

Reexamining the Effect of Refugees on Civil Conflict: A Global Subnational Analysis*

Yang-Yang Zhou[†] and Andrew Shaver[‡]

May 12, 2021

Supplementary Information

Contents

I	Supplementary Information	1
S1	Summary Statistics	2
S2	Data Sources	4
S3	Analysis of Other Conflict Types	5
S4	Analysis using Panel Match	6
S5	Additional Heterogeneous and Subgroup Effects	7
	S5.1 Heterogeneous Effects of Large versus Small Refugee Populations	7
	S5.2 Heterogeneous Effects of Formal Camps versus Informal Settlements	9
	S5.3 Heterogeneous Effects of Provinces Near vs. Far from Border	12
	S5.4 Subgroup Effects by Region	15
S6	Robustness Checks	16
	S6.1 Drop One Country at a Time	16
	S6.2 Drop One Region at a Time	18
	S6.3 Analysis using Full Data of All Countries	19
	S6.4 Using the Dynamic Subset Data of only Refugee-hosting Countries	19
	S6.5 Leading the Dependent Variables by 5 Years	20
	S6.6 Analysis using wzoneData conflict outcomes	21
S7	Expert Interviews with UNHCR Officials	23
	S7.1 General null effects of refugee-hosting on conflict	23
	S7.2 On militarized refugees and prevention	23
	S7.3 Refugee-host relations	23
	S7.4 Effects of refugee-hosting on development, infrastructure, other services	24
	S7.5 Economic effects of hosting refugees	25
	S7.6 Effects of refugee-hosting on local state capacity	25
	S7.7 Differences between refugee camps and settlements	26
	S7.8 How contemporary refugee-hosting dynamics differ from the Cold War era	26
	S7.9 Suggestions for better policy	26
S8	Regression Tables	27
	S8.1 H2: Concentrated or Dispersed Refugee Presence on Conflict	27
	S8.2 H3: Conditional Risk Reduction through Increased Infrastructure	28
	S8.3 H1 Models Replacing Refugee Presence (binary) Variable with Refugee Population and Ratio of Refugee to Local Population	29
S9	Salehyan and Gleditsch (2006) Extension Exercise	30

*Replication files, including R code and data, are available at the APSR Dataverse (<https://doi.org/10.7910/DVN/1CNMHF>)

[†]Assistant Professor, Department of Political Science, University of British Columbia, yangyang.zhou@ubc.ca, www.yangyangzhou.com

[‡]Assistant Professor, Department of Political Science, University of California, Merced. Email: ashaver@ucmerced.edu, www.andrewshaver.com

S1 Summary Statistics

This section shows descriptive statistics for the variables we use in the analyses.

Table S1: Summary Statistics for Main Data of Only Countries that have Hosted Refugees

Statistic	N	Mean	St. Dev.	Min	Median	Max
Conflict Onset (until 2008)	14,960	0.08	0.30	0.00	0.00	1.00
Conflict Incidence (until 2008)	18,981	0.30	0.50	0.00	0.00	1.00
Violent Events	28,659	4.00	28.00	0.00	0.00	1,986.00
Violent Events State Based	28,659	3.00	22.00	0.00	0.00	1,367.00
Violent Events Non-State	28,659	0.50	8.00	0.00	0.00	959.00
Violent Events One Sided	28,659	0.80	6.00	0.00	0.00	356.00
Battle Deaths	28,659	64.00	2,051.00	0.00	0.00	324,269.00
Battle Deaths State Based	28,659	32.00	447.00	0.00	0.00	49,208.00
Battle Deaths Non-State	28,659	6.00	82.00	0.00	0.00	6,785.00
Battle Deaths One Sided	28,659	26.00	1,994.00	0.00	0.00	324,243.00
Nighttime Lights Mean	20,118	6.00	6.00	0.00	4.00	77.00
Refugee Presence	28,971	0.09	0.30	0	0	1
Other Province Refugee Presence	28,971	0.60	0.50	0	1	1
Refugee Camps	28,971	0.06	0.20	0	0	1
Other Province Refugee Camps	28,971	0.50	0.50	0	1	1
Refugee Settlements	28,971	0.04	0.20	0	0	1
Other Province Refugee Settlements	28,971	0.20	0.40	0	0	1
Neighboring Violent Events	28,971	12.00	63.00	0	0	3,472
Lagged Population (logged)	26,824	13.00	1.00	8.00	13.00	19.00
Lagged GDP	26,824	6.00	18.00	0.00	1.00	388.00
Excluded Ethnic Groups Mean	24,636	0.40	0.50	0.00	0.10	4.00
Terrain Ruggedness	28,971	98.00	84.00	0.20	76.00	460.00
Province size (sqkm)	28,971	38,524.00	72,647.00	11.00	14,807.00	818,013.00
Distance from border (km logged)	28,971	5.00	1.00	1.00	5.00	7.00
Distance from capital (km logged)	28,971	6.00	0.90	2.00	6.00	8.00
IDP Presence	28,971	0.01	0.09	0	0	1

Table S2 shows the number of province-year observations by number of refugee sites in the data (e.g. 26, 237 province-years have no refugee sites, 1421 province-years have 1 refugee site, etc.). Next, Table S3 breaks down the number of refugee camps, settlements, and sites (both camps and settlements) by region.

Refugee Sites	Province-Years	Refugee Sites	Province-Years	Refugee Sites	Province-Years
0	26237	13	4	38	1
1	1421	14	12	39	16
2	462	15	5	41	1
3	223	16	6	53	1
4	131	17	23	55	1
5	122	20	1	59	2
6	68	21	2	60	6
7	45	22	14	65	6
8	29	23	1	66	1
9	24	25	3	67	7
10	18	33	1	69	2
11	23	34	2	79	1
12	19	35	1	103	29

Table S2: Frequency of Province-Years by Number of Refugee Sites

Figure S1 shows the main conflict outcomes across provinces over time. Because of the spike in battle deaths in 1994 due to the Rwandan genocide, we log this outcome in our analyses.

	Refugee Camps	Refugee Settlements	Refugee Sites
Central Asia	2	65	67
Eastern Africa	108	143	251
Middle Africa	84	67	151
Northern Africa	65	11	76
South America	3	0	3
Southeastern Asia	211	0	211
Southern Africa	1	1	2
Southern Asia	91	114	205
Southern Europe	6	9	15
Western Africa	243	29	272
Western Asia	73	6	79
Central America	0	104	104
Eastern Europe	0	2	2
Melanesia	0	22	22

Table S3: Number of Refugee Camps, Settlements, and Sites (Camps + Settlements) per Region

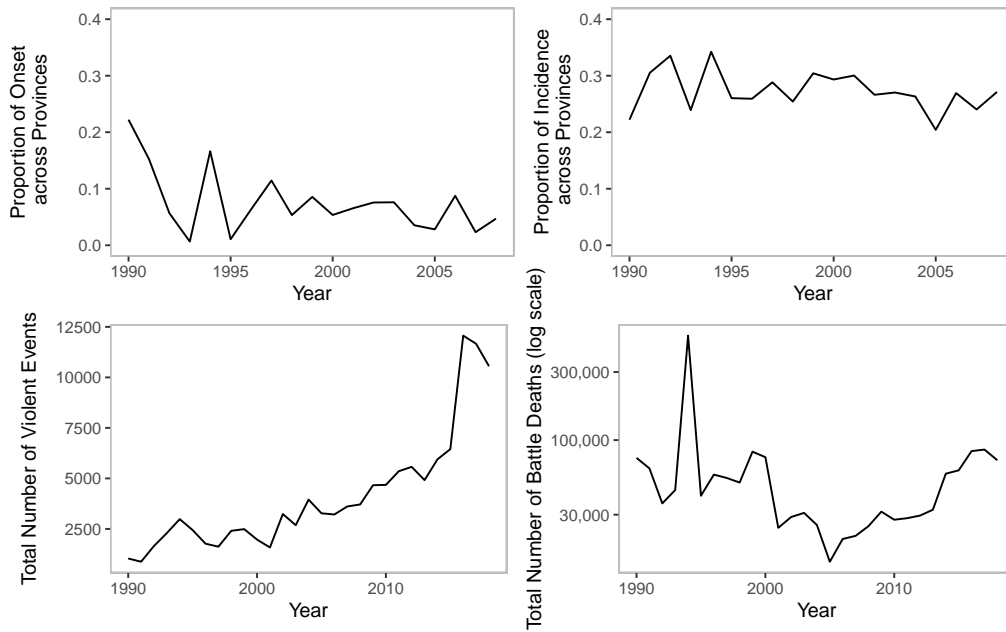


Figure S1: This figure shows the propensity of conflict onset and incidence (1990–2008) across time in plots 1 and 2; and the number of violent events and battle deaths (1990–2018), across time in plots 3 and 4, using the main data.

S2 Data Sources

The UNHCR georeferenced refugee locations dataset that we introduce in this paper can be accessed in our replication files. Specifically, the raw file from the UNHCR is `wr1_pp1_poc_07_01_2020.xls`, and the data that we additionally processed (see description in the manuscript) for use in the main analysis is `camps_settlements_processed.csv`.

Although we include our data cleaning code for transparency, we do not include within our replication files the following data that we use as inputs into the cleaning code, because they are either proprietary or already available publicly online. We detail how to find these data below.

UCDP GED

The Uppsala Conflict Data Program’s (UCDP) Georeferenced Event Dataset (GED) (version 19.1) Sundberg and Melander (2013) can be accessed at: ucdp.uu.se/downloads/olddw.html

UCDP’s GED separate Syria dataset (version 652.1601.1911) can be accessed at: ucdp.uu.se/downloads/olddw.html

PRIO Conflict Sites

The Peace Research Institute Oslo (PRIO) Conflict Site (1989-2008) data (version 3.0) Dittrich Hallberg (2012) can be accessed at: prio.org/Data/Armed-Conflict/Conflict-Site/

wzoneData

The wzoneData: Zones of Armed Conflicts data (corresponding to UCDP GED version 19.1) developed by Kikuta (2020) are available at: dataverse.harvard.edu/dataverse/kyosukekkt

CShapes Shapefile

The CShapes shapefile (version 0.6) (Weidmann, Kuse and Gleditsch, 2010) can be accessed at: nils.weidmann.ws/projects/cshapes/shapefile.html

Administrative Unit Shapefile

The 1998 1st level administrative (“province”) boundaries used in this project can be procured from ESRI: www.esri.com

Universal Transverse Mercator Shapefile

The global Universal Transverse Mercator grid shapefile (updated as of 11/11/2018) can be access at: hub.arcgis.com/datasets/utm-grid

PRIO-GRID

PRIO-GRID (Tollefsen, Strand and Buhaug, 2012) data (version 2.0) can be accessed at: grid.prio.org/download
The PRIO-GRID polygon (cell) shapefile can be accessed at: grid.prio.org/extensions
The PRIO-GRID centroids shapefile no longer appears on PRIO-GRID website. Individuals interested in obtaining these data might reach out to the following contacts listed on the PRIO-GRID website: andreas@prio.org; marteg@prio.org.

Terrain Ruggedness

Data on terrain ruggedness (original version) from Shaver, Carter and Shawa (2016) can be accessed at: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/WXUZBN>

S3 Analysis of Other Conflict Types

This section shows the main (H1) and secondary (H2) analyses on other conflict types included in the UCDP GED data: state-based, non-state, and one-sided violent events and battle deaths. Consistent with the analyses in the paper, Figure S2 shows that there is no effect of refugee presence on the various types of conflict, and Figure S3 shows that there are negative effects of concentrated refugee presence on several of the conflict types – state-based and one-sided violent events and state-based battle deaths and null effects for dispersed refugee presence. This is also consistent with our theorized predictions.

Note that two placebo test coefficients are positive, implying that provinces hosting refugee sites in the future may be experiencing more conflict in the present, which goes against our selection concerns that refugees are able to select into provinces with less conflict.

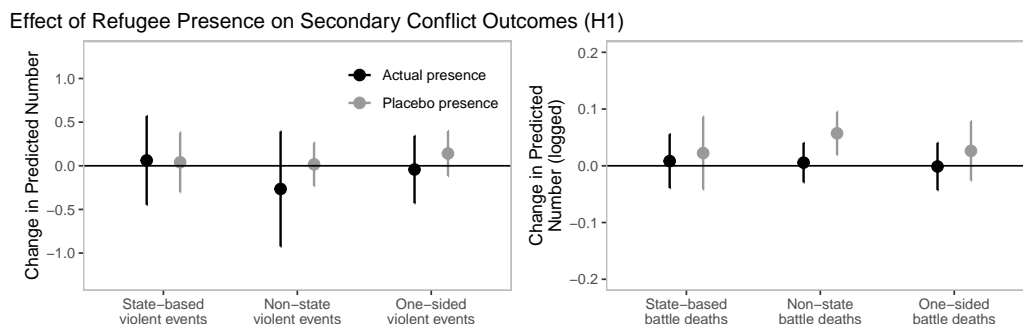


Figure S2: This figure shows the effect of refugee presence on state, non-state, and one-sided conflict outcomes (black), compared to their respective placebo estimates, i.e. effect of future refugee presence (gray). All point estimates include 95% CIs.

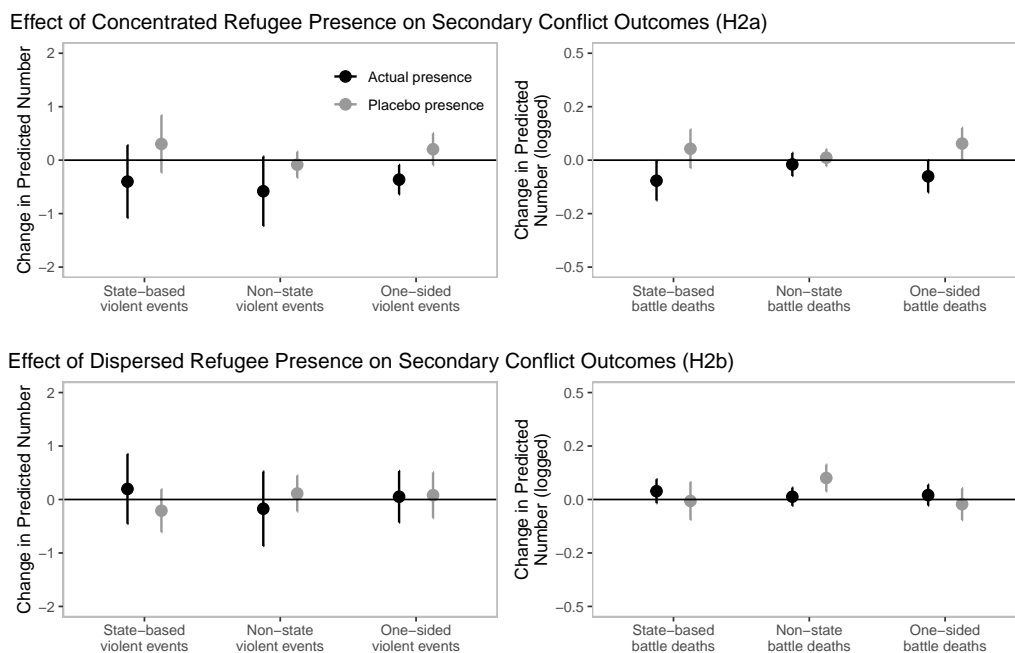


Figure S3: This figure shows the effect of refugee presence on state, non-state, and one-sided conflict outcomes (black) compared to their respective placebo estimates, i.e. effect of future refugee presence (gray), conditional on refugee presence in other provinces of the same country-year. All point estimates include 95% CIs.

S4 Analysis using Panel Match

We use a matching method for causal identification with time-series cross-section data by Imai, Kim and Wang (2018a). In this method, each treated observation is matched with a set of control observations that share identical treatment history up to the pre-specified number of periods. Next, this set of matched controls is further refined by adjusting for covariates. Finally, we simply calculate the difference-in-differences estimator in order to account for an underlying time trend. To do this, we use the R package `PanelMatch` (Imai, Kim and Wang, 2018b). Specifically, we create matched sets with 3 years (default) of treatment history, e.g. comparing “treated provinces” with refugee presence histories of $[0\ 0\ 1]$ observations to all “control provinces” of $[0\ 0\ 0]$ observations, without refinement. Results are substantively similar when refining on mahalanobis distance or CBPS propensity score. Similar to the results in the paper, these results using matching analyses show a null effect of refugee presence on conflict outcomes, and whether conditional on dispersed/concentrated or not.¹

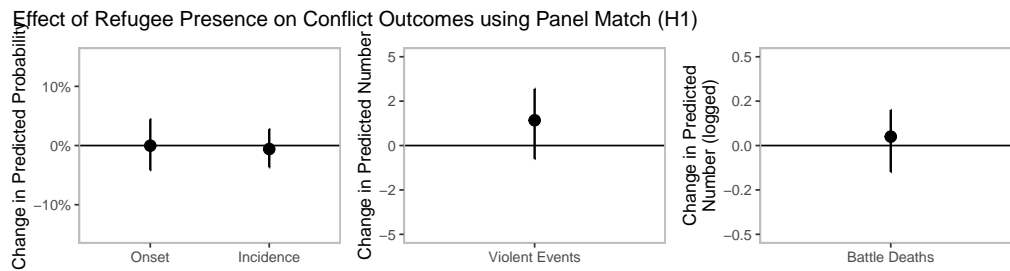


Figure S4: This figure shows the effect of refugee presence on conflict outcomes using PanelMatch. It corresponds to Figure 2 in the paper. All point estimates include 95% CIs.

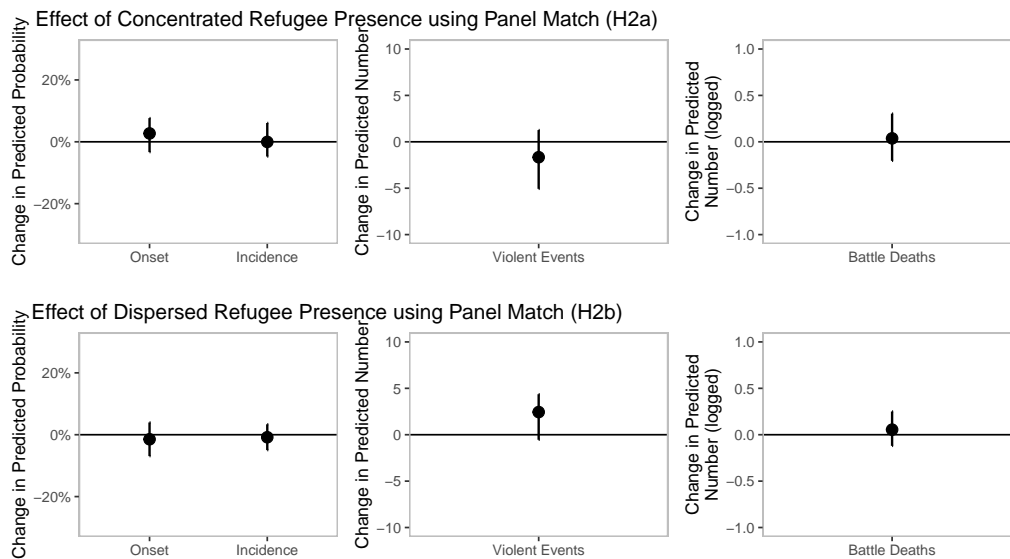


Figure S5: This figure shows the effect of refugee presence on conflict outcomes conditional on refugee presence in other provinces of the same country-year, using PanelMatch. It corresponds to Figure 3 in the paper. All point estimates include 95% CIs.

¹Note, these results are noisier than those in the main paper, since `PanelMatch` only analyzes those treated and control observations for which comparison units with identical treatment histories can be found. For example, in the first onset analysis, only 101 treated unit-year pairs have matches that can be used to estimate the ATT, and only 55 control unit-year pairs have matches that can be used to calculate the ATC. Refinement using pre-treatment covariates only reduces the size of these matched sets even further. Therefore, we urge caution when interpreting these results, due to the small subset of the data used to estimate causal effects.

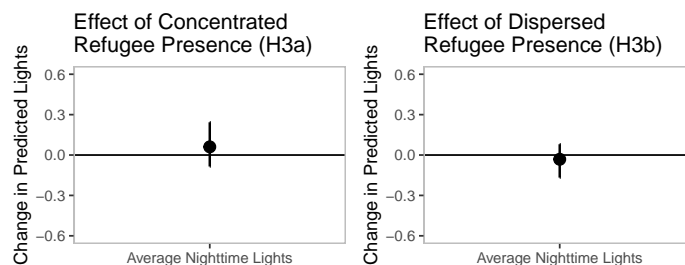


Figure S6: This figure shows the effect of refugee presence average nighttime lights conditional on refugee presence in other provinces of the same country-year, using PanelMatch. It corresponds to Figure 4 in the paper. All point estimates include 95% CIs.

S5 Additional Heterogeneous and Subgroup Effects

S5.1 Heterogeneous Effects of Large versus Small Refugee Populations

As described in the Data section of our paper, for a subset of our data, we have information from the UNHCR on refugee population, specifically the total number of refugees per site, which we aggregate to the province-year level. We only have this data for African countries from 2010 to 2015. Thus, for the heterogeneous effects analysis based on population size, we can only examine the violent events and logged battle death outcomes as they span this time period. In this section, we create a binary variable for heterogeneous effects analysis based on the size of refugee presence in a province-year in order to test H4. We designate refugee-hosting provinces that have a 5% ratio or greater of refugee population compared to local population as “large” and hosting provinces under 5% as “small.”

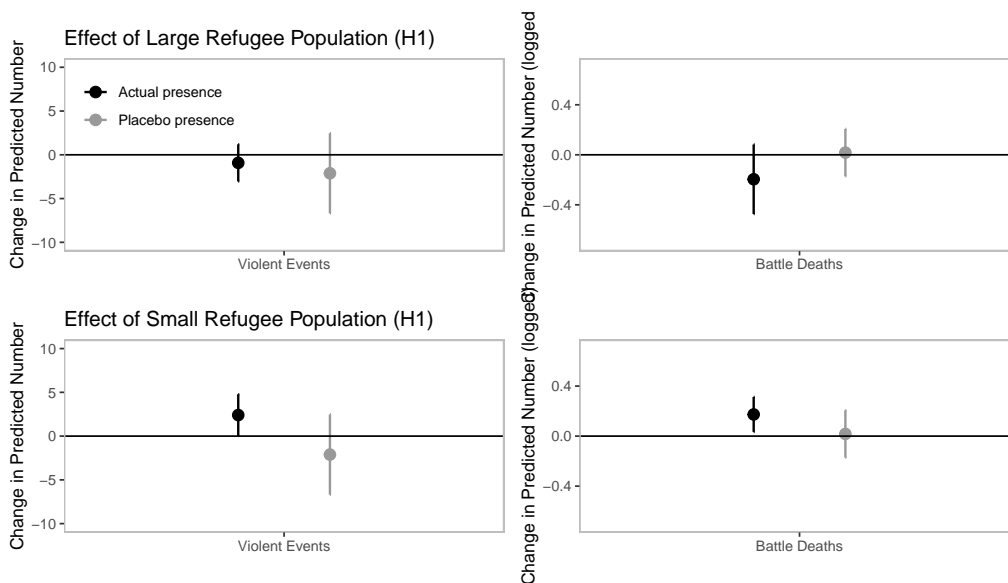


Figure S7: This figure shows the effect of refugee presence in provinces hosting relatively large versus small refugee populations (heterogeneous effects analysis) on violent events and battle deaths 2010–2015 in Africa (black), compared to their respective placebo estimates (gray). All point estimates include 95% CIs.

Figure S7 shows the heterogeneous effects for provinces hosting large relative populations (top plot) versus small (bottom plot), respectively, on violent events and battle deaths (onset and incidence fall outside the time period of this subset data). There is no effect of large relative populations on conflict. However, there is a positive effect of small populations on conflict. This may be due to small populations not bringing developmental benefits, and being easier targets for victimization. Future research should further interrogate this finding.

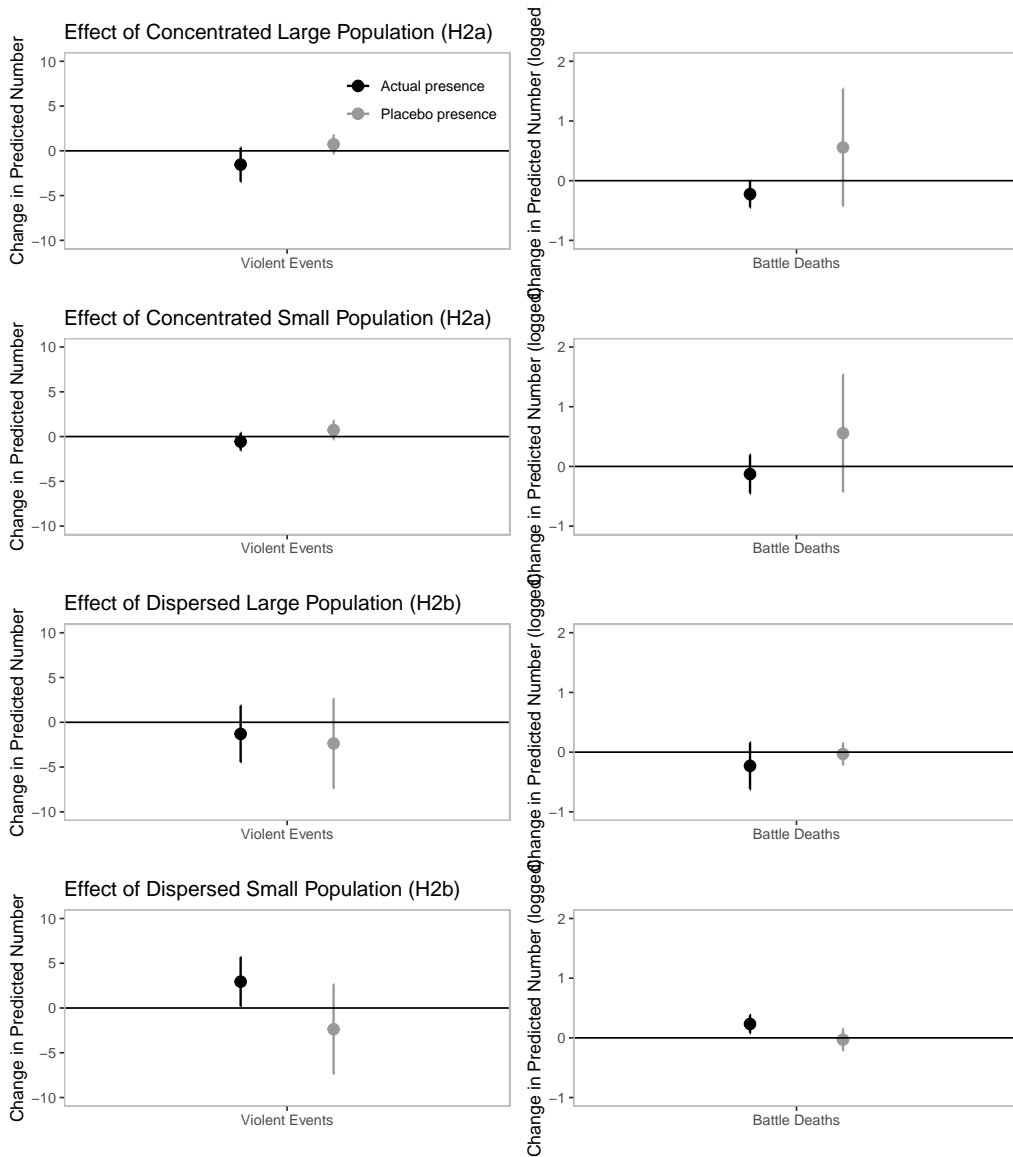


Figure S8: This figure shows the effect of refugee presence in provinces hosting relatively large versus small refugee populations (heterogeneous effects analysis) on violent events and battle deaths 2010–2015 in Africa (black) compared to their respective placebo estimates (gray), conditional on refugee presence in other provinces of the same country-year. All point estimates include 95% CIs.

Figure S8 shows the heterogeneous effects for provinces hosting large versus small relative populations respectively, conditional on refugee presence in other provinces of the same country-year. The top two plots compare large and small refugee presence when it is geographically concentrated, while the bottom two plots do so in geographically dispersed cases. It appears that the conditional risk reduction effects in the cases of large concentrated refugee presence. However, there is a positive effect of small dispersed presence on conflict.

Figure S16 shows the heterogeneous effects on nighttime lights for provinces hosting large versus small relative populations respectively, conditional on refugee presence in other provinces of the same country-year. The top two plots compare large and small refugee presence when it is geographically concentrated, while the bottom two plots do so in geographically dispersed cases. We cannot discern differential effects by relative population size; note the placebo estimates for concentrated cases do not perform well.

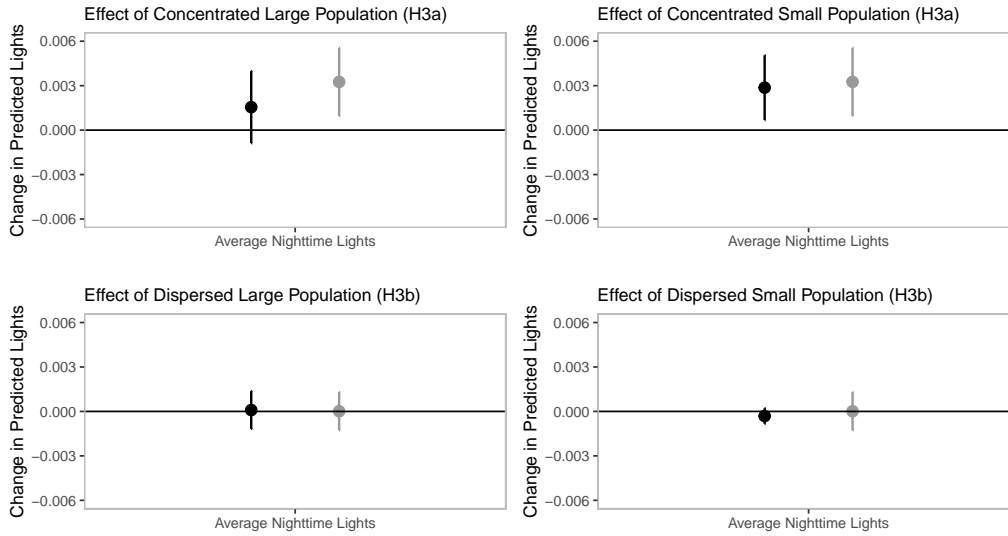


Figure S9: This figure shows the effect of refugee presence in provinces hosting relatively large versus small refugee populations (heterogeneous effects analysis) on nighttime lights 2010–2015 in Africa (black) compared to their respective placebo estimates (gray), conditional on refugee presence in other provinces of the same country-year. All point estimates include 95% CIs.

S5.2 Heterogeneous Effects of Formal Camps versus Informal Settlements

We assess whether refugee camps or settlements have different effects on conflict onset or incidence. According to our discussions with UNHCR officials, whether a refugee site is designated as a “camp” or a “settlement” is largely a political decision by the host state to recognize it formally, not a decision by the UNHCR. While camps are generally associated with more state control and international recognition and settlements are informal areas of refugee communities, there is no clear delineation between the two based on realities on the ground such as population sizes or levels of aid.² In fact, UNHCR operational protocols refer to camps and settlements together without explicitly differentiating the two.³ Nevertheless, since host states choose to recognize “settlements” as informal sites, they might understand that these areas are less contentious with the local population than sites that they call “camps.” Thus, we hypothesize that provinces with refugee settlements might experience less conflict onset and incidents than provinces with camps.

Figure S10 descriptively shows the number of provinces with only formal refugee camps (dotted line), informal refugee settlements (dashed line), and both (solid line) over time. Throughout the data of province-years, of all 2734 provinces with refugee sites, 57.9% are provinces with only refugee camps, 33.5% are provinces with only refugee settlements, and only 8.6% are provinces with both.

²Conducted with the Field Information and Coordination Support Section (FICSS) of UNHCR on December 30, 2013 by teleconference.

³www.unhcr.org/uk/448d6c122.pdf

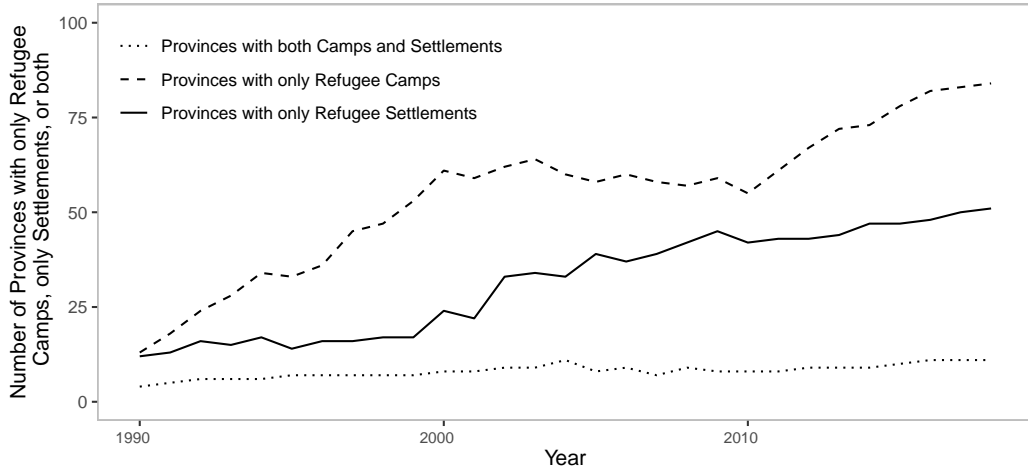


Figure S10: This figure shows the number of provinces with only formal refugee camps (dotted line), informal refugee settlements (dashed line), and both (solid line) over time. Throughout the main data of province-years, of all 2734 provinces with refugee sites, 57.94% are provinces with only refugee camps, 33.5% are provinces with only refugee settlements, and 8.56% are provinces with both.

Figure S11 shows the heterogeneous effects of the presence of refugee camps (top plot) versus settlements (bottom plot), respectively, on the main conflict outcomes. All plots confirm that regardless of the type of refugee site, there is no effect of refugee presence on conflict. In fact, there is a negative effect of camps on onset and incidence, which is consistent with our theoretical predictions of increased development in these cases.

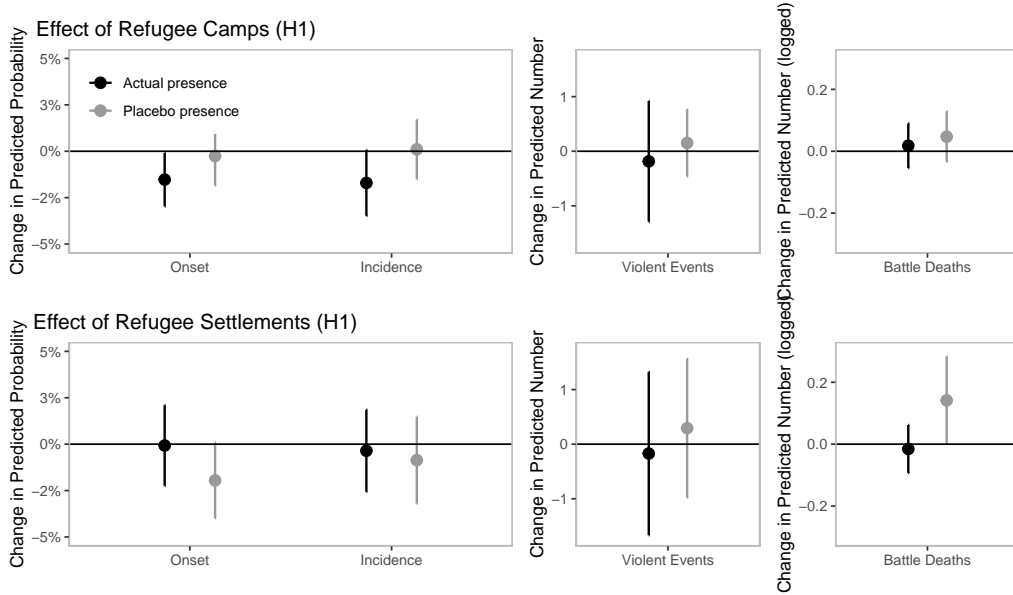


Figure S11: This figure shows the effect of refugee camps versus settlements (heterogeneous effects analysis) on conflict outcomes (black), compared to their respective placebo estimates, i.e. effect of future refugee camps vs. settlements (gray). All point estimates include 95% CIs.

Figure S12 shows the heterogeneous effects of the presence of refugee camps versus settlements, respectively, conditional on refugee camps or settlements in other provinces of the same country-year. The top two plots compare the presence of camps and settlements in geographically concentrated cases, while the bottom two plots compare the presence of camps and settlements in geographically dispersed cases. All plots confirm that there is no effect of refugee presence even if there are refugee sites, whether camps or settlements, elsewhere (dispersed) in the country. Additionally, with respect to the *conditional risk reduction effect* in the case of concentrated refugee

presence in the host country, for camp presence these negative effects are on conflict onset and incidence while for settlement presence, these negative effects are on conflict intensity outcomes, violent events and battle deaths.

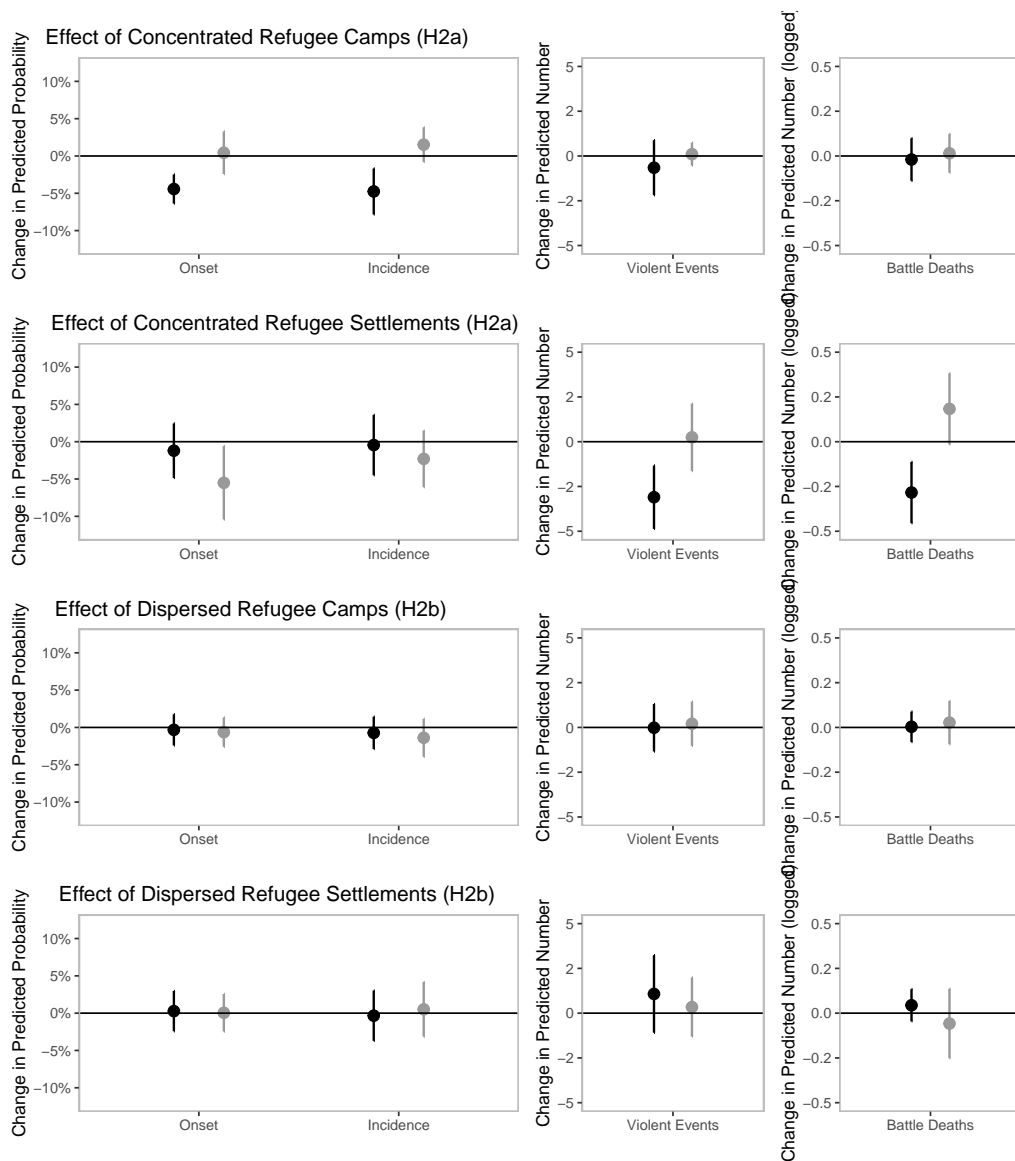


Figure S12: This figure shows the effect of refugee camps vs. settlements (heterogeneous effects analysis) on conflict outcomes (black) compared to their respective placebo estimates, i.e. effect of future refugee camps vs. settlements (gray), conditional on refugee camps vs. settlements in other provinces of the same country-year. All point estimates include 95% CIs.

Figure S13 shows the heterogeneous effects on nighttime lights of the presence of refugee camps versus settlements, respectively, conditional on refugee camps or settlements in other provinces of the same country-year. The top two plots compare the presence of camps and settlements in geographically concentrated cases, while the bottom two plots compare the presence of camps and settlements in geographically dispersed cases. Although not statistically significant, there appears to be positive effects of refugee camps and settlements in geographically concentrated cases; the positive effect for camps approaches statistical significance. This appears supportive of our argument and other findings on development.

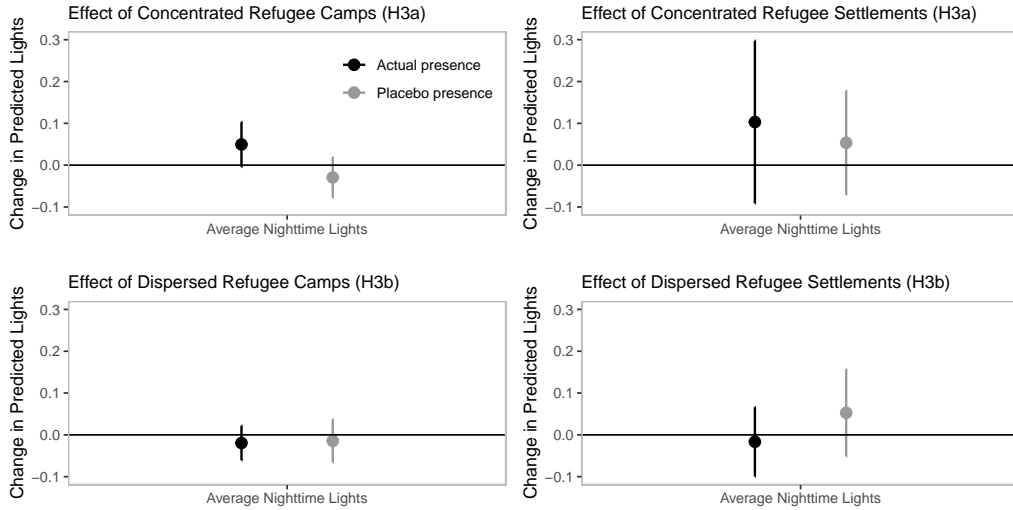


Figure S13: This figure shows the effect of refugee refugee camps vs. settlements (heterogeneous effects analysis) on average nighttime lights (black), compared to their respective placebo estimates, i.e. effect of future refugee presence (gray), conditional on refugee presence in other provinces of the same country-year. All point estimates include 95% CIs.

S5.3 Heterogeneous Effects of Provinces Near vs. Far from Border

This section shows heterogeneous effects by whether provinces are near or far from the international border. Near is coded as within or equal to 100km, which is about the median distance in our data, and far is coded as greater than 100km. Figure S14 shows the heterogeneous effects for provinces far from the border (top plot) versus those near the border (bottom plot), respectively, on the main conflict outcomes. All plots confirm that there is no effect of refugee presence on conflict. In fact, there is a negative effect for provinces in the interior of countries (far from border) on onset and incidence, which may also be consistent with our theoretical predictions of increased development in these cases.

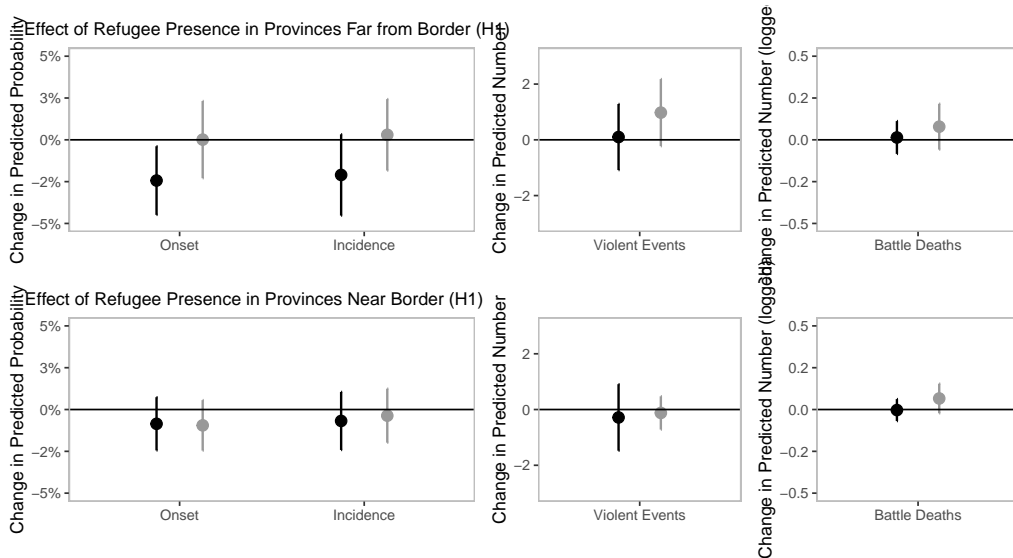


Figure S14: This figure shows the effect of refugee presence in provinces near vs. far from the border (heterogeneous effects analysis) on conflict outcomes (black), compared to their respective placebo estimates (gray). All point estimates include 95% CIs.

Figure S15 shows the heterogeneous effects for provinces far from the border versus those near the border respectively, conditional on refugee presence in other provinces of the same country-year. The top two plots compare far and near border provinces when refugee presence is geographically concentrated, while the bottom two plots do so in geographically dispersed cases. There are no effects of refugee presence in dispersed cases. Additionally, it appears that the conditional risk reduction effects in the cases of concentrated refugee presence are small but present for near border provinces, but prominent for provinces in the interior (far from border).

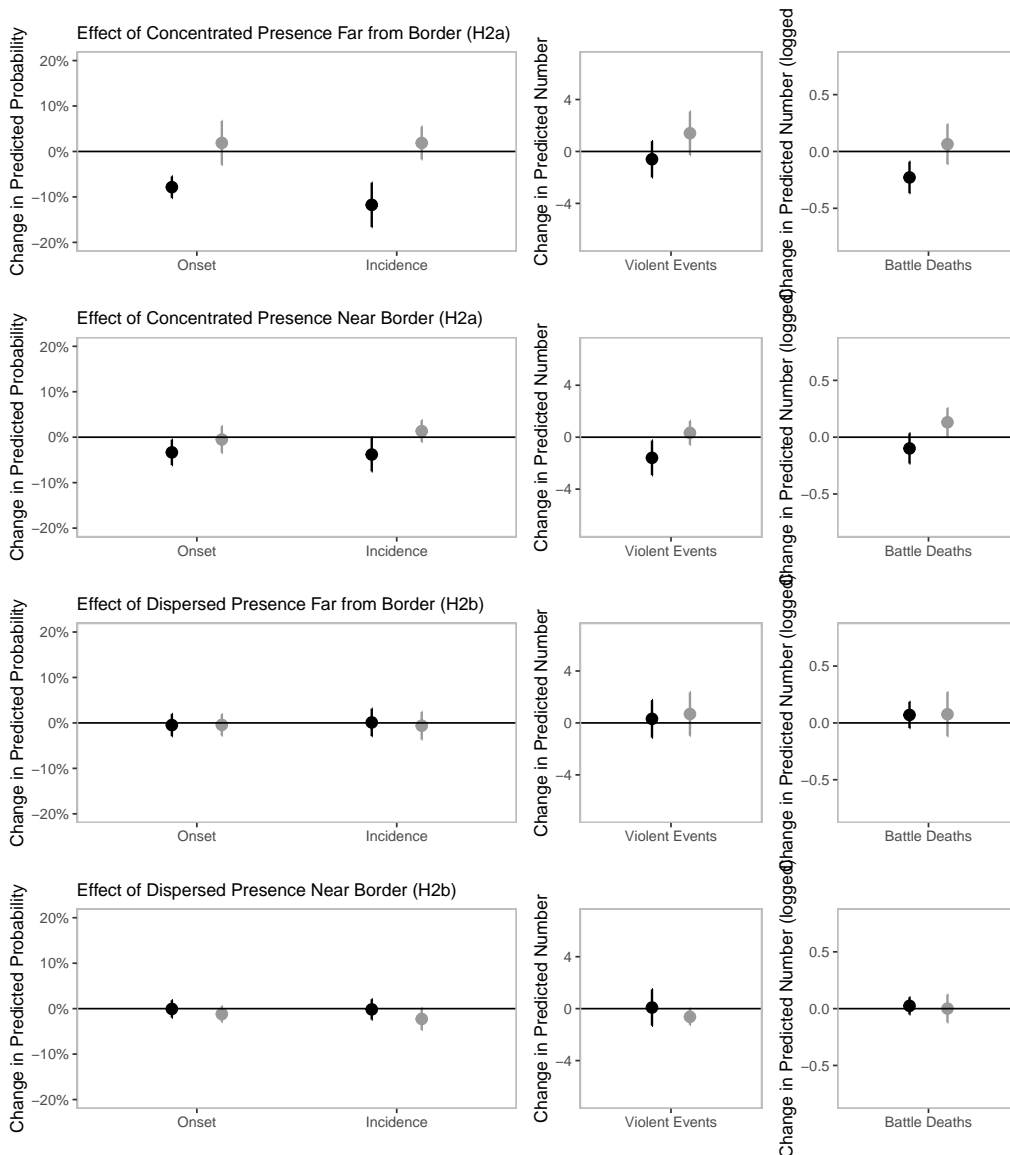


Figure S15: This figure shows the effect of refugee presence in provinces near vs. far from the border (heterogeneous effects analysis) on conflict outcomes (black) compared to their respective placebo estimates (gray), conditional on refugee presence in other provinces of the same country-year. All point estimates include 95% CIs.

Figure S16 shows the heterogeneous effects on nighttime lights for provinces near and far from the border, respectively, conditional on refugee presence in other provinces of the same country-year. The top two plots compare far and near border provinces when refugee presence is geographically concentrated, while the bottom two plots do so in geographically dispersed cases. There are no effects of refugee presence in dispersed cases. There is a positive and statistically significant effect of refugee presence for provinces far from the border (in the interior) when refugee presence is geographically concentrated. This appears supportive of our argument and other findings on development.

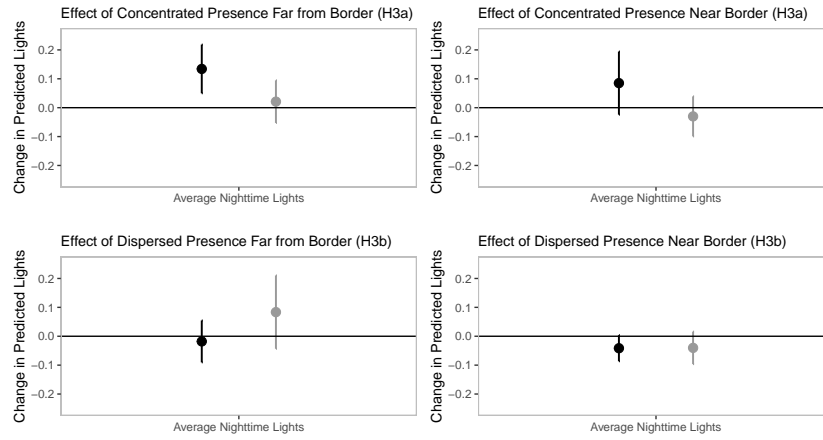


Figure S16: This figure shows the effect of refugee presence in provinces near vs. far from the border (heterogeneous effects analysis) on nightlights (black) compared to their respective placebo estimates (gray), conditional on refugee presence in other provinces of the same country-year. All point estimates include 95% CIs.

S5.4 Subgroup Effects by Region

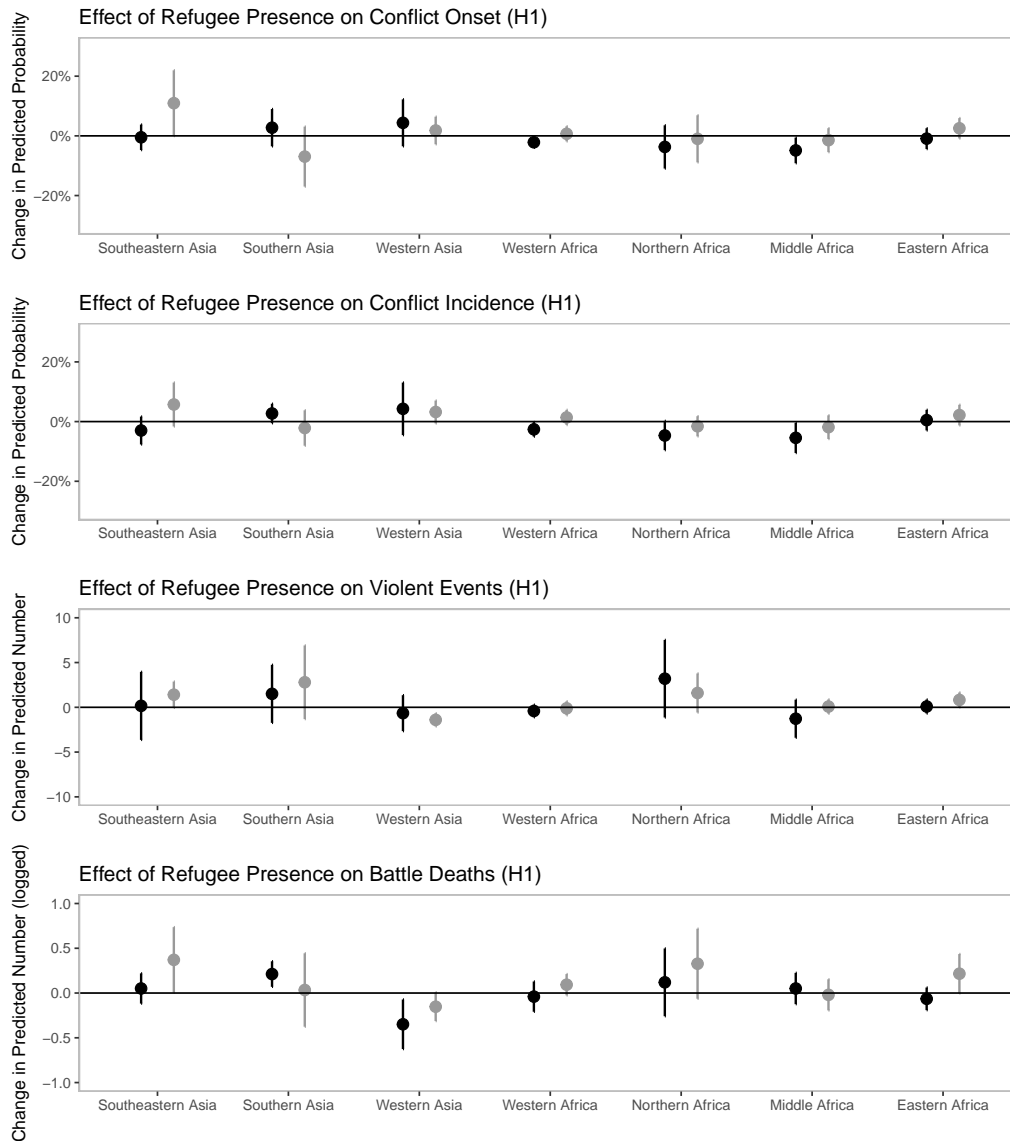


Figure S17: This figure shows the effect of refugee presence by region (subgroup effects analysis) on conflict outcomes (black), compared to their respective placebo estimates (gray). All point estimates include 95% CIs.

There are 14 regions in our data, but 6 of these regions either only had 1 country that hosted refugees or did not have enough observations for both the actual analysis or placebo version. Thus, we end up with 7 regions for subgroup analysis. Of these 7 regions, most show null effects. However, we observe a statistically significant positive effect of refugee presence on battle deaths (as well as positive estimates that do not reach statistical significance for the other three outcomes) in Southern Asia – Afghanistan, Bangladesh, India, Iran, Nepal, and Pakistan. On the other side, we observe a statistically significant negative effect of refugee presence on battle deaths in Western Asia, which includes countries that host Syrian refugees, i.e. Turkey and Jordan.

We do not carry out additional subgroup analysis for H2 (concentrated vs. dispersed presence on conflict outcomes) or H3 (concentrated vs. dispersed presence on nighttime lights) because once we subset further for HTE, there is not enough data to analyze results in a meaningful way. Nevertheless, showing these subgroup effects for H1 are useful for us to understand the overall null effect; in some areas it is a null effects, while in other areas, namely Southern Asia and Western Asia, the null masks positive and negative effects.

S6 Robustness Checks

Here, we detail the following robustness checks for the conflict analyses included in the paper: (1) dropping one country at a time, (2) dropping one region at a time, (3) using the full dataset of all countries, (4) using a dynamically subsetting dataset of only countries that have already hosted refugee sites at time t , (5) using 5 year lead dependent variables to confirm that there are no delayed effects of refugee presence on conflict, and (6) replacing our `Conflict Onset` and `Conflict Incidence` outcomes (1990 - 2008) from ? with `wzoneData`, conflict polygon data constructed by Kikuta (2020) using machine learning based on UCDP GED 19.1 conflict events data (1990 - 2018).

S6.1 Drop One Country at a Time

S6.1.1 Robustness of the Null Effect of Refugee Presence (H1)

Figure S18 shows the distribution of the main effect of the presence of refugee sites on conflict outcomes after dropping one country at a time. There are a total of 70 countries in the main data. All plots show that the predicted probability estimates generated when dropping a country tightly group around the actual estimate, in which no country is dropped, confirming that our findings are not due to outlier countries.

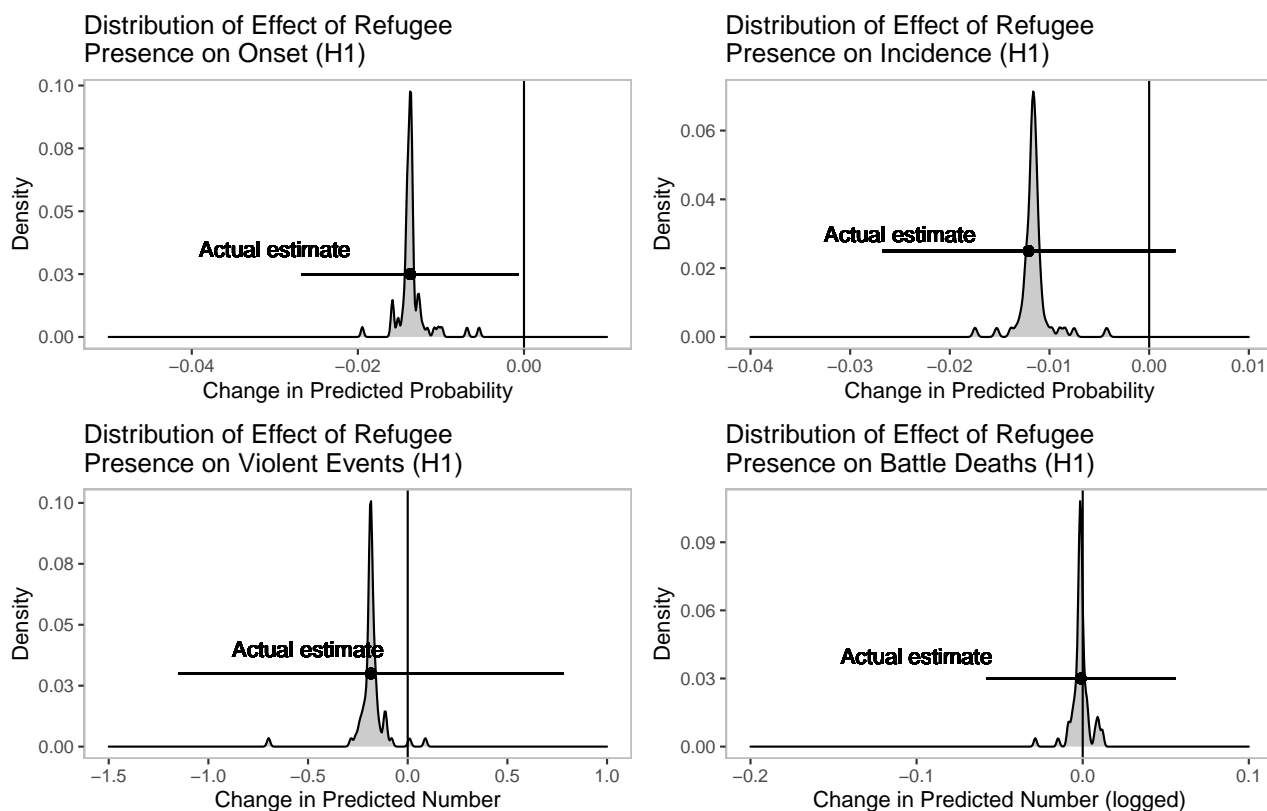


Figure S18: This figure shows the distribution of the main effect of the presence of refugees sites on conflict outcomes after dropping one country at a time compared to the actual estimate (no dropped countries) with 95% CIs. These estimates are grouped tightly around the actual estimates and do not substantively change the main results.

S6.1.2 Robustness of the *conditional risk reduction effect* conditional on Concentrated Refugee Presence (H2a)

Figure S19 shows the distribution of the *conditional risk reduction effect*, in the case of geographically concentrated refugee sites, for conflict outcomes after dropping one country at a time. Again, there are a total of 70 countries in the data. Overall, these distributions, which concentrate around the actual estimates (including all countries), show that the *conditional risk reduction effect* is not driven by any outlier countries.

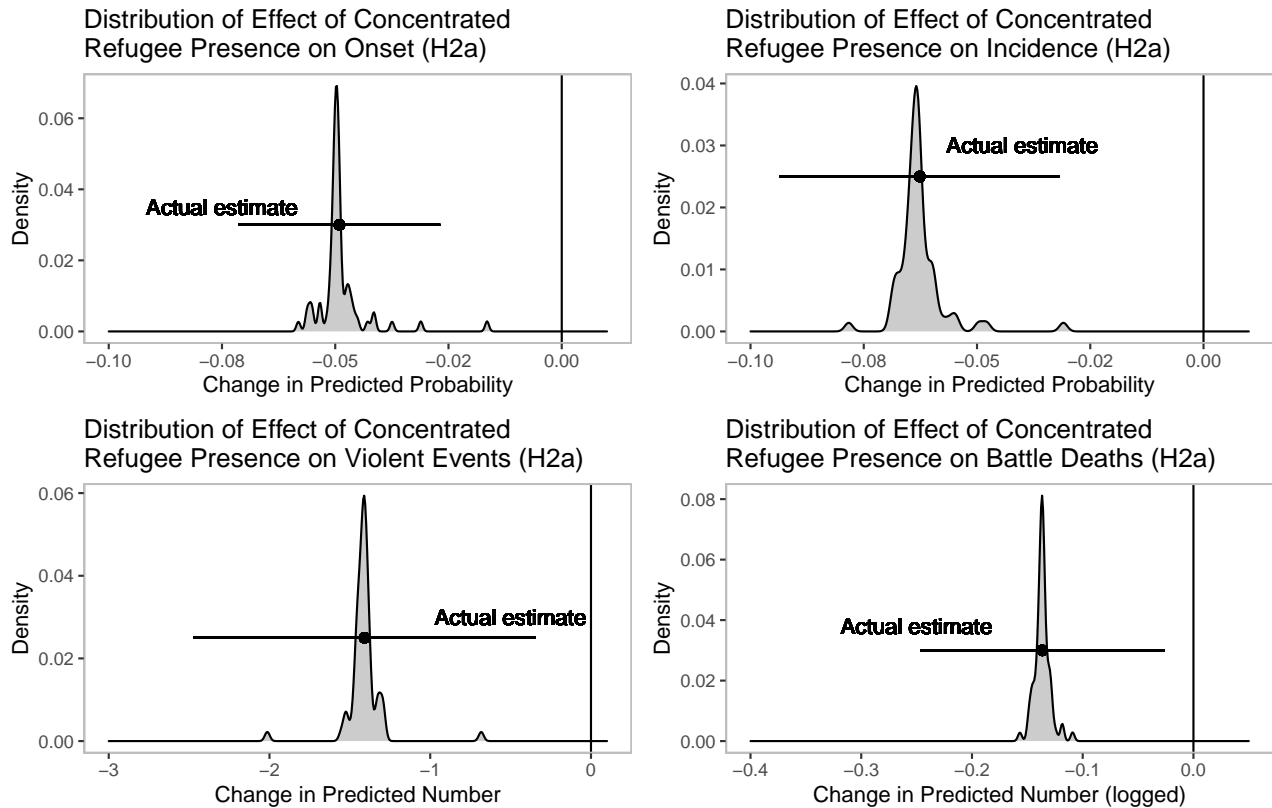


Figure S19: This figure shows the distribution of the *conditional risk reduction effect*, in the case of geographically concentrated refugee sites within a country, after dropping one country at a time compared to the actual estimate (no dropped countries) with 95% CIs. These estimates are grouped tightly around the actual estimates and do not substantively change the main results.

S6.2 Drop One Region at a Time

S6.2.1 Robustness of the Null Effect of Refugee Presence (H1)

Figure S20 shows the barplot of number of provinces dropped per region, and the main null effect of refugee presence once each region is dropped. Overall, these estimates confirm the consistency of the main result of the null effect of the presence of refugee sites on conflict, confirming that our findings are not driven by any particular region.

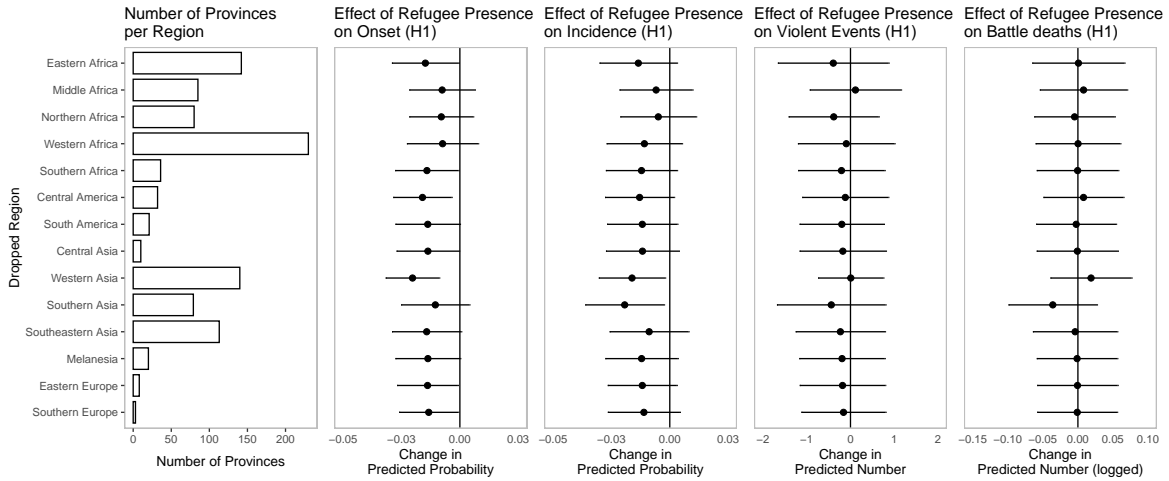


Figure S20: This figure shows the barplot of number of provinces dropped per region, and the effect of the presence of refugee sites on conflict outcomes once each region is dropped, with 95% CIs.

S6.2.2 Robustness of the conditional risk reduction effect conditional on Concentrated Refugee Presence (H2a)

Figure S21 shows the barplot of number of provinces dropped per region and the conditional risk reduction effect, in the case of geographically concentrated refugee sites, for conflict outcomes after dropping one region at a time. All estimates remain negative and aside from incidence after dropping Northern Africa, statistically significant. This analysis confirms that our findings are not driven by any particular region.

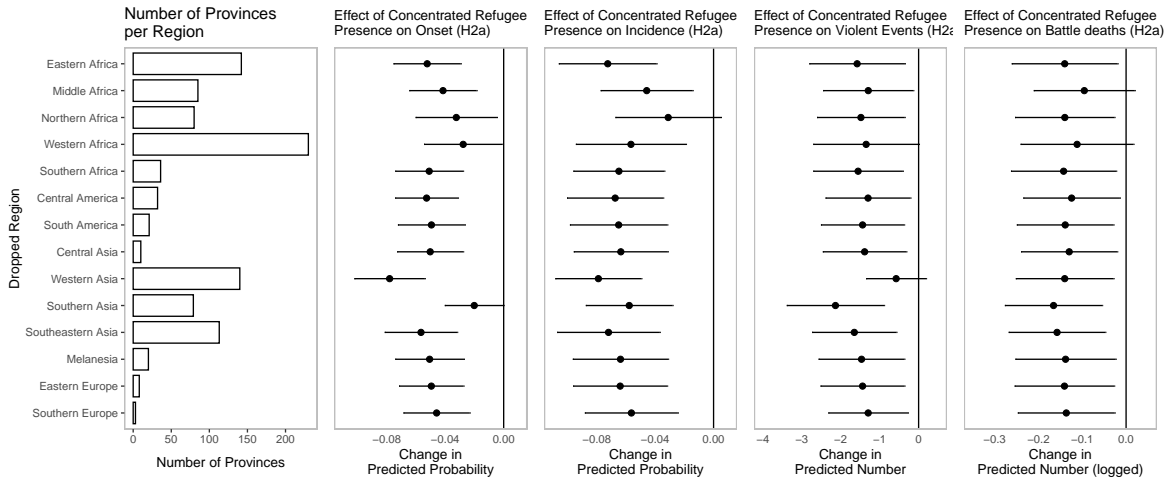


Figure S21: This figure shows the barplot of number of provinces dropped per region, and the conditional risk reduction effect once each region is dropped, with 95% CIs.

S6.3 Analysis using Full Data of All Countries

In this section, we rerun the main (H1) and secondary (H2) analyses in the paper using the full dataset of all countries, as opposed to subsetting down to only countries who have ever hosted refugees during the study period. While we do not believe including all countries, even those who never host refugees, is the correct counterfactual, we perform this analysis for completeness and as a robustness check. Results do not substantively change.

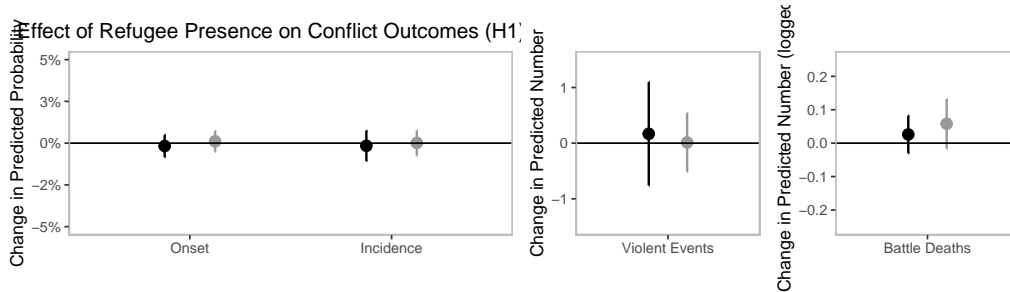


Figure S22: This figure replicates Figure 2 with the full dataset of all countries.

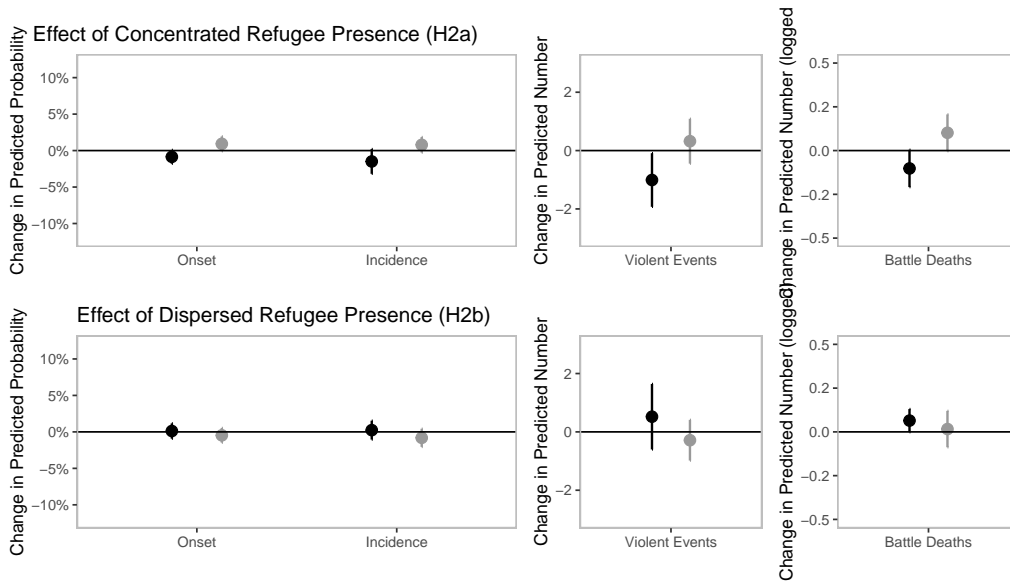


Figure S23: This figure replicates Figure 3 with the full dataset of all countries.

S6.4 Using the Dynamic Subset Data of only Refugee-hosting Countries

This section reruns the main analyses using a modified subset data of only countries that have ever hosted refugee sites. Unlike the original data which includes all countries that have ever hosted refugees *anytime* during our study time period 1990 - 2008, *this version* is subsetting dynamically to address possible concerns of collider bias, which can emerge when analysis at a certain time conditions on variables in the future (Morgan and Winship, 2014). In our case, this means at time t we are including the subset of all countries that not only have hosted refugees by time t but also will host refugees after time t . To address this possible concern, here we create a *dynamic subset* in which at any year t , we only include the subset of countries that have *already* hosted refugees in the years preceding t , so our variation comes from provinces within these already-hosted countries that have not or have not yet hosted a refugee site. However, using this dynamic subset data greatly decreases our number of observations from 28,971 province-years in the main data to 22,539 province-years. Results do not substantively change.

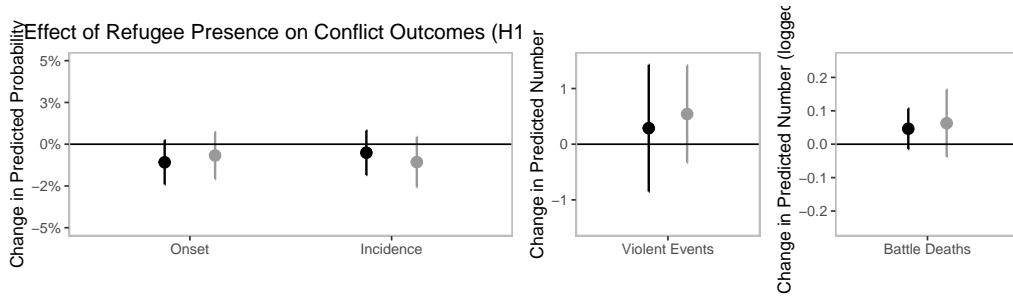


Figure S24: This figure replicates Figure 2 with the dynamic subset data.

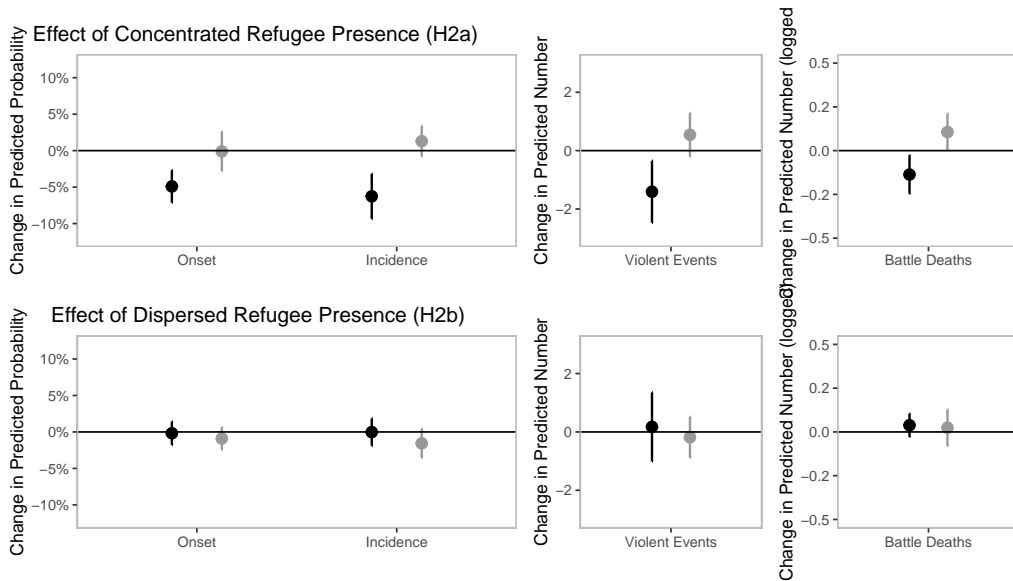


Figure S25: This figure replicates Figure 3 with the dynamic subset data.

S6.5 Leading the Dependent Variables by 5 Years

This section shows the main analysis replacing the conflict onset and incidence outcome variables with conflict onset and incidence five years into the future (lead 5 years) to confirm that there are no delayed effects of the presence of refugee sites on conflict. Note that several of the placebo tests cannot rule out selection effects; however, this would imply refugees selecting into more conflict-prone areas, not that they are selecting less conflict-prone areas.

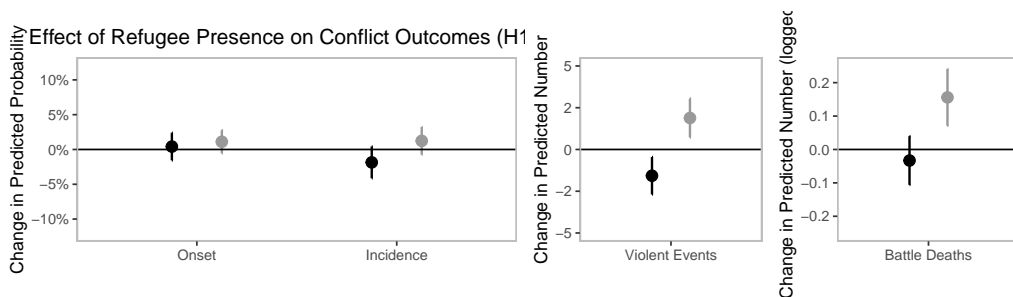


Figure S26: This figure replicates Figure 2 with a 5-year lead DV.

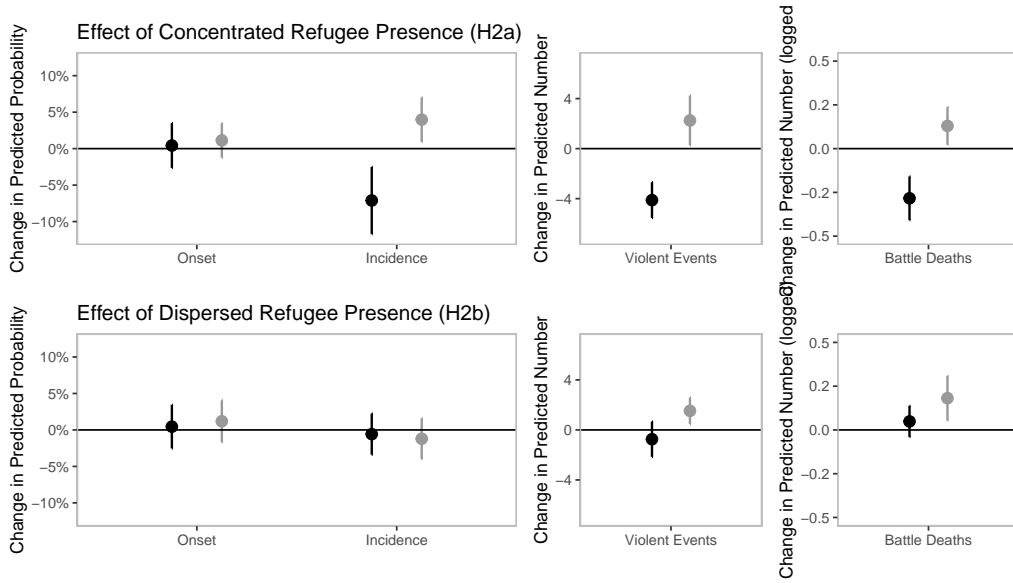


Figure S27: This figure replicates Figure 3 with a 5-year lead DV.

S6.6 Analysis using wzoneData conflict outcomes

For analyses examining conflict onset and incidence, we replace PRIO Conflict Circles (1990 - 2008) with wzoneData Conflict Polygons (1990 - 2018) (Kikuta, 2020). The conflict polygons from wzoneData are created from UCDP GED 19.1 conflict events data (Sundberg and Melander, 2013; Petterson and Öberg, 2020), using a machine learning method. Note that this version of wzoneData excludes conflict events from the Syrian Civil War, activity by the Islamic State, and the US War in Afghanistan and Iraq. The effects shown for number of violent events and the number of logged battle deaths are the same as in the paper, based on UCDP GED 19.1 conflict events data (Sundberg and Melander, 2013; Petterson and Öberg, 2020).

With this alternative measure of onset and incidence, Figure S28 confirms that there is still not effect of refugee presence on these outcomes. For H2, the *conditional risk reduction effects* are now null.

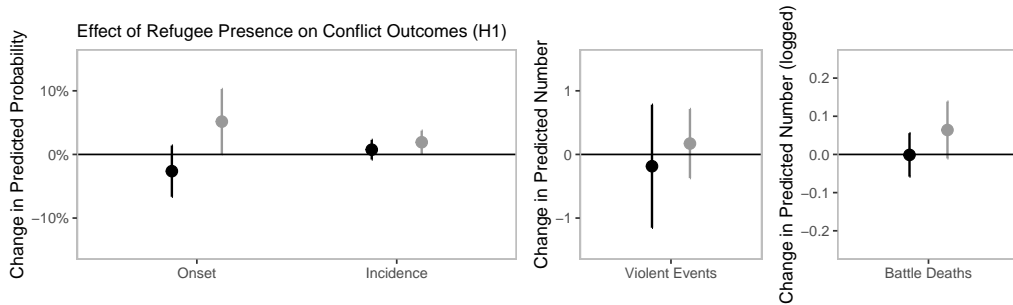


Figure S28: This figure replicates Figure 2 using wzoneData-based Conflict Onset and Incidence measures.

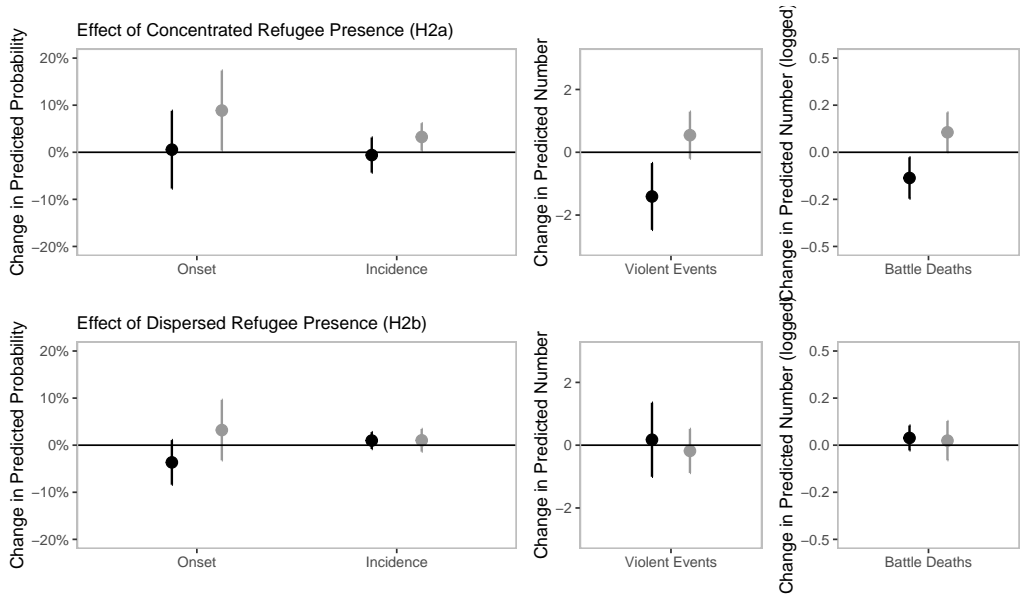


Figure S29: This figure replicates Figure 3 using wzoneData-based Conflict Onset and Incidence measures.

S7 Expert Interviews with UNHCR Officials

In 2018, we interviewed twelve senior officials of the UNHCR about their expertise knowledge on UNHCR current and historical policies and practices; refugee issues in general; and their observations of refugee-host relations and refugee hosting dynamics in the countries where they had worked in their capacity as UNHCR officials. To be clear, these were not elected public officials. Our institution’s IRB advised us that because we were interviewing experts on their professional expertise (external topic) rather than asking any personal information, we did not need to submit a protocol for IRB review. With respect to the specific principles outlined in the 2020 APSA Principles and Guidance for Human Subjects Research⁴, we obtained informed written (by scheduling over email with our project summary and list of possible questions) and oral (at the start of the interview) consent. Our research involved no to minimal harm or trauma. We did not use any deception in our interviews. We were also considerate of the experts’ time, by limiting our interviews to 1-2 hours. We do not expect their reputations to be negatively affected, especially since we keep their identities confidential. We do not name them, but rather only refer to them by their position or title. Below, we include additional longer-form quotes from our expert interviews categorized by topic.

S7.1 General null effects of refugee-hosting on conflict

“Despite all the hysteria of refugees and conflict, refugees account for a small part of the population. This is an eminently manageable problem. Those populations are predominantly found in the poorer countries of the developing world. The Syrian arrivals may have distorted that picture. Those kind of figures in Uganda and in Bangladesh, they have been managed without disastrous consequences by governments with much fewer resources. What occurred in Europe is a political crisis, it’s not a refugee crisis... I’m very familiar with the Afghan situation and in 2013 and 2014 with the Syrian, and I’ve never see a conflict at scale. Of course, like in Lebanon where the country’s situation has had to under normal demographic situations, the would only reach it in 2050 if not for the refugees. And they themselves emerged from a long civil war, and turbulent relations with Syria. It’s quite remarkable that there hasn’t been any major conflict.” – Senior Official at Policy Development and Evaluation, interview conducted on July 6, 2018.

“Certainly there are instances where the presence of refugees creates [non-violent] conflict with the local community like competition for resources, access to fire wood collection, environmental impact. You see those kinds of immediate impact. In a broader sense, I have never seen a situation where refugees have fueled [violent] conflict.” – Senior Policy Officer, interview conducted on July 23, 2018.

S7.2 On militarized refugees and prevention

“From the Kakuma perspective, the refugee camps are of a civilian nature and we try to maintain it that way. When we have identified armed elements, we tell the government and their security apparatus could separate them from the civilians. It is mainly the government who identifies them at the entry points, so the government interviews them. Armed elements are not simply child soldiers, but known commanders. It’s not of course perfect, but once they identify them. Child soldiers are a different ballgame, the civilian asylum system does not mean they are not ex-combatants. When they come and they are not engaged in any military activity then they are accepted. As children, once they come, in interviews and in schools, the teachers and social workers realize that there is something and when the method is brought to us, you try to pay much more closer attention to that. Not like in West Africa, we do not see many child soldiers. Our efforts try to prevent them from becoming combatants. They might choose to go back to the other side of the border to join military action. We are engaging in vocational training, sports, arts, we are creating much positive alternatives.” – Senior Official at Kakuma Refugee Camp, interview conducted on July 23, 2018.

S7.3 Refugee-host relations

“Based on my experience, what you see often is that refugees go to parts of the country, cross the border not too far where they came from, and the border that you see is artificial. You have on both sides of the border population that know each other from the same clan. Like Dadaab the Somalis on either side. That’s the first element, you have communities which know each other who have been moving up and down, not because of conflict, due to drought and famine. These communities know each other and are often homogeneous. In this

⁴<https://connect.apsanet.org/hsr/principles-and-guidance/>

kind of context, you don't go where you're not welcome or wanted. You move to parts of the region where you know people will not reject you. People who are in a position to provide you some support." – Senior Official at UNHCR Kenya, interview conducted on July 6, 2018.

"Second, when large populations across the borders take place, the shock of the demographic expansion can be to a greater or lesser extent, depending on the ethnic affinities across the border, most instances these populations move to areas with similar cultural, religious, linguistic affinities. With the way that the colonial boundaries were drawn in Africa, you find that the populations are very very similar." – Senior Official at Policy Development and Evaluation, interview conducted on July 6, 2018.

"We would assume that if different communities live side by side or live together, there will be conflict. The working assumption is that there will be conflict. So UNHCR and aid agencies try to put in place prevention mechanisms. And that means conflict between refugees and host populations. That normally is done over time and by creating informal structures, committees, where the local people like elders and chiefs would meet regularly. So that is the understanding that we have. There are other issues where UNHCR would advocate that camps should not be outnumbering local communities. That we limit camps in size. We would assume that in poor countries that the host population would need to be compensated in one way or another for hosting refugees, and that compensation is either giving them employment in camps or giving them, mobilizing other partners to invest in host communities. That is the normal situation. Every community has structures. Even the refugee community is not just some amorphous group of people but they have spokespeople, traditional leaders, more educated than others. We will find these people pretty fast and easily because they are our interlopers. On the side of the host communities, we also have existing institutional structures: state sponsored, traditional chiefs this and that. Depending on circumstances, we would try to sensitize the local government at an early stage on setting up these informal structures in the beginning... What I would call this is the unwritten rulebook. I doubt you will find a lot of in our archives and written instructions. This is in a way a part of the oral culture of UNCHR, colleagues who have been in the business for many years, try what has worked elsewhere and train young colleagues. In the end of the day this is all common sense... The working assumption is that conflict is going to happen and that it needs to be managed, and if not, it might become violent. So you better intervene early." – Senior Official at Division of Programme Support and Management, interview conducted on September 12, 2018.

S7.4 Effects of refugee-hosting on development, infrastructure, other services

"Good relationships between citizens and refugees is a phenomenon that is monitored, and a goal that drives protection interventions, such as infrastructure development or individual assistance programs that are allocated specifically to local communities rather than refugees alone. The intent is to help people (refugees) make a life, while benefitting local communities. Unfortunately there is not much data on this." – Senior Regional Protection Officer, interview conducted on June 25, 2018.

"There may be a hospital built for the refugees or clinic, and then it is opened to the host community. Schools, and things like that. There's an effort like that in some areas." – Senior Official 1 at FICSS, interview conducted on June 25, 2018.

"On the African continent, the assistance provided is very significant, in areas with limited government presence, areas that were not priorities. Garrison in northeastern part of Kenya, the level of illiteracy and poverty. Same thing in Turkana, the presence of refugees brought better schools, better healthcare, brought food assistance." – Senior Official at UNHCR Kenya, interview conducted on July 6, 2018.

"On the distribution of humanitarian resources, I suppose the main observation that in terms of when people are concentrated, the tendency is to invest in services and facilities that directly serve that population. I suppose this has been the tradition that when camps are set up the way that things were traditionally done, you basically set up a medical center in the camp, water infrastructure serving that camp, food distribution would be carried out in the camp. The aid would be much more focused on the refugee population. Increasingly over the years, there was a real push to set things up in such a way that the local community was also benefitting from them like health clinics for example, the nearby villages to also benefit from that facility. More recently, the shift has been to how you boost services more generally to areas where refugees are located, enhancing existing services. Where

you have a camp based program, what that does do is allow you to concentrated your resources in one location. Whereas if you're working in a context that is much more spread out, in smaller settlements or perhaps not in camps at all, the shift is much more is to boost national and local services rather than set up dedicated facilities or subsidizing refugees through cash assistance. Or you might have a program that targets the most vulnerable among the refugees, and UNHCR reimburses the cost of treatment in the local facilities." – Senior Policy Officer, interview conducted on July 23, 2018.

S7.5 Economic effects of hosting refugees

"When there is a reasonable number of refugees, it attracts resources of international community. There is a net economic gain in that area... The physical pretense of a large number of internationals might have some sort of calming effect, that might change the social and political dynamics." – Senior Official of Shelter & Settlements Section, interview conducted on July 3, 2018.

"If you look at rural sub-Saharan Africa, there is not much money circulating in the villages. If you have a camp where you pay staff, then you create an imbalance. The refugee camps become a big market where locals sell their produce there." – Senior Official at Division of Programme Support and Management, interview conducted on September 12, 2018.

"The winners, you have a larger number of consumers, you do have people with capital and skills that can contribute positively to the economy. And employers can find a ready supply of laborers at lower wages. The first 6 months a new equilibrium imposes itself. With resources from the international community, you have strong injections of finance into the local economy. So it's not altogether surprising that you go through the trajectory of at first shock, but then a period of stabilization, and even a changed environment. And things can evolve relatively positively." – Senior Official at Policy Development and Evaluation, interview conducted on July 6, 2018.

"[On Kakuma] Today if they want to have a job, they come here. Because there is a vibrant economy here, not only with humanitarian agencies. The refugees contribute to 3% economic growth every year in Kakuma. It is the second biggest city in Kenya, it's open. So it is totally 15 km radius, there are a quarter million people living who need food, washing detergent, whatever you need. They have driving schools, hardware stores, cinemas, restaurants, hotels, name it, bodas bodas, electronics, tailors, sports clubs. Long term, I think Kakuma it is the Kenyan government to maintain this camp. We are now integrating them into the Turkana government and population. And the government there realized that if they don't include this population (15%) into their development plans, they must because they are an economic engine." – Senior Official at Kakuma Refugee Camp, interview conducted on July 23, 2018.

"There are opportunities, particularly in camps, which creates a demand which often becomes a golden goose, an economic hub. Some of the camps in Kenya and Jordan. This is not something that authorities would agree with. They emphasize the negative effects as a funding strategy, and the negative rhetoric leads to negative attitudes because the refugees are being scapegoated." – Senior Official at UNHCR Syria, interview conducted on August 2, 2018.

S7.6 Effects of refugee-hosting on local state capacity

"I think the refugees arrive and there is a lack of state security, UNHCR asks for a significant increase in state security and the international community does, and we usually pay the state to provide additional security. Additional police, military, all kinds of things take place because there are conflicts within the community, but not necessarily because of conflicts surrounding, but just because of a large population that is displaced to ensure that there is some kind of law in order. It is not the responsibility of the humanitarian community to provide law and order, it is the responsibility of the host government, so we always ask for the host government to provide that law and order." – Senior Official 1 at FICSS, interview conducted on June 25, 2018.

"It's not only international humanitarian agencies, often times it's the military and the police presence are also there." – Senior Official 2 at FICSS, interview conducted on June 25, 2018.

"The other point worth making is that very often these populations have accumulated in border areas which are often the least developed parts of the country. So while the arrival of refugees evokes a security reflex, those

in charge of refugees in the government are placed under the ministry of security or ministry of the interior rather than the ministry of social, they tend to attract financing even through humanitarian channels into areas that generally did not have that interest in that part of the country. You have external financing that leverages government fiscal expenditures.“ – Senior Official at Policy Development and Evaluation, interview conducted on July 6, 2018.

S7.7 Differences between refugee camps and settlements

“In camps, more money is spent per capita on refugees, and there is increased ability to deliver assistance. Camps are more cost effective and provide physical security and services, but they are not meant for people to stay there forever. An example is Iraq, where refugees are hosted in camps for security reasons.” – Senior Regional Protection Officer, interview conducted on June 25, 2018.

“One thing I think is important to bear in mind, in some of these instances the government policy has been containment into camps. As long as the political will and support is there on the part of the donors to continue to subsidize those arrangements, then the issue of conflict with local citizens does not arise. My view is that [camp-based policy] is not helpful and it’s more costly than it needs to be.” – Senior Official at Policy Development and Evaluation, interview conducted on July 6, 2018.

“At UNHCR, we don’t advocate for camps. We would like no big concentrations of refugee populations in a place and yet, still the countries prefer a camp style arrangement for various reasons. Some of them might say that ‘I want more efficiency of programs, I want to make sure that there is no burden on my budgets so I can geographically distinct them, I want for security reasons to control the movements.’ There are many reasons why the governments prefer camps.” – Senior Official at Kakuma Refugee Camp, interview conducted on July 23, 2018.

S7.8 How contemporary refugee-hosting dynamics differ from the Cold War era

“Basically your [study] period is after Cold War. The Cold War period was a very distinct period because in Africa, there were a lot of freedom fighters.” – Senior Official at Kakuma Refugee Camp, interview conducted on July 23, 2018.

“There’s been quite an evolution since the 1980s, refugees camps in Pakistan were used as military bases. Cambodian camps were used as bases to oppose Cambodian occupation. Since that time, that form of politically motivated policy is no longer standard. UNHCR in particular has insisted as much as possible to maintain the civilian characteristic of these camps.” – Senior Official at Policy Development and Evaluation, interview conducted on July 6, 2018.

S7.9 Suggestions for better policy

“UNHCR has a standing policy that says camps are good for emergencies, but it’s better in the long run to give people freedom of movement. In addition, it could negatively impact a country, by pushing the country into restricting freedom of movement to greater degree by cloistering citizens into a singular province.” – Senior Regional Protection Officer, interview conducted on June 25, 2018.

“[In Kenya] we’ve had a number of situations where the government said that refugees were a threat to national security. We hear that around the world now. It’s clear that when you have a camp like Dadaab on the border with Somalia where you have an active conflict. It’s not a conventional war, it’s terrorist attacks. What we’ve been telling the government, you should have better asylum management and strategy. To register people better, to avoid strict encampment policy to lessen the radicalization of people. We’ve tried to discuss with the government with a policy towards socio-economic inclusion. Try to move away from strict encampment policy. This is what we’re trying to push for. It’s better for security than have a concentration of people. The idea is to have a more settlement type than a camp. Then you would still be able to identify the population but the type of assistance will be different. Instead of building a school further away from the host community, you can strengthen existing infrastructure. You look to see what the needs of the host populations are and taking into account. To a certain extent, that is what is being done in Uganda. What you then do is to augment the host community infrastructure.” – Senior Official at UNHCR Kenya, interview conducted on July 6, 2018.

S8 Regression Tables

S8.1 H2: Concentrated or Dispersed Refugee Presence on Conflict

	Onset	Onset (p)	Incidence	Incidence (p)	Events	Events (p)	Deaths	Deaths (p)
Refugee Presence	-1.26*** (0.35)		-0.95*** (0.26)		-1.41** (0.54)		-0.14* (0.06)	
Placebo Refugee Presence		-0.01 (0.23)		0.19 (0.14)		0.54 (0.38)		0.11* (0.05)
Refugee Presence in Other Provs	-0.53** (0.18)	-0.62** (0.22)	-0.17 (0.12)	-0.08 (0.14)	0.07 (0.32)	-0.15 (0.34)	-0.04 (0.03)	-0.05 (0.03)
Ref Pres x Ref Pres in Other Provs	1.20** (0.39)		0.95*** (0.28)		1.58* (0.77)		0.17** (0.07)	
Placebo Ref Pres x Ref Pres in Other Provs		-0.25 (0.29)		-0.43* (0.19)		-0.73 (0.49)		-0.08 (0.08)
Lagged Onset	-6.19*** (1.82)	-6.09*** (1.81)						
Lagged Incidence			2.45*** (0.07)	2.43*** (0.07)				
Lagged Events					0.64*** (0.11)	0.50*** (0.13)		
Lagged Deaths							0.59*** (0.01)	0.58*** (0.01)
Sum Events in Neighbor Provs	0.02*** (0.00)	0.02*** (0.00)	0.02*** (0.00)	0.02*** (0.00)	0.05* (0.02)	0.08* (0.04)	0.00*** (0.00)	0.00*** (0.00)
Lagged Population (logged)	0.12 (0.06)	0.07 (0.07)	0.13** (0.04)	0.12* (0.05)	0.51*** (0.15)	0.58*** (0.15)	0.05*** (0.01)	0.05*** (0.01)
Lagged GDP	-0.01* (0.00)	-0.01 (0.00)	-0.00 (0.00)	-0.01** (0.00)	-0.02** (0.01)	-0.01* (0.00)	-0.00 (0.00)	0.00 (0.00)
Terrain Ruggedness	0.00* (0.00)	0.00* (0.00)	0.00* (0.00)	0.00* (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00*** (0.00)	0.00*** (0.00)
Province Size (sqkm)	-0.00 (0.00)	-0.00 (0.00)	-0.00*** (0.00)	-0.00** (0.00)	0.00 (0.00)	0.00 (0.00)	0.00*** (0.00)	0.00*** (0.00)
Distance from border (km logged)	-0.01 (0.07)	0.00 (0.08)	-0.04 (0.05)	-0.02 (0.05)	-0.23 (0.15)	-0.25 (0.18)	-0.03** (0.01)	-0.03** (0.01)
Distance from capital (km logged)	-0.15* (0.07)	-0.13 (0.08)	0.10 (0.05)	0.13* (0.06)	-0.01 (0.19)	0.08 (0.22)	0.00 (0.01)	0.00 (0.01)
IDP Presence	0.88 (0.87)	0.00 (2.50)	0.13 (0.42)	-0.36 (0.66)	2.33 (4.91)	7.26 (9.09)	0.20 (0.12)	0.30 (0.16)
Intercept	-1.40 (1.22)	-1.04 (1.32)	-2.30* (0.90)	-2.53** (0.97)	3.65 (3.10)	1.06 (4.14)	0.56** (0.19)	0.53** (0.20)
AIC	3907.85	3345.18	7775.76	6593.46				
Log Likelihood	-1853.93	-1574.59	-3787.88	-3198.73				
Num. obs.	12801	11351	17244	15280	26512	22600	26512	22600
R ²					0.64	0.62	0.57	0.56
Adj. R ²					0.64	0.62	0.57	0.56

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table S4: Regression table for H2: Effects of Geographically Concentrated (coefficient for Refugee Presence) or Dispersed (coefficient for Refugee Presence x Other Prov Refugee Presence) Presence on Conflict Outcomes and their respective Placebo models (p). Note that for outcomes Onset and Incidence, the data is from 1990-2008, while Violent Events and Battle Deaths (logged) is 1990-2018. All models include control variables and Country and Year fixed effects.

S8.2 H3: Conditional Risk Reduction through Increased Infrastructure

	Nighttime Lights	Nighttime Lights (p)
Refugee Presence	0.10*	
	(0.05)	
Placebo Refugee Presence		-0.02
		(0.03)
Refugee Presence in Other Provs	0.08**	0.07*
	(0.03)	(0.03)
Ref Pres x Ref Pres in Other Provs	-0.13*	
	(0.05)	
Placebo Ref Pres x Ref Pres in Other Provs		0.02
		(0.04)
Lagged Nighttime Lights	96.65***	96.67***
	(0.64)	(0.67)
Ave Nighttime Lights in Neighbor Provs	0.04***	0.04***
	(0.01)	(0.01)
Sum Violent Events in Neighbor Provs	0.00	0.00
	(0.00)	(0.00)
Lagged Population (logged)	0.03***	0.03***
	(0.01)	(0.01)
Lagged GDP	-0.00	-0.00
	(0.00)	(0.00)
Terrain Ruggedness	-0.00	-0.00
	(0.00)	(0.00)
Province Size (sqkm)	-0.00	-0.00
	(0.00)	(0.00)
Distance from border (km logged)	-0.01	-0.01
	(0.01)	(0.01)
Distance from capital (km logged)	-0.08***	-0.09***
	(0.02)	(0.02)
IDP Presence	0.05	0.00
	(0.03)	(0.05)
Intercept	1.88***	1.98***
	(0.18)	(0.20)
R ²	0.98	0.98
Adj. R ²	0.98	0.98
Num. obs.	16200	13982

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table S5: Regression table for H3: Effects of Geographically Concentrated (coefficient for Refugee Presence) or Dispersed (coefficient for Refugee Presence x Other Prov Refugee Presence) Presence on Nighttime Lights and their respective Placebo models (p). All models include control variables and Country and Year fixed effects.

S8.3 H1 Models Replacing Refugee Presence (binary) Variable with Refugee Population and Ratio of Refugee to Local Population

For the 2010-2015 Africa subset data, Models 1 and 4 show the effects of logged Refugee Population on Number of Violent Events and logged Battle Deaths. Models 2 and 5 show the effects of the ratio of Refugee Population to Local Population in the Province-Year on the same conflict outcomes. And Models 3 and 6 show our main H1 models using Refugee Presence as the independent variable of interest; we include these to make sure that results are comparable since we are using a subset data of African countries from 2010 to 2015.

There are null effects for all three Refugee related independent variables. But there does appear to be a negative (not statistically significant) effect of greater ratio of refugees to locals on these conflict outcomes.

	Events	Events	Events	Deaths	Deaths	Deaths
Refugee Population (logged)	0.19 (0.12)			0.01 (0.01)		
Ratio Refugee-Local Population		-1.60 (1.75)			-0.47 (0.32)	
Refugee Presence (binary)			1.13 (0.79)			0.02 (0.06)
Refugee Presence in Other Provs	0.33 (0.33)	-0.20 (0.35)	0.33 (0.36)	0.04 (0.06)	-0.01 (0.06)	0.02 (0.06)
Lagged Events	0.68*** (0.20)	0.68*** (0.20)	0.68*** (0.20)			
Lagged Deaths				0.58*** (0.03)	0.58*** (0.03)	0.58*** (0.03)
Sum Events in Neighbor Provs	0.03** (0.01)	0.03** (0.01)	0.03** (0.01)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Lagged Population (logged)	0.63* (0.28)		0.63* (0.28)	0.03 (0.02)		0.03 (0.02)
Lagged GDP	-0.03 (0.10)	0.08 (0.10)	-0.03 (0.10)	0.03** (0.01)	0.03*** (0.01)	0.03** (0.01)
Terrain Ruggedness	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Province Size (sqkm)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Distance from border (km logged)	-0.47 (0.30)	-0.54 (0.32)	-0.48 (0.30)	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.02)
Distance from capital (km logged)	-0.71 (0.71)	-0.70 (0.68)	-0.66 (0.69)	-0.06* (0.03)	-0.06 (0.03)	-0.05 (0.03)
IDP Presence	-2.61 (3.15)	-2.16 (3.08)	-2.49 (3.12)	0.42 (0.25)	0.45 (0.25)	0.43 (0.25)
Intercept	-2.95 (5.90)	5.51 (5.26)	-3.19 (5.85)	-0.21 (0.40)	0.20 (0.25)	-0.24 (0.40)
R ²	0.59	0.59	0.59	0.66	0.66	0.66
Adj. R ²	0.58	0.58	0.58	0.65	0.65	0.65
Num. obs.	3252	3252	3252	3252	3252	3252

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table S6: Regression table for Effects of Refugee Population on Conflict Outcomes in Africa (2010–2015). All models include control variables and Country and Year fixed effects.

S9 Salehyan and Gleditsch (2006) Extension Exercise

Our study at the subnational level shows a null effect of refugee presence on conflict outcomes, and in cases when refugees are geographically concentrated in the host country, substantively large negative effects. These findings appear to run counter to the existing seminal quantitative research of Salehyan and Gleditsch (2006), who find that increased refugee populations from neighboring countries are correlated with increased likelihood of conflict onset. However, our study covers the period from 1990 to 2018, while their data are from the period from 1951 to 2001. To test whether the results of our analysis generally obtain if we carry out following the country-level approach that Salehyan and Gleditsch (2006) take, we replicate their study for the period 2002 through 2013.⁵

Table S7: SG Contemporary Comparison: Logistic regression results of Conflict Onset on Refugee Population (Logged)

	<i>Dependent variable:</i>			
	Original M1	Original M2	Extension M1	Extension M2
Refugees (logged)	.10* (.07)	.20** (.06)	-.03 (.07)	-.05 (.06)
Civil War in Neighbor	.50** (.20)	.30** (.20)	1.00 (.90)	.90 (.70)
Polity	.01 (.01)	.01 (.01)	-0.00 (.06)	-0.00 (.05)
Polity Squared	-.01*** (0.00)	-.01*** (0.00)	-.02 (.01)	-.02* (.01)
GDP (logged)	-.20* (.10)	-.20** (.09)	-.10 (.30)	-.07 (.20)
Population (logged)	.20*** (.05)	.30*** (.04)	.80*** (.20)	.70*** (.20)
Ethnic Heterogeneity	.02*** (.01)	.02*** (0.00)	-.80 (1.00)	-1.00 (1.00)
Ethnic Kin	.60*** (.20)		2.00* (1.00)	
Peace Years	-.50*** (.07)	-.50*** (.06)	.10 (.07)	.10* (.06)
Spline 1	-.01*** (0.00)	-.01*** (0.00)	-2.00 (2.00)	-1.00 (1.00)
Spline 2	0.00*** (0.00)	0.00*** (0.00)	-3.00* (2.00)	-4.00*** (1.00)
Spline 3	-0.00 (0.00)	-0.00 (0.00)	-14.00*** (4.00)	-14.00*** (3.00)
Constant	-3.00*** (.00)	-2.00*** (.90)	-14.00*** (5.00)	-10.00*** (3.00)
Observations	3,591	5,567	892	1,521
Log Likelihood	-658.00	-822.00	-55.00	-76.00
Akaike Inf. Crit.	1,343.00	1,669.00	137.00	175.00

Note:

*p<0.1; **p<0.05; ***p<0.01

This table shows the regression results of the Salehyan and Gleditsch (2006) reanalysis and extension exercise comparing their analysis during the period of their study (1946 - 2001) versus contemporary period (2002 - 2013). The first and second columns replicate Salehyan and Gleditsch (2006) using their data. The third and fourth columns show results for the same models in the contemporary period using our extended data. This table shows that in the contemporary period (although there are fewer observations because of the truncated time period), the coefficient for refugees becomes negative, although it is not statistically significant.

Before describing the results of this exercise, we briefly detail our replication and extension approach. We extend the authors' analysis to the more contemporary period up to 2013. We construct a panel dataset that matches the original on all variables. For conflict onset, the outcome of interest, we use the same conflict dataset that these scholars used and apply the same binary coding to the "main dependent variable [of] conflict onset, which is coded 1 for the first year of a conflict and 0 if no conflict takes place in the state in that particular year (Gleditsch et al., 2002). Subsequent ongoing years of the same conflict are dropped from the estimation sample" (Salehyan and Gleditsch, 2006, p. 350). We also follow these authors in their treatment of cases of conflict onset that occurred during the years when another conflict was ongoing – "In cases where there were multiple conflict onsets in a country, data on a new onset was included if it occurred during the years when another conflict was ongoing" (Salehyan and Gleditsch, 2006, p. 350). For their primary explanatory variable, refugee population, we follow these authors in using country-level data with refugee population values developed by the Population Data Unit of the Office of the United Nations High Commissioner for Refugees (UNHCR). Using the country distance

⁵2013 was the last year for which we were able to obtain country-year data for all variables included in their model.

measures developed by the authors for their analysis, we similarly use their “restrictive definition [of qualifying refugee sending states] in which neighbors are defined by borders falling within a distance of 100 kilometers or less (including contiguity)” (Salehyan and Gleditsch, 2006, p. 351). Like the authors, we also “take the natural log of the number of refugees after adding one to the base... and include a measure weighting the total number of refugees from neighboring states relative to the size of the host countries” (Salehyan and Gleditsch, 2006, p. 351).

With respect to control variables, we include all variables used by the authors. Our sources of data differ only when the datasets used by Salehyan and Gleditsch (2006) do not cover the 2002-2013 additional years. Polity scores⁶, indicators for the presence of transborder ethnic groups⁷ and, separately, of civil war in neighboring countries, and country distances are extracted from the same data sources and processed as described by Salehyan and Gleditsch (2006) in their paper. For gross domestic product, we use World Bank data (The World Bank, 2017) because the data that the authors provide does not cover as many contemporary years. Similarly, we use World Bank data for population data.⁸ We calculate size of the largest ethnic group using ethnic power relations data (Wimmer, Cederman and Min, 2009) because the source of ethnicity data used by Salehyan and Gleditsch (2006) is not available for our study period. We also use the R package `splines` to create a cubic smoothing spline with three interior knots for peace years.

Table S8: Logistic regression results of Conflict Onset on Refugee Population (Logged)

	<i>Dependent variable:</i>			
	Outcome: Civil War Onset			
	Cold War M1	Cold War M2	Post-Cold War M1	Post-Cold War M2
Refugees (logged)	.10 (.10)	.10 (.09)	-.03 (.04)	-.04 (.03)
Civil War in Neighbor	.60*** (.20)	.40** (.20)	.40 (.40)	.60 (.40)
Polity	.02 (.02)	.01 (.01)	0.00 (.03)	.01 (.03)
Polity Squared	-.01*** (0.00)	-.01*** (0.00)	-.01* (.01)	-.02*** (.01)
GDP (logged)	-.07 (.10)	-.10 (.10)	-.10 (.10)	-.20 (.10)
Population (logged)	.20*** (.06)	.20*** (.05)	.50*** (.10)	.40*** (.10)
Ethnic Heterogeneity	.01* (.01)	.01** (.01)	-.70 (.70)	-.80 (.60)
Ethnic Kin	.70*** (.20)		1.00** (.50)	
Peace Years	-.50*** (.08)	-.50*** (.07)	-0.00 (.05)	.02 (.04)
Spline 1	-.01*** (0.00)	-.01*** (0.00)	-3.00*** (.90)	-2.00*** (.70)
Spline 2	.01*** (0.00)	0.00*** (0.00)	-2.00 (2.00)	-4.00*** (1.00)
Spline 3	-0.00*** (0.00)	-0.00*** (0.00)	-10.00*** (2.00)	-11.00*** (1.00)
Constant	-3.00** (1.00)	-3.00** (1.00)	-7.00*** (2.00)	-5.00*** (2.00)
Observations	2,880	4,268	1,735	2,920
Log Likelihood	-480.00	-578.00	-147.00	-218.00
Akaike Inf. Crit.	986.00	1,180.00	321.00	461.00

Note:

*p<0.1; **p<0.05; ***p<0.01

This table shows the regression results of the Salehyan and Gleditsch (2006) reanalysis and extension exercise comparing their analysis during the Cold War and post-Cold War periods. The first and second columns show Cold War results with and without Ethnic Kin, respectively, replicating Salehyan and Gleditsch (2006) using their data from 1951 - 1989, excluding 1990 and onwards. The third and fourth columns show post-Cold War results with Ethnic Kin and without Ethnic Kin, respectively, during the post-Cold War period for 1990 - 2013. This table shows that the positive effect of refugees on Conflict Onset in Salehyan and Gleditsch (2006) appears to be driven by (country-year) observations during the Cold War period. When their analysis is carried out for the current post-Cold War period, the coefficient for refugees becomes negative, although it is not statistically significant.

⁶Although we use the same data source, the most current version of that data has been modified since Salehyan and Gleditsch (2006) used it in their study. In their analysis, they convert “special transition codes” to 0 in their analysis. Since then, however, Marshall and Jagers (2002) has since taken steps to deal more carefully with these codes. We, therefore, use the “Polity2” variable, which “modifies the combined annual POLITY score by applying a simple treatment, or ‘fix,’ to convert instances of ‘standardized authority scores’ (i.e., 66, 77, and 88) to conventional polity scores...”(Marshall, Jagers and Gurr, 2015, p. 8).

⁷Given the number of missing observations on this variable, we follow Salehyan and Gleditsch (2006) in estimating the models with and without this variable.

⁸This may or may not be the source of population data used by Salehyan and Gleditsch (2006). The source is not specified in their manuscript.

Results of this analysis are consistent with our sub-national analysis, as shown in S7, potentially suggesting that conflict dynamics between Salehyan and Gleditsch (2006)'s study period and the more contemporary period are different. While an analysis of the differences between those periods is beyond the scope of this paper, we offer a few thoughts and some exploratory analysis to guide potential future research into heterogeneous treatment effects of refugee communities across time. Noting that one difference between Salehyan and Gleditsch (2006)'s study period and ours is that their study captures many Cold-War years, we carry out an additional analysis whereby we separately replicate their study for the Cold War and contemporary periods.⁹¹⁰

Table S8 shows the regression results of the second analysis. Comparing columns 1 and 2 which cover the Cold War period with columns 3 and 4 which cover the post-Cold War period, the positive relationship between refugees and conflict onset in Salehyan and Gleditsch (2006) is constrained to the Cold War period. The coefficient of (logged) refugee population from neighboring countries remains positive when the analysis is carried out using only Cold War years. For the post-Cold War period, the estimated coefficient is negative. Yet, neither set of coefficients is statistically significant. Because Salehyan and Gleditsch (2006)'s coefficient loses statistical significance when truncated, this may suggest that, rather than Cold War dynamics accounting (at least entirely) for the positive refugee-conflict relationship that they observe, there may be something specifically related to refugee outflows in the 1990s (perhaps the various displacement dynamics that have been documented as the Soviet Union collapsed and regional conflict intensified).

⁹For the Cold War analysis, this simply involves replicating their analysis using the subset of their original dataset covering only the Cold War period (1951-1989). For the contemporary analysis, we extend the replication carried out above to run from 1990 through 2013.

¹⁰Possible differences in Cold War, post-Cold War dynamics that might affect the relationship between refugees and conflict emerged during discussions with UNHCR officials. One noted, for instance, that our study that “[t]he Cold War period was a very distinct period because in Africa, there were a lot of freedom fighters.” (Senior Official at Kakuma Refugee Camp, interview conducted on July 23, 2018). Another mentioned that “[t]heres been quite an evolution since the 1980s, refugees camps in Pakistan were used as military bases. Cambodian camps were used as bases to oppose Cambodian occupation. Since that time, that form of politically motivated policy is no longer standard. UNHCR in particular has insisted as much as possible to maintain the civilian characteristic of these camps.” (Senior Official at Policy Development and Evaluation, interview conducted on July 6, 2018).

References

- Dittrich Hallberg, Johan. 2012. "PRIO Conflict Site 1989–2008: A Geo-Referenced Dataset on Armed Conflict." *Conflict Management and Peace Science* 29(2):219–232.
- Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.
- Imai, Kosuke, In Song Kim and Erik Wang. 2018a. Matching methods for causal inference with time-series cross-section data. Technical report.
- Imai, Kosuke, In Song Kim and Erik Wang. 2018b. PanelMatch for nonparametric generalization of the difference-in-differences estimator. Technical report.
- Kikuta, Kyosuke. 2020. "A new geography of civil war: a machine learning approach to measuring the zones of armed conflicts." *Political Science Research and Methods* pp. 1–19.
- Marshall, Monty G and Keith Jagers. 2002. "Polity IV Project: Political Regime Characteristics and Ransitions, 1800–2002.".
- Marshall, Monty G, Keith Jagers and Ted Robert Gurr. 2015. "Polity IV Project: Dataset Users' Manual." *College Park: University of Maryland* .
- Morgan, Stephen L and Christopher Winship. 2014. *Counterfactuals and Causal Inference*. Cambridge University Press.
- Pettersson, Therése and Magnus Öberg. 2020. "Organized violence, 1989–2019." *Journal of peace research* 57(4):597–613.
- Salehyan, Idean and Kristian Skrede Gleditsch. 2006. "Refugees and the Spread of Civil War." *International Organization* 60(2):335.
- Shaver, Andrew, David B Carter and Tsering Wangyal Shawa. 2016. "Terrain Ruggedness and Land Cover: Improved Data for Most Research Designs." *Conflict Management and Peace Science* .
- Sundberg, Ralph and Erik Melander. 2013. "Introducing the UCDP georeferenced event dataset." *Journal of Peace Research* 50(4):523–532.
- The World Bank. 2017. "World Development Indicators.".
- Tollefsen, Andreas, Havard Strand and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* 49(2):363–374.
- Weidmann, Nils B, Doreen Kuse and Kristian Skrede Gleditsch. 2010. "The Geography of the International System: The CShapes Dataset." *International Interactions* 36(1):86–106.
- Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set." *American Sociological Review* 74(2):316–337.