Climate-conflict in Sub-Sahara Africa: Examining predictive power*

Stijn van Weezel
Department of Economics, Royal Holloway
University of London
pwte054@rhul.ac.uk

First draft: 11 November 2014
This draft: 6 October 2015

Abstract

Within the quantitative literature on violent armed conflict there is a heated debate on the effect of climate change on the incidence of civil conflict. While this discussion predominantly focuses on the robustness of results to different model specifications and extended datasets, generated predictions are largely ignored and not included in the analysis. This is surprising given the relevance of predictive power in the field of conflict studies and concerning climate change forecasts. This study illustrates that examination of a model's predictive power can be a useful analytical tool to scrutinise a model's performance. Focussing on the link between temperature and conflict in Sub-Saharan Africa between 1981-2001 I find that although the regression model has large in-sample predictive power, the performance of the temperature variable is low. Generating out-of-sample predictions for 2003-2013 illustrates that temperature has almost no predictive power.

JEL-Classification: D74, O55, Q54

*Many thanks to Burke et al. for making their data publicly available, and Michael Spagat and Halvard Buhaug for helpful comments. This draft version presents preliminary results.

Keywords: civil conflict, Africa, climate variability, statistical models

1 Introduction

There is an expanding literature that attempts to disentangle the factors that contribute to the incidence of violent armed conflict. Within this research area much attention is given to the robustness of results in terms of replicability, subject to the use of more recent data and different model specifications. However, surprisingly little attention is given to testing and analysing the predictive power of the models. Although for some scientific research having a model with good descriptive power is sufficient, considering conflict predictive power is particularly relevant given the damaging consequences of war and its long lasting effects (Abadie and Gardeazabal, 2003; Chamarbagwala and Morán, 2011; Serneels and Verpoorten, 2013). It is therefore remarkable that within this body of research seemingly so little effort is made at prediction (see Schrodt (2013) for a comprehensive critique).

Despite the increase in research effort, which has lead to a general understanding of some of the main factors that underlie conflict (Hegre and Sambanis, 2006; Blattman and Miguel, 2010), there seems to be no record of improvement in the accuracy of predictions in the conflict literature. As Ward et al. (2010) observe:

> too much attention has been paid to finding statistically significant relationships, while too little attention has been paid to finding variables that improve our ability to predict civil wars

Predictions generated by models are often ignored while they provide useful information on the model's performance as shown in O'Loughlin et al. (2012) for instance. There are some other notable exceptions to this trend, such as the recent studies by Goldstone et al. (2010), Weidmann and Ward (2010), Gleditsch and Ward (2013), and Blair et al. (2014).

This study contributes to this literature by attempting to predict war focussing on the link between climate and conflict. There is a growing body of research examining how climate influences economics outcomes and affects human well-being (see Dell et al. (2014) for an overview). There are large concerns about the human impacts of global warming and it is therefore important to understand the relation in order to assess the risks (Stern, 2006; German Advisory Council on Global Change, 2008). I focus on this particular field as it is of great interest due to climate change forecasts (IPCC, 2014) and because of the potential climate-conflict nexus. I say potential here because within the academic community there is an ongoing debate on the scientific evidence for a link between climate change and the risk of violent armed conflict (see for instance Scheffran et al. (2012a); Hsiang et al. (2013); Hsiang and Burke (2013); Klomp and Bulte (2013); Buhaug et al. (2014); Hsiang et al. (2014); Hsiang and Burke (2014); Raleigh et al. (2014)).[1]

From this literature I use the model from the seminal work by Burke et al. (2009) who found that higher temperatures are linked to increased probabilities of war in Africa. Simulation based predictions of their model show

[1]Edward Miguel, one of the leading scientists in this field of research, likened the research on climate-conflicts to the research on smoking about 50 years ago, arguing that we figured out that climate causes conflict but don't know how to prevent it (TEDxBerkeley, 2014).This analogy seems a bit premature given the lack of consensus in the literature. Other overview articles, not mentioned in the text, include Bernauer et al. (2012), Gleditsch (2012),Scheffran et al. (2012b), Meierding (2013).

that, conditional on historical patterns, by 2030 there will be a 54% increase in armed conflict incidence relative to the 1981-2002 baseline. This increase could even be an underestimate according to a more recent study (Burke et al., 2014).

The reason for focussing on this particular model is because of its central place in the literature as well as the simplicity of the model, which should be beneficial to generating accurate predictions.[2] There has been some debate with regard to the robustness of the results, where the critiques have predominantly focussed on sensitivity to model specification and re-estimation using more recent data (Buhaug et al., 2010; Sutton et al., 2010; Burke et al., 2010; Hsiang and Meng, 2014). The reply from Buhaug et al. (2010) seems to be the most forceful critique on the original study, claiming that climate variability is a poor predictor of armed conflict and that African civil wars can be explained by other more structural and contextual factors such poor economic performance for instance.

Hsiang and Meng (2014) disagree with this conclusion arguing that there is no statistical difference between the model specifications used by Burke et al. (2009) and Buhaug et al. (2010).[3]

To the best of my knowledge, no study has attempted to use the model for generating predictions for real out-of-sample data. This study is probably most similar to Buhaug et al. (2010), who use the original data of Burke et al., shorten the time period for model estimation to 1981-1998 and generate

[2]Their research has been cited 385 times since its publication in 2009 (according to Google Scholar, 18 April 2015).

[3]Hsiang and Meng (2014) found that Buhaug's results are not different from those in Burke et al. work based on the implementation of various statistical test, but do not attempt to verify whether the findings of Burke et al. were correct. See O'Loughlin et al. (2014b) for a critique on the Hsiang and Meng (2014) study and Buhaug (2014) for Buhaug's reply.

predictions for 1999-2002. In contrast, using the original model estimations I generate predicted probabilities for conflict using the most recent available data covering 2003-2013 which is a more extensive cross-validation of the model.[4]

The regression analysis shows that the model seems to do exceptionally well for in-sample prediction and in general has high descriptive power based on an Area Under the Curve (*AUC*) value of around 0.98.[5] This result seems to be an overfit however:

Examining the results in closer detail I find that the predicted values are almost entirely contingent on the inclusion of the country fixed effects and country-specific time trends.

For the out-of-sample predictions the model produces a large number of false positives (for countries with high levels of conflict prevalence), while at the same time failing to correctly predict war incidence in countries with slightly lower levels of past conflict prevalence. It seems that the in-sample performance on the training data exaggerates minor fluctuations in the data where the results do not generalise beyond the cases studied. Both the regression and prediction results show that the performance of the model is mainly driven by the inclusion of country fixed effects and country-specific time trends, whereas the inclusion of variables capturing climate variability adds little to no predictive power.

[4]This study thus provides a replication of the original work and an extension according to the proposed terminology of Clemens (2015).

[5]This is the *AUC* of the Receiver/Operator Curve (ROC) which measures the false positive rate versus the true positive rate. *AUC* values closer to 1 indicate a more accurate model. The *AUC* values for the three models are: model 1, 0.981; model 2, 0.981; model 3, 0.936. Omitting all explanatory variables and estimating the LPM with only country fixed effects and a country-specific time trend, the results still report a high *AUC* statistic of 0.980 which is a reduction in predictive power of just 0.09% compared to the preferred model (model 1).

2 Results

2.1 In-sample predictions

I start with replicating the results in order to check the in-sample predictive power of the model. As I try to argue in this paper that cross-validation should be part of the analysis, it could of course be that for some reason this isn't possible due to lack of data. When data is sparse one can always examine the in-sample predictive power of the model as I illustrate in this section. In most quantitative studies there is a lot of focus on the statistical significance of the estimated parameters based on the traditional p-value thresholds. This litmus test, so to speak, was never meant to be a definitive test when it was introduced by Fisher but intended as an indicative test for evidence worth a second look (Nuzzo, 2014). The original data covers 41 countries in Sub-Saharan Africa for the period 1981-2002. The preferred model (model 1) has the following functional form:

$$War_{it} = \beta_1 Temperature_{it} + \beta_2 Temperature_{i,t-1} + c_i + d_i year_t + \epsilon_{it}$$

The regression equation links civil war to current and lagged temperature levels ($Temperature$), conditional on country fixed effects (c_i) and time trends ($d_i year_t$).[6]

[6] Additional models include measures for current and lagged rainfall levels (model 2), and explicit lagged country controls for income per capita and regime type (model 3). Model 3 also includes year indicators rather than a country-specific time trend. Model 1 and 2 do not include year indicators meaning that they do not account for worldwide changes (Couttenier and Soubeyran, 2014). Civil wars are conflicts with >1,000 battle-related

Although War is a binary outcome variable, the model is estimated using Ordinary Least Squares (OLS). It seems to make little sense fitting a continuous regression to an outcome variable that can only take two values (Gelman and Hill, 2007). However, due to the use of time-series cross sectional data and the inclusion of fixed effects, a Linear Probability Model (LPM) might be preferred over non-linear models (Angrist and Pischke, 2008).[7]

Nonetheless, there are some concerns with using a LPM such as heteroskedasticity due to the binary outcome variable and the fact that the predicted probabilities are not constrained by the 0-1 interval.[8] Figure 1 shows the predicted probabilities for each model and illustrates that there are indeed predicted values outside the unit interval. This might not be of too much concern as long as not too many values fall outside of the interval. However, in this case the results are a reason of concern:

Examination of the preferred model (model 1) shows that the predicted probabilities for 39% of the observations have a value smaller than 0 or larger than 1. As a result, the estimates can be biased and inconsistent (Horrace and Oaxaca, 2006).[9]

deaths, temperature is measured in Celsius.

[7]Beck (2011) shows that estimating a LPM is as good as logit as long as the explanatory variables are normally distributed. An additional concern, not addressed in this study, is that due to the inclusion of fixed effects there is the concern of Nickell bias (Nickell, 1981) with $N = 41$ and $T = 21$. I refer to Gaibulloev et al. (2014) for an extensive discussion on the inconsistencies arising from the use of fixed effects in dynamic panel data.

[8]The heteroskedasticity will be accounted for by the use robust standard errors clustered by country.

[9]Model 1 generates 15 values above 1 and 334 below 0. Because of the probability of biased and inconsistent estimates, the LPM should be used as a basis for comparison with more appropriate models such as logit. See the appendix for model estimations done using logit.

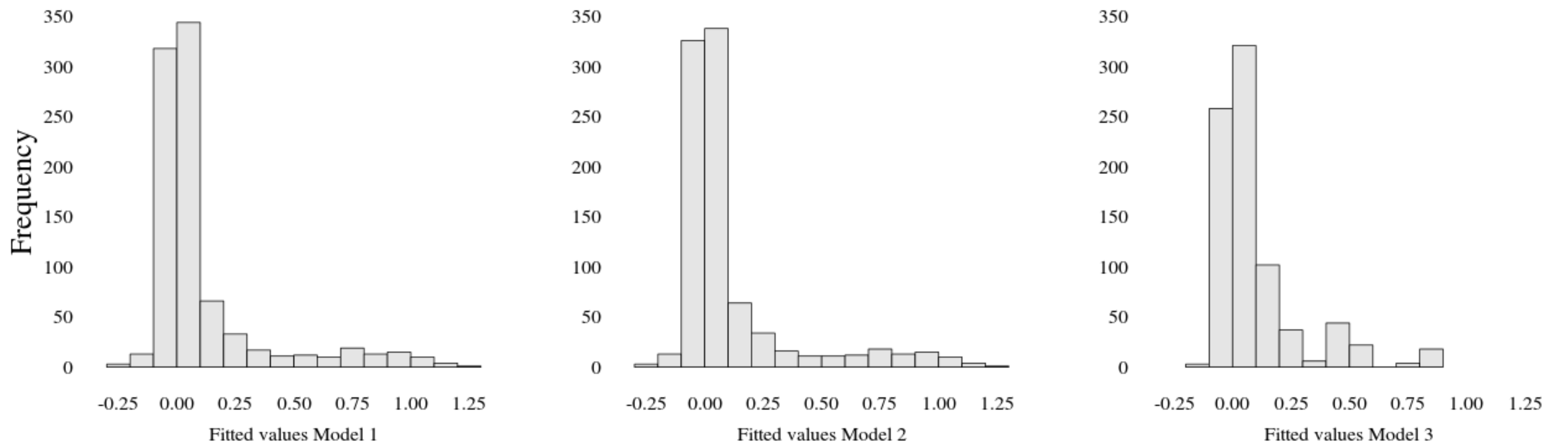

Figure 1: Fitted values linear probability model

Given that warming increases the risk of war we would expect that the fitted probabilities follow a trend where higher temperatures correspond with higher fitted probabilities. To examine this I plot the fitted probabilities against the temperature deviations from the country mean, as shown in figure 2. I use deviations from the country mean in order to better compare across countries.[10] The figure illustrates that there is no clear trend with regard to higher temperatures being associated with higher probabilities of war. The figure does show that in general the model is accurate in correctly predicting the incidence of armed conflict, as most observations with war correspond with higher fitted probabilities.

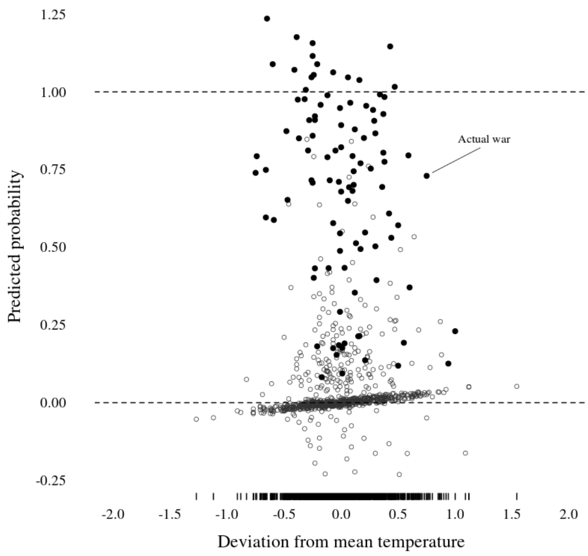

Figure 2: Fitted probabilities versus temperature deviations from the mean for preferred model

[10]This also parallels the model as it is estimated using country fixed effects which demeans the temperature variables using within-transformation (i.e. $T_{it} - \overline{T_i}$).

To measure the model's in-sample predictive power I use a separation plot, shown in figure 3. In a separation plot the predicted outcomes of the model are ordered along the x-axis from low to high (also indicated by the line through the plot), and the dark (light) panels correspond with incidences of war (no war) (Greenhill et al., 2011). [11] If a model is capable of matching high-probability predictions with actual events, and low-probability predictions with no events, we should observe a plot separating the dark and light panels.

Figure 3 shows that all three models perform well in separating the cases from the controls. Looking at the Area Under the Curve (AUC) statistic confirms the high in-sample predictive power with values of about 0.98 for model 1 and 2 and 0.94 for model 3. Similar to Buhaug et al. (2010) I find that the high predictive power of the model seems to be predominantly driven by the inclusion of the fixed effects and that the temperature variable has little explanatory power (see appendix).

[11] The triangular marker indicates the expected number of events (from right to left) based on the cumulative of predicted probabilities.

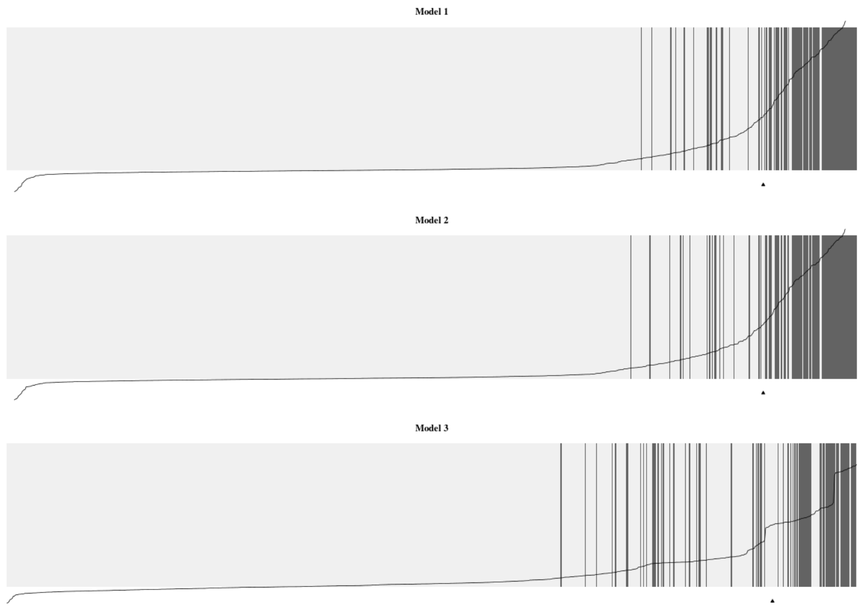

Figure 3: In-sample separation plots

2.2 Out-of-sample predictions

For the out-of-sample predictions I update the dataset to cover the period between 2003-2013 using the most recent data on conflict, temperature and precipitation, income, and regime type.[12] The outcome variable in the model is the incidence of major conflicts, i.e. conflicts resulting in $>$ 1,000 battle-related deaths in a given year, identical to the original work.

Figure 4 shows the average temperature for each country and the whole sample along with the proportion of countries that experience civil war. The figure illustrates that for all countries there has been increase in temperature over time. For 2003-2013 the average temperature is about $0.3°C$ higher

[12] Conflict data is taken from UCDP/PRIO Armed Conflict Dataset (Gleditsch et al., 2002). Climate data comes from the Climatic Research Unit of the University of East Anglia (Harris et al., 2014). Income data is taken from the World Development Indicators (World Bank, 2012) and data on regime type from Polity IV Project (Marshall et al., 2013).

than in 1980s. The upward trend in temperature is what we would expect given the results from the climate change research (IPCC, 2014).[13] At the same time the figure also illustrates that there has been a decreasing trend in the number of civil wars in line with the thesis of Pinker (2011). So on aggregate although on average Sub-Sahara Africa has been getting warmer, there has been a decline in the number of civil wars. This challenges the idea that warming increases the risk of civil war.

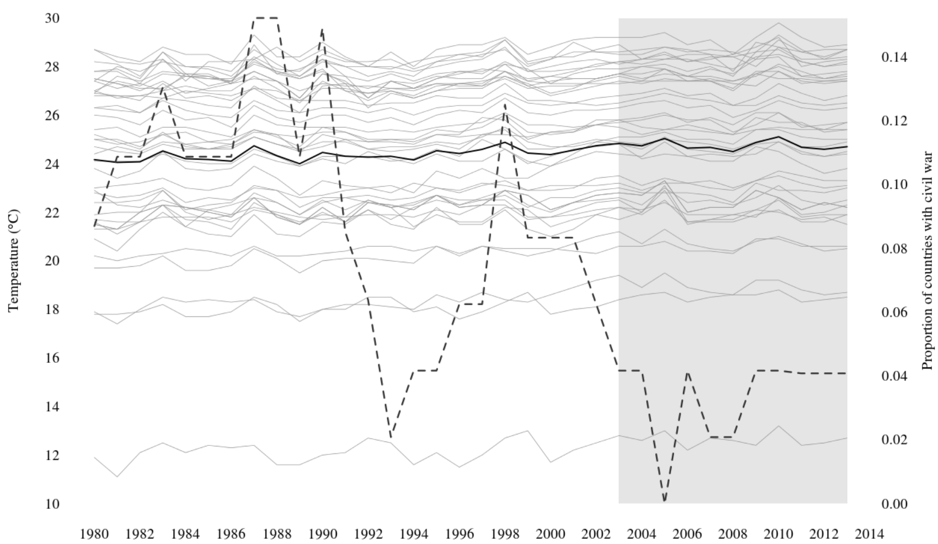

Figure 4: Temperature over time and proportion of countries with civil war, 1980-2013 The light shaded lines represent the temperature time series for each individual country while the darker shaded line represents the sample average. The dashed line shows the proportion of countries in civil war.

[13]See figure B1 for a graph of only the temperature data.

The main model is cross validated by generating predictions for 2003-2013. Similar to the in-sample predictions, the LPM generates predicted probabilities outside the unit interval, ranging from -0.97 to 1.22 (upper panel figure 5) with about 37% of all values outside the interval.[14] With an *AUC* statistic of 0.80, the predictions seem reasonably accurate.[15] Inspection of the separation plot shows however that the model is not entirely accurate in its predictions (upper panel figure 5). At the higher end of the probability distribution the model correctly separates observations with war from those without but there are also clusters of false negatives. This is not entirely surprising as war is relative rare with just 4% of the observations coded as experiencing civil war.

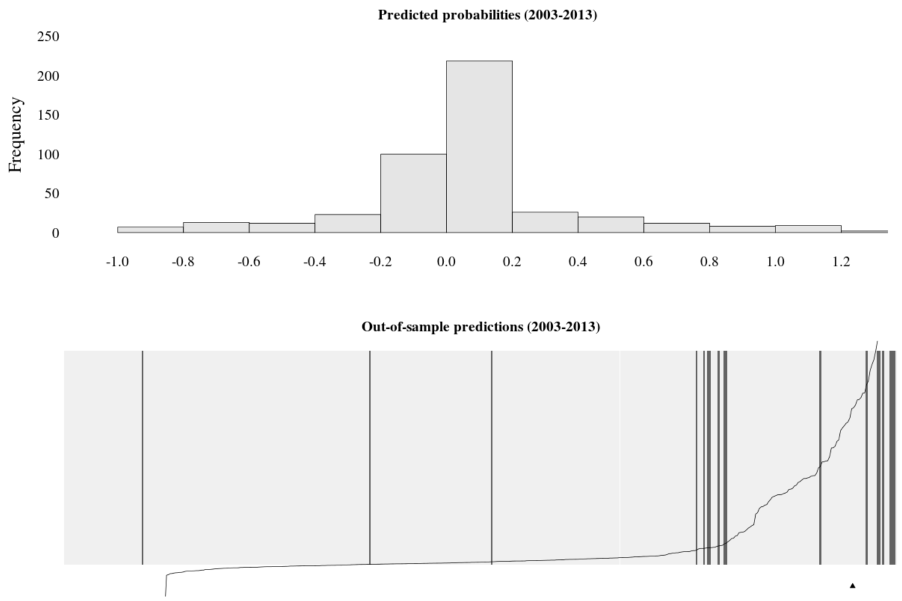

Figure 5: Out-of-sample predicted values and separation plot (2003-2013)

[14]I don't report results for predictions using the other two models as the predicted probabilities fall well outside the unit interval. For example, when including rainfall the predicted probabilities range from -11 to 14. Analytically this does not make much sense.

[15]$N = 451$, 18 cases of war.

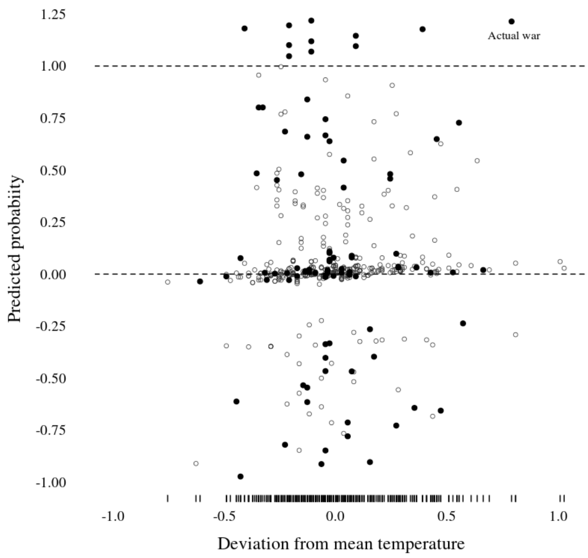

Figure 6: Predicted probabilities versus temperature deviations from the mean for preferred model

This is also illustrated by figure 6. In contrast with the in-sample predictions (figure 2), the model is less accurate in correctly predicting the outcome leading to more false positives. Similar to the in-sample predictions the results show that there no clear association between higher temperatures and higher predicted probabilities of war.[16]

In general the model seems to over-predict the incidence of violent armed conflict. Using $\hat{p} > 0.5$ as a threshold for forecasting war, the model predicts wars in 37 observations. Out of these 37 observations 7 actually experience war leaving 30 false positives as well as generating 11 false negatives (all summarised in table 1).[17] If we only consider the predicted wars than then out-of-sample prediction success rate drops considerably: only 18.9% of the

[16]As the graph shows there is one exception which is Sudan in 2010, but this was an ongoing war.

[17]The logit model performs similarly but includes more false positives.

predicted wars actually experienced a war compared with a 68.2% success rate for the in-sample data.

Interestingly the model identifies a number of countries in the Great Lakes region such as Burundi, DRC, and Rwanda as high risk countries. The model also consistently generated high probabilities for Sudan which to some extent resonates with the results by Maystadt et al. (2014).

Table 1: Actual versus predicted wars out-of-sample for 2003-2013

	$\hat{p} \leq 0.5$	$\hat{p} > 0.5$
No war	403 obs.	Angola (2003-2005) Burundi (2003-2013) DRC (2003-2012) Rwanda (2013) Sudan (2005, 2007-2009, 2013)
War	Chad (2006) Liberia (2003) Nigeria (2013) Rwanda (2009) Somalia (2007-2012) Uganda (2004)	DRC (2013) Sudan (2003, 2004, 2006, 2010-2012)

A possible explanation for the over prediction could be that over the past decade African wars have become less sensitive to climate, an argument also brought forth by Burke et al. (2010).[18] For 15 of the 30 false positives there was a conflict in the predicted country-year, although at a lower intensity level (<1,000 battle-related deaths).[19] This could suggest that maybe the conflicts that are responsive to temperature have become less intense.[20] Accounting for this by including minor conflicts with between 25 and 999 battle-related deaths does triple the correct prediction rate for conflicts (from 7 to 22) and reduces the false positive rate by half (from 30 to 15). However, there is also a steep increase in the false negative rate which stands at 69 now.

Rather than focussing on predicted conflict in specific country-years we can also consider the general forecast of civil wars. For each year I tally all the predicted probabilities and compare it with the actual number of wars, as shown in figure 7 which includes both the in-sample and out-of-sample forecasts. Save for some local reversals the actual number of civil wars in a given year has declined from around 5 per year between 1981-2002, to 2 per year for 2003-2013. This trend in the data is captured by the model through the inclusion of the country fixed effects and specific time trends. The forecast number of wars moves towards zero over time. In the training data (1981-2002) the temperature variables capture some of the local reversions which is also reflected in the out-of-sample data. In contrast with the previous exercise, based on predicted probabilities alone the model performs slightly better. The forecast number of wars is on average only

[18] Burke et al. (2010) state that "African conflict appears less sensitive to climate over the past decade, a change likely related to the unprecedented growth and democratization that most of the continent has recently experienced".

[19] There were civil conflicts in the following country-years: Angola 2004; Burundi 2003-2006, 2008; DRC 2006-2008, 2012; Sudan 2005, 2007-2009.

[20] See figure B2 for trends in conflict over time.

0.5 off the actual number of wars and 0.3 if I exclude 2005 which seems like an outlier.[21] A worry with regard to the predicted outcomes is that these

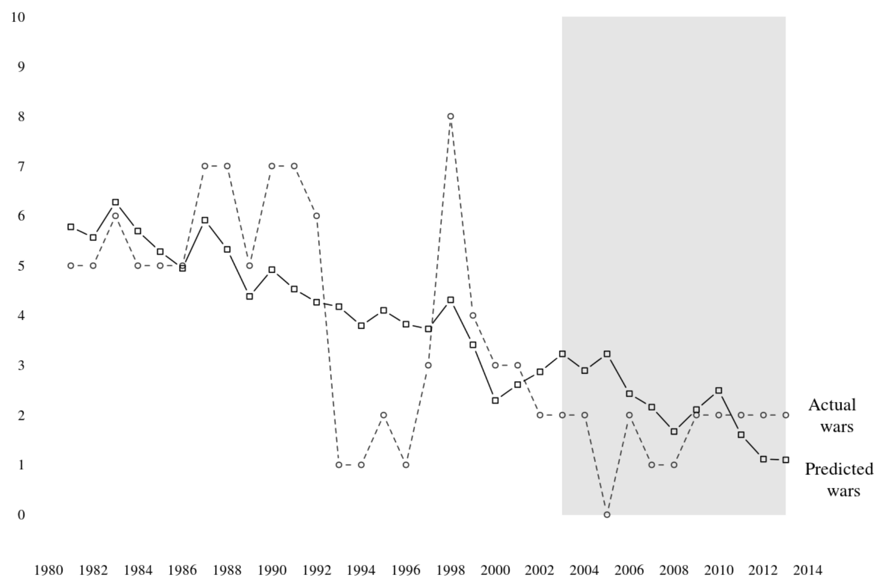

Figure 7: Forecast

could be predominantly driven by the inclusion of country fixed effects and country-specific time trends, as the predicted probabilities with values >0.5 are all in countries with high levels of past conflict prevalence. Omitting country fixed effects from the model, estimating only with the temperature variable and country-specific year trend leads to exactly the same predictions as the full model.[22] It therefore seems that the effect of factors such as geographical characteristics and colonial history on conflict risk are comparable across countries.

To test the predictive power of the temperature variable I generate pre-

[21]For 2003-2013 the model forecast 24 wars. The actual number was 18. For the training data the sum of predicted probabilities equals the actual number of wars (98).

[22]Using a model omitting the country-specific time trend reduces the range of predicted values and leads to fewer predicted wars. The number of correctly predicted wars decreases by 1, false positives by 3 while false negatives increase by 1.

dictions based on a model estimated with fixed effects and time trends only, the results for which are very similar to the full model (including temperature variables). The main difference is that this model leads to a slightly larger number of false positives, 34 versus 30, but it correctly predicts the same wars.[23] The results are identical when generating predictions from a model with country-specific time trends only, indicating that the predicted probabilities are almost exclusively driven by the inclusion of these effects. Figure 8 illustrates these differences in predictive power for three different models: A model specified with the temperature variables and country-specific year trend, a model with only the temperature variables, and a model with only the country-specific year trend. The results show that including the country-specific time trend is beneficial for the predictive power of the model, and also that there is little difference between a LPM with or without temperature.

[23] See table B1 for results

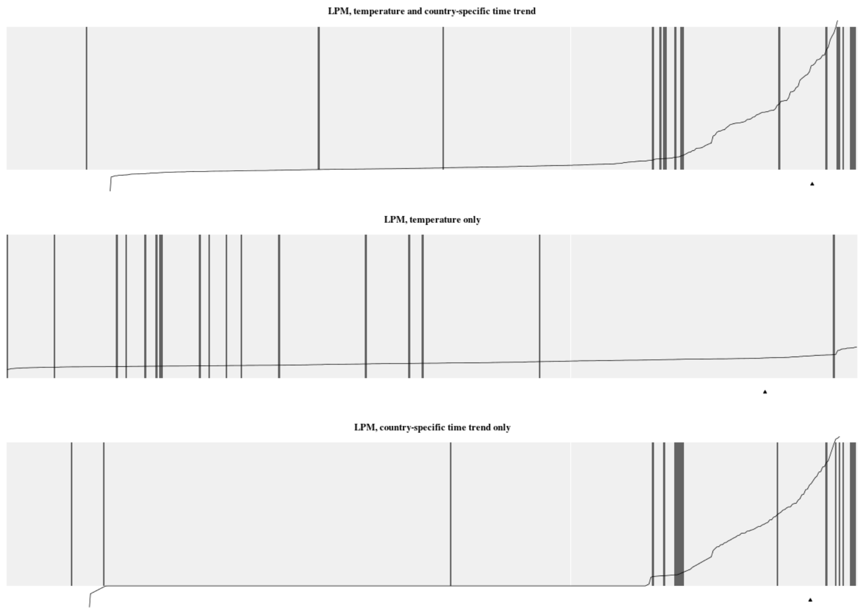

Figure 8: Predictive performance LPM

As an additional check I generate predictions using a benchmark civil war model including lagged time-varying covariates on GDP per capita, population, regime type, natural resource rents, and a fixed variable for ethnic polarisation.[24] The model is estimated using OLS as well as logit.

Accurately predicting war is difficult as illustrated by figure 9. In all estimations the predicted outcomes are below 0.5, which was used as threshold so far. The LPM model including the temperature variable is the only one producing predicted probabilities outside the 0-1 interval. In general the LPM and logit results are comparable and in both cases separation is not perfect. According to the *AUC* statistic, the LPM actually has a slight edge over logit (0.769 vs. 0.765) in correctly predicting the outcome (*note:* these are the models without the temperature variable). However, since there are a lot of zeroes in the data it is not too difficult to register high *AUC* statistics, and as shown both models have trouble correctly identifying the cases that experienced civil war.

Estimating the benchmark model including the temperature variables leads to a 8 percentage point reduction in the *AUC* statistic and reduces the predictive power of the model as illustrated in the lower panel of figure 9.

[24]Data on GDP, population, and natural resource rents all taken from World Bank Development Indicators. Resource rents are measured as percentage of GDP. Regime type is taken from the Polity IV project and data on ethnic polarisation from Garcia-Montalvo and Reynal-Querol (2005)

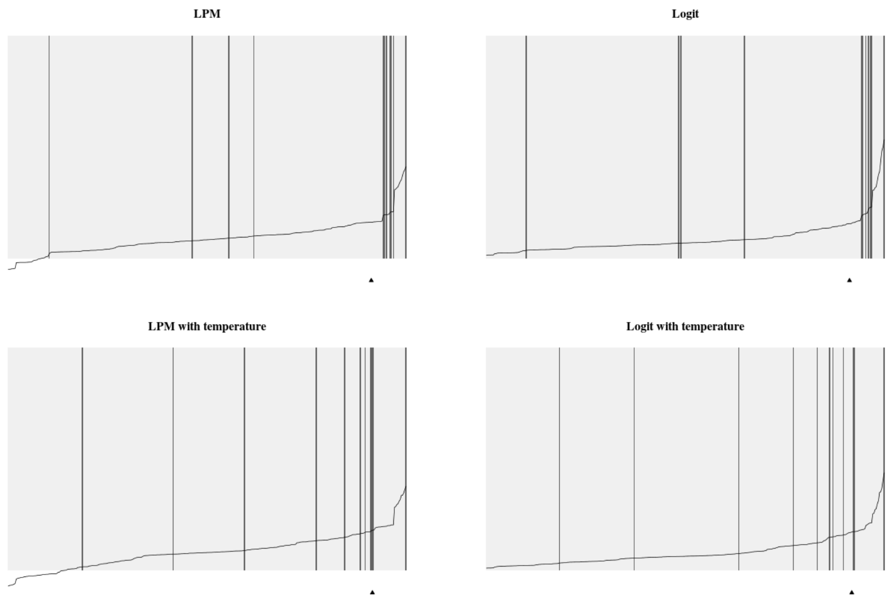

Figure 9: Separation plot benchmark civil war model

3 Discussion

Predicting conflict accurately is a difficult task. There are many complex underlying dynamics that are overlooked by the statistical models and there is always the issue that certain social constructions don't behave uniformly which makes predicting conflict very difficult indeed. In the quantitative literature on violent armed conflict there seems to be very little appreciation for the predictive power of the models as most analyses focus on the statistical significance of the variables. This studies follows Ward et al. (2010) and argues that examining predicted probabilities and generating out-of-sample predictions can provide useful insights in the performance of the models.

To illustrate this I focus on the highly contested research area of climate-conflict. Using one of the standard models linking temperature to climate, I find that it produces inaccurate out-of-sample predictions which are mainly driven by the inclusion of country fixed effects and country-specific time trends rather than the variable of interest. Using a benchmark civil war model I find that out-of-sample prediction is difficult in general: the time-varying covariates have little predictive power and this model performs worse when including temperature.

As I attempt to gauge the predictive power of temperature variation on conflict, there are two issues that are not accounted for and that might be of interest for future research on conflict prediction and are also described in O'Loughlin et al. (2014a) and O'Loughlin et al. (2014a). First, the model estimates the link between weather variation and conflict using the country-year as unit of analysis. This means that it neglects within-country as well as within-year variability in both climate and conflict. Second, using a crude binary indicator means that a lot of information on conflict intensity is lost. Additionally, the assumption that the relationship observed between climate and conflict in the past will be the same in the future might be misleading. These are factors worth taking into account for further examination of the predictive power of climate variability on violence.

References

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review 93*(1), 113–132.

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

Beck, N. (2011). Is OLS with a binary dependent variable really OK ?: Estimating (mostly) TSCS models with binary dependent variables and fixed effects. Unpublished working paper NYU.

Bernauer, T., T. Böhmelt, and V. Koubi (2012). Environmental changes and violent conflict. *Environmental Research Letters 7*(1), 015601.

Blair, R. A., C. Blattman, and A. Hartman (2014, October). Predicting Local Violence.

Blattman, C. and E. Miguel (2010, March). Civil War. *Journal of Economic Literature 48*(1), 3–57.

Buhaug, H. (2014). Concealing agreements over climate-conflict results. *Proceedings of the National Academy of Sciences of the United States of America 111*(6), E636.

Buhaug, H., H. Hegre, and H. Strand (2010). Sensitivity Analysis of Climate Variability and Civil War.

Buhaug, H., J. Nordkvelle, T. Bernauer, T. Böhmelt, M. Brzoska, J. Busby, A. Ciccone, H. Fjelde, E. Gartzke, N. Gleditsch, J. Goldstone, H. Hegre, H. Holtermann, V. Koubi, J. Link, P. Link, P. Lujala, J. O'Loughlin, C. Raleigh, J. Scheffran, J. Schilling, T. Smith, O. Theisen, R. Tol, H. Urdal, and N. von Uexkull (2014). One effect to rule them all? a comment on climate and conflict. *Climatic Change 127*(3-4), 391–397.

Burke, M., J. Dykema, D. Lobell, E. Miguel, and S. Satyanath (2010). Climate and civil war: Is the relationship robust? NBER Working Paper No. 16440.

Burke, M., J. Dykema, D. B. Lobell, E. Miguel, and S. Satyanath (2014). Incorporating Climate Uncertainty into Estimates of Climate Change Impacts. *The Review of Economics and Statistics forthcoming*(October), 1–33.

Burke, M., E. Miguel, S. Satyanath, J. a. Dykema, and D. B. Lobell (2009, December). Warming increases the risk of civil war in Africa. *Proceedings of the National Academy of Sciences of the United States of America 106*(49), 20670–4.

Burke, M. B., E. Miguel, S. Satyanath, J. A. Dykema, and D. B. Lobell (2010, June). Reply to Sutton et al.: Relationship between temperature and conflict is robust. *Proceedings of the National Academy of Sciences 107*(25), E103–E103.

Chamarbagwala, R. and H. E. Morán (2011). The human capital consequences of civil war: Evidence from Guatemala. *Journal of Development Economics 94*(1), 41–61.

Clemens, M. (2015). The Meaning of Failed Replications : A Review and Proposal.

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

Dell, M., B. F. Jones, and B. a. Olken (2014). What Do We Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature 52*(3), 1–70.

Gaibulloev, K., T. Sandler, and D. Sul (2014). Dynamic Panel Analysis under Cross-Sectional Dependence. *Political Analysis 22*, 258–273.

Garcia-Montalvo, J. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *The American Economic Review 95*(3), 796–816.

Gelman, A. and J. Hill (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

German Advisory Council on Global Change (2008). *World in Transition: Climate Change as a Security Risk*. Earthscan.

Gleditsch, K. S. and M. D. Ward (2013, January). Forecasting is difficult, especially about the future: Using contentious issues to forecast interstate disputes. *Journal of Peace Research 50*(1), 17–31.

Gleditsch, N. P. (2012). Whither the weather? Climate change and conflict. *Journal of Peace Research 49*(1), 3–9.

Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002, September). Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research 39*(5), 615–637.

Goldstone, J., R. Bates, D. Epstein, T. R. Gurr, M. B. Lustik, M. G. Marshall, J. Ulfelder, and M. Woodward (2010). A global model for forecasting political instability. *American Journal of Political Science 54*(1), 190–208.

Greenhill, B., M. D. Ward, and A. Sacks (2011, October). The Separation Plot: A New Visual Method for Evaluating the Fit of Binary Models. *American Journal of Political Science 55*(4), 991–1002.

Harris, I., P. Jones, T. Osborn, and D. Lister (2014, March). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology 34*(3), 623–642.

Hegre, H. and N. Sambanis (2006, August). Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of Conflict Resolution 50*(4), 508–535.

Horrace, W. C. and R. L. Oaxaca (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters 90*, 321–327.

Hsiang, S. and M. Burke (2014). Climate, conflict, and social stability: what does the evidence say? *Climatic Change 123*(1), 39–55.

Hsiang, S., M. Burke, and E. Miguel (2013). Quantifying the influence of climate on human conflict. *Science 341*(6151).

Hsiang, S., M. Burke, and E. Miguel (2014). Reconciling climate-conflict meta-analyses: reply to buhaug et al. *Climatic Change 127*(3-4), 399–405.

Hsiang, S. M. and M. Burke (2013, October). Climate, conflict, and social stability: what does the evidence say? *Climatic Change 123*(1), 39–55.

Hsiang, S. M. and K. C. Meng (2014). Reconciling disagreement over climate-conflict results in Africa. *Proceedings of the National Academy of Sciences 111*(6), 2100–2103.

IPCC (2014). *Climate Change 2014: Synthesis Report. Contribution of*

Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC.

Klomp, J. and E. Bulte (2013). Climate change, weather shocks, and violent conflict: A critical look at the evidence. *Agricultural Economics 44* (s1), 63–78.

Marshall, M. G., K. Jaggers, and T. R. Gurr (2013). Polity IV Project: Political Regime Characteristics and Transitions, 1800-2012.

Maystadt, J.-F., M. Calderone, and L. You (2014). Local Warming and Violent Conflict in North and South Sudan. pp. 78.

Meierding, E. (2013). Climate change and conflict: Avoiding small talk about the weather. *International Studies Review 15*, 185–203.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica 49* (6), 1417–1426.

Nuzzo, R. (2014, February). Scientific method: Statistical errors. *Nature 506*, 150–152.

O'Loughlin, J., A. M. Linke, and F. D. W. Witmer (2014a). Effects of temperature and precipitation variability on the risk of violence in sub-Saharan Africa, 1980–2012. *Proceedings of the National Academy of Sciences 111* (47), 16712–16717.

O'Loughlin, J., A. M. Linke, and F. D. W. Witmer (2014b). Modeling and data choices sway conclusions about climate-conflict links. *Proceedings of the National Academy of Sciences of the United States of America 111* (6), 2054–5.

O'Loughlin, J., F. D. W. Witmer, A. M. Linke, A. Laing, A. Gettelman, and J. Dudhia (2012, November). Climate variability and conflict risk in East Africa, 1990-2009. *Proceedings of the National Academy of Sciences of the United States of America 109*(45), 18344–9.

Pinker, S. (2011). *The better angels of our nature: The decline of violence in history and its causes.* Viking.

Raleigh, C., A. Linke, and J. O'Loughlin (2014). Extreme temperatures and violence. *Nature Climate Change 4*, 76–77.

Scheffran, J., M. Brzoska, J. Kominek, P. M. Link, and J. Schilling (2012a). Climate Change and Violent Conflict. *Science 336*(6083), 869–871.

Scheffran, J., M. Brzoska, J. Kominek, P. M. Link, and J. Schilling (2012b). Disentangling the climate-conflict nexus: Empirical and theoretical assessment of vulnerabilities and pathways. *Review of European Studies 4*(5), 1–13.

Schrodt, P. a. (2013). Seven deadly sins of contemporary quantitative political analysis. *Journal of Peace Research 51*(2), 287–300.

Serneels, P. and M. Verpoorten (2013). The impact of armed conflict on economic performance: Evidence from rwanda. *Journal of Conflict Resolution.*

Stern, N. (2006). *Stern Review: The economics of climate change.*, Volume 30. London: HM Treasury.

Sutton, A. E., J. Dohn, K. Loyd, A. Tredennick, G. Bucini, A. Solórzano, L. Prihodko, and N. P. Hanan (2010, June). Does warming increase the risk of civil war in Africa? *Proceedings of the National Academy of Sciences of the United States of America 107*(25), E102; author reply E103.

TEDxBerkeley (2014). Climate, conflict, and African development: Edward Miguel at TEDxBerkeley.

Ward, M. D., B. D. Greenhill, and K. M. Bakke (2010, March). The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research 47*(4), 363–375.

Weidmann, N. B. and M. D. Ward (2010, July). Predicting Conflict in Space and Time. *Journal of Conflict Resolution 54*(6), 883–901.

World Bank (2012). World development indicators. http://data.worldbank.org/data-catalog/world-development-indicators.

A Complementary results

A.1 Logit estimation

Re-estimating the model using logit (results in table A1) shows that for model 1 and 2 the effect of temperature ceases to be statistically significant within the traditional boundaries.[25] Model 3 does report a statistically significant coefficient for temperature but the magnitude of the effect is lower than that of GDP: At the upper bound, a unit increase in temperature corresponds with a 29.5% positive difference in the probability of observing war, while a unit increase in GDP per capita decreases the likelihood of war by 79.5%.

Table A1: Results model estimation using logit

	Model 1	Model 2	Model 3
Temperature	1.3	1.4	1.2
	(0.9)	(0.9)	(0.6)**
Temperature$_{(t-1)}$	0.4	0.6	0.6
	(0.9)	(1.1)	(0.7)
Precipitation		−0.9	−0.8
		(2.3)	(1.6)
Precipitation$_{(t-1)}$		2.0	0
		(2.1)	(1.6)
GPD per capita$_{(t-1)}$			−3.2
			(1.6)***
Regime type$_{(t-1)}$			0
			(0.1)
Deviance	165.52	164.45	255.96
AIC	333.52	336.45	345.96
AUC	0.9805	0.9805	0.936
N	889	889	815
Number of wars	98	98	81
Country FE	Yes	Yes	Yes
Country-specific time trend	Yes	Yes	–
Time FE	–	–	Yes

Notes. Intercept not reported. FE, fixed effects; AIC, Akaike information criterion. Robust standard errors clustered at country level (given in parentheses). *** p≤ 0.01, ** p≤ 0.05, *≤ 0.1

[25] Note that due to the inclusion of the country and year indicators perfect prediction occurs in countries with little variation in the outcome variable.

A.2 Explanatory power variables

As stated in the main text, p-values are not meant as a definitive test (Nuzzo, 2014) but serve as an indicator for results that are worth closer examination. As Ward et al. (2010) argue we should not rely too much on these p-values to determine the strength of a particular model and be cautious about the implied effects.

A useful measure for the explanatory power of variables of interest is the change in the AUC statistics. To gauge the in-sample predictive power of temperature I re-estimate model 3 omitting one variable at a time to measure the change in the AUC statistic.

The results shown in figure A1 illustrate that in general most variables have very little explanatory power, and in some cases the model is even better off not including them in the model specification. Although temperature seems to be the best predictor of armed conflict, its explanatory power is marginal. Omitting temperature from the model leads to a reduction of 0.002 in predictive power.

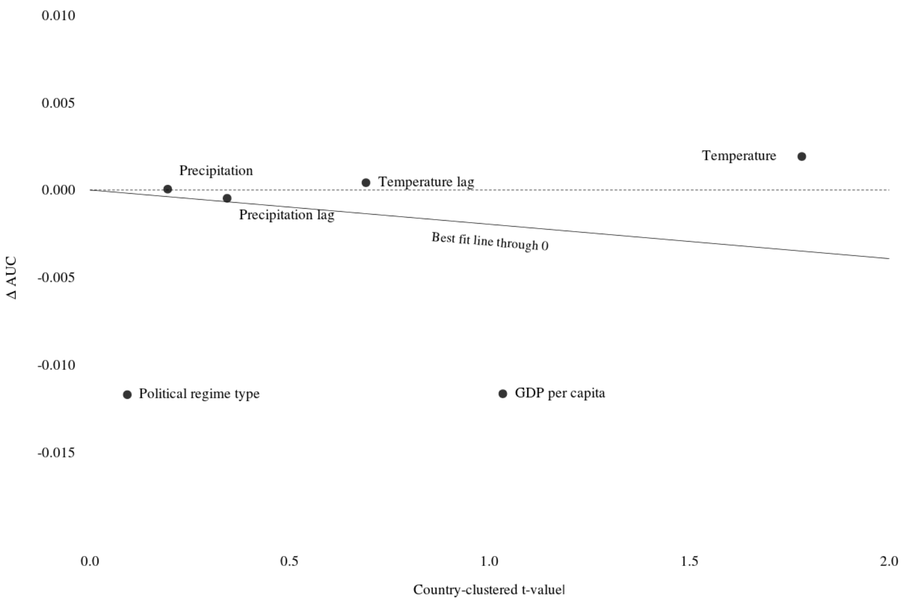

Figure A1: Statistical significance versus predictive power of variables.

A.3 In-sample predictions

Using $\hat{p} > 0.5$ as the threshold for predicted war, the main model (model 1) correctly predicts 73 conflicts while generating 11 false positives and 25 false negatives. The specific cases are shown in table A2.

Most of the false positives are generated for countries that have high levels of conflict prevalence such as Sudan which experienced conflict in 18 out of the 22 years in the sample and Angola (17 out of 19).

The model has more trouble with correctly predicting conflict in countries with relatively lower levels of war prevalence such as Sierra Leone, Rwanda, and Chad.

These results seem to suggest that the fitted probabilities of this model are driven largely by country-specific trends rather than the temperature variable. Estimating the model omitting the temperature variable and only including fixed effects and country-specific time trends leads to almost exactly the same predictions. The only difference is that South Africa-1987 has a lower predicted probability leading to an additional false negative.

Table A2: Actual versus predicted wars in-sample for 1981-2002

	$\hat{p} \leq 0.5$	$\hat{p} > 0.5$
No war	780 obs.	Angola (1996-1997) DRC (2002) Ethiopia (1986) South Africa (1984-1985) Sudan (1981-1982, 1993-1994) Uganda (1990)
War	Burundi (1998) Chad (1987, 1990) DRC (1997-2000) Congo (1997-1998) Ethiopia (1991) Guinea-Bissau (1998) Liberia (1990-1992) Rwanda (1991-1992, 1998, 2001) Sierra Leone (1998-1999) Somalia (1988, 1990-1992) South Africa (1988) Uganda (1991)	Angola (1981-1995, 1998-1999) Burundi (2000-2002) Ethiopia (1981-1985, 1987-1990) Mozambique (1981-1992) South Africa (1981-1983, 1986-1987) Sudan (1983-1992, 1995-2002) Uganda (1981-1989)

B Additional tables and figures

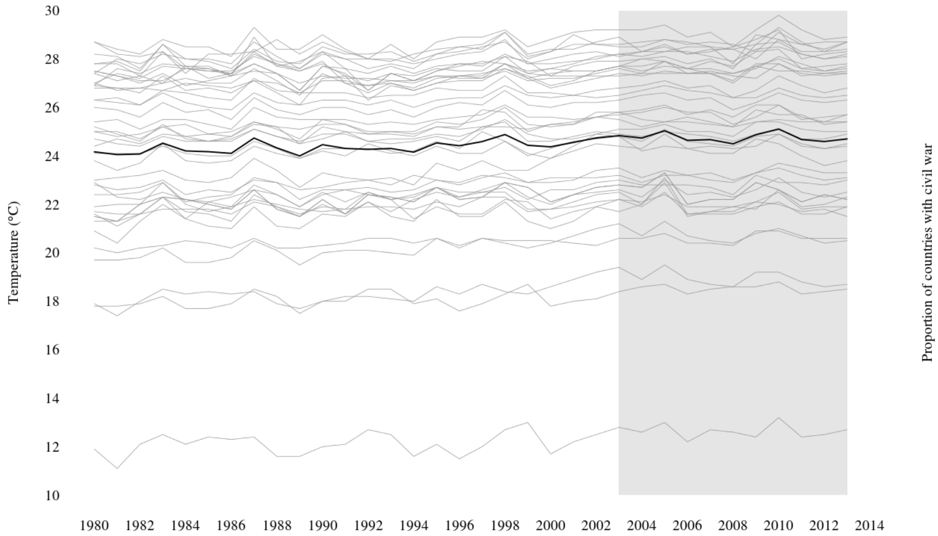

Figure B1: Average temperature over time for 1980-2013. The light shaded lines represent the country average, the dark shaded line the continent average. Data: Climatic Research Unit, University of East Anglia.

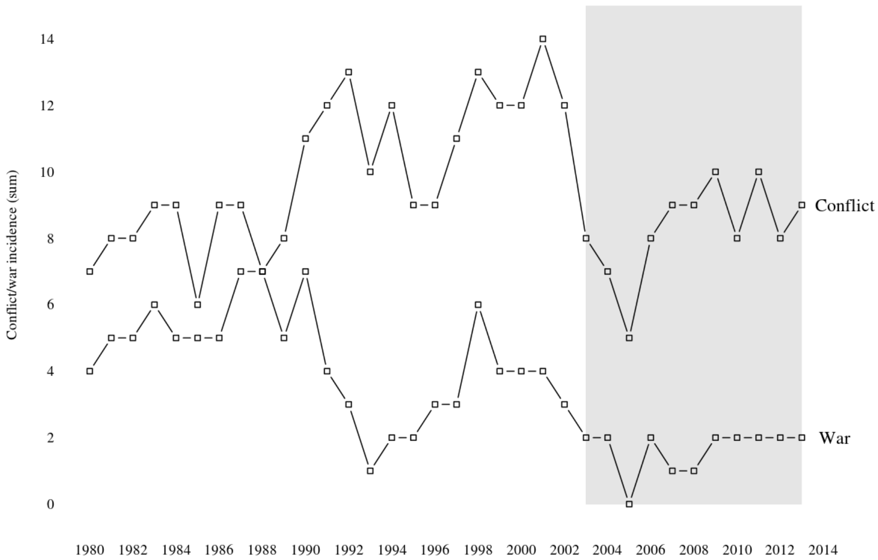

Figure B2: Number of civil conflicts and war over time for 1980-2013. *"Conflict"* includes all conflicts with the number of battle-related deaths between 25-999, while *"War"* includes all conflicts with the number of battle-related deaths above 1,000. Data: UCDP/PRIO.

Table B1: Actual versus predicted wars out-of-sample for 2003-2013 (Model with country fixed effects and country-specific time trends only)

	$\hat{p} \leq 0.5$	$\hat{p} > 0.5$
No war	403 obs.	Angola (2003-2007) Burundi (2003-2013) DRC (2003-2012) Rwanda (2010-2013) Sudan (2005, 2007-2009)
War	Chad (2006) Liberia (2003) Nigeria (2013) Rwanda (2009) Somalia (2007-2012) Uganda (2004)	DRC (2013) Sudan (2003, 2004, 2006, 2010-2013)