Appendix for: They're Still There, He's All Gone: American Fatalities in Foreign Wars and Right-Wing Radicalization at Home

Richard J. McAlexander* Rob Williams† Michael A. Rubin‡

09 August, 2021

Contents

1 Videos

This section provides more detail on the videos connected to the metadata. While the metadata forms the basis of our analysis, inspecting the videos serves to bolster our claims that they are original content recorded and uploaded by Parler users. 0.71% of videos have a date before Parler's launch in August, 2018, suggesting that the data are not perfect, but errors are minimal. The last date in the archive is 2021-01-10, indicating that no videos were uploaded after Parler's hosting was shut down on that date.

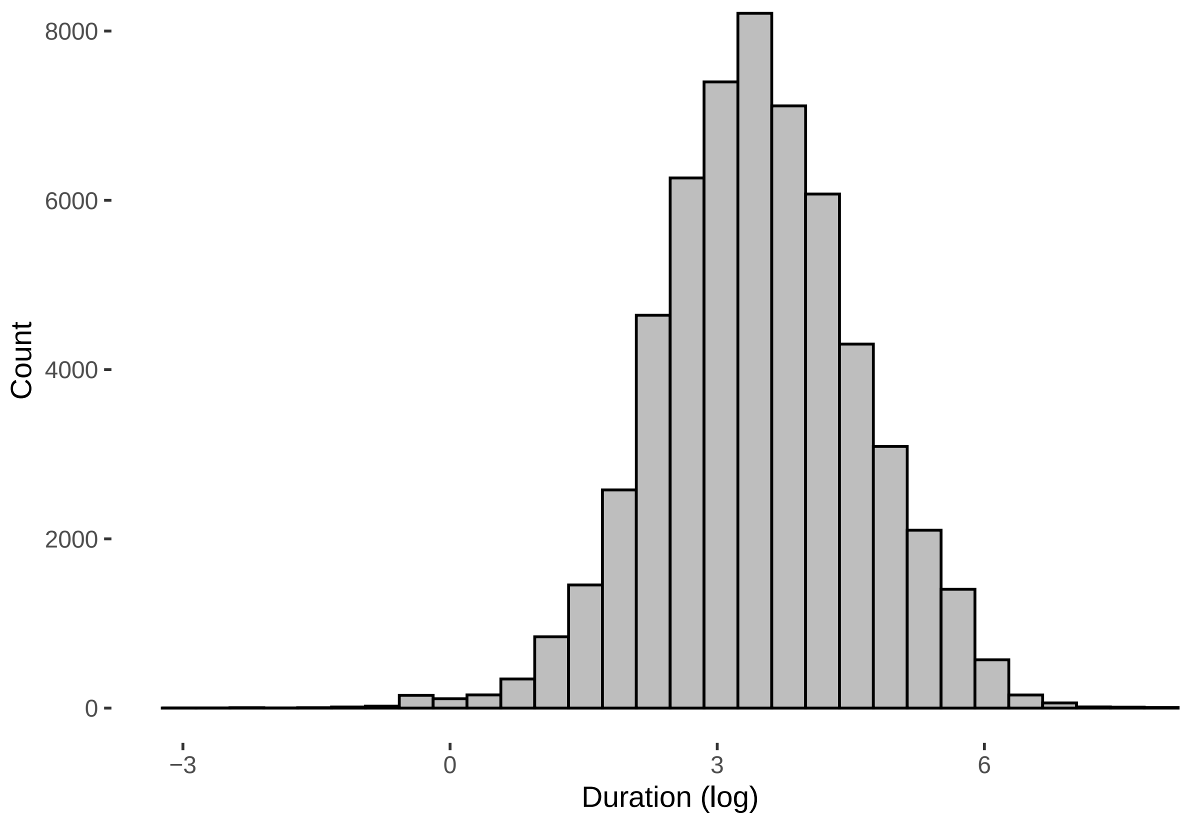


Figure 1: duration

1.1 Duration

68 of the geocoded videos were not available on the server.[1] The mean length of a video is 59.90 seconds, while the median is 31.10.[2] Figure 1 presents the histogram of logged video duration in seconds.

1.2 Originality

98.00% of geocoded videos are shot in a vertical orientation, where the image's vertical dimension is greater than its horizontal one. Professionally produced videos rarely use this 'portrait' orientation because it conflicts with both the horizontal arrangement of our eyes and over a century of established videographic convention (Pogue 2018). Thus, it is likely that a majority of the 57,181 videos in our sample represent original content recorded by the uploader.

To further assess this claim, we conduct a human review of a random 1% sample of videos with geolocation information. Videos were coded on three criteria: whether they were filmed as a 'selfie' (subject of the video clearly holding a device and using a front-facing camera to record), whether they contained images of a television or other screen, and whether they were at a political rally.[3]

12.59% were coded as being selfies, 27.62% as footage of screens, and 11.36% as being filmed at a rally or protest. These videos frequently contained footage of television new programs, suggesting that the uploader used the recording as a way to share the message within that given segment. Other videos contained images of what appear to be residential dwellings. Taken together, these patterns indicated that a large proportion of videos were shot and uploaded in members homes. As a result, using demographic and political information measured in geographic units (census block groups, counties) is inferentially valid because many of these videos were uploaded from members' homes.

1.3 Location

72.31% of the 65 videos that were coded as being filmed at a rally do not appear in our analysis dataset, indicating that our same day 25 km rule is reasonably successful in excluding videos filmed at a rally. ACLED defines a protest as "individuals and groups who peacefully demonstrate against a political entity, government institution, policy, group, tradition, businesses or other private institutions" (ACLED 2019, 12), so pro-Trump rallies would not appear in ACLED unless there were counterprotesters at the event to oppose it, which likely explains the videos uploaded at rallies that were not removed by our geographic exclusion procedure.

*University of Pennsylvania, Perry World House.

†Washington University in St. Louis.

‡University of Connecticut, Human Rights Institute.

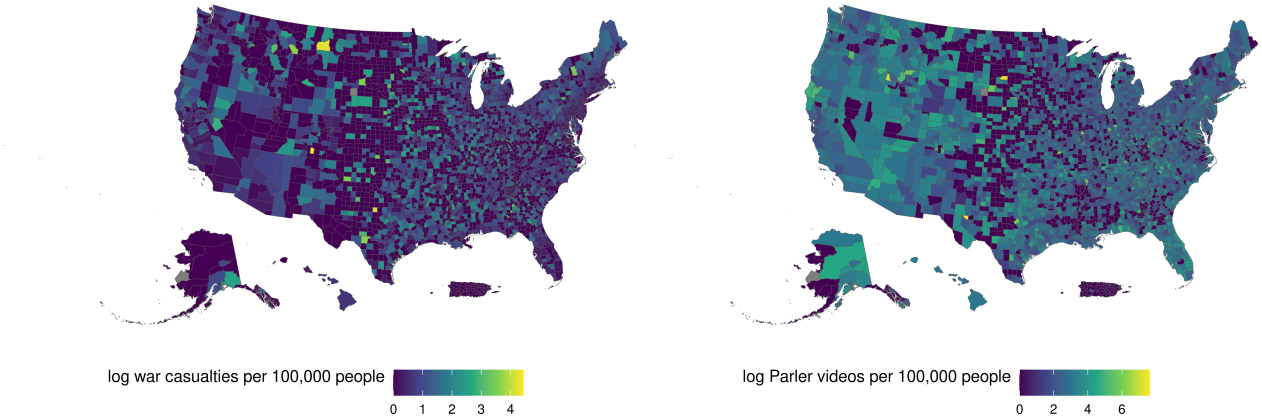
[1] These videos are included in the metadata, but it is not clear whether the archivists were unable to download them alongside the metadata or if they were removed after the fact.

[2] The data are heavily right-skewed with 7 videos over 40 minutes in length.

[3] All of the 572 videos sampled were available for download on the server.

2 Maps

This section presents the choropleths maps in the paper but with Alaska, Hawai'i, and Puerto Rico included.



(a) War fatalities (b) Parler videos

Figure 2: Choropleth map of Parler videos and war fatalities in contiguous 48 US states

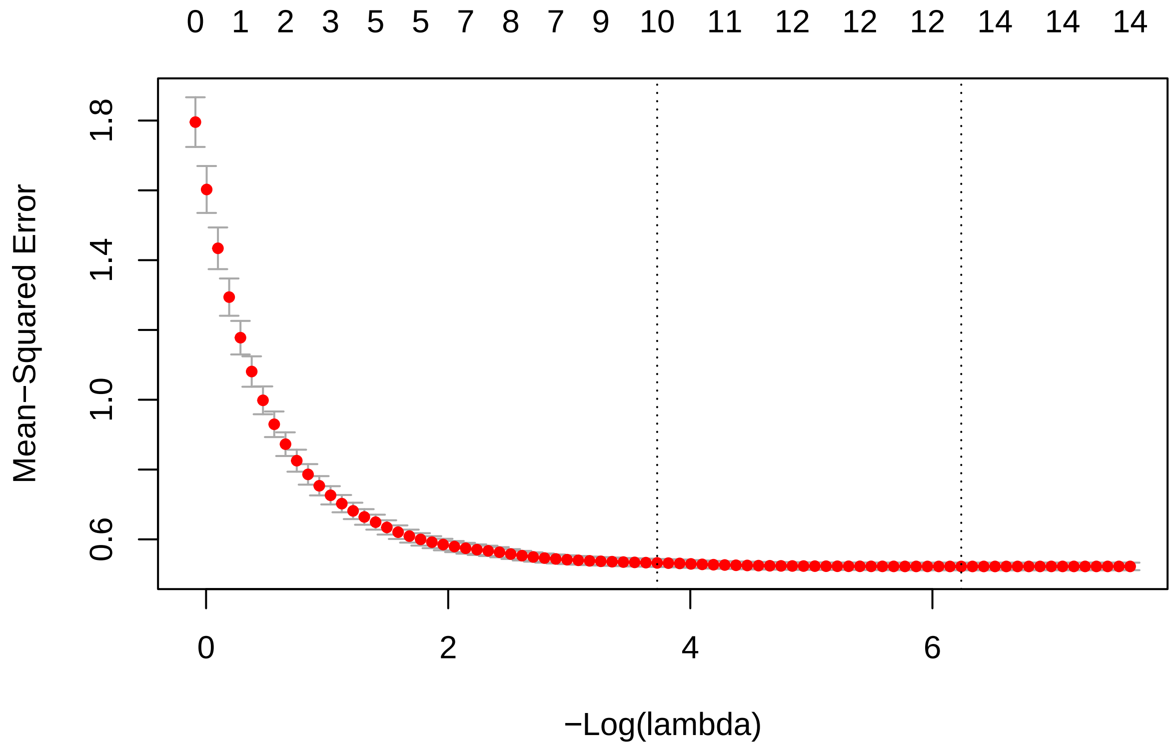
3 Predictive power

To further assess the link between fatalities in foreign wars and far-right radicalization, we evaluate the predictive power of the covariates in our main model. Using LASSO regression for covariate selection allows researchers to discover which variables in a regression have the most the predictive power Hastie, Friedman, and Tibshirani (2009).

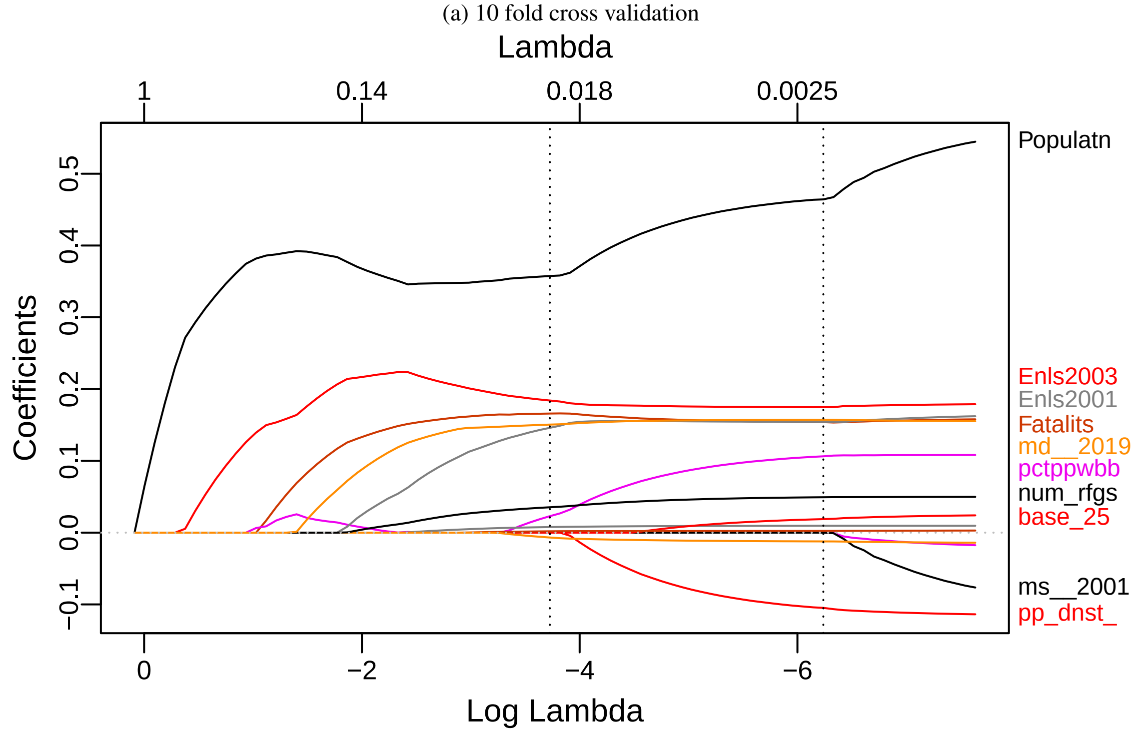
We perform 10-fold CV in Figure 3a to pick the λ value that minimizes mean squared error (λ is a tuning parameter in LASSO regression commonly chosen this way). Mean squared error (MSE) monotonically decreases as $ln(\lambda)$ becomes more negative, which means that accuracy decreases as covariates are removed from the model. By examining the order in which covariates are dropped as $ln(\lambda)$ approaches zero, we can get a sense of their importance within the model in terms of predictive accuracy.

Figure 3b depicts the order in which these covariates drop out of the model. Population (Populatn) is the last to drop out, which makes sense given the importance of population in predicting the count of Parler videos. The number of individuals enlisted in the Army since 2003 (Enls2003) is the next to last, followed by percent without broadband internet (pctppwbb). War fatalities (Casualts) is next, but for most of the range of $ln(\lambda)$, its coefficient is significantly greater than broadband access.

War fatalities is more important to predictive accuracy than the other demographic controls included in the main model, indicating the strength of the association between it and Parler videos. At the largest value of $ln(\lambda)$ where MSE is still within one standard deviation of the minimum MSE (denoted by the left dotted line in Figure 3), War Fatalities is the largest coefficient.



(a) 10 fold cross validation



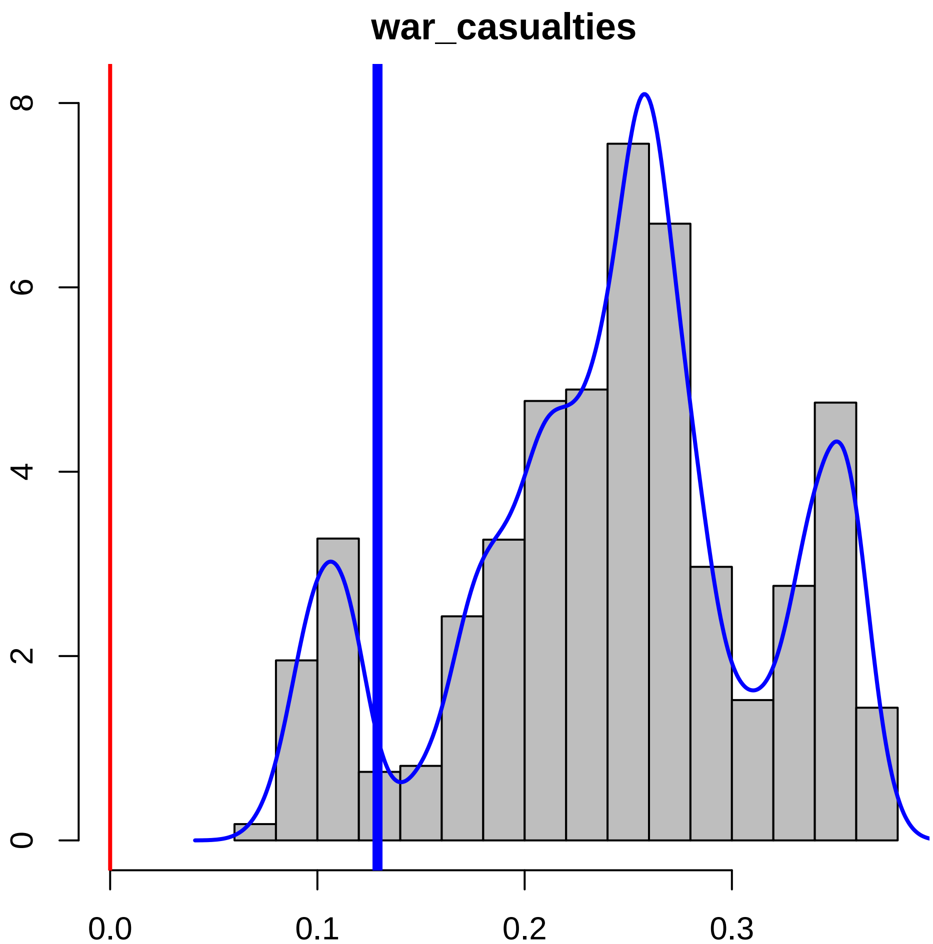
(b) Coefficient estimates

Figure 3: LASSO

4 Extreme Bounds Analysis

We also run an extreme bounds analysis to probe the robustness of our results and see if the relationship between war fatalities and parler participation is affected by the model specification. We run 8,474 regression models with 38 control variables randomly selected for inclusion. Since both our independent and dependent variables are counts, we force all models to include population as a control. The results are remarkably robust: the average effect for war fatalities on parler video uploads at the county level is 0.223. The minimum coefficient value is 0.065 and the maximum coefficient value is 0.376. In all of the models ran, the effect of war fatalities is positive and statistically significant at the .95 level.

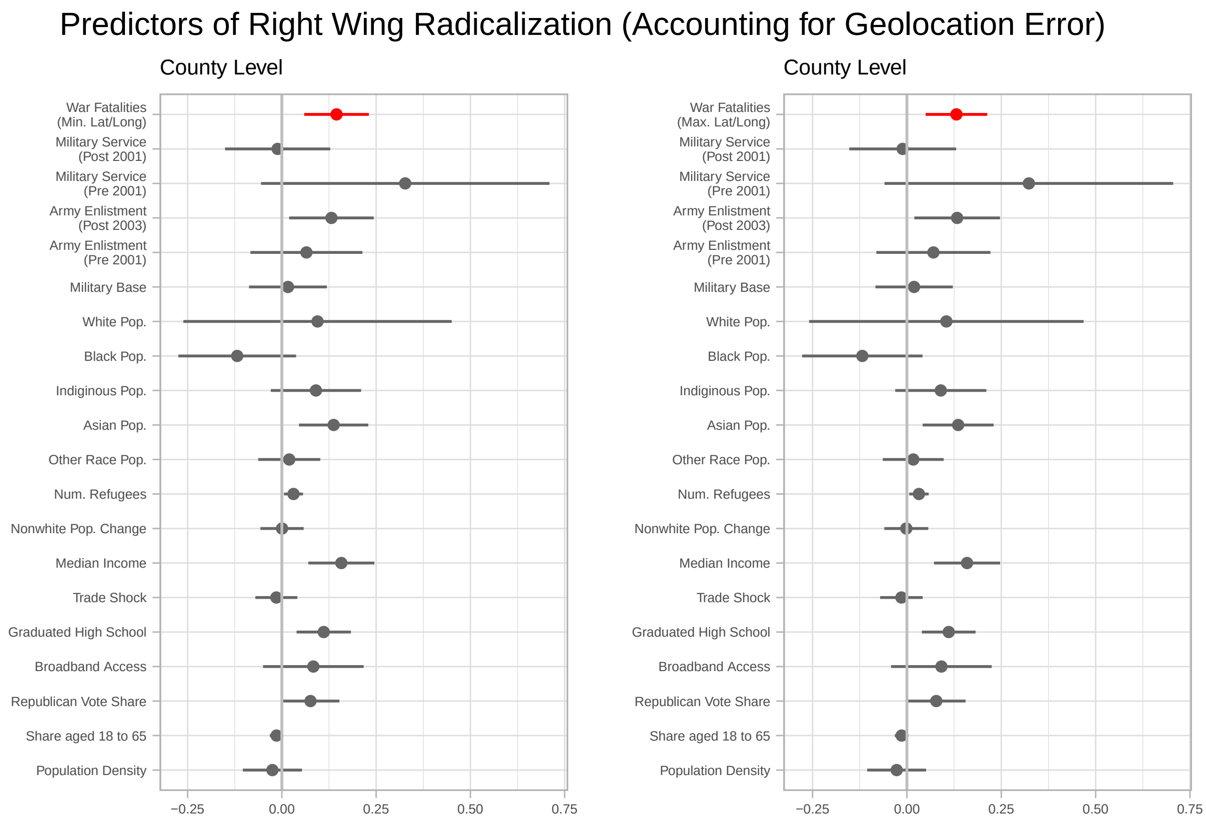
Below we present a histogram of the effects estimated for war fatalities and a list of additional variables we included. The thick blue vertical line indicates the coefficient in our county level model presented in the main text.



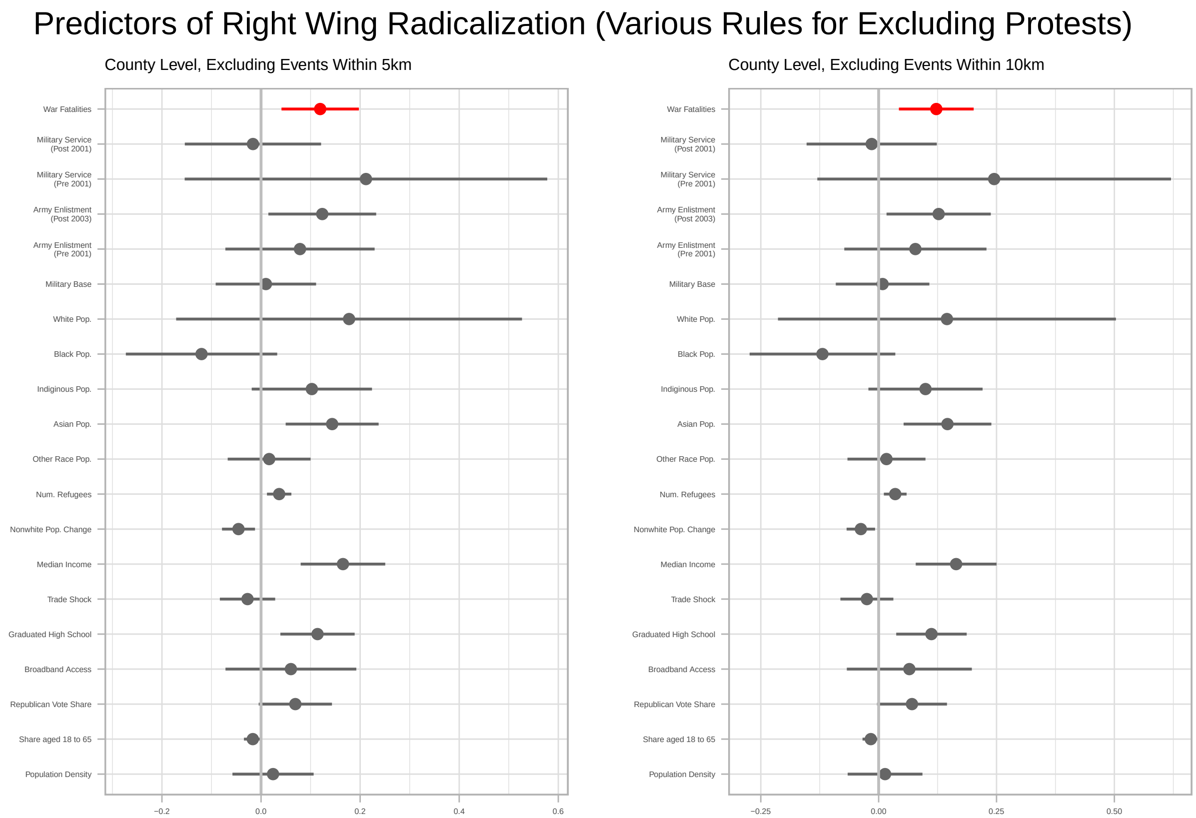
Variables Included in Extreme Bounds Analysis
War Fatalities
Military Service (After 2001)
Military Base
Median Income
Military Service (Before 2001)
Army Enlistment Rate
Num. Refugees
Army Enlistment (Before 200)
Layoffs due to Trade
Union Membership
Dem. Exposure to Dems.
Rep. Exposure to Dems
Dem. Exposure to Reps.
Rep. Exposure to Reps.
Trade Assistance
Republican Vote Share (2016)
Ethnic Fractionalization
Percent Unemployment
Import Shock
Export Shock
Broadband Access
Male Percent
Rural Percent
High School
Some College
Dentists Per Cap
Mental Health Providers Per Cap
Poor/Fair Health Per Cap
Adult Obesity Percent
Diabetes Percent
Smoking Percent
Drinking Percent
Car Death Percent
Age 18 to 65 Percent
Share Long Commute Driving
Share Drive Alone to Work
Share Housing Problems
Physical Inactivity Index
Food Insecurity

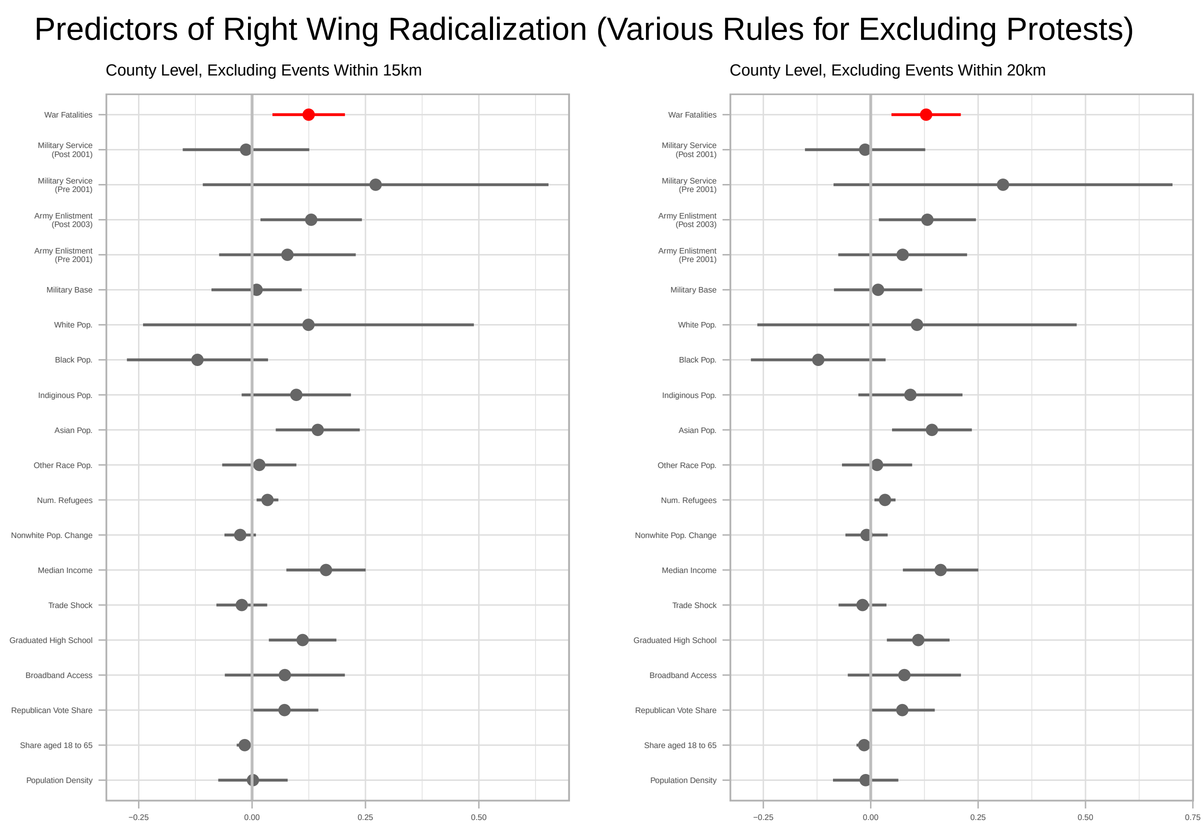
5 Combined Models for Mismatch Between Data.

As stated in the main document, there is geolocation error for about 8% of our fatalities. This results from ambiguity in matching hometown names that may have multiple matches within a state (for example, two entries for Springfield in Virginia). In our main models, we aggregate the fatalities with multiple matches by taking the mean of the latitude and longitude of the entries matched. We do this because many of the matches are for entities with overlapping territorial jurisdictions. However, in cases where this does not happen, we may be assigning a casualty to the wrong county. To address this, we run two models where we take the maximum and minimum value of the latitude and longitude for entities that have multiple locations. In effect, one model will be run on data where a casualty is attributed to Springfield, VA located on the northwestern most part of the state, while the other model will be run on data where a casualty is attributed to Springfield, VA located on the southeastern most part of the state. The results are presented below and almost identical to our main results. This is to be expected, as the geolocation error is quite small.



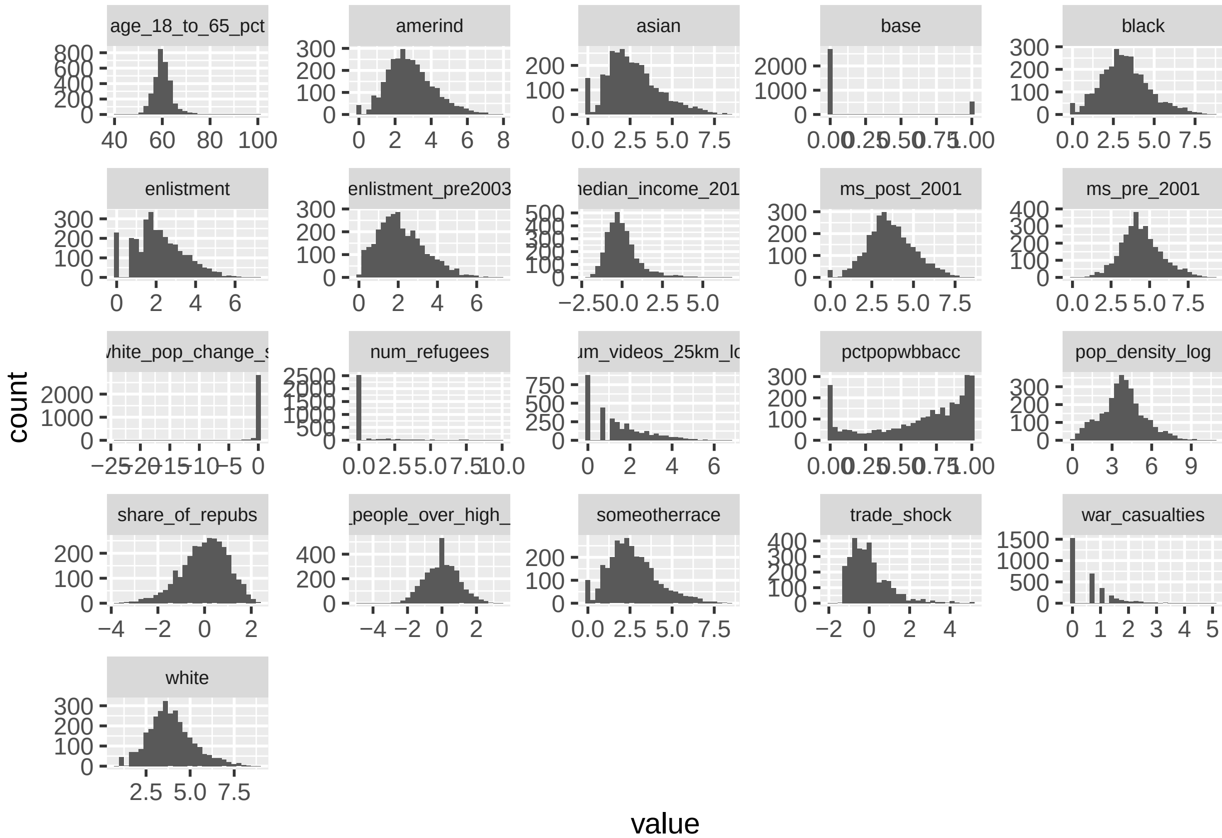
6 Models with Different Exclusions for Protests/Events



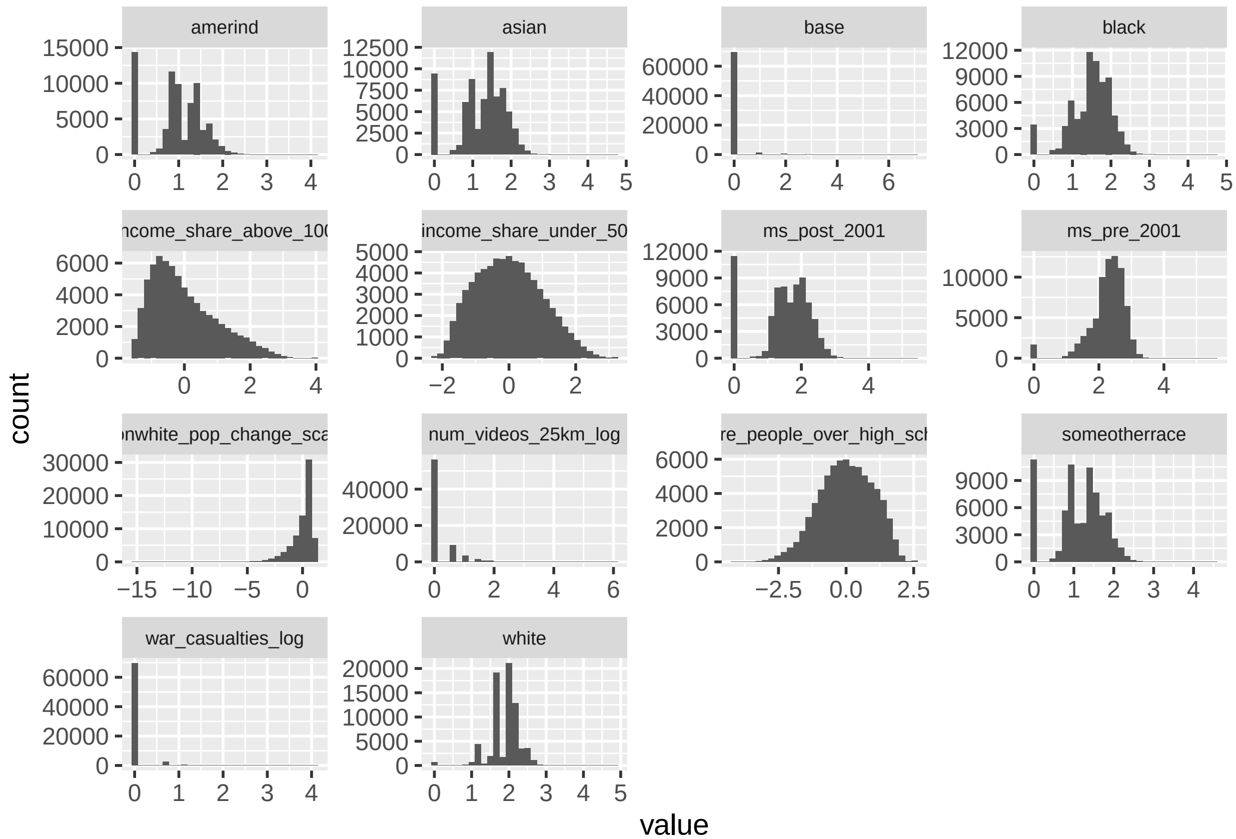


7 Histogram of Variables

7.1 County Level Variables



7.2 Tract Level Variables



References

ACLED. 2019. "Armed Conflict Location & Event Data Project (ACLED) Codebook."

Hastie, Trevor, Jerome H. Friedman, and Robert Tibshirani. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second edition, corrected 12th printing. New York: Springer.

Pogue, David. 2018. "Video Looks Most Natural Horizontally, but We Hold Our Phones Vertically." *Scientific American*, March.