The Geometry of Misinformation: Embedding Twitter Networks of Users Who Spread Fake News in Geometrical Opinion Spaces

Pedro Ramaciotti Morales[1,2], **Manon Berriche**[1,2], **Jean-Philippe Cointet**[1]
[1]médialab, Sciences Po, Paris
[2]LPI, Université Paris Cité
pedro.ramaciottimorales@sciencespo.fr, manon.berriche@sciencespo.fr, jeanphilippe.cointet@sciencespo.fr

Abstract

To understand why internet users spread fake news online, many studies have focused on individual drivers, such as cognitive skills, media literacy, or demographics. Recent findings have also shown the role of complex socio-political dynamics, highlighting that political polarization and ideologies are closely linked to a propensity to participate in the dissemination of fake news. Most of the existing empirical studies have focused on the US example by exploiting the self-reported or solicited positioning of users on a dichotomous scale opposing liberals with conservatives. Yet, left-right polarization alone is insufficient to study socio-political dynamics when considering non binary and multi-dimensional party systems, in which relevant ideological stances must be characterized in additional dimensions, relating for example to opposition to elites, government, political parties or mainstream media. In this article we leverage ideological embeddings of Twitter networks in France in multi-dimensional opinions spaces, where dimensions stand for attitudes towards different issues, and we trace the positions of users who shared articles that were rated as misinformation by fact-checkers. In multi-dimensional settings, and in contrast with the US, opinion dimensions capturing attitudes towards elites are more predictive of whether a user shares misinformation. Most users sharing misinformation hold salient anti-elite sentiments and, among them, more so those with radical left- and right-leaning stances. Our results reinforce the importance of enriching one-dimensional left-right analyses, showing that other ideological dimensions, such as anti-elite sentiment, are critical when characterizing users who spread fake news. This lends support to emerging accounts of social drivers of misinformation through political polarization, but also stresses the role of the entanglement between fake news, anti-elite polarization, and the role of scientific authorities in public debate.

Introduction

In recent years, the spread of misinformation online has become a salient issue. The term fake news may be assigned to a wide variety of online content, ranging from satire or parody to misleading or fabricated content (Wardle and Derakhshan 2017; Jack 2017; Tandoc Jr, Lim, and Ling 2018; Rogers 2020). However, most empirical research has converged toward a simple and operational definition of the concept: any information published by a media, a website, or social platform which has been classified as unreliable or deceptive by fact-checking agencies constitutes fake news. Using this convention, most recent large-scale studies seeking to assess the extent of the phenomenon have concluded that individual exposure to fake news was not as high as initially feared (Guess, Nyhan, and Reifler 2018; Guess, Nagler, and Tucker 2019; Grinberg et al. 2019). For instance, in the US, fake news represents only 1% of people's news media diets (Allen et al. 2020). Despite these reassuring results, the issue of fake news still merits closer examination. Indeed, even though the volume of fake news has been inflated in public discourse, many experts have argued that the impact of fake news is very hard to assess (Lazer et al. 2018), and that it may be a symptom of larger socio-political troubles whose roots are not yet sufficiently understood (Bennett and Livingston 2018). As distrust in the media, institutions, and experts is currently on the rise (Nichols 2017; Bennett and Livingston 2020), and is subverting the crucial function that a shared information ecosystem plays in sustaining democracy (Benkler, Faris, and Roberts 2018), it remains critical to conduct further research to better understand the underlying dynamics behind the spreading of fake news on social media.

Scholars have suggested various hypotheses to explain the individual propensity for sharing fake news, which includes: a lack of careful reasoning (Pennycook and Rand 2019, 2020, 2021), older age and a digital media literacy deficit (Brashier and Schacter 2020), or politically motivated reasoning (Kunda 1990; Miller, Saunders, and Farhart 2016). Combining digital trace data from social media and surveys, numerous studies have found that the users who share fake news tend to come from a highly politicized audience (Grinberg et al. 2019; Guess, Nagler, and Tucker 2019; Hopp, Ferrucci, and Vargo 2020). While it is becoming increasingly clear in the literature that ideology is the most critical to explaining the spread of fake news online (Osmundsen et al. 2021), it is also important to note that the emphasis has been put only on polarization along the left-right ideological dimension. This is a natural choice, especially in the case of the US, where most empirical studies have been produced. Nevertheless, focusing solely in the left-right axis risks neglecting other dimensions and cleavages that structure socio-political dynamics. For example, public opposition to government and political parties (Uscinski et al. 2021), or distrust towards the media, institutions and experts, have proven to be decisive in explaining the spread of political rumors (Petersen, Osmundsen, and Arceneaux 2020) or adhesion to conspiracy discourses (Miller, Saunders, and Farhart 2016), thus making a need for analyzing misinformation in several relevant opinion dimensions.

In this article we set out to study misinformation in geometrical opinion spaces including, for the first time, these identified dimensions that have been lacking so far and are beyond the left-right dimension. To do so, we leverage recently proposed ideological embedding methods (Ramaciotti Morales, Cointet, and Muñoz Zolotoochin 2021), which embed social networks in spaces where dimensions act as indicators of ideological traits of users. We use this method on Twitter data, in conjunction with survey data of stances along several dimensions that have been identified with distinguishable issues of public debate, including left-right stances and attitudes towards elites among others. With the help of this new method, we are able to embed Twitter users in opinion spaces where dimensions stand for indicators of opinions for identifiable issues. We will use this method to study the spread of misinformation in France, a country where trust in the news is now among the lowest in Europe (Newman et al. 2020). By positioning users in opinion spaces, and identifying who participates in misinformation, we seek to answer two research questions: **Q1)** Is the left-right polarization related to the participation in the spread of misinformation?, and **Q2)** Do attitudes towards elites relate to the spread of misinformation? Q1 expands on literature about the US that previously examined the role of the Democrat-Republican polarization in the spread of fake news, by examining a European setting. Q2 seeks to enrich this political characterization of users by adding another opinion dimension that is hypothesized by recent studies as a promising lead in understanding the socio-political dynamics of misinformation. One of our contributions is the proposal for a method for mapping latent ideological stances in social networks onto explicit scales given by opinion surveys, which allows us to 1) disentangle dimensions that capture left-right and anti-elite cleavages, and 2) to measure positions in predefined, externally provided opinion scales, removing the need for us to interpret the meaning of dimensions of latent spaces.

Related Work

Types of Users Spreading Misinformation

Numerous research works have found that fake news accounts for a very small part of most people's media diet (Allen et al. 2020), and is consumed and shared by a tiny minority of users on social media (Guess, Nagler, and Tucker 2019). For instance, 0.1% of users were responsible for 80% of the fake news shared on Twitter during the 2016 US presidential campaign (Grinberg et al. 2019). An important challenge for scholars then has been to understand the distinctive characteristics of these scarce fake news spreaders. Competing theoretical frameworks have suggested various hypotheses to explain the main individual drivers. On the one hand, research in experimental psychology has shown how a lack of attention and careful reasoning could fuel the sharing of fake news (Pennycook and Rand 2019, 2020, 2021). On the other hand, numerous studies linking digital traces to survey data have shown that the vast majority of fake news sharers tend to belong to the extreme fringes of the political spectrum (Grinberg et al. 2019; Guess, Nagler, and Tucker 2019; Guess, Nyhan, and Reifler 2018). To disentangle whether the spread of fake news is spurred by a lack of attention or by partisan motivation, a recent study investigated which variables was the most predictive to explain the sharing of fake news on Twitter. The study shows that cognitive skills do not play a major role with respect to explanations involving the political partisanship of individuals (Osmundsen et al. 2021). These accounts invite further analyses of misinformation as a socially-motivated phenomenon (Kunda 1990).

Ideology and Misinformation

One important challenge in advancing the understanding of the role of ideology in fake news spreading is the lack of frameworks on which to distinguish the variables that would explain ideological motivations in participating in fake news. While multiple political motivations could potentially drive the sharing of fake news, in most studies, the underlying socio-political framework is a binary classification of individuals as Democrat- or Republican-leaning, with "polarization" being an independent variable ranging from most liberal to most conservative. Political extremism in two-party systems such as that of the US, however, is difficult to translate to general settings that might need several opinion dimensions to characterize political competition (Benoit and Laver 2012). In particular in European settings, it is known that more opinion dimensions (*e.g.*, attitudes towards the EU, immigration, or trade openness) are needed to explain political choices (Hix, Noury, and Roland 2006): *e.g.*, while most political stances in the UK might be represented with a single ideological variable, they would need at least three in Finland, with France somewhere in the middle (Bakker, Jolly, and Polk 2012). It has also been shown (Ramaciotti Morales, Cointet, and Muñoz Zolotoochin 2021) that at least three political opinion dimensions participate in structuring Twitter networks in France, including a Left-Right dimension, but also a dimension of opinions towards internationalization (including issues such the EU or trade openness), and towards immigration. Besides the lack of Left-Right cleavages to capture sufficiently relevant dynamics of misinformation, there is additional motivation to include specific new opinion dimensions. For instance, as "social media can lend a voice to anti-system forces that actively seek to undermine liberal democracy" (Tucker et al. 2017), it becomes relevant to examine to which extent the sharing of fake news may be driven by populism and anti-elites sentiment (Humprecht, Esser, and Van Aelst 2020; Petersen, Osmundsen, and Arceneaux 2018) beyond traditional factors such as partisanship. These analyses, however, require the ability to map and study social systems on dimensions beyond traditional Left-Right, of Democrat-Republican cleavages, to include dimensions related to populism, associated with attitudes towards elites. "Elites", in this sense, alludes to a social group referenced in most populist discourses, according to which, regardless of how they stand to test or quantification, society may be divided into "the people" and "the elites", two homogeneous and antagonistic groups (Mudde 2004). More than measuring belonging to any of these two groups, no matter how they might be defined, the question of the attitudes towards elites measures an individual (or collective, *e.g.*, in the case of a party) disposition towards "the rhetoric that construct politics as the moral and ethical struggle" between them (De la Torre 2010). Different political surveys measure self-position or expert positioning along people-elite cleavages, for example, positioning political parties on axes measuring "salience of anti-establishment" and "anti-elite rhetoric" (Bakker et al. 2020).

Estimating Ideologies and Opinions in Networks

Finally, our work relates to methods seeking to embed individuals in *opinion* spaces, where dimensions are informative of attitudes, and that go back to the NOMINATE method (Poole and Rosenthal 1984), used to infer a liberal-conservative scale on which to position US members of parliament (MPs) based on how they vote on bills. Thanks to social network data, researchers have sought to apply this procedure to users, replacing how *MPs vote bills* by how *users follow politicians* online. Drawing on this principle, researchers have been able to embed and validate (using polls or voting records) millions of Facebook (Bond and Messing 2015) and Twitter (Barberá 2015) users in the US liberal-conservative one-dimensional scales. These methods were further developed for multi-dimensional European settings, using political survey data to compute meanings of spatial dimensions (Ramaciotti Morales, Cointet, and Muñoz Zolotoochin 2021) and validated using text utterances (Ramaciotti Morales and Muñoz Zolotoochin 2022). This method, called *ideological embedding*, takes its name from the fact that dimensions are proven to stand for indicators of attitudes towards sets of grouped issues that are identified with an ideology. For example, one dimension captures attitudes towards a set of issues related to globalization (*e.g.*, European integration, trade protectionism). These methods have been subsequently used in other applications in computational social sciences, such as embedding other entities like cited news articles (Cointet et al. 2021) or even online groups affiliated with social movements (Ramaciotti Morales et al. 2021). One of the limitations of ideological embedding –that we aim at overcoming in this article– is this aforementioned "issue entanglement": inferred dimensions may measure attitudes towards several simultaneous issues of debate, preventing the examination of isolated issues such as attitudes towards elites.

Social Networks Opinion Space Data

Twitter Network Embedded in Ideological Space

To study the position of users participating in misinformation we first embed parts of the Twitter network in opinion spaces. We take the Twitter sub-graph of the neighbors of French parliamentarians (MPs) active on Twitter [1] To this graph we apply the so-called ideological embedding procedure. We replicate the procedures described by Ramaciotti Morales, Cointet, and Muñoz Zolotoochin (2021) as a starting point for our opinion spaces. Here we report what is needed for this study; see the cited article for further details.

We take manually annotated 831[2] (out of 925) accounts of French MPs that are active on Twitter, affiliated to 10 political parties, and their collected followers (4 487 430 by May 2019). Following (Barberá 2015), users that follow less than 3 MPs or that have less than 25 followers are filtered out. Users with repeated sets of followed MPs are also filtered to obtain 368 831 followers ensuring a full-rank adjacency matrix. This MPs-followers bipartite network is embedded in a latent space taking the Correspondence Analysis (Greenacre 2017) of the graph, producing an homophily embedding (Lowe 2008; Barberá et al. 2015): users close in space have higher probability of following the same MPs. To provide meaning for dimensions, the positions of reference points (political parties) along each dimension are correlated with positions in opinion scales available in political opinion surveys. For each party, the spatial position is computed as the mean of its MPs (see Fig. 1) and compared to the position of parties in more than 40 attitudinal scales (ranging from 0 to 10) provided by the Chapel Hill Expert Survey (CHES, Bakker et al., 2020). The relation between positions in these two different sources (ideological embedding and CHES) is assessed with a Pearson correlation. Only the first three latent ideological dimensions show statistically significant correlations (p-value$<$0.05) with CHES attitudinal scales: the first is related with Left-Right cleavages (called the "Left-Right" dimension by Ramaciotti Morales, Cointet, and Muñoz Zolotoochin, 2021), the second is related with attitudes towards trade protectionism and European integration (called "Local-Global" dimension), and the third is related with attitudes towards immigration and multiculturalism (called "Immigration & Multiculturalism" dimension). These latent dimensions account for attitudes towards a set of grouped issues, and were inductively interpreted by Ramaciotti Morales, Cointet, and Muñoz Zolotoochin, (2021) using CHES data.

For these 368 831 users, the subtended social graph (who *follows* whom) was collected. Some users have disabled permissions to have their followers collected, thus resulting in a graph of 230 911 users and 67 217 556 edges (density=0.00126). This is the social graph embedded in the so-called ideological space, reproduced from Ramaciotti Morales, Cointet, and Muñoz Zolotoochin (2021), which will be our starting point in building our opinion spaces.

[1] These data have been declared, the 19 Mars 2020 at the registry of data processing at the *Fondation Nationale de Sciences Politiques* (Sciences Po), and respect the General Data Protection Regulation 2016/679 (GDPR) and Twitter's policies.

[2] Obtained from http://www2.assemblee-nationale.fr/deputes/liste/reseaux-sociaux for deputies, and http://www.senat.fr/espace_presse/actualites/201402/les_senateurs_sur_twitter.html for senators.

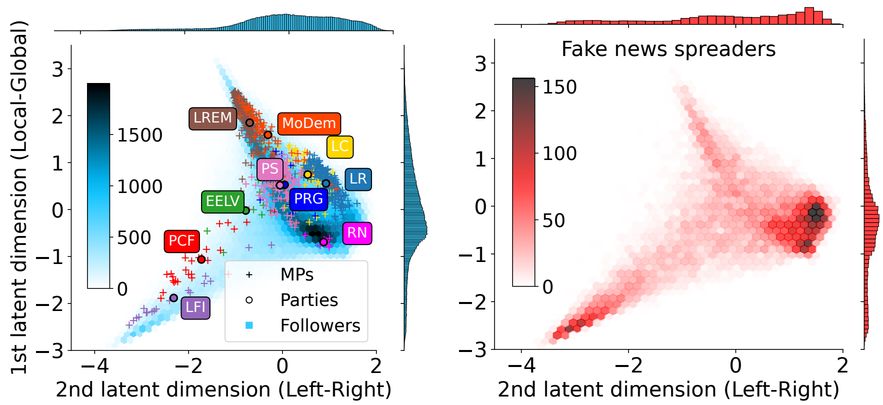

Figure 1: First two dimensions of the bipartite fraction of the Twitter network of French parliamentarians and their followers embedded in an ideological space (left). Embedding was achieved using Correspondence Analysis and meaning of latent dimensions was computed with the Chapel Hill Expert Survey (CHES) data (Bakker et al. 2020). Results reproduced from Ramaciotti Morales, Cointet, and Muñoz Zolotoochin (2021). Spatial distribution of the 15.018 users (out of 230 911) that shared an URL signaled as being misinformation (right).

Identifying Users that Share Misinformation

While most previous works have operationalized fake news by labeling media outlets according to content sanctioned by fact-checkers (Grinberg et al. 2019; Osmundsen et al. 2021; Guess, Nagler, and Tucker 2019), we use a higher-granularity level by considering labeled URLs. We define a fake news spreader as an individual who has shared at least once an URL that has been fact-checked. We leverage Meta's (formerly Facebook's) Third-Party Fact-Checking Program to identify a list of such URLs. The program partners with the International Fact-Checking Network (IFCN) to establish a list of fact-checking agencies worldwide. In France, 4 agencies participate in the IFCN, hosted by: *20 Minutes*, *Agence France Presse*, *Le Monde* (*Les Décodeurs du Monde*), and *France 24* (*Les Observateurs de France 24*). Suspicious French URLs are sent to these media houses who may decide to publish a fact-check. To track this process, we make use of the URL shares dataset. This dataset (Messing et al. 2018) consists of web page addresses that have been shared on Facebook starting January 1, 2017 through to and including February 28, 2021, more than 100 times. Among the various metadata attached to each URL, one variable indicates whether the URL was fact-checked, and what its final rating is, including "mixture or false headline", "satire", "false", "true" among others. We only consider the 1 786 unique French URLs (meaning mostly shared by French users) categorized as "False" or "mixture or false headline" by fact-checkers. We reviewed them manually and removed 22 URLs which did not correspond to a specific news story but pointing to an entire web domain. We then proceeded to collect all tweets sharing the resulting 1 764 URLs using the Twitter API v2. This resulted in 340 747 identified tweets. We filter *replies* and *quotes* which may signal a critical stance at the original content (Roth, St-Onge, and Herms 2021) to focus on original tweets and retweets (321 821) citing 1 457 distinct URLs, which resulted in 134 524 Twitter users. Among these users, 15 018 are part of our ideologically positioned dataset (6.5% or our 230 911 users; see Fig. 1, right side).

Social Network Attitudinal Embedding

These ideological spaces have relevant limits that we must address before tackling our research questions. First, they have no explicit reference center from which users could be deemed polarized by virtue of their distance: the mean position on each dimension is only the center of our sample and cannot be assured to be close to centered opinions. Second, even if there was an agreed upon dimension-wise central position, distance consistency across space cannot be assured: *e.g.*, if 0 is the center, users at -2 in the Left-Right dimension cannot be said to be twice as radical as users at -1. Finally, being ideological dimensions, they stand as indicators for attitudes towards sets of correlated, grouped issues; we cannot readily inspect attitudes towards isolated issues, such as attitudes towards elites (which is required to test Q2).

To overcome these limits, we propose mapping the ideological positions of users, from ideological space, onto an *attitudinal reference frame*: an opinion space from an external and explicit instrument, such as the CHES, which has explicit and bounded numerical scales where respondents position parties based on predefined questions, such as, *e.g.*, "On a scale from 0 (most left-wing) to 10 (most right-wing), where do you position party X?". Value 5 explicitly codes for the political center. The CHES contains two dimensions of particular interest in the study of fake news: a left-right dimension (pointed to as relevant by literature linking misinformation and political polarization) ranging from 0 to 10, and linked to Q1, and an explicit dimension where respondents place parties answering "What is the salience of anti-establishment and anti-elite rhetoric of the party?" on a scale from 0 ("not important at all") to 10 ("extremely important"), linked to Q2. We use an affine transformation to map positions from the identified 3 dimensions of ideological space onto these 2 attitudinal CHES dimensions. This choice results from the observation that party positions along these latent ideological dimensions are correlated with positions along several CHES issue dimensions. The aforementioned study by Ramaciotti Morales, Cointet, and Muñoz Zolotoochin (2021) discovered nearly 30 CHES dimensions having statistically significant correlations (p-value$<$0.05) with latent ideological dimensions. In particular the two CHES attitudinal dimensions of interest for us, "Left-Right" and "Anti-elite saliance", also display significant correlations with the computed latent dimensions: "Anti-elite saliance" with the first dimension, "Left-Right" with the second (see Fig 2). We thus consider an affine transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ mapping ideological space onto a two-dimensional CHES attitudinal reference frame for Left-Right and Anti-elite salience dimensions. Let $Y \in \mathbb{R}^{2\times 8}$ be the position of the 8 political parties that exist both in our annotated dataset and in the CHES dataset, along the two selected CHES dimensions. Let $X \in \mathbb{R}^{3\times 8}$ be the position of these 8 parties in the first 3 ideological dimensions, where Pearson correlation with CHES dimensions were deemed of significance. We seek to determine the optimal affine transformation T^* that minimizes the error in Frobenius norm $E = \|Y - T(X)\|_F$. To determine T^* we express T, slightly abusing notation, in augmented matrix form (also called *homogeneous coordinates*), as $\widetilde{T}(\widetilde{X}) = \widetilde{T} \cdot \widetilde{X}$, with $\widetilde{T} \in \mathbb{R}^{3\times 4}$,

$$\widetilde{X} = \begin{pmatrix} X \\ 1 \end{pmatrix} \in \mathbb{R}^{4\times 8}, \qquad \widetilde{Y} = \begin{pmatrix} Y \\ 1 \end{pmatrix} \in \mathbb{R}^{3\times 8},$$

where $\widetilde{X}$ and $\widetilde{Y}$ are augmented by adding a row of ones to X and Y.

Finally, we compute optimal augmented transformation $\widetilde{T}^*$ that minimizes E as the pseudo-inverse as $\widetilde{T}^* = \widetilde{Y}\widetilde{X}^T\left(\widetilde{X}\widetilde{X}^T\right)^{-1}$, and we retrieve the optimal affine transformation T^* by removing the last row from $\widetilde{T}^*$ (Penrose, 1956; see Dokmanić and Gribonval, 2017, for further details). After we computed parameters of transformation T^*, we applied it to all users, MPs, and political parties to position them in our two-dimensional CHES attitudinal reference frame (see Fig. 3). As expected, Marine Le Pen's RN party is in the anti-elite far-right corner, Jean-Luc Mélenchon's LFI party in the anti-elite far-left corner, while Emmanuel Macron's LREM party in the center of the Left-Right dimension, with low anti-elite stance.

Before using our two CHES dimensions to study misinformation, we seek to gain confidence in the reliability of the newly computed positions of users. We do this in two different ways. First, we control that the new party positions in estimated attitudinal reference dimensions correspond with the party positions of the CHES data which served for guiding the transformation T^*. We do this by looking at the minimized mean error in positioning parties along each dimension. We find that the mean error of party positioning for the CHES Left-Right attitudinal dimension is 0.895 (on a 0 to 10 scale, or 8.14% relative error), and 0.522 (on a 0 to 10 scale, or 4.75% relative error) for the CHES Anti-elite salience attitudinal dimension. We further tested the stability of the

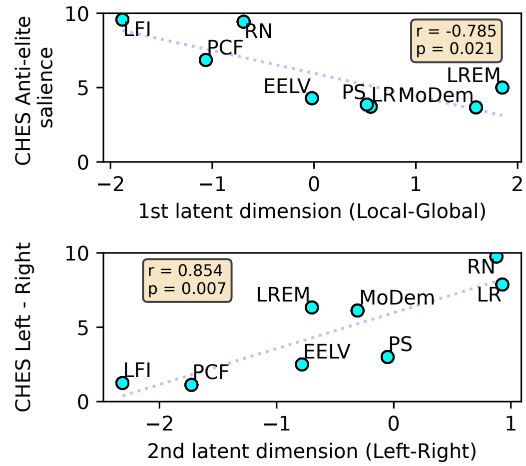

Figure 2: Correlation of party positions according to latent ideological dimensions and as provided in the Chapel Hill Expert Survey (CHES) along two relevant attitudinal dimensions: Left-Right cleavages and Anti-elite salience.

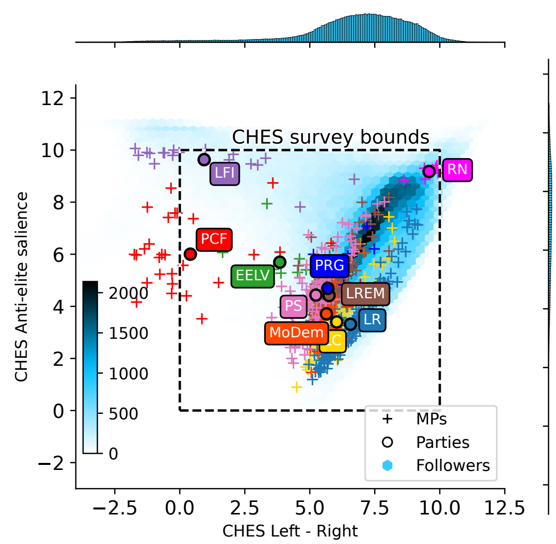

Figure 3: Positions of parliamentarians (MPs), political parties, and followers of MPs mapped from latent ideological space of Fig. 1 onto two Chapel Hill Expert Survey (CHES) dimensions of interest: Left-Right cleavages and Anti-elite salience.

position of parties bootstrapping the attitudinal embedding method with different sample fraction sizes (see Section of the Annex). This provides further confidence that transformation T^* produces a reliable attitudinal embedding of political parties.

We then seek to gain insight on whether our 230 911 are rightly positioned by T^*. To do so, we rely again on the methods proposed by Ramaciotti Morales and Muñoz Zolotoochin (2022), and examine the bios self-description texts written by users in their profiles. We consider three different user classification *labels* to test their positions: *Left-wing* (identified by the use of the word "gauche", "left" in French), *Right-wing* (identified by the use of the word "droite", "right" in French), and the label *Anti-elite* (identified by the use of the words "elite", "politicien" or "peuple", correspondingly "elite", "politician" and "people" in French). For "elite", "politicien" and "peuple" we included common orthographic spelling errors and plural forms. For labels *Left-wing* and *Right-wing*, we filtered out individuals that used the corresponding words in a critical sense as done by Ramaciotti Morales, Cointet, and Muñoz Zolotoochin (2021); we filtered out users that also use the three most frequently used words (as found in a term frequency analysis) associated with critique: "anti", "contre", and "déteste". This allowed us to label (or not) each one of the 230 911 user with these three labels. Fig. 4 shows the density of users having these labels by regions of space on the two dimensions of our CHES attitudinal reference frame. We observe that the affine transformation T^* positions users in such a way that the density of those having these three labels grows monotonically. For example, the density of users having the label *Anti-elite* (thus suspected of having salient opinions towards "elites" and the "establishment", as formulated by the corresponding CHES question), grows monotonically as we move further up on the CHES Anti-elite salience attitudinal dimension, from 0 to 10. Some users (including MPs) are outside the bounds of the CHES scales. This is not surprising, as the scale is proposed for positions of political parties, and it is natural to consider that some MPs, and more so some users, are, *e.g.*, more to the left, or more to the right than the leftmost or rightmost positions considered by respondents for the position of parties. To illustrate and test the general interest and validity of the method, we applied the attitudinal embedding method and its text-based validation to a network using annotated Twitter accounts of Italian MPs with similar results (see Section of the Annex).

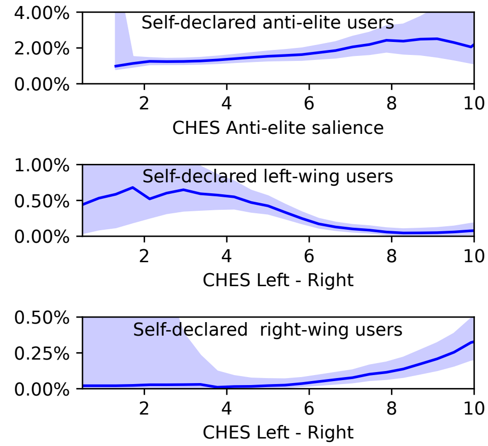

Figure 4: Density of users using words defining three labels (*Left-Wing*, *Right-wing*, and *Anti-elite*) on their profile descriptions, mapped along the two dimensions of the chosen attitudinal reference frame, consisting on the dimensions "Left-Right" and "Anti-elite salience" of the Chapel Hill Expert Survey (CHES) data. Confidence intervals (α=0.05) shown in light blue.

Geometrical Analysis of Misinformation

So far we have produced an *attitudinal* embedding of our social network in a two-dimensional frame where two dimensions stand as indicators of attitudes towards left-right cleavages, and towards elites. This positioning, because it is projected onto scales of the CHES survey, has an explicit meaning as to where the "center" and the "extremes' are, and what dimensions measure (*e.g.*, users at value 5 of Left-Right axis are judged to be part on the center). All surveys and polls are subject of critique, but by mapping users onto the scales of survey, we remove the need for expert interpretation of latent space, and make explicit the interpretation of dimensions and numerical scales and distances.

We may now finally undertake the analysis of misinformation in this two-dimensional attitudinal CHES reference frame. We begin by addressing question Q1. In previous research on U.S. subjects, it has been observed that "polarization" is the main driver of participation in misinformation. In the work by Osmundsen et al. (2021) polarization is taken to mean that a user has extreme Democrat- or Republican-leaning positions. To test this hypothesis in France, we take interest in the density of Twitter users along the CHES Left-Right dimension. To do so, we produce a uniform spatial partition along this dimension and compute the percentage of users that shared fact-checked URLs on each bin of the partition. We explore only positions in which partitions have at least 100 users per bin. This goes well beyond outside the [0,10]x[0,10] bounds of the CHES frame, except for positions on the least anti-elite part of the plane, where there were too few users. We observe that, as in the US, few users in average left-right positions share misinformation, and that the density of those that share fake news grows towards extreme left and right positions. At the center of the CHES Left-Right scale (value = 5), 5.58% of our users share fake news (with 3.9%-7.27% CI bounds at $\alpha = 0.1$). In contrast, 13.64% of our users shared fake news (with 9.81%-17.48% CI bounds at $\alpha = 0.1$) at the rightmost end of the CHES Left-Right scale (value = 10), and 21.53% shared fake news (with 14.15%-28.0% CI bounds at $\alpha = 0.1$) at the leftmost end of the CHES Left-Right scale (value = 0). More extreme positions have even higher percentages of fake news spreaders. According to these results, users distant from the CHES Left-Right center (relative to the 0-10 scale), approximately below 4.5 and above 8, are more likely to share misinformation than the global ratio of signaled users: at 6.5%. Fig. 5

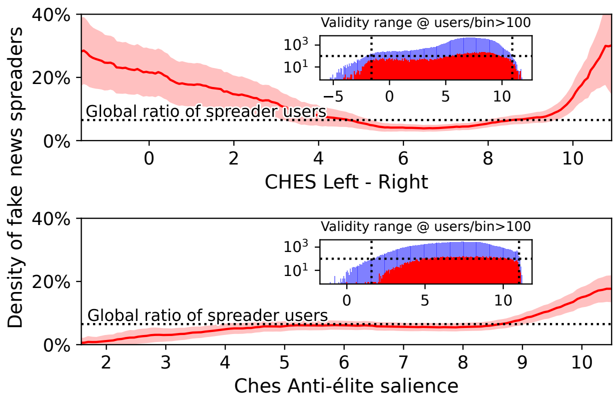

Figure 5: Ratio of fake news spreaders according to their positions in attitudinal dimensions of the Chapel Hill Expert Survey (CHES). In insets, net quantities of users (blue) compared to users that shared misinformation (red).

illustrates this ratio at different attitudinal positions.

We now take interest in how users that shared fake news are positioned along the CHES Anti-elite salience dimension (relating to Q2), for which we produce a partition with the same binning. We observe that the ratio of fake news spreaders grows monotonically with the position along this dimension. At the least anti-elite end of the CHES bounds (value = 0) 0.83% share misinformation (with 0%-2.86% CI bounds at $\alpha = 0.1$). At the most anti-elite end of the CHES bounds (value = 10) 14.98% share misinformation (with 11.99%-17.97% CI bounds at $\alpha = 0.1$). These results show that only users with high anti-elite sentiments (values above 9, approximately) have a higher probability of sharing misinformation, when compared with the global ratio threshold of fake news spreaders (6.5%).

The existence of trends for the ratio of fake news spreaders along both dimensions opens the possibility for more complex geometrical patterns in the attitudinal reference frame. To further inspect the relation between stances in both axes and the tendency to spread fake news, we plot the level curves of percentage of these users by region of space in Fig 6. We observe that attitudes towards elites are correlated with the tendency to share misinformation, but that this attitude alone cannot fully account for the observed positions of fake news spreaders. Indeed, we observe that the highest ratio of users that share fake news occur among those that hold strong anti-elite sentiments, but in so far as they are *also* polarized (holding extreme positions) in the left-right scale. In fact, users that hold extreme (values greater than 10) anti-elite sentiments have less propensity to participate in spreading fake news if they are not polarized in the left-right scale (values near 5). We computed the density of fake news spreaders in a second country, Italy (see Section of the Annex), to confirm the emergence of the same pattern: users who spread misinformation are more densely packed in anti-elite regions and are polarized on the left-right dimension.

Network Activity Analysis of Misinformation

In this section we explore activity of users in the network with greater detail to further characterize the link between stances and participation in misinformation. Using the social graph we compute their clustering coefficient (Watts and Strogatz 1998), their friends, and their followers. From the collected user meta-data, we obtain their number of tweets and creation date of the account, from which we compute an estimation of mean tweets per day for each user. Because we also know the number of tweets containing fact-checked URLs, we can estimate the mean ratio of tweets per users that contain misinformation. We observe that fake news spreaders are comparatively more active and more connected than the rest of users (see Fig. 7). Their clustering coefficient is lower, occupying less dense neighborhoods than those that do not spread fake news, suggesting a limited degree of shared acquaintances between their friends and followers. We will connect this observation to qualitative analysis of types of fake news spreaders in the next section.

Because fake news spreaders are more active, the question remains whether the observed positions with higher densities of spreaders are just positions of hyperactive users. The possibility remains that these users have the same propensity to share fake news (as measured in individual ratio of tweets with misinformation), but they post more tweets per day and thus are more easily detected as fake news spreaders. To answer this question we examined how the activity (mean tweets per day) of each user is distributed in our attitudinal space. We found little evidence that users in extreme left-right, or extreme anti-elite stances post more tweets (See Fig. 12 in Section of the Annex). There are more fake news spreaders in extreme positions, and they post more contents, but, overall (including all users), frequency of posting is not comparatively superior at these extremes. We computed the mean value of the ratio of fake news per users, aggregated by bins in our two-dimensional space (we use the same binning as in Fig. 6) and corroborated that it has a similar distribution (see Fig. 13 in the Appendix). We also re-ran the structural comparisons of spreaders versus non-spreaders using a nearest available pair-matching method (Rubin 1973) to guarantee that the ideological distribution of non-spreaders map the one from spreaders. Still we observe the same discrepancy between the two, indicating that the higher the number of followers and friends, the higher activity and lesser clustering coefficient that characterize spreaders are not determined by the position of those individuals in the ideological space.

Qualitative Analysis of Fake News Spreaders

We then proceed to a more qualitative in-depth analysis of the fake news spreaders accounts identified in the various regions of the attitudinal CHES space. More precisely, we examined the accounts located in the vicinity of three remarkable points A, B and C shown in Fig. 6. The points A and C are located where the propensity to spread fake news is the highest in the left- and right-leaning regions. Conversely we also consider point B, situated at (5,10) corresponding to anti-elite centrist fake news spreaders. Note that the propensity to share fake news at point B is below the global ratio

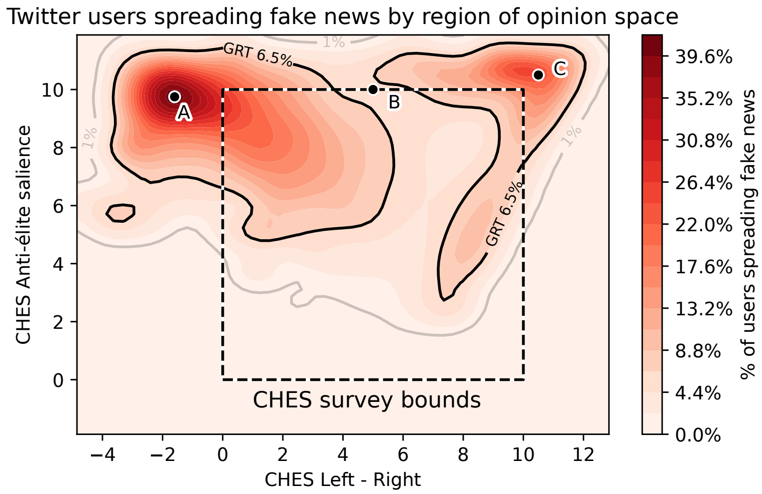

Figure 6: Percentage of users that spread misinformation in different regions of the Chapel Hill Expert Survey (CHES) reference frame made of a left-right, and an anti-elite salience dimension. The global ratio threshold (GRT) of users that share misinformation in our sample is 6.5%.

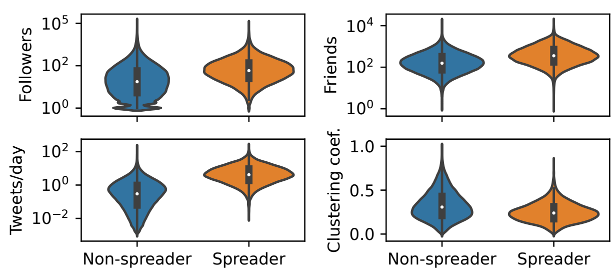

Figure 7: Distribution of followers, friends, mean tweets per day, and clustering coefficient of spreader and non-spreader users.

threshold. We then identify every fake news spreader positioned at a distance smaller than 1 in the attitudinal space, resulting in 536 users near A, 211 near B and 393 users near C. We perform a qualitative analysis of these accounts based on their description as provided by their account metadata and the type of fake news they shared.

Unsurprisingly, anti-elite users, both far-left and far-right, ostensibly express their ideological support in their Twitter profiles (description field), notably by displaying their membership in a political organization or their support to a political figure. For instance, hashtags and keywords linked to far-left candidate Jean-Luc Melenchon and his political movement *La France Insoumise* were very common in the descriptions of users next to A (*e.g.*, #JLM2022, #LFI2022). Symmetrically, references to the far-right candidate Marine Le Pen and her party *Rassemblement National* or her competitor Eric Zemmour were very frequent in the descriptions of users next to C (*e.g.*, #RN2022, #Marine2022, #Zemmour2022). Interestingly, we also found that the majority of anti-elite far-left users were political activists with no official political mandate, while a number of anti-elite far-right users reported being local councilors or elected representatives as well. These observations allow us to better understand findings from our precedent section: since many fake news spreaders are political activists, or even elected officials, they leverage multiple opportunistic strategies, connecting with users beyond their neighborhoods (explaining the lower clustering coefficient), and showing higher activity levels. Furthermore, we observe that accounts in the vicinity of points A and C, share mostly fake news promoting their own political agenda. For instance, fake news criticizing big companies, the government and the media, or denouncing police violence, financial scandals, poverty and social injustices have been widely spread by users next to point A. Conversely, accounts next to C mainly disseminated fake news related to terrorism or immigration and eliciting anti-Muslim and anti-Arab sentiments. Users in the vicinity of B appear to be less politically engaged. Some of them even forcefully reject the political game and claim to be “without a political label” or to “hate corrupt politicians and journalists”, suggesting that fake news sharing, when fueled by anti-elite sentiments, can still be compatible with a non-partisan stance. We also noted that fake news spreaders in this region (which are pretty rare in comparison to points A and C), tend to share fake news regarding health, ecology, and conspiracy theories at large.

Discussion and Conclusions

In this article, we produced the first *attitudinal embedding* of social network users onto the dimensions of a survey instrument (the CHES) to observe their distribution along two distinct opinion dimensions and how they spread fake news. In contrast with previous works, treating fake news at the level of the source, we were able to identify users sharing fact-checked URLs and characterize a clear-cut population of fake news spreaders along the two aforementioned dimensions: 1) Similar to what has been observed in the US for Democrat-Republican polarization, left-right polarization in France (despite not being the only societal cleavage in politics) is also correlated with the spread of fake news: the larger the distance from centered stances, the more likely it is that accounts will share fake news (answering Q1). According to our measurement, around 5.6% of users at the center of the CHES Left-Right dimension shared fake news (below the 6.5% global ratio of fake news spreaders), while this figure is more than double at the extreme left and right positions of this CHES scale. 2) Thanks to our new attitudinal embedding method, we were also able to relate anti-elite sentiments with propensity to share misinformation, linking an important stream of theoretical research in social sciences with empirical observations at massive scales for the first time (answering Q2). According to our measurements, users in the most anti-elite stance of the CHES scale are more than 15 times more likely to share at least one fake news story than those at the least anti-elite positions. Furthermore, we unveil complex relations between these two dimensions and misinformation: it is the combination of both, extreme left-right polarization and anti-elite sentiments, that is most predictive of the propensity to share fake news. Anti-establishment accounts with no strong left or right leaning are actually less likely to spread fake news than the global average of 6.5%.

Furthermore, we were able to examine fake news shared from users positioned in different regions of our attitudinal space to gain insight about the diversity of the phenomenon both in terms of social positions occupied by users and through the themes of fake news stories they disseminate. In line with the literature, our qualitative analysis reveals that most extreme left- and right-wing accounts who share fake news exhibit an explicit link to political activities (whether as an elected official or as a party supporter). As a result, fake news stories they share are directly linked to the political agenda they defend (financial scandals, police violence on one side, and terrorism and migration on the opposite side). However, we also observed that other kinds of fake news related to health or ecology can be "popular" in other regions (*e.g.*, centrist and anti-elite) of the attitudinal space. The structural analysis of the fake news spreader accounts in the social graph reveals that these individuals are hyper-active but also potentially less socially integrated.

Finally, our method comes with the obvious limitation that we can only observe positions of users that satisfy the criteria needed for spatial positioning (following enough MPs). Only 11.2% of the users who shared fact-checked URLs are embedded in the ideological space. Nevertheless our hypothesis is that the general population obeys to the same pattern. Complementary data analysis would be required to test it, but the observed geometrical pattern fits the theoretical expectations from a large number of previous works interested in the drivers of fake news. This can be achieved with other methods from the literature, allowing to embed larger sets of users, and not only those that follow reference users (MPs in our case), exploiting ideological coherence of some followers, which can be leveraged in inferring ideological positions of their neighbors (Ramaciotti Morales, Cointet, and Laborde 2020). Overall, our geometrical approach of misinformation offers a leap in the understanding of the social drivers for spreading fake news over a significant number of Twitter users.

Stability of Attitudinal Positions

We test the stability of the attitudinal embedding method with bootstrapping. Once we have assembled the bipartite graph of 831 MPs and their 368 831 followers, we randomly chose a subsample of edges among them to compute the attitudinal positions on the CHES Left-Right and CHES Anti-elite salience dimensions, as described in Section . We do this selecting first a percentage of the edges that will be sampled, and drawing 100 samples with which to compute attitudinal positions independently. We compute the positions for percentages 40% and 80%. Lower fractions of used edges increase the dispersion variance of the estimated positions of parties, but nonetheless with sufficiently accurate results. Figure 8 shows the estimated positions of the MPs with bootstrapping, the computed positions of parties as means of MPs, and the positions of parties computed with the full dataset for comparison. For ease of comparison we compute the ellipse of the level curve of a fitted Gaussian distribution at 3 standard deviations.

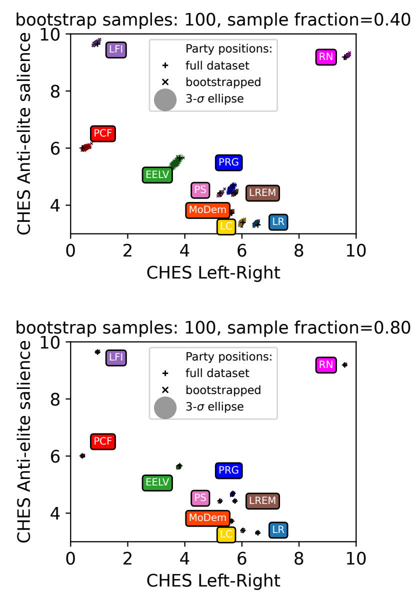

Figure 8: Positions of parties in CHES attitudinal reference space measured as the mean position of those of the MPs, computed with 100 bootstrapped samples, using 40% of all *following* links (top), and 80% (bottom).

Comparison with the Case of Italy

We tested the attitudinal embedding method with annotations of Twitter accounts of MPs from Italy to further illustrate its general interest. To do this, we rely on a dataset collected in October 2020. As done with the data accessible from France, we manually annotated the Twitter accounts of 791 Italian MPs (belonging to 10 parties) out 951. We collected their 5 639 305 followers, and we applied the same filtering criteria described in Section , resulting in 377 067 followers. Using the CHES data for Italy and the first 3 latent dimensions, we produced an attitudinal embedding as described in Section , shown in Figure 9.

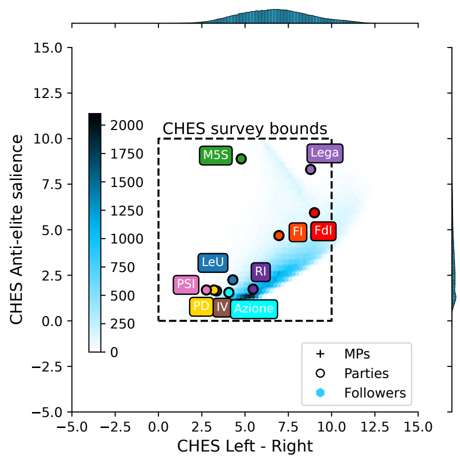

Figure 9: Positions of Italian parties and followers of MPs in two Chapel Hill Expert Survey (CHES) dimensions of interest: Left-Right cleavages and Anti-elite salience.

We collected Twitter bio profile texts of the 377 067 followers and labeled them using keywords “popolo”, “elite” and “politici” (people, elite, and politicians) including gender and number variants. We also labeled according to whether they used the keywords “sinistra” and “destra” (left and right) with positive sentiment and report the density of labeled user along the CHES Left-Right and CHES Anti-elite salience dimensions in Figure 10. Similar to Section , we used Meta’s dataset to identify 251 fact-checked URLs rated as “false” and “mixture or false headline”, and we collected 7 800 Twitter users who shared them, of which 4 796 were also among our 377 067 attitudinally-positioned users. We report their density in attitudinal space in Figure 11.

Activity and Ratio of Fake News Tweets by Regions of Space

Fig. 12 shows the dispersion for the CHES Anti-élite dimension and for political “polarization” measured as distance from the center of the the CHES Left-Right dimensions. Fig. 13 shows the computed mean value of the ratio of fake news per users, using the same binning as in Fig 6.

Ethics Statement

Our analyses were performed on pseudonymized data. A Data Management Plan was filled with our university’s Data Protection Officer as required by GDPR regulation, and a Potential Risks Analysis report was also filled to further assess the implications of our dataset. The related

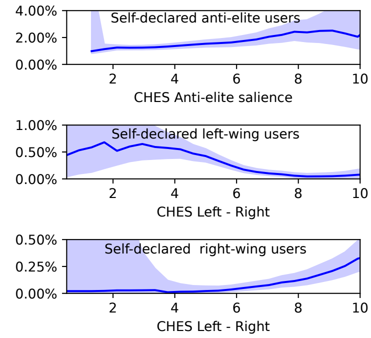

Figure 10: Density of Italian users with three labels (*Left-Wing*, *Right-wing*, and *Anti-elite*) on their profiles, on two CHES dimensions: “Left-Right” and “Anti-elite salience”. Confidence intervals (α=0.05) shown in light blue.

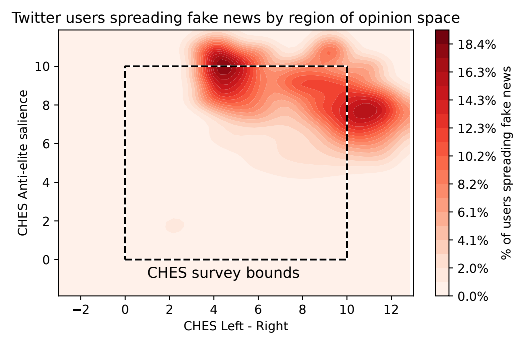

Figure 11: Percentage of users that spread misinformation in different regions of the CHES reference frame made of a left-right, and an anti-elite salience dimension for the case of Italy.

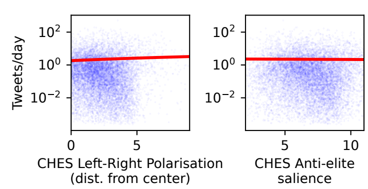

Figure 12: Mean daily tweets and ratio or tweets containing fake news in the French dataset, by polarization (distance from the center of the CHES Left-Right dimension) and attitudes towards élites.

Legal Notice and additional documents can be found at https://medialab.sciencespo.fr/en/activities/epo/.

Our study leverages political opinions of users, and we have taken all necessary steps described in GDPR to work with, and manage separate pseudonymized data to limit the risks associated with sensitive information. Working with these data provides a crucial advantage as it allowed us to unveil previously-unobserved socio-political dynamics underlying fake news, opening new paths to understanding,

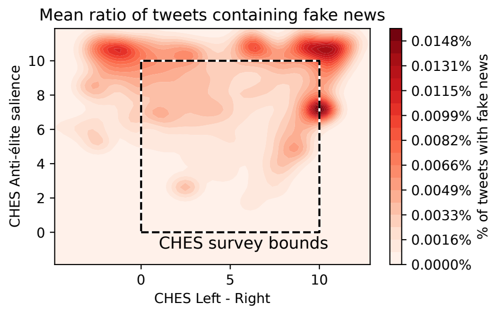

Figure 13: Mean individual ratio of tweets containing fake news by region of attitudinal space using the French dataset.

monitoring, and managing this phenomenon.

Acknowledgments

This work has been partially funded by the Learning Planet Institute, by the Reboot Foundation, by European Polarisation Observatory (EPO) of the CIVICA Consortium, by the Data Intelligence Institute of Paris (diiP), by the French National Agency for Research (ANR) under grants ANR-19-CE38-0006 Geometry of Public Issues (GOPI) and ANR-18-IDEX-0001 IdEx Université de Paris, and by project *De Facto* Observatory of Information (Connecting Europe Facility - 2020-FR-IA-0291). URL Shares dataset was made accessible through "I read it on Facebook" a project funded by the SSRC's Social Science One program. Finally, we would also like to thank Guillaume Plique for his assitance collecting the data from the URL Shares dataset, and Shaden Shabayek for her support for collecting tweets through Twitter v2 API.

References

Allen, J.; Howland, B.; Mobius, M.; Rothschild, D.; and Watts, D. J. 2020. Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14): eaay3539.

Bakker, R.; Hooghe, L.; Jolly, S.; Marks, G.; Polk, J.; Rovny, J.; Steenbergen, M.; and Vachudova, M. A. 2020. 2019 Chapel Hill Expert Survey. *Chapel Hill*. Www.chesdata.eu.

Bakker, R.; Jolly, S.; and Polk, J. 2012. Complexity in the European party space: Exploring dimensionality with experts. *European Union Politics*, 13(2): 219–245.

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.; Jost, J. T.; Nagler, J.; Tucker, J. A.; and Bonneau, R. 2015. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, 26(10): 1531–1542.

Benkler, Y.; Faris, R.; and Roberts, H. 2018. *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Bennett, W. L.; and Livingston, S. 2018. The disinformation order: Disruptive communication and the decline of democratic institutions. *European journal of communication*, 33(2): 122–139.

Bennett, W. L.; and Livingston, S. 2020. *The Disinformation Age: Politics, Technology, and Disruptive Communication in the United States*. SSRC Anxieties of Democracy. Cambridge University Press.

Benoit, K.; and Laver, M. 2012. The dimensionality of political space: Epistemological and methodological considerations. *European Union Politics*, 13(2): 194–218.

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.

Brashier, N. M.; and Schacter, D. L. 2020. Aging in an era of fake news. *Current directions in psychological science*, 29(3): 316–323.

Cointet, J.-P.; Ramaciotti Morales, P.; Cardon, D.; Froio, C.; Mogoutov, A.; Ooghe-Tabanou, B.; and Plique, G. 2021. What colours are the yellow vests? An ideological scaling of Facebook groups. *Statistique et Société*.

De la Torre, C. 2010. *Populist Seduction in Latin America*. Ohio University Press.

Dokmanić, I.; and Gribonval, R. 2017. Beyond Moore-Penrose Part I: generalized inverses that minimize matrix norms. *arXiv preprint arXiv:1706.08349*.

Greenacre, M. 2017. *Correspondence analysis in practice*. CRC press.

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

Guess, A.; Nagler, J.; and Tucker, J. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances*, 5(1): eaau4586.

Guess, A.; Nyhan, B.; and Reifler, J. 2018. Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council*, 9(3): 4.

Hix, S.; Noury, A.; and Roland, G. 2006. Dimensions of politics in the European Parliament. *American Journal of Political Science*, 50(2): 494–520.

Hopp, T.; Ferrucci, P.; and Vargo, C. J. 2020. Why do people share ideologically extreme, false, and misleading content on social media? A self-report and trace data–based analysis of countermedia content dissemination on Facebook and Twitter. *Human Communication Research*, 46(4): 357–384.

Humprecht, E.; Esser, F.; and Van Aelst, P. 2020. Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, 25(3): 493–516.

Jack, C. 2017. Lexicon of lies: Terms for problematic information. *Data & Society*, 3(22): 1094–1096.

Kunda, Z. 1990. The case for motivated reasoning. *Psychological bulletin*, 108(3): 480.

Lazer, D. M.; Baum, M. A.; Benkler, Y.; Berinsky, A. J.; Greenhill, K. M.; Menczer, F.; Metzger, M. J.; Nyhan, B.; Pennycook, G.; Rothschild, D.; et al. 2018. The science of fake news. *Science*, 359(6380): 1094–1096.

Lowe, W. 2008. Understanding wordscores. *Political Analysis*, 16(4): 356–371.

Messing, S.; State, B.; Nayak, C.; King, G.; and Persily, N. 2018. Facebook URL Shares. https://doi.org/10.7910/DVN/EIAACS. Accessed: 2021-09.

Miller, J. M.; Saunders, K. L.; and Farhart, C. E. 2016. Conspiracy endorsement as motivated reasoning: The moderating roles of political knowledge and trust. *American Journal of Political Science*, 60(4): 824–844.

Mudde, C. 2004. The populist zeitgeist. *Government and opposition*, 39(4): 541–563.

Newman, N.; Fletcher, R.; Schulz, A.; Andi, S.; Robertson, C. T.; and Nielsen, R. K. 2020. Reuters Institute digital news report 2020. *Reuters Institute for the Study of Journalism*.

Nichols, T. 2017. *The death of expertise: The campaign against established knowledge and why it matters*. Oxford University Press.

Osmundsen, M.; Bor, A.; Vahlstrup, P. B.; Bechmann, A.; and Petersen, M. B. 2021. Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *American Political Science Review*, 1–17.

Pennycook, G.; and Rand, D. G. 2019. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188: 39–50.

Pennycook, G.; and Rand, D. G. 2020. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality*, 88(2): 185–200.

Pennycook, G.; and Rand, D. G. 2021. The psychology of fake news. *Trends in cognitive sciences*.

Penrose, R. 1956. On best approximate solutions of linear matrix equations. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 52, 17–19. Cambridge University Press.

Petersen, M. B.; Osmundsen, M.; and Arceneaux, K. 2018. A "need for chaos" and the sharing of hostile political rumors in advanced democracies. *American Political Science Association, Boston, MA*, 30.

Petersen, M. B.; Osmundsen, M.; and Arceneaux, K. 2020. The "need for chaos" and motivations to share hostile political rumors. *PsyArXiv (preprint)*.

Poole, K. T.; and Rosenthal, H. 1984. The polarization of American politics. *The journal of politics*, 46(4): 1061–1079.

Ramaciotti Morales, P.; Cointet, J.-P.; Benbouzid, B.; Cardon, D.; Froio, C.; Metin, O. F.; Ooghe, B.; and Plique, G. 2021. Atlas Multi-Plateformes d'un Mouvement Social: Le cas des Gilets Jaunes. *Statistique et Société*.

Ramaciotti Morales, P.; Cointet, J.-P.; and Laborde, J. 2020. Your most telling friends: Propagating latent ideological features on Twitter using neighborhood coherence. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 217–221. IEEE.

Ramaciotti Morales, P.; Cointet, J.-P.; and Muñoz Zolotoochin, G. 2021. Unfolding the dimensionality structure of social networks in ideological embeddings. In *2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*.

Ramaciotti Morales, P.; and Muñoz Zolotoochin, G. 2022. Measuring the accuracy of social network ideological embeddings using language models. In *2022 International Conference on Information Technology & Systems (ICITS22)*.

Rogers, R. 2020. Research note: The scale of Facebook's problem depends upon how 'fake news' is classified. *The Harvard Kennedy School (HKS) Misinformation Review*.

Roth, C.; St-Onge, J.; and Herms, K. 2021. Quoting is not Citing: Disentangling Affiliation and Interaction on Twitter. *arXiv preprint arXiv:2112.00554*.

Rubin, D. B. 1973. Matching to remove bias in observational studies. *Biometrics*, 159–183.

Tandoc Jr, E. C.; Lim, Z. W.; and Ling, R. 2018. Defining "fake news" A typology of scholarly definitions. *Digital journalism*, 6(2): 137–153.

Tucker, J. A.; Theocharis, Y.; Roberts, M. E.; and Barberá, P. 2017. From liberation to turmoil: Social media and democracy. *Journal of democracy*, 28(4): 46–59.

Uscinski, J. E.; Enders, A. M.; Seelig, M. I.; Klofstad, C. A.; Funchion, J. R.; Everett, C.; Wuchty, S.; Premaratne, K.; and Murthi, M. N. 2021. American politics in two dimensions: Partisan and ideological identities versus anti-establishment orientations. *American Journal of Political Science*, 65(4): 877–895.

Wardle, C.; and Derakhshan, H. 2017. Information disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe*, 27.

Watts, D. J.; and Strogatz, S. H. 1998. Collective dynamics of 'small-world'networks. *nature*, 393(6684): 440–442.