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.

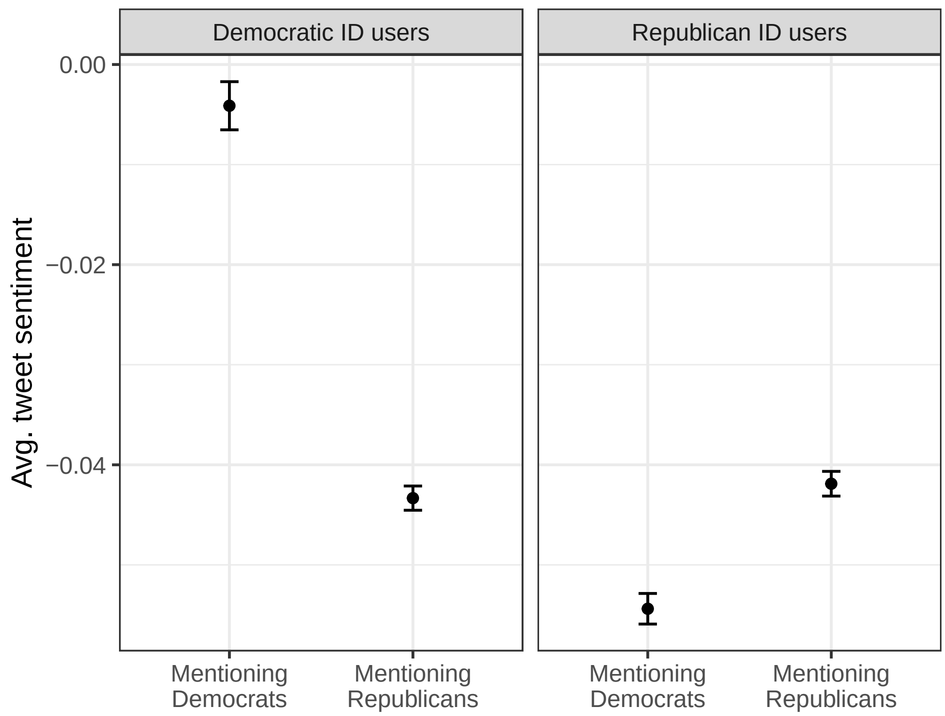

Figure B1: Average sentiment of tweets mentioning either `democrat*` or `republican*` for Democrat- and Republican-identifying users. Tweets referring to both parties are omitted.

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.

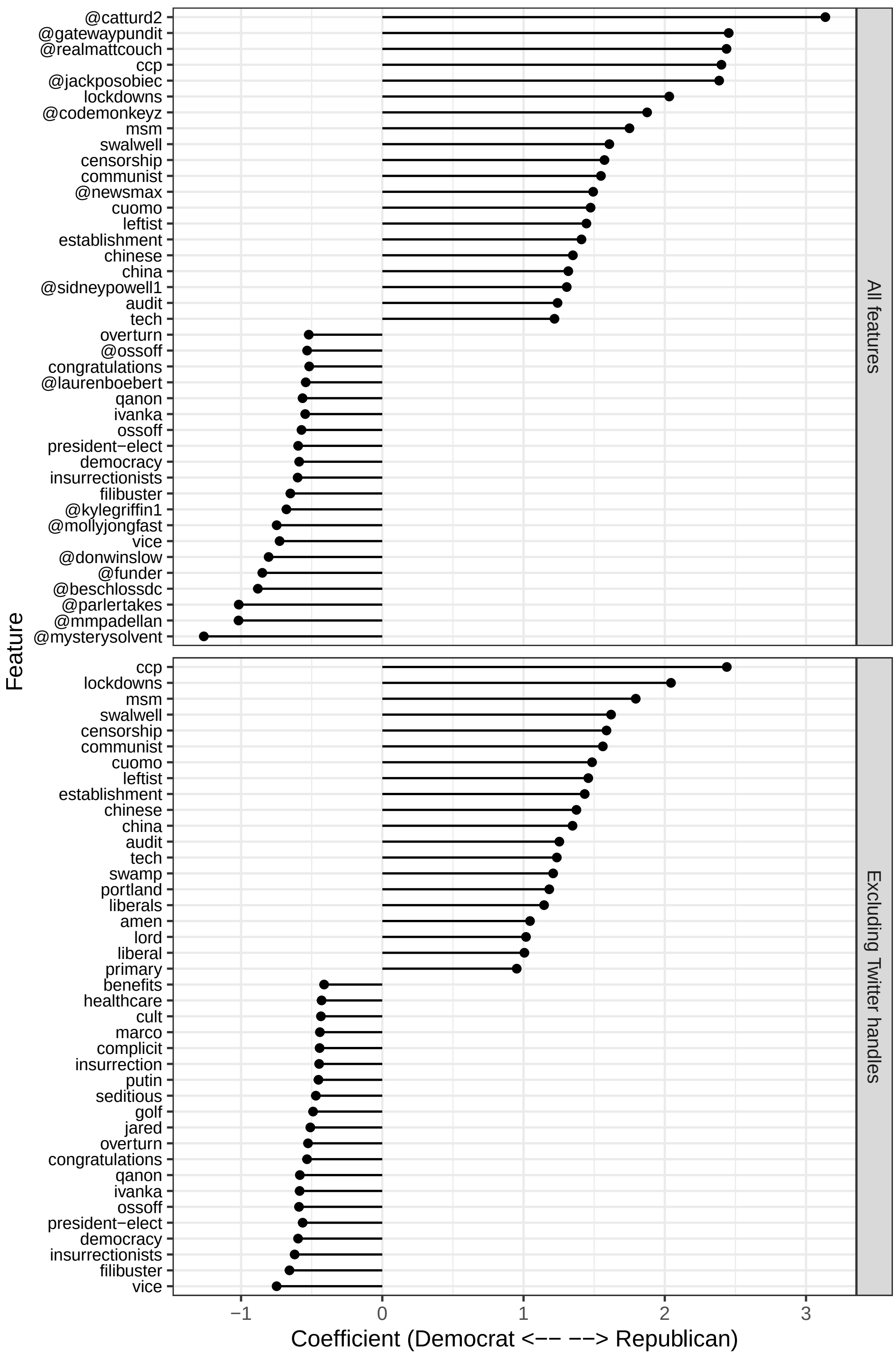

Figure B2: Top 20 terms in user tweets most predictive of identifying as a Republican (positive coefficients) or a Democrat (negative coefficients), based on a regularized regression model. Top panel shows results for all terms. Bottom panel shows results when omitting Twitter handles.

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.

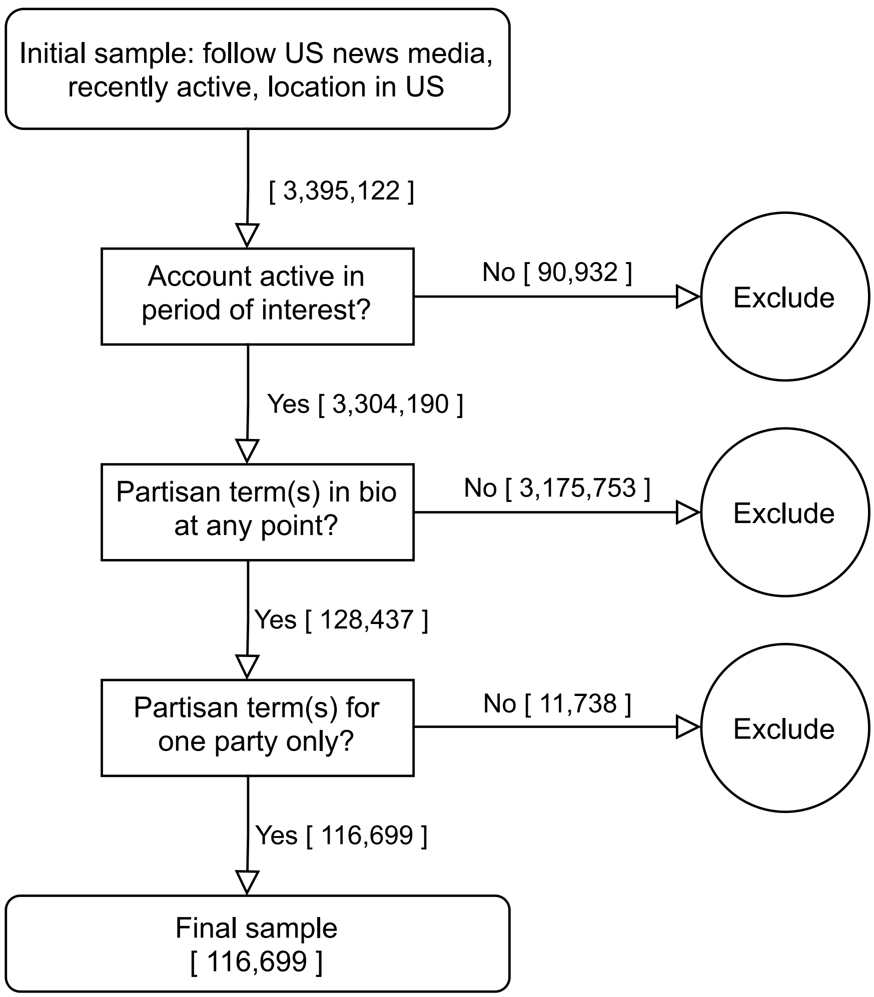

Figure C1: Flowchart outlining sample selection. Sample size at various stages of the selection process shown in brackets.

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

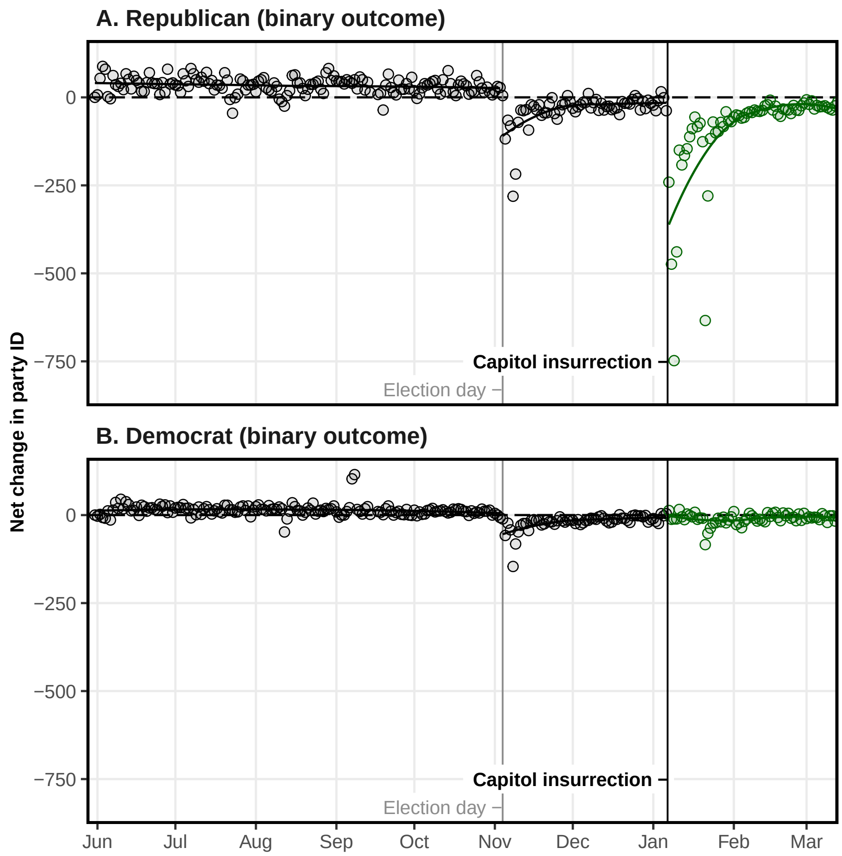

Figure D1: Change in Republican Party and Democratic Party identification over time (binary outcome).

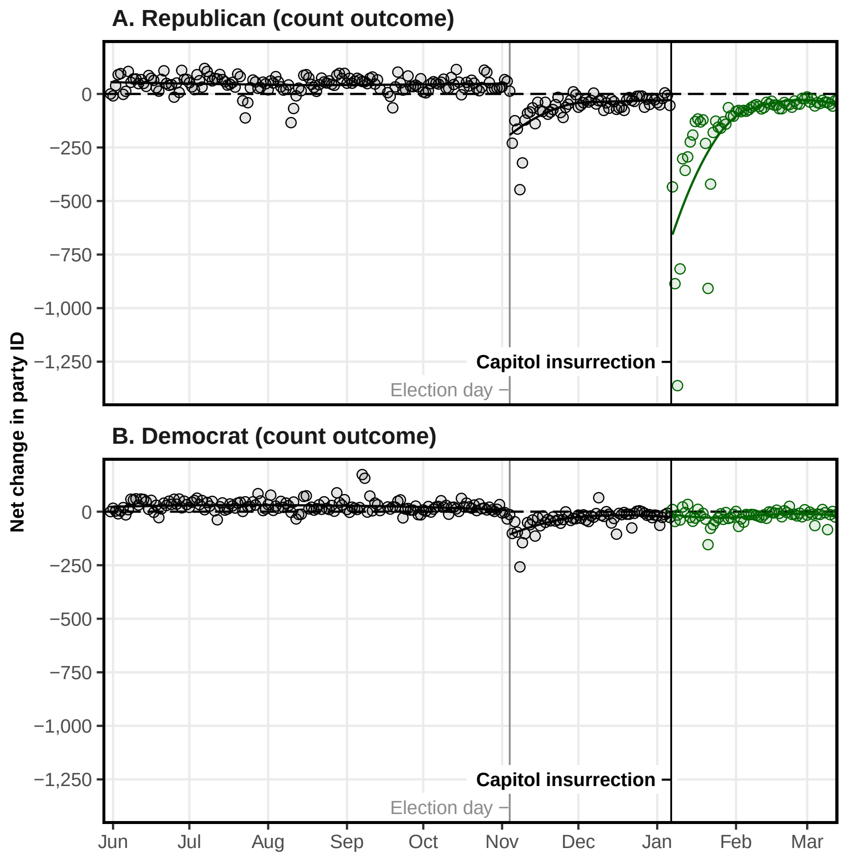

Figure D2: Change in Republican Party and Democratic Party identification over time (count outcome).

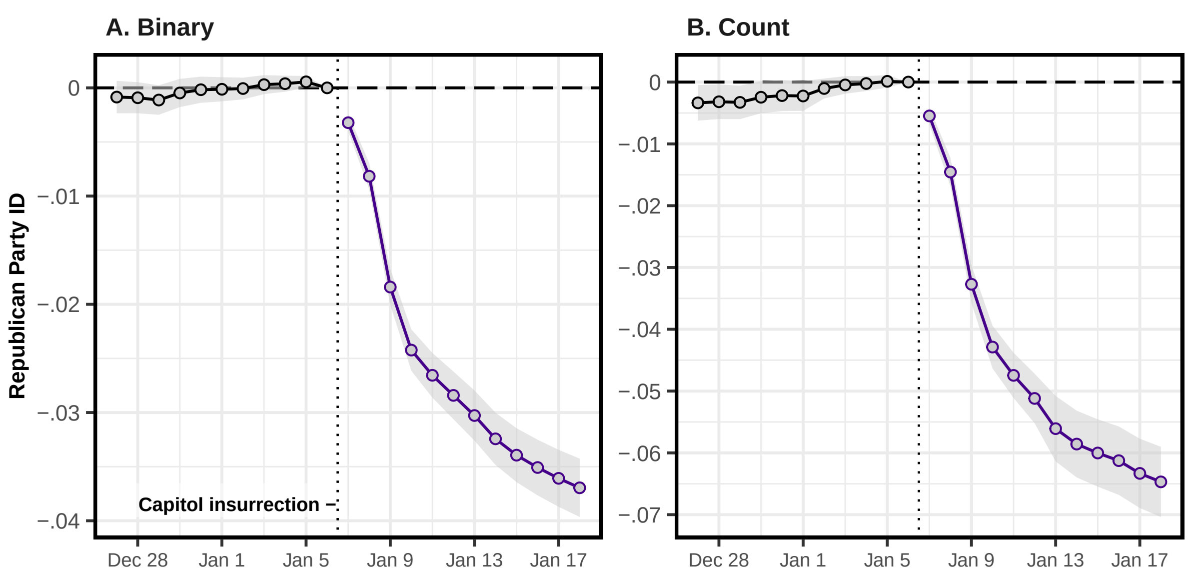

Figure E1: Primary event study results for both binary and count outcomes overview of how the two types of terms are distributed across the user bios in our data. For simplicity, we consider only the data from the first day of the time frame presented in Figure 2 in the main text, i.e. December 27$^{\text{th}}$.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

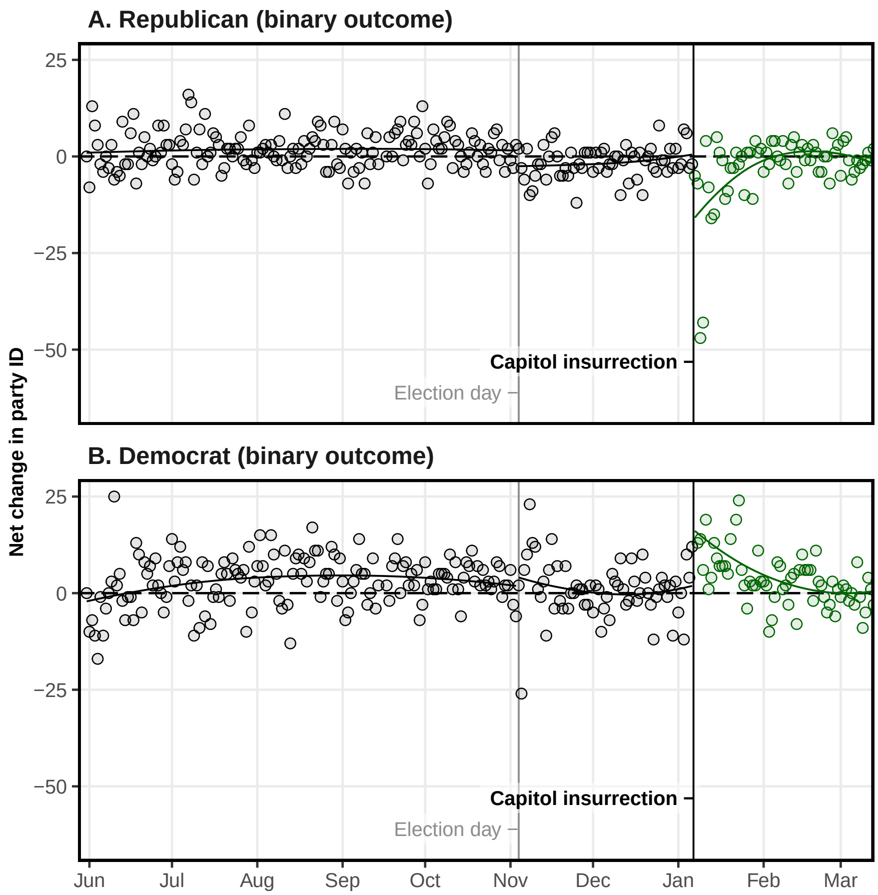

Figure F1: Net change in (binary) party identification over time ('republican' & 'democrat' only) findings regarding de-identification are robust to focusing exclusively on expressed party identification *per se* and not only movements associated with Donald Trump. Within a few days, Republican users 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.

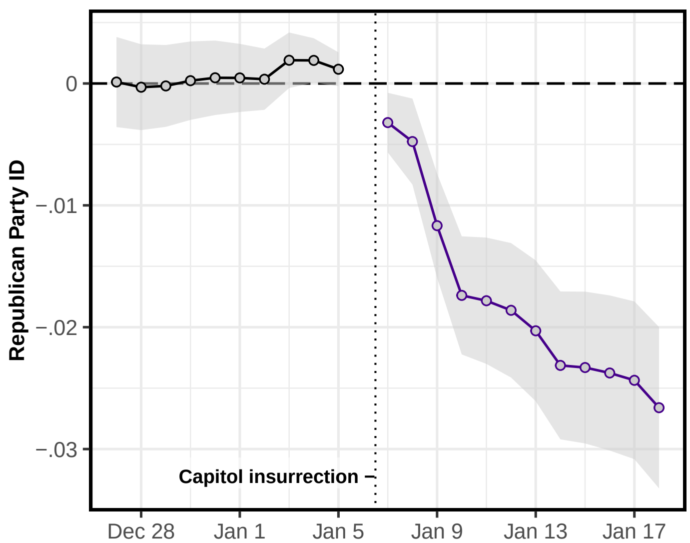

Figure F2: Event study models with identification defined only as "Democrat" or "Republican"

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.

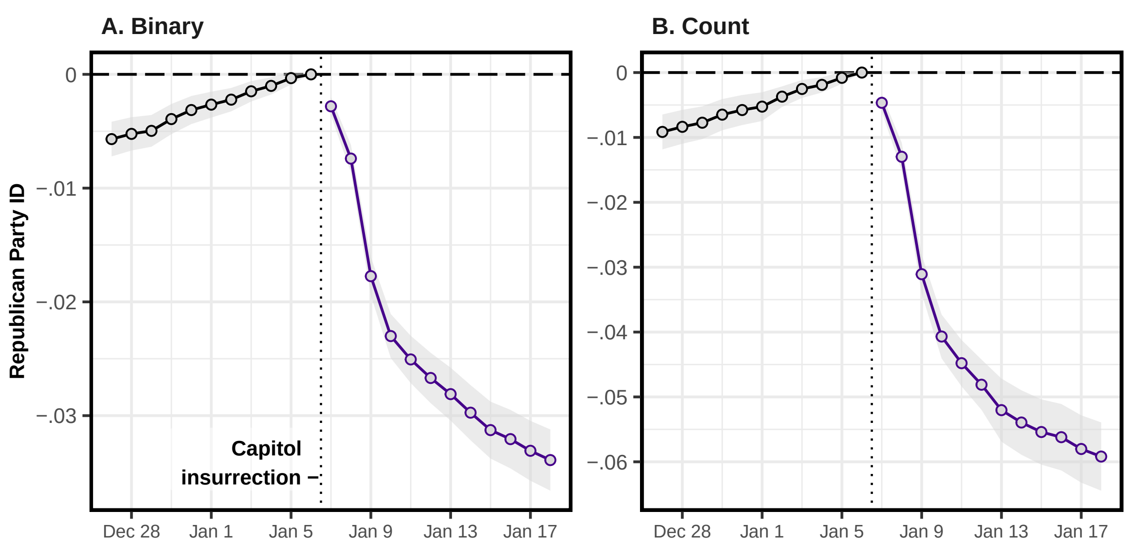

Figure H1: Event study estimates including the term `biden` in the Democratic ID keywords list

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.

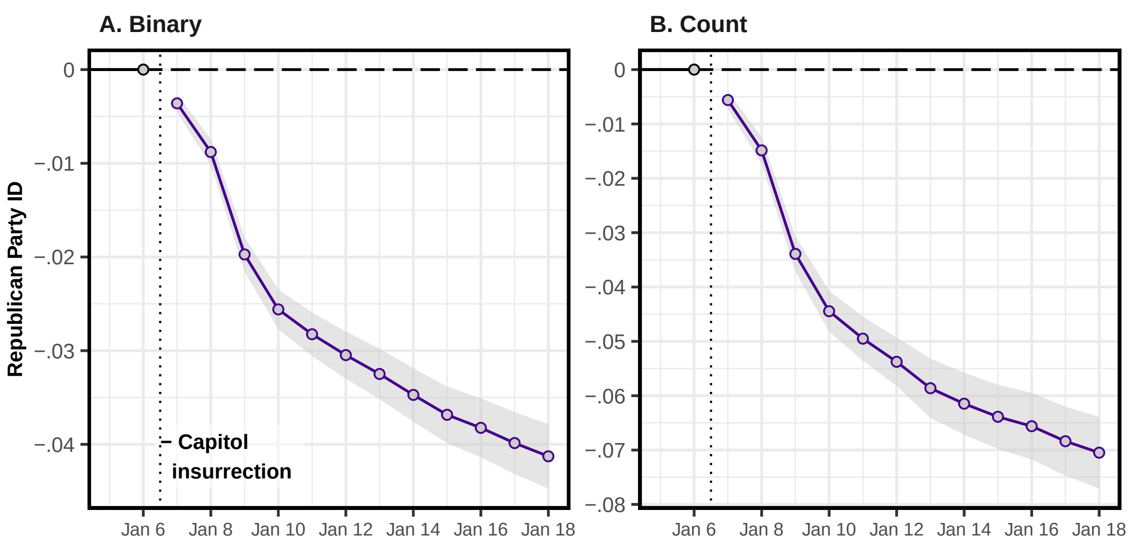

Figure H2: Event study estimates including the term `biden` in the Democratic ID keywords list (with modeled trend)

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

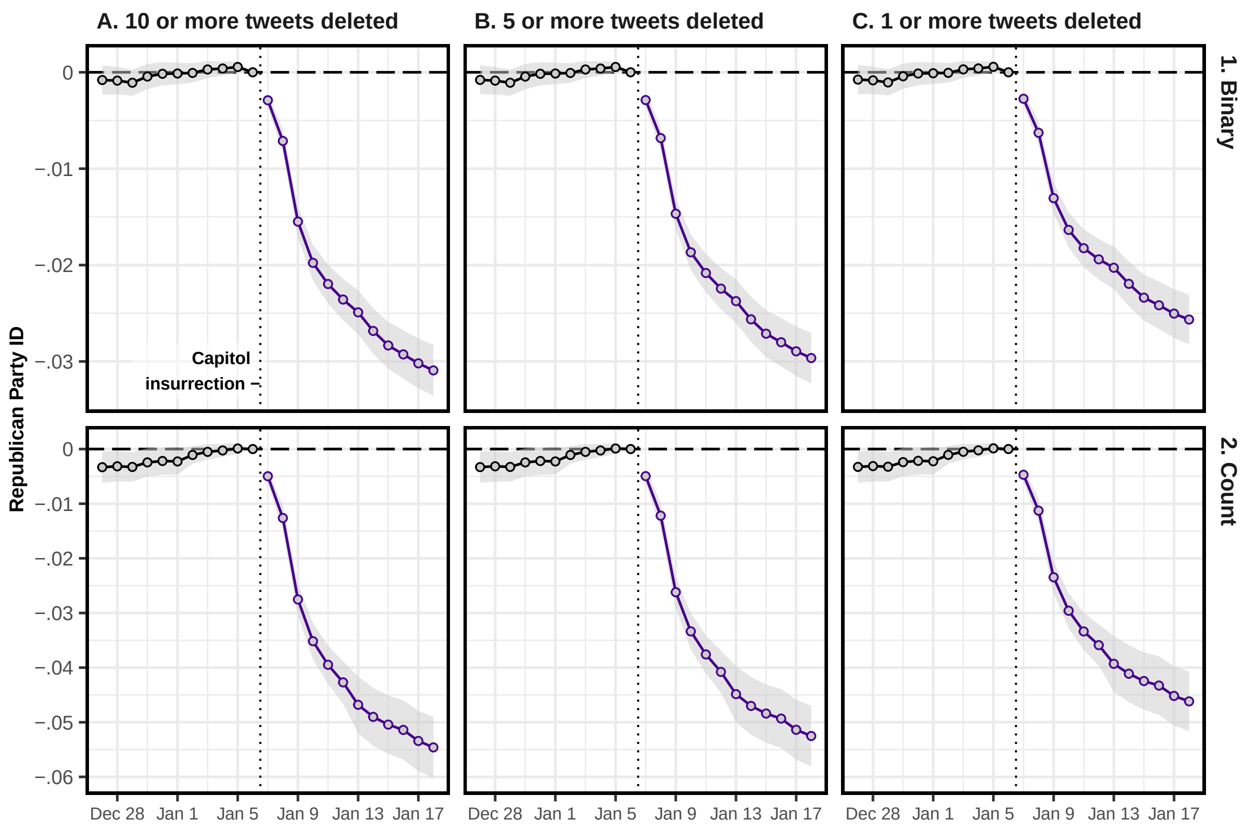

Figure I1: Event study estimates excluding users who deleted/scrubbed their timeline

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

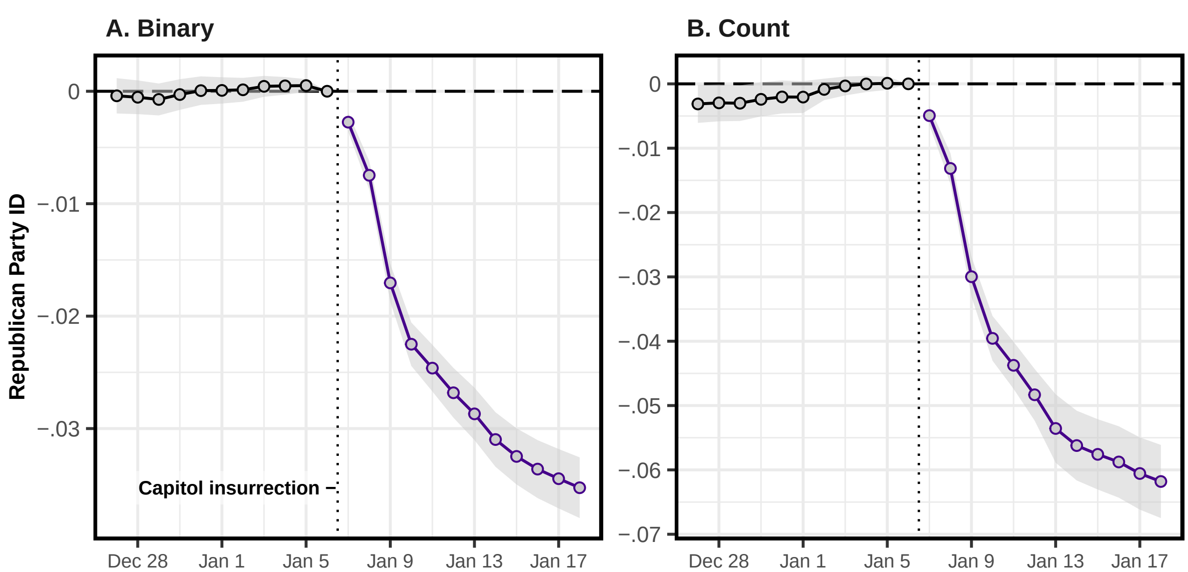

Figure J1: Event study estimates excluding users whose accounts were deleted by Twitter, were deleted by the users themselves, or were set to private whom we have complete data in the time period of interest, i.e. whose account was not deleted, suspended, or set to private. The results are presented 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.

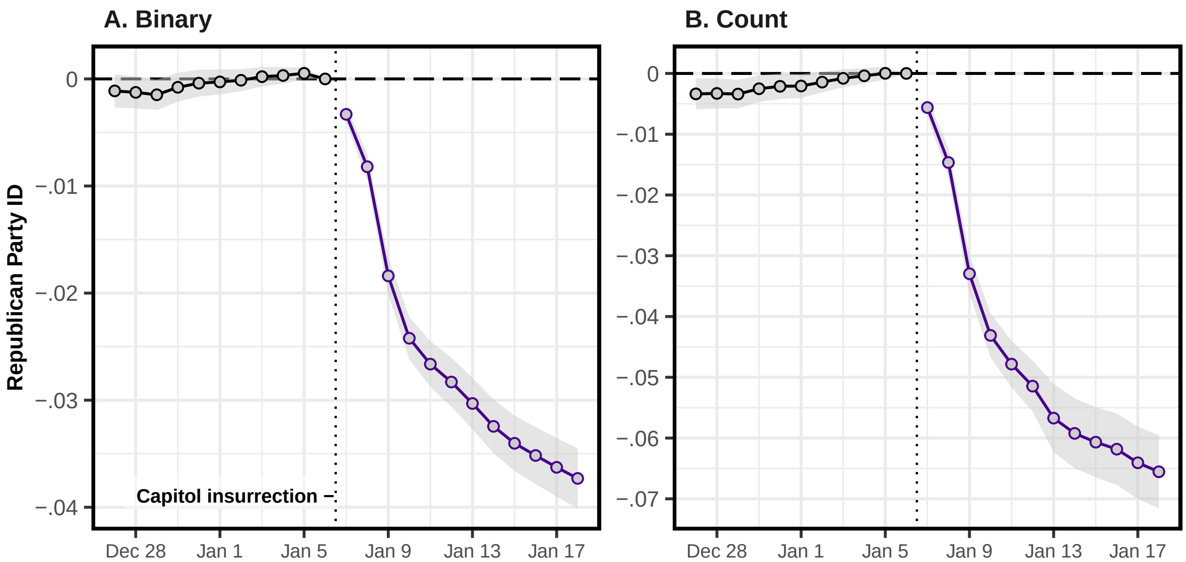

Figure K1: Event study results for binary and count outcomes excluding users are geo-located to the D.C. area (DC, MD, VA).

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.

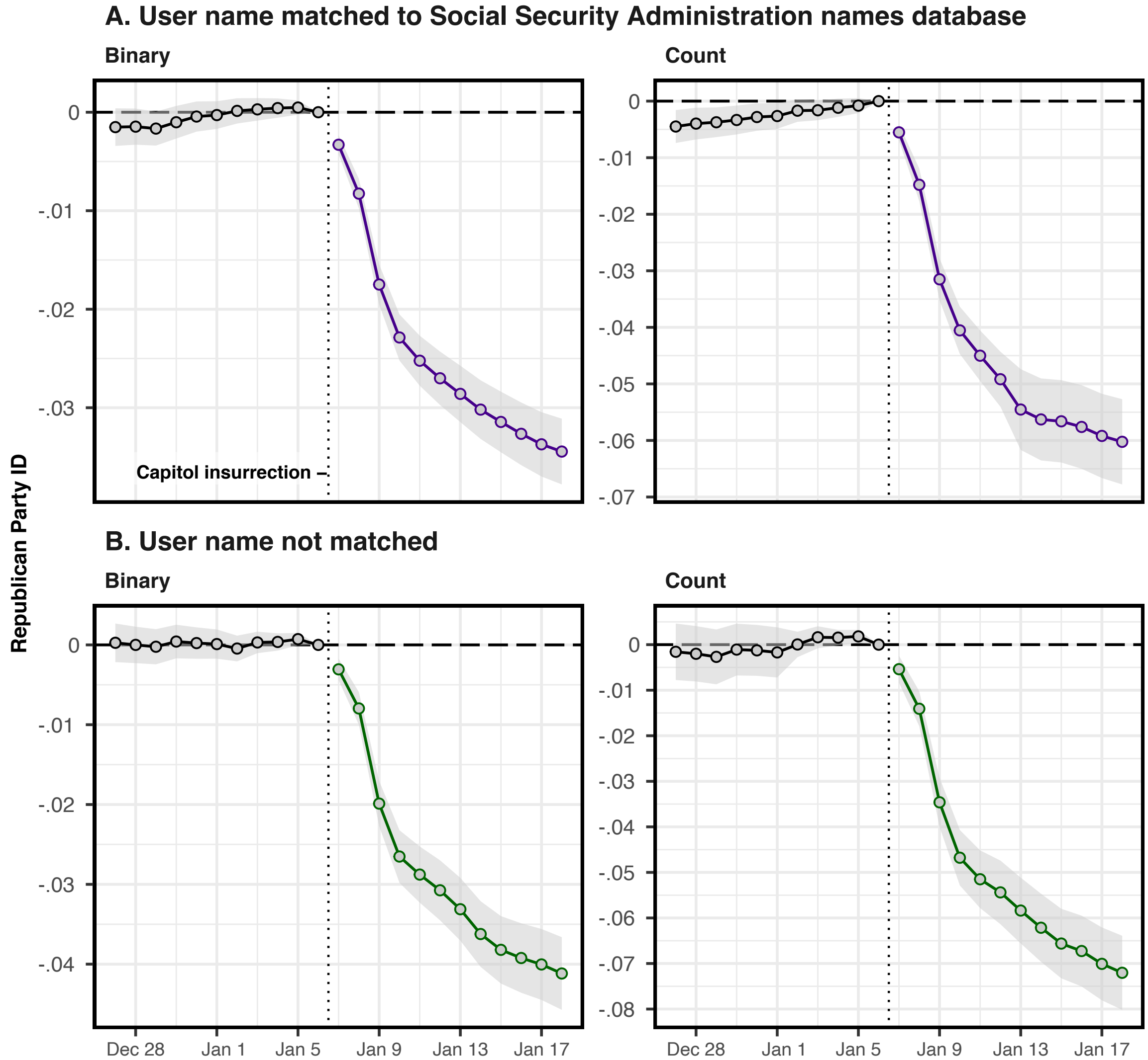

Figure L1: Event study results for binary and count outcomes among users whose user names have been matched to the Social Services Administration name list.

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.