Reproduction of Warming Increases the Risk of Civil War in Africa, Burke et al. (2009)*

Ethan Ignoffo, Myles Berry, Andrew Barker, Joseph Selfani, Andrew Song

April 16th, 2025

Abstract

Burke et al. (2009) examined the effects of both environmental and non-environmental factors on civil war incidence across 43 African countries between the years of 1981 and 2002. Temperature was used as the main independent variable in the original study, and was chosen due to the fact that prior research suggested higher temperatures ultimately harm a country's economic productivity in most lower income and agriculture dependent nations. Additional variables were added to the models to help explore the relationships between climate change and social discontent. Burke et al. finds that a 1°C increase in temperature leads to a 4.5 percentage point increase in civil war incidence in the same year. In this study, we directly replicate Burke et al. (2009), using the original datasets and statistical methods. However, we exclude the supplementary tables in the original study that involve alternative model specifications and variable transformations. We find that the original constant term estimates are not reproducible (which is most likely due to undocumented subset selection methods), the key estimates for temperature are confirmed for sign, magnitude, and statistical significance.

Keywords: Climate change, Civil war

*Authors: Ignoffo: University of California, Berkeley. E-mail: ethanignoffo@berkeley.edu. Berry: University of California, Berleley. E-mail: mylesberry@berkeley.edu. Barker: University of California, Berkeley. E-mail: ajbarker@berkeley.edu. Selfani: University of California, Berkeley. E-mail: selfanijo@berkeley.edu. Song: University of California, Berkeley. E-mail: andrewsong@berkeley.edu. For the replication of this study, we received no financial support and have no conflict of interest. All authors in this replication are not affiliated with the original authors of the study.

Team Participation:

Member Name	**Participation**	**Reason (If necessary)**
Ethan Ignoffo	20%	N/A
Myles Berry	20%	N/A
Andrew Barker	20%	N/A
Joseph Selfani	20%	N/A
Andrew Song	20%	N/A

1 Introduction

Burke et al. (2009)'s original study examines how temperature changes affect the risk of civil war incidence in Sub-Saharan Africa across 43 countries in the years 1981 to 2002. The original research analyzes how these temperature changes (and various other factors such as rainfall) influence the probability of a civil war happening. The study's hypothesis is that warming temperatures increases the risk of civil war breaking out, independent of economic or political factors. Their analysis incorporates conflict data from the Oslo International Peace Research Institute, Norway, & the University of Uppsala, Sweden (PRIO/Uppsala) database, as well as climate and precipitation data from the Climatic Research Unit of the University of East Anglia (CRU), Global Precipitation Climatology Project of NASA's Goddard Space Center (GPCP), and the National Center for Environmental Prediction/National Center for Atmospheric Research (NCC).

In this report, we seek to replicate their empirical models by using the dataset *climate_conflict.dta* provided in the authors' replication package. Our replication focuses on three progressively augmented models of civil war onset (war_prio_new) regressed on climate variables. We begin with a simple model of temperature (temp_all) and its one-year lag, including country-year fixed effects. We then add precipitation (prec_all) and its lag, and finally incorporate lagged GDP per capita (gdp_l) and political regime score (polity2_lag), following the authors' extended specification. Our estimation uses OLS with standard errors clustered by country, consistent with the original study.

Our empirical strategy reproduces their structure using Pandas and Python, and we present side-by-side results using Stargazer to highlight coefficient estimates on key climatic variables. We find that the estimated coefficients on temperature and its lag remain statistically significant and positive, corroborating the central finding of Burke et al. (2009). However, the magnitude and precision of these estimates vary across specifications and are sensitive to the inclusion of additional covariates and how standard errors are clustered. We record these effects within our regression tables, and discuss the implications that results have for both statistical significance and policy applications.

2 Reproduction

In the next section, we describe our computational reproduction of Burke et al. (2009) study. We replicate this study by using the same data and methods, but we implement the replication in Python instead of Stata and R. Specifically, we focus our replication on the main regression coefficients, tables, and figures found in the main text of the original study. Secondly, we do not conduct a sensitivity analysis of Burke et al. (2009), as the main focus of our replication is the verification of the original findings, and not on testing for model robustness. While we were able to reproduce the original coefficients with success, we found small variations in the constant terms from the original models. This variation is further described in this section.

2.1 Regression models: Table 1

For our analysis, we run the same three OLS regressions as in Table 1 of Burke et al. (2009). Despite the original authors running their regressions in Stata, we conduct our modeling in Python. We regress civil war incidence on temperature and lagged temperature, adding controls for model 2 and 3. The inclusion of lagged variables is to show whether conflict risk is immediate or delayed; but also to isolate the effect of current year shocks rather than long year trends. All models use country fixed-effects; however, only Models 1 and 2 include country time trends to control for time-varying country characteristics. The specifications are as follows:

Model 1:

$$Y_{it} = \alpha + \beta_1 Temperature_{it} + \beta_2 LaggedTemperature_{it} + \zeta X_t + \delta Z_{it} + \epsilon_{it} \quad (1)$$

Where Y is the incidence of the Civil War, the i and t subscripts correspond to the country and year respectively, *Temperature* is the average annual temperature and $LaggedTemperature$ is the average annual temperature from the previous year, X_t are country fixed effects and Z_{it} are country time trends. This corresponds with column 1 in Table 1 of the original study.

Model 2:

$$Y_{it} = \alpha + \beta_1 Temperature_{it} + \beta_2 LaggedTemperature_{it} + \\ \beta_3 Precipitation_{it} + \beta_4 LaggedPrecipitation_{it} + \zeta X_t + \delta Z_{it} + \epsilon_{it} \quad (2)$$

Model 2 extends the previous model by adding climate controls such as average annual precipitation (*Precipitation*) and its one year lag (*LaggedPrecipitation*). This corresponds to column 2 in Table 1 from the original study.

Model 3:

$$Y_{it} = \alpha + \beta_1 Temperature_{it} + \beta_2 LaggedTemperature_{it} + \\ \beta_3 Precipitation_{it} + \beta_4 LaggedPrecipitation_{it} + \beta_5 LaggedPerCapitaIncome_{it} \\ + \beta_6 LaggedPoliticalRegime_{it} + \zeta X_t + \epsilon_{it} \quad (3)$$

Model 3 additionally includes the climate controls as well as one year lagged GDP per capita (*LaggedPerCapitaInc*) and one year lagged democracy (LaggedPolitical-Regime), which captures the potential effects of the country's political and economic situation on civil conflict. This corresponds to column 3 in Table 1 from the original study.

The coefficient of *Temperature* is β_1, which we estimate to be **0.045** in Model 1, **0.043** in Model 2 and **0.049** in Model 3, suggesting a 4.3–4.9% increase in Civil War incidence from each additional 1°C temperature increase. Our estimates for β_1 are statistically significant at the 5% level in Models 1 & 2 and at the 10% level in Model 3, suggesting robustness of this effect with additional controls. The coefficient of

Lagged Temperature is β_2, which we estimate to be **0.0087** in Model 1, **0.0132** in Model 2 and **0.0206** in Model 3, suggesting around a 0.9% to 2% increase in Civil War incidence from each additional 1°C temperature increase in the previous year. However, our estimates of β_2 were not statistically significant at the 5% or 10% level in Models 1, 2 or 3, suggesting that civil conflict is driven more by current temperature effects, than delayed effects of temperature changes.

The coefficient of *Precipitation* is β_3, which we estimate to be **-0.0230** in Model 2 and **0.0165** in Model 3, and the coefficient of *Lagged Precipitation*, β_4 we estimated to be **0.0250** in Model 2 and **0.0278** in Model 3. However, our estimates for precipitation controls were not statistically significant at the 5% or 10% level in Models 2 or 3, suggesting that precipitation may not have a real effect on civil war incidence. This seems to align with the original paper that current temperature, rather than precipitation, is the driver of civil conflicts.

In Model 3, the coefficient of *LaggedPerCapita*, β_5, we estimated to be **-0.00002663**, and the coefficient of *LaggedPoliticalRegime*, β_6 we estimated to be **-0.0005**, suggesting a negligible effect of lagged economic or political scores on civil war conflicts. And, we notice that none of these were statistically significant.

In our regression analysis we encountered a few challenges and found some discrepancies with the original findings. Firstly, the coefficient estimates that we replicate in Table 1 show slight differences from those in the original study. This is due to rounding that occurred when creating the regression table from the models. Secondly, the constant estimates in the regression models differ from the original study. We believe that this might be occurring because of unspecified subset selection methods that occurred in the original analysis. More specifically, we believe that the original authors may be dropping a year from the dataset, possibly arbitrarily, which may explain the difference in the constant estimates. However, original Stata code specifies the regression models and does not elaborate on any subsets or dropped variables that were used in those models. Finally, the Residual Mean Square Error (RMSE) values that are in the original Table 1 were not reproduced (although our R^2 scores were almost identical). These values could not be replicated through the original Stata code, as the original authors did not specify how these values were calculated or what packages were used to conduct these calculations. Additionally, as our R^2 values were nearly identical, the omission of these RMSE values was not deemed an issue as they are just additional indicators of model fit.

But despite these minor observed differences in the constants and RMSE values, the economic interpretation of the results remains the same. In all of our replicated models, the current year's temperature had a positive and significant impact on civil war incidence, while precipitation, lagged variables, and political-economic factors remained insignificant. This is consistent with the original study's finding that precipitation, lagged variables, and economic and political causes were not the driver of civil war conflicts. Civil war incidence only seems to increase in the warmer years, and this effect is robust and higher in magnitude, even after considering country's fixed effects, time trends, and additional controls like precipitation and socio-economic factors.

2.2 Emissions scenarios: Table 2

In order to complete Table 2 we replicated the projected changes in civil war incidence through 2030. We used three emission scenarios (A1B, A2, B1), following the original paper's approach representing a balanced emissions scenario, a high emissions/fragmented world scenario, and a low emissions scenario with an environmentally focused world. First, we had to create a weighting system that links countries to regions in Africa, so that we could normalize our data for varying land-size of regions in the climate data. We calculated the average monthly temperature (converted to yearly numbers) and historical average monthly rainfall (to get yearly predicted precipitation) data from a climate dataset derived from climate models and used in the paper (altogether.monthly in the code) across five regions in Sub Saharan Africa. Using the weights from the previous step we created a single set of temperature and rainfall projections for the whole continent that we used in the bootstrapped regression models to generate the conflict risk forecasts in table 2. Because economic and political conditions also impact the likelihood of civil war, we replicated the paper's efforts and used historical GDP and polity score data under an optimistic scenario and a constant historical growth scenario (which we later used in our regressions in Model 3 that feed into the bottom of figure 2). Next, we replicated the nine bootstrap models, sampling the data 10000 times for each, and ran a regression for each sample and saved the first six coefficients each time. The distribution of these coefficient estimates accounts for the uncertainty of the impact of climate changes on civil war incidence in the future, so we took the median values for Model 1 and 2 across the three scenarios, as well as a 5-95% confidence interval.

Using this data we replicated Table 2, but our results were slightly different than the actual paper. However, our results differ very slightly. Most likely these differences come from the bootstrapping process because they are so close to the actual results in the paper. Because bootstrapping is a sampling process, some variation in replication results is expected due to the random resampling of the data, version differences in the software, or undocumented processing steps from the original authors. These small differences are a natural consequence of the method and do not reflect substantive deviations from the original study's approach. We feel confident that given the similarity in results, there is no difference in the economic implications from Table 2. The impact of climate change is still positive and within the range suggested by the paper across both models and all three scenarios.

In the actual coding of the replication it was difficult to translate the R code used into python. The data cleaning process was not straightforward, and there wasn't clear enough direction. We also had difficulty with the bootstrapping process. One simple, but real, difficulty was the time required to run the bootstraps over and over again as we checked our results against the paper. In our bootstrap we also were a little more strict with dropping NaN values than in the R code in the paper (which focused on column 2 NaN values). Another area we address in our code is we set the seed for reproducibility, but we didn't see that in the paper's R code. We also weren't able to visualize Table 2 exactly like in the paper. Our table is grouped by Model 1 and Model 2, as opposed to by scenario (like in the paper). Overall, the difficulties in replicating Table 2 paled in comparison to Figure 1 and Figure 2.

2.3 Projected changes in climate & conflict: Figure 1

In order to replicate Figure 1 from the Burke et al. (2009) study, we needed to evaluate the relationship between the climate projections under the A1B emissions scenario that we modeled in our previous bootstrapping procedures and civil war risk in Sub-Saharan Africa. Figure 1 showcases two panels with one representing the climate projections and one representing the conflict projections.

We began with the left panel, overlaying two horizontal boxplots for each region's climate projections to the year 2030. The first boxplot showcases changes in precipitation (ΔP), that we calculated using the following formula:

$$\Delta P_r^{(j)} = \Big(\frac{\sum_{m=1}^{12} \big(1 + \frac{p_{r,m}^{(j)}}{100}\big) C_{r,m}}{\sum_{m=1}^{12} C_{r,m}} - 1\Big) \times 100,$$

In this formula, $p_{r,m}^{(j)}$ represents the percent change in monthly precipitation, and $C_{r,m}$ represents the baseline climatological rainfall. The second boxplot calculated with the following formula:

$$\Delta T_r^{(j)} = \frac{1}{12} \sum_{m=1}^{12} t_{r,m}^{(j)},$$

shows temperature changes, identified by $10\,\Delta T$, where the monthly temperature anomalies ($t_{r,m}^{(j)}$) in °C for each draw (j) were averaged and scaled by 10 in order to produce an axis range of $0.0 - 2.0$°C. We then recreated the inset boxplots, by calculating the median (the dark line), the inter-quartile range (colored box), and the $5 - 95\%$ whiskers. The inset positions corresponded to the regional centroids, with an extra row added for "All-SSA" or the unweighted averages across r=1,. . . ,5

We then focused on replicating the right panel of Figure 1, which visualizes the conflict projections to 2030. In order to replicate this panel, we let $\beta^{(b)}$, $b =$

$1, \ldots, 10000$ act as the bootstrap draws of Model 1's coefficient vectors $(\beta_T, \beta_{\ell T}, \beta_P, \beta_{\ell P}, \ldots)$. For each region r and climate draw j, we defined the projected civil-war change as:

$$\Delta Y_{r,j,b} = \beta_T^{(b)}\, \Delta T_r^{(j)} + \beta_P^{(b)}\, \Delta P_r^{(j)}.$$

Then, we formed three distributions of ΔY (in percentage-point changes):

1. *Climate-only uncertainty:* set $\beta^{(b)} \equiv \overline{\beta}$ (the median coefficients) and vary j.

2. *Model-only uncertainty:* set $\Delta T, \Delta P \equiv \overline{\Delta T}, \overline{\Delta P}$ (the median climate draws) and vary b.

3. *Joint uncertainty:* vary both j and b.

For each region, the boxplots created (1, 2, & 3), display the calculated median, interquartile range, and $5 - 95\%$ whiskers of the resulting distributions $\{\Delta Y_{r,j,b}\}$. Then to mimic the original study, we decomposed the uncertainty in our conflict projections into three categories: climate-only uncertainty, model-only uncertainty, and joint-uncertainty which were all represented by different boxplots.

While recreating Figure 1, we struggled to accurately capture the uncertainty in both our climate and conflict response models. We believe that this is due to the fact that the climate projections are subject to inherent variability, as the different models produce different predictions for temperature and precipitation. Similarly, we believe that the conflict response model comes with its own variability with how the climate variables impact the civil war risk. This required us to try and replicate the bootstrapping referenced in the previous section as closely as possible in order to attempt to produce similar results. Additionally, we faced another challenge with identifying the exact methods that the original authors used to create this figure. Both the climate and conflict models rely heavily on the data and subsets used, therefore we faced inherent difficulty replicating this figure as these subset methods were not accurately documented by the original authors.

Even though we followed the original methods as closely as possible, we ultimately noticed a degree of variability within our conflict and climate projections.

This was expected due to struggles we faced while conducting this replication. As previously mentioned, we believe that this is due to the differences in sampling and bootstrapping that occurred in previous steps. This resulted in our boxplots on both panels of our Figure 1 showing more variation than we initially expected. Economically, our results in replicating Figure 1 add to the original study's findings and further the understanding of how climate change may influence conflict risk in vulnerable regions of the world such as Sub-Saharan Africa. Our results indicate that both temperature and precipitation changes can significantly affect the risk of civil war incidence, and the uncertainty in the climate projections further complicates predicting these future conflict outcomes.

2.4 Impact of climate change: Figure 2

Lastly, we focused on replicating Burke et al. (2009)'s Figure 2. Figure 2 showcases a comparison of the projected change in civil war risk under two specific categories of uncertainty: (1) the climate-only models, and (2) the combined climate-economic-political models. In order to recreate this figure, we included Models 1 & 2 from our replication of Table 1, which included both the temperature-only and the temperature + precipitation models. We also included Models 3 through 6 from our bootstrap models that accounted for variations such as first differences, deviations from long-term trends, alternative climate datasets (e.g., CRU temperature and GPCP precipitation for Models 3 - 5), and NCC data (Model 6). Our combined models extended Model 3 by adding projections for income and democracy. In particular, we included two scenarios that were part of the original study's Figure 2: one where income and polity scores followed a linear extrapolation based on 1981-2002 trends, and another which represented an optimistic growth path.

In order to replicate this in Python, we loaded our nine bootstrap arrays that we had saved as (`bootstrap_1.Rdata`–`bootstrap_9.Rdata`), into a list structure saved as `hold[1:9]`. For the first six models, which exclude income and democracy variables, we chose to zero out those variable coefficients. Additionally, for Models

1 & 3 we zeroed out the precipitation coefficients where necessary. For the final two entries, we used Model 4's base coefficients, and readded the income and democracy effects. We then multiplied the income coefficients by either the historical or optimistic GDP growth rates ('incchg[1]' or 'incchg[0]'), or the mean polity change ('polchg'). Then, for each model i and each of the 20 climate $\Delta T/\Delta P$ draws, we computed $10,000$ predictions defined with the equation:

$$\Delta \mathrm{War}_{ij} = \Delta T_j \big(\beta_T + \beta_{T,-1}\big) + \Delta P_j \big(\beta_P + \beta_{P,-1}\big) + \mathrm{inc}_i \, \beta_{\mathrm{inc}} + \mathrm{polchg} \, \beta_{\mathrm{pol}} \, .$$

We then pooled our 200,000 estimates per model, sorted through them, and trimmed to capture the central 90th percentile, which allowed us to create eight final distributions. These eight distributions were then used to generate the horizontal boxplots in our replicated Figure 2. To further mimic the original study, we reversed the plot order that our climate-only models appeared at the top, and added further plot elements such as the vertical dashed line at zero, shaded the two model groups differently for visual clarity, and replicated the labels from the original graph.

Ultimately, our replication of Figure 2 produced results that were extremely similar to those found in the original study. Our climate-only models clustered between 3 and 8 percentage-point increases in civil war incidence, while our combined income and democracy scenarios spanned a wider range from 0 to 15 percentage points. These results confirm that our reproduction accurately reproduces Burke et al. (2009)'s original figure, and validate the robustness of the original study's findings across the various data sources. Economically, this replication reinforces the original conclusion that climate variables alone already predict a notable increase in civil war incidence across Sub-Saharan Africa.

3 Conclusion

Overall, our reproduction of Burke et al. (2009) was successful but not without a few challenges. The large amount of datasets, limited specificity in methods, and required translation of the original code made the replication process quite complex. As the original study was conducted through Stata and R, we needed to identify comparable Python packages in order to conduct the original methods as closely as possible. Luckily, the use of Python documentation, package-specific resources, and ChatGPT, made it possible to find appropriate tools to conduct our replication. We found that the OLS models were the most straightforward to reproduce, while the bootstrapping procedure and figure replication posed the greatest challenges.

Despite these challenges, our results aligned closely with those of the original study: a 1°C increase in annual temperature is associated with a 4.5 percentage point increase in the risk of civil war in the same year. The primary differences between our replication and Burke et al. (2009) all stem from the results of our bootstrapping procedure. These differences we believe are most likely due to sampling variability, differences between software versions, or undocumented processing steps conducted in the original analysis. Overall, we do not believe that these discrepancies change the interpretation of the original study's findings. We believe that they reinforce the robustness of the central conclusions, that rising temperatures increase the likelihood of civil conflict in Sub-Saharan Africa.

References

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

4 Figures

Figure 1 replication

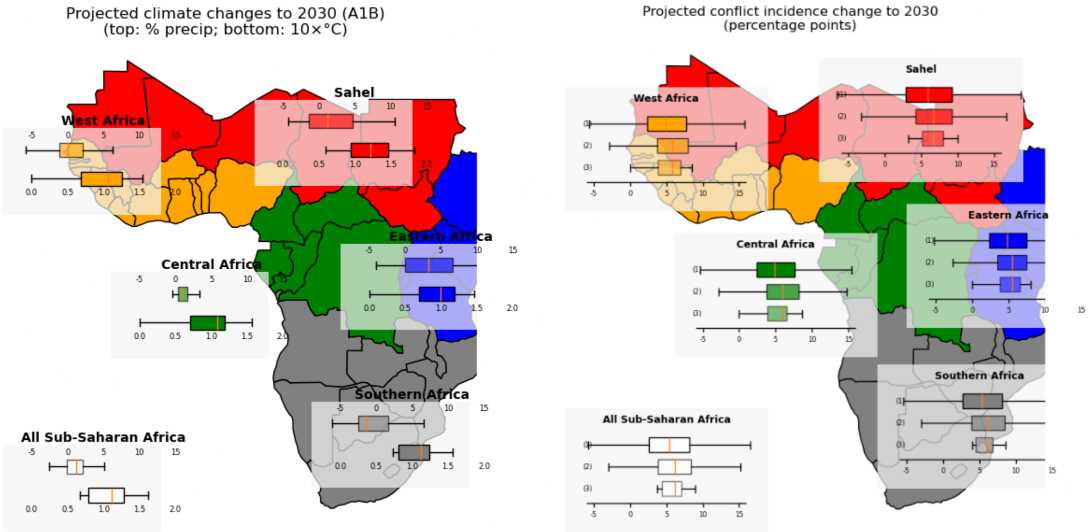
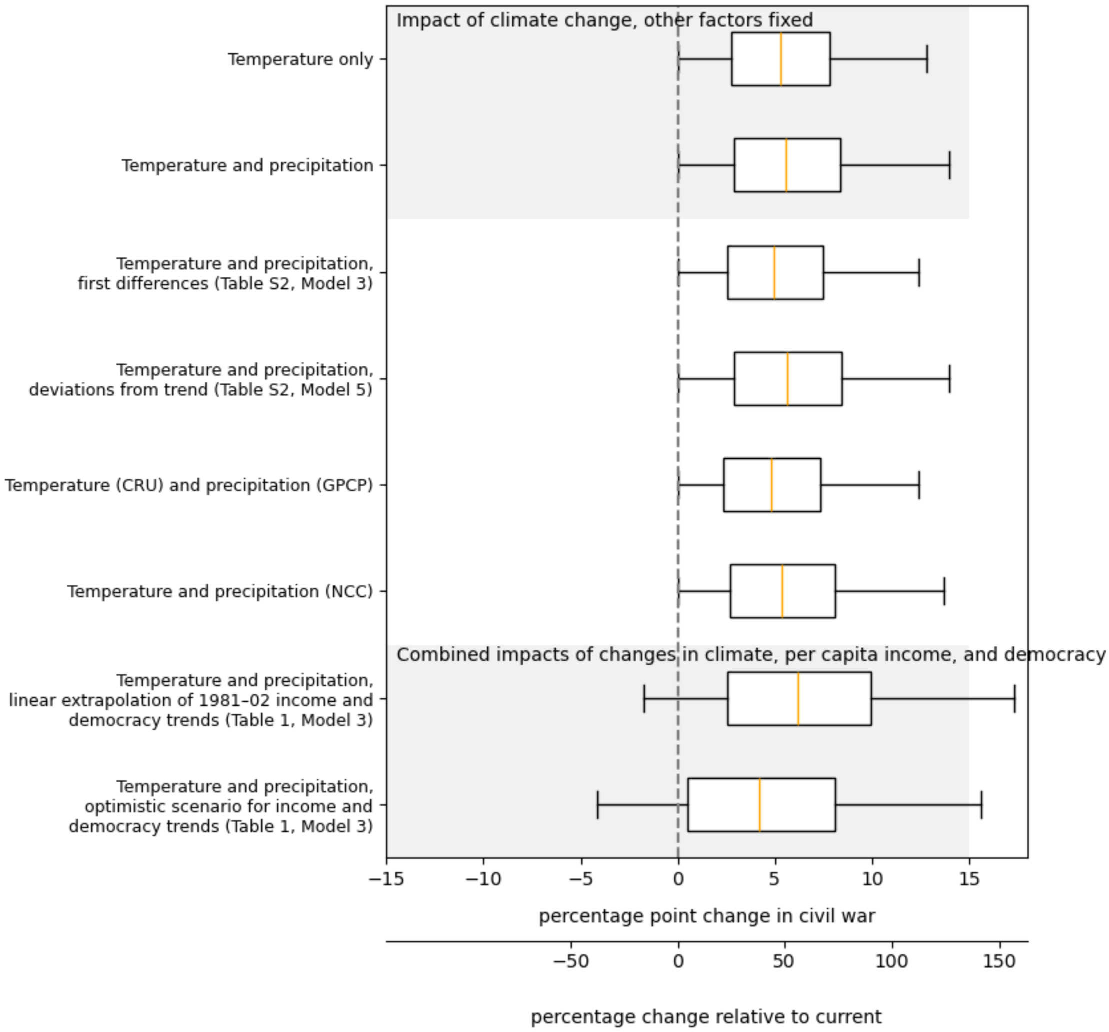


Figure 2 replication



5 Tables

Table 1 replication

Table 1. Regression coefficients on climate variables, with civil war as a dependent variable

	Dependent variable: war_prio_new		
	Model (1)	Model (2)	Model (3)
Temperature	0.045**	0.043**	0.049*
	(0.022)	(0.022)	(0.029)
Temperature lagged 1 year	0.009	0.013	0.021
	(0.021)	(0.023)	(0.030)
Precipitation		-0.023	0.016
		(0.052)	(0.075)
Precipitation lagged 1 year		0.025	0.028
		(0.049)	(0.074)
Per capita income lagged 1 year			-0.027
			(0.021)
Political regime type lagged 1 year			-0.001
			(0.002)
Constant	-1.110	-1.177	-1.521
	(0.769)	(0.772)	(0.983)
Observations	889	889	815
R^2	0.657	0.657	0.389
Adjusted R^2	0.622	0.621	0.354
Residual Std. Error	0.193 (df=805)	0.193 (df=803)	0.241 (df=770)
F Statistic	34.677*** (df=83; 805)	19.457*** (df=85; 803)	11.124*** (df=44; 770)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2 replication

```
Final Table 2:
```

Scenario	Model	Median % change	% increase relative to baseline	5th–95th percentile	% obs < 0
A1B	**Model 1**	5.9	53.1	5.0–118.8	3.5
A2	**Model 1**	5.2	46.8	4.4–101.4	3.5
B1	**Model 1**	4.7	42.9	4.0–98.9	3.5
A1B	**Model 2**	6.2	56.2	3.4–129.6	3.9
A2	**Model 2**	5.5	49.6	2.9–110.7	4.0
B1	**Model 2**	5.0	45.6	2.6–107.8	4.0