(D.2)}$$

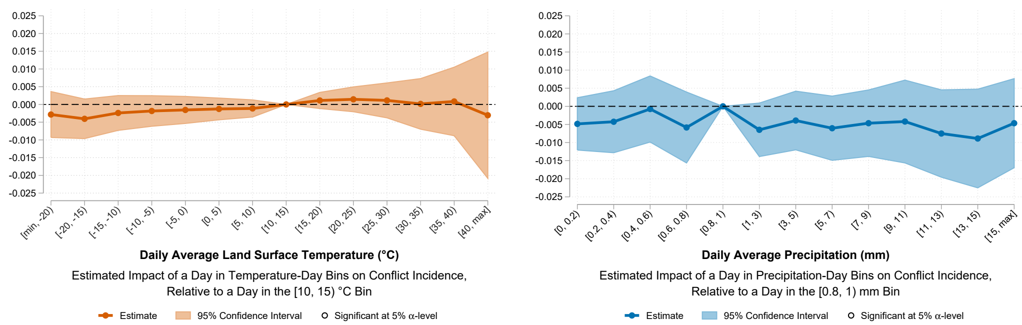
where $T^i_{\pi d,t-l}$ ($P^j_{\pi d,t-l}$) for $\pi > 0$ denote the average number of days with temperature (precipitation) levels in bin i (j) of districts whose centroids lie within a $(\pi - 1, \pi]$ one hundred kilometer band from the centroid of district d and $T^i_{0d,t-l} \equiv T^i_{d,t-l}$ ($P^j_{\pi d,t-l} \equiv P^i_{d,t-l}$). The coefficients $\{(\beta_i^{\pi l}, \gamma_j^{\pi l})\}_{(i,j)}$ for $\pi > 0$ capture spatial spillovers such that, e.g., $\beta_i^{\pi l}$ is the estimated impact on conflict incidence of—in all districts within a $(\pi - 1, \pi]$ one-hundred kilometers radius—exchanging one day with temperature levels in the omitted bin to a day in bin i in l year-months past to the current year-month.

Figure D.4a and D.4b show the baseline estimates when including average climate exposure within 0 to 300 kilometers in temperature and precipitation. The baseline results are by and large robust to this alternative specification. Figure D.5 and D.6 plot the spatially lagged temperature and precipitation response functions in the current and prior year-month. I find that spatial spillovers are insignificant.[55]

[55] I note that results are similar if one drops the province-year-month fixed effects except that contemporaneous average temperature exposure within a $(0, 100]$ km radius from a district d significantly affect current conflict incidence in district d (not shown). This finding provides an additional argument for the inclusion of province-year-month fixed effects in the baseline specification (1).



A. CURRENT YEAR-MONTH

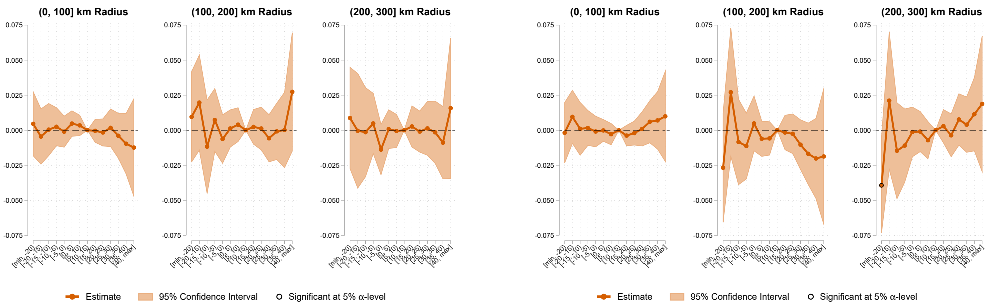


B. PRIOR YEAR-MONTH

FIGURE D.4. CONFLICT INCIDENCE TEMPERATURE AND PRECIPITATION RESPONSE FUNCTIONS—CONTROLLING FOR REMOTE TEMPERATURE AND PRECIPITATION VARIATION

Notes: The left and right figure of Panel A (B) plots estimates and 95 percent confidence bands of the contemporaneous (one year-month lagged) conflict incidence temperature and precipitation response functions obtained by estimating (D.2). The omitted temperature- and precipitation-day bins are $[10, 15]$ °C and $[0.8, 1)$ mm. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Standard errors two-way clustered at the district- and year-month-level.

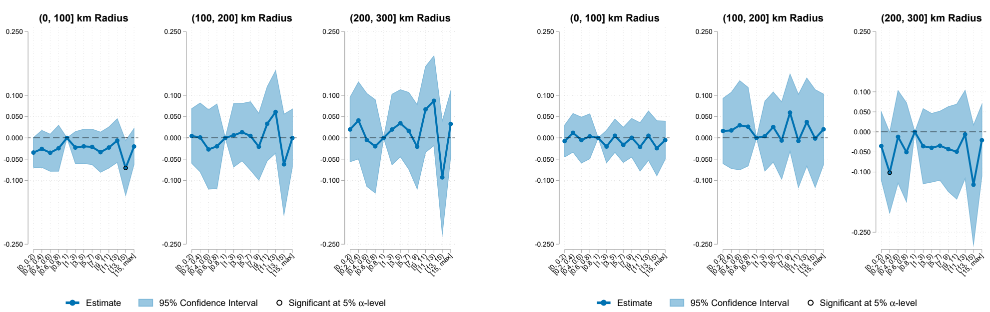
Grid Cell Fixed Effects.—Provinces and their boundaries are politically defined. These may therefore not suitably capture differences in agro-climatic zones that better describe the regional effects of climate change. To try to account for this, I distribute the 398 districts of Afghanistan into 30 2.5 by 2 lon-



A. CURRENT YEAR-MONTH B. PRIOR YEAR-MONTH

FIGURE D.5. CONFLICT INCIDENCE SPATIAL TEMPERATURE LAGS

Notes: These figures plots estimates and 95 percent confidence intervals of $\{\beta_i^{\pi l}\}_i$ for all $\pi \in \{1,2,3\}$ and $l \in \{0,1\}$ obtained by estimating (D.2). The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.



A. CURRENT YEAR-MONTH B. PRIOR YEAR-MONTH

FIGURE D.6. CONFLICT INCIDENCE SPATIAL PRECIPITATION LAGS

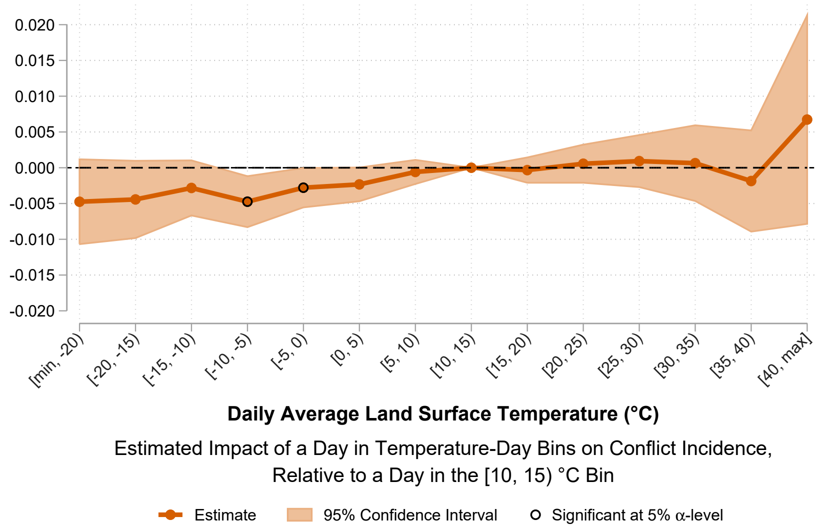
Notes: These figures plots estimates and 95 percent confidence intervals of $\{\gamma_i^{\pi l}\}_i$ for all $\pi \in \{1,2,3\}$ and $l \in \{0,1\}$ obtained by estimating (D.2). The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

gitude by latitude grid cells. Then I replace π_{pm} with π_{gm}, where g denotes a 2.5 by 2 longitude by latitude grid cell in (1). These cells are disjoint and cover all districts of Afghanistan. I find that precipitation still plays no role and that fewer coefficients are significant at the 5 percent α-level for the temperature conflict incidence relationship, although some are still significant at the 10 percent α-level. However, the caveat with the grid cell fixed effects is that since I assign district centroids to grid cells, there are grid cells that cover districts but not their centroids. This problem does not occur for the province-year-month fixed effects as provinces are by construction polygons that cover each of its districts perfectly. I, therefore, prefer results based on the baseline set of fixed effects.

Controlling for District-Year Fixed Effects.—Climate change adaptation may be the result of observed or expected extreme weather events. If there is inter-annual district-level climate change adaptation (e.g., improved water resource management) that affect opportunity costs to conflict and peace, my baseline estimators may be subject to omitted variable bias. Though this seems unlikely to be the case as Afghanistan only recently began to prepare rural communities for climate change (NEPA, UNEP and WFP, 2016), I here control for this potential omission by adding district-year fixed effects to my baseline specification. Figure D.8 reproduce the baseline results under this alternative specification. I find that the baseline results are by and large robust to the inclusion of district-year fixed effects.[56]

Conflict Intensity.—My focus on conflict incidence reflects my interest in ex-

[56]Note that the fact that the estimates in Figure D.8 are closer to zero compared with the baseline estimates may be a result of exacerbated attenuation bias if there were no district-year-level omitted variables in the baseline specification.



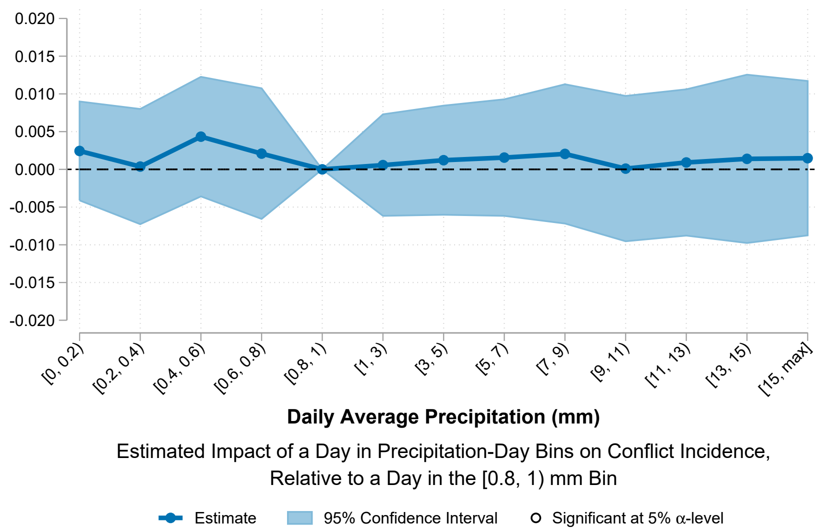
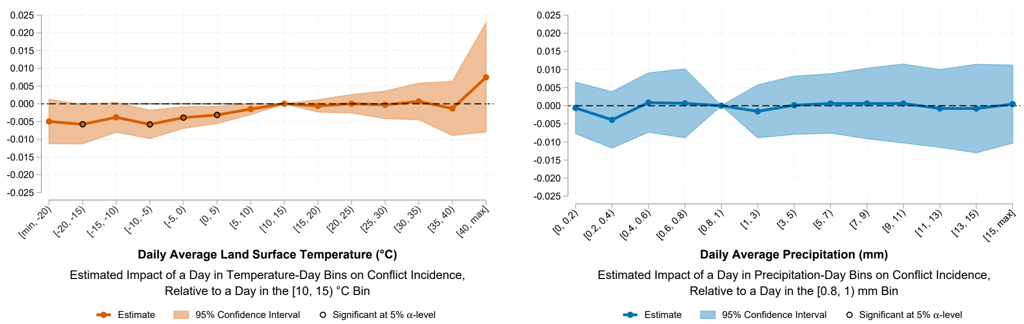
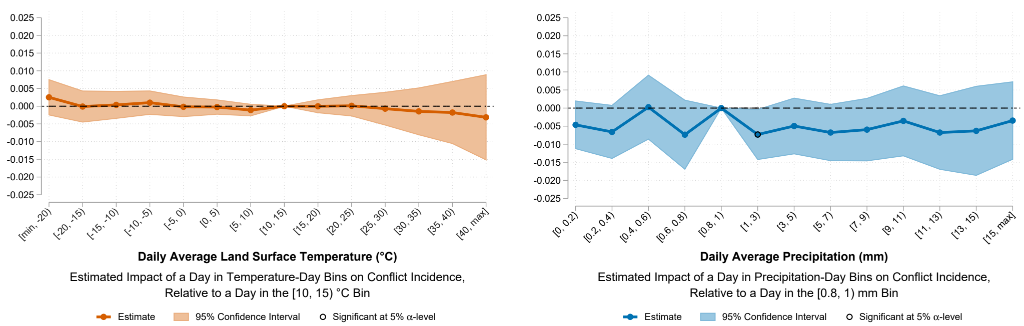


FIGURE D.7. CONFLICT INCIDENCE RESPONSE FUNCTIONS—ALTERNATIVE REGIONAL FIXED EFFECTS

Notes: These figures show temperature and precipitation response functions in the current year-month obtained by estimating (1) after replacing province-year-month fixed effects with year-month-varying 2.5 by 2 longitude-latitude grid cell fixed effects. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.



A. CURRENT YEAR-MONTH



B. PRIOR YEAR-MONTH

FIGURE D.8. CONFLICT INCIDENCE RESPONSE FUNCTIONS—CONTROLLING FOR DISTRICT-YEAR FIXED EFFECTS

Notes: The left and right figure of Panel A (B) plots estimates and 95 percent confidence bands of the contemporaneous (one year-month lagged) conflict incidence temperature and precipitation response functions obtained by estimating (1) after adding district-year fixed effects. The omitted temperature- and precipitation-day bins are $[10, 15]$ °C and $[0.8, 1)$ mm. The number of observations is 54,526 from a balanced panel of 398 districts and 137 year-months. Mean conflict incidence is about 15 percent. Standard errors two-way clustered at the district- and year-month-level.

plaining the general presence of conflicts. An alternative definition of violence is conflict intensity measured as the number of battle-related deaths. If a significant relationship between weather and conflict incidence but not between weather and conflict intensity, or vice verse, is found, this does not imply that one or the other model does not identify the parameters of interest. However, if results stand in stark contrast to each other, this may cast doubt on my baseline result. Furthermore, the relationship between weather and conflict intensity is in itself interesting.

To estimate the relationship between weather and conflict intensity I fit the number of battle-related deaths to a Poisson regression model. Specifically, for the number of battle-related deaths BRD_{dt} in district d year-month t I impose the following probability density function:

$$\mathbb{P}(BRD_{dt} = n|\boldsymbol{x}_{dt}) = \exp\left(\mu(-\boldsymbol{x}_{dt})\right)\frac{\mu(\boldsymbol{x}_{dt})^n}{n!}, \quad n = 0, 1, \ldots, \tag{D.3}$$

where the link function $\mu(\boldsymbol{x}_{dt}) \equiv \mathbb{E}(BRD_{dt}|\boldsymbol{x}_{dt})$ provides a parametric form for the conditional mean of conflict intensity given all covariates $\boldsymbol{x}_{dt}$:

$$\mu(\boldsymbol{x}_{dt}) = \exp\left(\sum_{l=0}^{1}\Big(\sum_i \beta_i^l T_{d,t-l}^i + \sum_j \gamma_j^l P_{d,t-l}^j\Big) + \delta_{dm} + \pi_{pt}\right). \tag{D.4}$$

For this model, $100\beta_i^l$ $(100\gamma_j^l)$ approximate the percentage change in the conditional mean number of battle-related deaths from exchanging a day in the omitted temperature-day (precipitation-day) bin to a day with temperature (precipitation) levels in bin i (j).

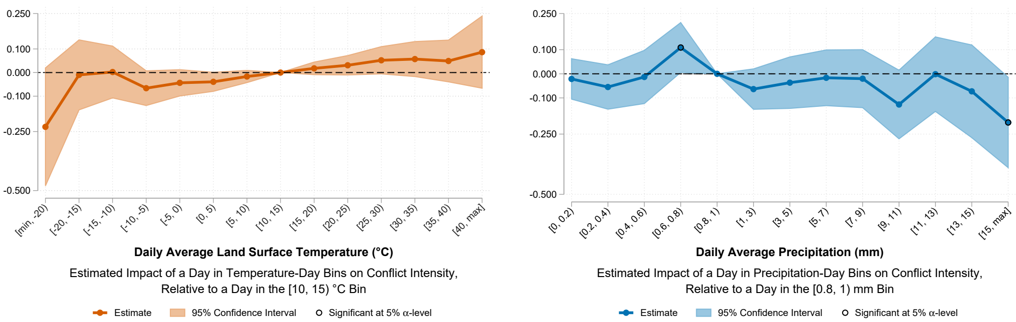
Estimating a Poisson regression model is more suitable than a log-linear model of conflict intensity for three reasons. First, it accounts for the fact that the number of battle-related deaths is a count variable and easily handles observed values of conflict intensity equal to zero.[57] Second, even if the Poisson distributional assumption does not correctly describe the empirical distribution of the number of battle-related deaths, maximum likelihood estimation produces unbiased estimates of the coefficients if (D.4) correctly describes the conditional mean of conflict intensity (Wooldridge, 1997, 1999). Third, the Poisson regression model does not suffer from the incidental parameters problem (Cameron and Trivedi, 2005).

There are however three primary drawbacks to my Poisson regression model. First, I have to omit district-month-specific yearly trends $q_{dm}(y)$. Second, I can only account for serial correlation in the error terms by clustering at the district-level, thereby ignoring spatial correlation. Third, estimation of Poisson fixed effects models tend to lead to some loss of data as observations who do not vary within a group specified by a fixed effect (e.g., a district-month) are dropped as these do not contribute to maximizing the likelihood (Cameron and Trivedi, 2005). In any case, I believe that my Poisson model is the best among sub-optimal solutions to study conflict intensity and treat results from my Poisson model as highly indicative only.

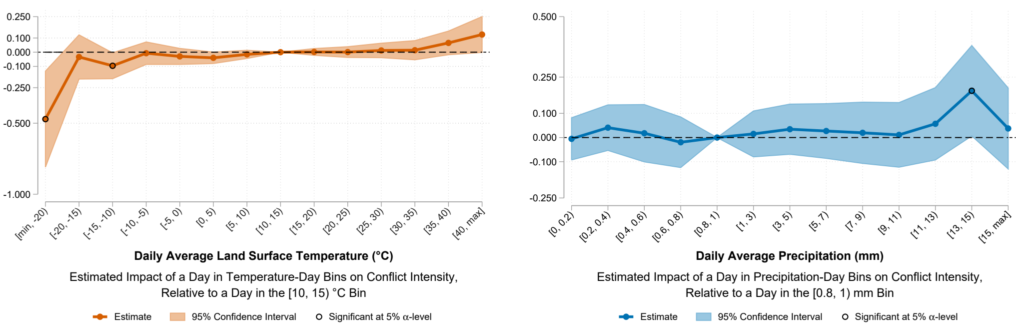
Figure D.9 show response functions for conflict intensity obtained by estimating (D.3). The estimation method drops observations for which the number of battle-related deaths is constantly zero within a fixed effects group. Conse-

[57] This, combined with the fact that the number of observations with zero values on the number of battle-related deaths is high, explains why I do not estimate a log-linear model as a robustness check. Some authors (e.g., Dube and Vargas (2013)) handle such a problem by replacing the variable with the log of the variable plus a small number ε (e.g., 0.001) to account for zero values. However, results from replacing BRD_{dt} with $\log(BRD_{dt} + \varepsilon)$ are highly sensitive to the choice of ε in the case of a high number of observations for which BRD_{dt} is zero. Indeed, I find that estimates can mechanically change and switch sign when experimentally varying ε over $(0.001, 0.0001)$.

quently, a loss of about 55 percent observations occurs. This substantial loss of data makes these results highly suggestive. Keeping this in mind and evaluating the results as they are I find that the role of precipitation and temperature in driving the number of battle-related deaths is barely significant. Also, three coefficients on contemporaneous and lagged precipitation seem spuriously significant as the pattern seem inexplicable. Nonetheless, the model prediction for contemporaneous temperature is consistent with the baseline results as higher contemporaneous temperature levels tend to increase the number of battle-related deaths. The predictions are also quantitatively meaningful, with, e.g., a decrease by about 3.9 percent (s.e. about 2.1 percent) in the mean number of battle-related deaths from contemporaneously exchanging a day with temperature levels between 10 and 15 °C to a day with temperature levels between 0 and 5 °C.



A. CURRENT YEAR-MONTH



B. PRIOR YEAR-MONTH

FIGURE D.9. CONFLICT INTENSITY RESPONSE FUNCTIONS

Notes: The left and right figure of Panel A (B) plots estimates and 95 percent confidence bands of the contemporaneous (one year-month lagged) conflict intensity temperature and precipitation response function obtained by estimating (D.3) with link function (D.4). The omitted temperature- and precipitation-day bins are $[10, 15]$ °C and $[0.8, 1)$ mm. The number of observations is 24,530 from an unbalanced panel of 350 districts and 137 year-months. Mean number of battle related deaths is about 3.14. Standard errors are clustered at the district-level.

E Estimation Method

I estimate my Poisson regression model with the Stata package `poi2hdfe` by Paulo Guimarães based on a procedure outlined in Figueiredo, Guimarães and Woodward (2015). The remaining multidimensional fixed effects regression models are estimated with a computationally efficient iterative and graph-theoretic estimation method developed by Correia (2016). The Stata package `reghdfe` that implement the estimation method by Correia (2016) allow for, e.g., two-way clustering and interaction of fixed effects with continuous variables. More notable is that the `reghdfe` package automatically drops singletons, i.e., fixed effects groups with only one observation. In this thesis, I only drop singletons for the conflict onset and ending model in Section 5.2. The reason is that maintaining singletons can overstate statistical significance (Correia, 2015). Though this may seem to be a non-optimal solution as we lose some information, there is currently no general solution on how to treat singletons.

F References of Appendix

Arino, Olivier, Jose Julio R. P., Vasileios Kalogirou, Sophie Bontemps, Pierre Defourny, and Eric van Bogaert. 2012. "Global Land Cover Map for 2009 (GlobCover 2009)." *European Space Agency and Université catholique de Louvain*.

Baysan, Ceren, Marshall Burke, Felipe Gonzalez, Solomon M. Hsiang, and Edward Miguel. 2015. "Economic and Non-Economic Factors in Violence: Evidence from Organized Crime, Suicides and Climate in Mexico." *Unpublished manuscript, University of California, Berkeley, California, USA*.

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45.

Conley, Timothy G. 2010. "Spatial Econometrics." In *Microeconometrics*. 303–313. London: Palgrave Macmillan.

Correia, Sergio. 2015. "Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix." *Unpublished manuscript, Duke University, North Carolina, USA*.

Correia, Sergio. 2016. "A Feasible Estimator for Linear Models with Multi-Way Fixed Effects." *Unpublished manuscript, Duke University, North Carolina, USA*.

Dell, Melissa. 2009. "GIS Analysis for Applied Economists." *Unpublished manuscript, Massachusetts Institute of Technology, Massachusetts, USA*.

Dube, Oeindrila, and Juan F. Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *The Review of Economic Studies*, 80(4): 1384–1421.

Fetzer, Thiemo. 2014. "Social Insurance and Conflict: Evidence from India." *Unpublished manuscript*.

Figueiredo, Octávio, Paulo Guimarães, and Douglas Woodward. 2015. "Industry Localization, Distance Decay, and Knowledge Spillovers: Following the Patent Paper Trail." *Journal of Urban Economics*, 89: 21–31.

Hijmans, Robert J., Luigi Guarino, Mariana Cruz, and Edwin Rojas. 2001. "Computer Tools for Spatial Analysis of Plant Genetic Resources Data: 1. DIVA-GIS." *Plant Genetic Resources Newsletter*, 127: 15–19.

Hsiang, Solomon M. 2016. "Climate Econometrics." *Annual Review of Resource Economics*, 8: 43–75.

Hsiang, Solomon M., Kyle C. Meng, and Mark A. Cane. 2011. "Civil Conflicts Are Associated with the Global Climate." *Nature*, 476(7361): 438.

Kienberger, Stefan, Raphael Spiekermann, Dirk Tiede, Irmgard Zeiler, and Coen Bussink. 2017. "Spatial Risk Assessment of Opium Poppy Cultivation in Afghanistan: Integrating Environmental and Socio-Economic Drivers." *International Journal of Digital Earth*, 10(7): 719–736.

United Nations Office on Drugs and Crime (UNODC). 2008. "Afghanistan Opium Winter Rapid Assessment Survey 2008." Vienna: United Nations, Office on Drugs and Crime.

United Nations Office on Drugs and Crime (UNODC). 2013. "Afghanistan Opium Risk Assessment 2013." Vienna: United Nations, Office on Drugs and Crime.

Wooldridge, Jeffrey M. 1997. "Quasi-Likelihood Methods for Count Data." *Handbook of Applied Econometrics Volume 2: Microeconomics*, 321–368.

Wooldridge, Jeffrey M. 1999. "Distribution-Free Estimation of Some Nonlinear Panel Data Models." *Journal of Econometrics*, 90(1): 77–97.