American Political Science Review (2023) 117, 3, 1151–1157
doi:10.1017/S0003055422001058 © The Author(s), 2022. Published by Cambridge University Press on behalf of the American Political Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Letter

Do Violent Protests Affect Expressions of Party Identity? Evidence from the Capitol Insurrection

GREGORY EADY *University of Copenhagen, Denmark*
FREDERIK HJORTH *University of Copenhagen, Denmark*
PETER THISTED DINESEN *University College London, United Kingdom, and University of Copenhagen, Denmark*

The insurrection at the United States Capitol on January 6, 2021, was the most dramatic contemporary manifestation of deep political polarization in the United States. Recent research shows that violent protests shape political behavior and attachments, but several questions remain unanswered. Using day-level panel data from a large sample of US social media users to track changes in the identities expressed in their Twitter biographies, we show that the Capitol insurrection caused a large-scale decrease in outward expressions of identification with the Republican Party and Donald Trump, with no indication of reidentification in the weeks that followed. This finding suggests that there are limits to party loyalty: a violent attack on democratic institutions sets boundaries on partisanship, even among avowed partisans. Furthermore, the finding that political violence can deflect copartisans carries the potential positive democratic implication that those who encourage or associate themselves with such violence pay a political cost.

The insurrection at the US Capitol on January 6, 2021, is widely considered one of the most remarkable examples of a violent attack on democratic institutions in a mature democracy in recent times (Bright Line Watch 2021). Yet, even though many politicians and pundits condemned the insurrection and mass approval of President Trump decreased in its aftermath (Bump 2021), we know little about the broader effects of political violence such as this episode on mass political behavior.

In this research letter, we take an important first step in uncovering the consequences of the Capitol insurrection for political affiliations by investigating changes in expressions of self-identification with the Republican Party and President Trump in the days immediately following the event. Specifically, drawing on recent work that uses social media self-descriptions as indicators of political identities (Rogers and Jones 2021), we study changes in identification with the Republican Party and then-President Donald Trump in the personal "bios" of around 117,000 users on the microblogging platform Twitter, all of whom express a partisan identity prior to the insurrection. Using panel data that track these users each day, we apply a flexible difference-in-differences model to estimate the causal effect of the insurrection on expressed partisanship. Our findings demonstrate that the insurrection caused an exceptionally clear immediate decrease in expressions of identification with the Republican Party and "Trumpism" that persists, at least in the short term, and is consistent across a wide series of robustness checks.

Our analysis contributes to a nascent literature concerning the consequences of violent protests for political behavior. Some prominent recent studies have brought this research agenda to the fore. One study finds that proximity to violent Black-led protests in the 1960s caused an increase in endorsements of "social control" and support for the Republican Party (Wasow 2020). Conversely, another study finds that proximity to the 1992 Los Angeles riots led to a liberal shift in policy support and an increase in Democratic Party voter registration (Enos, Kaufman, and Sands 2019). This resonates with other findings showing that protests around liberal issues correlate positively with subsequent local vote share for Democratic candidates and vice versa for conservative issues and Republican candidates (Gillion and Soule 2018). Beyond these few but important studies, however, "we know little about the effect of these events on political behavior" (Enos, Kaufman, and Sands 2019, 1012).

Gregory Eady, Assistant Professor, Department of Political Science and the Center for Social Data Science (SODAS), University of Copenhagen, Denmark, gregory.eady@gmail.com.
Frederik Hjorth, Associate Professor, Department of Political Science and the Center for Social Data Science (SODAS), University of Copenhagen, Denmark, fh@ifs.ku.dk.
Peter Thisted Dinesen, Professor, Department of Political Science, University College London, United Kingdom, and Professor, Department of Political Science, University of Copenhagen, Denmark, p.dinesen@ucl.ac.uk and ptd@ifs.ku.dk.

Received: May 25, 2021; revised: November 30, 2021; accepted: September 20, 2022. First published online: October 25, 2022.

https://doi.org/10.1017/S0003055422001058 Published online by Cambridge University Press

Our study advances our knowledge about the effects of violent protests in several important ways. First, in contrast to the aforementioned studies' focus on historical cases, we investigate how a contemporary protest shapes political attachments. This is relevant because increased affective polarization and associated politically motivated reasoning (Iyengar et al. 2019) may have reduced the role of events—even extreme ones—in shaping political attachments. In other words, are violent political protests still consequential for voters' political affiliations in a time of high political polarization? Second, we study whether reactions to violent protests extend beyond their immediate geographical locus. Given the increasingly nationalized nature of American political behavior (Hopkins 2018), more widespread effects seem plausible. Third, in contrast to two of the studies highlighted above (Enos, Kaufman and Sands 2019; Wasow 2020), which investigated protests associated with the political left, we examine the consequences of a right-led violent protest. This is important in light of known asymmetries between left- and right-wing movements (Grossmann and Hopkins 2016; but see Gillion and Soule 2018, who find similar reactions to historical right- and left-led protests). Because the right-led insurrection arguably differs from typical left-led protests on many accounts (e.g., guiding motive and the racial and social composition of protesters) it is difficult to make any clear predictions from existing studies ex ante (Manekin and Mitts 2022). Our study thus broadens our understanding of the immediate political behavioral consequences of political violence.

Our study also has implications beyond the developing literature on the political behavioral consequences of violent protests. First, it connects to the related literature on political violence and more specifically to the costs and benefits of violence to political actors in the comparatively rare setting of a developed democracy (Rosenzweig 2021). In essence, public reactions to the Capitol insurrection indicate whether political violence is an attractive strategy for political elites to appeal to US voters. Second, our study addresses the limits (or lack thereof) of partisanship in the US, at least for the Republican Party. Often attributed to increases in political polarization (Iyengar et al. 2019; Mason 2018), the strength of partisanship is now so socially and politically consequential in the US that, in the words of a recent study, "it is difficult to overstate the importance of party loyalty" (Barber and Pope 2019, 39). Indeed, partisanship has been shown to drive economic behavior (McConnell et al. 2018) and is itself linked to violence (Kalmoe and Mason 2022). Thus, by analyzing whether (expressed) partisans are willing to forego identifying with their party in the face of exceptional political violence, we examine the scope conditions of the "unmovable" character of partisanship in the United States. Finally, by investigating the immediate consequences of the insurrection, our study addresses subsequent struggles within the Republican Party over how to respond to the insurrection (Cheney 2021) and thus contextualizes later efforts and reversals by Republican politicians to mitigate the insurrection's political costs.

RESEARCH DESIGN

Data and Sample

We collected data daily from the Twitter bios of 3.4 million geolocated US users starting seven months prior to the Capitol insurrection until approximately two months afterward (June 1, 2020–March 15, 2021). This sample was drawn from a population defined as active US social media users who are at least minimally politically engaged. We defined this as any user who followed at least one of an ideologically diverse set of major US news media accounts including *MSNBC*, *Huffington Post*, *New York Times*, *Washington Post*, *CNN*, *Wall Street Journal*, *FOX News*, and *Breitbart News.* We first collected the profiles of followers of each media account. Then, to identify active users (and reduce the likelihood of collecting bot accounts), we included only users who sent at least one tweet in the past year, sent at least 25 tweets ever, and had at least 10 followers. Finally, we included only US-based users based on geocoordinates and text location information.

In total, this yields an initial sample consisting of 3.4 million users. Because our group of interest is partisans, we reduce this sample to the subset of users with expressed partisanship (see the next section) at any point during our study period. This yields a final sample of around 117,000 users—around 3.5% of the initial sample—roughly equally divided between Republican and Democratic partisans. We show each step of our sampling process in Appendix C, where we also report descriptive statistics on Twitter metadata demonstrating that users in our final sample are relatively more active and connected across a variety of metrics (e.g., number of Tweets and likes). This implies that our final sample of expressed partisans plausibly consists of users considerably more engaged than Twitter users overall, who are themselves more engaged than the average American (Blank 2017). Although we cannot infer to the population more broadly, this should make for a comparatively hard test of the malleability of expressed partisanship, as party identification is generally more stable among more politically engaged individuals (Green, Palmquist, and Schickler 2004). Due to the General Data Protection Regulation adopted by the European Union, we are unable to publicly share individual-level data, but we make aggregate-level data available as part of our replication materials (Eady, Hjorth, and Dinesen 2022). For more on data availability, including how to potentially access the individual-level data, see the Data Availability Statement below.

Measuring Expressed Partisanship

We measure expressed partisanship based on partisan terms in users' bios. To identify these terms, we apply a keyword expansion algorithm, which is shown to be superior to ad hoc keyword selection for social media data (King, Lam, and Roberts 2017). Beginning with a minimal set of seed words ("Democrat" and "Republican"), we identify relevant terms that users would include in their profiles to explicitly indicate

https://doi.org/10.1017/S0003055422001058 Published online by Cambridge University Press

their partisanship (or remove to deidentify from it). We detail this procedure in Appendix A. For robustness, we also run analyses using only the terms “Democrat” and “Republican”.

By relying on users’ Twitter profiles, our measure of partisanship departs from traditional survey-based measures. To highlight its expressive quality, we refer to it as “expressed partisanship.” However, there are reasons to expect the measure to track users’ partisan loyalties. First, our measure has high face validity: it stands to reason that publicly expressing support for a party or associated movement is an indication of identification. Second, in Appendix B we present two analyses validating our measure against the content of users’ tweets, another behavioral manifestation of partisanship. We first collect a total of 16.5 million tweets from a random subsample of Republican- and Democrat-identifying users. We then demonstrate a strong correlation between expressed partisanship and tweet sentiment: users identifying as Republican tweet less negatively about their own party than about the Democratic Party and vice versa. Then, in a more exploratory approach using supervised machine learning, we demonstrate that the terms in users’ tweets that are most predictive of Republican and Democratic expressed partisanship are highly politically loaded (e.g., among the terms most predictive of identifying as Republican are “msm” (“mainstream media”), “communist,” and “swamp”). These checks indicate that our measure of expressed partisanship picks up meaningful variation in users’ partisan loyalties.

Finally, the removal of Republican partisan terms reflects a distancing from the Republican Party but may be animated by different motives. Changes in these expressions may reflect a weakening identity, but they may also reflect that the social costs of associating with the Republican Party has increased (i.e., an act of “preference falsification”). Although we cannot address these motives directly, we do provide some tentative evidence in auxiliary analyses.

Difference-in-Differences Model

To estimate the effect of the Capitol insurrection on Republican partisan deidentification, we apply a flexible difference-in-differences (event study) model to data collected within a 10-day window around the event from users whose profiles include a Republican or Democratic keyword on at least one day within this period. This allows us to capture the dynamics of the effect of the insurrection on Republican partisan identification relative to Democrats, the natural counterfactual group, in the immediate aftermath of the event. Event study estimates also allow us to visually assess the parallel trends assumption for causal identification. More formally, we use the following model:

$$y_{it} = \alpha_i + \lambda_t + \sum_{t=-10}^{10} \beta_t \text{Republican}_i \times \text{Day}_t + \varepsilon_{it}, \qquad (1)$$

where the outcome y_{it} is a binary variable indicating whether user i’s profile contains a partisan-identity keyword on day t, and Republican$_i$ is a binary variable indicating whether user i’s partisan identity as measured by keyword use during the period is Republican (Republican = 1) or Democratic (Republican = 0). The parameters of interest, β_t, are day-specific interaction coefficients that capture the difference in differences between partisan-identifying Democrats and Republicans on a given day t relative to the day before the insurrection (t=0), which we set as a baseline. User and day fixed effects are α_i and λ_t, respectively. Finally, it is possible that any observed effect is driven by an *increase* in Democratic identity. However, as we show in Appendix D, we observe no major discontinuity in Democratic identification. The results below thus appear wholly driven by changes in expressions of Republican identification.

RESULTS

To begin, Figure 1 shows descriptively the daily net changes in the number of users who indicate a Republican identification across the entire data collection period. Immediately after the insurrection, we see a dramatic net decrease in users identifying with the Republican Party and President Trump, with a continued net deidentification over the following two months. In the three weeks immediately following the insurrection, a substantial 1 in 15 users (7%) remove Republican-identifying terms. This compares with just one in 108 users (1%) with Democratic terms. For comparison, during the three weeks before the insurrection, deidentification for Republican and Democratic terms was essentially equivalent (~0.5%). The postinsurrection drop is also far more pronounced than that following the 2020 presidential election, which is roughly equal among Republicans (2%) and Democrats (2%).

Because Figure 1 considers only Republicans, observed changes may reflect cross-partisan alienation from politics rather than reactions specific to Republicans. We thus fit our difference-in-differences model to examine changes in expressions of party identification among Republican- versus Democrat-identifying users.[1]

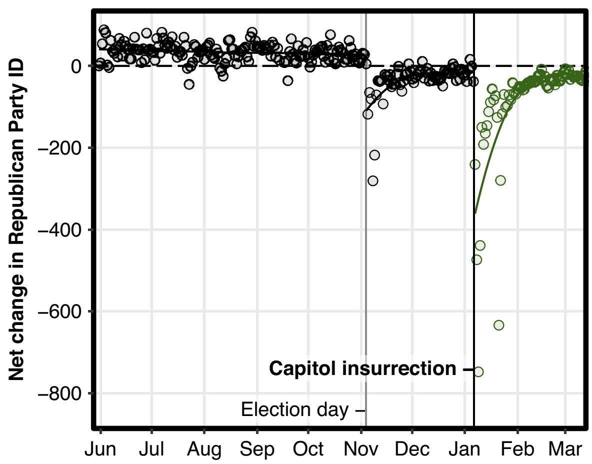
Figure 2 presents the model estimates, using data from a 10-day window around the insurrection. Each point represents the difference between Republican and Democratic users in the predicted probability of party identification relative to the preinsurrection baseline. Negative values indicate that Republican users deidentify more (i.e., drop all party-related terms) on a given day relative to Democrats.

Figure 2 clearly shows that before the insurrection, Republican and Democratic users changed their

[1] A smaller subset of users (~9%) who ever included both partisan terms in their bio is excluded because these users cannot be assigned to one group.

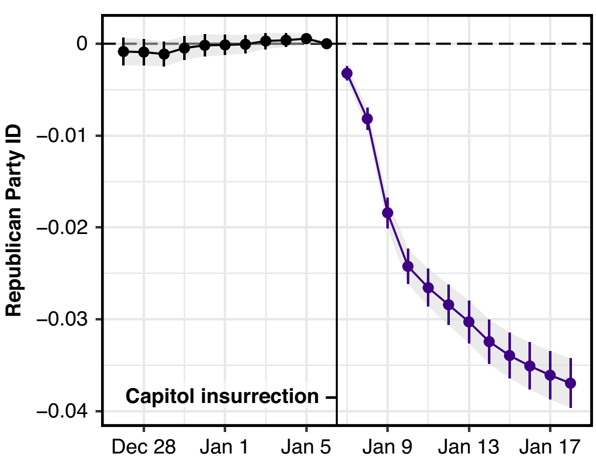
https://doi.org/10.1017/S0003055422001058 Published online by Cambridge University Press

FIGURE 1. Daily Net Change in Republican Party Identification from June 2020 to March 2021



Note: Values below zero indicate a net decrease in users with Republican identity terms compared with the previous day. LOESS regression included for reference.

FIGURE 2. Event Study Estimates (with 95% CIs)



Note: Data were collected each morning, and thus observations on January 6 (before vertical line) are preinsurrection. Standard errors are clustered at the user level.

expressed identification to a similar extent. This implies an absence of preinsurrection differential deidentification, which substantiates the parallel trends assumption. After the insurrection, the change is clear and dramatic. Within a few days, Republican users were on average 2 percentage points less likely to express a partisan identity relative to Democrats than they were before the insurrection. This relative difference increases to around 4 percentage points within a week and a half. These estimates, based on a narrower time frame than shown in Figure 1, imply that within this short period, roughly one in 25 Republican identifiers removed markers of partisan identification from their biographies. This result is substantively equivalent when examining the *count* of partisan terms in users' bios (Appendix E), and it is robust to alternative choices of partisan keywords (Appendices F and H).

Partisan terms indicating identification with Trumpism (e.g., "Trump," "MAGA") are more frequent than those referencing the Republican Party itself, which raises the question of whether the effect is wholly driven by deidentification with Trump rather than the party. Unsurprisingly, users overlap in their use of these terms (see Appendix F). However, by considering only the terms "Republican" and "Democrat", we can gauge whether the effect is driven solely by deidentification with Trumpism. As we report in Appendix F, the effect is similar, although diminished, when using only party labels.

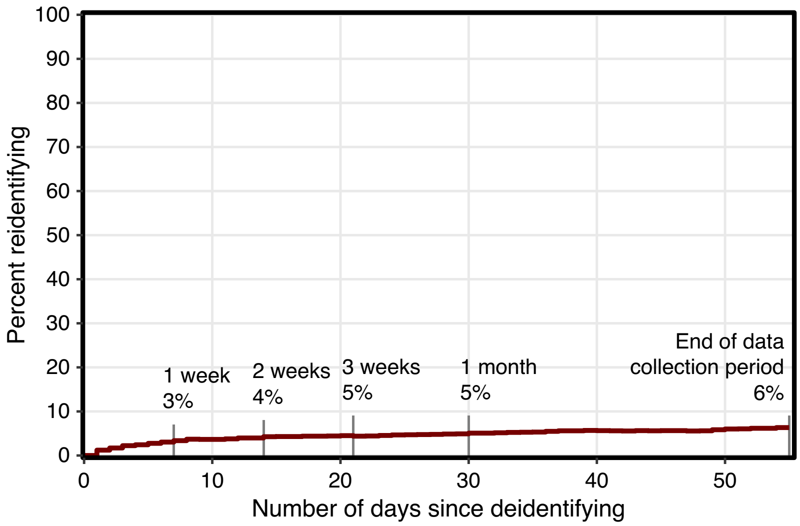
A natural follow-up question to this set of findings is whether those who deidentified as a result of the insurrection *re*identified shortly afterward. Figure 3 demonstrates that this is not the case by presenting the frequency of reidentification among Republican users who deidentified in the first week after the insurrection. By the end of our data collection period—almost two months after the insurrection—only 6% of deidentifiers had reidentified. The observed Republican deidentification was thus not merely an ephemeral shift but appears to be persistent, at least in the short term. Last, as we show in Appendix M, users identifying by more moderate terms (using only the term "Republican") before deidentifying were relatively less likely to reidentify, although the difference is not statistically significant.

We conducted a number of robustness tests and auxiliary analyses, most of which are reported in the appendix. First, we consider whether the certification of the presidential election may have driven the observed deidentification, perhaps out of aversion to being on the losing side. We cannot entirely rule out this alternative mechanism, but two observations speak against it. First, when restricting the data to those collected before the official certification (early morning on January 7), we find that deidentification begins before the certification, and the observed effect precertification on January 7 is essentially identical to that of the full sample used in Figure 2. This suggests that the trend toward deidentification started prior to the certification. Second, as Figure 1 shows, Republican deidentification in the wake of the election, where Trump's defeat was widely announced, was much smaller than was the subsequent deidentification after the insurrection. Although some supporters may still not have accepted the outcome at this point, we find it implausible that the certification could lead to Republican deidentification on a much larger scale than the election, especially because the most committed partisans are less likely to change their partisan identity in general (Green, Palmquist, and Schickler 2004).

Second, another alternative explanation is that users changed their bios out of fear of legal prosecution (including but not limited to actual rioters). Finding a significant drop when considering only the party label

https://doi.org/10.1017/S0003055422001058 Published online by Cambridge University Press

FIGURE 3. Percentage of Republican Deidentifiers who *Re*identified by the End of the Data Collection Period



Note: Data include any Republican-identifying user who deidentified within a week after the Capitol insurrection.

speaks prima facie against this alternative explanation. However, we can address this more directly by excluding users who scrubbed their timelines of potentially incriminating tweets, here proxied by removing tweets on the same day that they dropped Republican partisan terms from their bio. In Appendix I we show that the findings are robust to excluding these users.

Third, one concern might be that the results could be driven by Twitter's deletion of accounts related to QAnon—a loosely knit group of political conspiracy theorists, some of whose profiles may overlap with the set of Republican-identifying users—in the weeks after the insurrection (Singh 2021). This is not the case because deleted accounts do not affect our event study estimates on any given day, as we only exploit *within*-user variation and deleted users who drop out of the sample are not coded as deidentifying. Another possibility is that users may have preemptively scrubbed their timelines and profiles to potentially prevent detection. In Appendix J, we show that our findings are substantively unaffected when excluding any user-day observations for users whose accounts were deleted, suspended, or made private following the insurrection. Moreover, excluding users whose bios on the eve of the insurrection include terms "qanon," "wwg1wga" (a QAnon acronym), and the related "#StopTheSteal," produces effectively identical results to those in Figure 2.

Fourth, as noted in the introduction, we interpret the effect as a national-level shock rather than one driven by users geographically close to the insurrection. In Appendix K, we substantiate this by demonstrating that the effect is unchanged when excluding users geolocated to Washington, DC, and neighboring states.

Fifth, to gauge whether the observed deidentification is driven primarily by increased social costs of affiliating with the Republican Party (i.e., an act of preference falsification) as opposed to a weakened party identity, we compare event study models among users whose user names match and do not match a first name in US Social Security Administration records (as a proxy for potentially being identifiable, see Appendix L). If increased social costs were the primary animating motive, one would expect users who use a real name—and therefore bear higher costs due to potentially being identifiable—to be more likely to deidentify. We do not find this to be the case. This suggests that the observed effect is at least partly driven by a weakening of identification with the Republican Party.

Sixth, in Appendix G we consider mentions of political parties in users' tweets as an alternative outcome. Consistent with our main finding, we show that Republican deidentifiers make fewer references to parties in their postinsurrection tweets, a decrease driven by fewer references to Republicans.

Seventh and finally, in Table C1 in Appendix C we show that Republican deidentifiers are more active (e.g., tweet more) and more connected (e.g., have more followers) on Twitter than do the average Republican identifiers. This implies that this subset of users may have a broader influence than is reflected by their raw numbers.

CONCLUSION

Studying the effect of the US Capitol insurrection on expressed partisanship, we find that the insurrection caused a substantial number of Republican partisans to actively remove expressions of identification with the Republican Party and Donald Trump in its immediate aftermath. Our findings add to our understanding of

https://doi.org/10.1017/S0003055422001058 Published online by Cambridge University Press

the effects of violent protests on mass political behavior in several ways. Complementing studies of historical cases of left-wing protest, we provide evidence of the effects of violent protests in a contemporary setting and on the political right. Furthermore, we show that deidentification in response to the insurrection is nationalized—that is, it extends beyond its immediate geographical context. Last, we document that this immediate effect persists in the short term, with only a small minority of deidentifiers reidentifying during the following two months.

More broadly, our findings suggest that extreme events, such as those that violate democratic norms, can drive even some avowed partisans to distance themselves from their party. In the context of the ongoing debate about the negative consequences of polarization in the United States (Finkel et al. 2020; Iyengar et al. 2019), this finding is encouraging, as it suggests that there are limits to partisan loyalty. Our results thus complement recent work finding that exposure to incivility in same-party media leads partisans to distance themselves from their party (Druckman et al. 2019). Furthermore, this carries the positive democratic implication that political violence potentially deflects and demobilizes at least some copartisans, raising the political cost of using such tactics.

Nevertheless, the potentially positive conclusions from our study should not be overstated. Expressed deidentification is a potential indicator of distancing from the party and its leader, but it does not imply that partisan identities are no longer salient or consequential. Republican politicians' responses to the insurrection, for example, resulted in a struggle over the meaning of the party's identity, rather than its abandonment. A minority of radical partisans, as in the Capitol insurrection, may also use violence for their own ends despite its potential costs to a political party. Therefore, our results should be seen as contributing to an evolving understanding of the conditions under which partisanship may be curbed or amplified in an age of polarization (Druckman et al. 2019; Finkel et al. 2020; Iyengar et al. 2019; Kalmoe and Mason 2022).

Relatedly, because this study examines the short-term consequences of the insurrection, it helps contextualize the longer-term efforts by Republicans to minimize the insurrection's political costs. In its immediate aftermath, for example, Senate majority leader Mitch McConnell was highly critical of the insurrection and Donald Trump's role in fomenting it. Five months later—citing the insurrection's political costs—he sought to block a bipartisan commission designed to investigate it (Fandos 2021). Many rank-and-file Republican politicians, furthermore, sought to deemphasize the violence of the insurrection and obscure its partisan origins. A pressing question for future research thus concerns whether and how political elites are able to minimize the political costs of violence by strengthening the partisan identities that may have been weakened as a result of antidemocratic violence.

It is also worth considering the scope conditions of our findings. For one, they reflect the affordances and user base of a popular social media platform, Twitter. Understanding expressed partisanship and other behaviors on emerging fringe platforms (e.g., Parler), where some users may have moved their political communication after Twitter's deletion of QAnon-related accounts in the aftermath of the insurrection, is an important task for future research.

Moreover, as we have employed a novel behavioral measure of expressed partisanship, future research might examine the effect of the insurrection on other (behavioral) measures of partisanship (e.g., party registration) or investigate the downstream consequences of changes in expressed partisanship for other political behaviors and attitudes (e.g., voting behavior or policy support).

Given the short aftermath of the insurrection studied herein, it is also relevant to further investigate whether deidentification proved long-lasting, especially in light of changes in high-profile Republican politicians' approach to the event.

Last, the generalizability of our findings across the political spectrum is also a pertinent question. Put more substantively, would we expect to see a parallel deidentification among Democrats in response to the Black Lives Matter protests following the May 2020 police killing of George Floyd? Although our data collection began too late to test this directly, we conjecture, based on recent work, that this is not the case. For one, some recent work suggests that political responses to protest are contingent on protesters' group status (Manekin and Mitts 2022). Moreover, the George Floyd protests specifically appear to have prompted strong Democrats to become more liberal in their evaluations after the protests (Reny and Newman 2021), suggesting that, if anything, we may even expect an uptick in identification with Democrats after the Black Lives Matter protests. Future research would do well to further scrutinize these and other contemporary violent protests to provide an understanding of the political behavioral consequences of such significant events.

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit http://doi.org/10.1017/S0003055422001058.

DATA AVAILABILITY STATEMENT

Documentation and data for replicating results from the paper are available at the American Political Science Review Dataverse: https://doi.org/10.7910/DVN/EHH9XT.

The posted replication code and data reproduce the aggregate-level results in the article. As a consequence of the EU-wide General Data Protection Regulation, we are unable to share individual-level data in the replication materials as they are personal data, which enable identification of individuals. Thus, although we provide the code necessary to replicate all figures and tables in the article, we cannot make public the data

https://doi.org/10.1017/S0003055422001058 Published online by Cambridge University Press

needed to replicate the individual-level results. For more details, please see the documentation at the Dataverse. The individual-level data can potentially be accessed by other researchers for replication if (1) permission is obtained from the Danish Data Protection Agency and (2) a data transfer agreement is signed with the University of Copenhagen.

ACKNOWLEDGMENTS

The authors would like to thank participants at the Annual Conference of the NYU Center for Social Media and Politics, members of the Aarhus Research on Online Political Hostility project, members of the Political Communication group at Vrije Universiteit Amsterdam, and Charles Breton and Evelyne Brie for their thoughtful comments and suggestions on early versions of this manuscript.

FUNDING STATEMENT

This research was funded by the Independent Research Fund Denmark (grant number: 9038-00123B).

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

ETHICAL STANDARDS

The authors affirm this research did not involve human subjects.

REFERENCES

Barber, Michael, and Jeremy C. Pope. 2019. "Does Party Trump Ideology? Disentangling Party and Ideology in America." *American Political Science Review* 113 (1): 38–54.

Blank, Grant. 2017. "The Digital Divide among Twitter Users and Its Implications for Social Research." *Social Science Computer Review* 35 (6): 679–97.

Bright Line Watch. 2021. "American Democracy at the Start of the Biden Presidency January–February 2021 Surveys." Bright Line Watch Report. https://brightlinewatch.org/american-democracy-at-the-start-of-the-biden-presidency/.

Bump, Philip. 2021. "Disapproval of Trump Soars in Wake of Insurrection Attempt." *Washington Post*, January 15.

Cheney, Liz. 2021. "The GOP is at a Turning Point. History is Watching Us." *Washington Post*, May 5.

Druckman, James N., S. R. Gubitz, Ashley M. Lloyd, and Matthew S. Levendusky. 2019. "How Incivility on Partisan Media (De) Polarizes the Electorate." *The Journal of Politics* 81 (1): 291–95.

Eady, Gregory, Frederik Hjorth, and Peter Thisted Dinesen. 2022. "Replication Data for: Do Violent Protests Affect Expressions of Party Identity? Evidence from the Capitol Insurrection." Harvard Dataverse. Dataset. https://doi.org/10.7910/DVN/EHH9XT.

Enos, Ryan D., Aaron R. Kaufman, and Melissa L. Sands. 2019. "Can Violent Protest Change Local Policy Support? Evidence from the Aftermath of the 1992 Los Angeles Riot." *American Political Science Review* 113 (4): 1012–28.

Fandos, Nicholas. 2021. "House Backs Jan. 6 Commission, but Senate Path Dims." *New York Times*, May 19.

Finkel, Eli J., Christopher A. Bail, Mina Cikara, Peter H. Ditto, Shanto Iyengar, Samara Klar, Lilliana Mason, *et al.* 2020. "Political Sectarianism in America." *Science* 370 (6516): 533–36.

Gillion, Daniel Q., and Sarah A. Soule. 2018. "The Impact of Protest on Elections in the United States." *Social Science Quarterly* 99 (5): 1649–64.

Green, Donald P., Bradley Palmquist, and Eric Schickler. 2004. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, CT: Yale University Press.

Grossmann, Matt, and David A. Hopkins. 2016. *Asymmetric Politics: Ideological Republicans and Group Interest Democrats*. Oxford: Oxford University Press.

Hopkins, Daniel J. 2018. *The Increasingly United States: How and Why American Political Behavior Nationalized*. Chicago, IL: University of Chicago Press.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2019. "The Origins and Consequences of Affective Polarization in the United States." *Annual Review of Political Science* 22:129–46.

Kalmoe, Nathan P., and Lilliana Mason. 2022. *Radical American Partisanship: Mapping Violent Hostility, Its Causes, and the Consequences for Democracy*. Chicago, IL: University of Chicago Press.

King, Gary, Patrick Lam, and Margaret E. Roberts. 2017. "Computer-Assisted Keyword and Document Set Discovery from Unstructured Text." *American Journal of Political Science* 61 (4): 971–88.

Manekin, Devorah, and Tamar Mitts. 2022. "Effective for Whom? Ethnic Identity and Nonviolent Resistance." *American Political Science Review* 116 (1): 161–80.

Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity*. Chicago, IL: University of Chicago Press.

McConnell, Christopher, Yotam Margalit, Neil Malhotra, and Matthew Levendusky. 2018. "The Economic Consequences of Partisanship in a Polarized Era." *American Journal of Political Science* 62 (1): 5–18.

Reny, Tyler T., and Benjamin J. Newman. 2021. "The Opinion-Mobilizing Effect of Social Protest against Police Violence: Evidence from the 2020 George Floyd Protests." *American Political Science Review* 115 (4): 1499–507.

Rogers, Nick, and Jason J. Jones. 2021. "Using Twitter Bios to Measure Changes in Self-Identity: Are Americans Defining Themselves More Politically over Time?" *Journal of Social Computing* 2 (1): 1–13.

Rosenzweig, Steven C. 2021. "Dangerous Disconnect: Voter Backlash, Elite Misperception, and the Costs of Violence as an Electoral Tactic." *Political Behavior* 43 (December): 1731–54.

Singh, Kanishka. 2021. "Twitter Suspends Tens of Thousands of Accounts Dedicated to Sharing QAnon Content." *Reuters*, January 12.

Wasow, Omar. 2020. "Agenda Seeding: How 1960s Black Protests Moved Elites, Public Opinion and Voting." *American Political Science Review* 114 (3): 638–59.

https://doi.org/10.1017/S0003055422001058 Published online by Cambridge University Press

Online Appendix for

"Do Violent Protests Affect Expressions of Party Identity? Evidence from the Capitol Insurrection"

Gregory Eady, Frederik Hjorth, and Peter Thisted Dinesen

Contents

A Obtaining keywords that identify partisan identity

To select the keywords in users' profiles that indicate Democratic and Republican identity, we use the keyword expansion algorithm developed by King, Lam and Roberts (2017) (KLR). The algorithm works by starting with a set of seed words that are chosen by a researcher to define a set of texts—in this case, Twitter profiles—that indicate a concept or class of interest, i.e. partisan identity. It then applies a supervised learning model to predict texts of that class. Unlike in a standard supervised learning setup, however, the trained model is used to provide the researcher with a list of candidate keywords that are associated with the concept or class of interest. This allows for human input in the selection of the keywords that are most appropriate to defining that concept or class. This approach is useful for our purposes because it allows us to interpret and identify the terms that expressly signal a user's partisanship, and thus the terms that users would add to their profile to explicitly indicate their political identity, or remove from their profile to de-identify themselves from it.

We apply the KLR algorithm by first defining partisan identity minimally using the seed words `republican` and `democrat`, and applying the keyword expansion algorithm for each seed word separately. As recommended by KLR (King, Lam and Roberts, 2017), for each seed word we create a reference set R that contains all Twitter profiles that include the relevant seed word, and a search set S that includes all other profiles that do not contain the seed word. We then fit a Naïve Bayes classifier to a training set that consists of both sets R and S, and obtain a predicted probability that each profile in the search set S is a member of the class R. The basic idea is that there are profiles in the search set that do not contain the seed word(s), but that nevertheless—based on other terms within a profile—appear as if they might belong in the reference set. Profiles in the search set that are predicted to have a high probability of being in the reference set will contain terms associated with membership in that set, and thus contain keywords that a researcher might deem suitable for defining the concept or class of interest. To classify the profiles that appear as if they belong in the reference set R, we then partition the search set S into sets T ('target') and $S \setminus T$ (not

Table A1: Keyword target list based on `republican` seed word

	Feature	Likelihood	p	n_target	n_reference
1	`!`	189.61	0.00	210.00	230.00
2	🇺🇸	125.85	0.00	84.00	49.00
3	🌊	81.77	0.00	38.00	10.00
4	`#maga`	70.38	0.00	34.00	10.00
5	`.`	67.88	0.00	790.00	2432.00
6	`trump`	58.54	0.00	28.00	8.00
7	`conserv`	58.39	0.00	25.00	5.00
8	`love`	54.93	0.00	74.00	95.00
9	`god`	52.95	0.00	40.00	28.00
10	`,`	40.38	0.00	629.00	2025.00
11	`#resist`	32.39	0.00	25.00	18.00
12	`maga`	31.12	0.00	11.00	0.00
13	`america`	29.80	0.00	16.00	5.00
14	`#kag`	29.64	0.00	12.00	1.00
15	`#trump2020`	29.16	0.00	14.00	3.00
16	`wife`	26.59	0.00	35.00	44.00
17	`dog`	26.44	0.00	26.00	25.00
18	`vote`	26.22	0.00	12.00	2.00
19	`christian`	23.73	0.00	17.00	11.00
20	`mother`	21.86	0.00	26.00	30.00

'target') based on a probability threshold $p = 0.1$. Finally, following KLR (King, Lam and Roberts, 2017), we rank the keywords in the search set S based on a likelihood ratio based on their frequency in T and $S \setminus T$. For the specifics of each step in the algorithm, see Table 1 in KLR (King, Lam and Roberts, 2017).

The words with the highest likelihood ratio scores are presented in Table A1 for the seed word `republican`, and in Table A2 for the seed word `democrat`. Keywords in these tables indicate the terms that are relatively more common to user bios in the reference set R that is defined, respectively, by the seed word `republican` and the seed word `democrat`.

Based on the top terms for the seed word `republican` in Table A1, we select the following keywords in addition to `republican`: `#maga`, `trump`, `maga`, `#kag`, and `#trump2020`. As will be relatively well-known, the term `#maga` refers to Donald Trump's campaign slogan "Make Again Great Again", and the term `#kag`, the related term "Keep America Great".

Table A2: Keyword target list based on `democrat` seed word

	Feature	Likelihood	p	n_target	n_reference
1	.	230.28	0.00	2762.00	7231.00
2	`trump`	220.46	0.00	115.00	30.00
3	🌊	188.21	0.00	90.00	18.00
4	,	167.52	0.00	2306.00	6165.00
5	`democrat`	147.41	0.00	51.00	1.00
6	`#resist`	139.98	0.00	73.00	19.00
7	`polit`	91.59	0.00	103.00	92.00
8	`proud`	79.15	0.00	104.00	107.00
9	`vote`	78.04	0.00	33.00	4.00
10	`conserv`	74.89	0.00	60.00	36.00
11	`mother`	73.42	0.00	74.00	59.00
12	`liber`	73.15	0.00	43.00	15.00
13	`#maga`	70.06	0.00	63.00	44.00
14	🇺🇸	63.19	0.00	155.00	241.00
15	`patriot`	60.91	0.00	52.00	34.00
16	`mom`	60.56	0.00	126.00	179.00
17	`progress`	58.94	0.00	38.00	16.00
18	`wife`	52.63	0.00	90.00	113.00
19	`feminist`	51.94	0.00	26.00	6.00
20	`countri`	50.58	0.00	50.00	39.00

Based on the top terms for the seed word `democrat` in Table A2, we select the following keywords in addition to `democrat`: 🌊 and `#resist`. We note that the emoji 🌊 is commonly used on Twitter to indicate support for a 'blue wave' (Democratic) election; the term `#resist`, resistance to Republican leadership, primarily in reference to Donald Trump. Notably, unlike the term `trump`, which is highly associated with Republican identity, the term `biden` does not appear in the top-ranked list of keywords when using the seed word `democrat`. Because the Democratic Party leader and president are nevertheless linked for theoretical reasons to Democratic identity, we reproduce the main results of the article by including the term `biden` in the list of Democratic keywords. The results, shown in Appendix H, are nevertheless effectively identical to those presented in the main article.

Finally, comparing the two tables of candidate keywords, it is notable that several terms overlap, with Republican terms such as `trump` occurring in the Democratic target list. The reason is that the KLR algorithm picks up on co-occurrences of terms in profiles, that are likely to be political in nature, and can be critical of out-partisans. Some Democratic-identifying users, for example, might include the term “trump” in their profile to criticize the former Republican president. We select only terms in each list that are specific to the partisan identity of interest. As we note in the main article, we remove from the analysis the set of users who include terms from both parties. This does not, however, meaningfully affect the results.

B Validating the partisan identity measure using tweet content

In this section, we validate our Twitter bio-based measure of partisan identity by examining if partisan identities expressed in bios are reflected in the tone and content of users’ tweets. To analyze tweets from party identifiers, we collected the most recent tweets from the timelines of both Democratic and Republican identifiers using the Twitter REST API. Because the REST API sets limits on the number of tweets that one can collect per user, we used the Twitter Academic API to ensure that all tweets were collected back to at least December 1, 2020 (over a month prior to the insurrection). This is necessary for users who tweet extremely frequently, and thus for whom the limits of the REST API are insufficient: at present, the Twitter REST API allows one to collect the 3,200 most recent tweets from a given user.

Using Twitter’s Academic API, we collected the full tweet history back to at least December 1, 2020, for a random selection of 25,000 Republican-identifying users and 10,000 Democratic-identifying users.[1] In total, we collect $16,535,233$ tweets from this sample of users.

[1]We oversample Republican-identifying users to be able to compare de-identifiers and non-de-identifiers within this subset with more precision.

Using this sample of tweets, we are able to validate our Twitter bio-based measure of party identification in two ways. First, as a basic validation, we compare the sentiment of tweets mentioning either Democrats or Republicans in both groups of partisans. We measure sentiment using the default sentiment scoring function in the `sentimentr` package in R. If Twitter bios reflect party identification, we should expect partisans to speak more negatively when mentioning the out-party than when mentioning their own party.

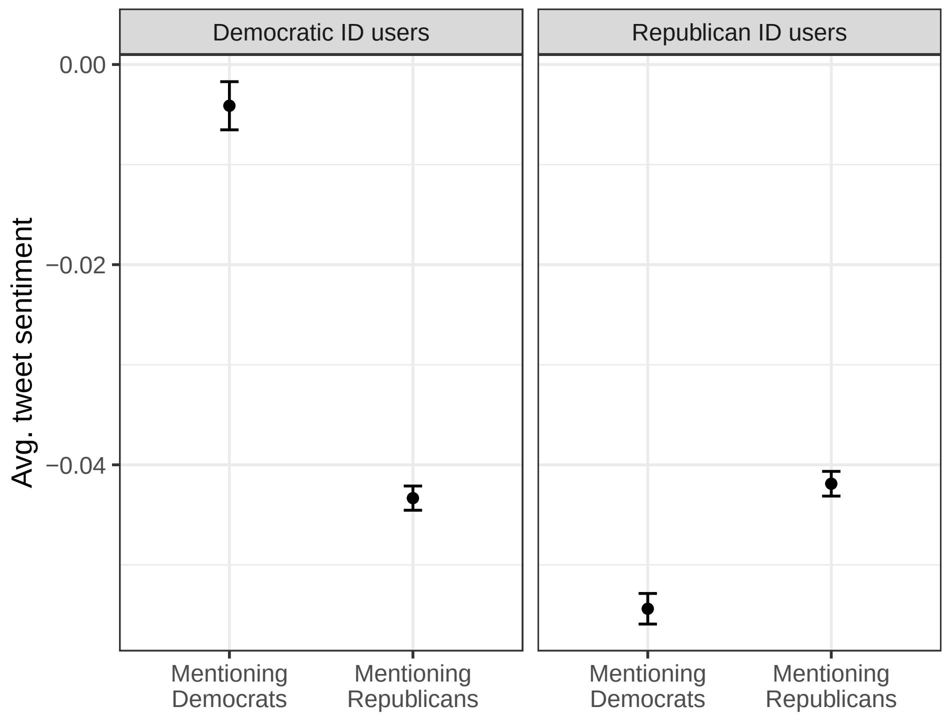


Figure B1: Average sentiment of tweets mentioning either `democrat*` or `republican*` for Democrat- and Republican-identifying users. Tweets referring to both parties are omitted.

Figure B1 demonstrates that this is indeed the case: Democrat-identifying users use more positive (less negative) language when talking about Democrats compared to when they talk about Republicans, and vice versa. Hence, the partisanship users express in their bios is systematically reflected in the valence with which they speak about the two major political parties.

As a second validation, we use supervised machine learning to explore which terms in users' tweets that are most predictive of the expressed partisanship in their bios. Specifically, we fit a regularized logistic regression model (a lasso model) (Friedman, Hastie and Tibshirani, 2010) to the tweet data, where the outcome is a user's partisan identification as defined by their bio (Democratic = 0, Republican = 1), and the features (independent variables) are the term frequencies (e.g. words, hashtags) from their tweets.

Regularized regression models are used to avoid overfitting and for cases in which there are large numbers of features (variables), e.g. with text data. These models thus make fitting regression models to Twitter data tractable, given the large number of terms in individual Tweets (e.g. Mitts, 2019). Regularized (lasso) regression penalizes large coefficients such that only features with the most predictive power are assigned meaningfully large coefficients.[2] The model thus selects those features (e.g. terms in tweets) that are the most important. This is referred to as the 'selection property' of these models. In our case, the model thus selects the subset of terms in tweets that are most strongly associated with users who have Democratic or Republican terms in their bios.

We present the key results from the regularized regression model, in the form of the 20 most predictive terms in either direction, in Figure B2. The top panel presents the most predictive terms overall. Because many of the most predictive terms are names ('handles') of Twitter users (e.g., the right-wing news site @`gatewaypundit` or the presidential historian @`beschlossdc`), the bottom panel presents the top results when omitting Twitter handles.

As shown in Figure B2, the terms most predictive of either Republican or Democratic partisanship are clearly politically loaded. For example, among the terms most predictive of identifying as a Republican are `msm` (i.e., 'mainstream media'), `communist`, and `swamp`. Conversely, among the terms most predictive of identifying as a Democrat are `healthcare`, `president-elect`, and `complicit`, all recognizable terms from late Trump-era partisan political discourse. In other words, Twitter users identifying as Republican or Democrat use terms in their tweets meaningfully reflecting their partisanship.

Jointly these two validations demonstrate that Twitter users' partisan identification is

[2]In practice, other features are assigned coefficients near zero.

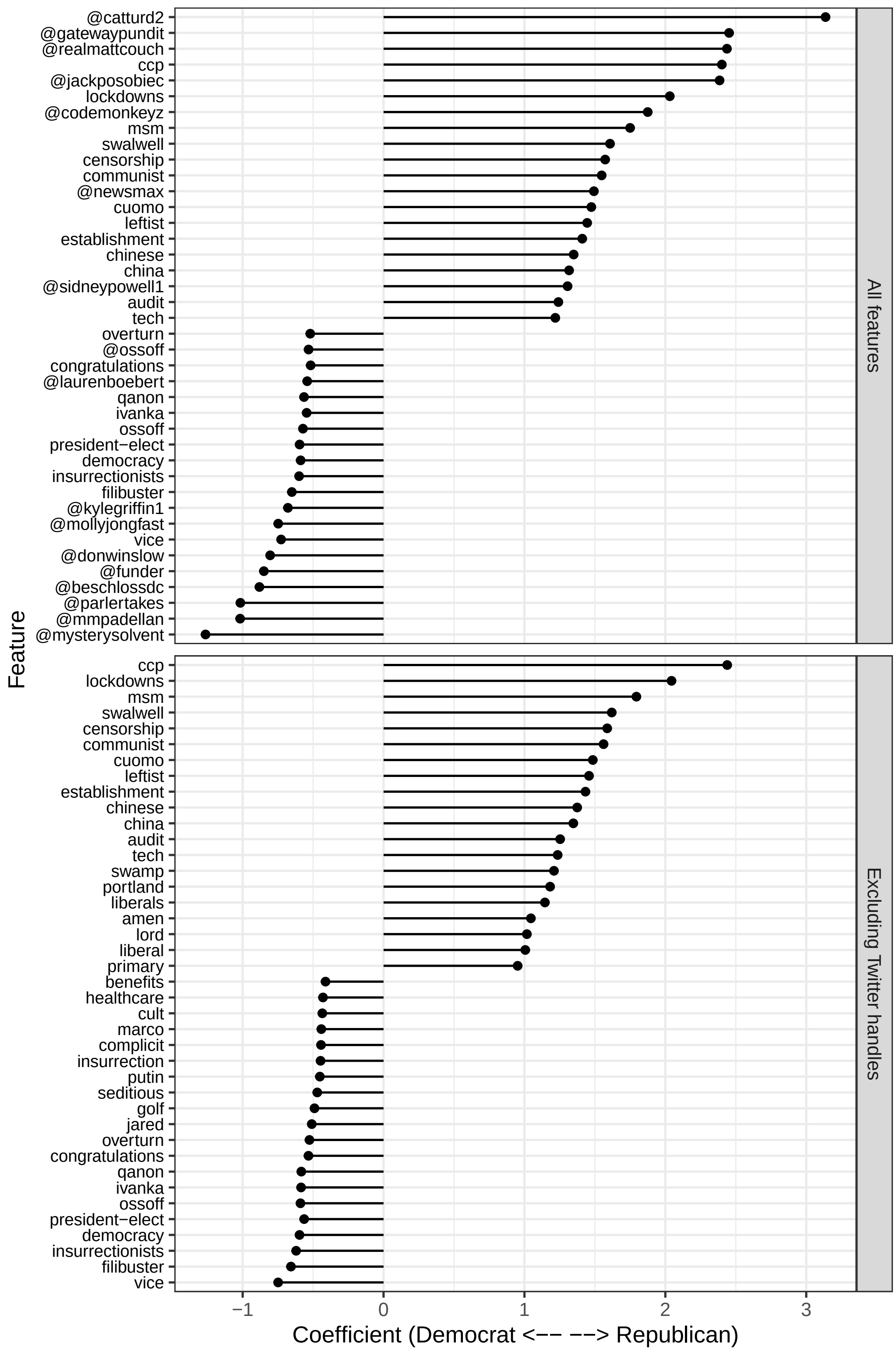


Figure B2: Top 20 terms in user tweets most predictive of identifying as a Republican (positive coefficients) or a Democrat (negative coefficients), based on a regularized regression model. Top panel shows results for all terms. Bottom panel shows results when omitting Twitter handles.

reflected in their tweets both with respect to tone (Figure B1) and content (Figure B2). This shows that partisan identities expressed in bios are tightly connected to users' everyday (political) behavior on Twitter, which in turn indicates that they are reflections of party political identities.

C Sampling and sample characteristics

Figure C1 presents an overview of our sampling process. See also the subsection "Obtaining user biographies" in the paper for more details.

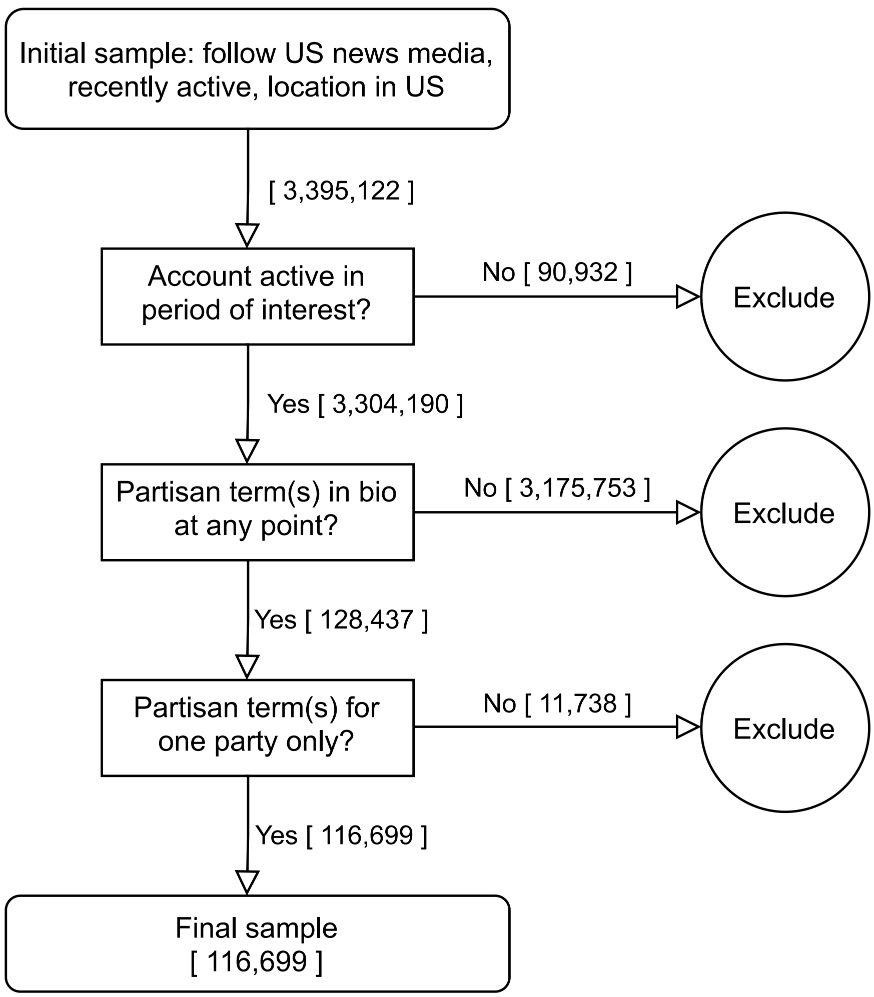


Figure C1: Flowchart outlining sample selection. Sample size at various stages of the selection process shown in brackets.

Table C1 presents summary statistics of Twitter meta-data (median number of tweets, median number of friends etc.) about the sample of (identifying) users (n = 116,699) used in the sample compared to all users (n = 3,395,122). As shown, users in the final sample are considerably more active and connected on Twitter across all the available metrics. It is thus reasonable to assume that they are a (politically) engaged subset of users.

Column 3 and 4 report meta-data for Republican identifiers and de-identifiers to give an idea about the characteristics of how those de-identifying differ from the larger pool of Republican identifiers. Here we see that the Republican users that de-identified in the wake of the January 6 insurrection are generally more active and more connected on Twitter.

Table C1: Summary information about users included in the sample used for analysis

	Initial sample	Final sample	Republican identifiers on January 6	Republican de-identifiers⋆
Median number of tweets	900	3,137	3,052	4,818
Median number of likes	779	4,651	3,880	7,208
Median bio length	66	119	120	117
Median number of followers	148	437	458	857
Median number of friends†	425	879	847	1,348
N	3,395,122	116,699	58,630	2,459
Proportion of initial sample	100%	3.4%	1.7%	0.1%
Proportion of sample in previous column	—	3.4%	50.2%	4.2%

⋆ Republican de-identifiers are defined as any user who had a Republican-associated term in their bio the day before the Capitol insurrection, but removed those terms within a week afterward.

† In Twitter parlance, "friends" are the accounts that a user follows. "Followers" are the accounts that follow a user.

Finally, we note that on a few dates (June 5, July 11, August 19, September 14-16), the data are incomplete for technical reasons. However, because all these are well before the insurrection, this does not affect our results.

D Net change in expressed party ID for Republicans and Democrats

In Figure D1, we present the daily net changes in the number of users whose profile includes a Republican (Panel A) or Democratic (Panel B) partisan identity term at any time during the time series. Panel A of Figure D1 is equivalent to Figure 1 in the main article. In both panels, vertical lines indicate election day (November 3) and the day of the Capitol insurrection (January 6). As shown, there is a decrease in both Republican and Democratic term following the election day. In the immediate aftermath of the Capitol insurrection, however, we see a large decrease in Republican terms, without any clear evidence of a similarly marked drop in Democratic terms. In Figure D2, we present substantively equivalent results for the count outcome for the number of Republican (Panel A) and Democratic (Panel B) terms in users' profiles.

E Results for binary and count outcomes

In Figure 2 of the main article, we present results for the binary indicator for whether a user includes a partisan-identifying term in their social media profile. In Figure E1, we also present event study results for both the binary indicator of identification (Panel A, as in Figure 2 in the main article) as well as a count indicator of the *number* of partisan identity terms (Panel B). As is clear from the estimates presented in Panel B, the results with the count outcome are substantively equivalent to those from Figure 2 in the main article (i.e. Panel A) for the binary outcome.

F Results for "`democrat`" and "`republican`" keywords only

We measure Republican Party identification by searching for the term `republican` as well as the Trump-related terms `trump`, `maga`, `#maga`, `#kag`, and `#trump2020`. Here, we present an

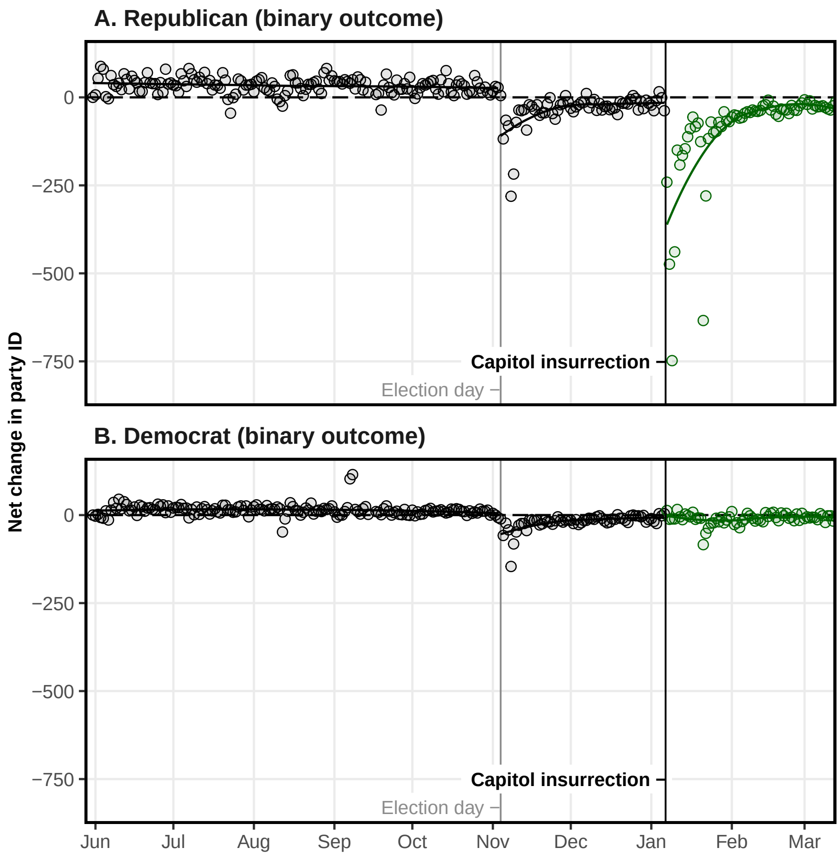


Figure D1: Change in Republican Party and Democratic Party identification over time (binary outcome).

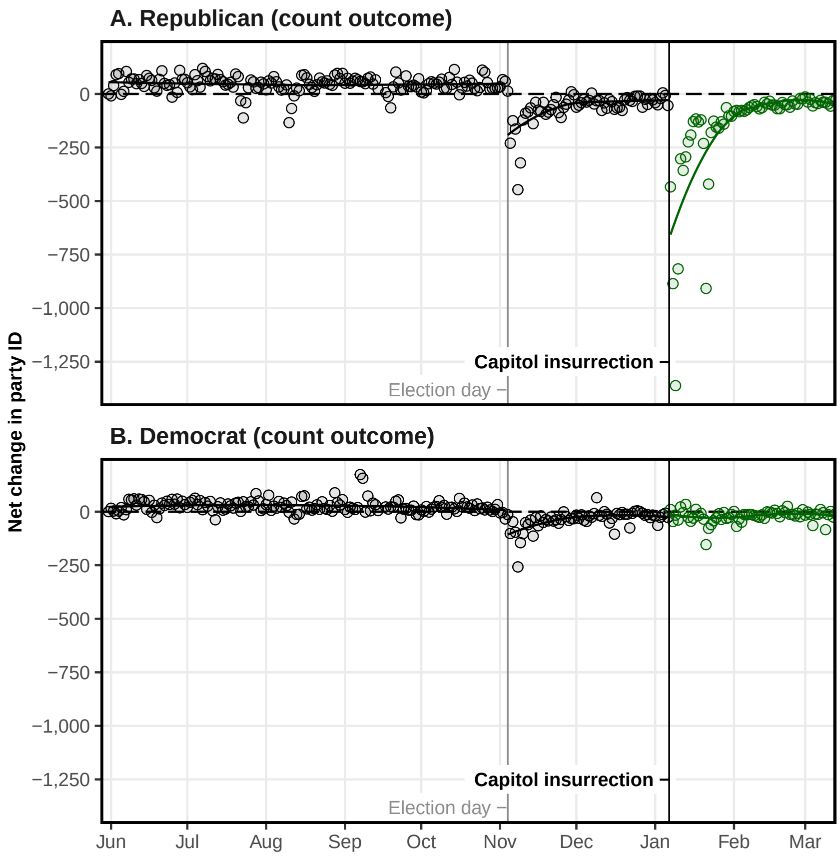


Figure D2: Change in Republican Party and Democratic Party identification over time (count outcome).

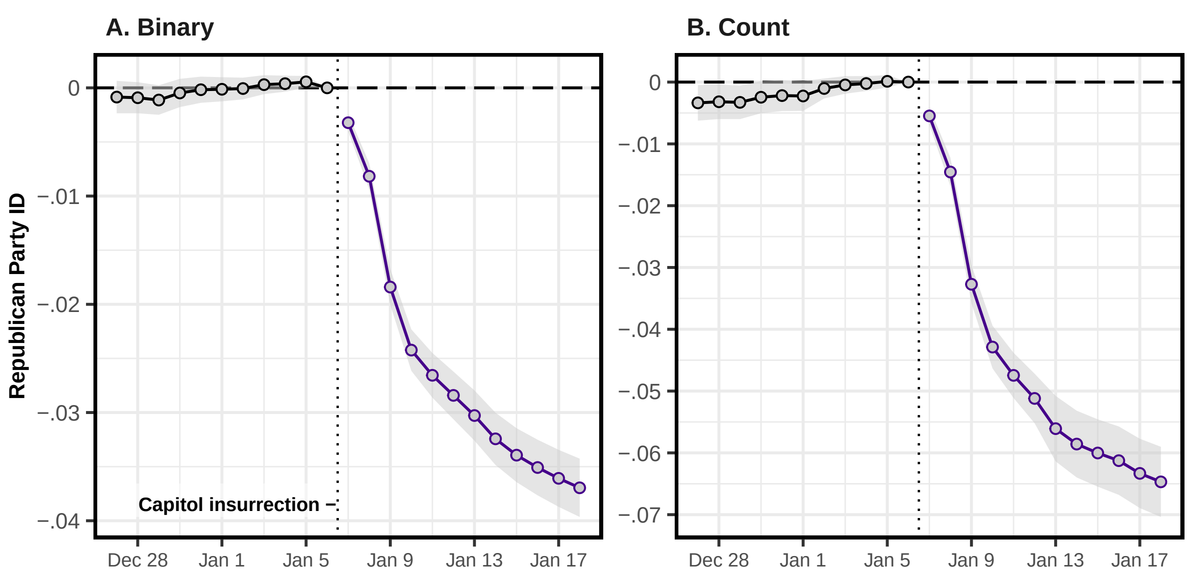


Figure E1: Primary event study results for both binary and count outcomes overview of how the two types of terms are distributed across the user bios in our data. For simplicity, we consider only the data from the first day of the time frame presented in Figure 2 in the main text, i.e. December 27$^{\text{th}}$. On this day, the two types of terms are distributed across the 3,301,900 tracked users as shown in Table F1.

Table F1: Usage of 'Republican' and Trump-related terms in a cross-section of users.

	0 Trump terms	1+ Trump terms
0 'Republican'	3,211,752	80,843
1+ 'Republican'	6,401	2,904

As shown, most user bios feature neither type of term, but as one would expect, usage of 'Republican' and Trump terms are positively correlated. Among users with one or more Trump terms, 3.5 percent include 'Republican' in their bio; the same is true of just 0.2 percent of users without Trump terms. The association is statistically significant ($\chi^2 = 31,022$, $p < .001$).

Because we use sets of terms of different lengths to capture Republican and Democratic identification (see Appendix A), count measures of identification are not directly comparable. Moreover, since there is inevitably some discretion involved in selecting the exact terms used to capture party identification, a reasonable concern could be that the main result reflects only this particular set of terms (King, Lam and Roberts, 2017). Here, we show that the main result holds using a maximally restrictive definition, comparing only usage of the seed words `democrat` and `republican` in users' profiles.

In Figure F1 we present the daily net changes using partisan identity defined solely by the seed words `democrat` and `republican`. Because we examine only a single term, the results in each panel are relatively noisy. However, we observe an immediate net decrease in the use of the term `republican` in the first few days following the insurrection (1.9% in the three weeks after the insurrection). For the term `democrat`, we see an increase of 1.6%.

Figure F2 presents the results defined solely by the seed words `democrat` and `republican` for the event study model.[3] More substantively, this result also implies that our main

[3] Results for the count outcome (not shown) are effectively identical, because identification is only defined

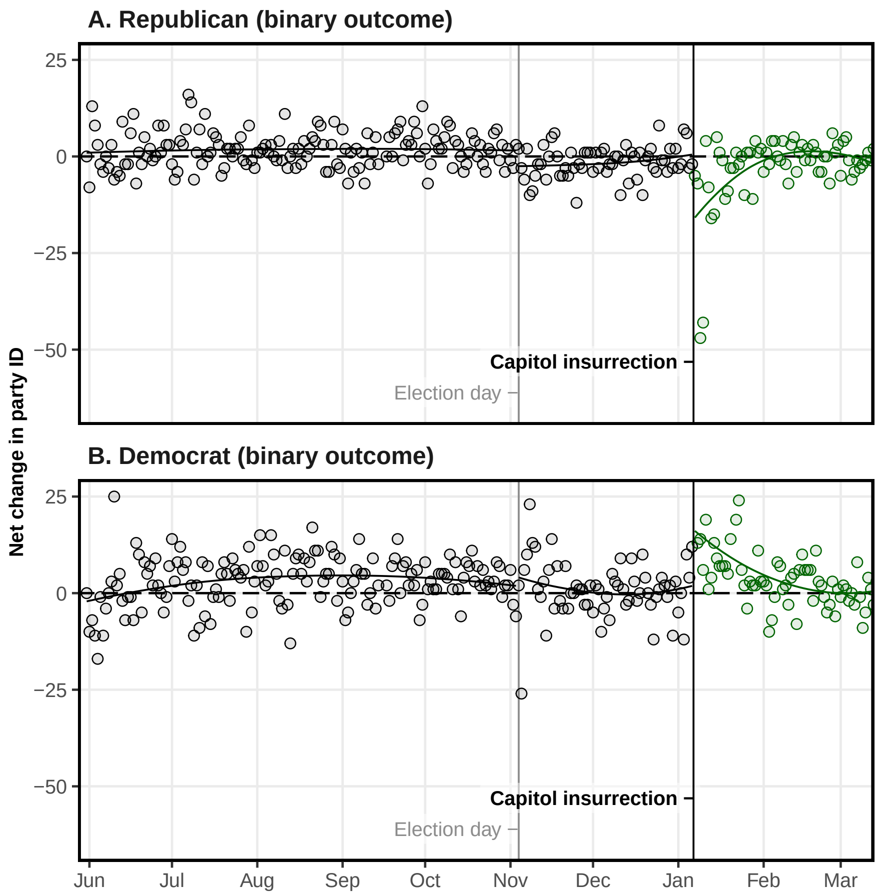


Figure F1: Net change in (binary) party identification over time ('republican' & 'democrat' only) findings regarding de-identification are robust to focusing exclusively on expressed party identification *per se* and not only movements associated with Donald Trump. Within a few days, Republican users were on average 2 percentage points less likely to express a party identity relative to Democrats than they were before the insurrection. Over the 10-day post-insurrection window, this relative difference increases further to around 2.7 percentage points.

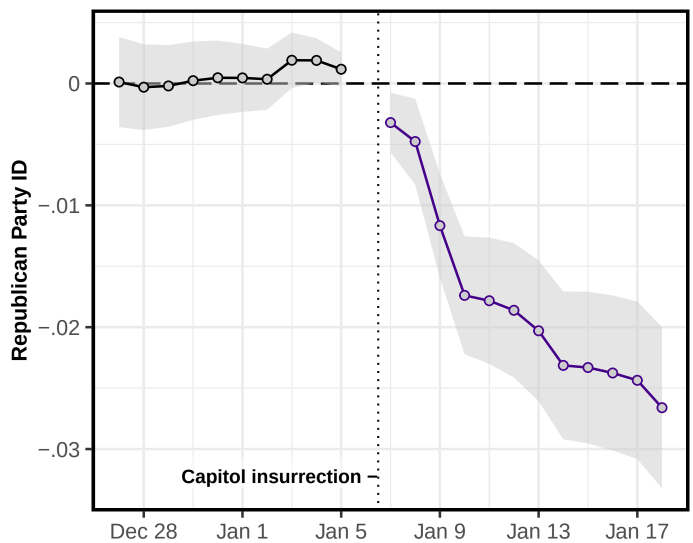


Figure F2: Event study models with identification defined only as "Democrat" or "Republican"

G De-identifiers & partisan tweeting behavior

As we note in Appendix B, we used the Twitter Academic API to collect the tweets of 25,000 Republican-identifying users back to (at least) December 1, 2020, to validate our bio-based measure of expressed partisanship. In this section we use these tweet data to investigate whether the users who removed Republican terms from their bios also subsequently sent fewer tweets containing references to the Republican Party, thereby indicating a disassociation from the party. One can think of this as a conceptual replication of our main findings, although

by the single term "`republican`".

it is important to note that this comparison is descriptive because by comparing those who de-identified to those who did not, we are examining two groups defined by whether they were affected by the insurrection (i.e. it is a post-treatment variable). Nevertheless, this comparison may shed some light on how de-identification potentially affects other partisan behaviors, and help open an avenue for future research.

To examine this, we identify all tweets among the sample of 25,000 Republican-identifying users that contain party-related terms `"republican"`, or `"democrat"`, or both. We then calculate the count of tweets sent in the pre- and post-insurrection period for each user, and the count of the total number of tweets sent. To investigate whether those who de-identify subsequently sent fewer tweets concerning the Republican Party, we fit a binomial model to the data to predict how many tweets contain a party term among all tweets sent by each user in the pre- and post-insurrection period. As predictors, we include a binary variable indicating whether a user de-identified, whether the observation is from the post-insurrection period, and an interaction between these two variables. The interaction term captures whether de-identifiers sent fewer tweets concerning the party in the post-insurrection period compared to non-de-identifiers, relative to the pre-insurrection period.

Results are presented in Table G1. In the first model, we see a decrease in the proportion of de-identifiers' tweets that contain references to `"democrat"` or `"republican"` in the post-insurrection period relative to those who do not de-identify. In the second model, we see a larger such relationship when we confine that term to be `"republican"`; in the third model, we observe no decrease in the use of the term `"democrat"`.[4] In sum, users who de-identified by removing Republican-related terms from their bios in the immediate aftermath of the insurrection, would also go on to reference the party less frequently in their tweets compared to those users who did not de-identify.

[4] Results are effectively equivalent with a quasi-binomial model that accounts for over-dispersion.

Table G1: Change in tweeting of partisan terms before and after the insurrection among de-identifiers and identifiers

	Dem. or Rep. terms	Rep. term	Dem. term
	(1)	(2)	(3)
De-identifier × post-insurrection	−0.047***	−0.107***	−0.005
	(0.013)	(0.019)	(0.017)
De-identifier	−0.043***	−0.134***	0.043***
	(0.011)	(0.015)	(0.014)
Post-insurrection	0.167***	0.137***	0.182***
	(0.005)	(0.007)	(0.007)
Intercept	−2.901***	−3.530***	−3.549***
	(0.004)	(0.006)	(0.006)
Observations	28,822	28,822	28,822
Log Likelihood	−84,321.670	−62,145.930	−65,886.590

*p<0.1; **p<0.05; ***p<0.01

H Results including "`biden`" in the Democratic keyword list

In the main article, we present results from party identification keywords that were selected using the keyword expansion technique described in Appendix A with the seed words `democrat` and `republican`. Given the relative infrequency/co-occurrence of the term `biden`, it was not selected from the keyword expansion technique itself. However, because identification with the Democratic Party leader and president, Joe Biden, is itself linked to party identification for theoretical reasons, we also examine the effect of the Capitol insurrection using the same set of keywords as described in Appendix A, but also including the term `biden`.

Results from an event study model with this expanded keyword list are presented in Figure H1. The results demonstrate an effect and dynamics similar to that presented in Figure 2 of the main article. However, unlike the estimates in Figure 2 (and the party-only estimates in Figure F2), it is clear that there are not parallel trends between Democratic Party and Republican Party identification prior to the Capitol insurrection when the term `biden` is included in the keyword list. The validity of difference-in-differences models relies on the assumption of parallel trends: that prior to an intervention, the outcome variable for the groups of interest would move in sync to the extent that, counterfactually, these trends would continue in parallel in the post-intervention period were it not for the intervention itself (Cunningham, 2021). Indeed, given the linear trend in the pre-treatment period, it is clear visually that the effect of the Capitol insurrection will be *under*-estimated relative to a counterfactual in which the pre-treatment trend continued in the post-treatment period.

To address this, we use a semi-parametric event study model in which we model the pre-treatment trend linearly such that we fit the following model:

$$y_{it} = \alpha_i + \lambda_t + \sum_{t=1}^{T} \beta_t \text{Republican}_i \times \text{Day}_t + \delta t \times \text{Republican}_i + \epsilon_{it}, \tag{H1}$$

where the outcome variable y_{it} is a binary variable indicating whether user i's profile includes a keyword representing their partisan identity at time t, and Republican_i is a binary variable indicating whether the user has identified as a Republican (Republican $= 1$) or Democrat (Republican $= 0$). The parameters α_i and λ_t are user and time fixed effects respectively.

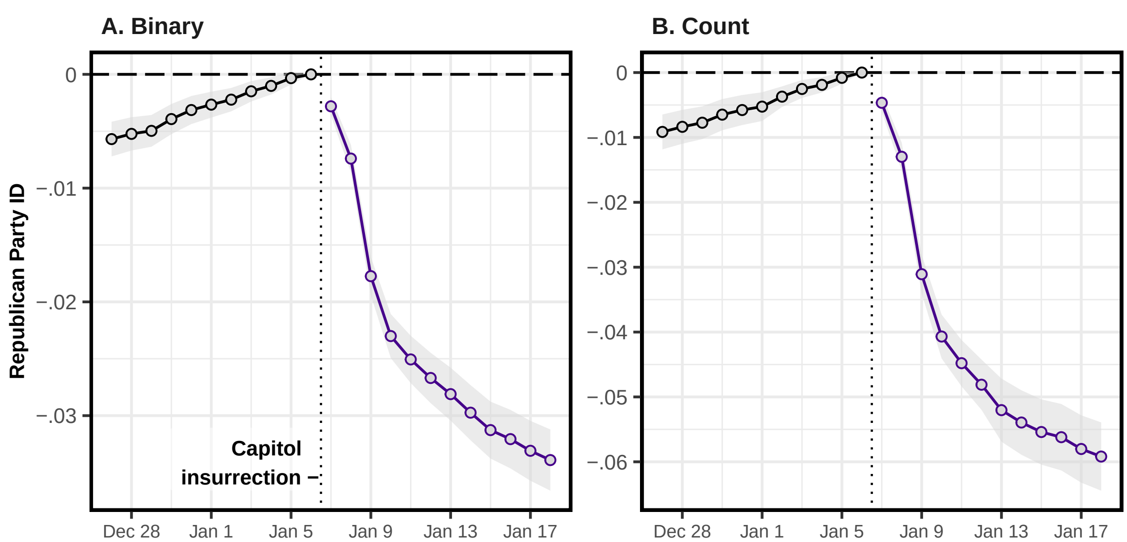


Figure H1: Event study estimates including the term `biden` in the Democratic ID keywords list

Importantly, the parameter δ captures the difference in the trend in party-identifying terms in Republican profiles relative to Democratic profiles. Our parameters of interest, β_t, thus capture post-treatment deviations from the existing pre-treatment trend in differences in profiles between Republican- and Democratic-identifying users.

As is clear from Figure H1, these deviations are visually obvious, and results from the event study model as defined in Equation H1 and presented in Figure H2 bear this out (including for the count outcome). Relative to the expected difference in the differences between Republican- and Democratic-identifying profiles (dashed line at 0), we see a substantial decrease in Republican-identifying users as a result of a Capitol insurrection. Thus the effect of the insurrection on Republican identification when using the additional term "`biden`" in the Democratic keyword list is effectively equivalent to that shown in Figure 2 of the main article.

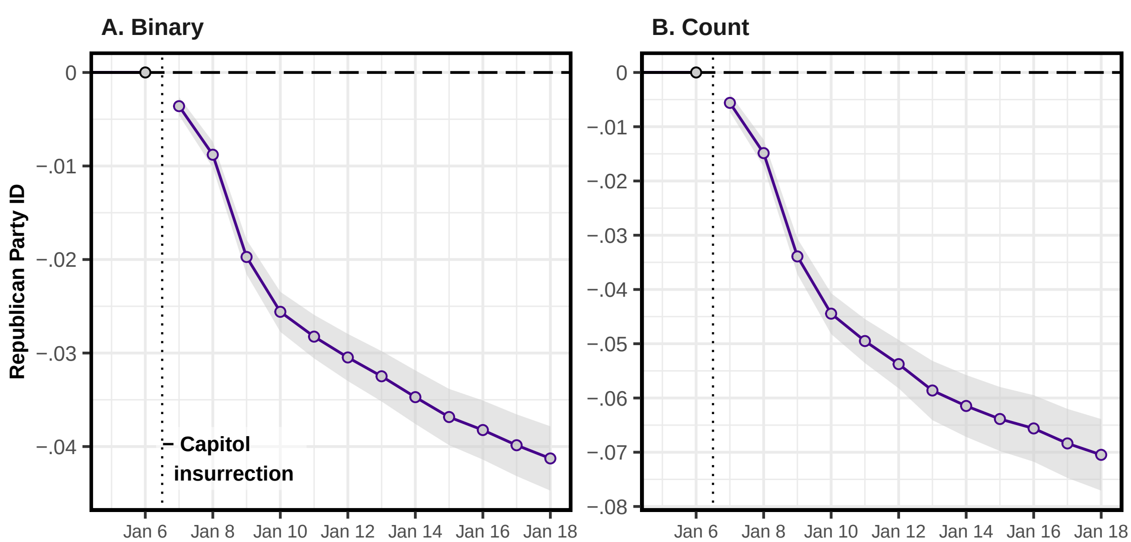


Figure H2: Event study estimates including the term `biden` in the Democratic ID keywords list (with modeled trend)

I Results excluding users who deleted or scrubbed their Twitter timelines

In this section, we examine whether de-identification with the Republican Party and Donald Trump was driven by users' fear of prosecution. It is possible, for example, that some users who participated in the insurrection scrubbed their social media accounts of any potentially incriminating information in the insurrection's aftermath. Although users' Twitter profile (bio) information alone is unlikely to contain incriminating material, some users may nevertheless have removed political information from their accounts in general, including any that can signal partisan affiliation. If a substantial number of such users are in our data, the results might better be interpreted as driven by fear of prosecution rather than de-identification itself.

To test whether our results are robust to this possibility, we use information captured in the day-level user profile data that indicate how many tweets are in each user's timeline. By examining day-to-day changes in the number of tweets in users' timelines, we can identify and remove from the analysis any user who deleted (i.e. potentially scrubbed) a substantial number of tweets on the day that they removed any political party-related identification from their profile. It may be the case, of course, that users also delete tweets from their timelines for innocuous reasons or remove tweets that indicated support for the January 6 "Stop the Steal" rally before it turned violent, thus removing them to disassociate themselves from support for the event. Nevertheless, removing users from the analysis who deleted tweets at the same time that they deleted party identifying information from their profile provides a useful indication of the extent to which the results are driven by users' fear of prosecution.

We implement this robustness check by setting three thresholds for users who potentially scrubbed their profiles: those who deleted at least 10 tweets; those who deleted at least 5 tweets; and, most conservative, those who deleted at least one tweet.[5] We then fit the main

[5]Although deletion of a single tweet may appear as an extremely conservative threshold, it may be the case that some users heavily scrub their profiles and then increase the number of tweets on their timeline by

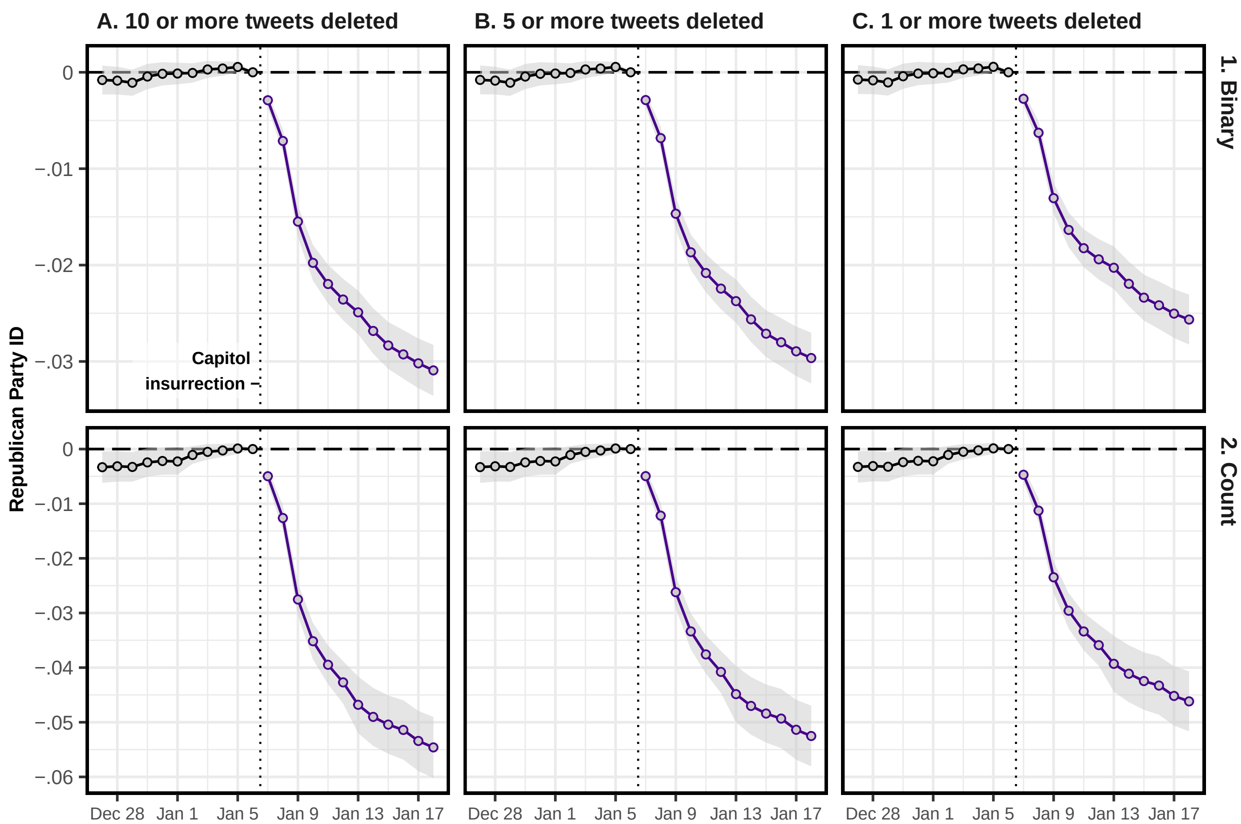


Figure I1: Event study estimates excluding users who deleted/scrubbed their timeline

event study model (equivalent to that in Figure 2 in the article) to the data by excluding each of these sets of users in turn. Results are presented in Figure I1, for both the binary (first row) and count (second row) outcomes. As the figure demonstrates, the results are highly robust to the exclusion of users who potentially scrubbed their profile of incriminating information. The magnitude of the effect of the Capitol insurrection decreases slightly the lower the threshold for exclusion, which results from the fact that the users removed from the data are, by definition, Republican-identifying users who engaged in profile de-identification. As the figure makes clear, however, removing users who potentially scrubbed their profile has effectively no meaningful effect on the main result.

tweeting liberally on that day.

J Results excluding users whose accounts were deleted

In this section, we examine whether the results might be driven by profile changes among users who Twitter identified after the Capitol insurrection as supporters of QAnon, a loosely knit group of conspiracists concerning US politics. Twitter sought to purge adherents of the group in the weeks after the insurrection (Singh, 2021).

Users who are deleted by Twitter (or users who delete their account themselves) do not themselves affect our estimates because the event study model includes a unit-level fixed effect and a deleted user is not coded as politically de-identifying. Nevertheless, some users may pre-emptively remove information in their timeline and profile concerning their conspiracist leanings or party identification to potentially avoid deletion (despite Twitter having data on deleted information nonetheless). In Appendix I, we found that the results are robust to a relatively strict threshold for users who delete/scrub timeline information. We complement this check by removing from the analysis any user who, in the time period of interest, had their account deleted or suspended by Twitter, deleted their own account, or set it to private. To do so, we fit the event study model only to data from users for

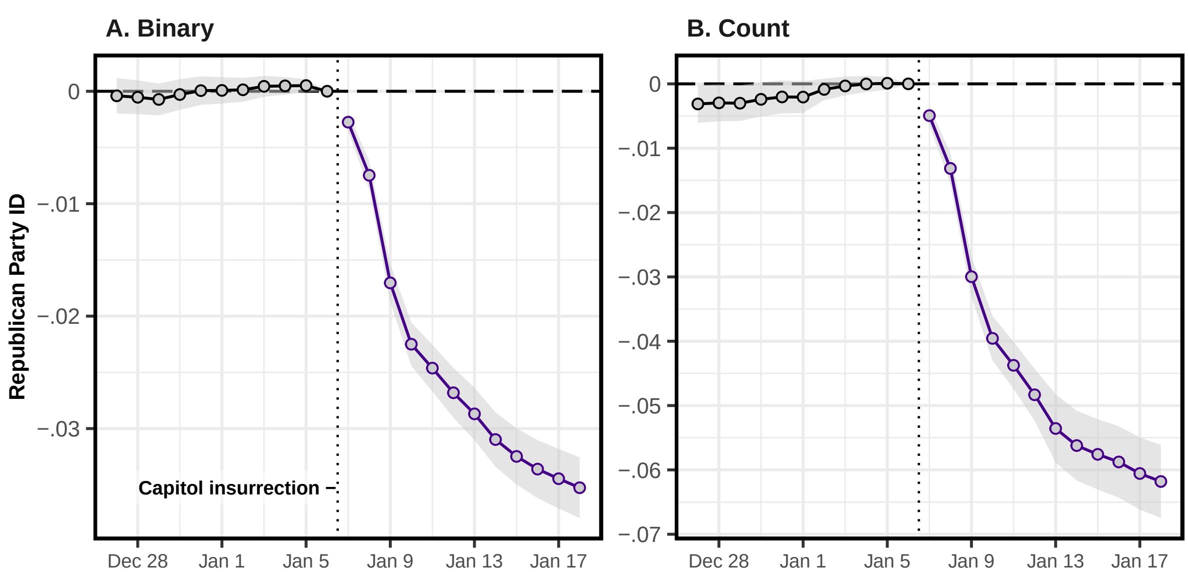


Figure J1: Event study estimates excluding users whose accounts were deleted by Twitter, were deleted by the users themselves, or were set to private whom we have complete data in the time period of interest, i.e. whose account was not deleted, suspended, or set to private. The results are presented in Figure J1. As the figure demonstrates clearly, the results are robust to the exclusion of users whose accounts were deleted or otherwise unavailable at any time during the period of interest.

K Results excluding users in states close to the Capitol insurrection

In this section, we examine whether the results are driven by users who indicate living geographically close to Washington D.C., the location of the insurrection. To examine this, we exclude from the model any user who was geo-located to Washington D.C. or the two adjacent states, Virginia and Maryland. We note that our sample contains a location, by state, for 95% of users (as noted in the article, we collect panel data only for users who can be geo-located to the US, Dredze et al., 2013).[6] The sample, furthermore, contains a relatively geographically representative set of users. The correlation between the proportion of users in our sample from each state and the 2020 census population is 0.95. Results from the models excluding users who are located close to the insurrection are presented in Figure K1. As results in the figure show, the results are effectively equivalent to those from the full sample (see Figure E1). The effect of the Capitol insurrection on de-identification, in other words, is not local to the site of the insurrection.

[6] In other words, 5% of users can be geo-located to the United States in general, but without state-specific information.

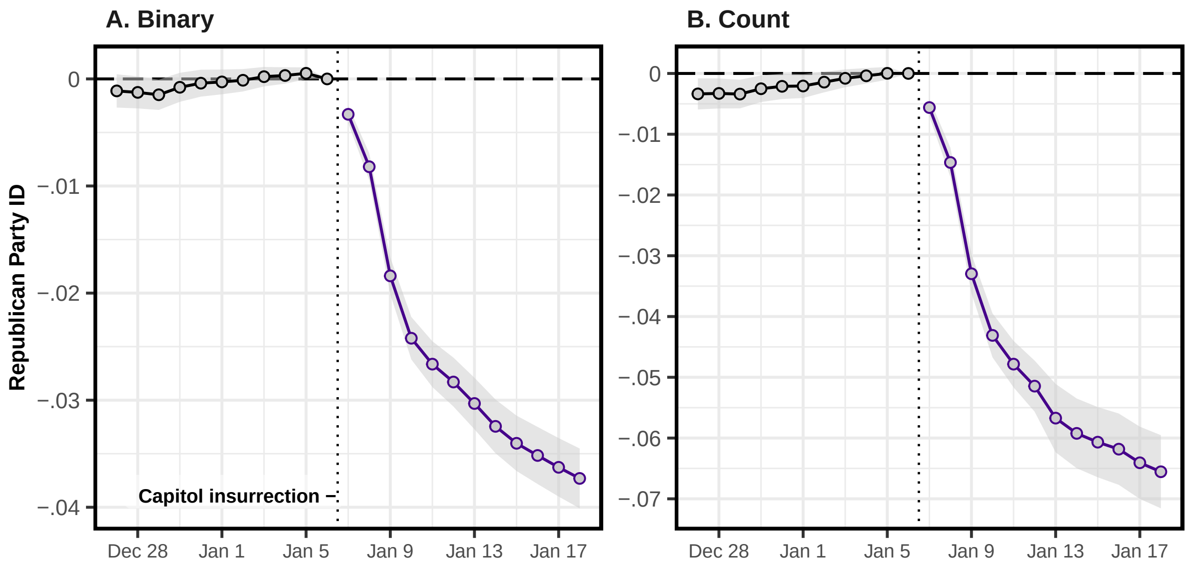


Figure K1: Event study results for binary and count outcomes excluding users are geo-located to the D.C. area (DC, MD, VA).

L Results among users who use real names as their user name

As explained in the main text, we tentatively address whether the observed de-identification is driven primarily by increased social costs of affiliating with the republican party (i.e. an act of preference falsification), as opposed to a weakened party identity by subsetting the event study models by whether a user name matches a first name in US Social Security Administration records. If increased social costs were the primary animating motive, we would expect users who use a real name—and therefore potentially bear higher costs—to be more likely to de-identify than those who use an alias. An important caveat to this test is that users may go by real name pseudonyms. As a consequence, the number of users who we identify as using a "real" name on Twitter is an upper bound on the number of users who truly use their real name.

More specifically, we first use regular expressions to isolate the first word in a user's name (absent, for example, leading emoji or other characters). Second, we match the first word in a user's name to the names database from the US Social Security Administration (birth names as recorded per year between 1920 and 2012). In total, 72% of users in the sample use a real first name as their user name.[7]

We then examine differences between users who use real names and those who use aliases by fitting event study models to each group of users, both for the binary outcome (any partisan term) and the count outcome (number of partisan terms) in users' profiles. Results are presented in Figure L1. As the figure shows, there are no meaningful differences in the effect of the Capitol insurrection on Republican de-identification among users whose user name can be matched to a real name in the Social Security Administration names database (Panel B), as compared to those users whose user names cannot (Panel B). If anything, the effect may be slightly larger among users who use aliases. This result suggests that the observed de-identification is at least partly driven by a weakening of identification with the Republican Party.

M Are those with Trump terms more likely to re-identify?

In the main article, we show in Figure 3 that very few social media users who de-identified in the week immediately following the Capitol insurrection went on to *re*-identify in the weeks afterward. Despite the small sample of re-identifiers, we seek here to investigate who re-identified. More specifically, we examine whether those with only the term "republican" in their bio before de-identifying were less likely to re-identify than more hard-core supporters using other terms before de-identifying (e.g. "Trump" or "MAGA"). To test this, we examine the subset of users who de-identified in the week following the insurrection. As an outcome variable, we code as 1 each user-day in which a user re-identifies, i.e. uses one or more

[7] This may appear a relatively high percentage, but is likely due to the fact that the sample includes only users who can be geo-located to the United States. Geo-location requires users to provide more personal information in general, for example, by allowing GPS-location of social media posts, or manually writing one's location (e.g. "Los Angeles, CA"). Those users who include more personal information such as this can be presumed to be more likely to also present themselves non-anonymously on Twitter more generally. For example, by contrast, 62% of users who cannot be geo-located use a real name.

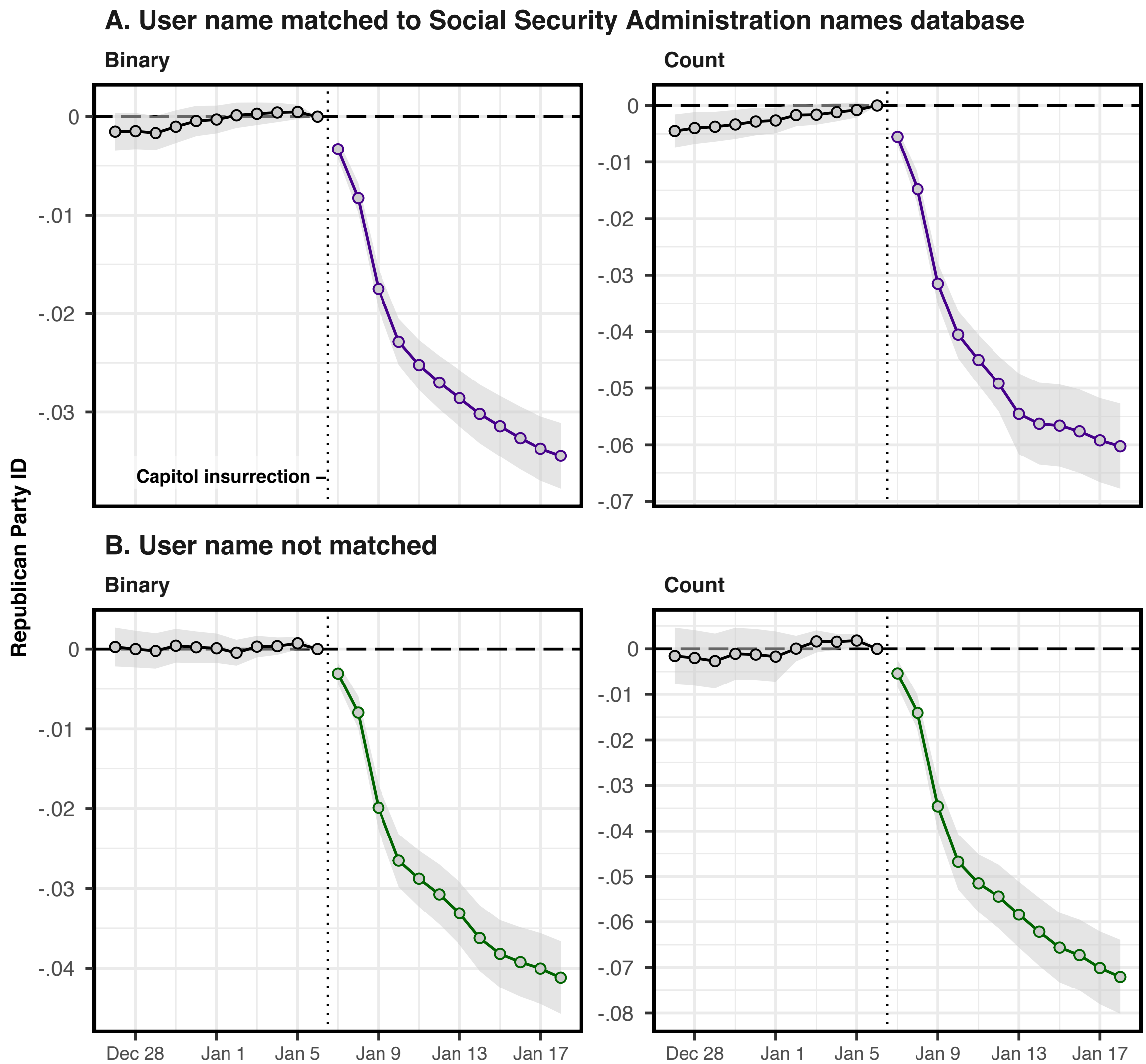


Figure L1: Event study results for binary and count outcomes among users whose user names have been matched to the Social Services Administration name list.

Republican terms their Twitter profiles, and 0 otherwise. We then regress this measure of re-identification on a dummy variable indicating whether a user had only a party-specific term. The coefficient on this variable tests whether users identifying with only party-specific terms (as opposed to terms like 'Trump' or 'MAGA') are more or less likely to re-identify.

Results are presented in Table M1. The results show that those with a party-only term in their bios were 1 percentage point less likely to re-identify in the two months thereafter, thus indicating that those who are presumably more moderate supporters, distance themselves

	Prob. of reidentifying
Intercept	0.05^{***}
	(0.00)
Trump & party term	0.00
	(0.03)
Party-only term	-0.01
	(0.01)
R^2	0.00
Adj. R^2	0.00
Num. obs.	137872
RMSE	0.21
N Clusters	2462

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table M1: Regression model predicting whether those using Trump terms in their bios before de-identifying are more likely to re-identify than those with Republican Party-only terms

more from the Republican Party after the insurrection. Note that this is a relatively large difference because on the average day in the post-insurrection time frame, only 4.6% of de-identifiers had re-identified. The difference, however, is not statistically significant, which is not unexpected given the small number of users overall who re-identify in the time period.

References

Cunningham, Scott. 2021. *Causal Inference: The Mixtape.* New Haven, CT: Yale University Press.

Dredze, Mark, Michael J. Paul, Shane Bergsma and Hieu Tran. 2013. "Carmen: A Twitter Geolocation System with Applications to Public Health." Association for the Advancement of Artificial Intelligence (AAAI) Workshop on Expanding the Boundaries of Health Informatics Using AI (HIAI).

Friedman, Jerome, Trevor Hastie and Rob Tibshirani. 2010. "Regularization Paths for Generalized Linear Models via Coordinate Descent." *Journal of Statistical Software* 33(1):1–22.

King, Gary, Patrick Lam and Margaret E Roberts. 2017. "Computer-Assisted Keyword and Document Set Discovery from Unstructured Text." *American Journal of Political Science* 61(4):971–988.

Mitts, Tamar. 2019. "From Isolation to Radicalization: Anti-Muslim Hostility and Support for ISIS in the West." *American Political Science Review* 113(1):1–22.

Singh, Kanishka. 2021. "Twitter suspends tens of thousands of accounts dedicated to sharing QAnon content." *Reuters*, January 12.