FB Analysis

Kim, Brew, Zilinsky

- After removing duplicate texts, we have 43611 observations for the US House, and 16828 observations for the US Senate.

Pre-processing

- We download the Facebook ads for both chambers of US Congress.
- We remove duplicates (in the ad-text sense)
- We merge in covariates, namely:
  - Party ID
  - Office
  - State
  - Vote share (a proxy for competitiveness)
- We create a corpus with the above mentioned document variables appended as covariates.
- We tokenize the text, removing English and Spanish stopwords, lower-casing the words, a removing most punctuation.
- We currently do not apply a stemmer.
- We create a document feature matrix, stacking the House and Senate data into one large matrix.
- The dimensions of this matrix are: 60439 x 46400.
- That is, there are 46400 features (types). The total number of tokens is 1835742.

Some of the analyses are at the ad level, but we also produce candidate level analyses. When each "document" is a candidate, and we prepare a document-feature matrix whose dimensions are: 721 x 46400.

The median number of words produced per candidate is 1110, the average number is 2546.1, but the maximum is 74402, so we will generally report "% of words" for a given candidate that meet some condition (e.g. belong to a particular dictionary).

Mentions of salient topics/figures

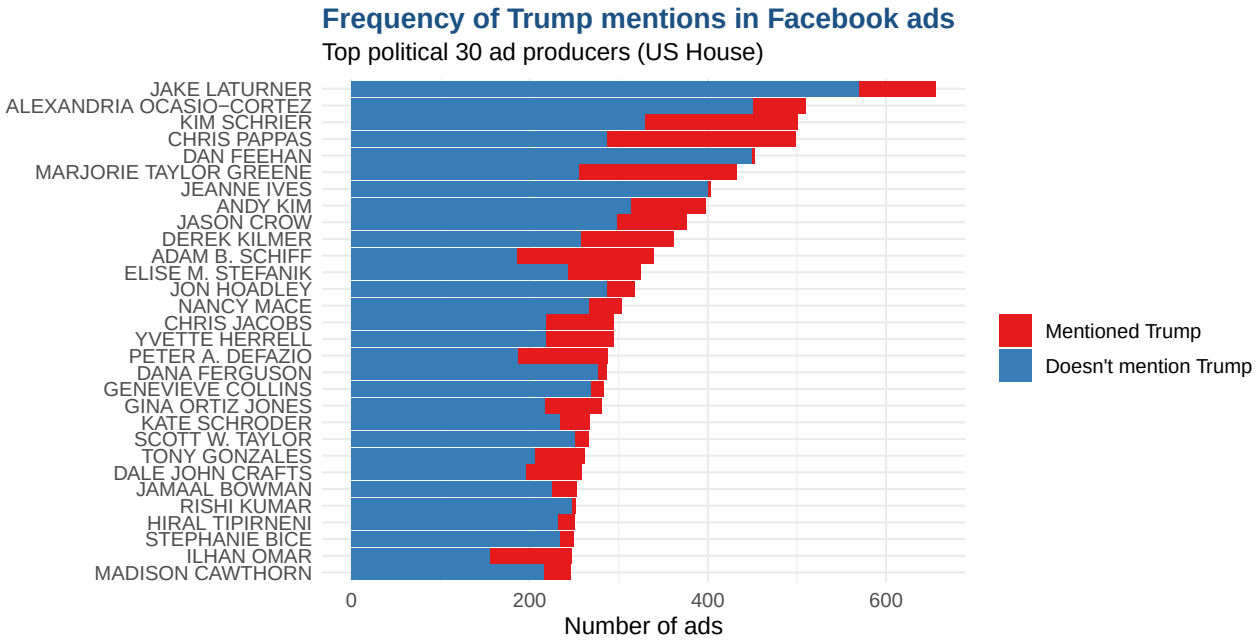


Figure 1: Mentions of Trump

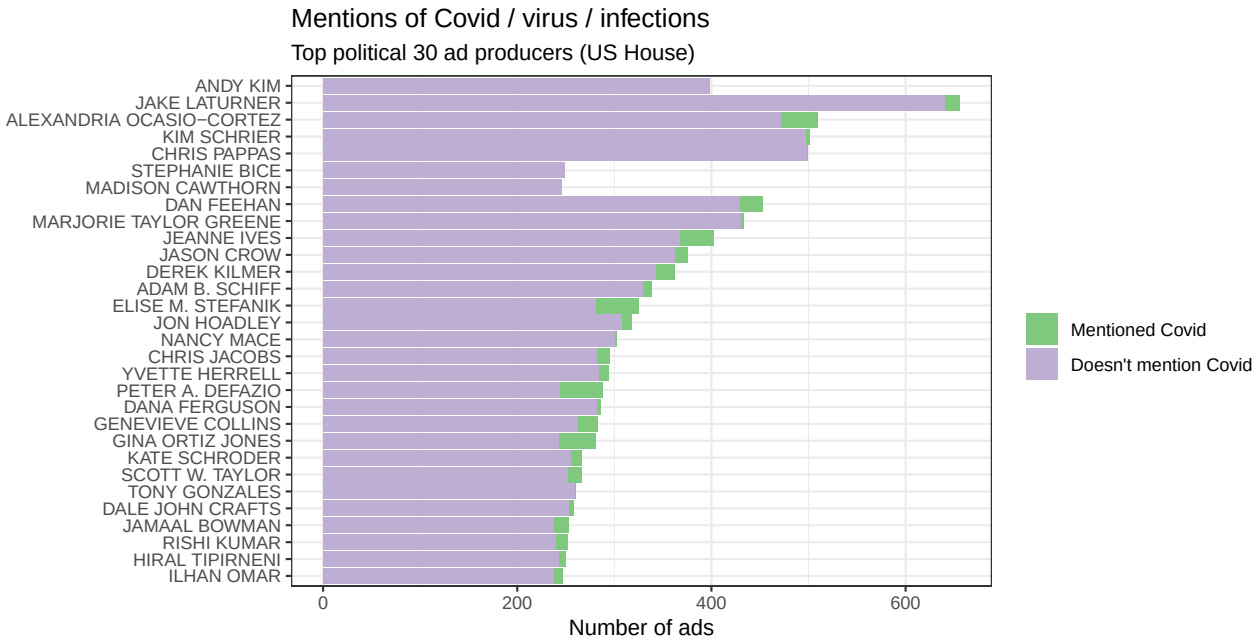


Figure 2: Mentions of Covid

Usage of salient words, broken down by PID

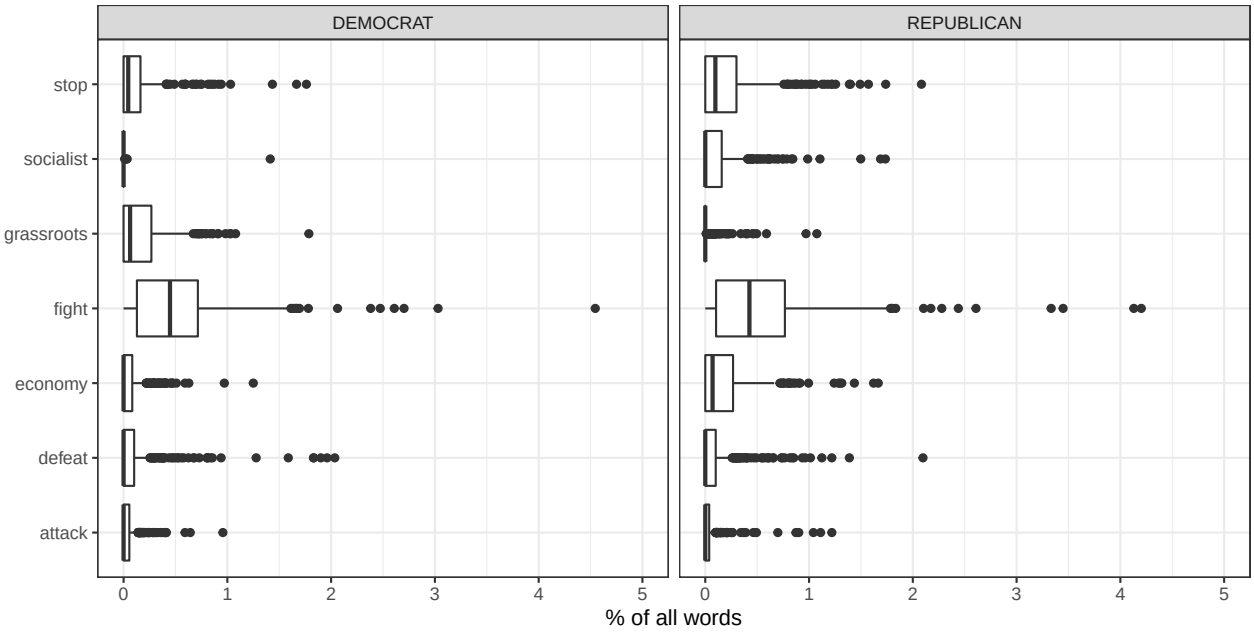


Figure 3: Candidate-level usage of selected words (% of all words used in ads)

Who produced most words for FB ads?

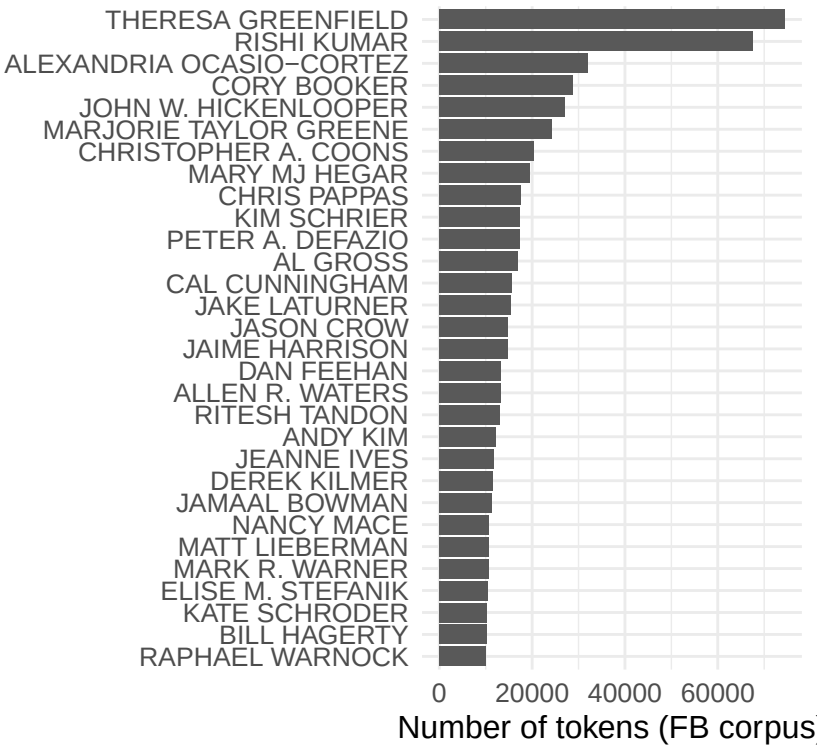


Figure 4: Number of tokens by the 30 most prolific candidates

Top features (20 most frequently occurring tokens)

All FB ads

```
##      help       can         $       now        us      need  congress  campaign
##     18198     15392     14848     13992     13590     13045     11491     10767
##     today     trump     fight    senate      back   support    people      chip
##     10538     10067      9586      9237      9006      8508      8446      7930
##      vote president      make      join
##      7789      7572      7329      7323
```

Democrats

```
##     help      can        $      now     need       us campaign    today
##    12627    11707    10854    10318     9956     9879     8374     7064
##   senate congress    fight     chip   people     back    trump     make
##     6934     6887     6424     6250     6188     5711     5573     5520
##     take     join     just  support
##     4966     4787     4622     4613
```

Republicans

```
##         help    president     congress        trump            $      support
```

```
##         5568         5077         4596         4494         3992         3889
##           us          now          can        today         back         vote
##         3677         3668         3659         3456         3291         3275
##        fight         need conservative        stand         join    democrats
##         3160         3084         2583         2579         2527         2418
##         like     campaign
##         2398         2389
```

Dictionary anlysis

Trolling words by ad type

When a donation link is included, the language is on average more aggressive:

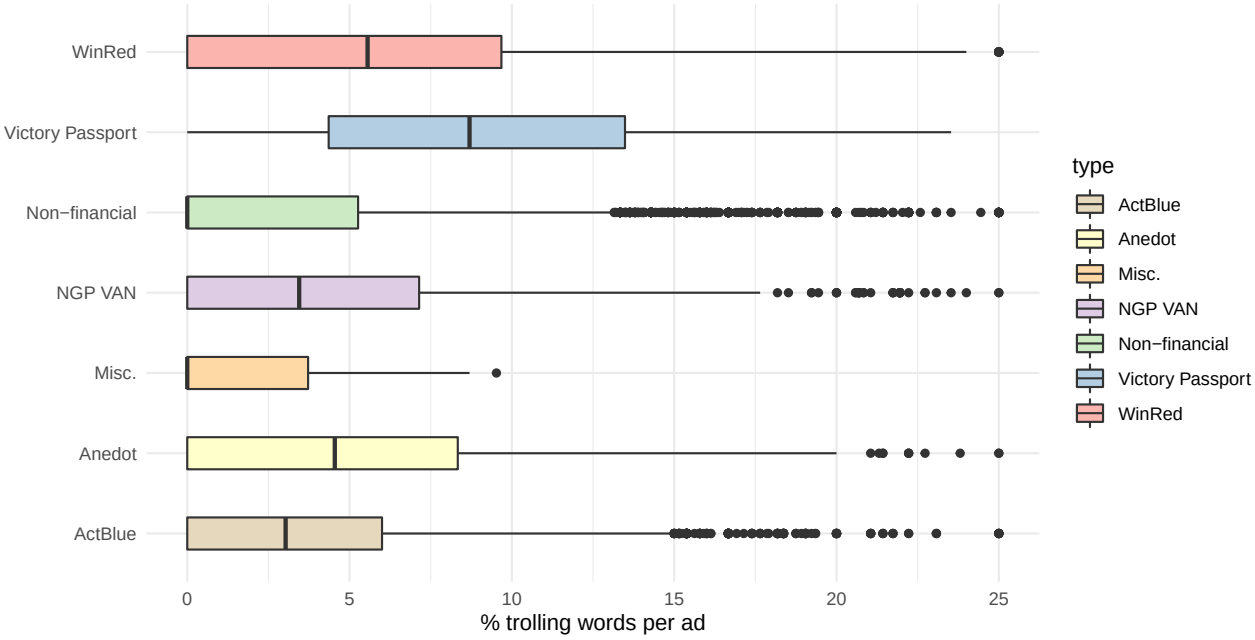


Figure 5: Percent of trolling words per ad, broken down by ad type

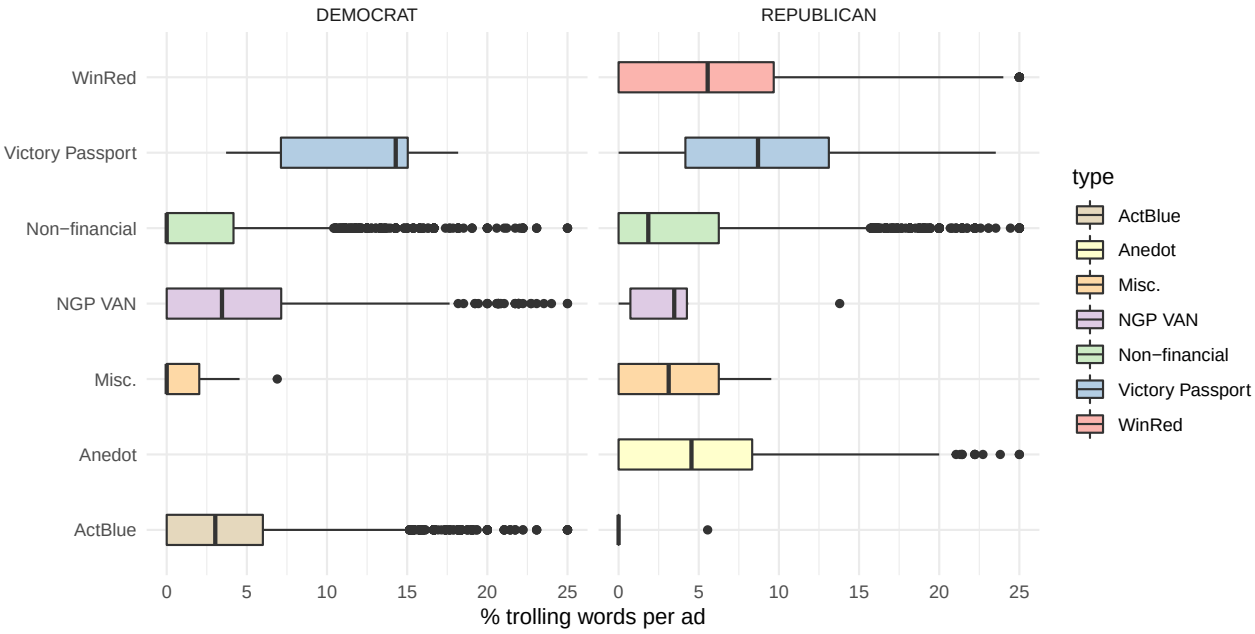


Figure 6: Percent of trolling words per ad, broken down by ad type and by Party ID

Trolling words [candidate-level analysis - PROPRTIONS]

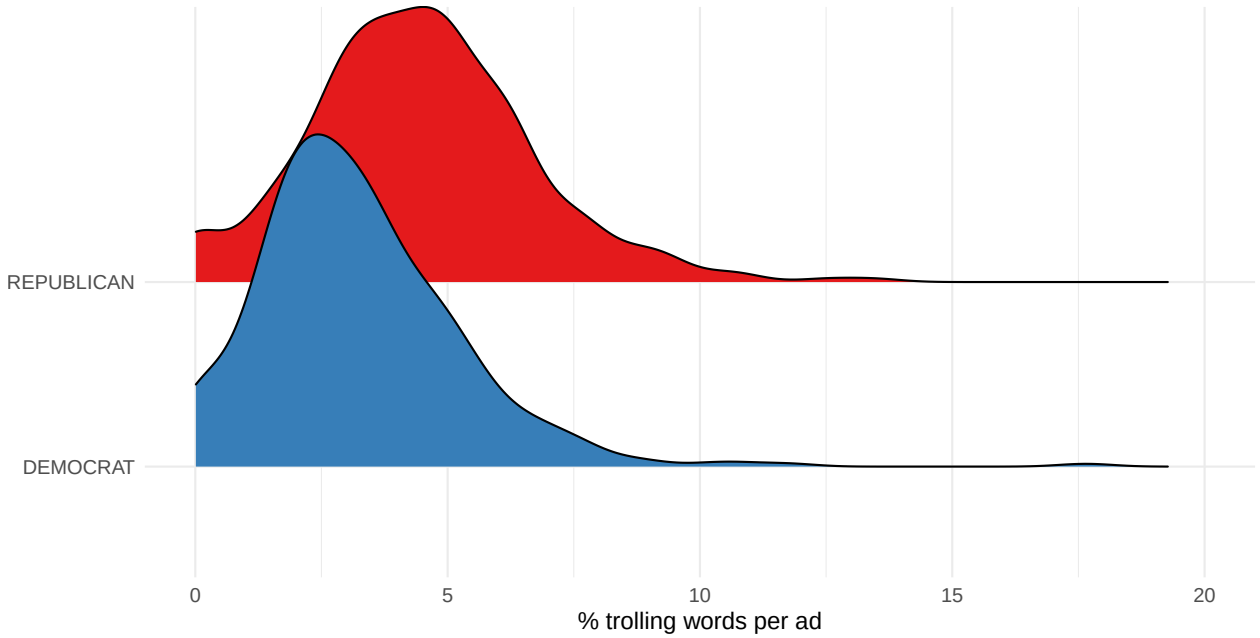


Figure 7: Distribution of candidate-level average of trolling usage (broken down by Party ID)

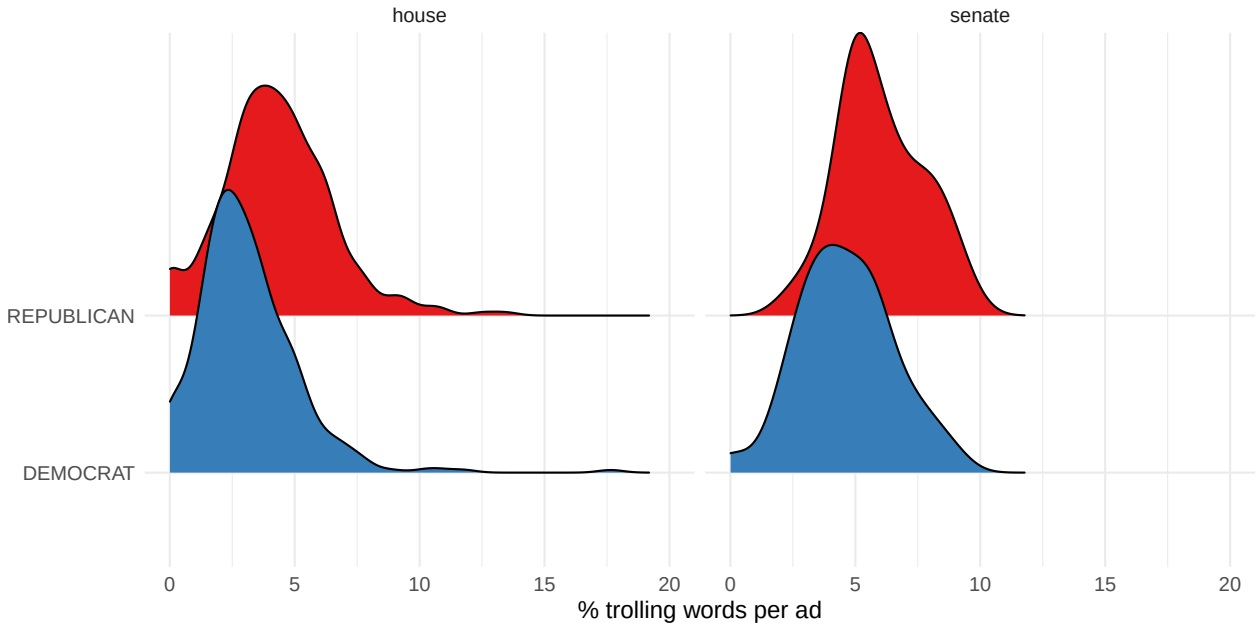


Figure 8: Distribution of candidate-level average of trolling usage (broken down by Party ID and chamber of Congress)

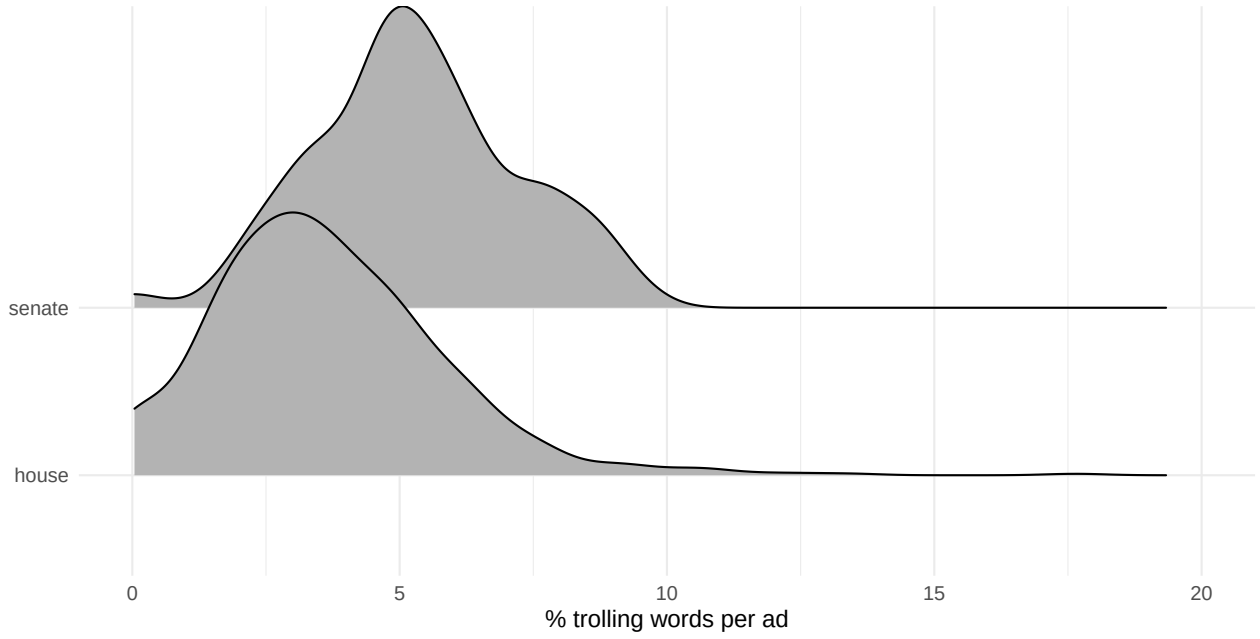


Figure 9: Candidate-level average usage of trolling words, broken down by chamber

Trolling words [ad-level analysis]

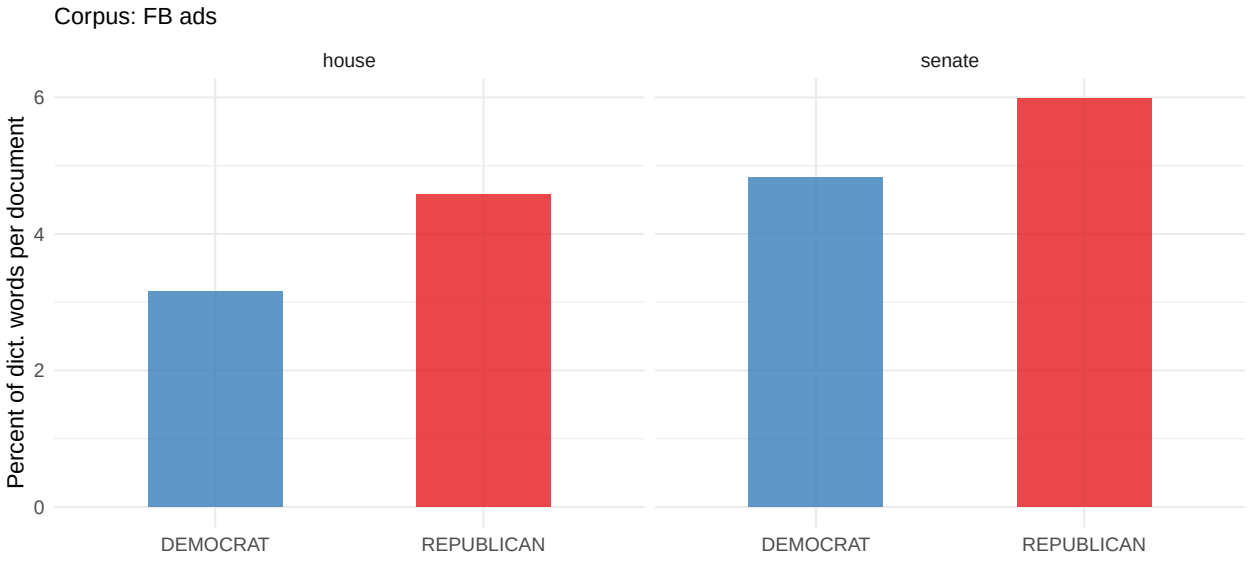


Figure 10: Average proportion of trolling words per ad, broken down by party ID and chamber

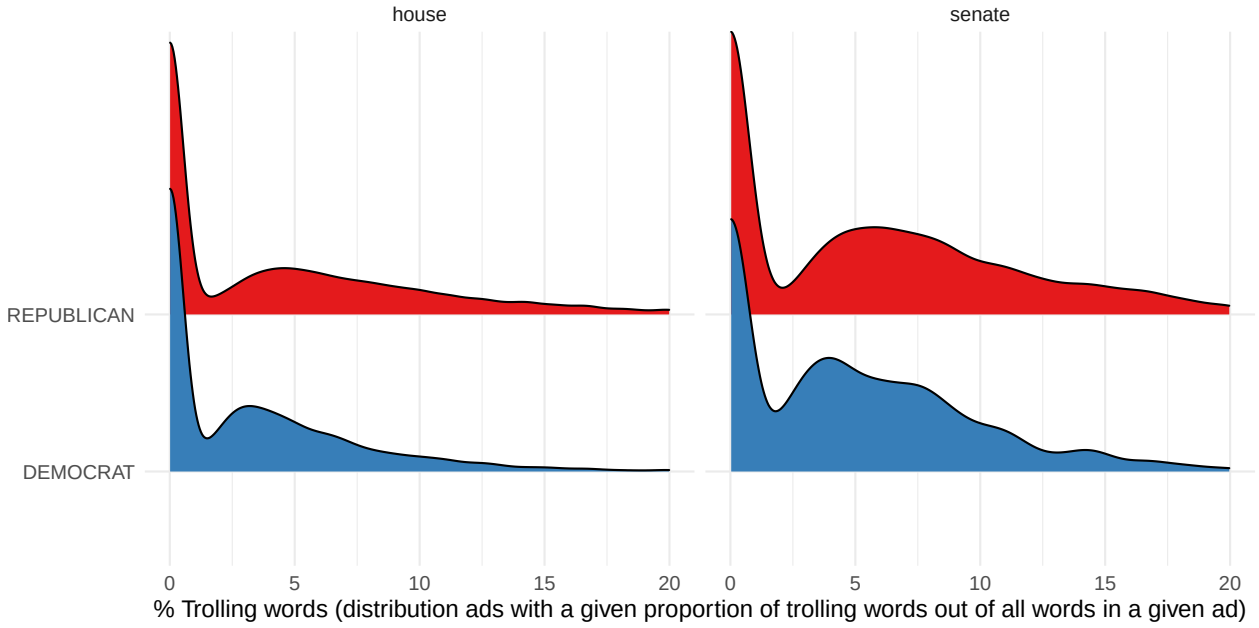


Figure 11: Distribution of trolling words per ad, broken down by party ID and chamber

Trolling words [ad-level analysis - COUNTS]

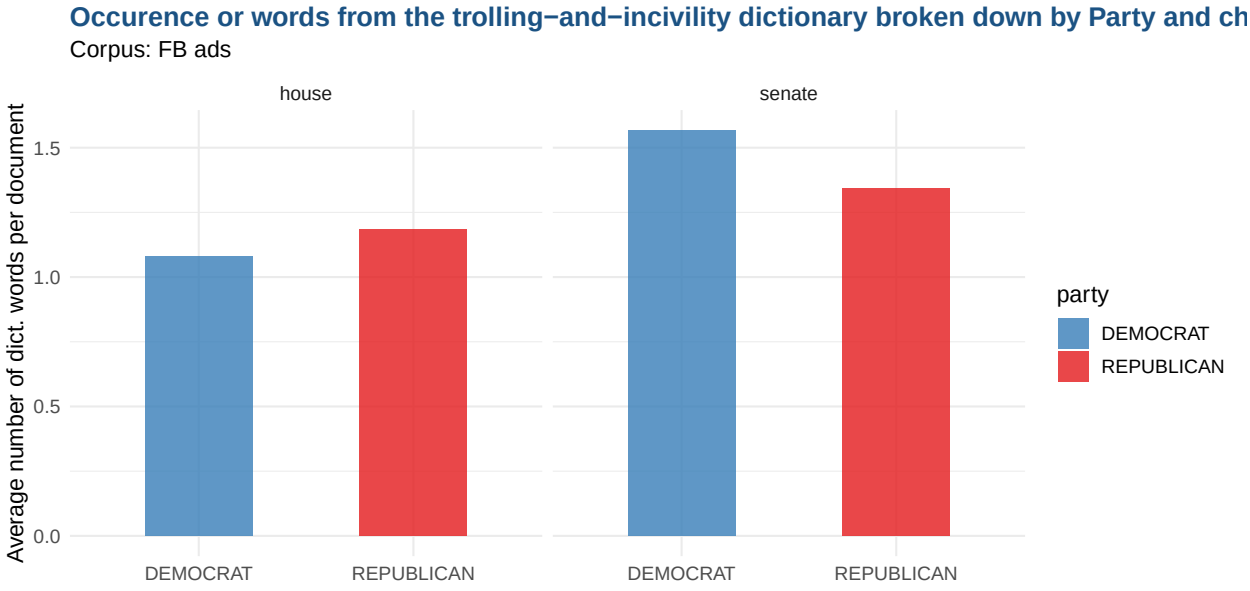


Figure 12: Average number of trolling words by party and chamber

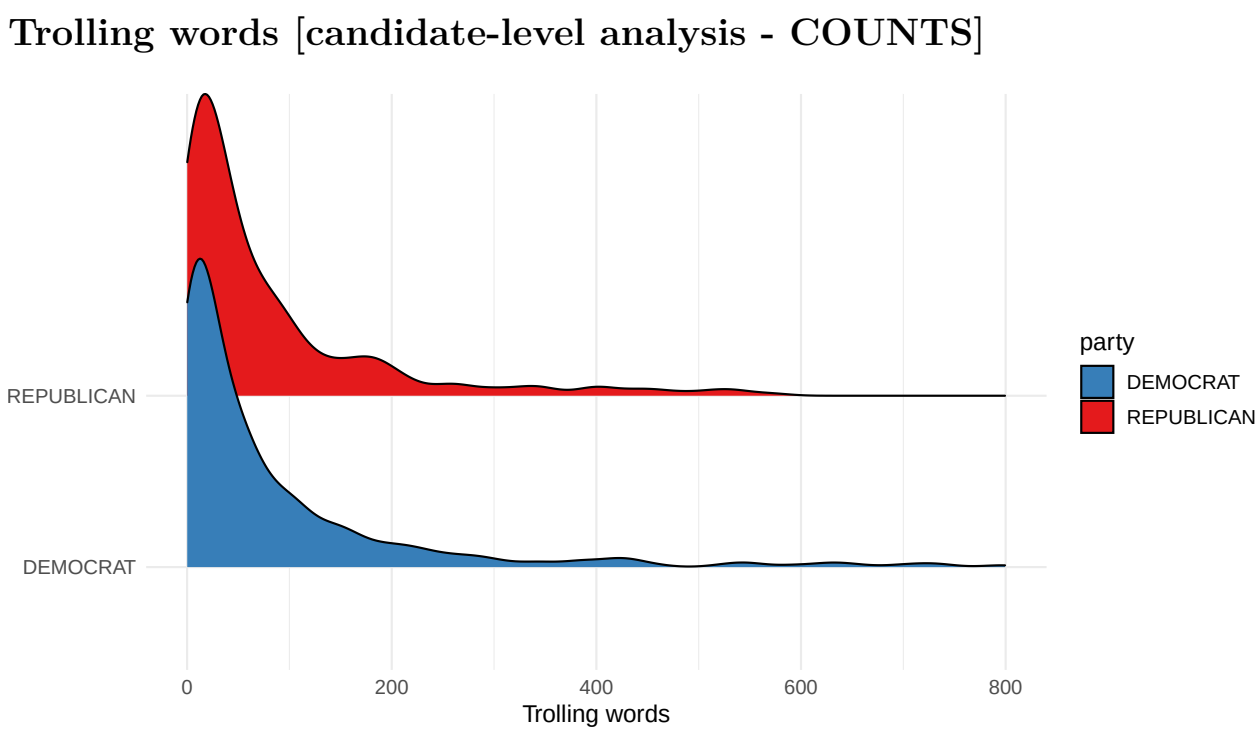


Figure 13: Distribtion of candidate-level trolling words counts (totals)

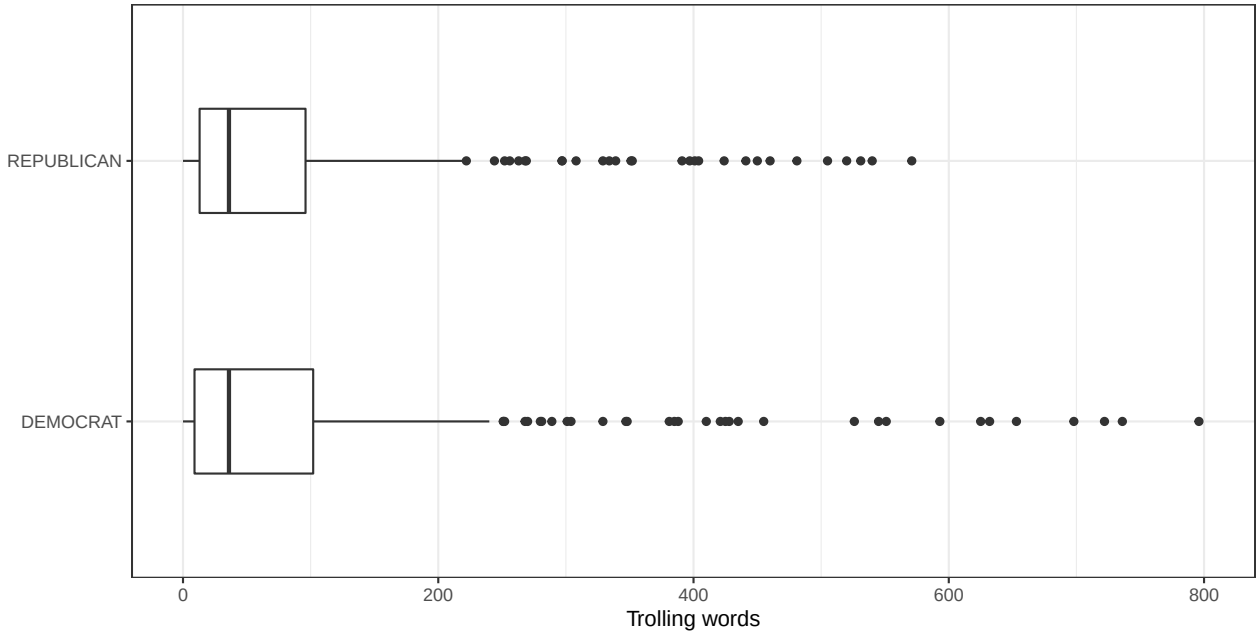


Figure 14: Distribtion of candidate-level trolling words counts (totals)

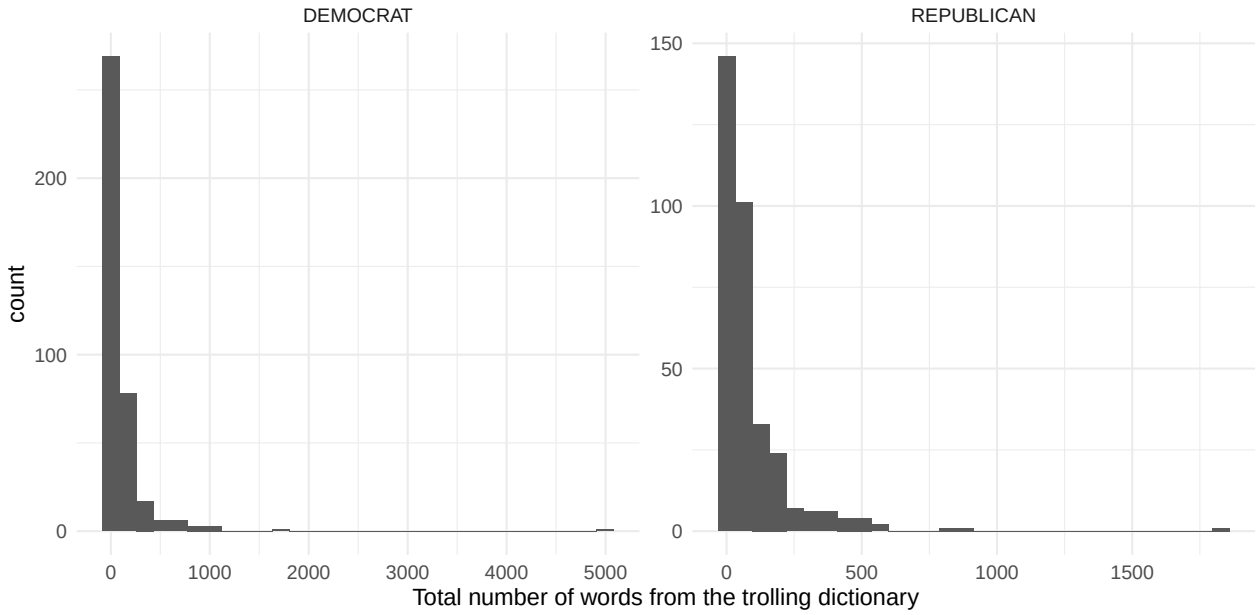


Figure 15: Total number of trolling words in the corpus produced by candidates

Moral foundations [candidate-level analysis]

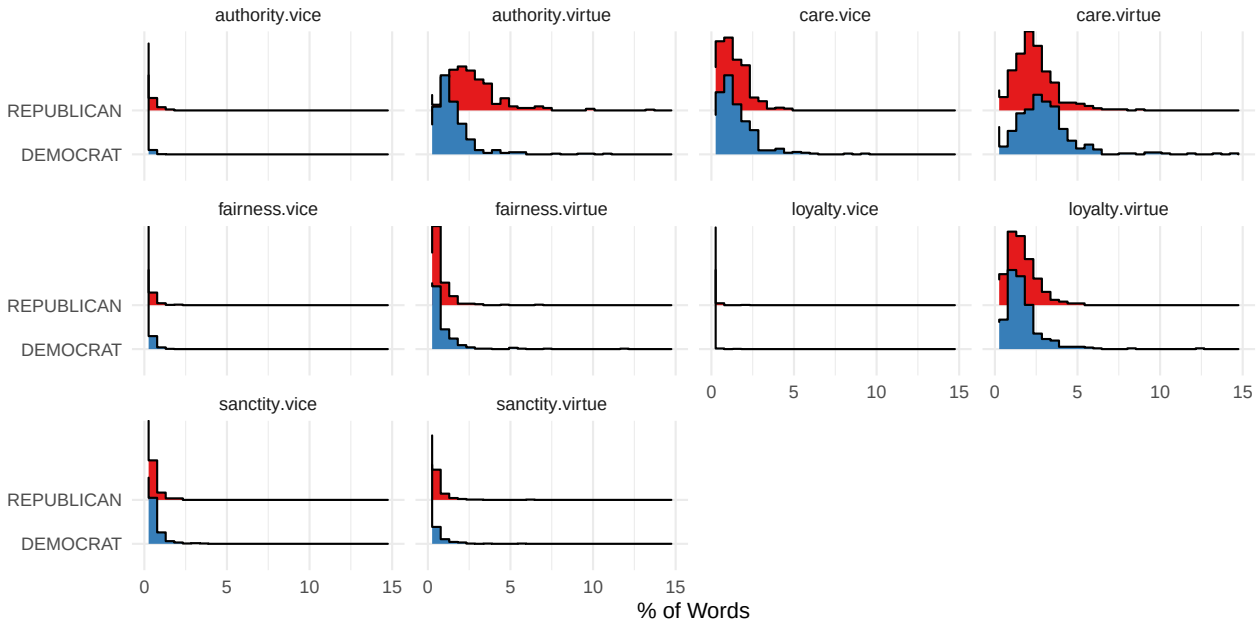
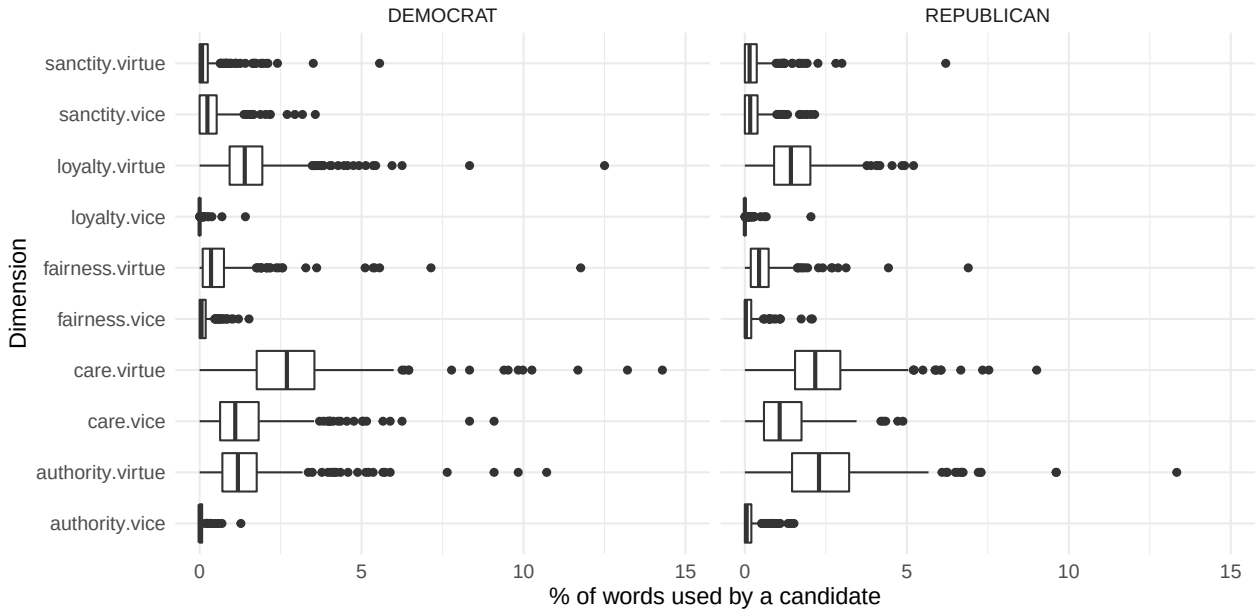


Figure 16: Distribution of Moral foundations words in FB ads, by party



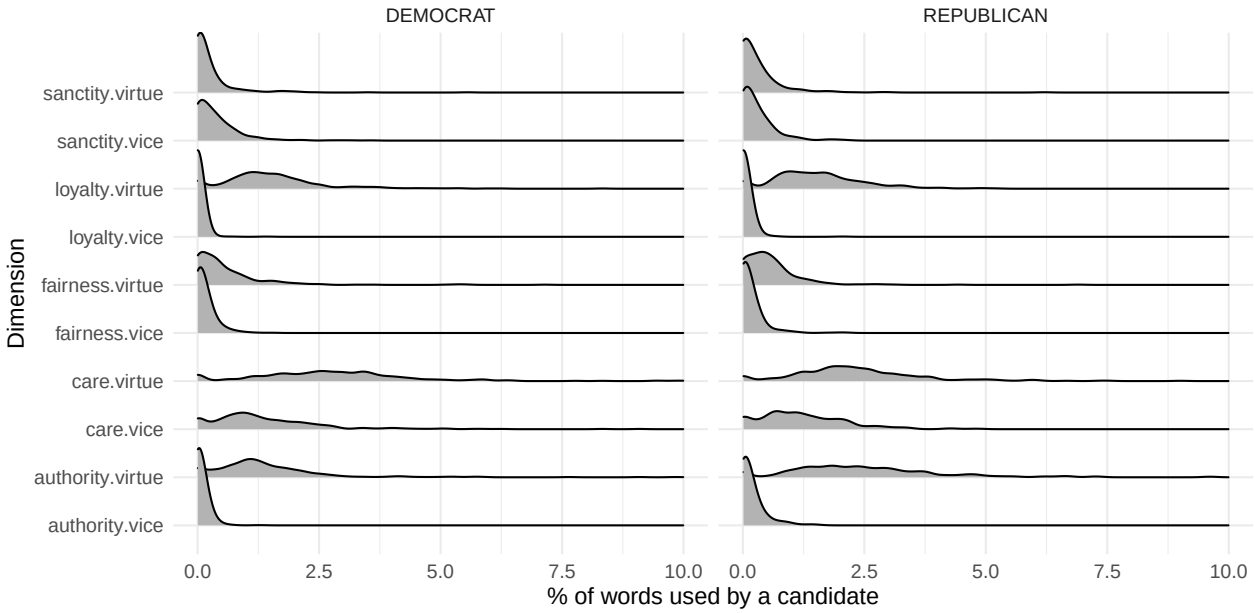


Figure 17: Candidate-level usage of moral words (MFD)

Table 1: Average Usage of words across candidates, broken down by Party ID

Average Usage of the dimension Per Candidate (in %)	DEMOCRAT	REPUBLICAN
authority.vice	0.06	0.16
authority.virtue	1.41	2.53
care.vice	1.37	1.22
care.virtue	2.93	2.33
fairness.vice	0.13	0.14
fairness.virtue	0.65	0.56
loyalty.vice	0.01	0.03
loyalty.virtue	1.57	1.54
sanctity.vice	0.38	0.27
sanctity.virtue	0.24	0.30