Causal Machine Learning for Forecasting the Effects of Weather Shocks on Conflict*

Joel Ferguson†

February, 2023

Preliminary and Incomplete.
Please do not cite or circulate.

Abstract

There is a large body of empirical evidence supporting the hypothesis that weather shocks are associated with greater risk of civil conflict in low-income settings. However, much less is known about how causal estimates of the effects of weather on conflict incidence can be used to enhance forecasts of conflict. To this end, I develop a class of causal machine learning estimators, re-centered neural networks (RcNNs), and apply them to the tasks of predicting and forecasting conflict. I show that RcNNs predict civil conflict incidence in Africa out-of-time better than a benchmark panel econometric method and exhibit similar performance to ordinary machine learning methods which cannot be interpreted causally. The results suggest that causal machine learning methods can potentially play a role in mitigating the effects of climate change on deadly conflict.

*I thank my advisors — Marco Gonzalez-Navarro, Ted Miguel, and Max Auffhammer — for their continued support and guidance. Any errors are my own.

†University of California, Berkeley. joel_ferg@berkeley.edu

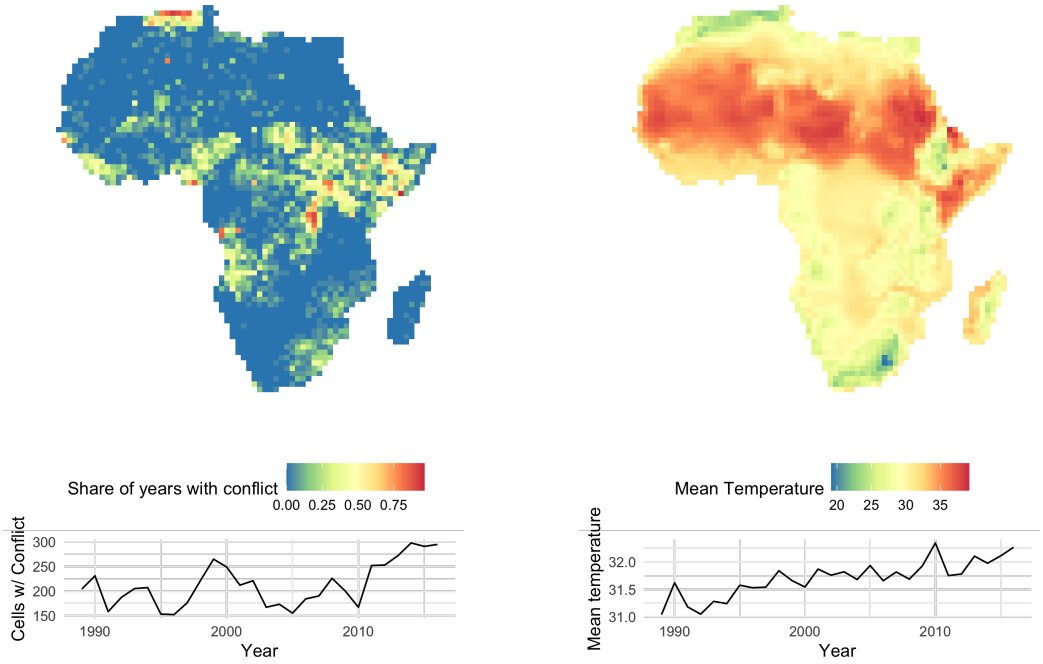
There is a large body of empirical research studying the causal relationship between weather and conflict (Hsiang et al., 2013; Burke et al., 2015). While there appears to be growing academic consensus that certain weather shocks causally increase conflict likelihood, some disagreement remains regarding this claim. Chapter 16 of the IPCC6 report, for example, states "Climate variability and extremes are associated with increased prevalence of conflict, with more consistent evidence for low-intensity organised violence than for major armed conflict (medium confidence)." Notably, there is substantial disagreement over the importance of weather as a determinant of conflict, as well as the mechanisms through which weather affects conflict likelihood (Mach et al., 2019).

Perhaps one reason for this lingering doubt and disagreement is the lack of evidence that causal effects of weather on conflict can be used to predict or forecast conflict with meaningful accuracy. While some studies have used their estimates to predict conflict under climate change (e.g. Burke et al. (2009); Harari and Ferrara (2018)), they generally do not evaluate the performance of their predictions out-of-sample. Importantly, the presence of a statistically significant causal effect of weather does not imply the model performs better at predicting than a model which omits weather entirely, as has been suggested is the case for the Climate-GDP growth relationship (Newell et al., 2021).

This paper aims to demonstrate that causal estimates of the effect of weather on conflict are useful for forecasting and prediction. To do so, I develop a class of causal machine learning estimators, re-centered neural networks (RcNNs), designed to explicitly leverage features of the weather data generating process. The RcNN architecture is built on a base network, which can be any feed-forward neural network for regression. This model is duplicated a number of times and the output of these duplicates is averaged. Importantly, the base network and all duplicates share weights, meaning that they all implement the same function on their respective inputs. As such, when the inputs to the duplicate networks are drawn from the same distribution as the input to the base network, the difference between the output of the base network and the averaged outputs of the duplicates has an expected value of 0. This (approximate) re-centering of the model ensures that only the exogenous shock components of weather are used to predict the outcome, granting the model causal identification (Borusyak and Hull, 2020).

I train a RcNN to predict yearly civil conflict incidence on a $1^\circ \times 1^\circ$ grid over Africa on daily temperature and rainfall data and compare the out of sample predictions to

Figure 1. Descriptive Statistics



Notes: The maps in this figure show mean conflict incidence (left) and temperature in degrees Celsius (right) for 1989-2016. The charts show total number of cells with any conflict (left) and mean daily temperature (right) by year from 1989-2016.

two baseline models: a multi-way fixed effects estimator, similar to those employed in a number of studies on the relationship between climate and conflict, and an ordinary neural network which omits the re-centering of the RcNN.

The results are striking. While the fixed effects estimator performs only slightly better than a random guesser out of sample, the RcNN can correctly identify a random conflict weather series from a random no-conflict weather series about 71% of the time. The RcNN also explains more than 240 times as much variation in-sample as mean weather and continues to explain more than 2% of out-of-sample variance, whereas the predictions using mean weather have negative out-of-sample within-R^2. The RcNN also approaches the performance of the ordinary neural net or even outperforms it in some cases.

1 Data

Following the recent climate-conflict literature I build the data on a $1° \times 1°$ grid over Africa, resulting in a set of 2,755 units. Figure 1 provides a description of the spatial and temporal variation in the data.

1.1 Conflict Data

Conflict data used in this study come from the Uppsala Conflict Data Program (UCDP) Global Event Dataset (GED) (Sundberg and Melander, 2013). A record in the GED corresponds to a conflict event, defined as "[an] incident where armed force was used by an organised actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date." Events are identified from news reports and validated against local sources. In addition to the date, latitude, and longitude of the event, the GED also contains information on the participants in the event, the number of deaths and casualties, and the precision with which this information is known. The GED contains data on events from around the world starting in 1989 through to the present.

To reduce measurement error, I drop events where the location is only known to the country, first administrative unit, or a non-administrative feature (14.1% of events in Africa). I then assign events to the grid cell that contains them and aggregate to the cell×year level, taking the sum of deaths, count of events, and creating an indicator for conflict incidence.

1.2 Weather Data

Weather data come from the Climate Hazards Center Infrared Temperature with Stations (CHIRTS) data set. CHIRTS is a daily weather re-analysis data set which combines weather station data with thermal infrared imagery to produce temperature estimates on a $0.05° \times 0.05°$ grid for 1983-2016. The CHIRTS prediction method explains about 40% more variation in observed temperature in Africa than the popular University of East Anglia Climate Research Unit product (Funk et al., 2019), making it especially suitable for this application.

I use a $0.25° \times 0.25°$ version of the CHIRTS data covering the African continent. I create a second grid which indicates cell-days with missing values (typically cells covered by water). Missing values in the CHIRTS data are replaced with the mean value for that day.

2 The Re-centered Neural Network Estimator

In this section, I describe the Re-centered Neural Network (RcNN) estimator and prove its consistency. Identification is built on results proven by Borusyak and Hull (2020). The main distinction between the RcNN and the re-centered instrument they

study is that the RcNN does not assume the econometrician knows how to combine the exogenous shocks ex-ante. Borusyak and Hull (2021) also explore this extension of their estimator, but approach it using efficient IV theory to develop their procedure rather the machine learning.

The RcNN is built on a base feed-forward regression network model. This model can be thought of as a nested collection of functions, with the innermost function evaluated on the input data and the outermost function predicting the outcome. Concretely, let there be d inputs, in this paper 730 days of weather data, $\mathbf{X}_i \in \mathbb{R}^d$. A layer, $l \in \{1, 2, ..., L\}$, of the network consists of a matrix $W_l \in \mathbb{R}^{H_l \times H_{l-1})}$, where $H_0 = d$, a constant term $b_l \in \mathbb{R}^{H_l}$, and an "activation function" $\sigma_l(.)$ which implements $\sigma_l(.)$ element-wise. The feed-forward neural network is then given by

$$\widehat{f}_{FFN}(\mathbf{X}_i) = \sigma_L(W_L\sigma_{L-1}(W_{L-1}\sigma_{L-1}(...\sigma_2(W_2\sigma_1(W_1X_i + b_1) + b_2)) + b_{L-1}) + b_L)$$

In the case of this paper, the activation function will always be the "rectified linear unit" (ReLU), $\sigma_l(x) := \max(0, x)$, henceforth denoted $\sigma(.)$. I also train convolutional neural networks, in which a number of the elements of the weight matrices are set to 0 and are not updated when training the network.

The RcNN takes as input $\mathbf{X}_i$ along with a number J of simulated inputs. These inputs are independent draws from the same distribution as the observed inputs, $\tilde{\mathbf{X}}_i^j \sim F_{\mathbf{X}_i}$. The RcNN prediction is then given by

$$\widehat{f}_{RcNN}(\mathbf{X}_i, \tilde{\mathbf{X}}_i^1, \tilde{\mathbf{X}}_i^2, ..., \tilde{\mathbf{X}}_i^J) = \alpha + \beta \left(\widehat{f}_{FFN}(\mathbf{X}_i) - \frac{1}{J}\sum_{j=1}^{J} \widehat{f}_{FFN}(\tilde{\mathbf{X}}_i^j) \right). \tag{1}$$

I now provide the intuition behind the proof of consistent causal estimation. Consider a feed-forward neural network with fixed weights. As implied above, the observed input is drawn from a distribution $\mathbf{x}_i \sim F_{\mathbf{X}_i}$, and is assumed to be independent of unobserved determinants of the outcome ε_i. Note however, that $\widehat{f}_{FFN}(\mathbf{x}_i)$ may still be correlated with ε_i

$$\mathbb{E}[\widehat{f}_{FFN}(\mathbf{X}_i)\varepsilon_i] = \mathbb{E}[\mathbb{E}_{\mathbf{X}_i}[\widehat{f}_{FFN}(\mathbf{X}_i)]\varepsilon_i],$$

which is not necessarily zero.

Subtracting the mean of the neural network applied to the simulated inputs ensures that the re-centered network output is not correlated with ε_i

$$\mathbb{E}[\left(\widehat{f}_{FFN}(\mathbf{X}_i) - \frac{1}{J}\sum_{j=1}^{J}\widehat{f}_{FFN}(\tilde{\mathbf{X}}_i^j)\right)\varepsilon_i] = \mathbb{E}[\mathbb{E}_{\mathbf{X}_i}\left[\left(\widehat{f}_{FFN}(\mathbf{X}_i) - \frac{1}{J}\sum_{j=1}^{J}\widehat{f}_{FFN}(\tilde{\mathbf{X}}_i^j)\right)\right]\varepsilon_i]$$
$$= \mathbb{E}[0\varepsilon_i] = 0.$$

This implies that estimating the effect of $\left(\widehat{f}_{FFN}(\mathbf{X}_i) - \frac{1}{J}\sum_{j=1}^{J}\widehat{f}_{FFN}(\tilde{\mathbf{X}}_i^j)\right)$ on the outcome via OLS provides consistent estimation of its causal effect.

Estimation proceeds as follows. First, the base feed-forward neural network is duplicated J times and two layers are added to the end. The first layer subtracts the average of the outputs of the J duplicates from the output of the original network. The second adds a scalar constant and scalar weight times the re-centered output to form the final prediction. The weights are optimized via stochastic gradient descent to minimize mean squared error. Once the weights have been optimized, the final layer (which makes the prediction) is removed, the re-centered output is calculated for all observations and α and β are estimated via OLS.

3 Model Comparison

I compare the RcNN to two baseline models. The first is a multi-way fixed effects model similar to those widely employed in the climate and conflict literature. The second is the base neural network for the RcNN without the re-centering. I evaluate all models on their performance within a training period 1989-2009 and a test period 2010-2016. All fitting, estimation, and hyperparameter selection is conducted using the training period data, whereas the test period data are only used to evaluate out-of-sample performance. I use two evaluation metrics. The first is adjusted within-R^2,

$$R^2(\mathbf{Y}, \widehat{\mathbf{Y}}) = \sum_{i=1}^{N}\left((Y_i - \bar{Y}) - (\hat{Y}_i - \bar{\hat{Y}})\right)^2 / \sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2.$$

The within-R^2 is preferable to ordinary R^2 in this case because we are interested in the predictive power of weather, rather than constant terms or fixed effects necessary for identification.

The second metric is the area under the receiver-operating characteristic (ROC) curve (AUC). The ROC curve is constructed by ordering observations by their predicted probability of conflict incidence with higher values first and calculating the

true positive and false positive rates if the first n observations are predicted to have conflict for $n \in \{1, ..., N\}$ and plotting the true positive rates against the false positive rates. The AUC is the area under the ROC curve and has the interpretation of the probability a classifier will accurately discern a random positive case from a random negative case. As such, a random classifier has AUC = 50 and a perfect classifier has AUC = 100.

Panel Data Estimator

The panel data estimator is specified as follows

$$Conflict_{it} = \alpha_i + \delta_t + \beta^0 Temp_{it} + \beta^1 Temp_{it-1} + \varepsilon_{it}$$

where $Temp_{it}$ is mean daily temperature in grid cell i in year t. Because the panel is balanced, β^0 and β^1 can be estimated by first subtracting the within-year means from each of the variables and estimating

$$\begin{aligned}(Conflict_{it} - \sum_{j=1}^{N} Conflict_{jt}) = \alpha_i + \beta^0 (Temp_{it} - \sum_{j=1}^{N} Temp_{jt}) \\ + \beta^1 (Temp_{it-1} - \sum_{j=1}^{N} Temp_{jt-1}) + \varepsilon_{it}\end{aligned} \tag{2}$$

via OLS. Predictions in the test data can then be formed without estimates of the year fixed effects for the test period by applying the same within-year differencing.

Neural Network and RcNN

The base neural network is a convolutional neural network taking as input the 730 day series of daily weather realizations ending on December 31 of the year of the observation for the four $0.25° \times 0.25°$ CHIRTS cells within the cell of observation. Missing data (typically occuring over water) are imputed as the mean for that day over the entire sample and an indicator for missing data is supplied as an input as well. Therefore, $\mathbf{x}_i \in \mathbb{R}^{730 \times 4 \times 4 \times 2}$. The first layer of the network consists of 12 $7 \times 4 \times 4 \times 2$ filters of weights which convolve along the time dimension. The second layer similarly consists of 48 $10 \times 1 \times 1 \times 12$ filters which also convolve along the time dimension. The final layer of the base neural network is a dense layer with scalar output and

ReLU activation.

For both the ordinary neural network and the RcNN the input weather data are standardized by subtracting the mean temperature in 2000 and dividing by the standard deviation of temperature in 2000 before being provided to the model. Both neural networks are optimized via Adam on mini-batches of 30 observations. Because conflict is a rare outcome, the first two epochs upsample conflict observations to ensure each batch has an equal number of conflict and no-conflict observations. Subsequent epochs do not upsample conflict observations for the ordinary neural network, but continue to do so for the RcNN. Both models are trained using data from 1989-2007 with data from 2008 and 2009 used for validation. Training continues until the validation R^2 does not improve for 10 epochs and the final weights used are those with the best validation R^2.

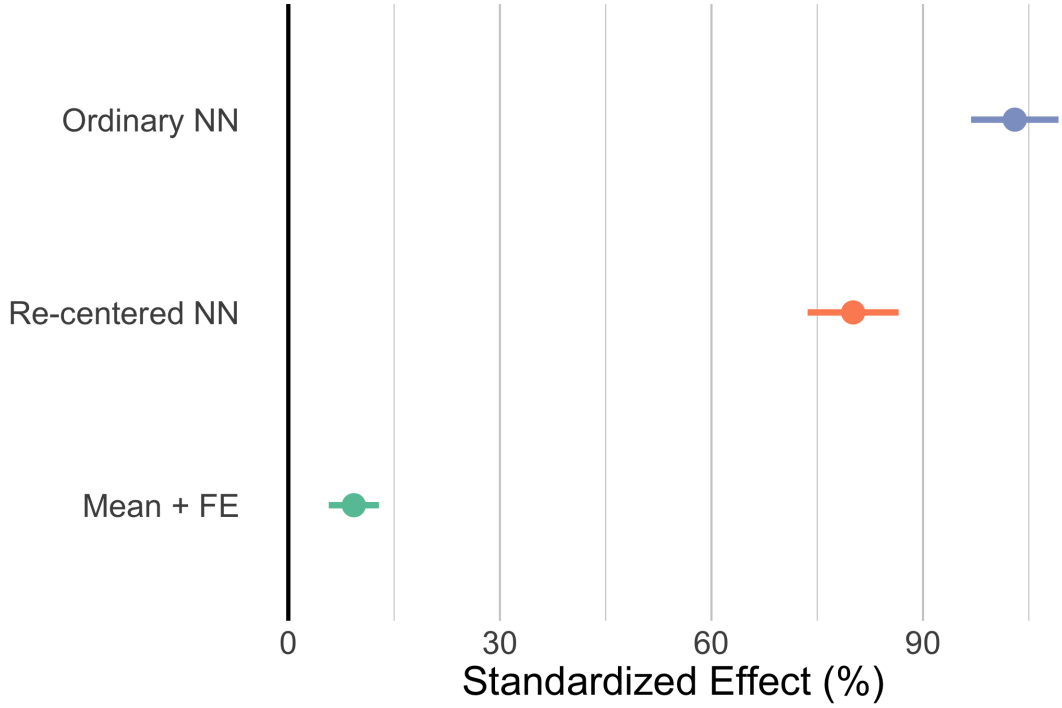
To form the simulated inputs for the RcNN I need to define a distribution of potential temperature realizations for each location-day. I approximate this distribution with the empirical distribution in the 30 days prior to the day of interest and the same 30 day period in the preceding year. For each day of the sample I draw 30 times with replacement from this distribution to form 30 simulated day series. Simulated temperature is for a given draw is then defined as the realized temperature on the simulated day series. This procedure preserves spatial dependence in weather realizations as every location has the same simulated day series.

4 Results

All methods studied succeed in finding a statistically significant relationship between temperature shocks and civil conflict incidence. The panel fixed-effects estimate suggests that a one standard-deviation increase in mean temperature (0.42 degrees Celsius) is associated with a 9.2% increase in conflict likelihood. This is in line with previous estimates, for example it is close to the meta-analysis estimate of 11.3% estimated by Burke et al. (2015).

However, the neural networks appear to find much stronger relationships. A one standard deviation increase in the ordinary neural network's prediction is associated with over 100% increase in conflict incidence. The analogous number for the RcNN is an 80% increase. While not necessarily tied to out-of-sample predictive performance, these relatively larger estimates suggest that the neural networks succeed in finding more meaningful temperature variation.

Figure 2. Estimated Standardized Effects by Method

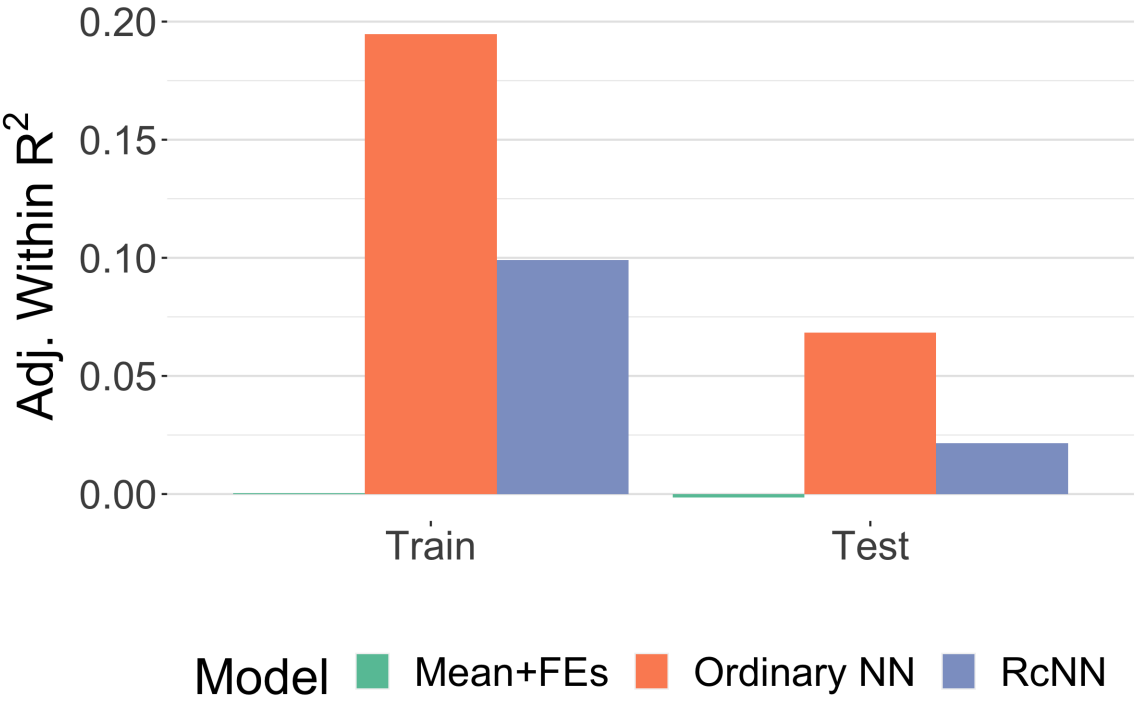


The main results shown in Figures 3 and 4. Figure 3 shows the adjusted within-R^2 of each method in the training and test sample. The standard plausibly causal panel estimator performs much worse than either of the neural networks, attaining a training adjusted within-R^2 of 0.0004. This is about 240 times lower than that of the RcNN or 473 times lower than the ordinary neural network. The RcNN explains about half of the within-variance explained by the ordinary neural network in the training sample.

In the testing sample, the traditional econometric method exhibits negative adjusted within-R^2, meaning that the residual variation in mean yearly temperature net of unit and fixed effects has essentially no predictive power out of sample. Unlike the linear panel data method, the neural networks both attain positive out of sample adjusted within-R^2s. The RcNN explains about 2.1% of within-variance out of sample, much more than the traditional method explains even in sample.

The reduction in out-of-sample predictive performance of the neural networks may be due to changing sensitivity of conflict likelihood to weather or trends in conflict incidence not related to conflict. In cases such as these, the predictions may remain highly correlated with conflict incidence and we might expect relatively little

Figure 3. Within R^2 by Method



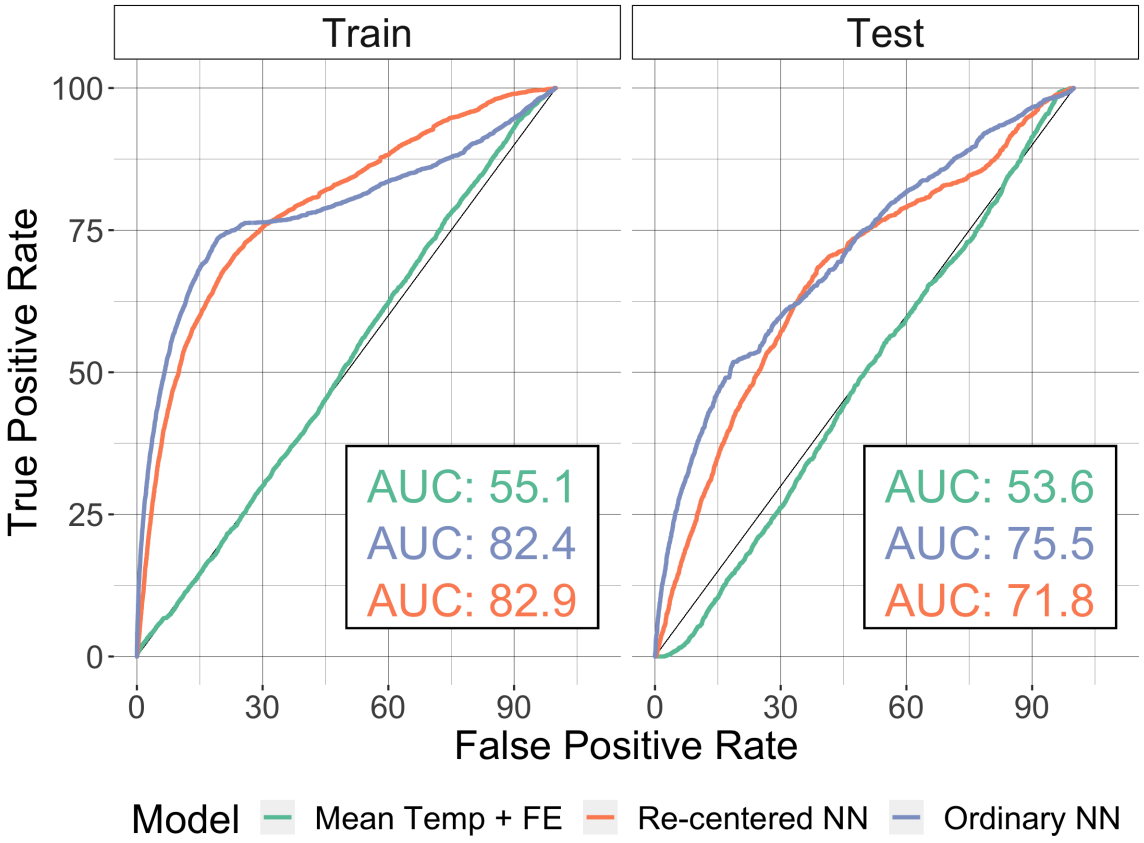
Notes: This figure shows adjusted within R^2 by method and sample. Observations from 1989-2009 were included in the training sample, while the test sample contains observations from 2010-2016.

movement in ROC curves.

Figure 4 shows that this appears to be the case. The neural networks see relatively small declines in AUC between training and test samples, with the ordinary neural network going from 82.4 to 75.5 and the RcNN going from 82.9 to 71.8. Interestingly, the RcNN outperforms the ordinary neural network in terms of AUC in the training sample, although most of the improvement comes at high false positive rates. While the RcNN has a slightly lower AUC than the ordinary neural network in the test sample, it performs slightly better over a small range of false positive rates from about 31%-45%.

In contrast to the neural networks, the traditional panel data method performs very poorly in both samples. The multi-way fixed effects estimator achieves an AUC of 55.1 in the training sample and 53.6 in the test sample. These results indicate that predictions formed from this estimator perform little better than a coin flip in distinguishing between a random conflict and no-conflict weather realization.

Figure 4. AUC by Method



Notes: This figure shows the ROC curve for each method by sample. Observations from 1989-2009 were included in the training sample, while the test sample contains observations from 2010-2016.

5 Conclusion

The results presented in this paper suggest that machine learning in general, and causal machine learning in particular, can be helpful in predicting civil conflict due to weather variation. While standard panel data methods employed in the literature succeed in finding a statistically significant relationship between weather shocks and conflict, they do not perform well in a prediction sense. Ordinary neural networks are very effective at using weather variation to predict conflict incidence, but do not have a causal interpretation. This lack of causality is of more than just academic interest. Causal estimators are demonstrated to have better out-of-sample predictive performance under covariate shift. Given that climate change induces covariate shift in the context of this problem, preserving causality may be desirable.

The RcNN architecture exhibits predictive performance comparable to, and at times better than, a traditional neural network while implementing a causal estimator. It is clear that causal estimates of the effects of weather shocks can be used to predict conflict reasonably well. However, the comparison of the RcNN performance to the multi-way fixed effects estimator shows that allowing for a flexible relationship between is essential for obtaining quality predictions.

Finally, while this empirical exercise focused on civil conflict as an outcome, the RcNN methodology is much more broadly applicable. Any statistical relationship between an outcome and weather can be implemented very similarly to the RcNN in this paper by simply replacing the outcome and potentially the weather variable. More generally, whenever the econometrician knows the distribution of the covariates of interest, or can reasonably approximate it, and these covariates are sufficiently high-dimensional the RcNN can be used to flexibly estimate causal effects.

References

Borusyak, Kirill and Peter Hull, "Non-Random Exposure to Exogenous Shocks: Theory and Applications," Technical Report w27845, National Bureau of Economic Research, Cambridge, MA September 2020.

_ and _ , "Efficient Estimation with Non-Random Exposure to Exogenous Shocks," *Working Paper*, December 2021.

Burke, Marshall B., Edward Miguel, Shanker Satyanath, John A. Dykema, and David B. Lobell, "Warming increases the risk of civil war in Africa," *Proceedings of the National Academy of Sciences*, December 2009, *106* (49), 20670–20674.

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

Funk, Chris, Pete Peterson, Seth Peterson, Shraddhanand Shukla, Frank Davenport, Joel Michaelsen, Kenneth R. Knapp, Martin Landsfeld, Gregory Husak, Laura Harrison, James Rowland, Michael Budde, Alex Meiburg, Tufa Dinku, Diego Pedreros, and Nicholas Mata, "A High-Resolution 1983–2016 Tmax Climate Data Record Based on Infrared Temperatures and Stations by the Climate Hazard Center," *Journal of Climate*, September 2019, *32* (17), 5639–5658.

Harari, Mariaflavia and Eliana La Ferrara, "Conflict, Climate, and Cells: A Disaggregated Analysis," *The Review of Economics and Statistics*, October 2018, *100* (4), 594–608.

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

Mach, Katharine J., Caroline M. Kraan, W. Neil Adger, Halvard Buhaug, Marshall Burke, James D. Fearon, Christopher B. Field, Cullen S. Hendrix, Jean-Francois Maystadt, John O'Loughlin, Philip Roessler, Jürgen Scheffran, Kenneth A. Schultz, and Nina von Uexkull, "Climate as a risk factor for armed conflict," *Nature*, July 2019, *571* (7764), 193–197.

Newell, Richard G., Brian C. Prest, and Steven E. Sexton, "The GDP-Temperature relationship: Implications for climate change damages," *Journal of Environmental Economics and Management*, July 2021, *108*, 102445.

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