

# Appendix:

## Mobilization, Persuasion, and the Partisan Fallout of the Gender Gap in U.S. Voting

Updated November 26, 2018

### NOTES ON ESTIMATION

#### *Categorical outcome variable*

The model described in our paper begins by counting up the number of votes (measured using ANES vote choice) that Republicans and Democrats obtain through different pathways. To compare survey waves with different sample sizes, we normalize the terms in the model by dividing each term by the sample size in their respective survey waves. After normalization, the terms in our model represent mutually exclusive proportions of the ANES sample. Because we have multiple exclusive proportions to estimate, we use a Categorical outcome model to estimate all proportions simultaneously.

We begin by constructing a categorical variable that indexes all combinations of party identification (Democrat, Republican, or unaffiliated), gender (man or woman), and vote choice (Democrat, Republican, or “other,” where “other” includes nonvotes as well as non-major-party votes). This totals 18 categories. We then use the sample data to estimate the probability of each category in the population of U.S. electorate. Operationally, we define the multichotomous outcome variable as  $y_i$  which takes any one of 18 possible values. We assume that  $y_i$  is categorically distributed according to the probability of each category in the population:

$$y_i \sim \text{Categorical}(\theta), \tag{1}$$

where  $\theta$  is a vector containing the probabilities of all 18 categories, whose elements sum to 1. We estimate the probability vector for each survey wave independently of other waves.<sup>1</sup>

Table 1 contains a description and coding of all 18 outcome categories. We use this set of probabilities to calculate various quantities in the model. For instance, we estimate the proportion of the

---

<sup>1</sup>We have suppressed the subscript  $t$  for the simplicity.

Table 1: Description of Categorical Outcome Variable

Outcome Category (gender $\times$ party ID $\times$ vote choice)	Index in $\theta$
Man, Democrat, Votes for Democrat	1
Man, Democrat, Votes for Republican	2
Man, Democrat, Votes “Other”	3
Man, Republican, Votes for Democrat	4
Man, Republican, Votes for Republican	5
Man, Republican, Votes “Other”	6
Man, Unaffiliated, Votes for Democrat	7
Man, Unaffiliated, Votes for Republican	8
Man, Unaffiliated, Votes “Other”	9
Woman, Democrat, Votes for Democrat	10
Woman, Democrat, Votes for Republican	11
Woman, Democrat, Votes “Other”	12
Woman, Republican, Votes for Democrat	13
Woman, Republican, Votes for Republican	14
Woman, Republican, Votes “Other”	15
Woman, Unaffiliated, Votes for Democrat	16
Woman, Unaffiliated, Votes for Republican	17
Woman, Unaffiliated, Votes “Other”	18

electorate that consists of Democrat-identifying men as  $\theta_1 + \theta_2 + \theta_3$  (i.e. summing across possible vote choices). The proportion of Republican-identifying women who are not mobilized to cast a partisan vote would be  $\theta_{13} + \theta_{15}$ . These probabilities can also be used to calculate Democratic advantages or gender gaps in each term of the model. For example, the Democratic advantage in persuasion among women would be  $\theta_{13} - \theta_{11}$ , and the gender gap in unaffiliated voters would be  $(\theta_{16} - \theta_{17}) - (\theta_7 - \theta_8)$ .

#### *Parameter uncertainty using Markov chain Monte Carlo*

Point estimates for the terms in the model can be constructed by simply adding and subtracting the elements in  $\theta$ , but we also care about the uncertainty associated with each term in the model. Rather than analytically derive the variance of each quantity in the analysis, we use a simulation approach to proliferate uncertainty in  $\theta$  throughout the model. We use Markov chain Monte Carlo to generate a distribution of plausible  $\theta$ s, which we then use to produce a distribution for every quantities of interest in our results.

Although Markov chain Monte Carlo is commonly used for Bayesian analysis, we use it for computational convenience rather than for including prior information. We therefore estimate  $\theta$  using a flat prior, so the “posterior” distribution is proportional to the likelihood of the data. A flat

prior for a set of probabilities is a Dirichlet distribution with a concentration parameter of 1.

$$\theta \sim \text{Dirichlet}(\alpha = 1) \quad (2)$$

We estimate the model using Stan (Carpenter et al. 2016). We run 4 chains with 2,000 iterations per chain, setting aside the first 1,000 iterations that were used to tune the sampling algorithm. Following Link and Eaton (2011), we do not thin the parameter chains.

Sample weights (variable VCF0009z from the ANES cumulative file) are incorporated by weighting the log-likelihood of each observation when calculating the probability of the data. This is akin to pseudo-likelihood approaches that apply weights in regression models. The log-likelihood of the data  $\mathbf{y}$  given  $\theta$  can be expressed as follows:

$$\ell(\mathbf{y} \mid \theta) = \sum_{i=1}^N \ell(y_i \mid \theta) w_i, \quad (3)$$

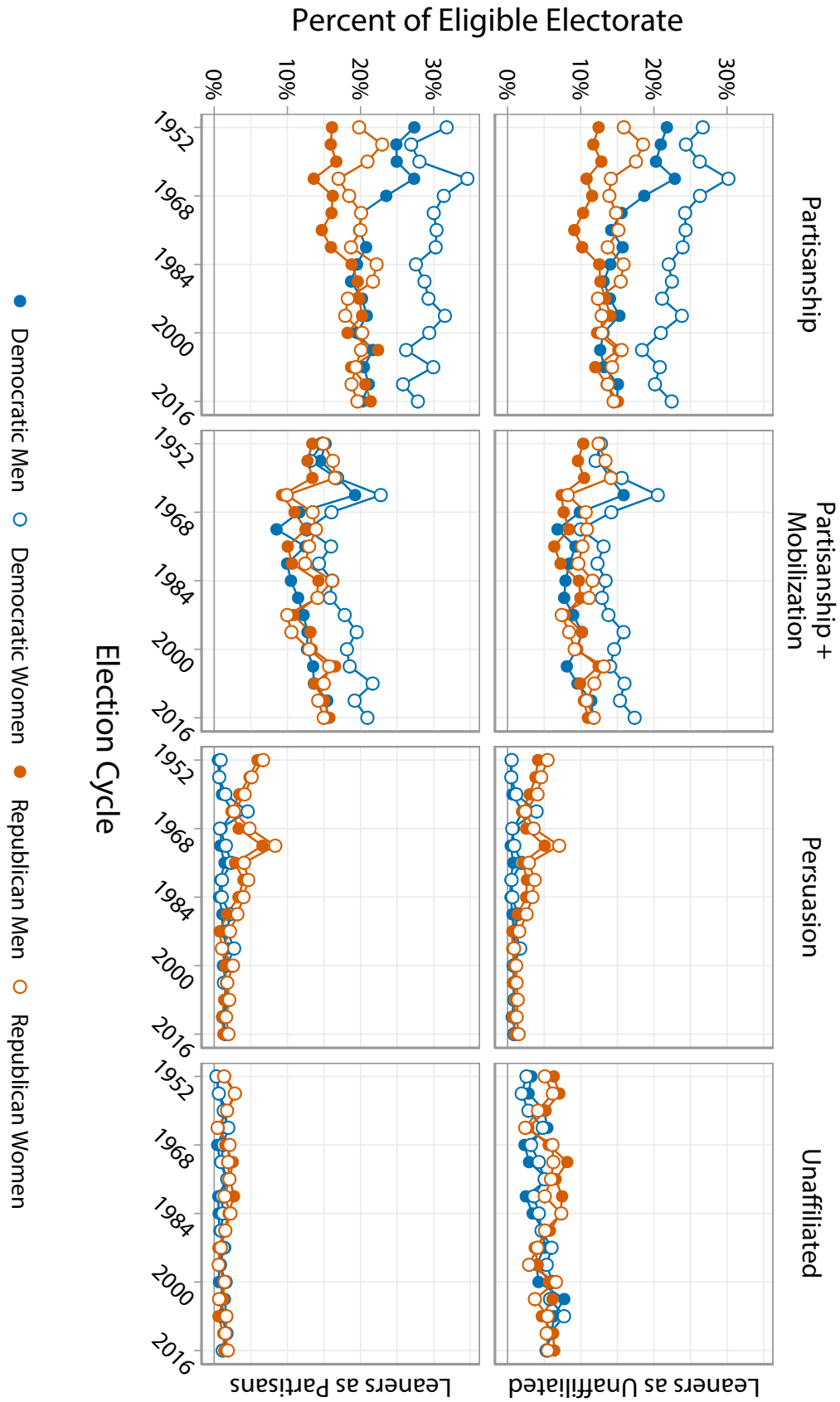
where  $\ell(y_i \mid \theta)$  is the log-likelihood of an individual  $y_i$  given  $\theta$  in the associated sample wave, and  $w_i$  is respondent  $i$ 's sample weight.

Estimating the model by weighting each observation does slow the estimation process somewhat. To speed up this routine, researchers could calculate the weighted number of respondents in each category, modeling the grouped data as a multinomial outcome variable. Multinomial modeling requires integer data, so non-integer values induced by sample weights must be rounded off. Although this method would be faster than estimating the model on individual data, our analysis uses the more intensive estimation method for the sake of exposition.

## LEANING INDEPENDENTS

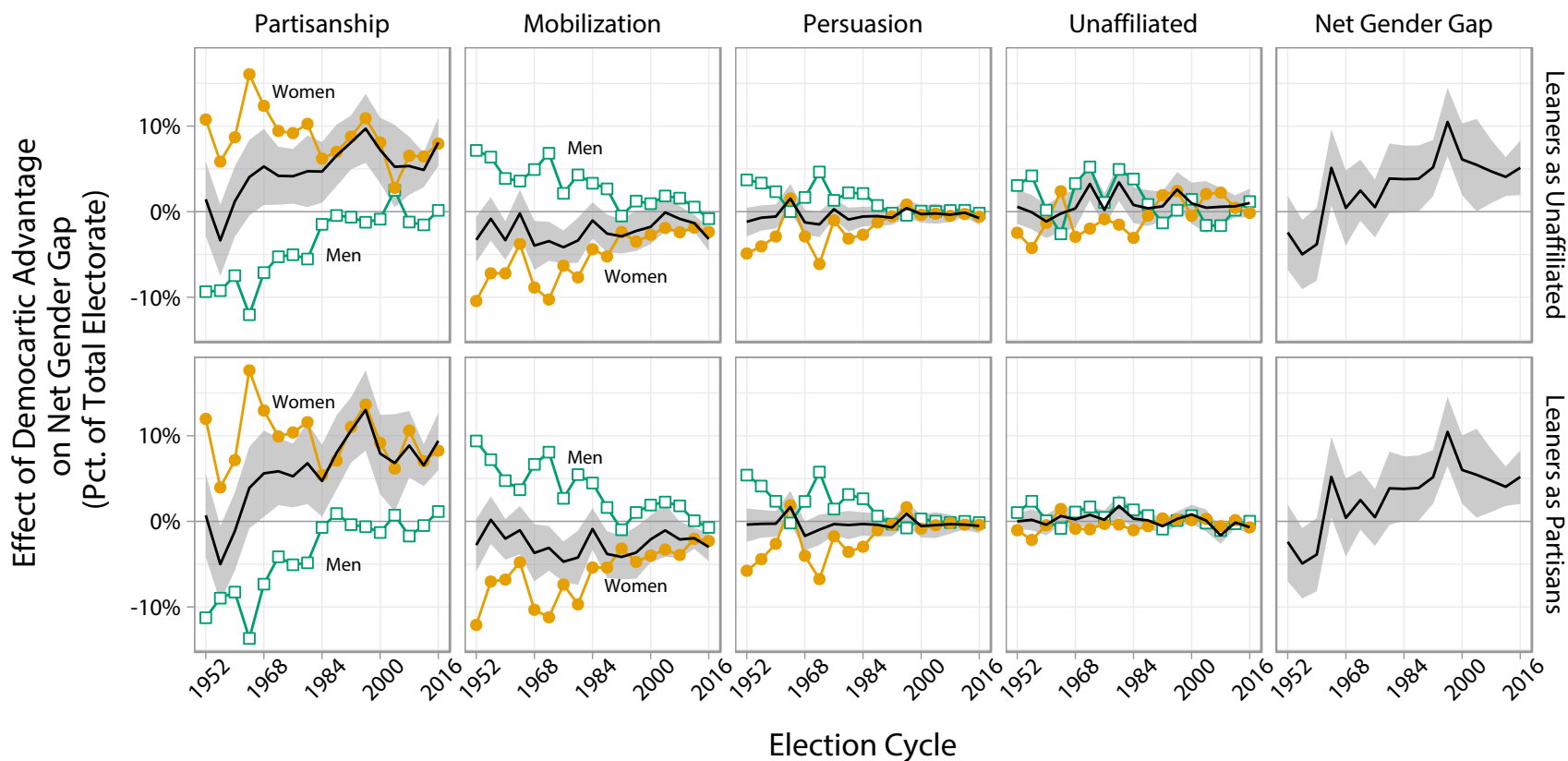
The main text of the paper presents only results with independent leaners coded as partisans. Figures 1 through 3 show results with leaners coded as unaffiliated (the top row in each figure) and as partisans (bottom row).

Figure 1: Measures of Mobilization, Persuasion, and Voting by the Unaffiliated 1952-2012



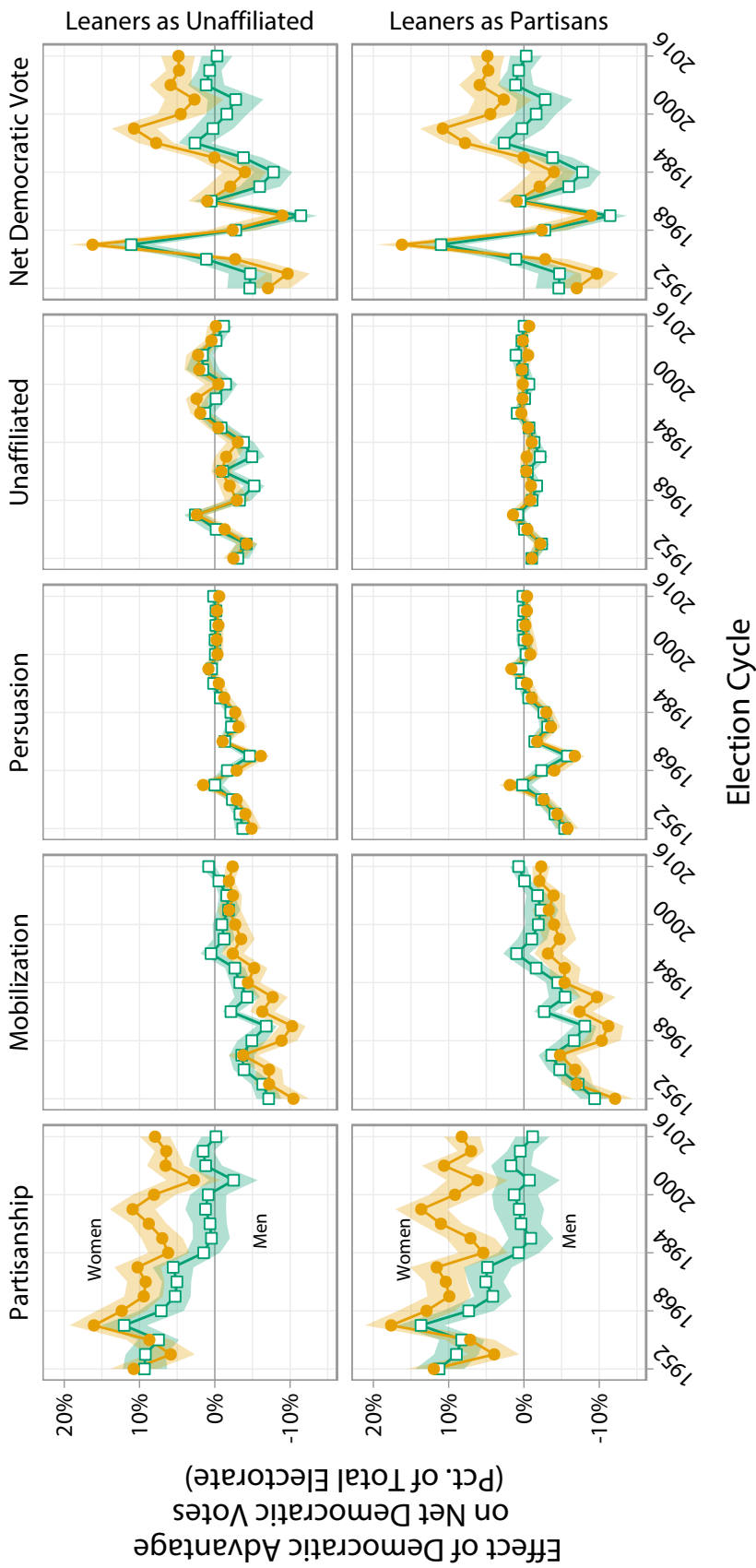
Notes: For presentational purposes, Column 2 displays the share of mobilized voters instead of the  $Mobilization_{prt}$  term from the model equation. The share of mobilized voters is the share of partisans minus the share of non-mobilized partisans. Leaning independents are coded as partisans.

Figure 2: How Each Mechanism Affects the Gender Gap



Notes: Democratic advantage among women exerts a positive effect on the net gender gap, while Democratic advantage among men exerts a negative effect. Women are shown with solid points, men with hollow points. The thick, solid line indicates the gender gap in each mechanism. The right-most panel is the sum of the net gender gap across all mechanisms.

Figure 3: Relating the Gender Gap and the Election Outcome: Gender differences in the sources of Democratic votes



Notes: Women are shown with orange solid points, men with hollow green points. The right-most panel is the Net Democratic Vote for each gender, summing across mechanisms.

## REFERENCES

- Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2016. “Stan: A probabilistic programming language.” *Journal of Statistical Software* 20: 1–37.
- Link, William A., and Mitchell J. Eaton. 2011. “On thinning of chains in MCMC.” *Methods in Ecology and Evolution* 3 (jun): 112–115.