Article

new media & society
1–27

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14614448221099860
journals.sagepub.com/home/nms
SAGE

Few voices, strong echo: Measuring follower homogeneity of politicians' Twitter accounts

Felix Rusche
University of Mannheim, Germany

Abstract

Politicians have discovered Twitter as a tool for political communication. If information provided by politicians is circulated in ideologically segregated user networks, political polarization may be fostered. Using network information on all 1.78 million unique followers of German Members of Parliament by October 2018, follower homogeneity across politicians and parties is measured. While the overall homogeneity is low, politicians of the AfD—a right-wing populist party—stand out with very homogeneous follower networks. These are largely driven by a small group of strongly committed partisans that make up around 7% of the party's but around 55–75% of the average AfD politician's followers. The findings add to the literature by showing potentially unequal distributions of network segregation on Twitter. Furthermore, they suggest that small groups of active users can multiply their influence online, which has important implications for future research on echo chambers and other online phenomena.

Keywords

Audience homophily, echo chambers, homophily, network analysis, network homogeneity, partisan media, political communication, selective exposure, Twitter

Introduction: investigating homogeneity of politicians' networks on Twitter

The advance of social media sparked a debate on whether new online platforms foster the contact with a diverse array of people and ideas, or whether the used technologies give rise to narrower, more homogeneous networks. Already in a 1997 working paper, Van

Corresponding author:
Felix Rusche, Department of Economics, University of Mannheim, L7 3-5, 68161 Mannheim, Germany.
Email: felix.rusche@gess.uni-mannheim.de

Alstyne and Brynjolfsson (1997, 2005) predicted the Internet's potential to create "Cyber Balkans" and their potential consequences: "Individuals empowered to screen out material that does not conform to their existing preferences may form virtual cliques, insulate them-selves from opposing points of view, and reinforce their biases" (pp. 865–866). While current empirical research casts doubt on such narrowing effects of social media (e.g. Scharkow et al., 2020), so-called "filter bubbles" and "echo chambers" frequently are the subject of scientific (see Section "Network homogeneity around politicians' accounts") and public debate (e.g. BBC, 2019). One major concern is that echo chambers may drive opinion polarization in society (Sunstein, 2001).

Politicians have long discovered the potential of social media to drive their agendas, particularly on the micro-blogging platform Twitter (Ausserhofer and Maireder, 2013; Jungherr, 2016). However, the impact of politicians' tweets depends on the audience that receives these messages. Tweets may unfold very different effects when communicated to different audiences (see Section "Network homogeneity around politicians' accounts"). Users who are committed to the politician's party may be more likely to support and believe the content and spread the message approvingly. Other users may follow a politician merely to stay up-to-date on current political debates, while yet others might deliberately seek for opportunities to critically respond to a politician's messages.

To assess the potential impact of politicians' tweets, one should thus carefully examine who is regularly reached by their messages on Twitter. Recent research suggests that followers on Twitter have the power to influence which discussions politicians pick up and how they set their agendas. In particular, Barberá et al. (2019) show that politicians in US congress are much more reactive to supporters' issue attention on Twitter than to the general public or media outlets. Given that followers are the first to receive and react to politicians' tweets, they can provide immediate feedback to their messages and behaviors, and shape the way in which politicians perceive their supporters.

However, while several studies on echo chambers examine Twitter users' tendency to communicate and connect with others that think alike (e.g. Barberá, 2015; Boutyline and Willer, 2017; Colleoni et al., 2014), individual politicians' follower networks and their structure have received little attention (see Section "Network homogeneity around politicians' accounts"). Studying these separately is important though, as for example, a conservative politician's average follower may differ substantially from the average conservative Twitter user. For example, if a small group of partisan accounts follows many conservative politicians, partisans get a low weight in analyses focusing on the average conservative Twitter user. At the same time, they may make up large shares of individual politicians' followers. Through high levels of activity, they can thus effectively multiply their influence online and create homogeneous follower networks. High levels of follower homogeneity may thus constitute a specific element of echo chambers, that is, of the segregation of networks into communities of like-minded individuals (Sunstein, 2001).

This article aims to carefully examine the network structure of individual politicians and to compare the results across politicians and parties. This is done in two steps: First, a measure of follower homogeneity is applied to estimate the degree to which each politician's followers are only interested in the politician's own party or also follow other political parties. The results are compared across politicians and political parties. Second, building on followers' connections across different parties, these are categorized according to their type and level of commitment to a given party. Drawing on this, the extent to which different user groups, such as partisans or users who are broadly interested in multiple parties, drive network homogeneity is analyzed.

To apply the above investigation, data on all 1.78 million followers of the 462 German politicians in the national parliament (Bundestag) active on Twitter was collected in October 2018. Drawing on their 5.4 million connections to Bundestag politicians, the homogeneity measure is implemented. The findings suggest that while politicians' follower networks tend to be homogeneous, followers remain interested in other parties as well, suggesting that there is no overall segregation. However, very strong homogeneity for politicians of the alternative for Germany (AfD), a right-wing populist party (Lewandowsky, 2015; Schmitt-Beck et al., 2017), is found. It is shown that the average follower of most AfD politicians only or mostly follows parliamentarians of the AfD, suggesting a high level of network homogeneity. More specifically, for most AfD politicians around 80% of the average follower's connections across all Bundestag politicians go toward the AfD. In the second step, individual followers are categorized depending on their following behavior across Bundestag politicians. It is shown that the strong homogeneity of followers encountered above is mainly driven by a small group of strongly committed partisans that make up around 7% of the AfD's followers, but around 55–75% of most AfD politicians' followers. The strong network homogeneity is thus driven by a small group of users that follow many AfD politicians (and few others). Furthermore, AfD politicians attract fewer users that are broadly interested in multiple parties, which further contributes to the high levels of network homogeneity. Using both frequentist and bayesian linear regressions, it is shown that the results are not driven by observable characteristics of politicians or their Twitter accounts, but largely accounted for by party affiliations. Applying a machine learning algorithm to classify profiles' probability to be bots further shows that the results are stable to the control for and exclusion of potential bots.

The remainder of the article is structured as follows: first, the article is discussed in the context of the existing literature and a number of hypotheses and research questions are derived. Section "Network homogeneity around politicians' accounts" further provides a short introduction to Germany's political and party landscape. In Section "Measuring Follower Homogeneity", the homogeneity measure is presented and discussed in the context of other measures of homophily and network homogeneity. Furthermore, the categorization of followers is introduced as well as methods to check the results' robustness. These data are discussed in Section "Data: politicians and their followers on Twitter". Section "Results" then presents the results and tests the robustness. Finally, Section "Conclusion" concludes and discusses potential paths for future research.

Network homogeneity around politicians' accounts

Psychological processes are likely to lead to the creation of homogeneous networks on social media. Social media, such as Twitter, allow people to actively construct their networks and choose their news diets. Users select the content that should appear in their news feed by following specific accounts (and not following others). This behavior may conform to “selective exposure,” the alleged tendency of individuals to tailor their own media environments to reflect their beliefs and predispositions, while avoiding opposing points of view (Stroud, 2018). Further and closely related, homophily, that is, individuals’ tendency to surround themselves with similar others (McPherson et al., 2001), has been suggested to drive network formation (e.g. Koiranen et al., 2019). This may further foster the creation of homogeneous networks, which is potentially amplified by audience homophily: Dvir-Gvirsman (2017) shows that individuals choose to consume media depending on the audience of an outlet to strengthen their identification with specific political groups.

Much of the existing literature studies network segregation at the level of individual users instead of focusing on influential accounts, such as news organizations or politicians and their followers. The main aim of such papers is to investigate whether individuals on social media tend to connect and discuss with like-minded others. For example, Bright (2018) shows that the interaction of users on Twitter decreases with their ideological distance. However, the literature also shows that users are nonetheless confronted with information that does not conform with their opinions (Barberá, 2015; Boutyline and Willer, 2017; Cinelli et al., 2021; Dubois and Blank, 2018; Eady et al., 2019; Flaxman et al., 2016; Lawrence et al., 2010; Lee et al., 2014; Vaccari et al., 2016; Yardi and Boyd, 2010). For example, Barberá et al. (2015) draw upon data of 3.8 million US Twitter users and their tweets. The authors measure a Twitter user’s ideological position based on her connections to popular partisan accounts, media outlets, and other users, using a latent space model. Following this, the authors focus on 12 political and non-political events to measure whether discussions are polarized, that is, mostly take place between like-minded individuals. They show that while discussions of political issues are highly polarized, those of non-political ones take place between individuals with different ideological positions.

More closely related to this article, a smaller strand of the literature focuses on measuring the homogeneity of specific outlets’ or media types’ audience. Starting with the latter, Gentzkow and Shapiro (2011) utilize an isolation index (see Section “Measuring Follower Homogeneity”) to study and compare ideological segregation between different types of media. Here, segregation is high if, for a certain media type (e.g. television), liberals and conservatives watch very different stations with little overlap in viewership. They show that while the Internet is more ideologically segregated than the market for television news or magazines, it is less segregated than offline social networks or readers of national newspapers. Similarly, Webster and Ksiazek (2012) analyze the audience segregation of different news outlets. They show that most outlets’ audience, such as NBC or Fox News, also read other news sites, though they do not take into consideration whether audiences of ideologically close outlets have stronger overlaps. In addition, a number of studies provide evidence regarding the homogeneity of specific news outlets’ followers. More specifically, three studies by Dvir-Gvirsman (2017), Flaxman et al. (2016), and Gentzkow and Shapiro (2011) start by categorizing individuals as conservative or not and show that in the United States and Israel websites with right-wing content have more conservative readers than those with center or left-wing content. For example, Flaxman et al. (2016) show that 31% of *New York Times* readers are conservative when compared to 59% of Fox News readers.

This study aims to combine both the follower and the followee perspective in a Twitter-based study by focusing on the following behavior of individuals as well as on the resulting follower composition and homogeneity of politicians' accounts. While a relatively large literature studies the tendency of individuals and user groups to communicate with like-minded others, research that extensively investigates the follower homogeneity and structure of individual politicians' or news outlets' followers remains scarce. To contribute to this, this study starts at the politician level, by estimating politicians' follower homogeneity and comparing it across parties. It then continues at the individual level, by sorting followers into groups, including "partisans," to measure what types of followers drive network homogeneity.

While, as discussed earlier, there is relatively little research on the homogeneity and composition of individual politicians' (and news outlets') followers on Twitter, the presented evidence in combination with individual-level tendencies to connect to and discuss with like-minded others suggests that follower networks are expected to be rather homogeneous. As discussed, this is likely driven by the mechanisms of selective exposure, homophily, and audience homophily, hence the first hypothesis:

H1. Politicians' follower networks are homogeneous.

In this context, more extreme politicians are likely to attract more homogeneous audiences. First, Dvir-Gvirsman (2017) suggests that the degree of audience homophily is stronger for individuals with more extreme views, that is, these tend to prefer news outlets with a more homogeneous audience. This implies that follower homogeneity is increased for outlets attracting such individuals. Here, selective exposure and homophily suggest that individuals are more likely to follow like-minded others, that is, politicians that reflect their own opinions (McPherson et al., 2001; Stroud, 2018). As a result, extreme parties' politicians are more likely to attract individuals with extreme views, thus increasing their follower homogeneity through higher levels of audience homophily. Second, individuals with more extreme ideologies have been shown to tend toward higher levels of selective exposure and homophily, given that their certainty of "holding the truth" is higher (Boutyline and Willer, 2017; Johnson et al., 2009). By the same logic as above, such individuals are more likely to follow extreme parties' politicians and thus increase politicians' follower homogeneity. Third, selective exposure is likely to reduce non-extreme users' willingness to follow extreme parties. While users may be willing to follow politicians of other parties that are relatively close to their own political affiliation, extreme parties are (by definition) further off, that is, non-extreme users may be less willing to expose themselves to such content. This is in line with research showing that user interaction on social media is reduced in the ideological distance between users and is particularly low in-between extremist and non-extremist users (Bright, 2018). Such mechanisms are also reflected in models where readers choose what sources to consume based on their prior beliefs about the truth (e.g. Mullainathan and Shleifer, 2005). Thus, the second hypothesis is as follows:

H2. Extreme parties' politicians have more homogeneous audiences than politicians of other parties.

Defining partisan followers as users that are only or mostly interested in a given party, extreme parties are expected to have the highest share of partisan followers. First, as discussed earlier, they are expected to attract fewer individuals that are rather broadly interested in politics, thus increasing the weight of other followers, including partisans. Second, given that extreme individuals have been suggested to have higher levels of selective exposure and audience homophily (Dvir-Gvirsman, 2017; Johnson et al., 2009), the share of partisans among an extremist party's unique followers is likely higher. Hence, hypotheses 3a and 3b are as follows:

H3a. The share of followers that are broadly interested in politics is lower for extreme parties' politicians than for politicians of other parties.

H3b. The share of followers that are partisan is higher for extreme parties' politicians than for politicians of other parties.

Combining the hypotheses on the follower homogeneity (*H1, H2*) with those on different user groups (H3a, H3b) leads to the question to what degree follower homogeneity is related to the share of the respective groups among followers. Both follower homogeneity and the shares of follower groups are defined on the politician level. At the same time, follower groups, such as partisans, can follow multiple politicians and thus affect the respective measures. Thus, even small groups of followers that follow a high number of politicians can affect the results regarding the previous hypotheses. By following many politicians, they would effectively have a higher weight on the respective outcomes than, for example, a user that only follows a single politician. As a result, the following research questions are put forward:

RQ1a. How is follower homogeneity related to the share of partisans among followers?

RQ1b. How is follower homogeneity related to the share of broad followers among followers?

RQ1c. How is follower homogeneity related to the share of other groups among followers?

Context: Germany's political landscape, party system, and political Twittersphere

This section aims to introduce the most general features of Germany's political and party landscape as well as political Twittersphere. While there is a multitude of parties, few make it into national parliament, given the 5% threshold. As a result, between unification and the 2013 elections, six parties have been considered "[. . .] the main 'actors' on the stage of the Bundestag" (Saalfeld and Schoen, 2015:. 106), of which most are placed at the center of the political spectrum. While the correlation between the economic and social left-right scale is lower than, for example, in the United States (Barberá, 2015), one can broadly classify parties on a left-right scale: SPD (social democrats) and Die Grünen (the greens) are placed at the center-left, the CDU and CSU (christian democratic/social union) at the center right, while the FDP (free democratic party), as a liberal, pro-market party, is often placed in-between both groups. CDU and CSU form a single parliamentary group and do not compete in elections, given that the CSU is exclusively present in Bavaria. Historically, the CDU(/CSU) and SPD have been the main parties (coalition leaders), with all former chancellors since the Second World War coming from either party. However, there have been several "grand coalitions" between both groups, including those under the leadership of Angela Merkel from 2005–2009 and 2013–2021 (Saalfeld and Schoen, 2015).

In addition, there are two extreme parties both on the left (Die Linke—the left) and the right (AfD—alternative for Germany). The AfD, which started as a single issue Euroskeptic party (Saalfeld and Schoen, 2015), developed into a right-wing populist party by 2017, when it first entered the Bundestag (Lewandowsky, 2015; Schmitt-Beck et al., 2017). On the other side of the political spectrum, Die Linke is considered less extreme and it remains contested whether it is a populist party (Chiocchetti, 2016; Fawzi et al., 2017). Here, the Varieties of Democracy (V-Dem) Project gives the party a populism score of 0.329 on a scale from 0 to 1, while the AfD has a score of 0.949 (all other parties in the Bundestag lie between 0.056 and 0.171). Similarly, for illiberalism, the AfD has a score of 0.671 while Die Linke is at 0.063 and thus slightly above the CSU at 0.061, while other parties are between 0.023 and 0.048 (Lührmann et al., 2020). The AfD is thus operationalized as the most extreme party. While Die Linke is less extreme, it may still be viewed as the second most extreme party on the opposite, that is, left, side of the political spectrum. Both parties are particularly successful in Eastern Germany, including during the 2017 elections. Furthermore, the AfD was stronger in South when compared to North–West Germany (Bundeswahlleiter, 2017a; Küpper et al., 2020).

Moving to Germany's Twittersphere, around 12% of Germans use Twitter, making the platform less popular than in France, the United States, or the United Kingdom (17%, 25%, and 31%; Newman et al., 2021). As in other contexts, Twitter users are not representative of the German public, but rather tend to be young, more partisan, and more politically active (Jungherr, 2016). Nevertheless, Twitter is an important tool for political communication, given the strong use among both politicians and journalists (Jungherr, 2014; Stier et al., 2018) and its effects on traditional media coverage (Gilardi et al., 2021) and political agendas (Barberá et al., 2019). For instance, in Switzerland, a setting with a very similar Twitter usage as Germany's (Newman et al., 2021), Gilardi et al. (2021) show that while the agendas of traditional media and those of parties and politicians on Twitter are generally congruent, parties on Twitter pushed the environmental issue into traditional media, showing Twitter's potential to drive agendas.

Measuring follower homogeneity

Introducing the measure

Scholars on social networks have developed a considerable number of measures for network segregation and polarization (Bojanowski and Corten, 2014). This article is strongly related to other studies on homophily on Twitter. As discussed earlier, studies in this context have increasingly relied upon latent space models (Barberá, 2015; Barberá et al., 2015; Bond and Messing, 2015). These start by placing media outlets and politicians on a political left-right scale and subsequently place other Twitter users on this scale as well, given their connections to politicians, media websites, and other users. However, as noted in Barberá (2015), such models might not translate well into the German political context, as a one-dimensional left-right scale performs relatively poorly in settings where the correlation between the social and economic left-right position is much lower than in the United States. Another difficulty arises, as there is no extensive data set scoring most German media outlets on a left-right scale.[1] Applying such models to Germany would thus require both a simplification of the German political sphere as well as assumptions on the positions of media outlets.

While the studies above place both users and politicians/media outlets on a political left-right scale, another strand of the literature uses isolation measures to measure the homogeneity in politicians'/media outlets' followers (see the discussion of Gentzkow and Shapiro (2011) in Section "Network homogeneity around politicians' accounts"). Here, Dvir-Gvirsman (2017) builds on the self-reported ideological position on a liberal-conservative scale, users' browsing behavior, and hand-coded ideological positions of websites. She calculates a measure on the user level, where users' networks are scored between −1 and 1. Here, a score of −1 for a conservative user suggests that all of the co-visitors of the websites she visits are liberal and a score of 1 suggests that all co-visitors are conservative, that is, that the user's browsing behavior is very homogeneous.

This article builds upon these ideas to overcome the problems encountered by latent space models in the German context, that is, the problem of the one-dimensional left-right scale and the lack of data (and potential bias thereof) on the position of media outlets on such a scale. Furthermore, it adds an important perspective by focusing on politicians' follower networks. To do so, it departs from the studies on individual users' tendency to connect to like-minded others, by exclusively focusing on politicians and by not estimating individuals' (or politicians') political ideology. Instead, the proposed measure simply relies on behavioral information on the network composition according to clearly given political affiliations of politicians.

The measure starts by calculating the distribution of individual followers' connections across Bundestag politicians, that is, the extent to which followers only follow one party or several different parties. This is similar to the measure of users' homophily in following behavior proposed by Colleoni et al. (2014), Siegel et al. (2021), and Halberstam and Knight (2016). For example, Colleoni et al. (2014) first estimate users' political affiliation (republican or democrat) and then continue by calculating the share of their connections going toward their own and other groups. The approach of this study differs from the above studies in the sense that no prior estimate of political ideology is required. Instead, for each user, the share of connections to either of the seven parties is calculated. One connection is represented by one politician she follows.

The key distinction to most existing studies, however, is the level of analysis. As discussed earlier, most studies conduct their analyses on the level of individual users. For example, Colleoni et al. (2014) make statements on the degree of homophily among republican when compared to democratic users. In comparison, this article's analysis is

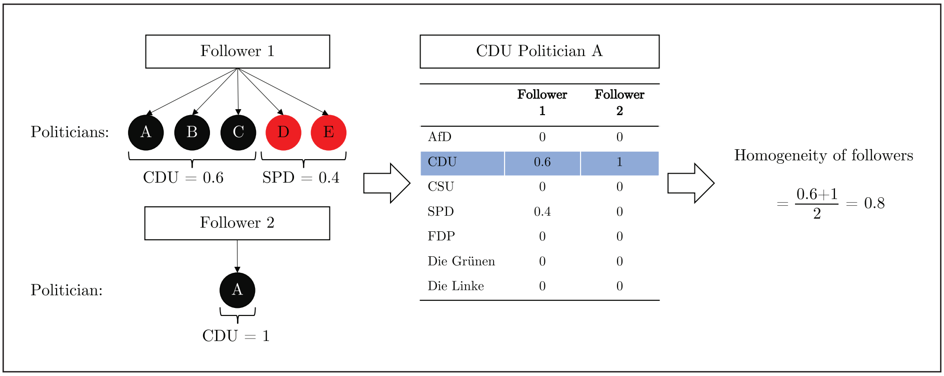

Figure 1. Example for homogeneity measure.

conducted on the politician level. Thus, what is measured is the degree to which a politician's average follower is also interested in other parties' positions or whether she mostly follows the politician's own party. This distinction is vital, as, for example, the average conservative user may differ a lot from the average follower of a given conservative politician. This is the case as particularly active users, that is, those that follow many politicians, gain a higher weight when analyzing follower networks on the politician level and can thus effectively multiply their influence online.

More precisely, the following measure is employed:

$$Homogeneity\ of\ Followers_{j,p=P} = \frac{1}{n_j} \sum_{i=1}^{n_j} \frac{edges_{i,\,p=P}}{\sum_{p=1}^{7} edges_{i,p}} \tag{1}$$

where the $Homogeneity of\ Followers$ of a politician j from party $p = P$ is an average across all n_j followers of politician j. Here, $edges_{i,p}$ describes the number of edges (i.e. connections) of follower i going toward party p, where p is any of seven German parties. The fraction $\frac{edges_{i,p=P}}{\Sigma_{p=1}^{7} edges_{i,p}}$ then describes the share of follower i's edges going toward the politician's party P (as a percentage of all edges going toward any Bundestag politicians). This is calculated for all n_j followers of politician j and averaged across these. The measure thus ranges from close to zero (the average follower follows all other parties' politicians but none of politician j's party aside from j) to one (the average follower only follows politicians from the politician's own party).

For intuition, Figure 1 provides an example. Here, *Follower 1* follows five politicians (A, B, C, D, E) from two parties: the SPD (D and E), and the CDU (A, B, C). Her share of connections going to either party is 0.4 and 0.6, respectively. Another user, *Follower 2*, only follows *Politician A* from the CDU. Moving to *Politician A*, she is followed by both followers and nobody else. As a result, her homogeneity measure is equal to the average share of her followers' connections going to her party: $\frac{0.6+1}{2} = 0.8$. This suggests that the politician's average follower follows the politician's own party with 80% of her connections across all Bundestag politicians.[2]

The approach presented earlier thus allows for the analysis of networks from the perspectives of both politicians and their followers: the composition of accounts that users follow and the compositions of accounts that follow a politician, containing people who may react to their posts and might thus influence what they post. This is discussed in the following section. At the same time, as discussed earlier, the approach does not require the estimation of individuals' ideological positions. Instead, it only relies upon observable following behavior.

Follower types

The measure proposed earlier investigates the average follower of each politician by drawing upon the share of connections to the politician's own party. This effectively gives more weight to profiles that follow multiple politicians. For example, *Follower 1* in Figure 1 affects the homogeneity measures of five politicians, while *Follower 2* only affects the measure of a single one. As discussed earlier, this constitutes a large difference in comparison to other studies that give equal weight to all users (e.g. Barberá, 2015; Colleoni et al., 2014). These may thus may overlook the extend to which single users can multiply their influence online.

To gain a better understanding for the degree to which users do so and for the resulting differences in follower homogeneity, users are categorized by type. This is done based upon the adjacency table that includes both the number and the share of connections going to each party. Four groups of followers are defined that mechanically affect the measure above and may thus explain differences in follower homogeneity: Starting with *Strongly committed partisan* followers, these are defined as all Twitter users who follow at least four politicians, with more than 80% of those they follow belonging to a single party. The interpretation as partisans is in line with many previous studies that infer ideological preferences from following behavior across prominent political accounts (see discussion of Colleoni et al., 2014, above). Similarly, *weakly committed partisan* followers are those that do not fall into the category of strongly committed partisans, while they follow at least two politicians with more than 50% of their connections going to a specific party. Both partisan groups are thus defined by party, for example, there are AfD and CDU partisans. While having many partisan followers leads to higher levels of follower homogeneity, followers interested in multiple parties' politicians decrease the measure. Here, *broad interest* followers are defined to have less than 50% of their edges going to a single party, that is, these also follow more than two politicians. Finally, users that only follow a single politician are hardly considered to be partisans. However, these would still have all edges going to a single party and thus drive up follower homogeneity. To account for this, *one-person* followers are defined as users that only follow a single politician in the Bundestag. All groups are mutually exclusive, that is, users can only belong to one group. Looking at the example in Figure 1, Follower 1 is a weakly committed CDU partisan while Follower 2 is a one-person follower.

After categorizing followers, the share of each group among each politician's followers is calculated to both investigate to what extend follower homogeneity is related to higher shares of each group and to compare follower groups' shares across politicians and parties.

Robustness: politician and account characteristics

Other variables than party membership may explain differences in follower homogeneity and partisan shares between politicians. While controlling for potential confounders does not fully rule out omitted variable bias, it can help to reduce it and allows for the analysis of differences in the outcome variables along other dimensions than party membership. To do so, this study performs a linear regression (OLS) with heteroskedasticity-robust standard errors on the level of politicians' Twitter accounts. The dependent variable is either provided by the politician's follower homogeneity measure or her share of partisan followers. The main independent variables of interest are dummies indicating party membership. In addition, a number of controls regarding politician characteristics and their Twitter accounts are included.

First, lesser known politicians, those new to Twitter, and thus, more generally, those with a smaller number of followers are likely to attract a more homogeneous audience and more partisan followers. Chen et al. (2014) investigate 250,000 US Twitter users that openly state their religion and analyze their religious in- and out-group connections. They show that more popular accounts attract more out-group followers, or as they nicely put it: "the pope is not a scaled up bishop" (Chen et al., 2014: 546). Closely related, but in a more general setup, Grabowicz et al. (2016) compare followers of similar expert accounts. They show that those of smaller accounts tend to be more similar to the account they follow. Furthermore, these are more likely to be experts themselves. Moving the results to this study's context, it is expected that smaller and more recently created accounts are more likely to be discovered by partisans, while larger accounts, such as those of prominent and longer serving politicians, have a higher share of out-group followers, such as broad interest followers and thus a lower homogeneity measure. To account for this, controls for the (log.) number of followers, the politician's account age, and whether she was re-elected in 2017 are included.

Second, users' attention is limited, that is, less politically interested users may shy away from politicians that tweet very frequently. This is in line with empirical results showing that a high tweet rate is predictive of unfollow behavior (Kwak et al., 2011; Maity et al., 2018). Very frequent tweeters may thus attract more politically interested followers. To account for this, the average number of tweets per year is included as a control.

Third, politicians' own characteristics may shape their behavior and success on Twitter as well as their follower structure. The literature suggests that men and women use Twitter differently, in particular, when discussing political issues (e.g. Hu et al., 2021). Similarly, politicians' age and education may play a role, as these have been suggested to affect other dimensions of their Twitter outcomes, such as the number of followers (Keller and Kleinen-von Königslöw, 2018). To control for this, politicians' age and gender are included as controls. Furthermore, as a proxy for a politician's education, a dummy indicating whether she has a PhD is included.

Fourth, controls for politicians' home regions are included to account for spatial differences in the respective measures. Despite social media's global reach, its geography has been shown to have a strong local dimension, with many ties formed between users at a short geographic distance (Takhteyev et al., 2012). Given the spatial variation regarding antidemocratic and populist attitudes across Germany (Küpper et al., 2020), politicians from regions with stronger populist attitudes may be more likely to attract partisans. To control for this, dummies for South and East Germany are included, while North–West Germany is the baseline category.[3]

Finally, Germany has a mixed-member electoral system, that is, politicians can either enter the Bundestag through a direct vote by their local constituency (first-past-the-post) or through party lists. Constituency candidates have been shown to more strongly focus on topics other than those of their party, on issues of their constituency, and to run more personalized campaigns (Gschwend and Zittel, 2015; Zittel and Gschwend, 2008). Given this in combination with Twitter's local nature (Takhteyev et al., 2012), the politician's type may affect her follower structure. The direction of the effect is unclear though. As a result, a control for constituency candidates is added.

The data underlying this study consist of the population of all German politicians on Twitter and their follower networks. To interpret confidence intervals of frequentist linear regressions, one has to view the resulting data as one draw from a "superpopulation" of politicians and their follower networks on Twitter. This appears appropriate in the given setup: First, while the data were gathered in October 2018, inference is done for a longer time-period, that is, the data represent one draw from the population of politicians and their follower networks over the past years. Second, unless a fully deterministic world is assumed, the data generating process, including decisions of users to follow politicians and of politicians to join Twitter, is driven by stochastic processes that affect both variables included in the model and those part of the error term. Thus, the data at hand are just one draw from a superpopulation of possible realizations of politicians and their networks on Twitter at the time of data collection (see Broscheid and Gschwend, 2003, for a discussion). Nevertheless, Bayesian Inference has been suggested as an alternative, given the differing interpretation of standard errors (Berk et al., 1995; Broscheid and Gschwend, 2005; Western and Jackman, 1994) while confidence intervals of frequentist regression models are interpreted as covering the true coefficient 95% of the time in repeated samples, Bayesian credible intervals quantify the certainty a researcher has, that is, these provide an interval in which the coefficient of the given data lies with 95% certainty. Thus, they do not require a theoretical "superpopulation." In addition to frequentist regressions, the main regression are repeated using Bayesian linear regressions with noninformative priors, assigning equal weight to all possible coefficient sizes (see regression tables for model specifications).

Robustness: bots

One important concern regarding robustness and the results' interpretation is the presence of bots, in particular, if these are unevenly distributed across parties. While Gallwitz and Kreil (2021) argue that the role of "social bots" on social media is vastly overestimated, other studies' results suggest that right-wing parties may particularly profit from these (Neudert et al., 2017). For example, Silva and Proksch (2021) show that when Twitter deleted millions of bots in July 2017, radical right parties' European members of parliament (MPs) lost the most followers, with an average loss of 5% of followers. One advantage in this context is that the data were collected 3 months after the purge. However, bots remain a concern.

At the time of data collection, commonly used tools, such as *botometer*, did not have the capacity to process large data sets with millions of profiles. Furthermore, they have been shown to work relatively poorly, especially outside of English speaking contexts (Feng et al., 2021; Rauchfleisch and Kaiser, 2020). Instead, to assess the data with respect to (w.r.t.) the presence of bots, a machine learning estimator is trained to distinguish bots from other profiles. This is built on Twibot-20, which currently provides the largest available annotated bot-detection data set, according to the authors (Feng et al., 2021). Compared to existing methods on bot-detection, the applied algorithm reaches high levels of accuracy, including compared to *botometer*. The estimator can correctly classify around 81% of profiles in the training data (see Appendix for details on the estimator, including a discussion of limitations and comparisons to existing algorithms).

To ensure that bots do not drive the results, two approaches are taken. First, the regressions including the controls above are re-calculated while excluding followers labeled as bots. Second, the baseline regressions are re-run while including controls for the estimated share of bots among followers and the average probability that a politician's follower is a bot.

As an additional precaution, all regressions control for the average follower's account age. The reasoning behind this is that Twitter frequently deletes bots (e.g. Silva and Proksch, 2021). Thus, bot accounts are expected to have a lower survival rate than other accounts and are thus younger. While this is a lower level control compared to those above, it is clearly imperfect as well, given that it may also reflect other mechanisms, such as differences in political preferences/extremism across Twitter cohorts or learning effects when following politicians and other accounts. However, average follower account age also represents a measure for the length of users' exposure to Twitter. Here, Jürgens and Stark (2022) suggest that longer exposure to Twitter is associated with a decrease in users' exposure diversity, which would translate into higher levels of homogeneity in this context.

Data: politicians and their followers on Twitter

Information on all followers of the members of the German national parliament ("Bundestag") is gathered by combining data from several sources. First, all politicians' names and their personal information is provided by the German electoral management body, the Bundeswahlleiter (2017b), and updated using information on politicians that left or joined the Bundestag since the election in 2017 (Deutscher Bundestag, 2017). Overall, information on 707 MPs is gathered.[4] Furthermore, politicians' Twitter accounts are obtained using Twitter lists, provided by several party accounts, which contain the accounts of MPs of each party present in the Bundestag. These profiles are verified using politicians' and party websites and linked to politicians' personal data, as provided by Bundeswahlleiter (2017b). Furthermore, aside from lists, all names of politicians without an account are searched for using Twitter's search engine.[5] The encountered profiles are again verified by checking politicians' and party websites.

Table 1. Summary statistics of adjacency table of all followers.

(1) Variable	(2) Avg.	(3) *SD*	(4) Min	(5) Max
AfD politicians followed	0.146	1.518	0	69
CDU politicians followed	0.441	1.523	0	83
CSU politicians followed	0.089	0.409	0	19
SPD politicians followed	1.084	2.163	0	104
FDP politicians followed	0.258	1.067	0	67
Grüne politicians followed	0.405	1.705	0	63
Die Linke politicians followed	0.635	1.488	0	57
Follows any AfD politician	0.051	0.220	0	1
Follows any CDU politician	0.249	0.432	0	1
Follows any CSU politician	0.071	0.257	0	1
Follows any SPD politician	0.636	0.481	0	1
Follows any FDP politician	0.193	0.394	0	1
Follows any Grüne politician	0.192	0.394	0	1
Follows any Die Linke politician	0.361	0.480	0	1

CDU: christian democratic union; CSU: christian social union; FDP: free democratic party; SPD: social democrats.
All observations are on the follower level with a total of 1.78 million unique followers.

Next, information on the latest tweet of each politician is obtained, in order to identify inactive profiles, given that these could bias the results. The data collection took place on 10 and 11 October 2018. MPs that did not tweet anything since 1 January 2018 are removed, a total of 33 politicians of which 11 had never posted anything (publicly). The resulting data show that 462 or 65.35% of politicians are identified as active on Twitter. Compared to the 261 or 41.36% of parliamentary Twitter users found by German news outlet *Spiegel Online* in the prior legislature period (2013–2017),[6] this represents an increase of 58% in the Twitter adoption rate. Figure 4 (in Appendix) shows politicians' distribution across parties, showing that five of the seven parties have around 60–80 politicians on Twitter. There are two outliers, with the SPD having more than 100 politicians while the Bavarian CSU (christian social union) has fewer than 20 (note that the CDU and CSU form one parliamentary group and do not compete in elections). Finally, Table 4 (in Appendix) describes the selection of politicians into Twitter, given observable characteristics.

For each politician, meta data on all followers as of October 2018 are gathered, using the rtweet package in R (Kearney, 2018; R Core Team, 2018). In total, 5,436,552 follower-politician edges connect 1,775,789 unique users to MPs across all parties, that is, each account follows three politicians on average. Drawing upon these data, an adjacency table is created as summarized in Table 1. For each of the 1.78 million users, the table describes whether she is connected to a specific party and how many connections she has to that party. More specifically, it shows that she follows 1.08 SPD politicians, but only 0.44 CDU politicians. At the same time, 24.9% of the 1.78 million users follow any CDU and 63.6% any SPD politician. As will be discussed later, this is driven by a couple of well-known politicians, such as Martin Schulz.

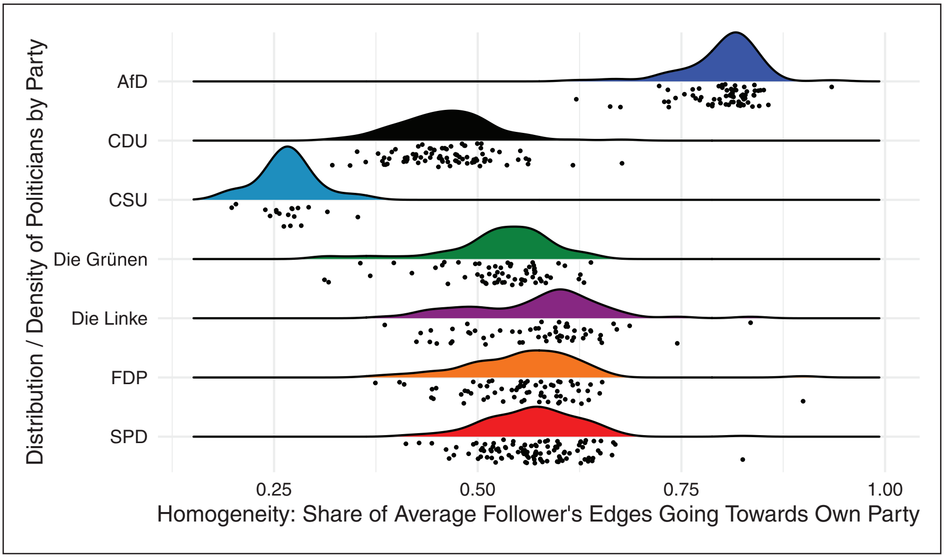

Figure 2. Density functions of politicians' homogeneity measure, by party.
The Figure shows density estimates of politicians' homogeneity measure by party, computed as suggested in Equation 1. In addition, the black dots indicate the homogeneity measure of individual observations. All density plots in this article are created using the ggridges package in R (Wilke, 2022).

Results

Follower homogeneity

Starting with the homogeneity measure from Equation 1, Figure 2 shows the density functions of politicians' follower homogeneity across all parties' politicians. Here, most parties' functions are rather broad and peak around 0.5 to 0.6. Here, a value of 0.5 means that the average follower of the given politician follows the politician's own party's politicians with 50% of her connections going toward any Bundestag politicians. The results thus suggest that the audiences of most politicians also follow other parties, suggesting that there is no overall segregation of follower networks. However, the results still suggest that users are more homogeneous than would be expected in the absence of any party preferences, that is, confirming H1. To see this, consider the case where users are homogeneous, that is, all behave the same (the effect of heterogeneous user types is analyzed below). Like the average user, these follow three politicians, which they choose by first randomly picking a party to follow (with equal probability across parties) and then randomly picking a politician of that party. If such a user follows a given politician, the expected share of the user's other two connections going to the politician's own party is $\frac{2}{7}$. The expected homogeneity measure of a given politician is thus $\frac{1+\frac{2}{7}}{3} = 0.429$.

Almost all parties and politicians' homogeneity measures are well above this value, as shown in Figure 2.

Table 2. Groups of followers in percent of parties' total (unique) followers.

Party	Strongly Committed Partisans	Weakly Committed Partisans	One-Person Flwrs	Broad Interest Flwrs	Total Flwrs (thsd)
AfD	6.97	18.41	34.01	24.07	91
CDU	0.42	5.22	18.35	41.28	442
CSU	0.03	1.03	11.35	54.07	127
SPD	0.57	7.81	50.86	19.30	1129
FDP	0.64	2.39	18.96	44.98	342
Grüne	1.20	6.59	18.73	41.53	342
Linke	1.16	13.35	26.62	30.27	641

CDU: christian democratic union; CSU: christian social union; SPD: social democrats; FDP: free democratic party.

Although, while most parties' density functions peak at around 0.5 to 0.6, two deviations can be observed. First, CDU and CSU politicians are embedded in the least homogeneous networks. This difference to other parties however largely disappears when treating CDU and CSU (who form one parliamentary group) as one party, as shown in Figure 5 (in Appendix). The lower levels of network homogeneity of the CDU/CSU is thus largely driven by politicians' followers following both parties' politicians. The AfD—a right-wing populist party—constitutes a more striking outlier. It stands out, given its very narrow density function that peaks around 0.8. Looking at the average AfD politician, this means that around 80% of her average follower's connections across all Bundestag politicians go toward her own party, while only 20% go to other parties. On the other extreme of the political spectrum, Die Linke is not an outlier. Thus, H2 can only be confirmed for the right-wing AfD.

Categorizing followers

To gain a better understanding for the underlying reasons of the above findings, the adjacency table is drawn upon to categorize followers by type. Table 2 shows the share of followers of each party that belong to the categories of followers defined in Section "Follower types". Around 7% of the AfD's followers can be considered to be strongly committed partisans of the AfD, followed by Die Grünen with 1.2%. Also, when compared to other parties, the AfD has a much higher share of weakly committed partisan followers, although it is followed relatively closely by Die Linke. Regarding one-person followers, the SPD has particularly many, with 50% of its unique followers falling into this category, followed by the AfD with 34%. Finally, the CSU has the highest percentage of broad interest followers, while only 11% of its followers are in the one-person category.

To investigate which of these groups drive the different levels of network homogeneity shown in Figure 2, the groups of followers above are summarized by politician to calculate the share of each group in each politician's follower network. Drawing upon density functions of every single party's politicians, these data are summarized in several graphs. Figure 7 (in Appendix) visualizes the percentage of broad interest followers in each politician's follower network. Approximately, 15% of the followers of most AfD politicians fall into this category. This value is very low when compared to other parties, whose densities peak at around 25% to 40%, while CDU and CSU peak at 50% and 60%. Die Linke also has a relatively low share of broad followers, though its density function is much closer to that of other parties. H3a can thus be confirmed, in particular for the AfD. These observations can partially explain why the CDU and CSU have rather low homogeneity values, while those of the AfD are higher. Second, Figure 8 (in Appendix) summarizes the data of one-person followers. Most parties' politicians have very few of these. AfD politicians' high homogeneity measure can thus not be explained by this group. The discrepancy between this graph and the much higher values observed in Table 2 is explained by two factors. First, one-person followers have a much lower weight when aggregating the data on the politician level. Second, they tend to follow the most popular politicians. For example, 67% of Martin Schulz's followers fall into this category. Given that he was the SPD's and Bundestag's most popular politician on Twitter, this drives up the percentage of SPD followers in this category, while not affecting other SPD politicians' values in Figure 2. To show that one-person followers do not affect the general results of the homogeneity measure, Figure 6 (in Appendix) compares the homogeneity measure when including and excluding one-person followers. The results show that excluding one-person followers does not fundamentally change politicians' (and thus parties') positions on the homogeneity measure.

Turning to weakly committed partisan followers, Figure 9 (in Appendix) shows that the CSU has few of these, as already suggested by Table 2. However, all other parties' density functions peak at around 20% to 26%. The CSU's low homogeneity can thus partially be explained by the absence of weakly committed followers, while the AfD's homogeneity cannot be explained by a high share of these. Finally, Figure 3 offers an explanation for AfD politicians' high levels of network homogeneity, showing that strongly committed partisan followers constitute between 55% to 75% of the followers of most AfD politicians, while they account for a much lower share in other parties. This confirms H3b, though the results do not translate to the extreme party on the left of the political spectrum. While the distribution of Die Linke is shifted to the right in comparison to other parties, such as the SPD and CDU, it remains below that of the FDP and generally not far off of other parties. At the same time, CDU and CSU politicians are barely followed by strongly committed partisan followers.

Including controls

The results presented earlier may potentially be driven by other confounding variables. To test this, the homogeneity measure and the share of strongly committed partisan followers are regressed on a number of controls.

In Table 3, both politician and account characteristics are regressed on the homogeneity measure. Furthermore, Table 10 (in Appendix) reports the results using Bayesian linear regressions instead. In Column (1), the measure is compared across parties, where the CDU provides the baseline category. It suggests that the average AfD

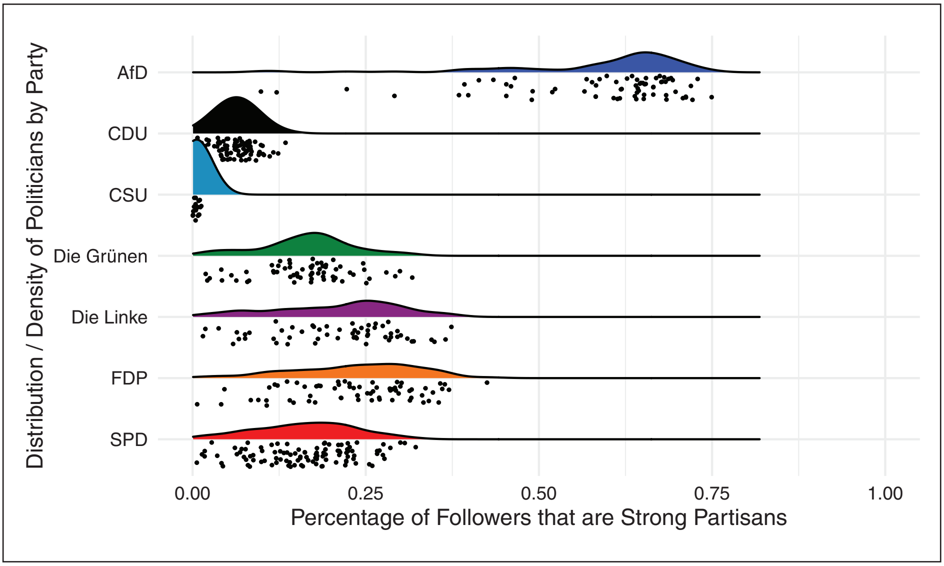

Figure 3. Percentage of strongly committed Partisan followers among politicians' followers across parties.
The Figure shows density estimates across parties' politicians' share of strong partisan followers. In addition, the black dots indicate the homogeneity measure of individual observations.

politician's average follower follows the party with 80.1% of her connections across the Bundestag. This value ranges between 26.7 (CSU) and 57.3% (Die Linke) across other parties. Introducing Twitter account characteristics in Column (2) and politician characteristics in Column (3) does not substantially change the results with respect to parties' coefficients. Moving to the controls, as expected, politicians with a high tweet frequency have a higher follower homogeneity. Furthermore, following Chen et al. (2014) and Grabowicz et al. (2016), more prominent politicians and accounts have a lower follower homogeneity. This is suggested by the negative coefficients of the number of followers and the incumbency dummy, though the coefficient of the politician's account age is insignificant. The relationship between the number of followers and their homogeneity measure is further visualized in Figure 10 (in Appendix) and holds across all parties. Finally, accounts with a higher average follower account age have a slightly increased homogeneity measure. This is first evidence that the results are unlikely to be driven by bots and/or that longer exposure to Twitter decreases users' diversity exposure. All other controls are insignificant, suggesting that there is a little heterogeneity in the outcome measure with respect to politicians' type, characteristics, or geography. In total, introducing controls does not substantially alter the results on the party level. More specifically, the AfD remains the party with the most homogeneous followers, followed by Die Linke and FDP.

Moving to the share of partisan followers, Tables 5 and 11 (both in Appendix) repeat the regressions above with the share of strong partisans as the dependent variable. Again,

Table 3. Regression: user homogeneity at politician level.

	Dependent variable: Homogeneity of followers		
	(1) Parties	(2) +Twitter characteristics	(3) +Politician characteristics
AfD	0.339***	0.375***	0.351***
	(0.009)	(0.011)	(0.020)
CSU	–0.195***	–0.187***	–0.193***
	(0.010)	(0.012)	(0.014)
FDP	0.094***	0.078***	0.060***
	(0.011)	(0.010)	(0.018)
SPD	0.107***	0.114***	0.122***
	(0.009)	(0.007)	(0.010)
Die Grünen	0.064***	0.059***	0.066***
	(0.011)	(0.010)	(0.016)
Die Linke	0.111***	0.127***	0.127***
	(0.013)	(0.011)	(0.018)
Log. followers		–0.015***	–0.014**
		(0.006)	(0.006)
Log. of average statuses per year		0.006**	0.007**
		(0.003)	(0.003)
Average follower account age (years)		0.024***	0.020***
		(0.006)	(0.006)
Politician account age (years)		0.000	0.001
		(0.002)	(0.002)
Female			–0.005
			(0.006)
Politician age			–0.000
			(0.000)
Re-elected 2017			–0.040***
			(0.013)
District candidate PhD			0.011
			(0.012)
			0.001
			(0.008)
South German			0.006
			(0.007)
East German			–0.001
			(0.007)
Constant	0.462***	0.427***	0.464***
	(0.007)	(0.045)	(0.042)
Observations	462	462	462
R^2	0.775	0.826	0.838

FDP: free democratic party; CDU: Christian democratic union; CSU: Christian social union; SPD: social democrats; AfD: alternative for Germany, Die Grünen: the greens, Die Linke: the left.
Standard errors in parentheses.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.
The unit of observation is a German Bundestag Politician's Twitter profile. The dependent variable is the homogeneity of followers, as measured using Equation 1. All regression are estimated using OLS with heteroskedasticity-robust standard errors.

the AfD stands out, with partisans making up 58.8% of the average politician's followers, a share that is far above that of any other party, whose values range between 0.6% (CSU) and 23.5% (FDP). Note here that, given the long tail of the distribution visible in Figure 3, the AfD's average is well below its median, which stands at 64%. Overall, the results are virtually unchanged after introducing Twitter and politician characteristics in Columns (2) and (3) and suggest strong party differences with respect to the share of partisan followers. In line with the literature, controls suggest a negative relationship between partisan shares and the number of followers, though the effects of incumbency and account age are insignificant. Similarly, accounts with a high tweet frequency have more partisan followers. Again, all other controls are insignificant, suggesting little heterogeneity across other dimensions of politicians' characteristics.

The definition of strongly committed partisans clearly leaves room for heterogeneity given its arbitrary definition of following at least four politicians with 80% of the connections going to a specific party. Thus, some parties' strong partisans may be less partisan than others. Table 6 (in Appendix) compares strong partisans across parties. The count of followed politicians from one's "own" party is, on average, highest for the AfD (mean: 15.21, median: 9) and lowest for the CSU (mean: 5.05, median: 5). The average partisan of the AfD is thus more partisan than that of other parties. In fact, while increasing the threshold of followed politicians would decrease the number of partisans across all parties, this would least affect AfD partisans. More specifically, the cumulative distribution over own-party connections among partisans is such that the share of profiles that remain defined as strongly committed is highest at any threshold chosen. The AfD's comparatively high level of partisanship will thus remain high, independent of the threshold.

Robustness: bots

Finally, as discussed in Section "Robustness: Bots," bots may be a concern, especially if these are disproportionately present among certain user groups and party-followers. Applying the mentioned random forest algorithm to all, 1.78 million users show that this is indeed the case. Here, Table 7 (in Appendix) presents the shares of bots across user groups presented in Table 2. As shown, the estimated share of bots among strong partisans is generally higher than across other groups. However, the AfD's share of bots within this group is lowest at 18%. In total, 18% of profiles are classified as bots. These make up around 22% of the 5.4 million follower-politician edges.

Taking this into consideration, Tables 8 and 9 (both in Appendix) repeat the regressions above while either excluding bots from politicians' followers in Columns (1) to (3) or controlling for the estimated share of bot followers and the average follower's probability to be a bot in Column (4). Starting with the homogeneity measure, Table 8 (in Appendix) suggest that bots do not substantially affect the results. The coefficients of all parties remain similar and in the same order regarding their size. Column (4) shows slightly reduced estimates for the AfD (and, to a lesser extend, most other parties), which remains the party with the highest homogeneity measure though. The same conclusions can be drawn when looking at the share of strongly committed partisans in Table 9 (in Appendix), thus suggesting that the article's results are not driven by bots.

Conclusion

The results offer important new insights to the study of political network homogeneity on Twitter. Focusing on politicians and drawing upon the network homogeneity measure presented in this article, it is shown that strong follower homogeneity is not an overall present phenomenon in networks of German MPs. However, differences in homogeneity are found across parties, and politicians of the AfD tend to be embedded in networks that display higher follower homogeneity than politicians of other parties. Diving deeper into the data, it is shown that this homogeneity is largely created by a small group of profiles of around 7% of the party's followers that make up around 55–75% of most politicians' followers.

The results and measures used have the following two main limitations: First, though potential mechanisms are discussed in Section "Network Homogeneity around Politicians' Accounts," the results remain descriptive in the sense that these describe an outcome rather than the underlying mechanisms causing it. Second, the homogeneity measure and the grouping of followers, solely rely on observed follower networks, while ignoring other available data, such as retweet networks and connections between followers. For example, Guerrero-Solé (2018) argue that such information is important to understand homophilic online behavior. However, the simplicity of the homogeneity measure employed and the way in which follower groups are defined, are also advantageous, given that they do not rely on assumptions and estimation strategies, such as sentiment analysis or the classification of tweets and users to political ideologies. They thus provide a simple way to better understand network structures that may constitute elements of echo chambers.

Subject to the above limitations, the findings have several important implications: First, adding to the empirical literature showing mixed evidence on echo chambers, the results show that political homogeneity is not present across all politicians' networks, but concentrated among politicians of the right-wing populist AfD. It is shown that even a small group of very active profiles can create homogeneous follower networks. Here, Dubois and Blank (2018) and Eady et al. (2019) show that only a few individuals can be considered to predominantly follow like-minded others or news. This article thus adds the important argument that these small groups can multiply their influence online and suggests that researchers are well advised not to underestimate the impact of small groups in this context. This is particularly important given evidence on Twitter that politicians are "more likely to follow, than to lead" (Barberá et al., 2019: 883) when setting agendas, and that they are more responsive to their supporters than the public. The findings thus open an important path for future research, namely the investigation of the impact of small but very active online groups on political outcomes and online interactions in general. For instance, if, as suggested by Barberá et al. (2019) and Stier et al. (2018), politicians are reactive to their followers, and if these tend to be very partisan, they may propagate and support more extreme views and policies.

In addition, the results may help to explain the gap between the strength of echo chambers as a narrative to explain increasing polarization (e.g. BBC, 2019; Sunstein, 2001) and the limited evidence on their existence, building on individual-level analyses with few users being embedded in homogeneous networks (e.g. Barberá et al., 2015;

Dubois and Blank, 2018). If small groups of users can create very homogeneous follower networks, while most individuals cannot be considered to pursue this behavior, then these may be perceived as insignificant in studies that give equal weights to users. Here, only few individuals may be identified as partisan or being embedded in homogeneous networks. As shown in this study, such individuals can however multiply their influence online and in fact create homogeneous follower networks around politicians' profiles. This adds a new perspective to the study of echo chambers, namely of a few followers creating homogeneous networks that potentially echo extreme politicians' messages.

Researchers may apply the proposed measures to other groups of profiles, additional online platforms, and different time periods. By focusing on well-defined and comparable profiles across a number of groups (such as parties or types of media), these provide a straightforward approach to measure network homogeneity. A major advantage of investigating politicians is that the ideological identification is usually clearly given. It would thus be of interest to apply the measure to other multi-party democracies. A first indication of where high levels of follower homogeneity and/or partisanship are to be expected is provided by Urman (2020), who analyzes and compares the audience overlap between different parties' official Twitter pages across multi-party countries. Van Vliet et al. (2021) and Praet et al. (2021), who study within-parliament Twitter relations across multiple countries, may provide additional starting points. In addition, the approach may also be applied to followers of different media types. Here, existing studies provide evidence that the audience of single media outlets, such as Fox News or NBC, have large overlapping audiences with other media outlets (Webster and Ksiazek, 2012). The picture may however differ, when first grouping outlets, for example, into right, center, and left media and then calculating their follower/audience homogeneity as well as the shares of partisans among followers. In this context, Gentzkow and Shapiro (2011) show that more extreme outlets have higher shares of republican/liberal visitors than moderate ones. Building on this and the results of this study, it is expected that smaller and more extreme outlets are likely to attract very homogeneous audiences, driven by fewer cross-ideological (i.e. broadly interested) followers and a higher number of partisans. However, additional research is required to study this in more detail, given that networks surrounding politicians may be different from those around other types of accounts.

Acknowledgements

The author thanks Olga Slivkó for her guidance, Franziska Scholl for her continuous support and feedback, Lion Behrens and Frederic Gerdon for their insights, and Antonio Ciccone for his continuous support. He also thanks three anonymous reviewers for their insightful feedback and participants of internal seminars at the Leibniz Centre for European Economic Research (ZEW), the University of Mannheim, and the Social Science Research Seminar of the German National Academic Foundation. He is further grateful for feedback from Thomas Gschwend, Vlad Achimescu, Florian Keusch, Sonja Malich, Patrick Schenk, and Alexander Wenz.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support by the German Research Foundation (DFG) through

CRC TR 224 (project A04) and by the University of Mannheim's Graduate School of Economic and Social Sciences (GESS) is gratefully acknowledged.

ORCID iD

Felix Rusche https://orcid.org/0000-0002-3593-583X

Supplemental material

Supplemental material for this article is available online.

Notes

1. But see, for example, population surveys on ideological profiles of media outlets (Pew Research Center, 2018).
2. Remark: the presented measure offers a clear interpretation that is the same across politicians. Furthermore, its expected value is equal across parties, assuming users behave the same and randomly pick parties to follow, with equal probability across parties (see Section "Follower homogeneity" for this thought experiment).

   An alternative "counterfactual" is one where users randomly allocate their connections between politicians, with equal probability across these. In this case, one may weight the number of edges by the number of politicians in a given party, to ensure equal expected values across parties in the absence of unequal distributions of network homogeneity. Here, connections to politicians of parties with fewer politicians on Twitter would receive a higher weight in the calculation of the homogeneity measure. However, this has the disadvantage that the measure, to some degree, loses its intuitive interpretation and is harder to compare across different parties' politicians. Given this, the fact that most parties have a similar number of politicians on Twitter, and the empirical finding that those with the highest number of politicians, CDU (+CSU) and SPD, have similar homogeneity measures as other parties in the political center (like FDP and Die Grünen), the chosen measure appears most appropriate in this context. In other settings, one may, however, consider weighting the connections.
3. "East German" indicates that a politician comes from a state that formerly belonged to the German Democratic Republic. "South German" indicates that the candidate is either from Bavaria or Baden-Württemberg. Politicians are assigned to a state according to the data provided by the Bundeswahlleiter (2017b).
4. As the analysis is on the party level and the politicians without party affiliation do not play an important role in German parliament, two politicians that left the AfD shortly after the 2017 election are excluded.
5. Each politician's name was searched and the first nine results in the "people" category were checked.
6. See https://www.spiegel.de/politik/deutschland/bundestagsabgeordnete-auf-twitter-wer-wie-viel-schreibt-und-mit-wem-a-1041402.html (accessed 12 May 2021).

References

Ausserhofer J and Maireder A (2013) National politics on Twitter. *Information, Communication & Society* 16(3): 291–314.

Barberá P (2015) Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis* 23(1): 76–91.

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

Barberá P, Jost JT, Nagler J, et al. (2015) Tweeting from left to right: is online political communication more than an echo chamber? *Psychological Science* 26(10): 1531–1542.

BBC (2019) The myth of the online echo chamber. Available at: https://web.archive.org/web/20220310122758/https://www.bbc.com/future/article/20180416-the-myth-of-the-online-echo-chamber (accessed 7 December 2020).

Berk RA, Western B and Weiss RE (1995) Statistical inference for apparent populations. *Sociological Methodology* 25: 421–458.

Bojanowski M and Corten R (2014) Measuring segregation in social networks. *Social Networks* 39: 14–32.

Bond R and Messing S (2015) Quantifying social media's political space: estimating ideology from publicly revealed preferences on Facebook. *American Political Science Review* 109(1): 62–78.

Boutyline A and Willer R (2017) The social structure of political echo chambers: variation in ideological homophily in online networks. *Political Psychology* 38(3): 551–569.

Breiman L (2001) Random forests. *Machine Learning* 45(1): 5–32.

Bright J (2018) Explaining the emergence of political fragmentation on social media: the role of ideology and extremism. *Journal of Computer-Mediated Communication* 23(1): 17–33.

Broscheid A and Gschwend T (2003) *Augäpfel, murmeltiere und bayes: Zur auswertung stochastischer daten aus vollerhebungen*. Technical report 03/7, MPIfG Working Paper. Cologne: MPIfG.

Broscheid A and Gschwend T (2005) Zur statistischen analyse von vollerhebungen. *Politische Vierteljahresschrift* 46(1): O16–O26.

Bundeswahlleiter (2017a) Bundestagswahl 2017: Ergebnisse. Technical Report, Bundeswahlleiter, Berlin.

Bundeswahlleiter (2017b) Endgültig gewählte Bewerberinnen und Bewerber. Technical Report, Bundeswahlleiter, Berlin.

Chen L, Weber I and Okulicz-Kozaryn A (2014) U.S. religious landscape on Twitter. In: Aiello LM and McFarland D (eds) *Social Informatics: 6th International Conference, Socinfo 2014*, Barcelona, Spain, November 11-13, 2014. Proceedings (Lecture Notes in Computer Science). Cham: Springer, pp. 544–560.

Chiocchetti P (2016) The German radical left: a success story. In: *The Radical Left Party Family in Western Europe, 1989-2015*. London: Routledge, pp. 81-122. https://www.taylorfrancis.com/chapters/mono/10.4324/9781315622057-13/german-radical-left-success-story-paolo-chiocchetti?context=ubx&refId=33659d1f-cfbf-4b89-ac93-5f43027843f2.

Cinelli M, Morales GDF, Galeazzi A, et al. (2021) The echo chamber effect on social media. *Proceedings of the National Academy of Sciences* 118(9): e2023301118.

Colleoni E, Rozza A and Arvidsson A (2014) Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication* 64(2): 317–332.

Deutscher Bundestag (2017) Vorzeitige Beendigung der Mitgliedschaft. Technical report, Deutscher Bundestag, Berlin.

Dubois E and Blank G (2018) The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society* 21(5): 729–745.

Dvir-Gvirsman S (2017) Media audience homophily: Partisan websites, audience identity and polarization processes. *New Media & Society* 19(7): 1072–1091.

Eady G, Nagler J, Guess A, et al. (2019) How many people live in political bubbles on social media? Evidence from linked survey and Twitter data. *SAGE Open* 9(1): 1–21.

Fawzi N, Obermaier M and Reinemann C (2017) Germany: is the populism laggard catching up? In: Aalberg T, Esser F, Reinemann C, et al. (eds) *Populist Political Communication in Europe*. New York: Taylor & Francis, pp. 111–126.

Feng S, Wan H, Wang N, et al. (2021) *Twibot-20: A Comprehensive Twitter Bot Detection Benchmark*. New York: Association for Computing Machinery.

Flaxman S, Goel S and Rao JM (2016) Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly* 80(S1): 298–320.

Gallwitz F and Kreil M (2021) The rise and fall of "social bot" research. SSRN scholarly paper, Social Science Research Network. Available at: https://papers.ssrn.com/abstract=2356199

Gentzkow M and Shapiro JM (2011) Ideological segregation online and offline. *The Quarterly Journal of Economics* 126(4): 1799–1839.

Gilardi F, Gessler T, Kubli M, et al. (2021) Social media and political agenda setting. *Political Communication* 39: 39–60.

Grabowicz P, Babaei M, Kulshrestha J, et al. (2016) The road to popularity: the dilution of growing audience on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media* 10: 567–570.

Gschwend T and Zittel T (2015) Do constituency candidates matter in German federal elections? The personal vote as an interactive process. *Electoral Studies* 39: 338–349.

Guerrero-Solé F (2018) Interactive behavior in political discussions on Twitter: politicians, media, and citizens' patterns of interaction in the 2015 and 2016 electoral campaigns in Spain. *Social Media + Society* 4(4): 1–16.

Halberstam Y and Knight B (2016) Homophily, group size, and the diffusion of political information in social networks: evidence from Twitter. *Journal of Public Economics* 143: 73–88.

Hu L, Kearney MW and Frisby CM (2021) Tweeting and retweeting: gender discrepancies in discursive political engagement and influence on Twitter. *Journal of Gender Studies*. Epub ahead of print 25 July. DOI:10.1080/09589236.2021.1995340.

Johnson TJ, Bichard SL and Zhang W (2009) Communication communities or "cyberghettos?" A path analysis model examining factors that explain selective exposure to blogs. *Journal of Computer-Mediated Communication* 15(1): 60–82.

Jungherr A (2014) The logic of political coverage on Twitter: temporal dynamics and content. *Journal of Communication* 64(2): 239–259.

Jungherr A (2016) Twitter use in election campaigns: a systematic literature review. *Journal of Information Technology & Politics* 13(1): 72–91.

Jürgens P and Stark B (2022) Mapping exposure diversity: the divergent effects of algorithmic curation on news consumption. *Journal of Communication*. Epub ahead of print 16 March. DOI: 10.1093/joc/jqac009.

Kearney MW (2018) *Rtweet: Collecting Twitter Data* (R package version 0.6.7). Available at: https://cran.r-project.org/package=rtweet

Keller TR and Kleinen-von Königslöw K (2018) Followers, spread the message! Predicting the success of Swiss politicians on Facebook and Twitter. *Social Media+ Society* 4(1): 1–11.

Koiranen I, Koivula A, Keipi T, et al. (2019) Shared contexts, shared background, shared values–homophily in Finnish parliament members' social networks on Twitter. *Telematics and Informatics* 36: 117–131.

Küpper B, Berghan W, Zick A, et al. (2020) Volkes Stimme—antidemokratische und populistische Einstellungen. In: Zick A and Küpper B (eds) *Die Geforderte Mitte: Rechtsextreme Und Demokratiegefährdende Einstellungen in Deutschland*, vol. 21. Bonn: Friedrich-Ebert-Stiftung, pp. 43–74.

Kwak H, Chun H and Moon S (2011) Fragile online relationship: a first look at unfollow dynamics in Twitter. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1091–1100. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.228.6826&rep=rep1&type=pdf

Lawrence E, Sides J and Farrell H (2010) Self-segregation or deliberation? Blog readership, participation, and polarization in American politics. *Perspectives on Politics* 8(1): 141–157.

Lee JK, Choi J, Kim C, et al. (2014) Social media, network heterogeneity, and opinion polarization. *Journal of Communication* 64(4): 702–722.

Lewandowsky M (2015) Eine rechtspopulistische Protestpartei? Die AfD in der öffentlichen und politikwissenschaftlichen Debatte. *Zeitschrift Für Politikwissenschaft* 25(1): 119–134.

Liaw A and Wiener M (2002) Classification and regression by randomforest. *R News* 2(3): 18–22.

Lührmann A, Düpont N, Higashijima M, et al. (2020) *Varieties of Party Identity and Organization (V-Party) Dataset*, v 1. Available at: https://www.v-dem.net/static/website/img/refs/vparty_codebook.pdf

McPherson M, Smith-Lovin L and Cook JM (2001) Birds of a feather: homophily in social networks. *Annual Review of Sociology* 27(1): 415–444.

Maity SK, Gajula R and Mukherjee A (2018) Why did they# unfollow me? Early detection of follower loss on Twitter. In: *Proceedings of the 2018 ACM conference on supporting groupwork*, pp. 127–131. Available at: https://arxiv.org/abs/1802.05091

Mullainathan S and Shleifer A (2005) The market for news. *American Economic Review* 95(4): 1031–1053.

Neudert L, Kollanyi B and Howard PN (2017) *Junk News and Bots during the German Parliamentary Election: What are German Voters Sharing over Twitter? Data Memo 2017*. Oxford: Project on Computational Propaganda.

Newman N, Fletcher R, Schulz A, et al. (2021) Reuters digital news report 2021. Reuters digital news report. Available at: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf

Pew Research Center (2018) *Datenblatt: Nachrichtenmedien und politische haltungen in deutschland*. Available at: https://www.pewresearch.org/global/fact-sheet/datenblatt-nachrichtenmedien-und-politische-haltungen-in-deutschland/ (accessed 1 April 2021).

Praet S, Martens D and Van Aelst P (2021) Patterns of democracy? social network analysis of parliamentary Twitter networks in 12 countries. *Online Social Networks and Media* 24: 100154.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Rauchfleisch A and Kaiser J (2020) The false positive problem of automatic bot detection in social science research. *PLoS ONE* 15(10): e0241045.

Saalfeld T and Schoen H (2015) Party politics and electoral behaviour. In: Colvin S and Taplin M (eds) *The Routledge Handbook of German Politics & Culture* (Routledge Handbooks). New York: Routledge, pp. 105–118.

Scharkow M, Mangold F, Stier S, et al. (2020) How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences* 117(6): 2761–2763.

Schmitt-Beck R, van Deth JW and Staudt A (2017) Die AfD nach der rechtspopulistischen Wende. *Zeitschrift Für Politikwissenschaft* 27(3): 273–303.

Siegel AA, Nagler J, Bonneau R, et al. (2021) Tweeting beyond tahrir: ideological diversity and political intolerance in Egyptian Twitter networks. *World Politics* 73(2): 243–274.

Silva BC and Proksch SO (2021) Fake it "til you make it: a natural experiment to identify European politicians" benefit from Twitter bots. *American Political Science Review* 115(1): 316–322.

Stier S, Bleier A, Lietz H, et al. (2018) Election campaigning on social media: politicians, audiences, and the mediation of political communication on Facebook and Twitter. *Political Communication* 35(1): 50–74.

Stroud NJ (2018) Selective exposure theories. In: Kenski K and Hall Jamieson K (eds) *The Oxford Handbook of Political Communication*. Oxford: Oxford University Press, pp. 1–19.

Sunstein CR (2001) *Republic.com*. Princeton, NJ: Princeton University Press.

Takhteyev Y, Gruzd A and Wellman B (2012) Geography of Twitter networks. *Social Networks* 34(1): 73–81.

Urman A (2020) Context matters: political polarization on Twitter from a comparative perspective. *Media, Culture & Society* 42(6): 857–879.

Vaccari C, Valeriani A, Barberá P, et al. (2016) Of echo chambers and contrarian clubs: exposure to political disagreement among German and Italian users of Twitter. *Social Media + Society* 2(3): 1–24.

Van Alstyne M and Brynjolfsson E (1997) Electronic communities: global village or cyberbalkans? Working paper. Available at: https://web.mit.edu/marshall/www/papers/CyberBalkans.pdf

Van Alstyne M and Brynjolfsson E (2005) Global Village or Cyber-Balkans? Modeling and measuring the integration of electronic communities. *Management Science* 51(6): 851–868.

Van Vliet L, Törnberg P and Uitermark J (2021) Political systems and political networks: the structure of parliamentarians' retweet networks in 19 countries. *International Journal of Communication* 15: 21.

Varol O, Ferrara E, Davis C, et al. (2017) Online human-bot interactions: detection, estimation, and characterization. In: *Proceedings of the international AAAI conference on web and social media*, vol. 11. Available at: https://arxiv.org/abs/1703.03107

Webster JG and Ksiazek TB (2012) The dynamics of audience fragmentation: public attention in an age of digital media. *Journal of Communication* 62(1): 39–56.

Western B and Jackman S (1994) Bayesian inference for comparative research. *American Political Science Review* 88(2): 412–423.

Wilke CO (2022) *ggridges: Ridgeline Plots in ggplot2* (R package version 0.5.3). Available at: https://cran.r-project.org/web/packages/ggridges/

Yardi S and Boyd D (2010) Dynamic debates: an analysis of group polarization over time on Twitter. *Bulletin of Science, Technology & Society* 30(5): 316–327.

Zittel T and Gschwend T (2008) Individualised constituency campaigns in mixed-member electoral systems: candidates in the 2005 German elections. *West European Politics* 31(5): 978–1003.

Author biography

Felix Rusche is a third-year economics PhD student at the University of Mannheim's Graduate School of Economic and Social Sciences (GESS). He obtained his bachelor's degree in Applied Economics in Osnabrück. During this time, he spent a semester abroad at Nelson Mandela University in Port Elizabeth, South Africa. Following his bachelor's studies, he completed a master's in Economics at the University of Mannheim and spent a semester abroad at Bocconi University in Milan, Italy. His research interests lie in the fields of development, political, and media economics.