Twitter and Affective Polarisation: Following Political Leaders in Spain

Javier Lorenzo-Rodriguez[1] & Mariano Torcal[2]

[1] *Universidad Carlos III de Madrid, Department of Social Sciences, Spain.*
ORCID: https://orcid.org/0000-0002-4473-8360
javier.lorenzo@uc3m.es
Twitter: @javilor

[2]*Universitat Pompeu Fabra, Department of Social Sciences, Spain*
ICREA Scholar, Director of RECSM Barcelona
ORCID: https://orcid.org/0000-0002-0060-1522
mariano.torcal@upf.edu
Twitter: @MarianoTorcal

Corresponding Author:

Javier Lorenzo-Rodríguez
javier.lorenzo@uc3m.es
c/Madrid 126, 28903 Getafe, Madrid. España.
Building: Carmen Martin Gaite
Office: 18.2D.09

Twitter and Affective Polarisation: Following Political Leaders in Spain

Abstract:

The present study addresses the effect of the discourse of elites on Twitter on citizens' affective polarisation through a quasi-experiment that was embedded in a survey panel. Participants were invited to follow one of the Twitter accounts of nine candidates from the main political parties during the 2019 European Parliament electoral campaign in Spain. Experiment compliance among participants was confirmed using web-tracking data (passive metre). The results show that exposure to candidates' Twitter accounts by self-selection does not increase affective polarisation. Although high levels of polarisation might contribute to building echo chambers, the polarising content contained in the partisan Twitter accounts has no effect on increasing affective polarisation, even among those who strongly identify with such parties. This finding confirms the so-called limited media effect hypothesis for social media.

Keywords: affective polarisation; social media; Twitter; survey experiment; passive metre

An extensive body of literature has primarily focused on highlighting the responsibility of social networking sites (SNSs) for the increased levels of 'affective politics'. Several studies maintain that social media tend to produce homogenous, self-reinforcing exposure to the most extreme radicals, thus over-representing levels of disagreement and activating social identities (Sunstein 2018). Furthermore, social media act as an instrument for increasing self-esteem contributing to such polarization (Bail 2021; Hutchens Hmieloski & Beam 2019; Tsfati & Nir 2017). However, scholars have increasingly disputed this argument by proposing that cross-cutting interactions on social media are more frequent than commonly believed (Barberá et al. 2015; Barnidge, Adèvol-Abreu & Gil de Zuñiga 2016), increasing the level of exposure to diverse news sources compared to other types of media (Barnidge 2016; Fletcher & Nielsen 2017) despite the presence of ranking algorithms (Bakshy, Messing & Adamic 2015). Nevertheless, said increased exposure to cross-cutting views might also be the reason underlying such polarising effects (Guess et al. 2021; Suhay, Bello-Pardo & Mauer 2018). Thus, despite a prolific amount of literature, scholars have yet to find clear, conclusive evidences regarding the direct consequences of online partisan social media (Barberá 2020, p. 35; Iyengar et al. 2019, p. 135).

Our study addresses this topic by conducting a quasi-experiment embedded in a survey panel in which participants were invited to follow the Twitter accounts of nine candidates from the main political parties during the 2019 European Parliament electoral campaign in Spain (Torcal et al . 2020). Following the lead of some innovative studies (e.g. Flaxman, Goel & Rao 2016; Gentzkow & Shapiro 2011), we employed a comprehensive data strategy that combined a survey experiment with web-tracking data to test the partisan SNS hypothesis in a non-US context. In this manner, individual data collected with a passive meter or behavioural tracker and some additional post-treatment check questions have allowed us to confirm the level of exposure and the degree of activism of each respondent on Twitter. In addition, we measured the content of these Twitter candidates' accounts, which allowed us to detect the different levels of anti-partisan messages as well as the polarising anti-national and subnational identity messages to which each respondent had been exposed.

This research design provides us with additional sources of information that are not typically available, which constitute one of the main challenges of investigating social media and democracy (Persily & Tucker 2020, pp. 314-20). In addition, this type of quasi-experiment has substantial theoretical external validity with respect to the distinctive effects of social media on individuals, as social media users can be passive or active participants and can choose

their preferred social media channels. Finally, we focus on the effect of message content unlike the majority of the aforementioned studies.

The examination of this phenomenon in Spain is well suited for several reasons. Spain presents high and increasing levels of affective polarisation (Gidron, Adams, & Horne 2020). Preceding studies in Spain (Cardenal et al. 2019, pp. 371-2) have discarded that this increase is linked to the presence of communication echo-chambers because of the tendency of citizens of this country to take advantage of its high levels of mass partisan media diversity. However, there has been a strong increase of partisan Twitter in Spain, being one of the top 10 countries worldwide (35 per cent) in terms of using this type of social media (Newman et al. 2020). In recent years all political parties has used SNS as an important mechanism of communication, which has been especially remarkable with the rising populist and radical right (Vox) and left (Podemos—We Can, more recently renamed as Unidas Podemos—United we can—UP). These type of parties, with increasing presence in Spain, are known in contemporary democracies for using SNS to attract attention and create distractions from mainstream party coverage via emotional and polarising messages (Enli & Rosenberg 2018; Haselmayer, Wagner & Meyer 2017). Finally, Spain presents two different polarising conflicts: the partisan one and the one emerging by the centre-periphery conflict which is based on the activation of different territorial/national identities. This setting gives us the opportunity to test the effect of social media on individual polarisation, not only in a country with increasing partisan polarisation, but also having an additional one which intensity has called the attention of all international media since 2014.

In the following sections, we argue that partisan Twitter exerts no effect on affective polarisation, even when individual self-selection based on existing political preferences produces a homogenous network of followers. This notion is seemingly true regardless of the polarising content of the majority of radical leaders to which citizens are exposed or how active individuals are on Twitter. As will be discussed in the concluding section, these findings have important empirical and theoretical implications for the comparative literature on this topic.

Social media, selective exposure and affective polarisation

SNSs such as Facebook, Twitter and—more recently—WhatsApp have been increasingly consolidated as a mechanism of political communication, especially among the youth (Fletcher & Jenkins 2019; Newman et al. 2020). Therefore, the possibility of direct interactions with representatives, parties, celebrities and influencers (Barberá et al. 2019) and discussions about

politics between peers, friends, relatives and media outlets (Hampton, Shin & Lu 2017) has reduced the classic gatekeeping role of mainstream media, facilitating, at the same time, the access to various informational sources (Chadwick 2013).

In this sense, most of the literature has identified SNSs as a political arena that triggers affective polarisation (Gerstlé & Nai 2019; Zollo et al. 2015), praises anger and anxiety (Rudolph, Gangl & Stevens 2000), posits illiberal opinions against democratic institutions and principles (Bail 2021), supports personal attacks on political opponents based on false information (Bradshaw et al. 2020; Vosoughi, Roy & Aral 2018) and enhances radical discourse, such as incivility or hate speech (Theocharis et al. 2016; Tucker et al. 2017).

Although there are some contradictory findings (Boxell, Gentzkow & Shapiro 2017; Johnson, Kaye, & Lee 2017), the evidence of the effects of SNSs on polarisation during electoral campaigns is abundant and concrete (Banda & Cluverius 2018; Hernández, Anduiza & Rico 2021; Iyengar et al. 2019; Suhay, Bello-Pardo & Maurer 2018). However, recent comparative literature has challenged the overall impact of social media on affective polarisation (Gil de Zúñiga, Barnidge & Diehl 2018; Barnidge et al. 2018; Yang et al. 2016). In the context of South European countries, although homophilic behaviour appears to be the modal outcome (Agathangelou et al. 2017; Cardenal et al. 2019; Giglietto et al. 2019; Marozzo & Bessi 2018), that homogeneity in individuals' networks does not seem to have the same increasing polarising effect (Bright 2018; Fletcher & Jenkins 2019; Urman 2020). Recent research in multi-party systems scenarios shows high polarisation among those supporting extreme or radical parties (Agathangelou et al. 2017) but more diffused among those ideologically alike or with moderated preferences (Bright 2018; Aragón et al. 2013; Vaccari & Valeriani 2015).

Overall, the aforementioned contradictory and inconclusive results regarding the effects of SNSs on affective polarisation could be due to the fact that most of this literature tends to overlook the motivation underlying the need to seek and consume pro- and counter-attitudinal information (Valentino et al. 2008) and—more concretely— the motivation resulting from the identification with a specific partisan group. This factor is certainly responsible for creating 'communities of like-minded discussants' (Barberá 2020, p. 37) and the 'partisan selective exposure' it generates; however, it is questionable whether all of this results in a 'breeding ground for extremism' (Sunstein 2018, p. 71), as it is the case in Spain, for instance, fosters greater polarisation at the individual level. In that sense, leaving aside the problem of the 'filter bubble', the effects of social media on individual polarisation might also

be highly conditioned by the same factors that result in self-exposure to 'communities of like-minded discussants' or by the social ties and sense of belonging that citizens share with certain groups that align with or challenge their existing beliefs' (Bail 2021; Bankert et al. 2017; Iyengar, Sood & Lelkes 2012; Mason 2016). This way, most individuals following liked-minded discussant or congenial forums on SNSs (Bakshy, Messing & Adamic 2015; Vaccari et al. 2016), as the 'motivated reasoning theory' presumes (Enders & Smallpage 2019; Parsons 2010), might only maintain their existing beliefs instead of refining (polarising) their political opinions (Heatherly, Lu & Lee 2017). This notion is particularly true for individuals with high levels of motivation for seeking more information (Cardenal et al. 2019; Prior 2013), which, thus, increases the need for validation and activates defence or partisan reinforcement motivation (Barberá 2020; Guess 2016).

In other words, despite the presence of some (inconclusive) evidence on the link between congenial media consumption and polarisation (Mutz 2006; Stroud 2007; Stroud 2010), partisan selection exposure in social media might have a more limited effect. This is because this type of media consumption implies a more active partisan role in selection exposure and therefore limits the effect of such exposure on polarisation, especially for the most radical or extremism partisan options (Bail 2021). This means that partisan in-group messages that are contained in social media—that is, messages from party candidates whose platforms align with an individual's preferences—might not influence affective polarisation, even among a minority of highly active and visible partisan individuals who might become more polarised as a consequence of a greater level of activism on Twitter (Barberá & Rivero, 2015; Shore, Baek & Dellarocas 2016; Yarchi, Baden & Kligler-Vilenchik 2020).

Finally, the prior inconclusive results on this topic may be due to the fact that the majority of the existing literature, with few significant exceptions (Theocharis et al. 2020), lacked consistent and significant efforts to consider the content of social media, which may be responsible for such polarisation. Instead of questioning the responsibility of social media, researchers should investigate how some activists use and abuse particular 'discourse practices' (Glapka 2019). Such tendencies contribute to the creation of 'relations of proximity, distance, affiliation and detachment, and inclusion and exclusion' (Wetherell et al. 2015, p. 58) with certain partisan social groups. In other words, we are far from identifying and measuring discourse practices that contain a visceral appeal that results in the framing of partisan social identities as the *we* who face threats from *them* (externally), which overrides any other rational considerations. This notion is characteristic of all polarising discourse practices (Breeze 2018,

p. 26). As noted out by Trepte and Loy (2017) in their study on social identity theory and social media, future research should 'take a more granular look at how exactly group memberships and media use are interrelated' (pp. 11–12).

Twitter and political elites

Finally, although scholars have argued that elite discourses produce affective politics (Lelkes 2021; Rogowski & Sutherland 2016), research on the conceptualisation and measurement of the type of content that produces such a polarisation is scarce (Cardenal et al. 2019, p. 372). Parties and elites not only share ideas and manners of speaking about them (i.e. discourses) on social media but also promote particular ways of feeling about different partisan identities and voters (Wetherell et al. 2015). Politicians who can embody and express feelings that resonate with large sectors of the electorate or who know how to influence voters at the affective level are typically highly successful, particularly in this age of social media (Frame & Brachotte 2016). The 'affective–discursive practices' of a party's spokespersons are likely to reflect the practices of its supporters and vice versa such that one reinforces the other, leading to escalating effects (Breeze 2018).

This discussion is essential for contemporary democracies, given that political elites, parties and candidates tend to behave strategically on SNSs, particularly during electoral campaigns (Bimber 2014; Cardenal 2011; Gibson 2015; Kruikemeier 2014; Lilleker Tenscher & Štětka 2015). Several scholars have observed heterogeneous behaviour among incumbent and challenging candidates or mainstream parties, concluding that those who are lesser known or have less media saliency tend to level the field of competition by attracting attention to their social media profiles (Boulianne, Koc-Michalska & Bimber 2020; Kahne & Bowyer 2018; Stier et al. 2018). In addition, populist and radical right parties are more likely to employ negative messages with a higher emotional charge, thereby increasing affective polarisation through SNSs (Haselmayer, Wagner & Meyer 2017). This seems especially true among individuals who are increasingly interested or mobilised. Thus, in order to assess the effects of SNSs on affective polarisation, we propose in this study the following hypotheses.

H_1: Being exposed to partisan out-group polarising messages on leaders' Twitter accounts increases affective polarisation.

H_2: Selecting leaders' Twitter accounts in accordance with one's party preferences increases affective polarisation.

H_3: Being more active on Twitter leads to an individual being more affected by elite polarising discourses (anti-partisan and anti-territorial out-group) identities, thereby increasing their affective polarisation.

The survey experiment design

In this study, we applied an embedded 'between-subjects-design' experiment in a general population survey that was distributed to participants who were recruited through an online panel managed by Netquest. (For details regarding the sample and additional technical information see online appendix section A.) The fieldwork for the survey was launched during the campaign for the European Parliament elections that took place in May 2019. The respondents were grouped as follows: (a) a control group that indicated their refusal to be exposed to some Twitter accounts (because, in many cases, because they were not Twitter users) at the beginning of the questionnaire; b) a second group that followed some Twitter accounts for at least three days within a total time period of seven days (17–24 May)[1]. Both group of respondents were invited to complete the same online questionnaire. Post-treatment checks identified participants who initially agreed to participate but failed to comply with the experiment, being assigned to the first group.

In addition, we captured the Internet activity of each respondent 15 days prior to the Election Day using a passive meter or behavioural tracker that was installed by each panellist in our sample after accepting the invitation to do so. This technology allows us to capture and measure each respondent's activity in internet. The information collected by this technology allows us to create two additional variables for each respondent. The first variable represents the total time (in seconds) that each respondent spent on Twitter during the 15 days before the election, which also encompasses the days of the fieldwork and the 7-day period that participants were granted to follow a Twitter account for the experiment. The second variable measures the daily average in seconds that each participant spent on Twitter but only during the days when they were invited to participate. These variables measure respondents' time spend on Twitter, constituting a good proxy to estimate the real compliance with the experiment and exposure to the treatment. Table 1 displays this data (time) for all respondents. Those who did not accept our invitation to participate in the experiment show considerably little time exposure to any Twitter account on average. In addition, the differences in the time spent on Twitter between those who complied with the experiment and those who did not are remarkable (almost 47 per cent more on total time and 24 per cent more on daily average, respectively).

(Table 1)

Evidently, this technology poses several methodological limitations in capturing the activities of all the respondents in Twitter. This instrument captures Internet activity only as the participants use a browser and fails to collect activities on the mobile app (Bosch & Revilla 2021). Certainly, it is true that most of the Twitter activity among our respondents using mobiles took place by the use of the Twitter App. This limitation may lead to the recording by the passive meter of missing or zero activity for respondents using the Twitter App in their mobiles. It is certain that most of activity in this social media is mostly via mobile devices, but 30% of their activity is via web browsers, and this is normally distributed across all respondents, indicating that access, use and, therefore, exposure occurs interchangeably across different access devices. Thus, we can affirm that the subsample of Twitter users captured by the passive meter constitutes a good estimation of respondents' time exposure for the entire sample. 79.1 per cent of the respondents who declared in the survey having an account on Twitter[2] recorded by the passive meter a total time spent on Twitter of 1,485 minutes on average during the fieldwork. Respondents answering the same in this question but who also declared compliance with the experiment in the post-treatment checks recorded a total of 1,588 minutes on average (see Table 1). In general, as it has been discussed in recent studies employing this instrument in both, web browsers and mobile applications have not identified different patterns (Guess 2021; Munzert et al. 2021).

Finally, participants in the experiment were randomly assigned to follow one of the two types of Twitter accounts. One group was assigned to a Twitter account that belonged to the main candidates of five national parties and four regional/nationalist parties (coalitions) running for the European Parliament. Then, respondents were asked to select one of those candidates' accounts. As a placebo, the other group was randomly assigned to the institutional accounts that shared content and information on the European Union, constituting the baseline to assess the average treatment effects of the treated (ATTs), and then asked to follow one of them. The list of Twitter accounts and the number of observations for each group and subgroup of treatments are provided in Table 2.

(Table 2)

Twitter account content

The Twitter activity of political candidate and EU institutional accounts was collected through the Twitter Representational State Transfer Application Programming Interface (REST API). The results presented herein reflect the activity of the candidate and institutional accounts from 17 to 24 May 2019, accounting for the last week of the campaign. In total 2,211 tweets and retweets (RTs) were posted by politicians and institutions (1,774 and 437, respectively). New and lesser-known candidates, such as the leaders of Partido Nacionalista Vasco (PNV; Basque Nationalist Party) and a left-wing regionalist party established in the region of Valencia, Compromis (Commitment Coalition), tweeted or retweeted 45.3 and 97.7 times per day, respectively (most of them RTs), while the leader of the left-wing pro-independence Catalan party Esquerra Republicana de Catalunya (ERC; Republican Left of Catalonia) was the least active, with 4.7 tweets or retweets per day.

Territorial identities

To classify tweets expressing pro- or anti-Catalan or Spanish sentiments, we designed four binary variables, thereby identifying a set of unambiguous expressions regularly used among elites, media outlets and public opinion for each sentiment[3]. Because these types of sentiments are not defined in Spanish, Catalan or Basque dictionaries and are often expressions rather than single words, the 2,211 tweets were manually coded.

(Table 3)

As Table 3 shows, considering the total tweets, the leaders expressing the greatest proportion of anti-Catalan sentiments belong to the national conservative party Partido Popular (PP; People's Party), the right-wing radical party Vox and, to a lesser degree, the moderate right-wing party Ciudadanos-Citizens (Cs). These accounts also share the largest percentage of pro-Spanish expressions. By contrast, the leaders with the highest percentage of anti-Spanish expressions and, simultaneously, the ones sharing the most pro-Catalan sentiments are Junqueras, the ERC candidate, and Puigdemont, the Junts per Catalunya (JxCat; Together for Catalonia) candidate and former president of the Catalan government who declared independence in 2017. Therefore, if H_1 is correct, we anticipate that the participants with the highest levels of ATTs belong to the group exposed to the political figures' Twitter accounts

containing the highest percentage of anti-Catalan or anti-Spanish messages, especially for those who match their selection with their party identification (PID) (H_2) and those who are specially active on that particular Twitter account (H_3).

Anti-partisan identities

Following the same method, we created a binary variable accounting for anti-partisan expressions, codifying any explicitly or implicitly negative mention of another political party or its members[4] as 1. The first two columns in Table 4 contain the total number of anti-partisan sentiments and their percentage of the total tweets. The next four columns provide the percentages of the tweets with anti-partisan messages. The political figures whose accounts have the highest percentages of anti-partisan messages are ERC, JxCat, PP, Vox and Cs. Given this information, we should expect that the participants exposed to the Twitter accounts of the ERC, JxCat, PP, Vox and Cs candidates would show greater levels of partisan affective polarisation (H_1), especially for those who match their selection with their PID (H_2) and those who are especially active on that particular Twitter account (H_3).

(Table 4)

In summary, the Twitter activities of the parties indicate a strategic-oriented behaviour, which confirms negative campaign theories on gaining media attention (Haselmayer, Wagner & Meyer 2019; Maier 2020). Furthermore, the results indicate a considerable inclination to post anti-Catalan, -Spanish or -partisan messages. Moreover, ERC, JxCat, PP, Vox and Cs stress their positions by posting against others in terms of partisan and territorial identities. Within the ideological dispute, those in the right-wing spectrum provide clear evidence of the fight to differentiate themselves among others in a more radical fashion compared to left-wing parties.

The outcome variable: polarisation

To assess the effect of this experiment, we considered two outcome variables. The first is a measure of partisan affective polarisation, which was proposed by Wagner (2020), and used feeling scales for voters' feelings as a more reliable measure of affective polarisation (Kingzette 2021; for further detail, see Section B of the Appendix) of the five national political

parties and the three most important Catalan and Basque regional parties, namely, Spanish Socialist Workers Party (Partido Socialista Obrero Español [PSOE]), PP, Cs, United We Can (Unidas Podemos), Vox, ERC, JxCat and PNV.

In addition, the E-DEM dataset contains feelings and trust scales towards different territorial groups in Spain, such as the Catalans, Basques, Madrileans and Andalusians. The last two were used together as a reference for Spanish identifiers. Moreover, we rescaled and summed up the two items (feelings and trust) and divided the resulting scale by 2, which ranges from 0 to 10. On the basis of these scales, we formulated an index that captures the extent to which an individual (on average) likes/dislikes other territorial groups compared to the individual's own group (i.e. the mean distance in sentiments from the territorial group of the respondents). This index ranges from −10 to 10. In this case, positive values indicate that the respondents hold more positive sentiments towards their in-territorial group than towards their out-territorial groups. The opposite is true for negative values. Finally, 0 indicates that the respondents hold the same sentiments towards the in- and out-territorial groups (for more details on these indices, see Section B of the Appendix).

¿Polarising effects among the exposed?

Territorial affective polarisation

What is the effect of being exposed to different Twitter accounts? In Figures 1 and 2, we display the ATTs for the two territorial polarisation indices that we used for those participants who successfully completed the experiment. In the upper part of these figures, we represent the ATTs of following each political figure's Twitter account, with the baseline being the polarisation average of the participants who were randomly assigned to choose one of the four EU institutional accounts (the average of the placebo condition). At the bottom of these figures are the ATTs of those following institutional Twitter accounts, taking the average of all respondents randomly assigned to condition 1 (political figures' accounts) as the baseline for comparison.

In Figure 1, the participants with higher levels of ATTs for territorial affective polarisation (measured by the territorial spread sentiments scores) were the respondents following political figures' accounts of the following political parties: PNV (2.04), JxCat (1.73), PP (1.71) and Vox (1.61). However, none of these ATTs are statistically significant at $p = 0.05$. For the rest, the ATTs are almost null, including for the followers of the pro-

independence party ERC. For PSOE (1.04) and UP (0.88), the ATTs are even lower than those of the placebo participants as these accounts did not share messages of this theme during the period of study. Finally, the bottom of Figure 1 shows that the ATTs of the institutional Twitter accounts' condition (the placebo treatment) are null for all of them.

(Figure 1)

Figure 2 presents the same results but with the indicator of mean group distance from Catalans as the outcome variable. This figure confirms the same initial conclusions reached by observing the results displayed in Figure 1.

(Figure 2)

Partisan affective polarisation

Figure 3, which presents the ATTs for out-group partisan polarisation, seems to confirm that leaders' out-group anti-party Twitter content did not have a consistent impact on participants. The ATTs are a little larger for the JxCat (8.24), PNV (8.09) and ERC (7.91) accounts. In addition, the difference in ATTs between the political figures' Twitter followers and the placebo are not significant. These results seem to question the importance of exposure to anti-partisan content via political figures' Twitter accounts on partisan affective polarisation.

(Figure 3)

Affective polarisation among party identifiers

The type of design presented herein reinforces the ecological validity of the experiment as it more realistically reflects the effects of political communication (Bennett & Iyengar 2008, p. 724), particularly Twitter use, on polarisation. However, it raises some important methodological issues. The acceptance of the participants to participate in the experiment, comply with it and expose themselves to different Twitter accounts depends on many factors that are endogenous to such decisions.

Thus, the ATTs reported earlier could be the product of other endogenous factors and are evidently not equal to the gold standard required to consider them equal to the sampling average treatment effects (SATEs) (Leeper 2017, p. 27). In addition, the effects of selection exposure to an account and its messages differ when respondents can self-select a Twitter account to follow compared to the effects for those who would and would not choose that particular account, resulting in the presence of treatment heterogeneous effects (Leeper 2017, p. 23). In our case, the ATTs are evidently related to some ideological and partisan preferences (partisan selection exposure). As Table 5 shows, the participants' choice of following a political figure's Twitter account is greatly informed by their ideological preferences, territorial identities and, obviously, party identification—something that, as per the same table, is absent when they have to choose an account under an institutional group condition.

(Table 5)

Moreover, the data in Table 5 raise a question concerning the effect of SNSs on affective polarisation among those initial (first choice) 'communities of like-minded discussants' that result from 'partisan selective exposure' (Mutz 2006; Stroud 2008). Can we expect a greater effect of polarisation among those respondents? (H_2).

To produce valid SATEs and respond to this question, we estimated an extended regression model in two steps. First, we introduced variables that might explain participant acceptance + compliance: a variable obtained with the passive meter measuring the respondents' activity on Twitter during the 15 days pre-election, party identification (dummy), education, gender and age. Second, we estimated the effect of each Twitter account exposure. In addition, we created an interactive term with the covariate on party identification to estimate the effect of partisan selective exposure or congenial social media consumption—more concretely, a dummy variable measuring whether the participant identified with the party of the respective political leader's Twitter account. In this manner, the regression coefficients represent the SATE of the respondents who were exposed to those leaders' accounts that corresponded to the party with which the respondent identified (the main cause of self-selection, as we have observed).

Table 6 presents the results of the estimation of two-step (extended) models, one for each affective polarisation index: territorial and partisan out-group affective polarisation. The

bottom part, which is intended to predict the acceptance + compliance of both models, shows that acceptance + compliance with the experiment is highly related to PID and age (much higher among younger people) and to gender albeit to a lesser degree (a lower number of women tended to accept and comply with the experiment than that of men) (Beltran et al. 2020; McGregor & Mourão 2016). These findings accord with the usual suspects in terms of who uses and participates in social media in general (Karlsen & Enjolras 2016).

At the top of each column in Table 6, we display the parameters obtained in the second step in this regression extended model with the polarisation indices as the dependent variable. The parameters represent the SATEs for exposure, PID and the interaction between exposure and this identification for each political figure's Twitter account. These results show that the exposure to a leader's Twitter account (treatment) has no significant effects on territorial polarisation by itself; more importantly, the SATEs among those who identify with a specific party are not significant (partisan selective exposure). The only significant estimators are the direct effects of identifying with a specific party.

This means that polarisation was present before the participants were exposed to leader's Twitter accounts and that they selected those based on their political preferences. However, the exposure to those accounts had no effect on affective polarisation, regardless of the anti-partisan or anti-Catalan/Spanish content. The participants who followed ERC comprise one exception; however, the SATE is only for out-group partisan polarisation and has considerably low levels of statistical significance. JxCat followers constitute another exception for territorial polarisation at low levels of significance. Therefore, we cannot confirm H_2. This means that in a context of polarisation, people are increasingly exposed to reinforcing homogeneous groups on Twitter by self-selection, but that exposure does not increase affective polarisation.

(Table 6)

Affective polarisation among party identifiers who are active on Twitter

What about going beyond simple exposure and evaluating those who actively engage with those Twitter accounts? Polarisation might also result from the freedom that these homogenous forums offer for expressing one's opinions (Shore, Baek & Dellarocas 2016). Table 7 reproduces the results of the out-group territorial polarisation but only for the

respondents who showed activity on their Twitter accounts during the days of the experiment[5]. As expected, 55.7 per cent of participants evinced no activity (posting tweets) during this time period, whereas 22.9 per cent of those who engaged on the platform did so at least once. The results in Table 7 also confirm that exposure among active participants does not substantively change the results on the lack of the effect of candidates' anti-territorial content in political figures' Twitter accounts on territorial polarisation, leading us to reject H_3 as well.

(Table 7)

Conclusions: the null effects of exposure among the motivated

This study arises from the ambition to contribute to the academic debate on social media's effect on affective polarisation. More specifically, it aims to shed some light on the effects of the discourses surrounding anti-partisan and anti-nationalist identities as produced by political elites on Twitter users in terms of affective polarisation in multiparty systems.

Our results align with those of previous investigations (Barnidge et al. 2016; Gil de Zúñiga, Barnidge & Diehl 2018; Yang et al. 2016) and point to a lesser or non-existent impact on affective polarisation for the individuals exposed to polarising discourses by political figures, even among the participants who are more polarised and active on these social media platforms. Hence, although the modal pattern for individuals, at least as the first choice, is to rely on communities of like-minded discussants, the results of our study confirm that the 'limited media effect theory' (Klapper, 1960) for social media in that said media act as 'reinforcers' or 'sustainers' of individuals' pre-existing beliefs, but not as triggers or enhancers of affective polarisation.

Thus, this study contributes to this debate in diverse directions. The study assesses the very limited or null effect of the *polarising content* that political elites may include in their Twitter accounts on the levels of polarisation of the citizens exposed to these messages. In addition, the results demonstrate that the possible null effects of polarising messages from partisan SNSs are extensive for the majority of active Twitter users, which contravenes existing arguments (Barberá & Rivero, 2015; Shore, Baek & Dellarocas 2016; Yarchi, Baden & Kligler-Vilenchik 2020). We argue that this observed null effect of polarising contents is a result of the individual motivations that are related to the need to consume and seek such information

(Valentino et al. 2008), a pattern also observed for other media (Cardenal et al 2019). The preceding finding are important given the absence of the potential moderating or counter-balancing effects of neutral social media, as it happens with mass media, and regardless of the neutralizing effects due to the active or passive exposure to other SNSs partisan sources (Barberá et al. 2015, Barberá 2020).

Nevertheless, we must interpret these results with caution owing to certain research limitations. First, our consideration of the polarising effect during the European Parliament may undermine such effects compared with what might happen during national elections (Hernández, Anduiza & Rico 2021). Second, some candidates and coalitions were new or unknown to the electorate, while the two Catalan pro-independence candidates from ERC and JxCat were in prison during the election campaign, which received significant attention during this time. These circumstances may have had an unexpected effect on the findings herein. Finally, the length of time that the participants were exposed to political figures' Twitter accounts was limited (to only the days of the experiment), and we could not assess the effect of longer or cumulative experiences on social media (Guess et al. 2021).

Declaration of interest statement

We wish to confirm that there are no known conflicts of interest associated with this publication. The funding received by the Spanish Ministry of Innovation and Science (CSO2016-79772-P and PID2019-106867RB-I00) and by ICREA has not influenced the outcome of this study.

References

Agathangelou, P., Katakis, I., Rori, L., Gunopulos, D., & Richards, B. (2017) 'Understanding online political networks: The case of the far-right and far-left in Greece', in *Social Informatics. SocInfo 2017. Lecture Notes in Computer Science*, eds. G. Ciampaglia, A. Mashhadi A., & T. Yasseri, vol 10539. Springer, Cham.

Aragón, P., Kappler, K., Kaltenbrunner, A., Laniado, D., & Volkovich, Y. (2013) 'Communication dynamics in twitter during political campaigns: the case of the 2011 Spanish national election', *Policy & Internet*, vol.5, no. 2, pp.183–206.

Bail, C. (2021). *Breaking the Social Media Prism*, Princeton University Press, Princeton.

Bakshy, E., Messing, S., & Adamic, L. (2015) 'Exposure to ideologically diverse news and opinion on facebook', *Science,* vol.348, no.6239, pp.1130–32.

Banda, Kevin K., & Cluverius, J. (2018) 'Elite polarization, party extremity, and affective polarization', *Electoral Studies,* vol 56, no. September, pp. 90-101.

Bankert, A., Huddy, L. & Rosema, M. (2017) 'Measuring partisanship as a social identity in multi-party systems', *Political Behavior,* vol. 39, no. 1, pp. 103-32.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015) 'Tweeting from left to right: Is online political communication more than an echo chamber?' *Psychological science,* vol. 26, no.10, pp.1531-1542.

Barberá, P. (2020) 'Social media, echo cambers and political polarization', in Persily, N & Tucker, J., eds., *Social media and democracy. Tthe state of the field and prospects for reform*, Cambridge University Press, Cambridge, pp 34-55

Barberá, P., Wang, N., Bonneau, R., Jost, J., Nagler, J., Tucker, J., & González-Bailón, S. (2015) 'The critical periphery in the growth of social protests', *PLoS ONE*, vol.10, no.11, pp.1-15.

Barberá, P. & Rivero, G. (2015) 'Understanding the political representativeness of twitter users', *Social Science Computer Review,* vol.33, no. 6, pp. 712-729.

Barnidge, M., Ardèvol-Abreu, A., & Gil de Zúñiga, H. (2016) 'Content-expressive behavior and ideological extremity: An examination of the roles of emotional intelligence and information network heterogeneity', *New Media & Society,* vol. 20, no.2, pp. 815-834.

Barnidge, M. (2016) 'Exposure to Political Disagreement in Social Media Versus Face-to-Face and Anonymous Online Settings'. *Political Communication,* vol.34, no.2, pp.302-321.

Barnidge, M. et al. (2018) 'Social media as a sphere for "risky" political expression: a twenty-country multilevel comparative analysis', *International Journal of Press/Politics,* vol. 23, no. 2, pp. 161-182.

Bradshaw, S., Howard, P., Kollanyi, B., & Neudert, L. (2020) 'Sourcing and automation of political news and information over social media in the united states, 2016-2018', *Political Communication,* vol.37, no.2, pp. 173–93.

Beltran, J. et al. (2021) 'Male and female politicians on Twitter: A machine learning approach', *European Journal of Political Research,* vol.60, pp. 239-250.

Bennett, W. L., & Iyengar, S. (2008) 'A new era of minimal effects? the changing foundations of political communication', *Journal of Communication,* vol. 58, no.4, pp. 707-731.

Bimber, B. (2014) ‘Digital media in the Obama campaigns of 2008 and 2012: adaptation to the personalized political communication environment’, *Journal of Information Technology & Politics*, vol.11, no.2, pp. 130–50.

Bosch, O. J., & Revilla, M. (2021). When survey science met online tracking: presenting an error framework for metered data. RECSM Working paper 62.

Boulianne, S., Koc-Michalska, K. & Bimber, B. (2020) ‘Right-wing populism, social media and echo chambers in Western democracies’, *New Media & Society*, vol. 22, no. 4, pp. 683-699.

Boxell, L., Gentzkow, M. & Shapiro, J. (2017) ‘Is the internet causing political polarization? evidence from demographics’, *National Bureau of Economic Research*, no. 23258, NBER Working Papers, National Bureau of Economic Research, Inc.

Breeze, R. (2018). 'Emotion in politics: Affective-discursive practices in UKIP and Labour', *Discourse & Society*, vol.30, no.1, pp.24-43.

Bright, J. (2018) ‘Explaining the emergence of political fragmentation on social media: The role of ideology and extremism’, *Journal of Computer-Mediated Communication*, vol.23, no. 1, pp. 17-33.

Cardenal, A., (2011) ‘Why mobilize support online? the paradox of party behaviour online’, *Party Politics,* vol.19, no. 1, pp. 83–103.

Cardenal, A. S., Aguilar-Paredes, C., Cristancho, C., & Majó-Vázquez, S. (2019) 'Echo-chambers in online news consumption: Evidence from survey and navigation data in Spain', *European Journal of Communication*, vol. *34*, no. 4, pp. 360–376.

Chadwick, A. (2013) *The hybrid media system: Politics and power*. Oxford University Press, Oxford.

Enders, A. M. & Smallpage, S. M. (2019) ‘Informational cues, partisan-motivated reasoning, and the manipulation of conspiracy beliefs’, *Political Communication*, vol.36, no. 1, pp. 83-102.

Enli, G., & Rosenberg, L. (2018) ‘Trust in the age of social media: populist politicians seem more authentic’, *Social Media + Society,* vol.4, no.1, pp. 1-11

Flaxman, S., Goel, S., & Rao, J. (2016) 'Filter Bubbles, Echo Chambers, and Online News Consumption', *Public Opinion Quarterly*, vol. 80, no. S1, pp. 298–320,

Fletcher, R., & Nielsen, R. K. (2017). 'Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication', *Journal of Communication*, vol.67, no.4, pp. 476-498.

Fletcher, R. & Jenkins, J. (2019) 'Polarisation and the news media in Europe', *European Parliamentary Research Service*. PE 634.413, March, pp 1-56.

Frame, A., & Brachotte, G. (2015) *Citizen participation and political communication in a digital world*, Routledge, New York.

Gentzkow, M., Shapiro, J. (2011) 'Ideological Segregation Online and Offline', *The Quarterly Journal of Economics*, vol. 126, no. 4, pp. 1799–1839,

Gerstlé, J. & Nai, A. (2019) 'Negativity, emotionality and populist rhetoric in election campaigns worldwide, and their effects on media attention and electoral success', *European Journal of Communication*, vol. 34, no. 4, pp. 410-444.

Gibson, R. K. (2015) 'Party change, social media and the rise of "citizen-initiated" campaigning', *Party Politics*, vol. 21, no. 2, pp. 183-197.

Gidron, N., Adams, J. & Horne, W. (2020) *American Affective Polarization in Comparative Perspective*, Cambridge University Press, Cambridge.

Giglietto, F., Valeriani, A., Righetti, N., & Marino, G. (2019) 'Diverging patterns of interaction around news on social media: insularity and partisanship during the 2018 italian election campaign', *Information Communication and Society* vol.22, no.11, pp. 1610–29.

Gil de Zúñiga, H., Barnidge, M. & Diehl, T. (2018) 'Political persuasion on social media: a moderated moderation model of political discussion disagreement and civil reasoning', *Information Society*, vol. 34, no. 5, pp. 302-315.

Glapka, E. (2019) 'Critical affect studies: On applying discourse analysis in research on affect, body and power', *Discourse & Society*, vol.30, no.6, pp. 600-621

Guess, A. M. (2021) '(Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets', *American Journal of Political Science*, vol. 65, no. 4, pp. 1007-1022.

Guess, A. M., Barberá, P., Munzert, S., & Yang, J. (2021) 'The consequences of online partisan media'. *Proc Natl Acad Sci U S A*, vol.118, no.14, pp.1-8.

Guess, A. (2016) 'Media choice and moderation: Evidence from online tracking data'. *ISPS* Yale,

Hampton, K., Shin, I., & Lu, W. (2017) 'Social media and political discussion: when online presence silences offline conversation', *Information Communication & Society*, vol.20, no.7, pp. 1090–1107.

Haselmayer, M., Wagner, M. & Meyer, T. M. (2017) 'Partisan bias in message selection: media gatekeeping of party press releases', *Political Communication*, vol. 34, no. 3, pp. 1-18.

Haselmayer, M., Meyer, T., & Wagner, M. (2019) 'Fighting for attention: media coverage of negative campaign messages', *Party Politics*, vol.25, no.3, pp.412–23.

Heatherly, K., Lu, Y., & Lee, L. (2017) 'Filtering out the other side? cross-cutting and like-minded discussions on social networking sites', *New Media and Society*, vol.19, no. 8, pp. 1271–89.

Hernández, E., Anduiza, E. & Rico, G (2021) 'Affective polarization and the salience of elections', *Electoral Studies*, vol. 69, no. 102203, pp. 1-9

Hobolt, S. (2014) 'A vote for the president? the role of spitzenkandidaten in the 2014 european parliament elections', *Journal of European Public Policy*, vol. 21, no. 10, pp. 1528–40.

Hutchens, M. J., Hmielowski, J. D., & Beam, M. A. (2019). 'Reinforcing spirals of political discussion and affective polarization', *Communication Monographs*, vol. 86, no.3, pp. 357-376.

Iyengar, S., Sood, G., & Lelkes, Y. (2012) 'Affect, not ideology: a social identity perspective on polarization', *Public Opinion Quarterly*, vol. 76, no. 3, pp. 405-31.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. (2019) 'The origins and consequences of affective polarization in the united states', *Annual Review of Political Science*, vol.22, pp. 129-46.

Johnson, T. J., Kaye, B. K. & Lee, A. M. (2017) 'Blinded by the spite? path model of political attitudes, selectivity, and social media', *Atlantic Journal of Communication*, vol. 25, no. 3, pp. 181-196.

Kahne, J. & Bowyer, B. (2018) 'The political significance of social media activity and social networks', *Political Communication*, vol. 35, no. 3, pp. 470-493.

Karlsen, R. & Enjolras, B. (2016) 'Styles of social media campaigning and influence in a hybrid political communication system: linking candidate survey data with twitter data', *The International Journal of Press/Politics*, vol. 21, no. 3, pp. 338-357.

Kingzette, J. (2021) 'Who do you loathe? Feelings towards politicians vs. ordinary people in the opposing party', *Journal of Experimental Political Science*, vol. 8, no.1, pp. 75–84.

Klapper, J. T. (1960) *The effects of mass communication*, Free Press, Glencoe.

Kruikemeier, S. (2014) 'How political candidates use twitter and the impact on votes', *Computers in Human Behavior*, vol. 34, pp.131–39.

Leeper, T. J. (2017) 'How does treatment self-selection affect inferences about political communication?', *Journal of Experimental Political Science*, vol. 4, no. 1, pp. 21-33.

Lelkes, Y. (2021) 'Policy over party: comparing the effects of candidate ideology and party on affective polarization', *Political Science Research and Methods*, vol.9, no.1, pp. 189-196

Lilleker, D. G., Tenscher, J. & Štětka, V. (2015) 'Towards hypermedia campaigning? Perceptions of new media's importance for campaigning by party strategists in comparative perspective', *Information, Communication & Society*, vol.18, no. 7, pp. 747-765.

Marozzo, F., & Bessi, A. (2018) 'Analyzing polarization of social media users and news sites during political campaigns', *Social Network Analysis and Mining*, vol. 8, no.1.

Maier, J., & Nai,A. (2020) 'Roaring candidates in the spotlight: campaign negativity, emotions, and media coverage in 107 national elections', *International Journal of Press/Politics*, vol. 25, no. 4, pp. 576-606.

McGregor, S. C. & Mourão, R. R. (2016) 'Talking politics on twitter: gender, elections, and social networks', *Social Media + Society*, vol. 2, no. 3, pp. 1-14.

Munzert, S., Barberá, P., Guess, A., & Yang, J. (2020) 'Do online voter guides empower citizens? Evidence from a field experiment with digital trace data', *Public opinion quarterly*, vol. 84, no. 3, pp. 675-698.

Mutz, C. (2006) 'How the mass media divide us', in P. S. Nivola & D. W. Brady eds., *Red and blue nation? Characteristics and causes of America's polarized politics* vol. 1, pp. 223–242, Stanford: Hoover Institution on War, Revolution and Peace, and Washington, DC: Brookings Institution Press.

Newman, N. et al. (2020) 'Reuters Institute Digital News Report 2020', *Oxford Internet Institute*. University of Oxford. Oxford.

Parsons, B. M. (2010) 'Social networks and the affective impact of political disagreement', *Political Behavior*, vol. 32, no. 2, pp. 181-204.

Persily, N., & Tucker, J. (2020) 'Conclusions: the challenges and opportunities for social media research', in Persily, N & Tucker, J.A., eds., *Social media and democracy. the state of the field and prospects for reform*. Cambridge: Cambridge University Press, pp. 313-31.

Prior, M. (2013) 'Media and political polarization', *Annual Review of Political Science*, vol. 16, pp. 101–127.

Rogowski, J. C., & Sutherland, J. L. (2015) 'How Ideology Fuels Affective Polarization', *Political Behavior*, vol.38, no.2, pp. 485-508

Rudolph, T. J., Gangl, A. & Stevens, D. (2000) 'The effects of efficacy and emotions on campaign involvement', *Journal of Politics*, vol.62, no. 4, pp. 1189-1197.

Shore, J., Baek,J., & Dellarocas, C. (2016) ‘Network structure and patterns of information diversity on twitter’, *Mis Quarterly*, vol.42, no.3, pp-849-872.

Stier, S. et al. (2018) ‘Election campaigning on social media: politicians, audiences, and the mediation of political communication on Facebook and Twitter’, *Political Communication,* vol. 35, no. 1, pp. 50-74.

Stroud, N.J. (2007) ‘Media effects, selective exposure, and *Fahrenheit 9/11*’, *Political Communication,* vol. 24, pp. 415-432.

Stroud, N.J. (2008) ‘Media use and political predispositions: Revisiting the concept of selective exposure’, *Political Behavior*, vol. 30, no. 3, pp. 341-366.

Stroud, N.J. (2010) ‘Polarization and partisan selective exposure’, *Journal of Communication,* vol. 60, pp. 556-576.

Suhay, E., Bello-Pardo, E., & Maurer, B. (2018) ‘The polarizing effects of online partisan criticism: evidence from two experiments’, *International Journal of Press/Politics* vol. 23, no. 1, pp. 95-115.

Sunstein, C.R. (2018) *#Republic: divided democracy in the age of social media*. Princeton University Press, Princeton.

Theocharis, Y. et al. (2016) ‘A bad workman blames his tweets: the consequences of citizens’ uncivil Twitter use when interacting with party candidates’, *Journal of Communication,* vol. 66, no. 6, pp. 1007-1031.

Theocharis, Y., Barberá, P., Fazekas, Z., & Popa, S. A. (2020) 'The Dynamics of Political Incivility on Twitter'. *SAGE Open*, vol.10, no.2, pp.1-15

Torcal, M., Santana, A., Carty, E., & Comellas, J. M. (2020). 'Political and affective polarisation in a democracy in crisis: The E-Dem panel survey dataset (Spain, 2018–2019)'. *Data in brief*, vol.32, pp. 1-12.

Trepte, S., & Loy, L. S. (2017), 'Social identity theory and self-categorization theory', in P. Rössler, C. A. Hoffner, & L. van Zoonen eds, *The international encyclopedia of media effects*, pp. 1-13, John Wiley & Sons, Malten.

Tsfati, Y., & Nir, L. (2017) ‘Frames and Reasoning: Two pathways from selective exposure to affective polarization’, *International Journal of Communication,* vol. 11, pp.301-322

Tucker, J. et al. (2017) ‘From liberation to turmoil: social media and democracy’, *Journal of Democracy*, vol. 28, no. 4, pp. 46-59.

Urman, A. (2020) ‘Context matters: political polarization on Twitter from a comparative perspective’, *Media, Culture and Society*, vol. 42, no. 6, pp. 857-879.

Vaccari, C., & Valeriani, A. (2015) ‘Follow the leader! direct and indirect flows of political communication during the 2013 Italian general election campaign’, *New Media and Society,* vol. 17, no. 7, pp. 1025-1042.

Vaccari, C. et al. (2016) ‘Of echo chambers and contrarian clubs: exposure to political disagreement among German and Italian users of Twitter’, *Social Media + Society,* vol. 2, no. 3, pp. 1-24.

Valentino N., Hutchings V., Banks A., & Davis, A. (2008) 'Is a worried citizen a good citizen? Emotions, political information seeking, and learning via the internet', *Political Psychology*, vol. 29, no.2, pp. 247–273.

Vosoughi, S., Roy, D., & Aral, S. (2018) ‘The spread of true and false news online’, *Science,* vol.359, no.6380, pp.1146–51.

Wetherell, M., McCreanor, T., McConville, A., Moewaka Barnes, H., & Le Grice, J. (2015) 'Settling space and covering the nation: Some conceptual considerations in analysing affect and discourse', *Emotion, Space and Society,* vol.16, pp.56-64

Yang, J. H. et al. (2016) ‘Why are “others” so polarized? perceived political polarization and media use in 10 countries’, *Journal of Computer-Mediated Communication*, vol. 21, no. 5, pp. 349-367.

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2020) ‘Political polarization on the digital sphere: a cross-platform, over-time analysis of interactional, positional, and affective polarization on social media’, *Political Communication*,vol.38, no.1-2, pp 98-139.

Zollo, Fabiana et al. (2015) ‘Emotional dynamics in the age of misinformation’, *PLoS ONE,* vol. 10, no. 9. pp.1-22.

TABLE 1: Behavioural information of respondents

	Control group (no participants)	No compliance group	Compliance group
Total Twitter time during the preceding 15 days (seconds)	232.9	1082.1	1588.4
Average time in Twitter per day in which respondent navigated	37.5	129.2	160.6
Total N	913	240	906

Source: E-DEM dataset. Own elaboration based on data collected by the passive meter.

TABLE 1: Behavioural information of respondents

	Control group (no participants)	No compliance group	Compliance group
Total Twitter time during the preceding 15 days (seconds)	232.9	1082.1	1588.4
Average time in Twitter per day in which respondent navigated	37.5	129.2	160.6
Total N	913	240	906

Source: E-DEM dataset. Own elaboration based on data collected by the passive meter.

TABLE 2: Composition of the treated groups

	Total observations	
	Politicians Twitter accounts	European institutional Twitter accounts
Leader Vox (Buxadé)	42	
Leader Partido Popular-PP (Montserrat)	51	
Leader Ciudadanos-Cs (Garicano)	60	
Leader Unidas Podemos-UP (Rodríguez Palop)	94	
Leader Partido Socialista Obrero Español-PSOE (Borrell)	133	
Leader Esquerra Republicana de Catalunya-ERC (Junqueras)	24	
Leader Junts per Catalunya- JxCat(Puigdemont)	19	
Leader Compromís (Sebastiá)	23	
Leader Partido Nacionalista Vasco-PNV (Izaskun Bilbao)	11	
Total N	457	
Spanish Office of the European Parliament		201
European Parliament (in Spanish)		87
European Commission in Madrid		36
EuroNews		124
Total N		448

Source: E-DEM dataset.

TABLE 3: Sentiment analysis of tweets about Catalans and Spaniards (Spain May 2019)

Catalan sentiment tweets					
Twitter account	**Total tweets**	**anti-Catalan**		**pro-Catalan**	
Institutional tweets		**total**	**%**	**total**	**%**
Euronews	289	5	1.73	0	0
European Commission in Madrid	55	0	0	0	0
European Parliament (in Spanish)	33	0	0	0	0
Spanish Office of the European Parliament	60	0	0	0	0
Leaders tweets					
Oriol Junqueras (ERC)	**38**	**0**	**0**	**14**	**36.8**
Carles Puigdemont (JxCat)	**78**	**0**	**0**	**13**	**16.7**
Dolors Montserrat (PP)	**80**	**15**	**18.7**	**2**	**2.5**
Jorge Buxadé (Vox)	**109**	**7**	**6.4**	**0**	**0**
Luis Garicano (Cs)	**62**	**9**	**14.5**	**0**	**0**
Neutral tweets					
Josep Borrell (PSOE)	114	0	0	2	1.7
María Eugenia Rodríguez Palop (UP)	148	0	0	0	0
Izaskun Bilbao (PNV)	363	0	0	4	1.1
Jordi Sebastiá (Compromís)	782	0	0	4	0.5
Spanish sentiment tweets					
Twitter accounts	**Total tweets**	**anti-Spanish**		**pro-Spanish**	
Institutional tweets		**total**	**%**	**total**	**%**
Euronews	289	0	0	1	0.35
European Commission in Madrid	55	0	0	0	0
European Parliament (in Spanish)	33	0	0	0	0
Spanish Office of the European Parliament	60	0	0	0	0
Leaders tweets					
Oriol Junqueras (ERC)	**38**	**19**	**50**	**0**	**0**
Carles Puigdemont (JxCat)	**78**	**26**	**33.3**	**0**	**0**
Jorge Buxadé (Vox)	**109**	**0**	**0**	**9**	**8.3**
Dolors Montserrat (PP)	**80**	**0**	**0**	**2**	**2.5**
Luis Garicano (Cs)	**62**	**0**	**0**	**6**	**9.7**
Neutral tweets					
Josep Borrell (PSOE)	114	0	0	3	2.6
María Eugenia Rodríguez Palop (UP)	148	0	0	0	0
Izaskun Bilbao (PNV)	363	1	0.3	0	0
Jordi Sebastiá (Compromís)	782	2	0.3	0	0

Source: E-DEM dataset.

*Own elaboration based on data retrieved from Twitter using REST API and hand coded sentiment analysis regarding Catalans and Spaniards expressed in percentage among the whole twitter activity of the referred accounts during the days of the experiment.

** Given the lack of dictionaries expressing these kinds of feelings in Spanish, Basque and Catalan, each tweet has been coded as 1 or 0 accordingly if it contains any of the following expressions:

TABLE 4: Sentiment analysis of tweets about other parties (Spain May 2019)

Twitter accounts	Total tweets	Antipartisan tweets	
Institutional tweets		**total**	**%**
Euronews	289	0	0.0
European Commission in Madrid	55	0	0.0
European Parliament (in Spanish)	33	0	0.0
Spanish Office of the European Parliament	60	0	0.0
Leaders tweets (by percentage of antiparty messages)			
Oriol Junqueras (ERC)	38	16	42
Carles Puigdemont (JxCat)	78	27	34.6
Dolors Montserrat (PP)	80	26	32.5
Jorge Buxadé (Vox)	109	31	28.4
Luis Garicano (Cs)	62	17	27.4
María Eugenia Rodríguez Palop (UP)	148	23	15.5
Jordi Sebastiá (Compromís)	782	93	11.9
Josep Borrell (PSOE)	114	12	10.5
Izaskun Bilbao (PNV)	363	37	10.2

Source: E-DEM dataset.

*Data collected through Twitter REST API from candidates and institutions Twitter accounts included in the experiment. Antipartisan tweets represents the number of tweets that accounts for any negative expression attached to a clear reference to the name of the political party, either the long version or the acronym, political leaders or well-known members of their parties.

TABLE 5: Ideological preferences, territorial identities and party identification among Twitters users

Twitter accounts	Average left-right position	Average identification with Spain	Highest party identification (%)
Euronews	3.60	6.69	20(UP)
European Commission in Madrid	3.91	5.67	26(PSOE)
European Parliament (in Spanish)	3.57	6.69	26(UP)
Spanish Office of the European Parliament	4.20	6.10	24(PSOE)
Oriol Junqueras (ERC)	3.12	2.33	71(ERC)
Carles Puigdemont (JxCat)	3	1.36	53(JxCat)
Dolors Montserrat (PP)	6	7.80	60(PP)
Jorge Buxadé (Vox)	5.8	7.88	74(Vox)
Luis Garicano (Cs)	5.31	7.93	83(Cs)
María Eugenia Rodríguez Palop (UP)	2.37	5.90	50(UP)
Jordi Sebastiá (Compromís)	3.30	5.43	41(UP)
Josep Borrell (PSOE)	3.77	7.24	71(PSOE)
Izaskun Bilbao (PNV)	3.54	3.63	n.i.
CONTROL GROUP (No participants+no compliants)	4.20		

Source: E-DEM dataset.

*Respondents who voluntarily gave their own twitter accounts self-placement in the ideological spectrum, identity belonging to Spain, and party identification.

TABLE 6: Extended linear regression for territorial and partisan affective polarisation

	Territorial affective polarisation (Spread of scores towards territorial groups)	**Partisan out-group affective polarisation**
Vox twitter Follower	0.056	-0.136
	(0.214)	(0.304)
Vox Identifier	1.059***	-0.693**
	(0.167)	(0.218)
Vox **Identifier+** Vox **twitter follower**	0.215	0.443
	(0.365)	(0.478)
PP twitter Follower	0.428*	-0.387
	(0.193)	(0.271)
PP Identifier	0.958***	-0.522*
	(0.166)	(0.215)
PP Identifier+PP twitter follower	-0.603+	0.213
	(0.335)	(0.441)
Cs twitter Follower	0.220	-0.059
	(0.180)	(0.259)
Cs Identifier	0.289+	-0.859***
	(0.150)	(0.184)
Cs Identifier+Cs twitter follower	-0.103	0.641
	(0.311)	(0.401)
PSOE twitter Follower	0.006	-0.200
	(0.135)	(0.181)
PSOE Identifier	0.057	-0.118
	(0.113)	(0.145)
PSOE Identifier+PSOE twitter follower	0.042	0.007
	(0.209)	(0.266)
UP twitter Follower	-0.174	-0.504*
	(0.173)	(0.230)
UP Identifier	0.084	-0.442**
	(0.111)	(0.142)
UP Identifier+UP twitter follower	0.163	0.414
	(0.241)	(0.307)
Compromis twitter Follower	0.379	-0.474
	(0.288)	(0.364)
UP Identifier	-0.342	0.770
	(0.482)	(0.591)
UP Identifier+Comprimis twitter follower	-0.362	0.324
	(0.316)	(0.423)
ERC twitter Follower	0.711***	-0.270
	(0.195)	(0.249)
ERC Identifier	0.250	-0.198
	(0.479)	(0.608)
ERC Identifier+ERC twitter follower	0.020	0.616+

	(0.305)	(0.371)
JxCat twitter Follower	1.296***	0.023
	(0.273)	(0.349)
JxCat Identifier	0.185	-0.510
	(0.620)	(0.763)
JxCat Identifier+JxCat twitter follower	0.717*	0.392
	(0.332)	(0.434)
PNV twitter Follower	1.511	-0.523
	(1.068)	(1.285)
PNV Identifier	-0.133	-0.535
	(1.585)	(1.866)
PNV Identifier+PNV twitter follower	0.056	-0.136
	(0.214)	(0.304)
Compliance	-1.773***	-0.278
	(0.137)	(0.472)
Intercept	2.368***	8.088***
	(0.116)	(0.358)
Compliance Equation		
Twitter average daily use	-0.000	0.000
	(0.000)	(0.000)
Education	0.052*	0.077*
	(0.023)	(0.033)
Age	-0.008**	-0.007+
	(0.003)	(0.003)
Female	-0.103	-0.200*
	(0.067)	(0.099)
Party Identification (PID)	0.332***	0.346***
	(0.078)	(0.098)
Intercept	0.808***	0.788**
	(0.221)	(0.305)
Error variance(territorial polarisation)	1.568***	
	(0.100)	
Correlation(e.compliance,e.territorial polarization)	0.795***	
	(0.035)	
Error variance (partisan polarisation)		1.662***
		(0.091)
Correlation(e.compliance,e.partisan polarization)		0.162
		(0.205)
Wald Chi^2	315.02***	59.95***
Log likelihood	- 3218.59	- 2192.80
N	1145	1017

Source: E-DEM dataset.
Standard errors in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 7: Extended linear regression for territorial affective polarisation, by level of activism in Twitters accounts

	Non-active	**Active**
Vox twitter Follower	-0.363	-0.383
	(0.918)	(0.526)
Vox Identifier	0.832^{+}	1.725^{+}
	(0.450)	(0.882)
Vox **Identifier+** Vox **twitter follower**	-0.673	No cases
	(1.132)	
PP twitter Follower	0.790^{+}	0.802
	(0.431)	(0.640)
PP Identifier	0.945*	0.740
	(0.425)	(0.542)
PP Identifier+PP twitter follower	-0.310	-0.317
	(0.782)	(1.022)
Cs twitter Follower	0.651	-0.012
	(0.543)	(0.651)
Cs Identifier	0.029	0.608
	(0.318)	(0.436)
Cs Identifier+Cs twitter follower	0.112	0.362
	(0.800)	(0.918)
PSOE twitter Follower	1.004^{+}	0.234
	(0.544)	(0.390)
PSOE Identifier	-0.158	-0.372
	(0.279)	(0.349)
PSOE Identifier+PSOE twitter follower	-0.667	-0.165
	(0.649)	(0.623)
UP twitter Follower	0.319	-0.035
	(0.349)	(0.380)
UP Identifier	0.564^{+}	-0.519
	(0.320)	(0.325)
UP Identifier+UP twitter follower	-0.880^{+}	0.665
	(0.509)	(0.561)
Compromis twitter Follower	-1.466	1.243*
	(0.936)	(0.614)
UP Identifier	0.564^{+}	-0.519
	(0.320)	(0.325)
UP Identifier+Comprimis twitter follower	No cases	No cases
ERC twitter Follower	No cases	0.272
		(0.663)
ERC Identifier	1.338**	-0.208
	(0.442)	(0.500)
ERC Identifier+ERC twitter follower	No cases	No cases
JxCat twitter Follower	-0.387	-0.728
	(0.665)	(1.074)
JxCat Identifier	3.200***	-1.280^{+}
	(0.894)	(0.777)

JxCat Identifier+JxCat twitter follower	-1.544	No cases
	(1.417)	
PNV twitter Follower	-1.850^{+}	1.556**
	(0.969)	(0.485)
PNV Identifier	No cases	0.355
		(1.117)
PNV Identifier+PNV twitter follower	No cases	No cases
	0.769	-1.932***
Compliance	(0.748)	(0.379)
Intercept	0.228	2.812***
	(0.623)	(0.365)
Compliance Equation		
Twitter average daily use	-0.000	-0.000*
	(0.001)	(0.000)
Education	-0.069	-0.007
	(0.099)	(0.101)
Age	-0.012	-0.023*
	(0.012)	(0.010)
Female	-0.520^{+}	-0.385
	(0.291)	(0.285)
Party Identification (PID)	0.491	0.346
	(0.304)	(0.301)
Intercept	2.531*	2.630**
	(1.053)	(0.917)
Error variance(territorial polarisation)	0.859***	1.097***
	(0.155)	(0.171)
Correlation(e.compliance,e.territorial polarisation)	-0.441	0.828***
	(0.387)	(0.102)
Wald Chi^2	48.47***	62.93***
Log likelihood	- 220.88	- 192.78
N	134	119

Source: E-DEM dataset
Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

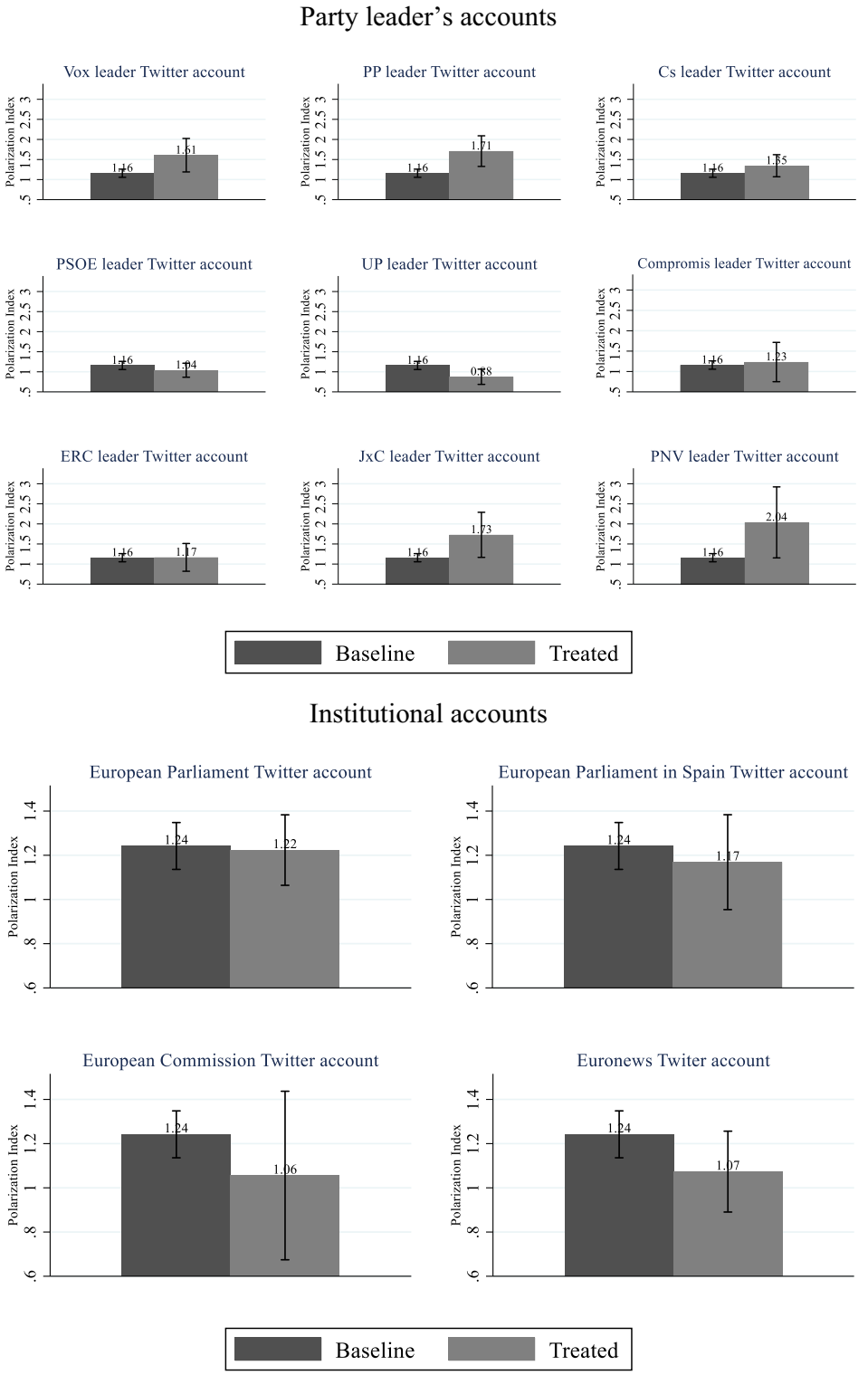

Figure 1. Average treatment of the treated (ATT) on territorial affective polarisation (Spread of sentiments scores towards territorial groups).
Source: E-DEM dataset.

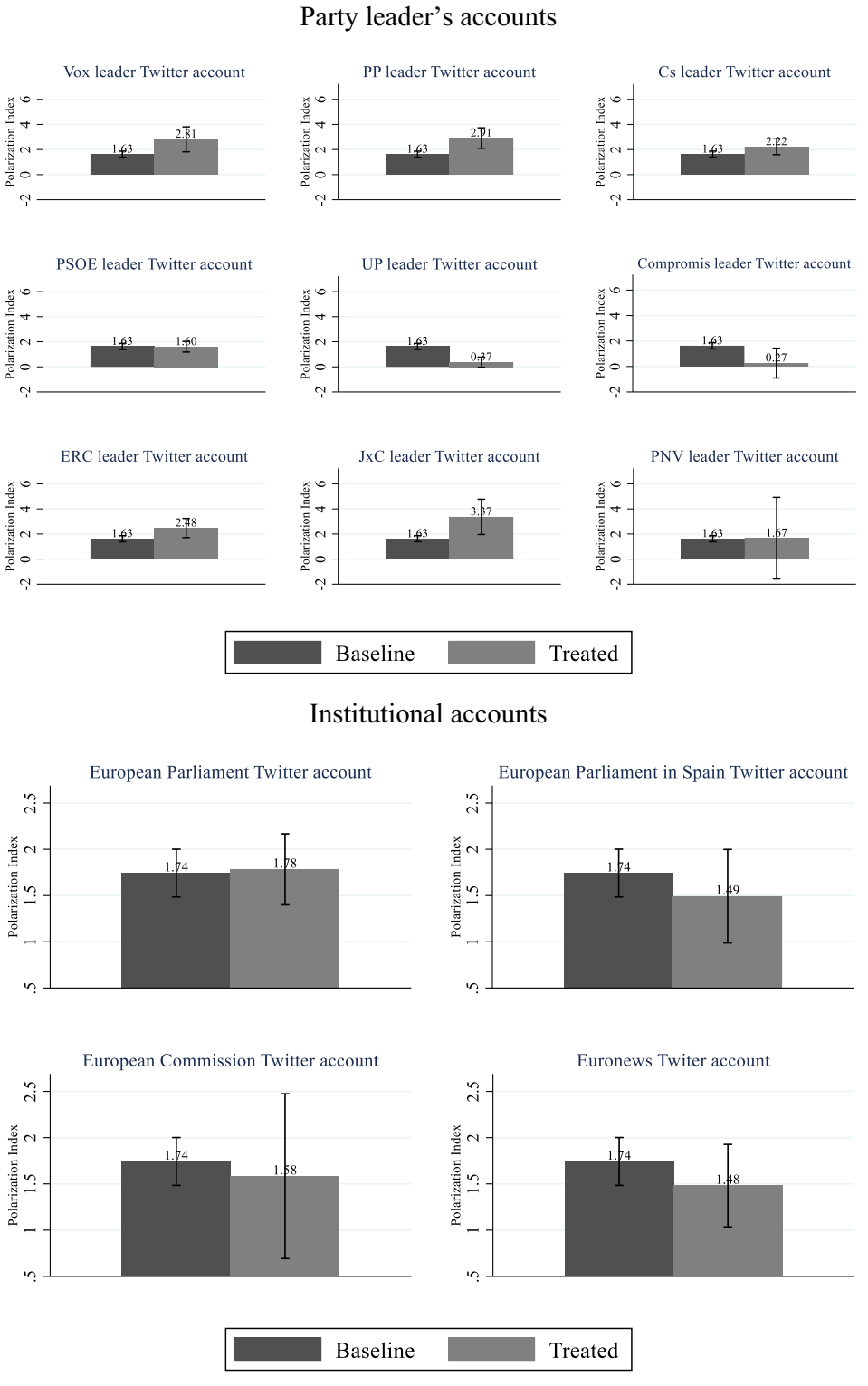

Figure 2. Average treatment of the treated (ATT) on territorial affective polarisation (Mean distance from Catalans).
Source: E-DEM dataset.

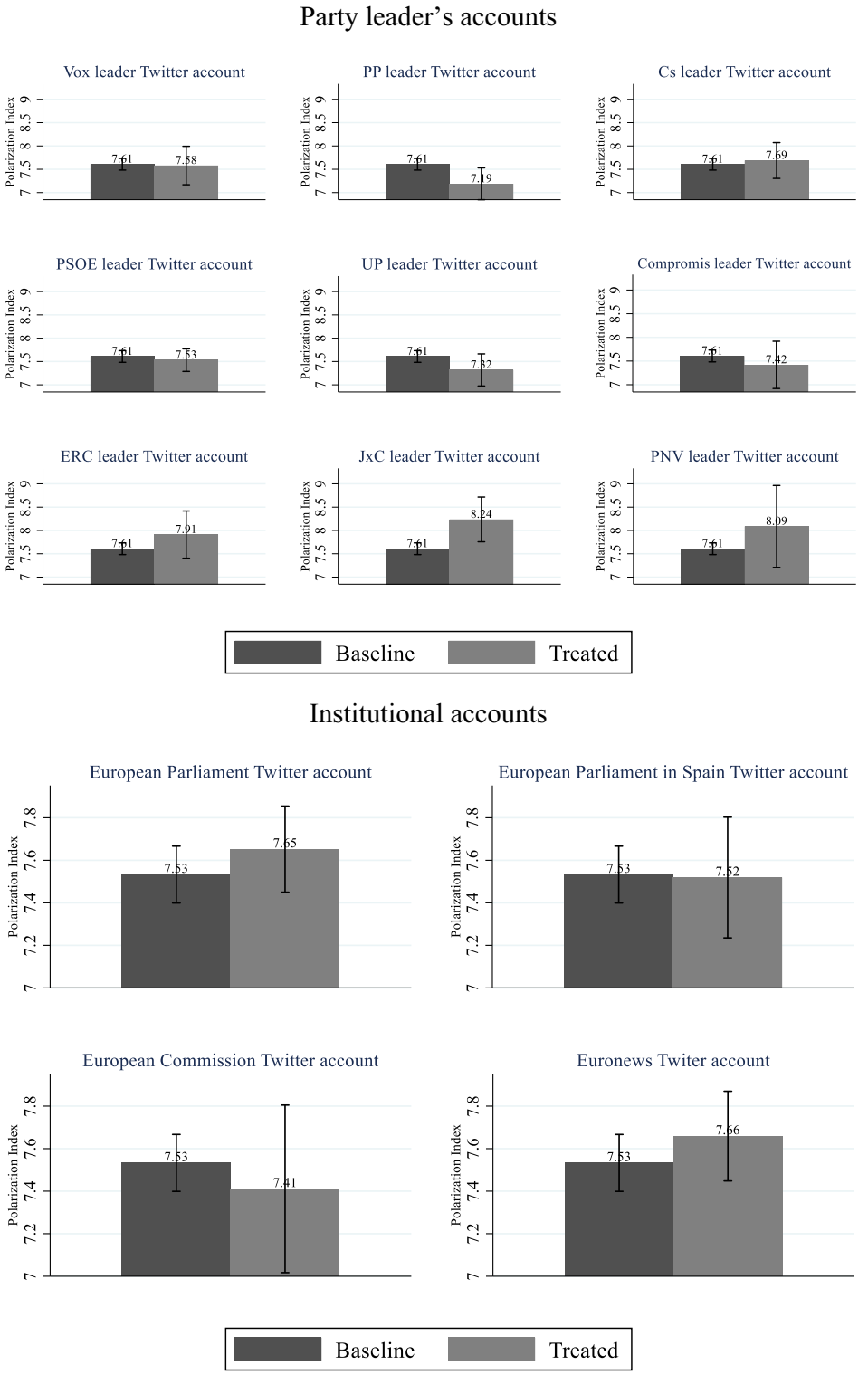

Figure 3. Average treatment of the treated (ATT) on partisan affective polarisation (out-group polarisation).
Source: E-DEM dataset.

[1] In Spain, Election Day was 26 May.

[2] Question p19a_4 in the codebook of the dataset.

[3] Those are as follow:

Anti-Catalan—'independentis*', 'criminal', 'separatis*', 'henchmen', 'want to break Spain' and 'seccesion*';

Pro-Catalan—'freedom fighters', 'our country', 'state recognition', 'proud of being Catalan', 'proud of our land', 'the fight will make us free' and 'long live Catalonia';

Anti-Spanish—'from the exile', 'exiles', 'Spanish injustice', 'nondemocratic country', 'repress*', 'Spanish nationalists', 'totalitarians' and 'court prosecution' and

Pro-Spanish—'proud to be Spanish', 'great Spain', 'Spaniards together are stronger' and 'long live Spain'.

[4] Negative expressions were coded as anti-partisan when attached to a clear reference to the name of the political party, either the long version or the acronym, political leaders or well-known members of their parties. In addition, we included regular expressions that were commonly employed in the media and among public opinion to refer to each party, such as 'socialistas' or 'sociatas' for PSOE; 'comunistas' for Unidas Podemos; 'conservadores' or 'peperos' for PP; 'naranjas' or 'naranjito' for Cs; 'abertzales' for EH-Bildu; 'fachas' or 'fascist*' for Vox; 'separatis*' 'independentis*' as an anti-partisan feeling against JxCat and ERC and 'PPPSOE' as a negative expression for both PP and PSOE.

[5] To capture each respondent's Twitter behaviour, we asked them in a previous survey for their usernames, collecting a total of 269 accounts. The resulted datasets have been completely anonymised using the respondent's numerical id for the survey as the matching id; all traces of usernames (either 'screen_name' or 'user_id') have been deleted.