Journal of Quantitative Description: Digital Media 1(2021), 1–94 10.51685/jqd.2021.021

Voting and Social Media-Based Political Participation

SASCHA GÖBEL

Goethe University Frankfurt, Germany

Does online political involvement reinforce or compensate participatory deficiencies at the polls? Extant survey evidence portrays online participation as a weapon of the strong, wielded by a highly politically involved, white, and affluent subset of the American electorate. Surveys face systematic sampling and measurement errors in the domain of political participation, however. In this study, I revisit this question using individual voter registration records that I integrate with observed Twitter activity. Based on a large sample that reflects Florida's voting-eligible population, I find that political involvement on Twitter is prevalent across the electorate and extends to those most likely to abstain from voting. Moreover, race and income, which are salient dividing lines in voting, do not structure social media-based political participation, and common turnout patterns for age and party subgroups are reversed, though especially among more engaged voters. These results offer a novel perspective on reinforcement theory and social media's compensatory potential for more inclusive representation. I discuss implications for political representation and future research examining political involvement.

Keywords: voting, online political participation, participatory inequalities, voter records, social media, Twitter

Göbel: sascha.goebel@soz.uni-frankfurt.de

Date submitted: 2021-02-22

Having ample opportunities for political engagement is fundamental to promoting equal political voice, representation, and legitimacy (Schlozman et al., 2018). Over the past two decades, the emergence of online means of participation, especially social media, has probably been the most important addition to citizens' participatory repertoire. Acts such as sharing or posting political content on social media have been shown to provide distinct mechanisms to communicate information, needs, preferences, and hold representatives accountable (Barberá et al., 2019; Gibson and Cantijoch, 2013; Oser et al., 2013). But who uses these opportunities?

Two competing theoretical expectations surround this question. One asserts that predictors of online participation are the same as for voting. Accordingly, online channels for political involvement benefit highly engaged voters equipped with the necessary resources, reinforcing participatory inequalities (Bimber, 1999; Norris, 2001). A more optimistic view anticipates that reduced costs and the interactive quality of online participation attracts less politically involved citizens, compensating for their inactivity at the polls (Carpini, 2000; Krueger, 2002).[1]

Extensive online participation by young adults and males, both commonly underrepresented in the US voting population, provides support for the compensation theory (Bode et al., 2014; Bekafigo and McBride, 2013). Yet survey evidence from the US in favor of the reinforcement thesis prevails (Hindman, 2009; Norris, 2001). Findings portray online participants as a subset of the most engaged part of the American voter population (Bode and Dalrymple, 2016; Gainous and Wagner, 2014; Oser et al., 2013; Schlozman et al., 2010) who are predominantly white and affluent (Best and Krueger, 2005; Schlozman et al., 2018).

Surveys face systematic sampling and measurement errors in the domain of political participation, however. Highly engaged voters are more inclined to participate in surveys than their counterparts (Brehm, 1993). If response behavior is driven by political engagement, on- and offline, it acts as a collider variable. Conditioning on this variable via removal of non-respondents introduces selection bias, which may lead us to overstate

[1] While the terminology overlaps with research on digital media use and political involvement (see Oser and Boulianne, 2020), this is not the focus of this article. Here, online participation refers to online *political* involvement, not to general internet and social media usage.

online participants' propensity to vote and the prevalence of characteristics that also predict turnout (Elwert and Winship, 2014). Additionally, political involvement is a socially desirable behavior and misreporting of both turnout (Jackman and Spahn, 2019) and political engagement on social media (Guess et al., 2019) is prevalent. Turnout overreporting, for instance, occurs in particular among politically non-involved respondents who share many characteristics with politically involved respondents, which distorts sociodemographic differences (Ansolabehere and Hersh, 2012). Further measurement errors enter through an assessment of individual voting proclivities based on only one or two elections. This conflates different voter types and is prone to exaggerate the share of highly engaged voters among online participants.

In this article, I revisit the question of whether online political participation reinforces or compensates participatory inequalities in voting. To overcome survey-specific problems, I focus on political involvement on Twitter and combine voter records from the state of Florida with Twitter accounts. For one, this yields a broad sample that closely approximates the voting-eligible population without overrepresenting voters at general elections. Second, it allows us to directly observe individuals' validated voting histories and online participation along with other political and sociodemographic characteristics. To account for electoral volatility and election-specific idiosyncrasies in the assessment of voter engagement, I use a measurement model that links observed turnout at several elections to voting propensities. I trace individuals' Twitter activity over an eight-month period surrounding the 2018 midterm elections and measure political involvement via a domain- and context-specific dictionary.

These data complement existing studies with a novel perspective on compensation and reinforcement. This additional angle is relevant, as it affords hitherto unexplored descriptive insights that build on unmediated observations of actual on- and offline behavior, which directly informs the internet's compensatory potential for political participation. Moreover, if constituents' political involvement on Twitter includes a more diverse set of participants than previously acknowledged, then this reveals potential for better substantive representation by political elites. In the broader sense this also speaks to what politicians as well as researchers can expect to learn about the electorate on social networking platforms.

Challenging the consensus in favor of reinforcement theory, I find that voters' political involvement on Twitter cannot be reduced to a highly engaged subset of the electorate. To the contrary, I document that low-propensity and irregular voters continuously participate politically on Twitter. Employing multilevel regression for subgroup estimates, I find very limited evidence that social media-based participation is structured along race or income. At the same time, results moderate expectations regarding young adults' disproportionately high online engagement, which emerges primarily among more engaged voters. Amidst survey evidence that dominantly depicts online means of participation as a weapon of the strong, this study shows that political involvement on Twitter, at least, exhibits potential for more inclusive representation.

Data Collection

The Florida voting-eligible electorate comprises this study's target population.[2] The reason for focusing on Florida is the availability of email addresses in its public voter files. Email addresses serve as unique keys to combine the voter records with social media accounts and are usually treated as confidential in other states. Florida's voter files, however, are easily obtained, cleared for non-commercial research, and among the richest in available information (Cooper et al., 2009). In substantive terms, Florida is also a politically diverse and perennial swing state referred to as a microcosm of the United States (MacManus et al., 2015).

I collected the state's voter records as of October 2017. The list records an individual's registration date, registered party affiliation, sex, date of birth, race, residence, email address, and turnout at 12 consecutive primary and general elections between 2006 and 2016. Turnout at the 2018 primary and general elections was updated using the file as of December 2018. Since individual income information is not available, I rely on the 2017 American Community Survey 5-year estimates of per capita income at small-scale census block groups as the closest possible surrogate.

[2] Unlike the US Twitter population, which includes persons who are not eligible to vote, this target population allows for a comparison of voting and online political involvement.

Linking Twitter Accounts to Voter Records

To assess social media-based political participation, I concentrate on the social networking service Twitter. The platform emphasizes communication and interaction and provides a major venue for Americans to take part in political discourse (Barberá et al., 2019; Nagler and Tucker, 2015).

In order to identify registered voters' Twitter accounts, I rely on self-reported email addresses in the voter records. Twitter users cannot be located via a simple platform search for their email addresses. Because of this, I programmed an automated routine that uses the platform's email account synchronization feature to return matching Twitter users. The approach is described in detail in Appendix A. To summarize, email addresses from the voter record were uploaded as contacts to an email account specifically created for this project. A Twitter account that was likewise created for this project was then synchronized with this email account's contacts to yield a list of Twitter users who have allowed others to find them by their email address.[3]

What complicated the matter was that matching Twitter users were returned in random order and detached from the corresponding set of email addresses. A second matching stage was therefore required. The only other indicator available for both registered voters and Twitter users are persons' names. Existing research shows that a majority of Twitter users are identifiable via reported first and last names (Longley et al., 2015; Peddinti et al., 2017). Of course, unlike email addresses, names are not generally unique identifiers. However, when synchronizing Twitter with just a subset of email addresses limited to persons with unique first and unique last names, names become unique identifiers within the second matching stage. To put it in simpler terms, email addresses uniquely identify Twitter accounts for a subset of persons with unique names at the first stage and names simply serve to put them together in the correct order at the second stage. Accordingly, Twitter accounts were linked to voter records iteratively by uploading, synchronizing, and matching small subsets of email addresses, one at a time and each including only persons with unique

[3] See https://help.twitter.com/en/using-twitter/upload-your-contacts-to-search-for-friends (last accessed July 2021). At the time of data collection, email discoverability was the default and Twitter users had to explicitly opt-out to disallow being found via their email address.

first and unique last names, until all email addresses were processed.[4]

Voter records have been integrated with social media data before (see Appendix A for an extended discussion). For Twitter, prior approaches rely on reported names and locations of active users collected through Twitter's streaming API to match registered voters with unique names within locations (Barberá, 2014; Grinberg et al., 2019). One novelty of the method used in this paper is that it neither excludes passive Twitter users nor persons with names that appear more than once. It thus increases the coverage of registered voters with Twitter accounts. A second advantage of the method presented above is that it is less vulnerable to mismatches. Here, mismatches are only possible in arguably unlikely cases. For instance, when a registered voter reports an email address that actually belongs to another person on Twitter with the same name or when an identified Twitter user displays a fake name that happens to match the name of another person in this specific unique-names-subset. Since prior approaches do not proceed from a unique identifier, mismatches are much more likely. For example, two persons could share the same name and place of residence, yet only one of them appears in the voter record while the other may be the one with the active Twitter account.

Using the method described above, 90,832 registered voters were successfully linked with a non-protected Twitter account. Their activity on Twitter was recorded daily between August 1, 2018 and March 31, 2019 using Twitter's REST API. In all, the full text of more than 6 million posts (users' own and shared) was collected. Appendix B offers further details on the collection of voter registration and Twitter data. Ethical and legal considerations are discussed in Appendix I.

Comparing Sample and Target Population

The specifics of the data sources and their linkage bring about several selection steps. Selection into the sample depends on having registered to vote, having reported an email address in the voter registration application (reported by 681,096 registered voters,

[4] Note that persons with names appearing more than once in the voter record were not generally dropped, their email addresses were simply processed in separate batches.

5.3%)[5], having a Twitter account that was created with or is linked with the reported email address, not having opted out on Twitter from being located via this email address, having reported one's actual name on Twitter, and not having a protected Twitter account. If these criteria are not met, a person is not included in the sample (detailed processing steps are reported in Appendix A).

Figure 1 compares this non-probability sample to different realizations of the Florida electorate, i.e., the registered-voter, citizen-voting age, and voting-eligible population along observable characteristics.[6] Note that the registered-voter and voting-eligible population are fairly similar, which lends additional support for focusing on Florida. Although there is some overrepresentation of the middle-aged at the expense of those 65-plus and of White over Black voters the non-probability sample closely approximates the voting-eligible electorate. The notable overrepresentation of young adults and Democrats in conventional Twitter samples (Sloan et al., 2015; Wojcik and Hughes, 2019) does not occur here. Party affiliation matches the distribution in the registered-voter population and turnout at the 2018 general election (64.5%) is close to official results (63%).[7] The overrepresentation of voters at the 2018 primary, however, cautions that this sample is also not entirely immune to the biases discussed earlier.

A central drawback of these data is that they exclude non-registered non-voters, who may differ from registered non-voters in important ways. Existing research shows, for instance, that racial and ethnic minorities are more likely than whites to be left out of the voter rolls due to incomplete voter registration applications (Merivaki, 2020). To the extent that those non-voters who (successfully) registered to vote differ in their online political involvement from unregistered non-voters, the results presented in this article will be biased (in unknown directions) as well. This means that the nonresponse-problem in surveys, discussed in the introduction, is potentially only partially alleviated (Nyhan et al., 2017) and that the findings presented here are certainly not the last word on this matter.

[5] Note that data privacy considerations that may also govern the reporting of email addresses are not related to political engagement (see Appendix A, section "Sample Processing Steps" for details.)

[6] See Appendix C for definitions and estimation of the different populations. For a geographic breakdown, showing broad geographic coverage of the sample, see Appendix A.

[7] Based on https://dos.myflorida.com/elections/data-statistics/elections-data/voter-turnout/ (last accessed July 2021).

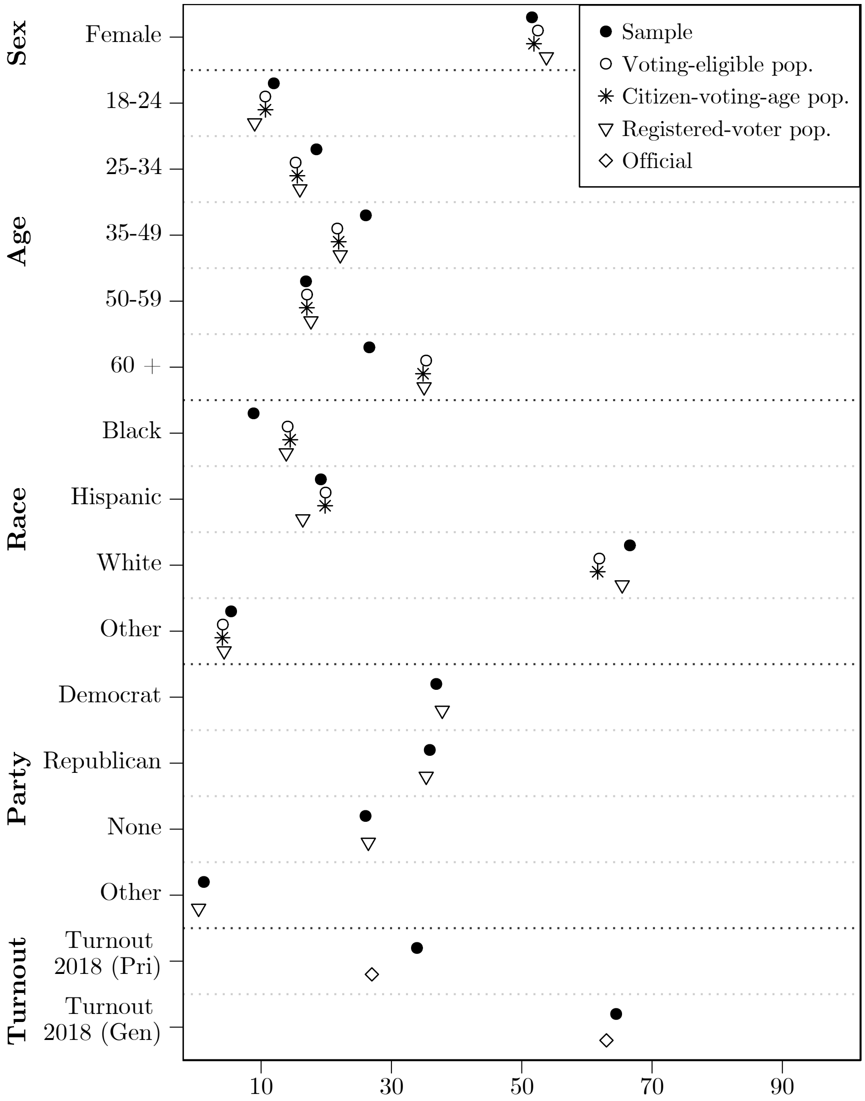


Figure 1. Sample and target population characteristics.

Since surveys seem to provide at least some coverage of unregistered non-voters (see Jackman and Spahn, 2018), future research may consider alternative data collection strategies that use both voter records and surveys in combination with social media accounts to harvest their complementary strengths and overcome remaining biases.

Measuring Political Participation

Integrating voter records with Twitter data helps mitigating some survey-specific problems due to selection bias and especially overreporting. Assessing whether political involvement on Twitter reinforces or compensates participatory deficiencies at the polls still requires a decision on how to measure an individual's inclination to vote and social media-based participation.

I begin by defining political participation as "a voluntary activity by citizens in the area of government, politics, or the state" (van Deth, 2014). To cast a vote is the most common form of participation meeting this definition (Schlozman et al., 2018). Therefore, it makes sense to draw on the decision to vote at an election as the benchmark against which to compare online political engagement. Measuring turnout at one specific election, however, ignores that participation varies depending on the importance ascribed to high- and low-stimulus elections (Campbell, 1960). Contextual factors, person- and election-specific idiosyncrasies additionally twist voter behavior (Sigelman and Jewell, 1986). Identifying voter types, such as low-propensity, marginal, or highly engaged voters, based on one or two elections is therefore prone to measurement error.

To reduce measurement error and empirically inform a fine-grained differentiation between voter types, I rely on latent variable modelling (Ansolabehere et al., 2008). Specifically, I use a two-parameter item response theory model to assess individuals' general inclination to vote (Clinton et al., 2004). The model represents each person's probability of voting in different elections as a function of an underlying voting propensity and two election-specific factors with reference to this latent trait, the discriminating power or weight of an election and the threshold at which voting is more likely than abstaining (Fowler et al., 2008). All of these quantities are jointly estimated from observed participation decisions at several elections. This allows for the information included in various elections and election types to differentially contribute to a general assessment of voter engagement, which is provided by the estimated voting propensity. In other words, the voting propensity measure takes into account that the decision to vote is not fixed but differs across elections.

To estimate the model, I rely on validated turnout at 14 consecutive elections (seven primary, four midterm, and three presidential) from 2006 to 2018. Note that individuals' turnout decisions are only included in the model for the subset of elections at which they were actually eligible to vote, i.e., registered and of legal age (See Appendix B for details on how election-specific voting eligibility was computed). This is possible because the item response theory model allows for unbalanced data, whereby voters with longer registration records contribute more to parameter estimates. To put the resulting scale into context: located around 0 are citizens with a voting probability that is high for presidential, moderate for midterm, and low for primary elections. Consistent non-voters (8.6%) score an average voting propensity of -1.36, consistent voters (6.2%) score 1.48. See Appendix D for the model, its implementation, and parameter estimates.

Drawing on an extension of the above definition (Theocharis, 2015), social media-based political participation is conceptualized as an activity targeted at the area of government, politics, or the state. Accordingly, I focus on own and shared posts with political content or a political recipient on Twitter. Using a keyword-based classifier, I categorize the text of all collected posts in line with this definition.

Rather than using off-the-shelf terms for classification, I rely on a computer-assisted algorithm for keyword discovery (King et al., 2017) to build a problem- and context-specific dictionary. The method uses machine learning algorithms to detect keywords based on coded examples and texts including relevant terms. Two political scientists hand-coded a random sample of 4,000 posts following the definition above.[8] A collection of 728,089 georeferenced Twitter posts from Florida, gathered daily throughout the period of investigation, provided texts with potentially relevant keywords. The resulting dictionary consists of 331 keywords. Based on this dictionary 1,525,672 (24%) posts belonging to 12,876 (14%) registered voters were categorized as political.[9] I measure political involvement on Twitter conditional on the occurrence of one or more political posts.[10] Details on text processing, the construction

[8] Interrater reliability based on Cohen's $\kappa = 0.91$ (95% confidence interval = 0.89, 0.93).

[9] Hughes and Asheer (2019) report 13% political posts related to 10% of U.S. adults. At least two reasons account for this discrepancy: (1) U.S. adults (Hughes and Asheer) vs. the Florida voting-eligible electorate (this study) as target population, (2) including only posts on national politics (Hughes and Asheer) vs. including posts on all levels of government and general political topics (this study).

[10] This is a coarse measure compared to the voting propensity. There exists no precedent for a similar measure or guidelines on which dimensions to incorporate, except for the frequency of

of the dictionary, and its validation with reference to statistical bias are given in Appendix E.

Results

The area-proportional Euler diagrams in Figure 2 show that 91% of voters politically engaged on Twitter also voted in the 2016 presidential election. However, high-stimulus elections lump together highly engaged, irregular, and low-propensity voters. Less salient elections, where the voting population narrows to core voters (Campbell, 1960),[11] are more informative about the claim that online participation is merely executed by the highly engaged. The fraction of the politically involved voters on Twitter who also voted in the election reduces to 73.5% in the 2018 midterms and drops below 50% in the last two primaries.[12] This contradicts prior survey evidence, which finds higher voting rates among politically engaged on Twitter even in midterm elections (e.g., Bode and Dalrymple, 2016, reporting 94% in 2010).

Low-stimulus elections roughly distinguish core voters from the rest of the electorate. Yet such elections confound likely voters who missed a particular election and less inclined voters who participated for a change (see Figure G2 in the Appendix). The estimated voting propensities, on the other hand, allow us to disentangle the degree of electoral participation among voters who are politically involved on Twitter and compare it to the underlying voting-eligible population. Figure 3 presents this comparison. Among the voters who are politically engaged on Twitter, the distribution of voting propensities is somewhat more left-skewed than for the overall sample. This implies that more engaged voters are, perhaps unsurprisingly, more likely to incorporate online means of participation into their repertoire. Other than predicted by reinforcement theory, however, voters who are politically involved

political posts which are not as clearly bounded as elections. Developing such a measure is outside the scope of this paper. However, as shown in Figures G3, G5, and Table G1 in the Appendix, adjustments based on the frequency of political posts reduce the overall amount of social media-based participation but otherwise yield substantively similar results.

[11] See Figure D2 in the Appendix.

[12] Table G1 in the Appendix shows similar results for sample subsets defined by levels of (political) Twitter activity, time windows, party registration, and voter status. Figure G1 in the Appendix additionally shows that results are independent of specific events. The proportion of politically active voters and non-voters on Twitter is strikingly constant over time.

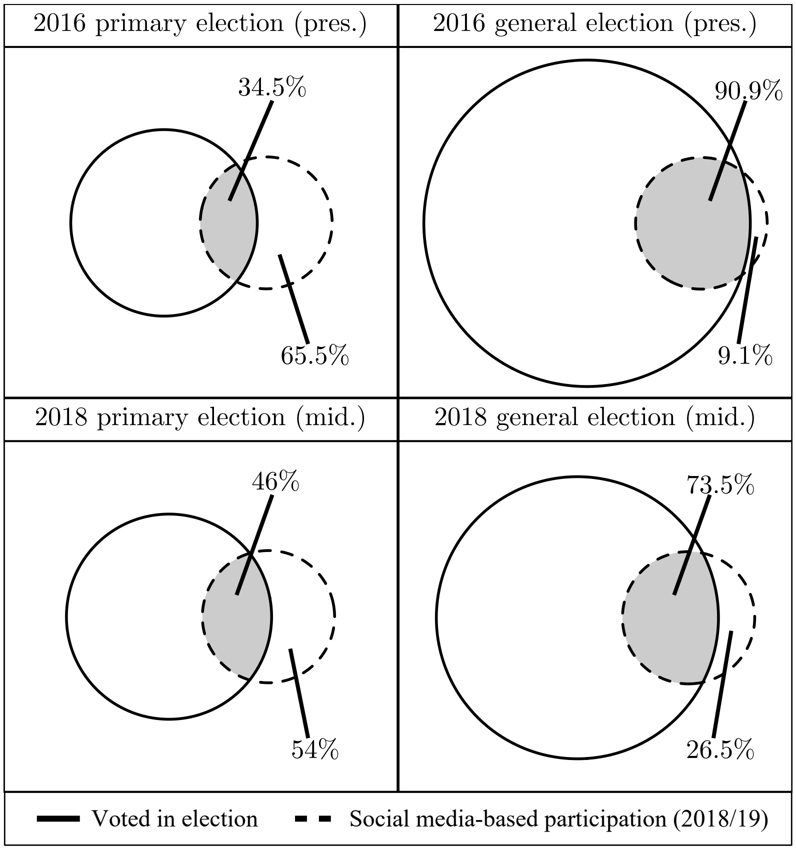


Figure 2. Area-proportional Euler diagrams of voting and social media-based participation.

on Twitter are not concentrated around higher voting propensities, i.e., are neither solely nor primarily composed of the highly engaged segment of the electorate. Instead, social media-based participation is spread across the electorate and extends to those who are least likely to vote.

The question remains whether this online engagement also extends to traditionally disadvantaged groups. Schlozman et al. (2018) report that online political participation is almost twice as prevalent among White compared to Black voters. The reported disparity between whites and Hispanics is even greater. They also highlight a noticeable increase in social media-based participation along socioeconomic status. To investigate how various

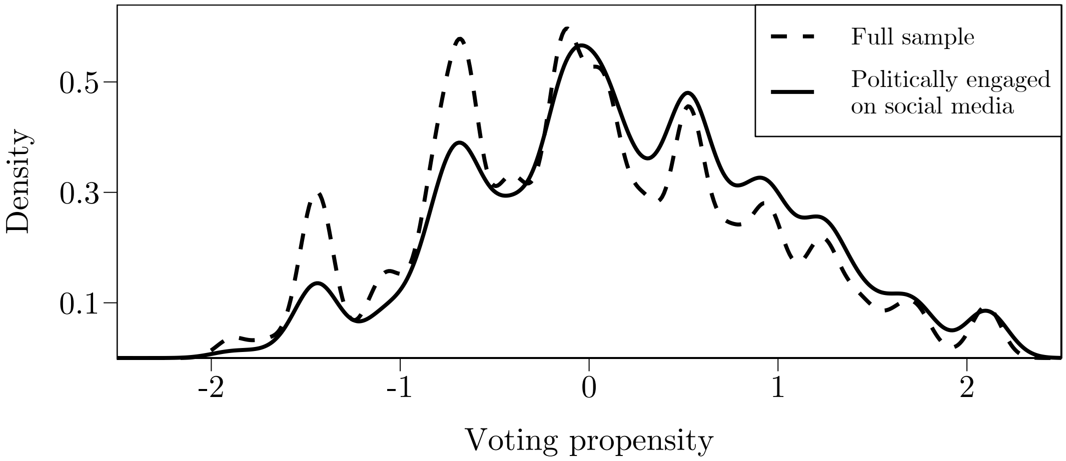


Figure 3. Voting propensities in the sample and among politically involved on social media.

subgroups differ in their political involvement on Twitter, I conduct a multilevel regression with varying intercepts for demographic and party groups along with corresponding two-way interactions (Ghitza and Gelman, 2013).[13]

I begin with results that pool all voters, regardless of their voting propensity. Corresponding estimates are shown in Figure 4. What stands out immediately is that political involvement on Twitter is substantially skewed towards males, younger adults, and Democrats. Established voting research (Fraga, 2018; Leighley and Nagler, 2013) and Florida voter turnout in 2018 (See Table G2 in the Appendix) depict these voter groups as less active at the polls than their respective counterparts. Differences along sex and age corroborate prior results on online political engagement (Bode et al., 2014; Bekafigo and McBride, 2013). Differences among race and income groups, however, are less discernible than existing studies suggest. Importantly, this finding does not confirm the long-standing narrative of strong racial and income disparities in online political activity.[14] At first sight, the results thus seem to uniformly support a compensation perspective.

This aggregate perspective, however, blends different voter types. It may mask evi-

[13] See Appendix F for the formal model, implementation, and postestimation.

[14] Figure G4 in the Appendix presents poststratified estimates that account for remaining differences between the sample and target population, leaving results unchanged.

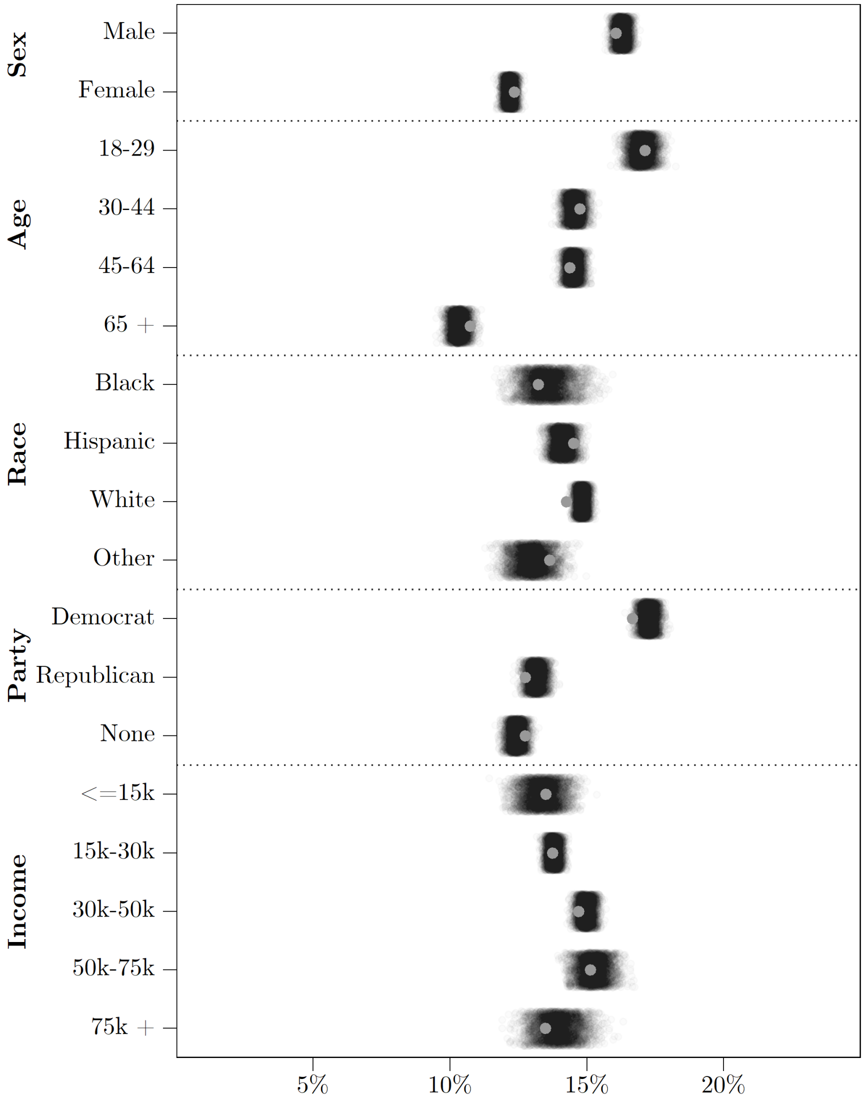


Figure 4. Social media-based participation in subgroups. Raw shares (gray circles) and population-averaged predictions based on 4,000 posterior draws (black circles).

dence of compensation or reinforcement, for example, if minority and low-income working-class voters politically involved on Twitter come disproportionately from low-propensity or core voters. For this reason, I replicate the multilevel regression, this time partitioned by four concentric voter types: low-propensity, irregular, regular, and core voters. The voter types are informed by individuals' average participation probabilities at primary, midterm,

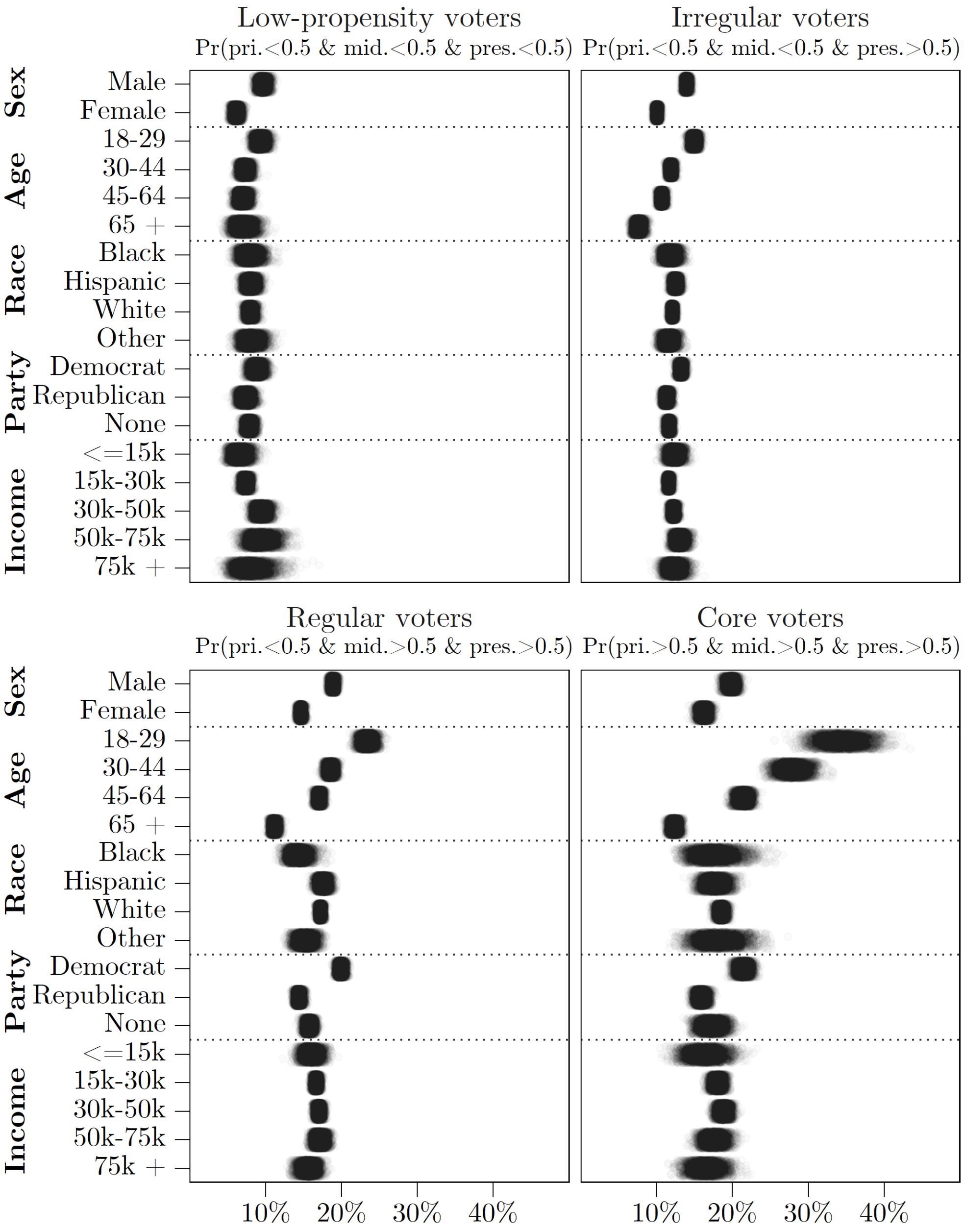


Figure 5. Social media-based participation in subgroups by voter types. Population-averaged predictions based on 4,000 posterior draws.

and presidential elections derived from the measurement model. Corresponding estimates are shown in Figure 5.

With regard to sex, race, and income groups, results remain largely unchanged.[15] Political participation on Twitter compensates minority and low-income voters' relative absence at the polls (Fraga, 2018; Schlozman et al., 2018, see Table G2 in the Appendix for the Florida electorate) insofar as it appears balanced across all voter types and among low-propensity and irregular voters in particular.

For age and party subgroups, however, the gap in political involvement on Twitter is much less pronounced towards low-propensity voters and increases substantially towards core voters. This limits the evidence in favor of the compensation perspective suggested by the pooled results in Figure 4. The findings for age subgroups are particularly interesting in this regard. Taking age as a proxy for digital literacy (Guess and Munger, 2020), engagement in political discourse on social media is not necessarily more prevalent among digitally more literate but politically alienated voter groups. Instead and contrary to what is frequently expected by compensation theorists, young adults' elevated political involvement on Twitter appears primarily among already engaged and frequent voters.

Implications

Social media continues to surface as a distinct addition to citizens' participatory repertoire. Yet survey evidence about its use for political engagement in the American electorate remains sobering. Online participation is largely perceived as a weapon of the strong, wielded by the highly politically involved, white, and affluent class.

Offering a new perspective on this research, this study departs from survey self-reports and combines administrative data with Twitter accounts. I find that constituents' political involvement on Twitter does not mirror persistent participatory inequalities in voting and extends to those who are least likely to turn out on election day. These results suggest an opposing view to reinforcement theory and highlight social media's compensatory value for more inclusive representation. However, the findings also point out that social media may not offer the much-anticipated remedy for America's youth participation gap.

[15] Disaggregation into interacted subgroups and different model specifications yield substantively similar results (see Figures G6 to G31 in the Appendix).

The results presented here speak to several domains. First, they add to a recent literature which challenges the enduring narrative of participatory deficiencies among non-white Americans (Anoll, 2018). Second, findings are consistent with studies that question the importance of individual resources for structuring political participation (Ansolabehere and Hersh, 2012). Third, the results have direct implications for political representation. Recent research indicates that American legislators' political agenda is responsive to preferences expressed on Twitter, although primarily with regard to strong partisan supporters (Barberá et al., 2019). Especially during primary and midterm elections, where the voting population narrows to core voters, non-voters keep exercising political voice on Twitter. Politicians can act on this knowledge to learn about and better represent a broad spectrum of the electorate, including those who are traditionally underrepresented at the polls, both during and between elections. Moreover, being more accountable to posts by these groups might have beneficial consequences for turnout among these voters. Finally, the findings raise questions as regards the relationship between digital literacy and online political involvement. Under what conditions do inequalities in digital literacy also result in unequal participation online? Are older adults who are less inclined to vote maybe more motivated to adapt digitally to alternative channels than their highly-engaged counterparts?

An important limitation to the external validity of this study is its sole focus on Twitter. Twitter is distinct from other social networking services, such as Facebook or Instagram, and likely has a different and potentially more politically active user base. Another caveat is the exclusion of unregistered voters (Nyhan et al., 2017). This means that the online political involvement of the most disengaged and chronic non-voters remains hidden to us. Both of these drawbacks are likely to account for the still relatively high voting rates found here and call for future research. In this sense, pairing commercially augmented voter registration lists with nationally representative surveys to increase coverage of unregistered non-voters (Ghitza and Gelman, 2020; Jackman and Spahn, 2018) while obtaining participants' consent to access their activity on various social media platforms, may offer a way to assess the validity and generalizability of my findings. The insights and methodological approaches presented in this article will hopefully help to support such efforts and to reinvigorate research on diverse forms of political participation.

Acknowledgments

This research was supported by a grant from the Excellence Initiative of the German federal and state governments at the University of Konstanz. I would like to thank Lada Rudnitckaia for her excellent research assistance and Fred Hockney for language editing and proofreading. For their helpful comments, I thank Christian Breunig, Benjamin Guinaudeau, Andreas Jungherr, Konstantin Käppner, Lukas Kawerau, Simon Munzert, Peter Selb, Susumu Shikano, participants of the panel "Using Big Data for Electoral Research" at the 2018 ESRA Big Data Meets Survey Science conference in Barcelona, the editors, and the anonymous reviewers.

References

Anoll, A. P. (2018). What makes a good neighbor? Race, place, and norms of political participation. *American Political Science Review*, 112(3):494–508.

Ansolabehere, S. and Hersh, E. (2012). Validation. What big data reveal about survey misreporting and the real electorate. *Political Analysis*, 20(4):437–459.

Ansolabehere, S., Rodden, J., and Snyder, J. M. J. (2008). The strength of issues. Using multiple measures to gauge preference stability, ideological constraint, and issue voting. *American Political Science Review*, 102(2):215–232.

Barberá, P. (2014). Birds of the same feather tweet together. Bayesian ideal point estimation using Twitter data. *Political Analysis*, 23(1):76–91.

Barberá, P., Casas, A., Nagler, J., Egan, P. J., Bonneau, R., Jost, J. T., and Tucker, J. A. (2019). Who leads? Who follows? Measuring issue attention and agenda setting by legislators and the mass public using social media data. *American Political Science Review*, 113(4):883–901.

Bekafigo, A. M. and McBride, A. (2013). Who tweets about politics? Political participation of Twitter users during the 2011 gubernatorial elections. *Social Science Computer Review*, 31(5):625–643.

Best, S. J. and Krueger, B. S. (2005). Analyzing the representativeness of internet political participation. *Political Behavior*, 27(2):183–216.

Bimber, B. (1999). The internet and citizen communication with government. Does the medium matter? *Political Communication*, 16(4):409–428.

Bode, L. and Dalrymple, K. E. (2016). Politics in 140 characters or less. Campaign communication, network interaction, and political participation on Twitter. *Journal of Political Marketing*, 15(4):311–332.

Bode, L., Vraga, E. K., Borah, P., and Shah, D. V. (2014). A new space for political behavior. Political social networking and its democratic consequences. *Journal of Computer-Mediated Communication*, 19(3):414–429.

Brehm, J. (1993). *The phantom respondents. Opinion surveys and political representation.* University of Michigan Press, Ann Arbor.

Campbell, A. (1960). Surge and decline. A study of electoral change. *Public Opinion Quarterly*, 24(3):397–418.

Carpini, M. X. D. (2000). Gen.com. Youth, civic engagement, and the new information environment. *Political Communication*, 17(4):341–349.

Clinton, J., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2):355–370.

Cooper, C. A., Haspel, M., and Knots, G. H. (2009). The value of voterfiles for U.S. state politics research. *State Politics and Policy Quarterly*, 9(1):102–121.

Elwert, F. and Winship, C. (2014). Endogenous selection bias. The problem of conditioning on a collider variable. *Annual Review of Sociology*, 40:31–53.

Fowler, J. H., Baker, L., and Dawes, C. T. (2008). Genetic variation in political participation. *American Political Science Review*, 102(2):233–248.

Fraga, B. L. (2018). *The turnout gap. Race, ethnicity, and political inequality in a diversifying America.* Cambridge University Press, Cambridge.

Gainous, J. and Wagner, K. M. (2014). *Tweeting to power. The social media revolution in American politics.* Oxford University Press, Oxford.

Ghitza, Y. and Gelman, A. (2013). Deep interactions with MRP. Election turnout and voting patterns among small electoral subgroups. *American Journal of Political Science*, 57(3):762–776.

Ghitza, Y. and Gelman, A. (2020). Voter registration databases and MRP. Toward the use of large-scale databases in public opinion research. *Political Analysis*, 28(4):507–531.

Gibson, R. and Cantijoch, M. (2013). Conceptualizing and measuring participation in the age of the internet. Is online political engagement really different to offline? *Journal of Politics*, 75(3):701–716.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425):374–378.

Guess, A. and Munger, K. (2020). Digital literacy and online political behavior. *Unpublished manuscript*.

Guess, A., Munger, K., Nagler, J., and Tucker, J. (2019). How accurate are survey responses on social media and politics? *Political Communication.*, 36(2):241–258.

Hindman, M. (2009). *The myth of digital democracy*. Princeton University Press, Princeton, NJ.

Hughes, A. and Asheer, N. (2019). National politics on Twitter. Small share of U.S. adults produce majority of tweets. *Pew Research Center*.

Jackman, S. and Spahn, B. (2018). Politically invisible in America. *Unpublished manuscript*.

Jackman, S. and Spahn, B. (2019). Why does the American National Election Study overestimate voter turnout? *Political Analysis*, 27(2):193–207.

King, G., Lam, P., and Roberts, M. R. (2017). Computer-assisted keyword and document set discovery from unstructured text. *American Journal of Political Science*, 61(4):971–988.

Krueger, B. S. (2002). Assessing the potential of internet political participation in the United States. A resource approach. *American Politics Research*, 30(5):476–498.

Leighley, J. E. and Nagler, J. (2013). *Who votes now? Demographics, issues, inequality, and turnout in the United States*. Princeton University Press, Princeton.

Longley, P. A., Adnan, M., and Guy, L. (2015). The geotemporal demographics of Twitter usage. *Environment and Planning A: Economy and Space*, 47(2):465–484.

MacManus, S. A., Jewett, A., Bonanza, D. J., and Dye, T. R. (2015). *Politics in Florida*. Peppertree Press, Sarasota, FL, 4 edition.

Merivaki, T. (2020). Who is left out? the process of validating voter registration applications. *American Politics Research*, 48(6):682–686.

Nagler, J. and Tucker, J. (2015). Drawing inferences and testing theories with big data. *Political Science and Politics*, 48(1):84–88.

Norris, P. (2001). *Digital divide. Civic engagement, information poverty, and the internet worldwide*. Cambridge University Press, New York.

Nyhan, B., Skovron, C., and Titiunik, R. (2017). Differential registration bias in voter file data. A sensitivity analysis approach. *American Journal of Political Science*, 61(3):744–760.

Oser, J. and Boulianne, S. (2020). Reinforcement effects between digital media use and political participation. A meta-analysis of repeated-wave panel data. *Public Opinion Quarterly*, 84(S1):355–365.

Oser, J., Hooghe, M., and Marien, S. (2013). Is online participation distinct from offline participation? A latent class analysis of participation types and their stratification. *Political Research Quarterly*, 66(1):91–101.

Peddinti, S. T., Ross, K. W., and Cappos, J. (2017). User anonymity on Twitter. *IEEE Security and Privacy*, 15(3):84–87.

Schlozman, K. L., Brady, H. E., and Verba, S. (2018). *Unequal and unrepresented. Political inequality and the people's voice in the new gilded age*. Princeton University Press, Princeton, NJ.

Schlozman, K. L., Verba, S., and Brady, H. E. (2010). Weapon of the strong? Participatory inequality and the internet. *Perspectives on Politics*, 8(2):487–509.

Sigelman, L. and Jewell, M. E. (1986). From core to periphery. A note on the imagery of concentric electorates. *Journal of Politics*, 48(2):440–449.

Sloan, L., Jeffrey, M., Burnap, P., and Williams, M. (2015). Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data. *PLoS ONE*, 10(3).

Theocharis, Y. (2015). The conceptualization of digital networked participation. *Social Media + Society*, 1(2).

van Deth, J. W. (2014). A conceptual map of political participation. *Acta Politica*, 49(3):349–367.

Wojcik, S. and Hughes, A. (2019). Sizing up Twitter users. *Pew Research Center*.

Supplementary Materials

Contents

Appendix A: About the Sample

Non-Probability Sampling Strategy

Florida's voter record as of October 2017 was used as basis for acquiring a sample of registered voters with a Twitter account. Persons in the voter record were selected into the sample if they reported an email address based on which they could be uniquely identified on the social networking service Twitter. This makes the sample strictly non-probability. Selection into the sample depends on having registered to vote, having reported an email address, having a Twitter account linked to the reported email address, not having opted out on Twitter from being located via the email address, and further factors specific to the matching approach introduced below.

I used a five-step strategy to identify registered voters Twitter accounts based on reported email addresses:

1. Randomly sample a small batch of voter record entries with unique first and last names.
2. Upload their email addresses to Google's Gmail and synchronize with Twitter.
3. Collect information about Twitter contacts from synchronization output.
4. Apply dynamic name matching between Gmail and Twitter contacts.
5. Repeat steps 1-4 until voter record is empty.

At the time, identifying Twitter users via email addresses was only possible via synchronization of Twitter with a Google Gmail account, not via a simple search.[16]

[16] Note that email addresses of registered voters can have any domain and are not restricted to Gmail. The Gmail account is only required on the part of the researcher to initialize the synchronization of contacts with Twitter.

As output of the synchronization process, Twitter offers a collection of accounts associated with the email addresses stored in the Gmail contacts. An email address is a unique identifier. The combination of user name and domain can be assigned only once. Hence, we know for sure that the subset of accounts provided by Twitter belong to persons in the Gmail contacts.

However, the output neither lists Twitter accounts next to email addresses nor in the original order. An additional matching step is thus required to link the identified accounts back to the respective email addresses. The only information in the output that can be used for this purpose are users' names (not handles). Since names are not unique identifiers, however, it becomes necessary to proceed iteratively along batches from the voter record with unique first and last names (step 1). For persons with the same first or last name, only one person is kept in a batch, the others are left to be drawn in next iterations. This way, names are unique identifiers in the source, i.e., the Gmail contacts.

In the synchronization output names are not necessarily unique. A person might choose to display a first and last name different from her true name. Accounts will not be matched back to the sample batch and are discarded in such cases. So selection into the sample also depends on persons reporting actual names on Twitter. Duplicate first or last names that match the Gmail contacts never occurred in the synchronization output. Mismatches are consequently only possible in cases where two identified accounts reported a wrong name and one of them happened to display the actual name of the other. This is arguably rather unlikely and considered noise. Otherwise, the synchronization via email addresses ensures that only Twitter accounts of persons in the Gmail contacts are returned. Unique first and last names in the Gmail contacts ensure that a Thomas in the synchronization output who matches a Thomas in the Gmail contacts are one and the same person.

The strategy was programmed in an algorithm that fully automatizes the procedure (see Figure A1 for pseudocode). After having drawn and filtered a sample

```
Algorithm: EmailToTwitter
begin
    $pool \leftarrow$ sequence 0 to row length of $E$;
    initialize two automated browser sessions via Selenium;                      } prepare
    while length of $pool > 0$ do
        if length of $pool \geqslant 1000$ then
            $batch \leftarrow$ sample of size 1000 from $p$ without replacement;
        else
            $batch \leftarrow pool$;
        remove from $batch$ indices of duplicates in $E$[, first name] and       } partition
          $E$[, last name];
        $B \leftarrow E[batch,]$;
        assign/add $B$ to $Blist$ and write $B$ to disk;
        $pool \leftarrow$ remove from $pool$ the set $pool \cap batch$;
        navigate and log in to Gmail and Twitter in browser sessions;
        import $B$ from disk in GMail then import contacts in Twitter;
        $T \leftarrow$ extract user data of Twitter contacts via XPath;          } extract
        assign/add $T$ to $Tlist$;
        clear Gmail and Twitter contacts;
        $names \leftarrow$ list(split elements in T[, name] into word vectors);
        $handles \leftarrow$ list(split elements in T[, handle] into word vectors);
        cut $T[names$[every first element] $\cap$ $B$[, first name]$,]$ from $T$ and
          integrate in $B$;
        cut $T[names$[every last element] $\cap$ $B$[,last name]$,]$ from $T$ and
          integrate in $B$;                                                     } match
        cut $T[names$[every second element] $\cap$ $B$[,first name]$,]$ from $T$
          and integrate in $B$;
        cut $T[names$[every second element] $\cap$ $B$[,last name]$,]$ from $T$
          and integrate in $B$;
        cut $T[handles$[every last element] $\cap$ $B$[,last name]$,]$ from $T$ and
          integrate in $B$;
        assign/add $B$ to $Mlist$
    return list$(Mlist, Blist, Tlist)$
end
```

Figure A1. Pseudocode for matching algorithm.

Note: Input – *user* = Gmail and Twitter user name, *key* = Gmail and Twitter password, E = array with columns 'email', 'first name', and 'last name' from voter record. Data – *pool* = index of email pool, *batch* = index of current email batch, *names* = index of Twitter screen name components vector, *handles* = list of Twitter handle components vectors, B = *batch* subset of E, T = array with columns 'name', 'handle', 'id' (from Twitter), *Blist* = list of Bs collected throughout iterations of the algorithm before the extraction step, *Tlist* = list of Ts collected throughout iterations of the algorithm, *Mlist* = list of Bs collected throughout iterations of the algorithm during the match step. Output – list(*Mlist, Blist, Tlist*).

of persons with unique first and last names from the voter record the algorithm launches a simulated web browser session via Selenium, a framework for web browser automation. The algorithm uploads the batch of persons from the voter record to a Gmail account and synchronizes with Twitter. Using the XPath query language the algorithm then extracts account information. Via dynamic name matching that proceeds with various combinations of a full name, the account information is finally linked back to the sample and stored in a separate file. The algorithm repeats this procedure until all individuals with a reported email in the voter registry have been processed. Several plausibility checks on random samples comparing email address, full name in the registration record, Twitter name, and Twitter handle support that matches are genuine.

Comparison of the Non-Probability Sampling Strategy to Prior Work

Strategies similar to the above have been adopted before. To identify registered voters' social media accounts researchers at the University of California San Diego collaborated with Facebook and devised a group-level matching procedure which assigns turnout behavior to Facebook users (Jones et al., 2013). Their strategy yields several potential turnout frequencies for each individual to guarantee Facebook users' anonymity. These frequencies are then used to predict individuals' probability to be unregistered, a voter, or an abstainer and classify them accordingly. The procedure hence produces a statistical match, as opposed to the exact match used in this paper (See Sakshaug (2018) for a distinction between exact, probability, and statistical linkage). Moreover, an implementation of this approach is dependent on a formal collaboration with Facebook. As others have noted, Facebook is rather reserved when it comes to collaboration with academia (Margetts, 2017) – a privilege enjoyed primarily by already tenured and well funded researchers (Ruths and Pfeffer, 2014). Another caveat is that Facebook prohibits linkage of information other than turnout behavior, such as individuals' place of residence (see Settle et al., 2016). In addition, the linkage method requires initial removal of entries with the same combination of individuals' first name, last name, and date of birth used for matching.

Another approach makes use of the fact that some Twitter users enable their posts to be geotagged (Barberá, 2014). Barberá used Twitter's live stream to accumulate messages sent with coordinates located in the U.S. over a long period. Metadata of the geolocated messages was then used to extract names of the respective account holders and identify their zip codes. Using a combination of first name, last name, and zip code, entries in voter records were merged to Twitter accounts. The initial reliance on Twitters' live stream selects users based on activity, systematically excluding the inactive population. The active Twitter population, however, is not representative of the general or voting-eligible population (Wojcik and Hughes, 2019). This strategy offers probabilistic linkage as matching at the zip code level introduces uncertainty. This is because Twitter users who opt into having their messages geocoded, will not have their residential address revealed but the location from which a message was sent. Accordingly, a match could turn out to be a person just visiting and posting from the respective location while living or being registered to vote elsewhere. Similarly, a Twitter user who just moved into a zip code area and does not appear in the voter record could be mistaken with a registered voter who wasn't active on Twitter or doesn't even have a Twitter account. Further, individuals on registration lists sharing the same first and last name within a zipcode area are excluded from the matching procedure. More systematic bias might occur because only very few users opt into having their posts geolocated and those who do differ systematically from those who don't (Klasnja et al., 2017; Sloan and Jeffrey, 2015).

Sample Processing Steps

The non-probability sampling strategy was implemented from December 2017 until February 2018, processing 681,096 (5.3%) registered voters who reported an email address. The long runtime is due to the algorithm operating in live browser sessions. In addition, the voter record could not be processed at once but only in smaller batches and frequent names piled up towards the end making samples with unique first and last names ever smaller. Twitter account information of 113,268 (16.7%) registered voters was returned by the algorithm.

An unnoticed behavior in Twitter's synchronization procedure resulted in multiple duplicates of account information for specific persons. If none of the email addresses in the Gmail contacts sample could be linked to a Twitter account, Twitter instead suggested account information of other persons. Given that some of these persons had very common names, their account information was falsely linked to persons in the Gmail contacts. Fortunately, the accounts suggested by Twitter in such instances were always the same. In consequence, the falsely matched Twitter handles occurred very frequently and were easily removed, reducing the sample to 109,491. Further duplicates occurred because of family members reporting the same email address. Manual inspection for the most frequent duplicates and matching on full names was used to identify the person associated with the Twitter account, reducing the sample further to 108,258. I also noticed some duplicate voter IDs and used the registration date to remove outdated entries in the record. This reduced the sample to 105,436. Remaining duplicates were likely couples using the same email address. In light of limited resources to deal with this, I decided to go for accuracy instead of sample size and removed all remaining duplicates, leaving the sample at 102,291. Next, observations with protected, i.e., non-public, as well as terminated Twitter accounts were removed, cutting the sample down to 90,895.

Protected accounts offer an opportunity to check whether data privacy considerations, which are potentially involved in the decision to report an email address and hence selection into the sample, are related to political engagement. This would again introduce selection bias. However, users with protected accounts hardly differ in voting propensity (mean $= -0.04$, sd $= 0.82$) from users with public accounts (mean $= 0.01$, sd $= 0.85$). Finally, I removed a few voters whose residence is not available as well as some who were not eligible to vote in 2016 and not registered in 2018 anymore, resulting in the final sample of 90,832 observations.

Geographic Composition of the Sample

Figure A2 depicts the geographic distribution of the sample. Small black dots represent voters and gray circles incorporated cities, the larger the circle, the larger the city. The left panel shows that registered voters in the sample are widely

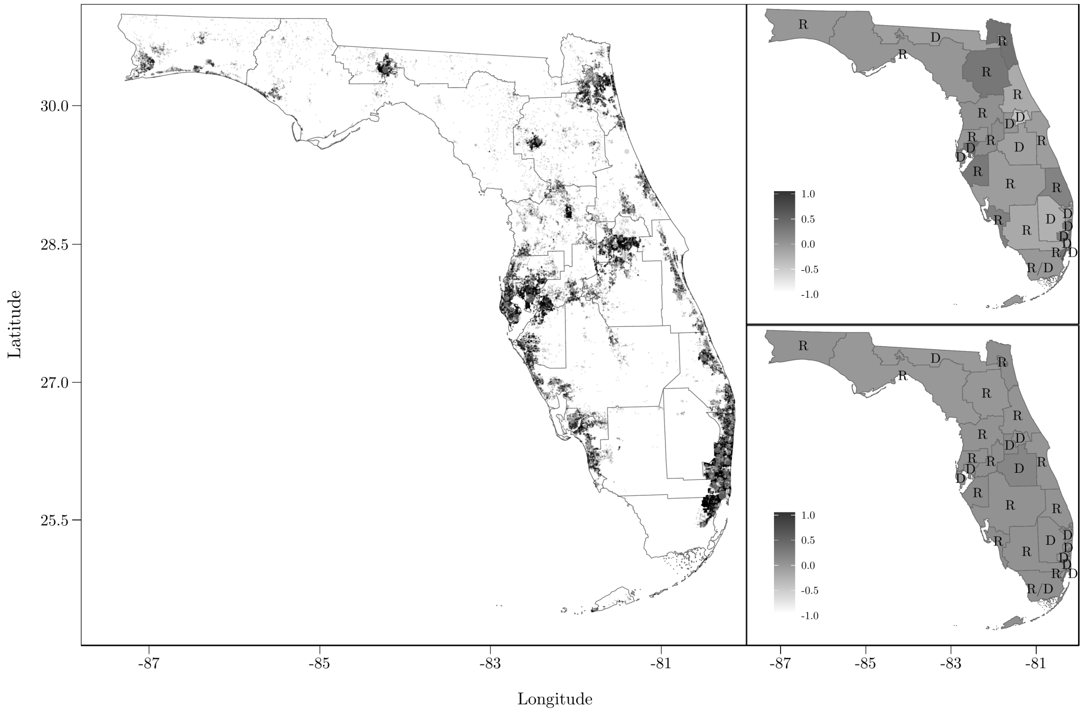


Figure A2. Geographic distribution of sample.

distributed across the state, covering both rural and urban areas. The sample covers every congressional district and almost every incorporated city, only a single county is not covered (Flagler county).

Congressional districts in the US are based on population. They are drafted with reference to latest census data to achieve an approximately equal population distribution. The lower right panel in Figure A2 shows actual deviations from an equal population distribution in Florida's congressional districts. The upper right panel shows the same for the sample. The capital letters "D" and "R" denote Democratic and Republican district incumbents. The sample overrepresents some of the urban areas, especially the regions surrounding Jacksonville, Tampa, and the Miami metropolitan area. This overrepresentation does not appear to be related to the party of the district incumbent, however. Similarly, wired broadband coverage in Florida is at 96%, mobile broadband access even at 100% and districts with lesser coverage are not underrepresented in the sample.[17] Still, the sample should not be taken as being representative down to specific regions or congressional districts. With view to the state as a whole, however, the sample provides a remarkably diverse geographic representation.

Appendix B: Details on Data Collection

Voter Registration Lists

A copy of monthly published voter registration lists including voting history information can be obtained directly from the Division of Elections of the Florida Department of State for a small processing charge[18] or downloaded from FL voters,[19] a collection of lists maintained by a former Republican state representative.

The voter registration lists are extracted from the Florida Voter Registration System. The associated voting histories come from the 67 county supervisors of election. The county supervisors of elections in Florida have several maintenance provisions in place, for instance, to identify voters that moved or are deceased (National Association of Secretaries of State, 2017) and researchers have found registration lists to be of generally high quality (Ansolabehere and Hersh, 2012). Nonetheless, some administrative errors do occur (Pettigrew and Stewart, 2018). As one measure to further clean the lists, I used the individuals' voter ID and registration date to remove outdated duplicate entries.

I collected the voter file as of October 2017, which served as starting point for matching Twitter accounts and constructing the sample. Information on biological sex was missing for 2,483 registered voters. To recover these missing values, I used software to predict sex from names and birth dates based on U.S. Social Security Administration baby name data and user profiles across major social networks (Mullen

[17] See https://broadbandnow.com/Florida (last accessed July 2021).

[18] See https://dos.myflorida.com/elections/data-statistics/voter-registration-statistics/voter-extract-disk-request/ (last accessed July 2021).

[19] See https://flvoters.com/downloads.html (last accessed July 2021).

et al., 2018a; Wais et al., 2019). Information on persons' race was missing from 1,064 observations. I recovered this information using software that predicts race based on surname, census tract, age, sex, and party affiliation (Imai and Khanna, 2016). Racial categories in the voter file are coded along the official categorization scheme by the United States Census Bureau. I grouped American Indian or Alaskan Natives, Asian or Pacific Islanders, and those with multiple or other races in the "other" category. In the analysis, I focus on non-Hispanic black or African American, Hispanic, and non-Hispanic white registered voters. Registered voters with a party affiliation other than Democratic, Republican, or no party affiliation were grouped in the "other" category. In the analysis, I focus on the first three. Age was calculated with leap year and leap second precision based on birth dates and with reference to February 2019 (when the data was formated).

I compute election-specific voting eligibility based on birth dates, election dates, and registration dates. Those who register to vote for the first time in a Florida county must do so 29 days before the election.[20] Florida allows preregistration after a person's 16th birthday, so that a person may vote in an election occurring on or after its 18th birthday. I hence categorize individuals as eligible to vote at a specific election if they reached their 18th birthday 29 days before bookclosing at that election and if their registration date lies before the election date. I did not distinguish between primary and general elections for assigning eligibility. Even though Florida is a closed-primary state, i.e., only those with a registered party affiliation are eligible to vote in partisan primaries, every registered voter is eligible to vote on nonpartisan offices and ballot issues in primary elections. Voting histories do not distinguish between partisan and non-partisan ballots at the primaries and accordingly show turnout of voters without party affiliation at primary elections. Appendix G includes additional checks for primary elections without non-affiliated voters in the sample.

[20] See https://dos.myflorida.com/elections/data-statistics/voter-registration-statistics/bookclosing/ (last accessed July 2021).

Table B1: Elections included in voting histories.

Type	Primary	General
Midterm	September 5, 2006	November 7, 2006
Presidential	August 26, 2008	November 4, 2008
Midterm	August 24, 2010	November 2, 2010
Presidential	August 14, 2012	November 6, 2012
Midterm	August 26, 2014	November 4, 2014
Presidential	August 30, 2016	November 8, 2016
Midterm	August 28, 2018	November 6, 2018

Registered voters who do not respond to an address confirmation notice or for whom the notice is returned as undeliverable are marked as inactive voters in the Florida registration list. Inactive voters are still eligible and registered to vote. They are purged from the voter registration list and have to reregister after failing to show voting activity or updates to their registration file for two subsequent general election cycles.[21] Given that inactive voters are still eligible to vote and need only show up at the polls, I do not remove them from the analysis sample. In 2018, for instance, I find that a substantial amount of inactive voters actually turned out to vote at the general election. Also, a majority of inactive voters was listed as active in 2016. These might be marginal voters who skip midterm elections. Appendix G includes additional checks without inactive voters in the sample.

Table B1 lists the primary and general elections included in the voting histories. Turnout at the 2018 primary and general elections was update by matching voter IDs to voting histories from the December 2018 voting registration list. The majority of the 67 county supervisors of election do not explicitly record non-attendance at an election. Instead, if a registered voter did not vote at a specific election, there is no record of that person for that election in the voting history. I hence recorded turnout for a registered voter at a specific election only it the person was mentioned for that

[21] See the Florida Statute 98.065 http://www.leg.state.fl.us/statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=0000-0099/0098/Sections/0098.065.html (last accessed July 2021).

election in the voting history of any of the 67 counties. It is important to look up every registered voter in each of the 67 county voting histories to account for prior turnout of those who moved within Florida. To not overestimate non-attendance, I always asses turnout conditional on election-specific voting eligibility in all analyses in the paper and supplementary materials.

Twitter Data

Twitter communication of individuals in the sample was collected using a combination of Twitter's "users/lookup" (Twitter, 2019b) and "statuses/user_timeline" (Twitter, 2019a) API endpoints. Beginning August 1, 2018 all publicly available Tweets (posts) and Retweets (shared posts), including replies, of the 90,832 users in the sample were collected by querying their Twitter IDs via the API. From that point onward a script ran automatically every day, which initiated a lookup for user activity, compared it to previous activity, and collected all new statuses and shared statuses, if any. The script kept track of activity counts every day and always compared back to activity counts of the previous day before collecting any user data. This approach avoided unnecessary redundancies in data collection, putting as little strain as possible on Twitter's servers.

The data collection script was scheduled to run every day at 12 a.m. Central European Time (6 a.m. Greenwich Mean Time 4, Tampa, Florida). Users were queried in random order every day. The actual time and day of statuses and shared statuses was later assigned based on the official time stamp attached to each activity. The script gathered the entire multilingual text of all statuses and shared statuses, each with a maximum of 280 characters. Data covered in the paper and supplementary materials are based on data collection that ran for 243 days between August 1, 2018 and March 31, 2019, assembling 6,379,966 status and shared status activities. The collection process was automatically monitored with programmed alarms based on HTTP status codes. No interruptions in data collection occurred during this time.

Throughout the studied period, 52,715 (58%) registered voters in the sample were active on Twitter. This figure is based on observed statuses as well as dynamic liking and friending behavior, which was collected in addition and similar to the data described above but is not central to analyses in the paper. If we also consider users who received followers, we count 67,396 (74%) active users. Considering activity before the studied period as well (statuses, friends, likes), 88,692 (98%) are active. Inactive Twitter users can be passive users but they can also represent people who abandoned their Twitter account. As concerns the latter, Twitter does have an inactive account policy that indicates the removal of accounts if no login is registered within six months time.[22] This falls well into the period under study and the sample of 90,832 is already cleaned of terminated accounts. In terms of turnout propensities, the inactive (mean = 0.04, sd = 0.84), with regard to the very first definition, and active subpopulation (-0.01, = 0.88) are fairly similar. Considering this, I do not find it plausible that we would find highly engaged/disengaged voters among falsely labeled passive Twitter users more or less politically active on social media than highly engaged/disengaged voters among observed active users if we were to uncover their "true" social media activity. Moreover, passive Twitter users are part of a study population that is seldom included since common means of Twitter data collection actively sample on user activity. For these reasons, inactive users are here taken as passive users who do not participate politically on social media and are not removed from the analysis sample. However, Appendix G includes a version of the multilevel model that focuses on the active subpopulation only.

Auxiliary Data

Approximating individual-level income using per capita income at small-scale census block-group level requires geographical information about voters residence. Latitude and Longitude coordinates of individuals' residence were determined based on reported addresses (city, street, zip code) in the voter registration list using the Bing Maps API (Microsoft, 2018). 63 geocoded addresses yielded low accuracy values

[22] See https://help.twitter.com/en/rules-and-policies/inactive-twitter-accounts (last accessed July 2021).

and were placed outside Florida – these were corrected through manual research. I used the coordinates to identify individuals' census block codes (15 digits) based on TIGER/Line shapefiles from the United States Census Bureau (Macfarlane and Kressner, 2018). Unfortunately, income estimates are not available at the census block level, so I shortened the codes to the block-group level (12 digits). Finally, the codes were used to collect 2017 American Community Survey 5-year estimates of per capita income at census block-group level from the census API and match it to individuals in the sample (Recht, 2019). Census block groups contain between 600 and 3,000 people and ideally around 1,500. While this is a rather crude surrogate for individual-level income, it does capture the larger neighborhood or social setting in which people live. Understood as such, income at the block-group level might be an even better indicator for persons' socioeconomic status than individual self reports (Hersh and Nall, 2016).

Appendix C: Estimation of Target Populations

Figure 1 in the paper compares the non-probability sample to different realizations of the Florida electorate on several characteristics. These realizations are all based on large probability samples.

Estimates of the registered-voter population are constructed from a simple random sample of 100,000 registered voters drawn from the October 2017 Florida voter registration list. The same list was also used to construct the sample. Age groups are constructed so that they align with estimates of the voting-eligible and citizen-voting age population. The grouping of voters into the other demographic categories and parties is described in detail in Appendix B.

The citizen-voting age population is comprised of US citizens age 18 and older. To arrive at estimates for the citizen-voting age population, I begin with the 2017 American Community Survey 1-year subject table on the Florida citizen voting-age population (Walker et al., 2019). In this subject table, only the racial category white excludes Hispanics, i.e., Non-Hispanic whites, all other races include Hispanics. For this reason, the summed total of all race categories exceeds the total citizen voting-age population. However, for comparability with the sample and the registered-voter population as well as to construct the voting-eligible population, it is necessary to correct for this by grouping all Hispanics together and removing excess Hispanics from the other categories. To achieve this, I rely on the 2017 1-year Florida Public Use Microdata Sample (PUMS) (United States Census Bureau, 2017). PUMS data code both race categories and Hispanic ethnicity, which allows to estimate the proportion of Hispanics among race groups and subtract it accordingly (Thaler, 2019). After removing excess Hispanics from black and other voters in the subject table the summed total of race groups aligns with the total citizen voting-age population. Age in the American Community Survey subject table is also not grouped in a way required to estimate and ensure comparability to the voting-eligible population. Hence, I fully rely on PUMS data to construct citizen-voting age population estimates for appropriate age groups. Estimating the distribution of all demographic characteristics based only on PUMS yields very similar estimates.

The citizen-voting age population is not necessarily the same as the voting-eligible population. In Florida, the former includes felons and mentally incapacitated persons who are not eligible to vote (Fraga, 2018). The correctional population, however, is not a random sample from the Florida population and largely comprised of black males (Shannon et al., 2017). To estimate characteristics of the voting-eligible electorate, I start from the above estimates of the citizen-voting age population and adjust them for the correctional population. This mirrors the current gold standard in estimating the voting-eligible population (Fraga, 2018; McDonald, 2017). I use data from the 2017–2018 Annual Report of the Florida Department of Corrections to quantify the demographic distribution of Florida's correctional population (Florida Department of Corrections, 2018). The correctional distribution includes both prisoners as well as those on parole and probation who are also barred from voting in Florida. The annual report lists corresponding population totals along categories as depicted in Figure 1 in the paper. Estimates for the voting-eligible population are constructed by removing these totals from the respective group-estimates of the citizen voting-age population.

Appendix D: Estimation of Voting Propensities

Formally, the two-parameter item response theory model[23] can be written as:

$$
\begin{aligned}
y_{ij} &\sim \text{Binomial}(1, \pi_{ij}) \\
\pi_{ij} &= \text{logit}^{-1}(\alpha_j(\theta_i - \beta_j)),
\end{aligned}
$$

where y_{ij} is person i's decision to vote or abstain in election j, provided eligibility, and assumed to follow a Binomial distribution. The probability to turnout at a specific election π_{ij} is a function of the latent trait θ_i, the voting propensity, the difficulty parameter β_j, and the discrimination parameter α_j. In this context, the difficulty parameter locates the threshold at which voting is more likely than abstaining and the discrimination parameter allows each election to additionally have a different weight in the latent trait (see Fowler et al., 2008). Accordingly, the participation decision is treated as distinct for every election so that each election contributes differently in discriminating between low and high-propensity voters.

Note that all of the parameters are unobserved and jointly estimated based on the observed participation choices. To identify the model, it is hence necessary to explicitly specify the direction, location, and scale of the latent dimension. For this, I rely on a Bayesian approach (Levy and Mislevy, 2016) with hierarchical prior

[23] Prior research in this context shows that a one-dimensional two-parameter solution is preferable to two-dimensional, one-parameter (Rasch), or three-parameter models (Fowler et al., 2008; Spahn and Hindman, 2014).

information as follows:

$$\begin{aligned}
\alpha &\sim \text{Lognormal}(0, \sigma_\alpha) \\
\beta &\sim \text{Normal}(\mu_\beta, \sigma_\beta) \\
\theta &\sim \text{Normal}(0, 1) \\
\mu_\beta &\sim \text{Cauchy}(0, 5) \\
\sigma_\alpha &\sim \text{Cauchy}(0, 5) \\
\sigma_\beta &\sim \text{Cauchy}(0, 5).
\end{aligned}$$

The latent trait θ is given a standard normal distribution to identify both location and scale. This ensures that the posterior will yield voting propensity estimates with a mean around 0 and a variance around 1. The discrimination parameter α_j is constrained to be positive via a lognormal prior to fix the direction. This prohibits elections that are "easier" for voters with lower voting propensity, which makes sense as we expect the relationship between observed participation choices and the underlying voting propensity to be monotonically increasing. In addition, the prior scales relative to the voting propensities. Similarly, the difficulty parameter β is given a normal prior. For efficiency reasons, the location parameter of β, μ_β is itself given a prior (centered parameterization). The scale parameters of α and β, σ_α and σ_β are also given priors. These hyperparameters are determined mainly from the data with weakly informative Cauchy priors, i.e., proper but barley informative with reference to the likelihood. Note that the priors for σ_α and σ_β are constrained to be positive by their declarations, effectively yielding half-Cauchy priors.

The data used for estimating voting propensities is cross-classified, holding one row for each voter-election pair. Rows for elections where voters were not eligible to vote are omitted amounting to 928,460 observations. The participation decision is coded as a binary choice. Data for registered voters who were ultimately removed from the analysis sample (see Appendix A) was included in estimating the item response theory model. These individuals hold much and valuable information about participation decisions that can only improve estimates of θ.

Another reason for choosing a Bayesian approach is given by the amount of parameters to be estimated. Since θ is estimated for every individual, we face more than 100,000 parameters together with the election-specific difficulty and discrimination parameters. With such complex models where the number of parameters is a function of sample size, Markov chain Monte Carlo approaches are likely more valid and efficient than conventional maximum likelihood estimators (Clinton et al., 2004). The model was implemented using Stan, a program for Bayesian statistical inference with Markov chain Monte Carlo sampling (Carpenter et al., 2017). Figure D1 shows the Stan code used to estimate the model. I ran four parallel Markov chains from random starting values with 2,000 iterations each. In each chain, The first 1,000 warm-up draws were discarded yielding estimates based on 4,000 posterior draws. In order to detect potential non-convergence and biased inference, I checked several diagnostics: the potential scale reduction statistic Split $\hat{R}$, effective sample size, autocorrelation plots, traceplots, divergent transitions, and energy plots. None indicated any pathological behavior in the chains. Detailed results of these diagnostics are available upon request.

Table D1 reports posterior medians and credible intervals for the difficulty and discrimination parameters. The estimated difficulty parameters are largely consistent with Campbell's (1960) concentric circle model. Presidential general elections appear as easiest, where even low propensity voters are still likely to participate. This is followed by midterm general elections where the participation threshold is broadly located at the center of the propensity scale. Primary elections are mostly the domain of higher propensity voters. At the same time, there are noticeable differences between elections even within election types. The estimated discrimination parameters additionally show that elections vary in how informative they are about individual voting propensities, but not systematically along election types. It is thus necessary to include multiple elections in an assessment of voter engagement to account for contextual factors as well as person-, and election-specific idiosyncrasies, as similarly

```
data {
  int<lower=1> I; // number of elections
  int<lower=1> J; // number of voters
  int<lower=1> N; // number of observations
  int<lower=1, upper=I> ii[N]; // observed election for observation n
  int<lower=1, upper=J> jj[N]; // observed voter for observation n
  int<lower=0, upper=1> y[N]; // observed turnout for observation n
}

parameters {
  vector<lower=0>[I] alpha; // discrimination parameter for election i
  vector[I] beta; // difficulty parameter for election i
  vector[J] theta; // ability for voter j
  real mu_beta; // average election difficulty
  real<lower=0> sigma_alpha; // scale of (log) discrimination
  real<lower=0> sigma_beta; // scale of difficulties
}

model{
  vector[N] pi;

  // priors on hyperparameters
  mu_beta ~ cauchy(0, 5);
  sigma_alpha ~ cauchy(0, 5);
  sigma_beta ~ cauchy(0, 5);

  // priors on parameters
  alpha ~ lognormal(0, sigma_alpha);
  beta ~ normal(mu_beta, sigma_beta); // centered parameterization
  theta ~ std_normal();

  // likelihood
  for (n in 1:N)
    pi[n] = alpha[ii[n]] * (theta[jj[n]] - beta[ii[n]]); // centered parameterization
  y ~ bernoulli_logit(pi);
}
```

Figure D1. Stan code for two-parameter logistic item response theory model.

noted before by Spahn and Hindman (2014). Item response curves shown in Figure D2 visualize that neither one election nor a specific election type are sufficient to allow a clean separation between low-propensity, marginal, and highly engaged voters. The Figure also shows the distribution of estimated voting propensities in the sample together with posterior medians and 80% credible intervals of the θ parameters.

Table D1: Posterior medians and 95% credible intervals for difficulty and discrimination parameters.

Election	Difficulty (β)	Discrimination (α)
2006 Primary (midterm)	1.37 [1.35, 1.39]	1.94 [1.89, 2.00]
2006 General (midterm)	0.23 [0.21, 0.24]	1.81 [1.76, 1.85]
2008 Primary (presidential)	1.41 [1.39, 1.44]	2.07 [2.02, 2.13]
2008 General (presidential)	−1.32 [−1.36, −1.29]	1.54 [1.49, 1.58]
2010 Primary (midterm)	1.07 [1.06, 1.09]	2.71 [2.64, 2.78]
2010 General (midterm)	0.08 [0.07, 0.10]	2.36 [2.31, 2.42]
2012 Primary (presidential)	1.12 [1.10, 1.13]	2.45 [2.39, 2.51]
2012 General (presidential)	−1.02 [−1.04, −1.00]	1.88 [1.83, 1.93]
2014 Primary (midterm)	1.17 [1.16, 1.19]	2.50 [2.45, 2.56]
2014 General (midterm)	−0.13 [−0.14, −0.12]	2.15 [2.10, 2.19]
2016 Primary (presidential)	0.78 [0.76, 0.79]	2.18 [2.14, 2.22]
2016 General (presidential)	−1.52 [−1.54, −1.49]	1.84 [1.79, 1.88]
2018 Primary (midterm)	0.55 [0.54, 0.56]	1.97 [1.93, 2.00]
2018 General (midterm)	−0.57 [−0.58, −0.55]	1.47 [1.44, 1.49]

Appendix E: Classification of Political Twitter Posts

Text Processing

Single Twitter posts describe the unit of analysis in all text-based operations conducted in this paper. All collected Twitter posts were pre-processed before even assembling the dictionary. The purpose of this was mainly to reduce dimensionality and remove features which were not considered relevant with regard to political participation or complicated further analyses.

I started with removing all emojis using various emoji dictionaries. A more common approach would have been to remove all non-ASCII characters but this would have resulted in the removal of posts in Chinese or other languages. Next, I removed all RT (retweet) tags and URLs. While URLs can link to political content, too, classifying domains as political is a whole separate challenge that would probably introduce uncertainty above all else. After all, the exact content type a URL points

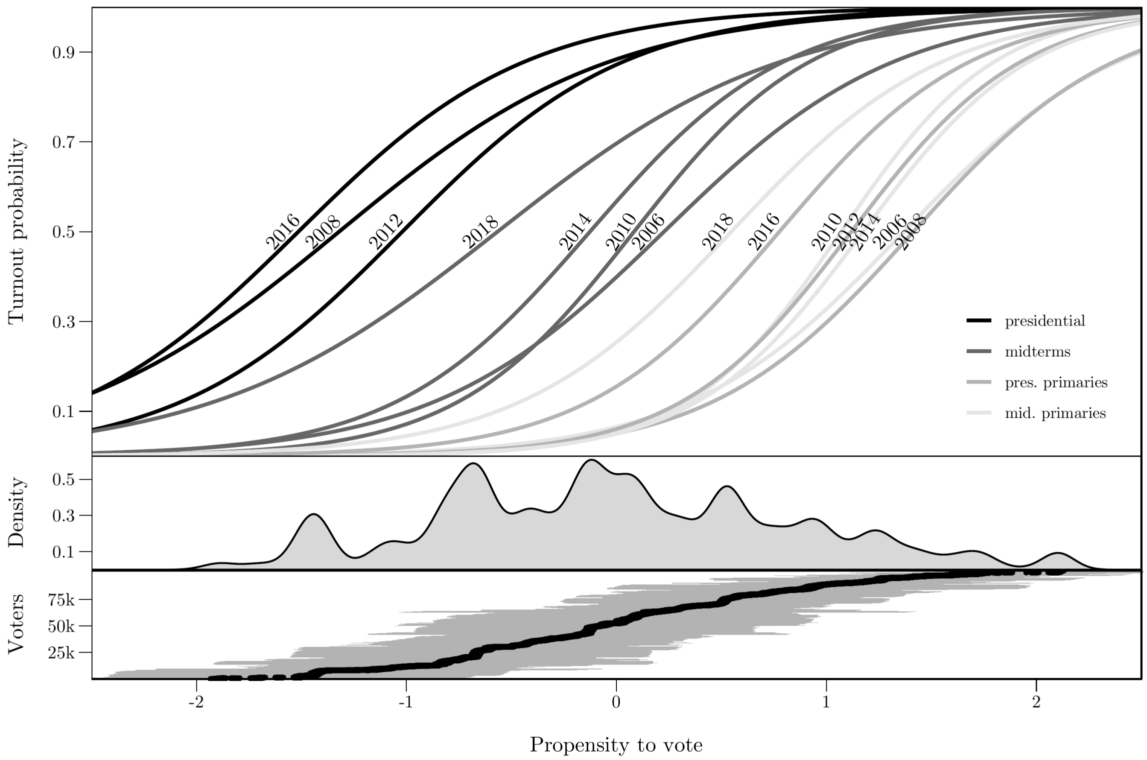


Figure D2. Item characteristic curves for measurement model of voting propensity.

to can only be determined reliably by processing and classifying the respective page's content. More importantly, we would not reasonably expect the behavior to post a URL with political content to differ from the behavior to post text with political content to such an extent that it biases inferences about social media-based participation across the electorate and in various subgroups. Following these first processing steps, the language of each post was determined using Google's Compact Language Detector 2 (Ooms and Sites, 2018). The texts were then transformed to lower case and stopwords, numbers, punctuation, and redundant whitespace was removed.

I rely on recently developed software to assess the potential impact of some of these processing steps on further analyses (Denny and Spirling, 2018). The preText algorithm was used to pre-process a random sample of 1,000 statuses in 128 different ways, including combinations of the use of ngrams (which I introduce in a later step), stemming, transformation to lower case, and the removal of stopwords, punctuation, numbers, and infrequent terms. Figure E1 plots the conditional effect of each processing step on the average preText score (normalized average rank order difference). Positive and statistically significant coefficients point towards increased risk of obtaining unusual results after applying the respective processing step. Here, this is the case for the removal of stopwords and infrequent terms. In a supervised context, however, the removal of stopwords is theoretically justified. The removal of stopwords serves not only to improve computation time by reducing the complexity of the vocabulary but also to get rid of meaningless terms (noise features) that have the potential to introduce misclassification due to overfitting to these terms. Infrequent terms were not removed as they might be substantively meaningful. Stemming was not applied to preclude words from being altered in such a way that substantive meaning was removed.

Computer-Assisted Keyword Discovery

King and colleagues' (2017) algorithm for computer assisted keyword discovery suggests a division of labor between the computer and a human coder. They show that human coders perform poorly at assembling a large collection of keywords that potentially represent a concept of interest – a job much better done by a machine. At the same time, using detailed contextual knowledge, human coders clearly outperform computers in recognizing and filtering appropriate keywords.

The algorithm optimizes on this. The researcher provides a set of documents that represent the concept of interest, known as the reference set, and a set of documents that may hold additional keywords of interest but which does not overlap with the reference set, also known as the search set. The algorithm then samples from these sets, takes texts' membership in reference and search set as outcome variable, and fits a range of classifiers (e.g., Naive Bayes, Logit, Support Vector Machine, Random Forest) to this training set. Documents from the search set, which were (mis)classified

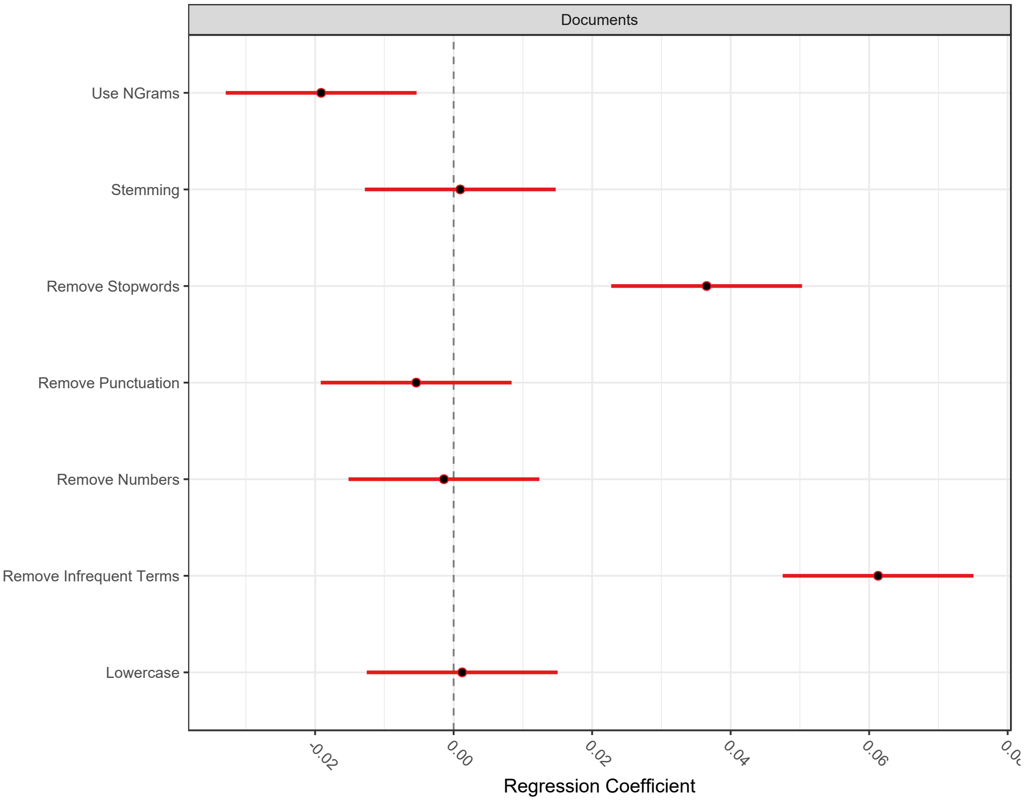


Figure E1. Sensitivity to text processing steps.

into the reference set identify the target set, i.e., the subset of the search set that likely contains yet undiscovered keywords representing the concept of interest. These keywords are subsequently ranked by how well they discriminate between the target and non-target set. A human coder evaluates the resulting keyword list and uses subject-specific knowledge to build and expand a dictionary. Finally, the keyword list can be used to refine and extend the initial reference set and iterate over the algorithm again in order to improve or further expand the dictionary.

I rely on this algorithm[24] to create a problem- and context-specific dictionary for binary classification of Twitter posts as either reflecting political engagement or

[24] The python code (keyword_algorithm.py, version 1.1) for implementing the algorithm was retrieved from the Harvard Dataverse https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FMJDCD (last accessed July 2021).

not. As reference set, I used a collection of Twitter posts that were manually coded according to the definition of social media-based political participation outlined in the paper. The coding was based on a simple random sample of 4,000 posts. The sample was drawn from a pool of 10,773,862 posts, which spans the entire period of investigation but also includes posts that were created before August 2018. Non-English posts in the sample were translated before the manual coding using Google's translator. Two coders categorized 541 (13.5%) posts as political participation. The interrater reliability based on Cohen's κ is 0.91 (95% confidence interval = 0.89, 0.93). Table E1 shows examples of posts in the initial reference set.

Throughout the period of investigation, Twitter's Search API was queried daily for up to 5,000 posts coming from Florida. Florida was located as the source of posts using "Florida, USA" as place identifier and the following set of bounding coordinates: sw.lon $= -87.63490$, sq.lat $= 24.39631$, ne.lon $= -79.97431$, ne.lat $=$ 31.00097. The resulting 728,089 posts were used as search set.

The algorithm for keyword discovery was run using these problem- and context-specific reference and search sets. After a first run, I evaluated the suggested keywords and used the initial selection to categorize a simple random sample of 100,000 posts. Among posts categorized as political, I retrieved a random sample of 2,500 and added them to the reference set. I then iterated over the algorithm again using the updated reference set. After three iterations of the algorithm, I collected 428 keywords and was not able to gain further terms reflecting the concept of interest.

At this point the dictionary was still messy. It certainly included terms that reflected the concept of interest. But some of these terms were so ambiguous that they were likely to produce a substantial amount of false positives. There will always be additional keywords that come to mind as fitting the concept of interest and as clearly missing in the final list. The task at hand, however, is not to arrive at an exhaustive list but to balance the list of keywords in a manner that minimizes statistical bias. To tune the dictionary, I relied on another sample of 4,000 manually coded Twitter posts

Table E1: Examples of political posts in initial reference set.

1. H.R.973 - SSI Fairness Act of 2015. If they get this passed we will continue to lose 40% of our social security.
2. @BernieSanders I've been without health insurance for years since losing a full time job.
3. @Lanna70115 @TolerForPres @TB_Times @BruceWStanley Don't bother, just flush your vote down the nearest toilet.
4. @nikkifried You are all we have in Florida, Commissioner, be the voice of all progressives and we will have your back.
5. The push for renewable energy is most emphasized in the reddest states.
6. @MoveOn The Supreme Court is a political institution and just as dangerous as Trump!
7. The #gop isn't even able to cut taxes. Makes you wonder why they even try. #GOPTaxScam
8. This #MuslimBan is just an #alternativefact that we all misunderstood, right?
9. @AlexPopov @jimgeraghty What is your understanding of "normal conservative"? This country has the most liberal abortion policies anywhere - including Europe.
10. RT @Jim_Jordan: Social media companies have incredible control over information. If they can restrict speech, they have unlimited power to influence elections and public policy. After all, social media is part of every American's life.
11. RT @jimv_ross: #GunControlSavesLives
12. They should have asked Cohen if a Trump hater/DNC was behind his despicable testimony. What was he getting promised?
13. #IBelieveChristineBlaseyFord
14. Nancy Pelosi has clearly passed her expiration date.
15. RT @holden2018: Francis Rooney and his colleagues in Congress aim to take away healthcare from millions of Americans without even offering an alternative. That's just wrong and has nothing to do with legislating. #SWFL deserves better.
16. Los republicanos retienen el control del Senado de EEUU. Los demócratas retoman el control de la Cámara de Representantes.
17. @SenKamalaHarris It was a just cause. Unlike you idiots of today, FDR had Americans in his best interest. The japanese were spies.
18. @marcorubio you should maybe sit this one out, you take money from the NRA.
19. RT @ellievan65: Senators voted in support of the resolution to end the war in Yemen. Marco Rubio, an outspoken defender of human rights, was one of those voting against ending the war.
20. @RepThomasMassie @JoePerticone Expect no congressional approval for the use of troops in the US. Does Whiskey rebellion ring a bell?
21. @democrat_proud I'd rather be prepared for war. We already appear weak to Europeans.
22. @TravelGov, I have a valid passport but I am repeatedly running out of visa pages.
23. The largest student loan company in America is being sued and under investigation by the Government.
24. RT @AOC: What? People accept when policy proposals that fight income inequality are obstructed? We can win public sentiment and stand our ground without having to be scared by GOP information.
25. I will give the first dollar for the wall. To Trump.
26. Já tentou conversar com um Trumpista?!
27. RT @LEBassett: 13% of all maternal deaths globally are accounted for by unsafe abortion. Trump's reinstating the 'Global Gag Rule' will be deadly.
28. @angelcintronjr Do you support $18 minimum wage? I do!
29. RT @NancyLeeGrahn: This may end in a recount, @AndrewGillum is now only down 49% to 49.7% and needs to un-concede. We need to find his winning votes that mysteriously disappeared in coincidentally Black districts.
30. President Obama, because of you we know we may someday have it again. Thank you for making America proud. Happy Birthday. #ObamaDay
31. RT @DavidCornDC: @realDonaldTrump should release his tax returns today.
32. @realDonaldTrump President Trump, I feel safer with you in charge. Keep up the good work. I just love you.
33. Vote for what you believe in and get informed on each candidate. Voting the party line all the way through is ignorance.
34. Those who scream the loudest have the most to hide. #LokThemAllUp #DrainTheSwamp #MAGA #QANon #WWG1WGA
35. RT @FLGovScott: Our law enforcement officers are working hard to keep people safe. We have more than 570 state troopers assigned to the Panhandle and Big Bend area of Florida to assist with response and recovery.

Note: Examples have been paraphrased to prevent identification of individuals in the sample.

Table E2: Confusion matrix for best-performing dictionary applied to validation set including 4,000 posts.

		Predicted	
		0	1
Actual	0	3462	51
	1	87	400

to test different versions. The best-performing version of the dictionary scored 96.6% accuracy applied to the validation set. The confusion matrix is shown in Table E2. The true positive rate (sensitivity) is 82% (18% false negatives), the true negative rate (specificity) is 98.5% (1.5% false positives). I optimized the dictionary with regard to false positives while keeping false negatives balanced. Missing some instances of political engagement (false negatives) means that we err more on the conservative side while falsely attributing political engagement brings us closer to the error we actually try not to make – overstating political engagement.

The chosen version of the dictionary contains 331 keywords (uni-, bi-, and trigrams) and is shown in Table E3. Even though text processing and keyword discovery was multilingual, only few non-English keywords made it into the dictionary. However, @mentions and hashtags are often language-neutral and allow to classify multilingual posts as well. Using Quanteda's (Benoit et al., 2018) "dfm" function, the dictionary was applied to classify a sparse document-feature matrix representation of 6,379,966 status and shared status activities. This dictionary-based classifier categorized 1,525,672 (24%) posts throughout the period of investigation as political engagement. 98.78% of these posts were in English language, followed by Spanish with 1.15%. Table E4 shows examples of posts categorized as political engagement. These examples come from the full spectrum of estimated voting propensities (mean = 0.08, sd = 1.3).

Table E3: Dictionary of social media-based political participation.

#americafirst	#antitrump	#backfiretrump	#betoforsenate	#bluetsunami
#bluewave	#bluewave2018	#bordersecurity	#brettkavanaugh	#bringithome
#buildthatwall	#buildthewall	#confirmkavanaugh	#democrats	#draintheswamp
#electionday	#endtheshutdown	#fairtax	#fakepresident	#flapol
#flgovdebate	#floridaelection	#floridaprimaries	#gop	#gopdebate
#govote	#guncontrol	#gunsense	#ibelievechristineblaseyford	#impeachtrump
#istandwithbrett	#ivoted	#kavanaugh	#kavanaughhearings	#kavanaughvote
#maga	#michaelcohen	#midterms	#muellertime	#nationalvoterregistrationday
#nevertrump	#paintourcountryred	#potus	#protrump	#redwaverising
#republicans	#scotus	#shawforflorida	#sotu	#speakerpelosi
#stopkavanaugh	#thewall	#thisisnotdemocracy	#traitortrump	#trump
#trumpaddress	#trumpresign	#trumprussia	#trumpshutdown	#vote
#votebeto	#voteblue	#votedem	#votegillum	#votered
#voteredtosaveamerica	#votethemout	#votingrights	#vpdebate	#walkaway
#walkawayfromdemocrats	#whatsatstake	@amyklobuchar	@andrewgillum	@barackobama
@bensasse	@berniesanders	@betoorourke	@brianschatz	@chuckgrassley
@corybooker	@dhsgov	@fladems	@flgovscott	@flotus
@gop	@gopchairwoman	@govhowarddean	@govmikehuckabee	@govrondesantis
@hillaryclinton	@housedemocrats	@housegop	@jeffflake	@jeffmerkley
@jimjordan	@kamalaharris	@lindseygrahamsc	@lisamurkowski	@marcorubio
@mattgaetz	@momsdemand	@nancypelosi	@nelsonforsenate	@ocasio
@ocasio2018	@orlandomayor	@potus	@presssec	@realdonaldtrump
@repadamschiff	@repmarkmeadows	@repmattgaetz	@repswalwell	@repthomasmassie
@rondesantisfl	@sarahpalinusa	@scottforflorida	@secnielsen	@secpompeo
@senatedems	@senategop	@senatemajldr	@senatorcollins	@senbillnelson
@senblumenthal	@senfeinstein	@sengillibrand	@senjoemanchin	@senjohnmccain
@senkamalaharris	@sensanders	@senschumer	@sentedcruz	@senwarren
@senwarrens	@senwhitehouse	@speakerpelosi	@speakerryan	@stabenow
@staceyabrams	@statedept	@tedcruz	@tedlieu	@thedemocrats
@votersincharge	@vp	@whitehouse	administration	administrations
alexandria ocasiocortez	america first	amy klobuchar	arming teachers	ballot
ballots	bernie	beto orourke	bipartisanship	blasey
blasey ford	blue wave	border security	border wall	brett kavanaugh
brett kavanaughs	brian kemp	build the wall	buildthedamnwall	chief staff
clinton	clintons	congress	congressional	congressman
congressmen	congresswoman	congresswomen	constituents	dem
democrat	democratic	democrats	dems	desantis
disenfranchised	disenfranchisement	donald trump	elect	elected
elections	élections	electoral	electoral college	electoral system
electorate	elegir	enfranchised	enfranchisement	federal
feinstein	gaetz	george bush	gillum	gobierno
gop	gov	governor	govmt	govt
grand old party	granted immunity	grassley	gubernatorial	gun control
gun law	gun laws	gun lobby	hillary	house judiciary
housegop	impeach	impeached	impeachment	jahana hayes
jeff flake	joe manchin	judiciary	judiciary committee	justice system
kavanaugh	kavanaughs	klobuchar	kyrsten sinema	lawmaker
leftwing	legislation	legislative	legislator	legislature
lindsey graham	maga	make america great	makeamericagreatagain	making america great
manafort	marco rubio	massgovernor	matt gaetz	maxine waters
mayor	mcconnell	mick mulvaney	midterm	midterms
mitch mcconnell	mueller	muellers	murkowski	national emergency
obama	obamacare	obamas	ocasiocortez	parties
partisan	partyline	paul manafort	pelosi	pence
pences	policymaker	policymakers	politician	politicians
politics	politique	politisyen	polling	polls
pompeo	potus	president trump	primaries	public office
rep	representatives	represented	republican	republicano
republicanos	republicans	rightwing	rubio	sanders
scotus	sec nielsen	sen	sen judiciary	senado
senador	senate	senategop	senatemajldr	senator
senator collins	senator john	senators	senorrinhatch	senrickscott
sessions	shutdown	statedept	susan collins	taxpayer
taxpayers	ted cruz	term limit	term limits	trump
trumpexpress	trumprussia	trumps	trumpshutdown	turnout
union address	vote blue	voter	voters	voting
white house				

Validation of Keyword-Based Classifier

Though widely applied in the social sciences, the reputation of dictionary methods for classification is ambiguous. This is largely because off-the-shelf dictio-

Table E4: Examples of posts categorized as political based on dictionary.

1. RT @GKCdaily: We need to make politics more local. Keep the politicians near enough to hold them accountable.
2. @RepAdamSchiff I cannot describe the shock at those who exploit tragedies to their own political benefit. We don't need gun legislation. We need legislators doing everything to ensure children have fathers in the home!
3. Don't allow @POTUS to roll back clean car standards. Make @EPA and @USDOT protect them. via @NRDC.
4. RT @Lawrence: The absentee ballot stealing scheme is no new invention. This secretly happened before all over the country?
5. @KamalaHarris People in OUR country are dying of hunger and homelessness. Why not care about them? Those parents should be deported and arrested in their own country. Get out of left field and into REALITY!
6. @realDonaldTrump They are chipping away slowly with nonsense and it may not end up well. You have to do something, it won't jeopardize your 2nd term. Bring out the big guns and show proof against the accusers for THEIR wrong doings. This has to happen NOW. #ConcernedLatinoTrumpSupporter
7. The Trump Economy Continues to rise #MAGA #VoteRED #walkaway #liberalismisamentaldisorder
8. @GOP @realDonaldTrump useless leadership has lead real wages to plummet.
9. Name one bill @SenSanders has written, co-sponsored or been involved in that has passed. (Besides the PO renaming.) I'm serious, what EXACTLY has he done for us or anybody but himself?
10. RT @RepMcGovern: Trump steaks - shut down. Trump magazine - shut down, Trump university - shut down, Trump casinos - shut down, Trump airlines - shut down. @realDonaldTrump kept his promise to run government as he ran his businesses: shut down.
11. Check his Senate record, yes, he is ineffective!
12. @AnnCoulter Because they have jobs, the majority of the people I know cannot afford healthcare under Obamacare. We live in a messed up country where lazy people who don't work get free healthcare.
13. I disagree with Tomi Lahren on her stance on right to life and other things. However, it's justified to ask whether the "prison reform" will be good for law-abiding Americans or whether it will just be a bid to make the GOP look hip?
14. @altNOAA @wthworld911wtf @realDonaldTrump Maybe you will understand why the southern border is completely open when you check back with border patrol or got there and camp for a few days.
15. The late term baby killer Democratic politician Bill's, Race Plan parenthood, and Abortion are a racist scam!
16. @GOP @VP I do not consider coal waste dumping in rivers, selling public lands to oil drillers and making our air and water dirtier an improvement of government functions.
17. RT @SpeakerPelosi: Today, Congress #ActOnClimate as we name those who serve on the Select Committee on the Climate Crisis.
18. @ewarren Democrats want you flipping burgers your whole life, so you can't challenge their system. This is not supposed to be a career job and minimum wage is not supposed to do that.
19. Shape Our Schools my__pcs @gchery @fladems baynews9 @ North Greenwood Recreation and Aquatic Complex
20. @HowardSchultz Life is good now and much better than in years past, why do we need so much change? The sure does seem to be running well country
21. @lisamurkowski Gambling debts, connections to Trump money, and Lies. Kavanaugh does not support women rights, he is not who we want.
22. RT @senatemajldr: I agree with @POTUS: #Coal needs to be part of our feature and is right here in #Kentucky. It employs thousands of hardworking Americans, is affordable, reliable, and powers the lights in our homes. Coal helped fuel our country's greatness
23. @johnnysez1 @SenSanders @CarmenYulinCruz Audit Trump and his staff!
24. @WayneDupreeShow People paying outrageous amounts for insurance because of Obamacare has devastated the middle class and mid-sized businesses. The judge has no clue, ridiculous, that should be unconstitutional.
25. @dguy53 To vote against Pelosi's leadership right now benefits trump. All the Republican anti-Pelosi advertisements, they fear her.
26. RT @SenatorLeahy: We could re-open the government today if Senate Republicans wouldn't hold Americans hostage for a wall. @senatemajldr said that we would pass a bill to fund the government already a month ago.
27. RT @AndrewGillum: Paying teachers what they're worth matters for our children's education and Florida's future.
28. RT @RWPUSA: No one would elect someone like Florida Gov. Scott with such company involvements and financial conflicts of interest.

Note: Examples have been paraphrased to prevent identification of individuals in the sample.

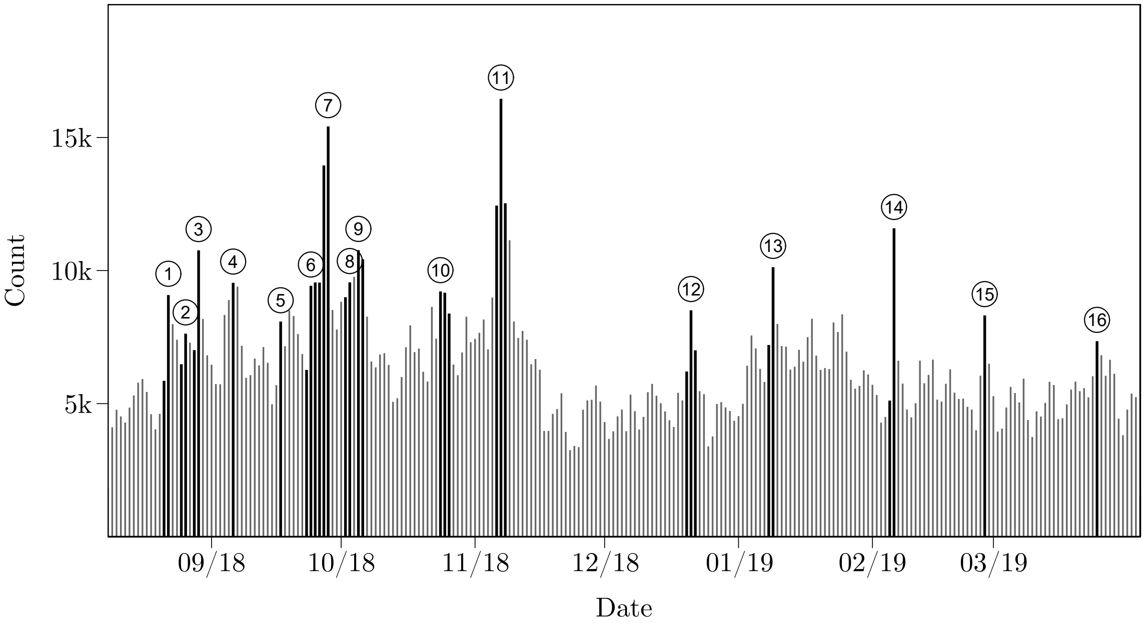
Table E5: Confusion matrix for dictionary-based classification applied to test set including 4,000 posts.

		Predicted	
		0	1
Actual	0	3459	48
	1	72	421

naries are often applied out of context, not validated, and not compared against alternative approaches (Grimmer and Stewart, 2013; Muddiman et al., 2019). The dictionary used in this paper is custom-built and problem- and context-specific. This section presents the validation of the dictionary based on human gold standards and compares its classification performance to a deep learning approach.

If categories are known in advance, which is the case in this project, a classifier's performance should be shown to reliably replicate human coding (Grimmer and Stewart, 2013). In addition to the training and validation sets used above, another simple random sample of 4,000 Twitter posts was manually coded. These posts serve as test set to evaluate the performance of the keyword-based classifier against human gold standards. Using the dictionary from Table E3, the keyword-based classifier scores a 97% accuracy applied to the test set, which is well above the no information rate (88%). The confusion matrix is shown in Table E5. The true positive rate (sensitivity) is 85% (15% false negatives), the true negative rate (specificity) is 99% (1% false positives). In addition, a time series of social media-based participation and external events (see Figure E2) signals predictive validity (Grimmer and Stewart, 2013). Peaks in political engagement correspond with notable political events.

A keyword-based classifier is only one of many approaches that can be chosen for the binary classification task at hand. To check whether the keyword-based approach is optimal for dealing with the problem confronted with, I compare it to a state of the art supervised machine learning method. Supervised methods are the chief competitor of dictionary methods when dealing with classification of known cat-



1 Michael Cohen and Paul Manafort convicted.

2 Mass shooting in Jacksonville. John McCain dies.

3 Midterm primary elections. Florida gubernatorial candidate Ron DeSantis involved in “Monkey this up” controversy.

4 Reports of Trump administration working against President’s agenda. Leaked documents on Supreme Court nominee Brett Kavanaugh’s time in White House.

5 Christine Blasey Ford accuses Brett Kavanaugh of sexually assaulting her in the 1980’s

6 More sexual misconduct claims against Brett Kavanaugh.

7 Kavanaugh hearing.

8 Washington Post journalist Jamal Khashoggi murdered inside the Saudi consulate in Istanbul.

9 Trump demands Kavanaugh confirmation. Kavanaugh confirmed as Supreme Court justice.

10 Mail bombing attempts targeting several political elites.

11 Midterm general elections.

12 Jim Mattis resigns. Congress fails to agree on a budget, partial federal government shutdown begins.

13 Donald Trump’s first TV address to the nation from the Oval Office.

14 State of the Union address.

15 Michael Cohen’s congressional testimony. The House blocks Trump’s declaration of a national emergency along the southern border.

16 Third reported suicide of a relative of a school shooting victim in one week.

Figure E2. Social media-based participation and external events over time.

Table E6: Confusion matrix for deep learning-based classification applied to test set including 4,000 posts.

		Predicted	
		0	1
Actual	0	3269	238
	1	455	38

egories (Grimmer and Stewart, 2013). Using the same split ratio as above (4,000 train, 4,000 validation, 4,000 test), I train, evaluate, and validate a sequential (linear stack of neural network layers) deep learning model with the Keras library (Chollet and Allaire, 2018). Each set was tokenized prior to training the model, keeping only the most common 10,000 words. Texts were then transformed into sequences of integers and padded to the same length of maximum 150. The model itself consists of one initial embedding layer followed by several densely-connected neural network layers with intermittent dropout to prevent overfitting. Against the test set, the deep learning approach scores a 83% accuracy, falling under the no information rate (88%). The confusion matrix is shown in Table E6. The true positive rate (sensitivity) is 8% (92% false negatives), the true negative rate (specificity) is 93% (7% false positives). Another run with a different split ratio (8,000 training, 2,000 validation, 2,000 test) yields similar results with 84% accuracy, which also scores below the no information rate (87%). The respective confusion matrix is shown in Table E7. The true positive rate (sensitivity) is 6% (94% false negatives), the true negative rate (specificity) is 95% (5% false positives). As it stands, the keyword-based classifier outperforms the supervised method. Sure, given much more training data, the deep learning approach may catch up to the keyword-based classifier. But it is unclear how much data is required and lacking the resources to additionally hand code thousands of Twitter posts pursuing this path any further is out of question. The success of the keyword-based classifier probably stems from the human evaluation and input while building the dictionary. This step adds detailed problem-specific knowledge absent to a deep learning approach.

Table E7: Confusion matrix for deep learning-based classification trained on larger training set and applied to test set including 2,000 posts.

		Predicted	
		0	1
Actual	0	1661	80
	1	244	15

Appendix F: Estimation of Multilevel Model

Multilevel logistic regression is used to derive precise estimates of social media-based participation for small demographic and political subgroups across the electorate. This approach is the preferred choice when interest is in group-specific estimates or in variation of individual-level predictors across groups (Gelman and Hill, 2007). Also, when groups get small, for instance due to group-interactions, multilevel modelling yields more precise estimates than classical regression by partially pooling estimates across groups.

Formally, a multilevel model with varying slopes and intercepts can be written as:

$$y_i \sim \text{Binomial}(1, \pi_i)$$
$$\pi_i = \text{logit}^{-1}\left(\sum_S \alpha_{S[i]} + \sum_S x_i \beta_{S[i]}\right)$$

where y_i is a person i's observed decision to participate politically on social media and assumed to follow a Binomial distribution. The probability of social media-based participation π_i is characterized by group-specific intercepts α over a set S of demographic and political groups and their two-way interactions. Accordingly, $\alpha_{S[i]}$ represents the intercept for the subgroup in S that includes unit i. This corresponds to the model in the main paper. For additional analyses in Appendix G, a vector of

Table F1: Variables in the multilevel model.

Stan declaration	Description	Type	Number of groups	Coefficient in model
y	Social media-based participation (1 = observed, 0 = not observed)	Output variable	–	–
theta	Voting propensity (only included in additional analyses in Appendix G)	Individual-level predictor/ varying-slope	–	Part of $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$
sex	Sex	Varying intercept	2	α_1
race	Race	Varying intercept	4	α_2
age	Age	Varying intercept	4	α_3
party	Registered party affiliation	Varying intercept	3	α_4
income	Block group per-capita income	Varying intercept	5	α_5
sex_race	Sex × race interaction	Varying intercept	$2 \times 4 = 8$	α_{1_2}
sex_age	Sex × age interaction	Varying intercept	$2 \times 4 = 8$	α_{1_3}
sex_party	Sex × party interaction	Varying intercept	$2 \times 3 = 6$	α_{1_4}
sex_income	Sex × income interaction	Varying intercept	$2 \times 5 = 10$	α_{1_5}
race_age	Race × age interaction	Varying intercept	$4 \times 4 = 16$	α_{2_3}
race_party	Race × party interaction	Varying intercept	$4 \times 3 = 12$	α_{2_4}
race_income	Race × income interaction	Varying intercept	$4 \times 5 = 20$	α_{2_5}
age_party	Age × party interaction	Varying intercept	$4 \times 3 = 12$	α_{3_4}
age_income	Age × income interaction	Varying intercept	$4 \times 5 = 20$	α_{3_5}
party_income	Party × income interaction	Varying intercept	$3 \times 5 = 15$	α_{4_5}

voting propensities x is added as individual-level predictor, with its slope β varying over groups in S as well. This allows group-specific estimates of social media-based participation to vary conditional on voting propensities. Table F1 summarizes the variables and combinations thereof included in the different models. Note that β varies over groups in S but not on the interactions additionally to varying intercepts. See Ghitza and Gelman (2013) for how a very similar cross-classified multilevel model is built up in stages starting from classical regression and with an application to estimating turnout.

Such models can be estimated rather quickly via maximum-likelihood using for instance the lme4 software (Bates et al., 2015). However, maximum likelihood estimation for such complex models and large datasets tends to be unstable and yield convergence errors. In addition, maximum likelihood does not capture uncertainty at all levels of the model as it relies on point estimates for hyperparameters. I thus resort to a Bayesian approach. The explicit specification of prior distributions for parameters and hyperparameters incorporates all levels of uncertainty in the model and helps in stabilizing computation. In the context of non-probability samples, frequentist confidence intervals are also theoretically incompatible. Bayesian inference is in any case inherently hierarchical and lends itself naturally to multilevel modeling.

I specify the following priors for parameters and hyperparameters:

$$\begin{gathered}
\alpha_S \sim t\left(5, \mu_\alpha, (\sigma_\alpha^S)^2\right) \\
\beta_S \sim t\left(5, \mu_\beta, (\sigma_\beta^S)^2\right) \\
\mu_\alpha \sim t(5, 0, 3) \\
\mu_\beta \sim t(5, 0, 1) \\
(\sigma_\alpha^S)^2 \sim t(4, 0, 2) \\
(\sigma_\beta^S)^2 \sim t(4, 0, 2)
\end{gathered}$$

In this centered parameterization all group-level models for α_S and β_S are given $t-$distributions centered at the global intercept μ_α and slope μ_β, which are themselves given hyperpriors. This locates constant terms in several places in the model and makes it nonidentifiable. Identifiability can be recovered by redefining the parameters, however. While more complex than placing one constant in the model and setting the location of the group-level models to 0, this redundant parameterization reduces the number of iterations for convergence and computation time considerably (Gelman and Hill, 2007), especially with large datasets. The scale parameters are given group-specific hyperpriors $(\sigma_\alpha^S)^2$ and $(\sigma_\beta^S)^2$. For the hyperpriors, I also use $t-$distributions.

The $t-$distribution strikes a balance between a Gaussian distribution with a strong peak at 0 and a Cauchy distribution ($t-$distribution with one degree of freedom) with very wide tails. This makes the $t-$distribution suitable as weakly informative prior.[25] To be sure, given the amount of data, a prior must be specified

[25] See also the prior choice recommendations of the Stan developer team at https://github.com/stan-

quite concentrated to influence posterior inference. While prior evidence does exist – usually derived from survey samples smaller than 2,000 observations and absent any interactions – it does not justify the priors required to dominate the likelihood in this case with more than 90,000 observations. Accordingly and in line with the above stated motivation for a Bayesian approach, weakly informative priors here serve primarily to regularize and assist convergence. Degrees of freedom were determined based on several performance tests with subsamples of the data. The scale parameters on the hyperpriors were specified to allow reasonable values (on logit scale) but restrict values from going off-scale. With a value of 5 covering 50% of the probability scale, a scale parameter of 3 (in both directions) for μ_α puts substantial probability mass on almost the whole scale, allowing for the most extreme values. For β coefficients we would usually not expect such extreme values so that a value of 1 makes values in the bottom and top 25th percentiles less likely without ruling them out completely. Similarly, a scale parameter of two on the variance hyperpriors allows for substantial variation between subgroups.

As before with the measurement model, the different versions of the multilevel model were implement using Stan (Carpenter et al., 2017). Figure F1 shows the Stan code used to estimate the largest model. The code for the main model in the paper excludes θ and all β parameters but is exactly the same otherwise. The "transformed parameters" block shows how identifiable parameters were recovered using the reparameterization suggested by Gelman and Hill (2007). I ran four parallel Markov chains from random starting values with 2,000 iteration each. In each chain, The first 1,000 warm-up draws were discarded yielding estimates based on 4,000 posterior draws. The maximum treedepth of the Stan sampler (default is $2^{10} = 1024$ steps per iteration) was increased to $2^{12} = 4096$ to preclude the sampler from terminating prematurely. In order to detect potential non-convergence and biased inference, I checked several diagnostics: the potential scale reduction statistic Split $\hat{R}$, effective sample size, autocorrelation plots, traceplots, divergent transitions, and

dev/stan/wiki/Prior-Choice-Recommendations (last accessed July 2021).

energy plots. None indicated any pathological behavior in the chains. Detailed results of these diagnostics are available upon request.

Population-averaged predictions are obtained using full posterior estimates from the multilevel model. Rather than evaluating predictions for specific or assumed representative cases – which for multilevel models means setting to 0 or deciding on specific varying intercepts and slopes – I consider the full distribution of parameter estimates in the data. This allows inferences about the underlying population instead of arbitrary or artificial cases. The procedure follows ideas outlined in Hanmer and Kalkan (2013) and Skrondal and Rabe-Hesketh (2009) and can be summarized as follows:

1. Set up the data for which population-averaged predictions shall be obtained, i.e., fix values for variables of interest across observations while holding all other variables as observed.

2. Evaluate the prediction for each observation in the data using observation-specific (according to individuals' group membership) parameter values from one common draw of the posterior.

3. Average the generated expected values over all cases in the data and store the result in a vector.

4. Repeat steps 1 to 3 for the remaining posterior draws to include estimation uncertainty.

```
data {
  int<lower=1> N; // number of observations
  int<lower=1> n_sex; // number of sexes
  int<lower=1> n_race; // number of racial groups
  int<lower=1> n_party; // number of party affiliations
  int<lower=1> n_age; // number of age groups
  int<lower=1> n_income; // number of income groups
  int<lower=1> n_sex_race; // number of sexes by racial group
  int<lower=1> n_sex_party; // number of sexes by party affiliation
  int<lower=1> n_sex_age; // number of sexes by age group
  int<lower=1> n_sex_income; // number of sexes by income group
  int<lower=1> n_race_party; // number of racial groups by party affiliation
  int<lower=1> n_race_age; // number of racial groups by age
  int<lower=1> n_race_income; // number of racial groups by income group
  int<lower=1> n_age_party; // number of age groups by party affiliation
  int<lower=1> n_age_income; // number of age groups by income group
  int<lower=1> n_party_income; // number of party affiliations by income group
  int<lower=1, upper=n_sex> sex[N]; // observed sex
  int<lower=1, upper=n_race> race[N]; // observed racial group
  int<lower=1, upper=n_party> party[N]; // observed party affiliation
  int<lower=1, upper=n_age> age[N]; // observed age group
  int<lower=1, upper=n_income> income[N]; // observed income group
  int<lower=1, upper=n_sex_race> sex_race[N]; // observed racial group with specific sex
  int<lower=1, upper=n_sex_party> sex_party[N]; // observed party affiliation with specific sex
  int<lower=1, upper=n_sex_age> sex_age[N]; // observed age group with specific sex
  int<lower=1, upper=n_sex_income> sex_income[N]; // observed income group with specific sex
  int<lower=1, upper=n_race_party> race_party[N]; // observed party affiliation with specific race
  int<lower=1, upper=n_race_age> race_age[N]; // observed age group with specific racial race
  int<lower=1, upper=n_race_income> race_income[N]; // observed income group with specific race
  int<lower=1, upper=n_age_party> age_party[N]; // observed party affiliation with specific age
  int<lower=1, upper=n_age_income> age_income[N]; // observed income group with specific age
  int<lower=1, upper=n_party_income> party_income[N]; // observed income group with specific party
                                                      // affiliation
  real theta[N]; // observed turnout propensity
  int<lower=0,upper=1> y[N]; // observed social media-based political engagement
}

parameters {
  real mu_alpha_raw; // global intercept
  real mu_beta_raw; // global effect for theta
  vector[n_sex] alpha_sex_raw; // varying intercept for sexes
  vector[n_race] alpha_race_raw; // varying intercept for racial groups
```

Figure F1. Stan code for logistic multilevel model with varying intercepts and slopes.

```
  vector[n_party] alpha_party_raw; // varying intercept for party groups
  vector[n_age] alpha_age_raw; // varying intercept for age groups
  vector[n_income] alpha_income_raw; // varying intercept for income groups
  vector[n_sex_race] alpha_sex_race_raw; // varying intercept for sex-race interaction
  vector[n_sex_party] alpha_sex_party_raw; // varying intercept for sex-party interaction
  vector[n_sex_age] alpha_sex_age_raw; // varying intercept for sex-age interaction
  vector[n_sex_income] alpha_sex_income_raw; // varying intercept for sex-income interaction
  vector[n_race_party] alpha_race_party_raw; // varying intercept for race-party interaction
  vector[n_race_age] alpha_race_age_raw; // varying intercept for race-age interaction
  vector[n_race_income] alpha_race_income_raw; // varying intercept for race-income interaction
  vector[n_age_party] alpha_age_party_raw; // varying intercept for age-party interaction
  vector[n_age_income] alpha_age_income_raw; // varying intercept for age-income interaction
  vector[n_party_income] alpha_party_income_raw; // varying intercept for party-income interaction
  vector[n_sex] beta_sex_raw; // varying slope for 'theta' among sexes
  vector[n_race] beta_race_raw; // varying slope for 'theta' among racial groups
  vector[n_party] beta_party_raw; // varying slope for 'theta' among party affiliations
  vector[n_age] beta_age_raw; // varying slope for 'theta' among age groups
  vector[n_income] beta_income_raw; // varying slope for 'theta' among income groups
  real<lower=0> sigma_alpha_sex; // variance parameter for the prior on alpha_sex
  real<lower=0> sigma_alpha_race; // variance parameter for the prior on alpha_race
  real<lower=0> sigma_alpha_party; // variance parameter for the prior on alpha_party
  real<lower=0> sigma_alpha_age; // variance parameter for the prior on alpha_age
  real<lower=0> sigma_alpha_income; // variance parameter for the prior on alpha_income
  real<lower=0> sigma_alpha_sex_race; // variance parameter for the prior on alpha_sex_race
  real<lower=0> sigma_alpha_sex_party; // variance parameter for the prior on alpha_sex_party
  real<lower=0> sigma_alpha_sex_age; // variance parameter for the prior on alpha_sex_age
  real<lower=0> sigma_alpha_sex_income; // variance parameter for the prior on alpha_sex_income
  real<lower=0> sigma_alpha_race_party; // variance parameter for the prior on alpha_race_party
  real<lower=0> sigma_alpha_race_age; // variance parameter for the prior on alpha_race_age
  real<lower=0> sigma_alpha_race_income; // variance parameter for the prior on alpha_race_income
  real<lower=0> sigma_alpha_age_party; // variance parameter for the prior on alpha_age_party
  real<lower=0> sigma_alpha_age_income; // variance parameter for the prior on alpha_age_income
  real<lower=0> sigma_alpha_party_income; // variance parameter for the prior on alpha_party_income
  real<lower=0> sigma_beta_sex; // variance parameter for the prior on beta_sex
  real<lower=0> sigma_beta_race; // variance parameter for the prior on beta_race
  real<lower=0> sigma_beta_party; // variance parameter for the prior on beta_party
  real<lower=0> sigma_beta_age; // variance parameter for the prior on beta_age
  real<lower=0> sigma_beta_income; // variance parameter for the prior on beta_income
}
```

Figure F1 (continued)

```
transformed parameters {
  real mu_alpha;
  real mu_beta;
  vector[n_sex] alpha_sex;
  vector[n_race] alpha_race;
  vector[n_party] alpha_party;
  vector[n_age] alpha_age;
  vector[n_income] alpha_income;
  vector[n_sex_race] alpha_sex_race;
  vector[n_sex_party] alpha_sex_party;
  vector[n_sex_age] alpha_sex_age;
  vector[n_sex_income] alpha_sex_income;
  vector[n_race_party] alpha_race_party;
  vector[n_race_age] alpha_race_age;
  vector[n_race_income] alpha_race_income;
  vector[n_age_party] alpha_age_party;
  vector[n_age_income] alpha_age_income;
  vector[n_party_income] alpha_party_income;
  vector[n_sex] beta_sex;
  vector[n_race] beta_race;
  vector[n_party] beta_party;
  vector[n_age] beta_age;
  vector[n_income] beta_income;

  // reparameterization
  mu_alpha = mean(alpha_sex_raw) + mean(alpha_race_raw) + mean(alpha_party_raw) +
             mean(alpha_age_raw) + mean(alpha_income_raw) + mean(alpha_sex_race_raw) +
             mean(alpha_sex_party_raw) + mean(alpha_sex_age_raw) + mean(alpha_sex_income_raw) +
             mean(alpha_race_party_raw) + mean(alpha_race_age_raw) + mean(alpha_race_income_raw)
      +
             mean(alpha_age_party_raw) + mean(alpha_age_income_raw) + mean(
     alpha_party_income_raw);
  mu_beta = mean(beta_sex_raw) + mean(beta_race_raw) + mean(beta_party_raw) +
            mean(beta_age_raw) + mean(beta_income_raw);
  alpha_sex = alpha_sex_raw - mean(alpha_sex_raw);
  alpha_race = alpha_race_raw - mean(alpha_race_raw);
  alpha_party = alpha_party_raw - mean(alpha_party_raw);
  alpha_age = alpha_age_raw - mean(alpha_age_raw);
  alpha_income = alpha_income_raw - mean(alpha_income_raw);
  alpha_sex_race = alpha_sex_race_raw - mean(alpha_sex_race_raw);
  alpha_sex_party = alpha_sex_party_raw - mean(alpha_sex_party_raw);
  alpha_sex_age = alpha_sex_age_raw - mean(alpha_sex_age_raw);
  alpha_sex_income = alpha_sex_income_raw - mean(alpha_sex_income_raw);
  alpha_race_party = alpha_race_party_raw - mean(alpha_race_party_raw);
  alpha_race_age = alpha_race_age_raw - mean(alpha_race_age_raw);
```

Figure F1 (continued)

```
  alpha_race_income = alpha_race_income_raw - mean(alpha_race_income_raw);
  alpha_age_party = alpha_age_party_raw - mean(alpha_age_party_raw);
  alpha_age_income = alpha_age_income_raw - mean(alpha_age_income_raw);
  alpha_party_income = alpha_party_income_raw - mean(alpha_party_income_raw);
  beta_sex = beta_sex_raw - mean(beta_sex_raw);
  beta_race = beta_race_raw - mean(beta_race_raw);
  beta_party = beta_party_raw - mean(beta_party_raw);
  beta_age = beta_age_raw - mean(beta_age_raw);
  beta_income = beta_income_raw - mean(beta_income_raw);
}

model{
  vector[N] pi;

  // priors on hyperparameters
  mu_alpha_raw ~ student_t(5, 0, 3);
  mu_beta_raw ~ student_t(5, 0, 1);
  sigma_alpha_sex ~ student_t(4, 0, 2);
  sigma_alpha_race ~ student_t(4, 0, 2);
  sigma_alpha_party ~ student_t(4, 0, 2);
  sigma_alpha_age ~ student_t(4, 0, 2);
  sigma_alpha_income ~ student_t(4, 0, 2);
  sigma_alpha_sex_race ~ student_t(4, 0, 2);
  sigma_alpha_sex_party ~ student_t(4, 0, 2);
  sigma_alpha_sex_age ~ student_t(4, 0, 2);
  sigma_alpha_sex_income ~ student_t(4, 0, 2);
  sigma_alpha_race_party ~ student_t(4, 0, 2);
  sigma_alpha_race_age ~ student_t(4, 0, 2);
  sigma_alpha_race_income ~ student_t(4, 0, 2);
  sigma_alpha_age_party ~ student_t(4, 0, 2);
  sigma_alpha_age_income ~ student_t(4, 0, 2);
  sigma_alpha_party_income ~ student_t(4, 0, 2);
  sigma_beta_sex ~ student_t(4, 0, 2);
  sigma_beta_race ~ student_t(4, 0, 2);
  sigma_beta_party ~ student_t(4, 0, 2);
  sigma_beta_age ~ student_t(4, 0, 2);
  sigma_beta_income ~ student_t(4, 0, 2);

  // priors on parameters (centered parameterization)

  alpha_sex_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_sex);
  alpha_race_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_race);
  alpha_party_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_party);
  alpha_age_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_age);
  alpha_income_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_income);
```

Figure F1 (continued)

```
  alpha_sex_race_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_sex_race);
  alpha_sex_party_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_sex_party);
  alpha_sex_age_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_sex_age);
  alpha_sex_income_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_sex_income);
  alpha_race_party_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_race_party);
  alpha_race_age_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_race_age);
  alpha_race_income_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_race_income);
  alpha_age_party_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_age_party);
  alpha_age_income_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_age_income);
  alpha_party_income_raw ~ student_t(5, mu_alpha_raw, sigma_alpha_party_income);
  beta_sex_raw ~ student_t(5, mu_beta_raw, sigma_beta_sex);
  beta_race_raw ~ student_t(5, mu_beta_raw, sigma_beta_race);
  beta_party_raw ~ student_t(5, mu_beta_raw, sigma_beta_party);
  beta_age_raw ~ student_t(5, mu_beta_raw, sigma_beta_age);
  beta_income_raw ~ student_t(5, mu_beta_raw, sigma_beta_income);

  // likelihood
  for (n in 1:N)
    pi[n] = alpha_sex_raw[sex[n]] + alpha_race_raw[race[n]] +  alpha_party_raw[party[n]] +
            alpha_age_raw[age[n]] +  alpha_income_raw[income[n]] +
            alpha_sex_race_raw[sex_race[n]] + alpha_sex_party_raw[sex_party[n]] +
            alpha_sex_age_raw[sex_age[n]] + alpha_sex_income_raw[sex_income[n]] +
            alpha_race_party_raw[race_party[n]] + alpha_race_age_raw[race_age[n]] +
            alpha_race_income_raw[race_income[n]] + alpha_age_party_raw[age_party[n]] +
            alpha_age_income_raw[age_income[n]] + alpha_party_income_raw[party_income[n]] +
            beta_sex_raw[sex[n]] * theta[n] + beta_race_raw[race[n]] * theta[n] +
            beta_party_raw[party[n]] * theta[n] + beta_age_raw[age[n]] * theta[n] +
            beta_income_raw[income[n]] * theta[n];
  y ~ bernoulli_logit(pi);
}
```

Figure F1 (continued)

Appendix G: Additional Tables and Figures

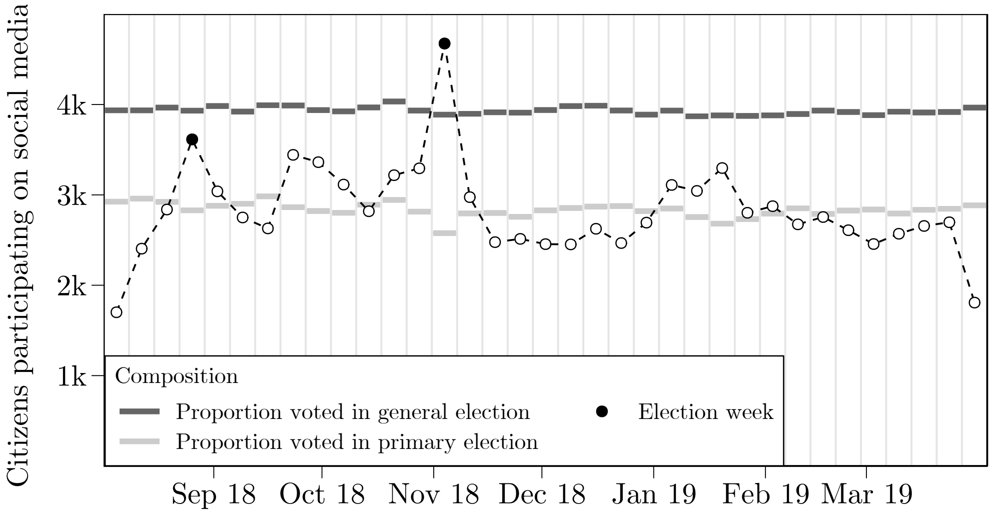


Figure G1. Social media-based participation and voter composition over time.

Note: Proportions in the background run from 0 to 1. The compensation of non-voting through online political involvement is not merely a consequence of some high-profile event that acts in place of a high-stimulus election. The number of citizens engaged on Twitter (circles) varies considerably and in response to elections. But the political voice of non-voters (area above the bars) is consistently represented

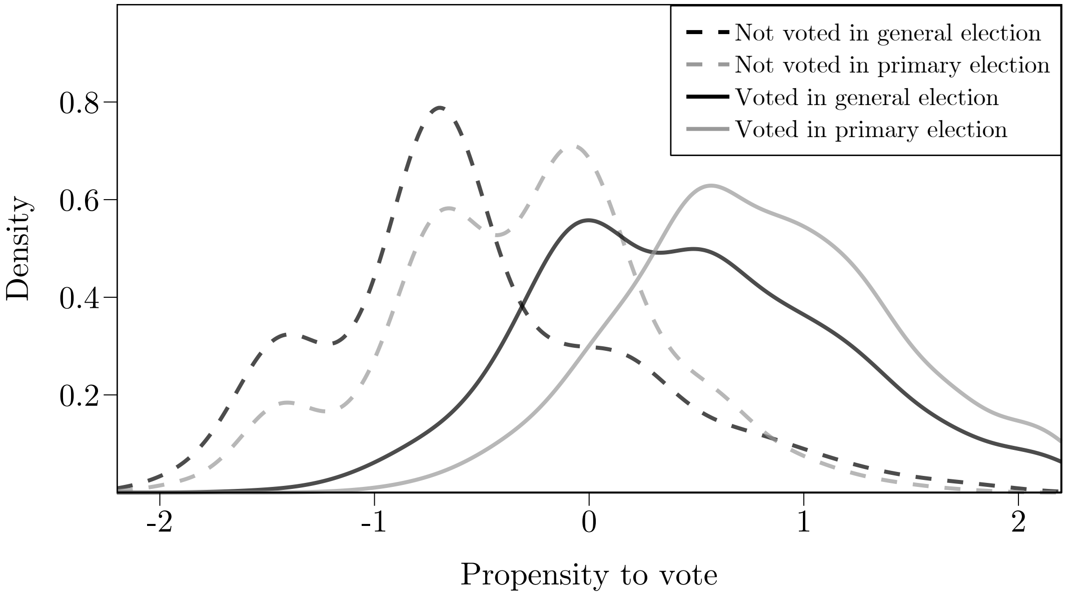


Figure G2. Voting propensity among 2018 voters and non-voters involved in social media-based participation.

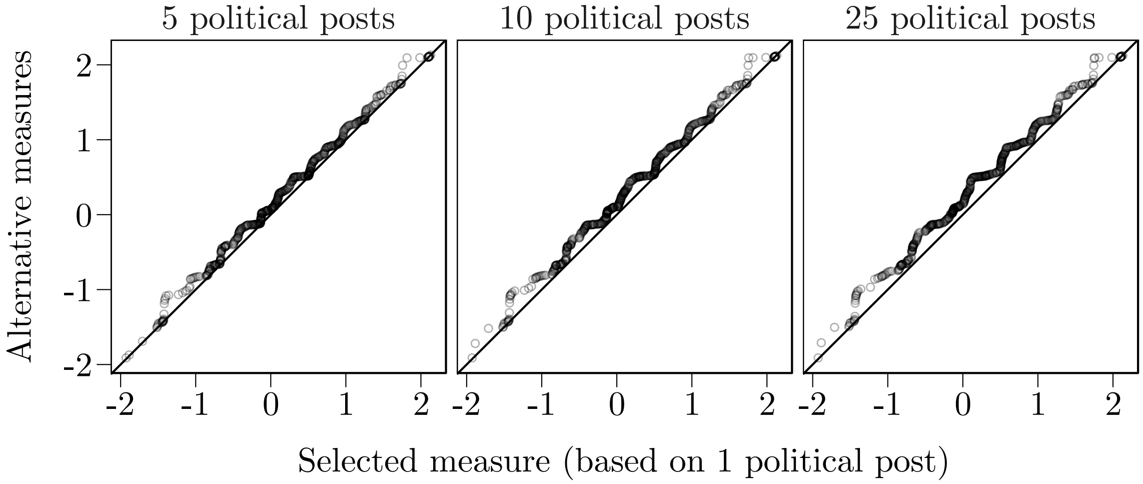


Figure G3. Quantile-quantile plot comparison of social media-based participation measures with respect to voting propensities.

Note: Adjusting measures of social media-based participation based on the amount of political posts barely affects the distribution of voting propensities among politically involved on social media. The distribution becomes slightly more left-skewed but remains similar to the selected measure based on one political post.

Table G1: Voting and social media-based participation for various sample subsets.

Sample subset	Not voted in 2018 primary election	Not voted in 2018 general election
Top 10% amount of posts	53.8%	28.2%
Top 5% amount of posts	51.2%	25.8%
Top 1% amount of posts	44.2%	23.7%
Bottom 10% amount of posts	42.3%	22.8%
At least 5 political posts	47.4%	23.1%
At least 10 political posts	43.9%	21.4%
At least 25 political posts	40.4%	20.1%
Election week	48.6%	22.5%
5 weeks prior to election week	43.9%	21.7%
5 weeks after election week	42.2%	20.4%
August 2018	46.6%	23.1%
September 2018	47.1%	22.9%
October 2018	47.6%	22.9%
November 2018	50.1%	24.1%
December 2018	48.1%	23.3%
January 2019	49.1%	24.4%
February 2019	48.4%	23.9%
March 2019	47.9%	23.8%
Registered party affiliation	47.1%	23.8%
No registered party affiliation	77.6%	35.4%
Only active voters	53.8%	26.1%
Voting propensity >= -1.5	53.5%	25.6%
Voting propensity >= -1	50.8%	22.1%
Voting propensity >= -0.5	42.7%	12.2%
Voting propensity >= 0	28.5%	10.6%
Voting propensity >= 0.5	15.0%	7.3%
Voting propensity >= 1	7.0%	6.4%
Voting propensity >= 1.5	2.0%	0.8%

Note: The voting propensity subsets indicate that to replicate prior survey evidence, the sample must be cut to include only persons with voting propensities >= 1.

Table G2: 2018 Turnout among subgroups of the Florida registered voter population

	2018 general election	2018 primary election
Male	54.9%	21.5%
Female	56.7%	22.7%
Age 18-29	34.6%	4.6%
Age 30-44	41.2%	8.3%
Age 45-64	58.5%	19.1%
Age 65+	69.4%	38.9%
Black	53.6%	20.5%
Hispanic	45.5%	12.1%
White	59.5%	25.7%
Other	48.7%	13.4%
Democrat	57.2%	24.0%
Republican	64.3%	30.5%
None	42.3%	7.8%
<=15k	44.3%	17.3%
15k-30k	51.1%	20.4%
30k-50k	62.7%	24.1%
50k-75k	64.7%	28.0%
75k+	62.2%	27.3%

Note: Based on a simple random sample (N = 100,000) of the 2018 Florida registered voter population and excluding voters marked as inactive and under voting age at the respective election. As in the paper, income is based on estimates of per capita income at small-scale census block groups.

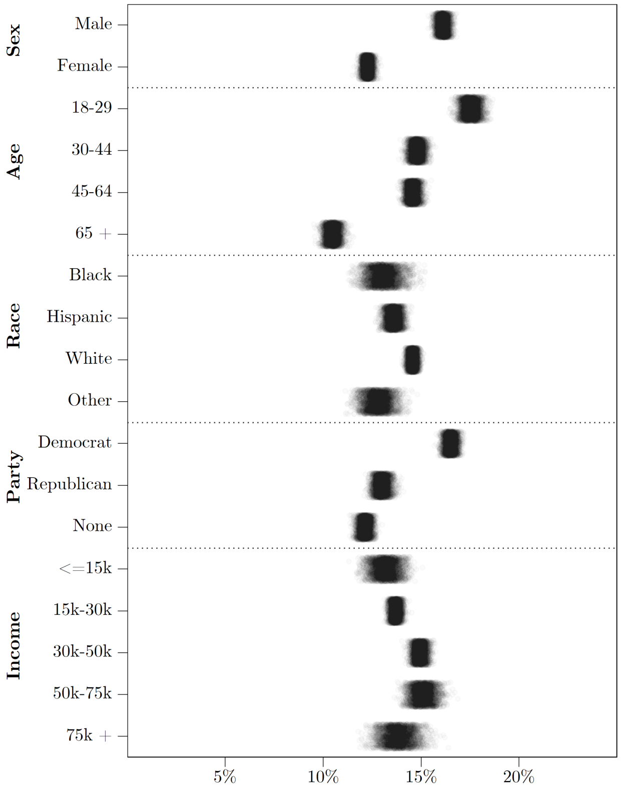


Figure G4. Social media-based participation in subgroups poststratified to a synthetic joint distribution of target population estimates.

Note: The poststratified estimate for a specific group $\theta_S = \sum_{j \in J_S} N_j \theta_j \Big/ \sum_{j \in J_S} N_j$, whereby θ_j is the estimate and N_j the known or estimated population in the interacted subgroup j in S. If estimates are generated using multilevel regression, this procedure is known as multilevel regression and poststratification or MrP. MrP requires the full joint population distribution, i.e., information about every population cell, which are 2(sex)×4(race)×4(age)×3(party)×5(income)= 480 in this case. Such detailed data is not available from census data for the target population including variables such as per capita income or party affiliation. However, Leeman and Wasserfallen (2017) show that poststratification with a simple synthetic joint distribution, constructed as the product of marginal distributions, performs as good as classical MrP. I rely on marginal distributions for sex, age, and race as estimated for the citizen-voting age population (see Appendix C). For party affiliation I rely on the marginal distribution in the registered-voter population. For per capita income, I collect information for all Florida census block groups, which I map and expand to citizen voting-age population totals in block groups. This yields the marginal distribution of block-group incomes in the citizen voting-age population. I use these marginal distributions to construct a simple synthetic joint distribution of the citizen voting-age population, to which I poststratify predictions of all 480 ideal types for the full posterior.

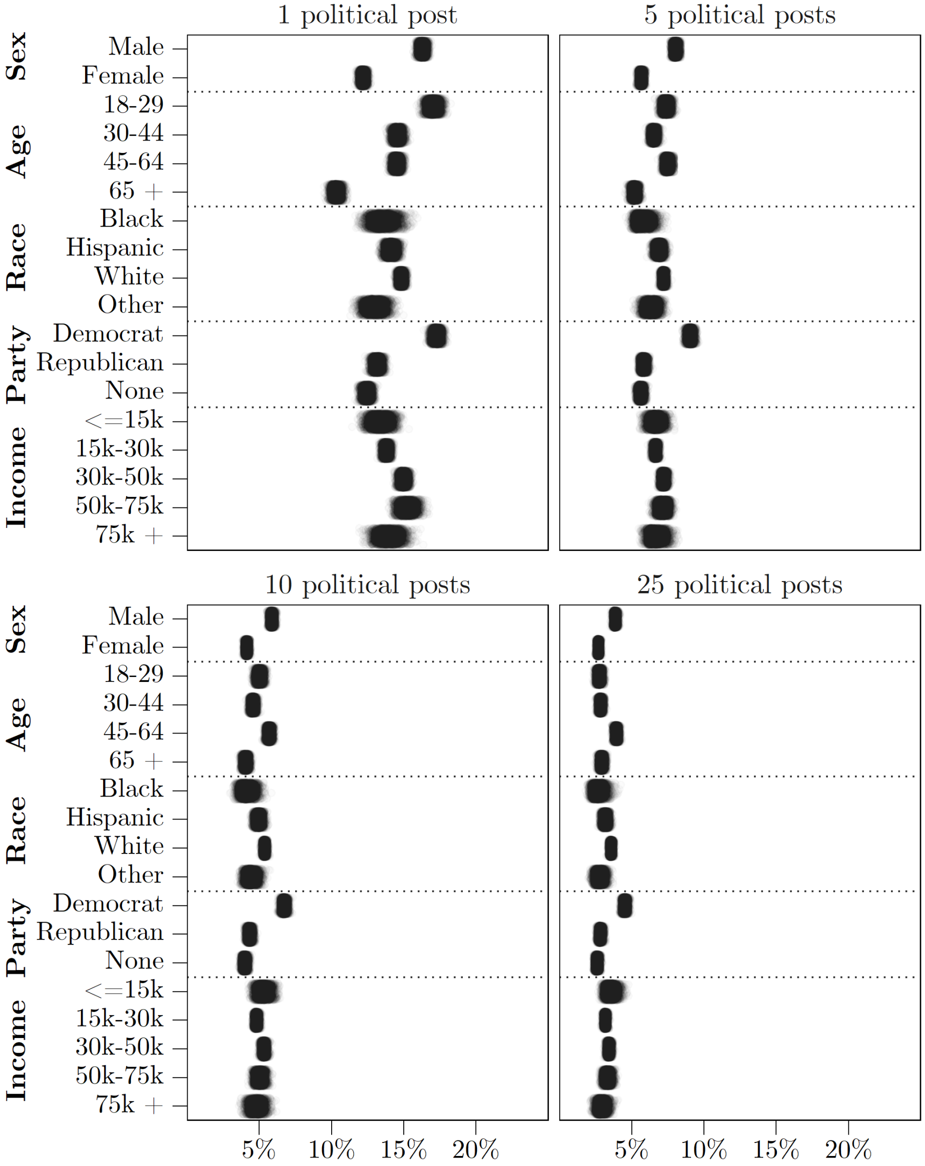


Figure G5. Social media-based participation in subgroups using alternative thresholds for measuring participation.

Note: Adjusting measures of social media-based participation based on the amount of political posts does not affect substantive conclusions regarding reinforcement theory. If anything, race and income differences become even less visible. Interestingly, young adults' disproportionately high online engagement vanishes, further attenuating compensation theorists' hopes concerning this particular group.

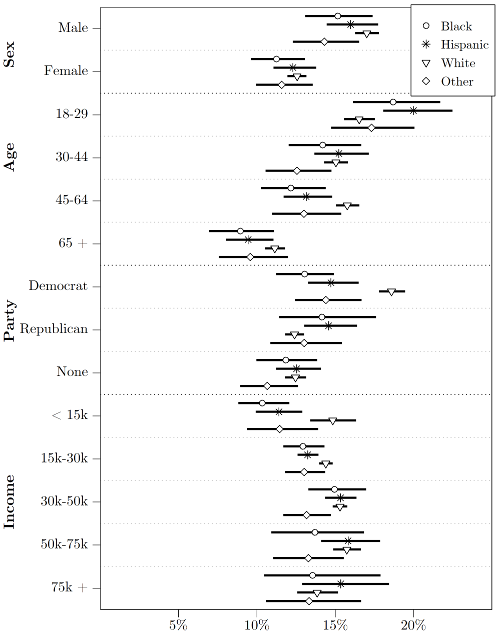


Figure G6. Social media-based participation in race-interacted subgroups.

Note: A breakdown of the whole sample in interacted subgroups reveals a notable stronger online activity on the part of poor democratic white males age 45–64, young adult minorities , and republican minorities. Yet, not only are most of these estimates surrounded by substantial uncertainty and non-significant, they also disappear for the most part after further disaggregating the sample by voter types (see Figures G16 to G31). Stronger engagement remains robust only among males and (mostly white) democrats.

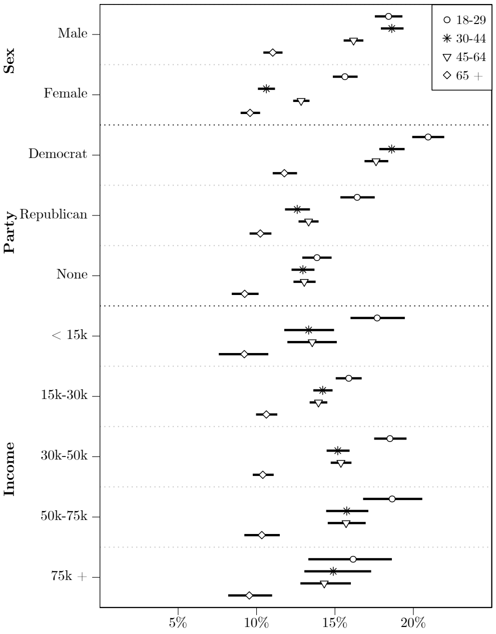


Figure G7. Social media-based participation in remaining age-interacted subgroups.

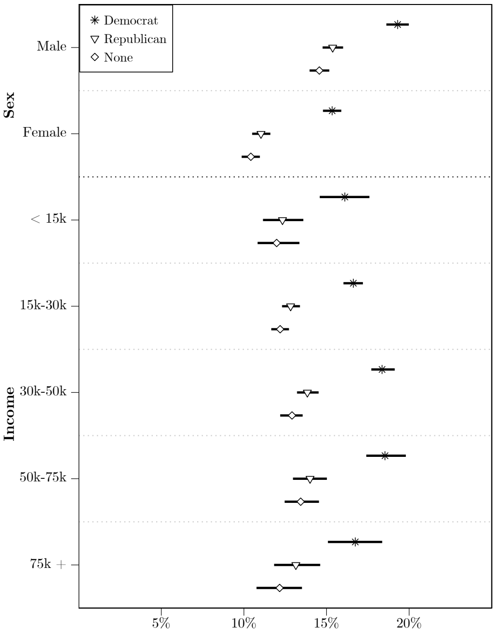


Figure G8. Social media-based participation in remaining party-interacted subgroups.

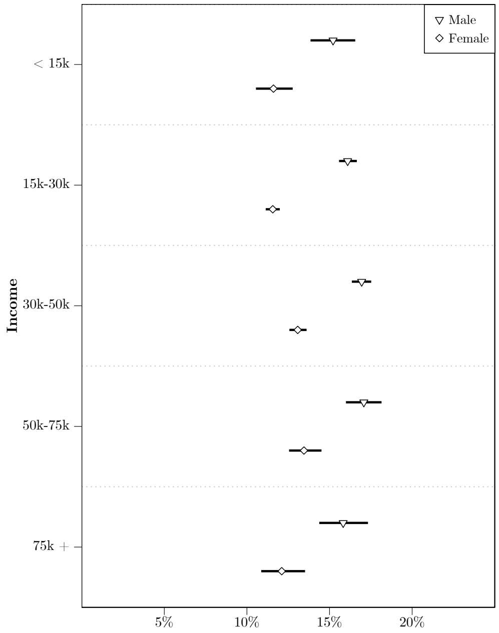


Figure G9. Social media-based participation in remaining sex-interacted subgroups.

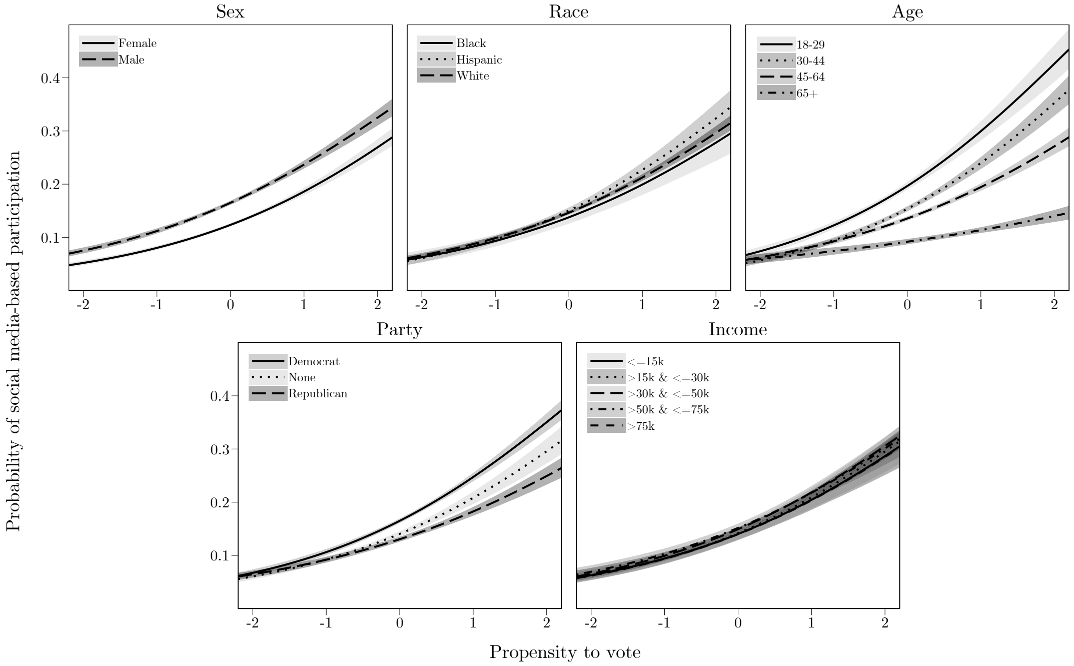


Figure G10. Social media-based participation in subgroups including voting propensity as individual-level predictor with varying slope.

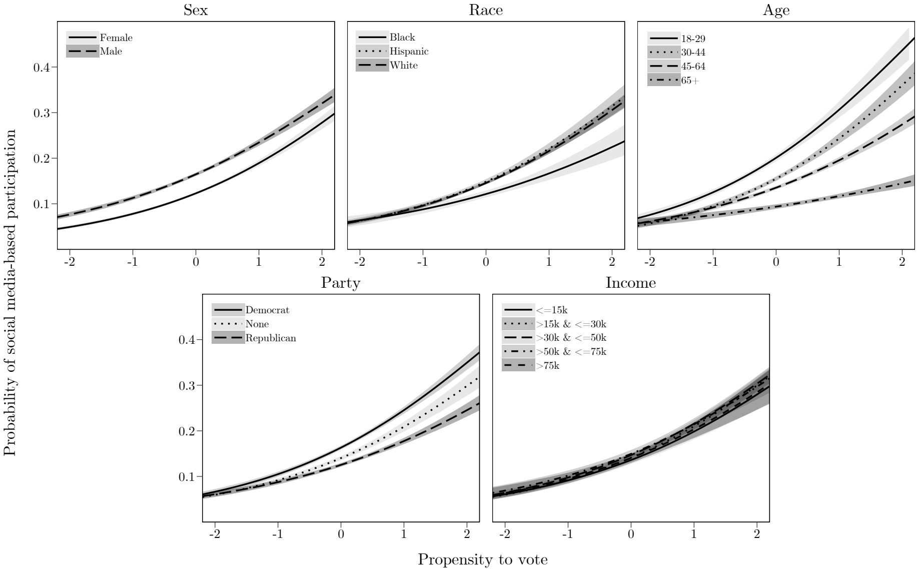


Figure G11. Social media-based participation in subgroups including voting propensity as individual-level predictor with varying slope and inappropriately excluding two-way interactions.

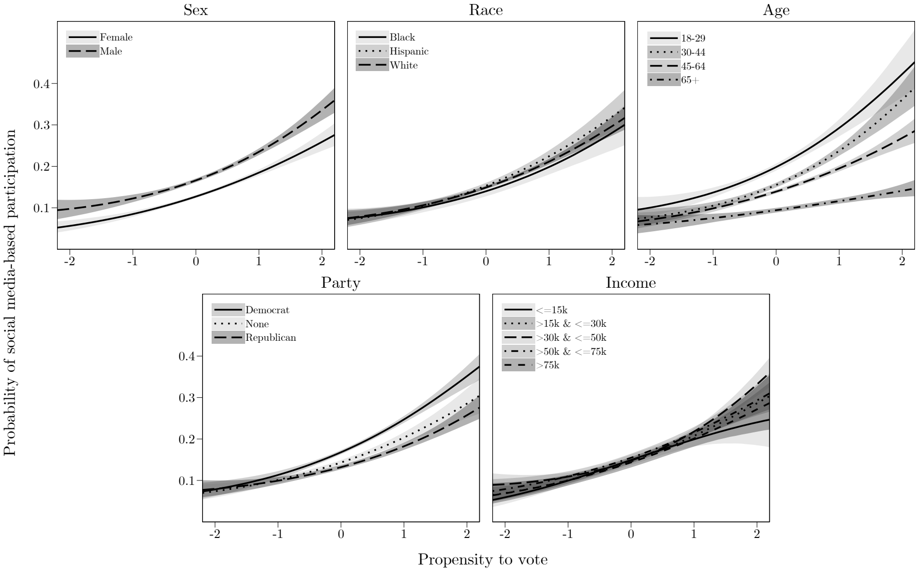


Figure G12. Social media-based participation in subgroups including voting propensity as individual-level predictor with varying slope and with additional cubic θ.

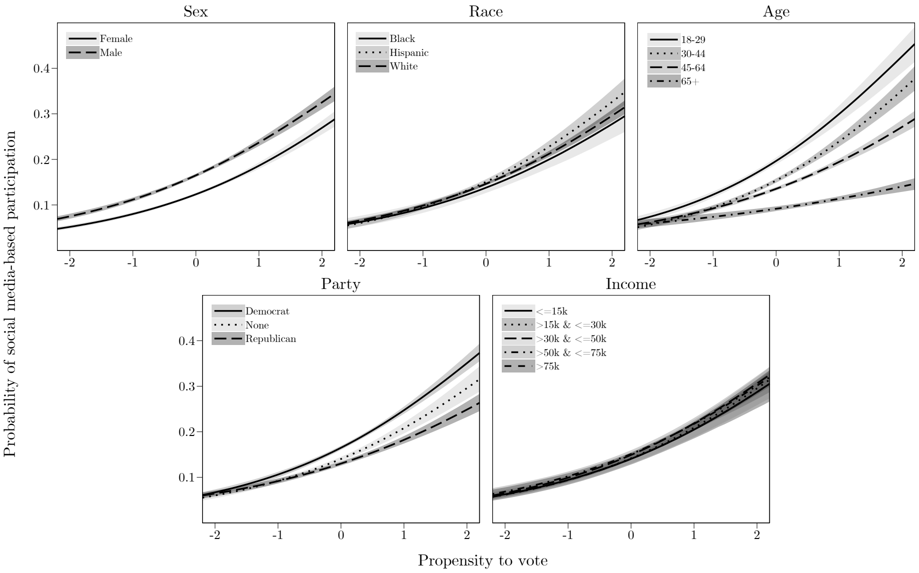


Figure G13. Social media-based participation in subgroups including voting propensity as individual-level predictor with varying slope and with less regularizing priors, $\mu_\alpha \sim t(5,0,5)$, $\mu_\beta \sim t(5,0,2.5)$, $(\sigma^S_\alpha)^2 \sim t(4,0,5))$, $(\sigma^S_\beta)^2 \sim t(4,0,5)$.

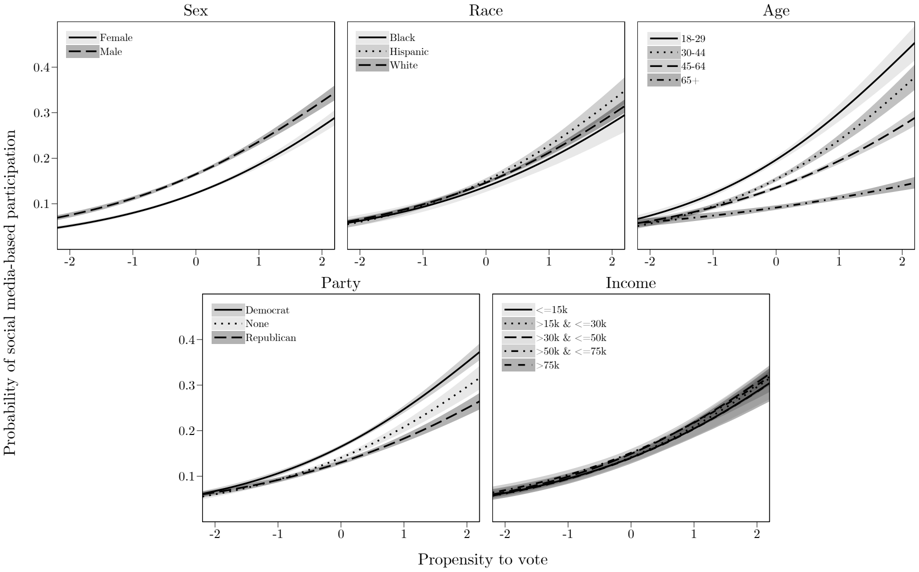


Figure G14. Social media-based participation in subgroups including voting propensity as individual-level predictor with varying slope and with very vague priors, $\mu_\alpha \sim t(5,0,100)$, $\mu_\beta \sim t(5,0,100)$, $(\sigma^S_\alpha)^2 \sim t(4,0,100))$, $(\sigma^S_\beta)^2 \sim t(4,0,100)$.

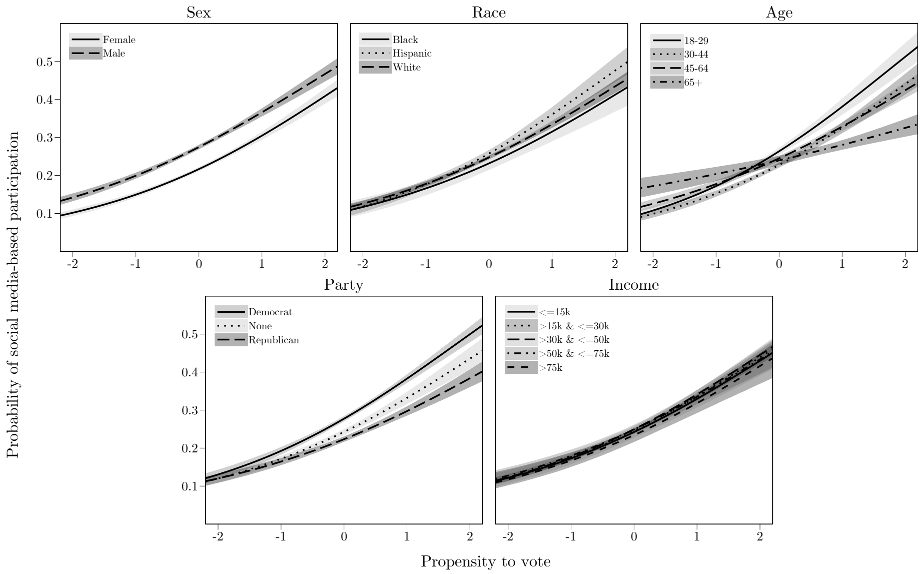


Figure G15. Social media-based participation in subgroups including voting propensity as individual-level predictor with varying slope and excluding inactive Twitter users.

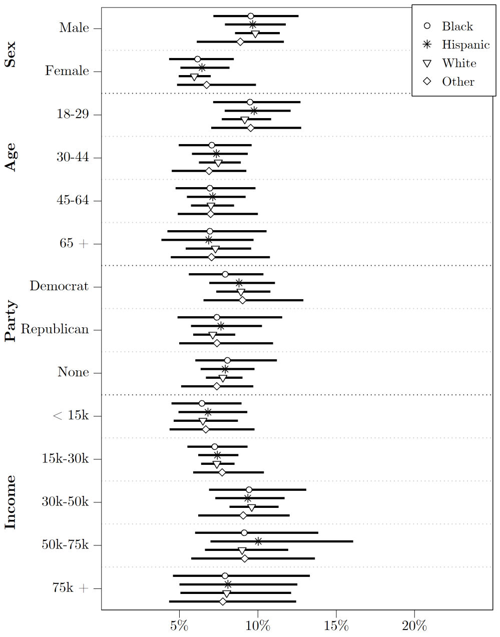


Figure G16. Social media-based participation of low-propensity voters in race-interacted subgroups.

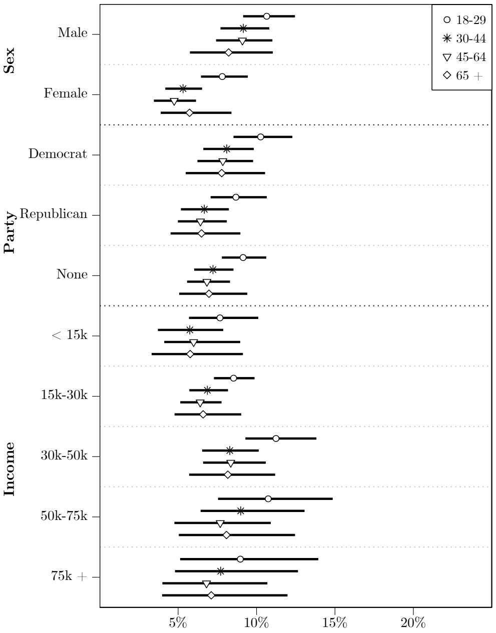


Figure G17. Social media-based participation of low-propensity voters in remaining age-interacted subgroups.

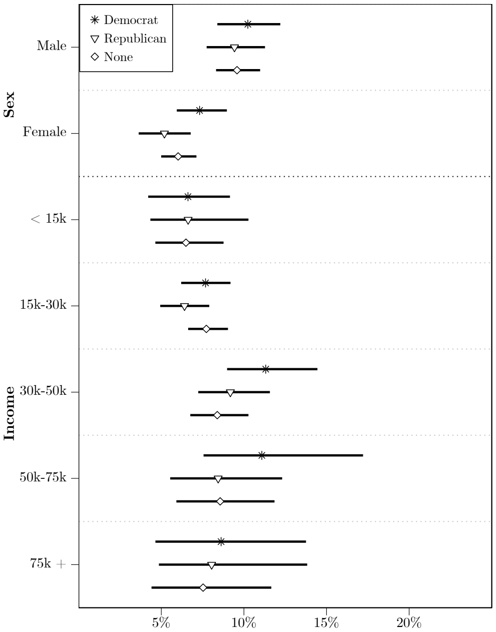


Figure G18. Social media-based participation of low-propensity voters in remaining party-interacted subgroups.

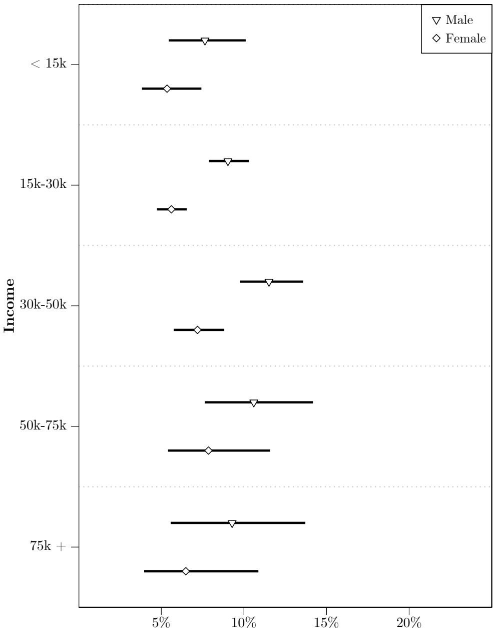


Figure G19. Social media-based participation of low-propensity voters in remaining sex-interacted subgroups.

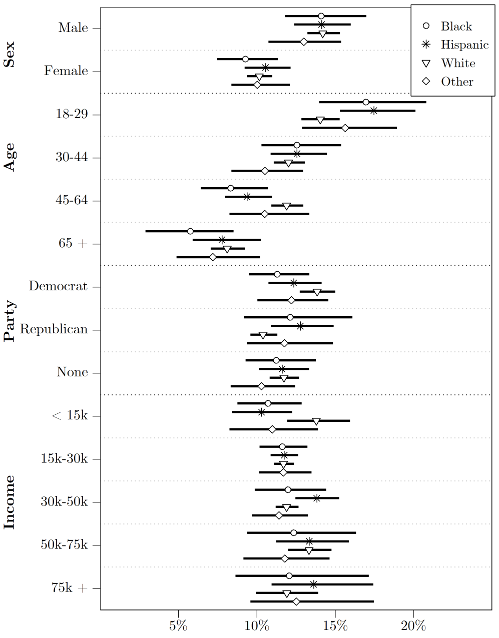


Figure G20. Social media-based participation of irregular voters in race-interacted subgroups.

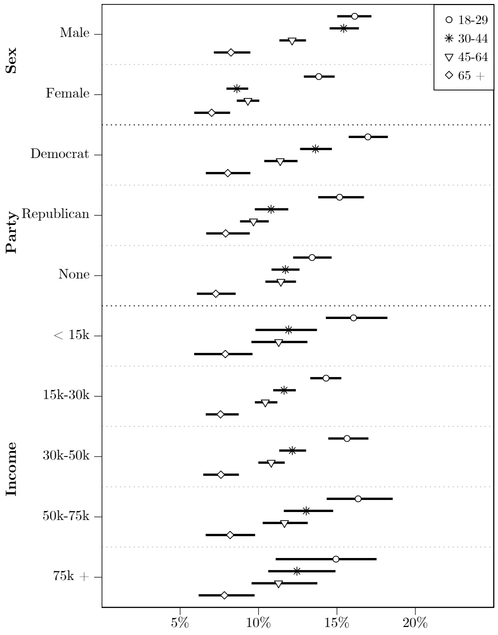


Figure G21. Social media-based participation of irregular voters in remaining age-interacted subgroups.

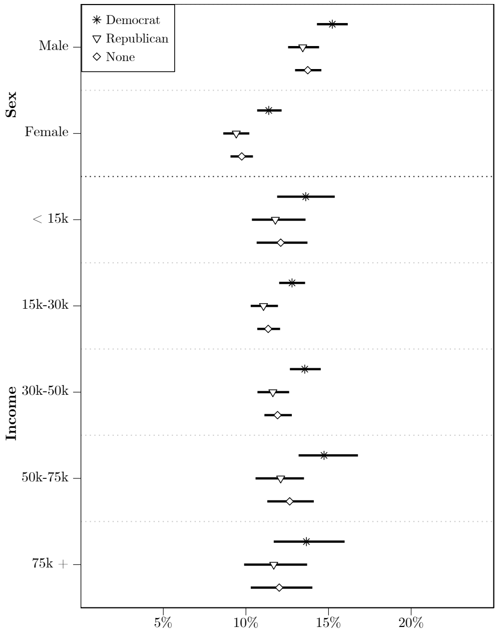


Figure G22. Social media-based participation of irregular voters in remaining party-interacted subgroups.

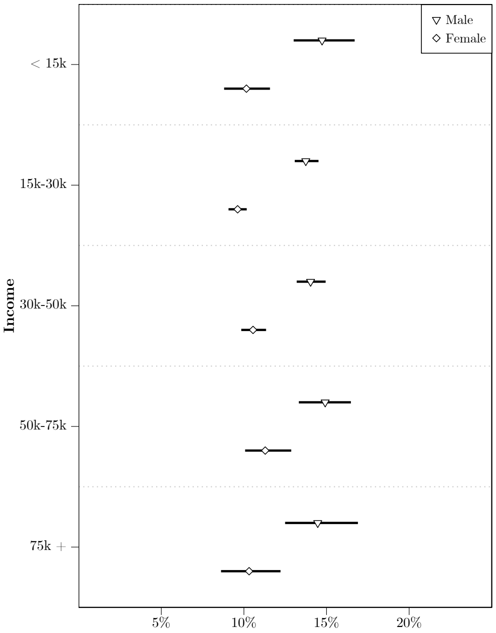


Figure G23. Social media-based participation of irregular voters in remaining sex-interacted subgroups.

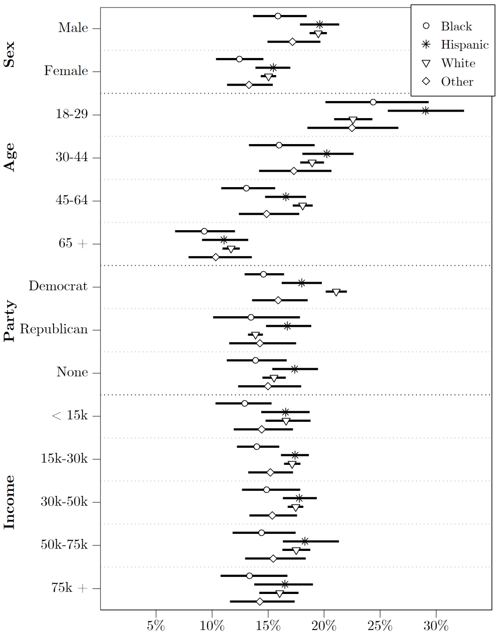


Figure G24. Social media-based participation of regular voters in race-interacted subgroups.

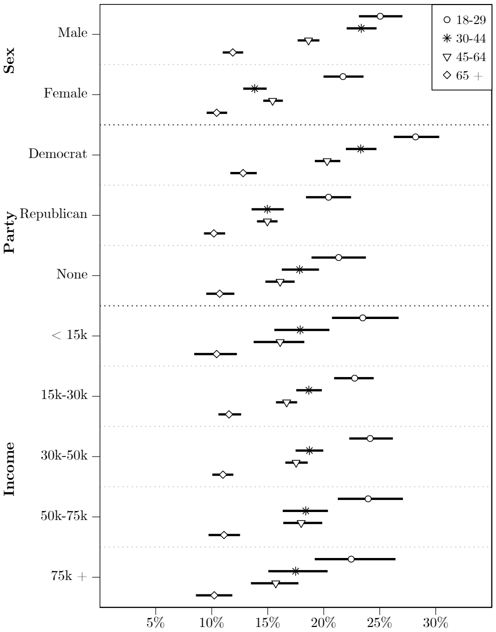


Figure G25. Social media-based participation of regular voters in remaining age-interacted subgroups.

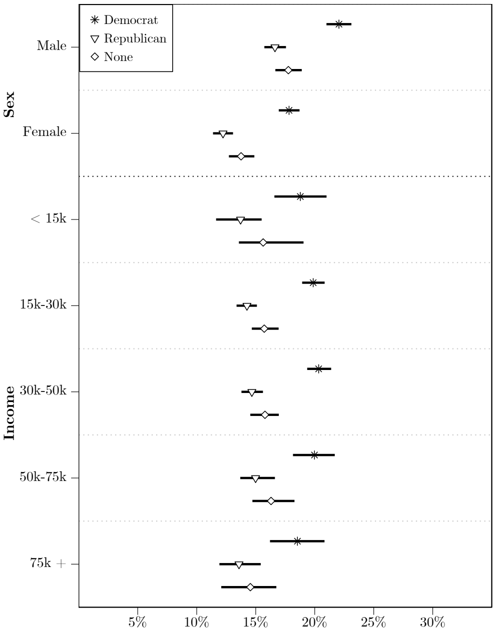


Figure G26. Social media-based participation of regular voters in remaining party-interacted subgroups.

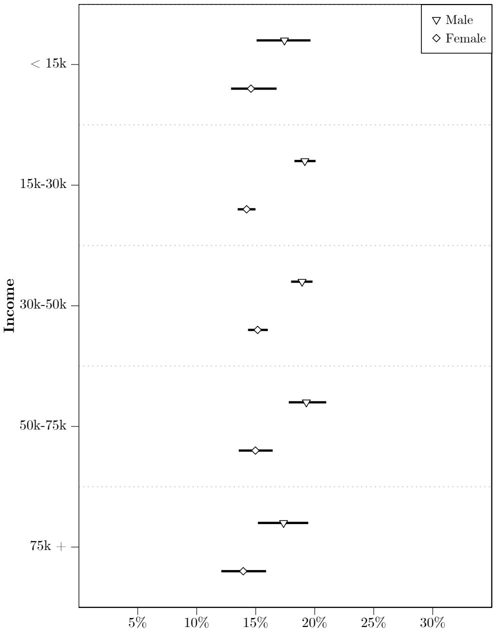


Figure G27. Social media-based participation of regular voters in remaining sex-interacted subgroups.

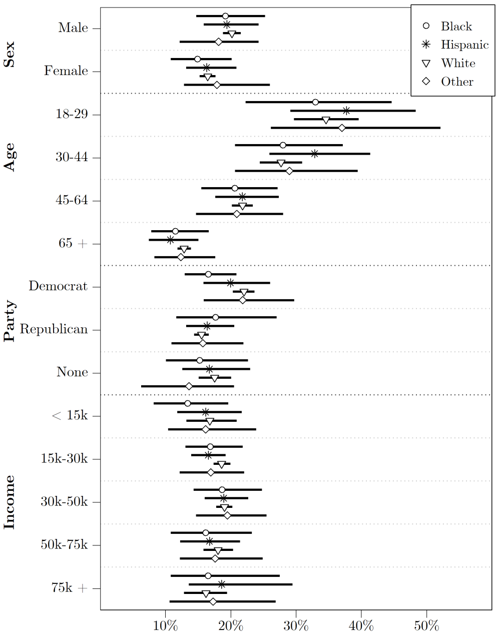


Figure G28. Social media-based participation of highly engaged voters in race-interacted subgroups.

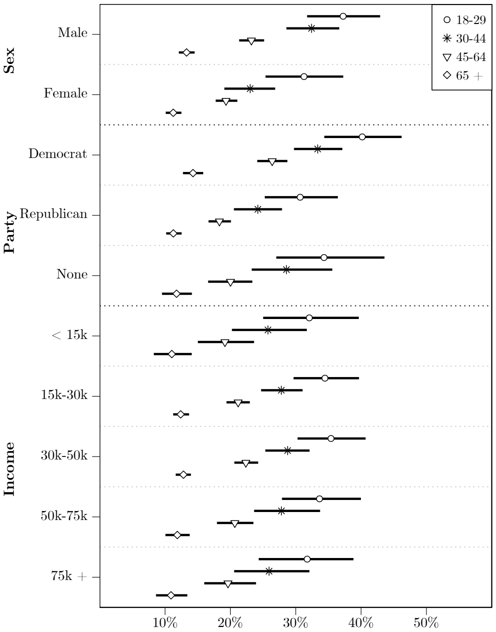


Figure G29. Social media-based participation of highly engaged voters in remaining age-interacted subgroups.

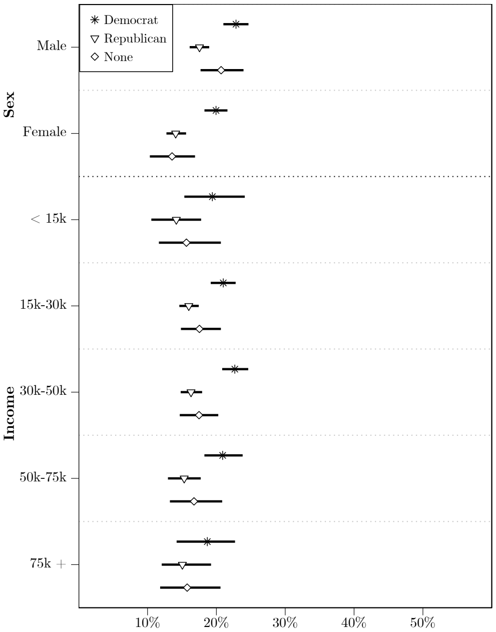


Figure G30. Social media-based participation of highly engaged voters in remaining party-interacted subgroups.

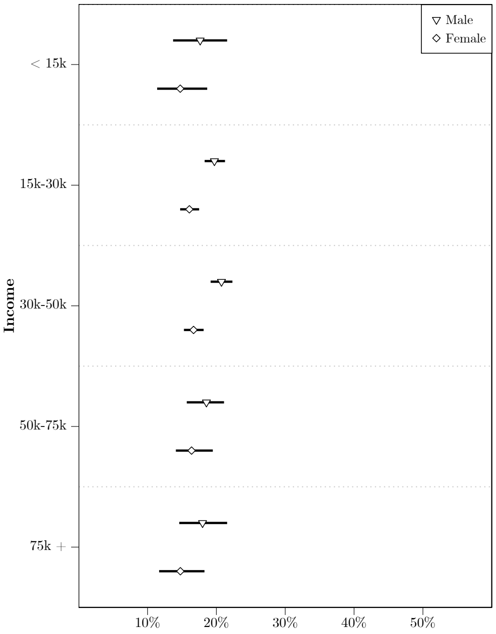


Figure G31. Social media-based participation of highly engaged voters in remaining sex-interacted subgroups.

Appendix H: Software Statement; Data and Code Availability

All code written for data collection, data processing, analyses, and graphics was run under Windows 10 x86-64 using R version 3.4.2. Table H1 lists R packages that were used. Code is divided into several scripts. These scripts are available at https://github.com/saschagobel/jqd-voting-and-sm-pol-participation together with anonymized replication data.

Appendix I: Ethical and Legal Considerations of Data

As of yet, there exists no common standard for the treatment of publicly available and passively observed data, such as it is available, for instance, from Twitter or official voter records (Steinert-Threkeld, 2018). However, past research practice offers some guidance for protecting such data against unanticipated secondary use and for compliance and transparency as regards the law (Salganik, 2018).

Table H1: R packages.

abind (Plate and Heiberger, 2016)
arm (Gelman et al., 2016)
bayesplot (Gabry et al., 2019)
censusapi (Recht, 2019)
censusr (Macfarlane and Kressner, 2018)
cld2 (Ooms and Sites, 2018)
cowplot (Wilke, 2020)
crayon (Csárdi and Gaslam, 2017)
data.table (Dowle et al., 2017)
DBI (R SIG on Databases et al., 2017)
dbplyr (Wickham et al., 2019b)
doParallel (Calaway et al., 2018)
dplyr (Wickham et al., 2019b)
eeptools (Becker and Knowles, 2019)
eulerr (Larsson et al., 2019)
extrafont (Chang, 2014)
foreach (Calaway et al., 2017)
gender (Mullen et al., 2018a)
genderizeR (Wais et al., 2019)
ggmap (Kahle et al., 2019)
ggplot2 (Wickham, 2009)
ggpubr (Kassambara, 2019)
gridExtra (Auguie and Antonov, 2017)
httr (Wickham, 2019)
keras (Falbel et al., 2019)
lubridate (Grolemund and Wickham, 2011)
magrittr (Bache and Wickham, 2014)
maptools (Bivand et al., 2019)
pacman (Rinker et al., 2017)
preText (Denny and Spirling, 2018)
pryr (Wickham and R Core team, 2018)
psych (Revelle, 2019)
PUMSutils (Thaler, 2019)
purrr (Henry et al., 2019a)
quanteda (Benoit et al., 2018)
R.utils (Bengtsson, 2019)
remoji (FitzJohn, 2015)
reticulate (Ushey et al., 2019)
rgeos (Bivand et al., 019b)
rlang (Henry et al., 2019b)
rlist (Ren, 2016)
ROAuth (Gentry and Lang, 2015)
RSelenium (Harrison and Kim, 2019)
RSQLite (Müller et al., 2019)
rstan (Guo et al., 2019)
rtweet (Kearney, 2020)
rvest (Wickham, 2016)
stopwords (Benoit et al., 2019)
stringi (Gagolewski et al., 2019)
stringr (Wickham, 2017)
tidycensus (Walker et al., 2019)
tidyr (Wickham et al., 2019a)
tm (Feinerer et al., 2018)
USAboundaries (Mullen et al., 2018b)
wru (Imai and Khanna, 2016)
XML (Lang and CRAN Team, 2019)

To protect the privacy of people included in this study and to guard against unanticipated malicious use of the data by others, identifying information and Twitter texts by individuals are not outright sent to or shared with others. Replication data is only offered in anonymized fashion, i.e., decoupled from data that would allow future re-identification, and exclude any Twitter texts, or shared conditional on a non-disclosure agreement and without identifying information. Twitter and voter record data are stored separately in a secure location and can be linked through a key that is not forwarded to others. Twitter texts presented in the paper or appendix was paraphrased to prevent re-identification.

As regards legal aspects, all data used in this research project is in the public domain and was not subjected to any kind of intervention or manipulation. Several published research papers have relied on same or comparable data and data collection approaches (Barberá, 2014; Grinberg et al., 2019; White, 2019). Per Florida Statute 97.0585, all information in voter registration lists is public record, including but not limited to sensitive information such as a person's name, address, date of birth, party affiliation, phone number, and email address.[26]

Per Twitter's Privacy Policy section 1.2, profile information, Tweets, Retweets, liked Tweets, Replies, and the respective creation date of these activities, as well as Followers and followed accounts are public information made publicly accessible via API's.[27] This does not apply to protected accounts, for which these data are not accessible and which are excluded here. According to Twitter Privacy Policy section 1.3, Twitter users who do not adjust their privacy settings accordingly (opt-out) allow others to find them via the email address they provided for account creation.[28]

In line with Twitter's Developer Policy (see Chapter 2, Section "Off-Twitter matching"), associating Twitter accounts and their public content with persons in external records is permitted, if this association is made based on publicly available data.[29] Here, such an association is made based on information from publicly available voter registration lists.

[26] See http://www.leg.state.fl.us/statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=0000-0099/0097/Sections/0097.0585.html and https://dos.myflorida.com/elections/for-voters/voter-registration/voter-information-as-a-public-record/ (last accessed July 2021).

[27] See https://twitter.com/en/privacy (last accessed July 2021).

[28] See https://twitter.com/en/privacy and https://help.twitter.com/en/safety-and-security/email-and-phone-discoverability-settings (last accessed July 2021).

[29] See https://developer.twitter.com/en/developer-terms/policy (last accessed July 2021).

References

Ansolabehere, S. and Hersh, E. (2012). Validation. What big data reveal about survey misreporting and the real electorate. *Political Analysis*, 20(4):437–459.

Auguie, B. and Antonov, A. (2017). *gridExtra. Miscellaneous functions for "grid" graphics.* R package version 0.2.3.

Bache, S. M. and Wickham, H. (2014). *magrittr. A forward-pipe operator for R.* R package version 1.5.

Barberá, P. (2014). Birds of the same feather tweet together. Bayesian ideal point estimation using Twitter data. *Political Analysis*, 23(1):76–91.

Bates, Douglas Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Becker, J. P. and Knowles, J. E. (2019). *eeptools. Convenience functions for education data.* R package version 1.2.2.

Bengtsson, H. (2019). *R.utils. Various programming utilities.* R package version 2.9.0.

Benoit, K., Muhr, D., and Watanabe, K. (2019). *stopwords. Multilingual stopword lists.* R package version 1.0.

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., and Matsuo, A. (2018). quanteda. An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30).

Bivand, R., Lewin-Koh, N., Pebesma, E., Archer, E., Baddeley, A., Bearman, N., Bibiko, H.-J., Brey, S., Callahan, J., Carrillo, G., Dray, S., Forrest, D., Friendly, M., Giraudoux, P., Golicher, D., Rubio, V. G., Hausmann, P., Hufthammer, K. O., Jagger, T., Johnson, K., Luque, S., MacQueen, D., Niccolai, A., Lamigueiro, O. P., Plunkett, E., Short, T., Snow, G., Stabler, B., Stokely, M., and Turner, R. (2019). *maptools. Tools for handling spatial objects.* R package version 0.9-5.

Bivand, R., Rundel, C., Pebesma, E., Stuetz, R., Hufthammer, K. O., Giraudoux, P., Davis, M., and Santilli, S. (2019b). *rgeos. Interface to geometry engine - Open Source ('GEOS')*. R package version 0.5-1.

Calaway, R., Microsoft, and Weston, S. (2017). *foreach. Provides foreach looping construct for R.* R package version 1.4.4.

Calaway, R., Microsoft, Weston, S., and Tenenbaum, D. (2018). *doParallel. Foreach parallel adaptor for the 'parallel' package.* R package version 1.0.14.

Campbell, A. (1960). Surge and decline. A study of electoral change. *Public Opinion Quarterly*, 24(3):397–418.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan. A probabilistic programming language. *Journal of Statistical Software*, 76(1).

Chang, W. (2014). *extrafont. Tools for using fonts.* R package version 0.17.

Chollet, F. and Allaire, J. J. (2018). *Deep learning with R.* Manning, New York.

Clinton, J., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2):355–370.

Csárdi, G. and Gaslam, B. (2017). *crayon. Colored terminal output.* R package version 1.3.4.

Denny, M. and Spirling, A. (2018). Text processing for unsupervised learning. Why it matters, when it misleads, and what to do about it. *Political Analysis*, 26(2):168–189.

Dowle, M., Srinivasan, A., Gorecki, J., Short, T., Lianoglou, S., and Antonyan, E. (2017). *data.table. Extension of 'data.frame'.* R package version 1.10.4.

Falbel, D., Allaire, J., Chollet, F., RStudio, Google, Tang, Y., Van Der Bijl, W., Studer, M., and Keydana, S. (2019). *keras. R interface to 'Keras'.* R package version 2.2.4.1.

Feinerer, I., Hornik, K., and Artifex Software, Inc. (2018). *tm. Text mining package.* R package version 0.7-6.

FitzJohn, R. G. (2015). *remoji. Fetch and search emoji.* R package version 0.1.0.

Florida Department of Corrections (2018). *Annual Report 2017–2018.* Tallahassee, FL.

Fowler, J. H., Baker, L., and Dawes, C. T. (2008). Genetic variation in political participation. *American Political Science Review*, 102(2):233–248.

Fraga, B. L. (2018). *The turnout gap. Race, ethnicity, and political inequality in a diversifying America.* Cambridge University Press, Cambridge.

Gabry, J., Mahr, T., Bürkner, P.-C., Modrák, M., and Barrett, M. (2019). *bayesplot. Plotting for bayesian models.* R package version 1.7.0.

Gagolewski, M., Tartanus, B., IBM, and Unicode, Inc. (2019). *stringi. Character string processing facilities.* R package version 1.4.3.

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

Gelman, A., Su, Y.-S., Yajima, M., Hill, J., Pittau, M. G., Kerman, J., Zheng, T., and Dorie, V. (2016). *arm. Data analysis using regression and multilevel/hierarchical models.* R package version 1.10-1.

Gentry, J. and Lang, D. T. (2015). *ROAuth. R interface for OAuth.* R package version 0.9.6.

Ghitza, Y. and Gelman, A. (2013). Deep interactions with MRP. Election turnout and voting patterns among small electoral subgroups. *American Journal of Political Science*, 57(3):762–776.

Grimmer, J. and Stewart, B. M. (2013). Text as data. The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3):267–297.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425):374–378.

Grolemund, G. and Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3).

Guo, J., Gabry, J., Goodrich, B., Lee, D., Sakrejda, K., Martin, M., Trustees of Columbia University, Sklyar, O., The R Core Team, and Oehlschlaegel-Akiyoshi, J. (2019). *rstan. R interface to Stan.* R package version 2.19.2.

Hanmer, M. J. and Kalkan, K. O. (2013). Behind the curve. Clarifying the best approach to calculating predicted probabilities and marginal effects from limited dependent variable models. *American Journal of Political Science*, 57(1):263–277.

Harrison, J. and Kim, J. Y. (2019). *RSelenium. R bindings for 'Selenium Web Driver'.* R package version 1.7.5.

Henry, L., Wickham, H., and RStudio (2019a). *purrr. Functional programming tools.* R package version 0.3.2.

Henry, L., Wickham, H., and RStudio (2019b). *rlang. Functions for base types and core R and 'tidyverse' features.* R package version 0.4.0.

Hersh, E. D. and Nall, C. (2016). The primacy of race in the geography of income-based voting. New evidence from public voting records. *American Journal of Political Science*, 60(2):289–303.

Imai, K. and Khanna, K. (2016). Improving ecological inference by predicting individual ethnicity from voter registration records. *Political Analysis*, 24(2):263–272.

Jones, J. J., Bond, R. M., Fariss, C. J., Settle, J. E., Kramer, A. D. I., Marlow, C., and Fowler, J. H. (2013). Yahtzee. An anonymized group level matching procedure. *PLoS ONE*, 8(2).

Kahle, D., Wickham, H., Jackson, S., and Korpela, M. (2019). *ggmap. Spatial visualization with ggplot2.* R package version 3.0.0.

Kassambara, A. (2019). *ggpubr. 'ggplot2' based publication ready plots.* R package version 0.2.3.

Kearney, M. W. (2020). *rtweet. Collecting Twitter data.* R package version 0.7.0.

King, G., Lam, P., and Roberts, M. R. (2017). Computer-assisted keyword and document set discovery from unstructured text. *American Journal of Political Science*, 61(4):971–988.

Klasnja, M., Barberá, P., Beauchamp, N., Nagler, J., and Tucker, J. (2017). Measuring public opinion with social media data. In Atkeson, L. R. and Alvarez, M., editors, *Oxford handbook of polling and polling methods*. Oxford University Press, Oxford.

Lang, D. T. and CRAN Team (2019). *XML. Tools for parsing and generating XML within R and S-Plus.* R package version 3.98-1.20.

Larsson, J., Godfrey, A. J. R., Gustafsson, P., Eberly, D. H., and Huber, E. (2019). *eulerr. Area-proportional Euler and Venn diagrams with ellipses.* R package version 5.1.0.

Leeman, L. and Wasserfallen, F. (2017). Extending the use and prediction precision of subnational public opinion estimation. *American Journal of Political Science*, 61(4):1003–1022.

Levy, R. and Mislevy, R. J. (2016). *Bayesian psychometric modeling.* CRC Press, Boca Raton, FL.

Macfarlane, G. and Kressner, J. (2018). *censusr. Collect data from the census API.* R package version 0.0.4.

Margetts, H. (2017). Political behaviour and the acoustics of social media. *Nature Human Behaviour*, 1(4).

McDonald, M. P. (2017). *Voter Turnout.* United States Elections Project. Accessed 21 October 2018.

Microsoft (2018). *Bing maps locations API.*

Muddiman, A., McGregor, S. C., and Stroud, N. J. (2019). (re)claiming our expertise. Parsing large text corpora with manually validated and organic dictionaries. *Political Communication*, 36(2):214–226.

Mullen, L., Blevins, C., and Schmidt, B. (2018a). *gender. Predict gender from names using historical data.* R package version 0.5.2.

Mullen, L., Bratt, J., and US Census Bureau (2018b). *USAboundaries. Historical and contemporary boundaries of the United States of America.* R package version 0.3.1.

Müller, K., Wickham, H., James, D. A., Falcon, S., SQLite Authors, Healy, L., R Consortium, and RStudio (2019). *RSQLite. 'SQLite' interface for R.* R package version 2.1.2.

National Association of Secretaries of State (2017). Maintenance of state voter registration lists. A review of relevant policies and procedures.

Ooms, J. and Sites, D. (2018). *cld2. Google's compact language detector 2.* R package version 1.2.

Pettigrew, S. and Stewart, C. I. (2018). Moved out, moved on. Assessing the effectiveness of voter registration list maintenance. *Massachusetts Institute of Technology Political Science Department Research Paper No. 2018-1.*

Plate, T. and Heiberger, R. (2016). *abind. Combine multidimensional arrays.* R package version 1.4.5.

R SIG on Databases, Wickham, H., Müller, K., and R Consortium (2017). *DBI. R database interface.* R package version 1.0.0.

Recht, H. (2019). *censusapi. Retrieve data from the census APIs.* R package version 0.6.0.

Ren, K. (2016). *rlist. A toolbox for non-tabular data manipulation.* R package version 0.4.6.1.

Revelle, W. (2019). *psych. Procedures for psychological, psychometric, and personality research.* R package version 1.8.12.

Rinker, T., Kurkiewicz, D., Hughitt, K., Wang, A., and Hester, J. (2017). *pacman. Package management tool.* R package version 0.4.6.

Ruths, D. and Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213):1063–1064.

Sakshaug, J. W. (2018). Methods of linking survey data to official records. In Vannette, D. L. and Krosnick, J. A., editors, *Palgrave handbook of survey research*, pages 257–261. Palgrave Macmillan, London.

Salganik, M. J. (2018). *Bit by Bit. Social research in the digital age.* Princeton University Press, New York.

Settle, J. E., Bond, R. M., Coviello, L., Fariss, C. J., Fowler, J. H., and Jones, J. J. (2016). From posting to voting. The effects of political competition on online engagement. *Political Science Research and Methods*, 4(2):361–378.

Shannon, S. K. S., Uggen, C., Schnittker, J., Thompson, M., Wakefield, S., and Massoglia, M. (2017). The growth, scope, and spatial distribution of people with felony records in the United States, 1948–2010. *Demography*, 54(5):1795–1818.

Skrondal, A. and Rabe-Hesketh, S. (2009). Prediction in multilevel generalized linear models. *Journal of the Royal Statistical Society*, 172(3):659–687.

Sloan, L. and Jeffrey, M. (2015). Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. *PLoS ONE*, 10(11).

Spahn, B. and Hindman, M. (2014). Eccentric circles and non-habitual voting. Turnout patterns in a 2-million-voter national panel dataset. *Paper presented at the 2014 Midwest Political Science Association Meeting.*

Steinert-Threkeld, Z. C. (2018). *Twitter as data.* Cambridge University Press, Cambridge.

Thaler, D. (2019). *PUMSutils.* R package version 0.3.2.

Twitter (2019a). *Twitter GET statuses/user_timeline API.*

Twitter (2019b). *Twitter GET users/lookup API.*

United States Census Bureau (2017). *2017 ACS 1-year Public Use Microdata Samples (PUMS). Florida population records.*

Ushey, K., Allaire, J., RStudio, Tang, Y., Eddelbuettel, D., Lewis, B., and Geelnard, M. (2019). *reticulate. Interface to 'Python'.* R package version 1.13.

Wais, K., VanHoudnos, N., Ramey, J., and Klebel, T. (2019). *genderizeR. Gender prediction based on first names.* R package version 2.1.1.

Walker, K., Eberwein, K., and Herman, M. (2019). *tidycensus. Load US census boundary and attribute data as 'tidyverse' and 'sf'-ready data frames.* R package version 0.9.2.

White, A. (2019). Family matters? Voting behavior in households with criminal justice contact. *American Political Science Review*, 113(2):1–48.

Wickham, H. (2009). *ggplot2. Elegant graphics for data analysis.* Springer-Verlag New York.

Wickham, H. (2016). *rvest. Easily harvest (scrape) Web pages.* R package version 0.3.2.

Wickham, H. (2017). *stringr. Simple, consistent wrappers for common string operations.* R package version 1.2.0.

Wickham, H. (2019). *httr. Tools for working with URLs and HTTP.* R package version 1.4.1.

Wickham, H., Henry, L., and RStudio (2019a). *tidyr. Tidy messy data.* R package version 1.0.0.

Wickham, H. and R Core team (2018). *pryr. Tools for computing on the language.* R package version 0.1.4.

Wickham, H., Ruiz, E., and RStudio (2019b). *dbplyr. A 'dplyr' back end for databases.* R package version 1.4.2.

Wilke, C. O. (2020). *cowplot. Streamlined plot theme and plot annotations for ggplot2.* R package version 1.1.1.

Wojcik, S. and Hughes, A. (2019). Sizing up Twitter users. *Pew Research Center.*