scientific reports

Check for updates

OPEN

Political context of the European vaccine debate on Twitter

Giordano Paoletti[1,2], Lorenzo Dall'Amico[1], Kyriaki Kalimeri[1], Jacopo Lenti[3,4], Yelena Mejova[1✉], Daniela Paolotti[1], Michele Starnini[3,5] & Michele Tizzani[1]

At the beginning of the COVID-19 pandemic, fears grew that making vaccination a political (instead of public health) issue may impact the efficacy of this life-saving intervention, spurring the spread of vaccine-hesitant content. In this study, we examine whether there is a relationship between the political interest of social media users and their exposure to vaccine-hesitant content on Twitter. We focus on 17 European countries using a multilingual, longitudinal dataset of tweets spanning the period before COVID, up to the vaccine roll-out. We find that, in most countries, users' endorsement of vaccine-hesitant content is the highest in the early months of the pandemic, around the time of greatest scientific uncertainty. Further, users who follow politicians from right-wing parties, and those associated with authoritarian or anti-EU stances are more likely to endorse vaccine-hesitant content, whereas those following left-wing politicians, more pro-EU or liberal parties, are less likely. Somewhat surprisingly, politicians did not play an outsized role in the vaccine debates of their countries, receiving a similar number of retweets as other similarly popular users. This systematic, multi-country, longitudinal investigation of the connection of politics with vaccine hesitancy has important implications for public health policy and communication.

Despite the success of vaccination in reducing mortality and eradicating diseases like smallpox[1], skepticism about vaccine safety and efficacy has persisted throughout history. The rapid development and global distribution of the COVID-19 vaccines spurred renewed apprehension. In February 2022, a Eurobarometer survey found that while most EU citizens supported vaccination, concerns about unknown long-term side effects of COVID-19 vaccines ranged from 47% to 70% across different countries, with those who were altogether against vaccination reaching 29% in Bulgaria, followed by 24% in Slovakia and 21% in Slovenia[2]. Even before COVID-19, the World Health Organization has indicated vaccine hesitancy as one of the top 10 threats to global health in 2019[3].

Vaccine hesitancy was found to correlate with a complex combination of psychological and sociological factors. Recent literature has linked such attitudes with alternative health practices[4], science denial[5], and conspiratorial thinking[6]. However, this must be contextualized in the personal experiences and beliefs, and in the broader societal, public health and communication environments[7]. A concerning trend around web-based communication channels, including social networks and those integrating recommendation systems, is the possible formation of echo-chambers, both at national[8,9] and global scales[10], as the tightly-knit, homogeneous communities in such echo-chambers provide a fertile ground for fringe narratives that oppose the mainstream[11]. These narratives often exclude traditional and authoritative sources[12] and instead are supported by low-quality information and misinformation[13], which undermines the trust in the public health authorities.

As governments around the world rushed to confront the COVID-19 pandemic, various political actors joined the discussion. At the same time, the increasing adoption of social media has coincided with its use by populist and anti-establishment politicians[14], who took advantage of the context collapse around bite-size units of communication to promote skepticism of authority[15]. Worldwide, vaccine hesitancy has been linked to political beliefs, including in France and Italy[16,17], where those backing right-wing parties had a higher unvaccinated rate. In Poland, around August 2021 vaccination rates were highest in the areas supporting the politician opposing the ruling conservative Law and Justice (PiS) party – a party that has been accused of "flirting" with "anti-vaxxers"[18]. In the U.S., by May 2022 Republicans and Republican-leaning independents (60%) were less likely than Democrats and Democratic leaners (85%) to be fully vaccinated[19]. Overall, studies find vaccine hesitancy and political populism are positively associated across Europe[20,21].

Although the connection between the use of social media and vaccine hesitancy has been documented[13,22], little attention has been paid to the connection between the politicized communication around vaccination on

[1]ISI Foundation, Turin, Italy. [2]Department of Control and Computer Engineering, Politecnico di Torino, Turin 10129, Italy. [3]CENTAI, Turin, Italy. [4]Department of Computer, Control, and Management Engineering Antonio Ruberti, Sapienza University of Rome, Rome, Italy. [5]Departament de Física, Universitat Politècnica de Catalunya, Campus Nord, Barcelona 08034, Spain. ✉email: yelenamejova@acm.org

social media and vaccine hesitancy. For instance, some studies on Twitter have found politics to be a part of the conversation: COVID-19 vaccine-related discussions in Japan included criticism of the government's handling of the vaccine rollout and the holding of the 2020 Olympic Games in Tokyo[23]. As politicians take actions impacting public health and the practice of medicine, including communicating to their constituents on the matters of personal health[24], it is urgent to understand the interplay between such political actors and their audience in the context of vaccination.

To address this research gap, we turn to one of the most popular social media platforms – Twitter – to gauge the relationship between political actors, political interest, and the consumption of vaccine-hesitant content in the period immediately before and during the onset of the COVID-19 epidemic. We propose a custom network analysis pipeline to identify users likely to retweet to vaccine-hesitant content in 17 European countries using a multilingual, longitudinal dataset. Our approach extends previous research[25] that assumes a two-sided controversy by also handling multiple "pockets" of opinion stances, making it generalizable to multi-sided discourse around the world.

Using these tools, we answer the following research questions:

RQ1 Are those more likely to endorse vaccine-hesitant content interested in specific political parties?
RQ2 Is a user's politicization (*i.e.* the extent of the interest in politics and the focus on few parties) related to their endorsement of vaccine-hesitant content?
RQ3 Are political actors more influential in the vaccination debate compared to non-political users?

Crucially, our multilingual dataset allows us to examine the national discussions in the native languages, as out of the 17 countries considered, 16 do not have English as an official language (unlike in previous studies that use English-only queries to analyze "worldwide" vaccine sentiment[26,27]). Further, we perform an extensive mapping between the Twitter accounts in these countries and their respective political figures, which we in turn enrich using *ParlGov*[28], a political science resource that provides detailed information about political parties, elections, cabinets, and governments in parliamentary democracies worldwide. The combination of these resources has allowed us to complete the first quantitative investigation of the relationship between political interest on social media and vaccine hesitancy on a European scale, as detailed below.

Results

In this study, we focus on the Twitter debates around vaccination immediately before and during the COVID-19 pandemic in 17 European countries, in their official languages. We begin by assigning each user in the vaccine debate a score that we call Vaccine Hesitancy Endorsement (VHE), spanning from 0 to 1 that gives an estimate of how likely the individual is to endorse a vaccine hesitant stance (VHE = 1). The score is computed by combining label propagation and community detection on the country endorsement (retweet) network. These networks are obtained for 18 European countries and are composed of posts published in the period 2019/10/01 - 2021/3/31. A number of tweets was manually annotated (see Methods for details) as "vaccine-hesitant", "pro-vaccine" or neutral. We define vaccine-hesitant tweets as those stating directly the user will not vaccinate, questioning their efficacy or safety, or espousing conspiratorial views around their creation or distribution. For each country and period, we then perform 100 partial randomizations of the retweet network and run a community detection algorithm on each of the shuffled networks. Each community is assigned a score based on the proportion of the pro-, anti-vax and neutral labeled tweets. Finally the VSE score is assigned to each user by averaging the 100 scores of the communities it belonged to in the different shufflings. Adding the shuffling step, we further quantify how tightly an individual is connected to a given community, allowing us to give a continuous score to each user.

We begin by examining the distributions of the VHE score for each country and one of four periods (the first before, while the remaining ones during COVID-19), as shown in Fig. 1. Note that some pairs country-period were excluded due to data sparsity. One can see that in most countries and periods, VHE distributions are closer to zero (pro-vaccine) than 1 (vaccine-hesitant), indicating that the majority of users are more likely to endorse pro-vaccine content. Notable exceptions are France, with VHE distributions centered around 0.5 in all periods, the Netherlands, with broad VHE distributions, and Poland, where the majority of users have a VHE score below 0.5, but there exists a non-negligible minority of users tend to anti-vaccine content. Moreover, we can find interesting trends within each country, across time periods. For instance, the distribution of VHE scores in Italy begins with most users likely to endorse roughly the same amount of vaccine-hesitant content as the pro-vaccine one. This trend then shifts towards 0, where more pro-vaccine content is easier to encounter for most users. On the other hand, Germany starts out with most users having VHE scores close to zero, and over time develops a minority of users with much higher scores. Yet in other cases, such as in France, the peak of the overall distribution remains stable over time. In aggregate, however, the peak of the VHE score comes in period 2, during the early days of the pandemic, with a macro-average of 0.39; the VHE score goes down to 0.30 by period 4, during the vaccine rollout.

RQ1   Are those more likely to endorse hesitant content interested in specific political parties?

To answer this question, recall that we model the vaccination debate in each country as a country-specific retweet network, one for each of the four time periods; for this analysis we choose only those that have at least 300 users (61 in total). For each country/period combination, we then perform an OLS regression to model a user's Vaccine Hesitancy Endorsement (VHE) score using the user's followership of politicians in different parties as predictors (as well as some control variables, see Methods). Out of these, 72% (44) had an Adjusted
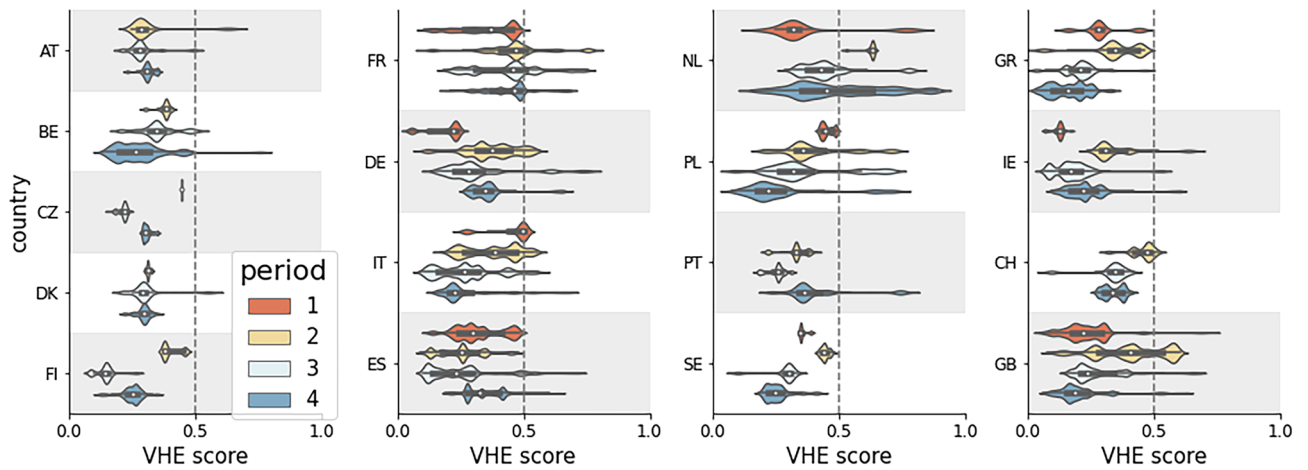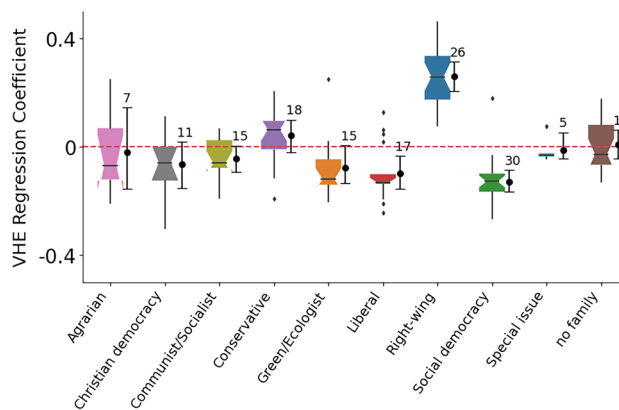
Figure 1. Vaccine Hesitancy Endorsement (VHE) score (see Algorithm 1) distribution across countries and periods. Dashed grey lines indicate scores where a user's exposure has equal shares of vaccine-hesitant and pro-vaccine content. Score near 1 indicates more hesitant, and 0 – more pro-vaccine. Country abbreviations are ISO 3166 standard: AT *Austria*, BE *Belgium*, CZ *Czech Republic*, DK *Denmark*, FI *Finland*, FR *France*, DE *Germany*, GR *Greece*, IE *Ireland*, IT *Italy*, NL *Netherlands*, PL *Poland*, PT *Portugal*, SE *Sweden*, ES *Spain*, CH *Switzerland*, GB *United Kingdom*.

R^2 score greater than or equal to 0.1, which we select for further analysis. In these models, 266 coefficients of parties (51.6%) were significant: 126 positive (having a positive relationship to the VHE score) and 140 negative, where positive (negative) coefficients indicate that users who follow these parties are more (less) likely to endorse vaccine-hesitant rather than pro content on Twitter. For example, the most positive score was found by *Alternative for Germany* in period 3, whereas the most negative by *Civic Platform* in Poland (period 1) (for full listing of coefficients for each party across periods, see Data availability section). Thus, we find mixed results – the relationship between party interest and VHE score may vary between party, country, and period.

To see if these findings generalize across countries, we group the parties according to the ParlGov classification (an extensive resource on political parties in parliamentary democracies[28]). Figure 2 shows the distribution of coefficients for user interest in parties grouped in families using ParlGov with the accompanying 99% bootstrapped confidence intervals. We find that the parties identified as Right-wing have the strongest, and most positive, relationship with the VHE score. On the other hand, following parties in the Social democracy and Liberal families have a negative relationship with users endorsing such vaccine-hesitant content. We find no other statistically robust relationships for other party groups.

To check how consistent these results are over time, in Fig. 3 we plot the significant coefficients and their confidence intervals for five countries that have models with Adjusted $R^2 > 0.1$ for all four periods. We find

Figure 2. Distribution (boxplots) of OLS coefficients modeling users' Vaccine Hesitancy Endorsement (VSE) score (see Algorithm 1) by their interest in parties, grouped by families using ParlGov, an extensive resource on political parties in parliamentary democracies. Positive (negative) coefficients indicate that users who follow these parties are more (less) likely to engage with vaccine-hesitant rather than pro-vaccine content on Twitter. Accompanying points and whiskers indicate a 99% bootstrapped confidence interval. Numbers indicate how many parties are in each group. The parties identified as Right-wing have the strongest, and most positive, relationship with the VHE score.
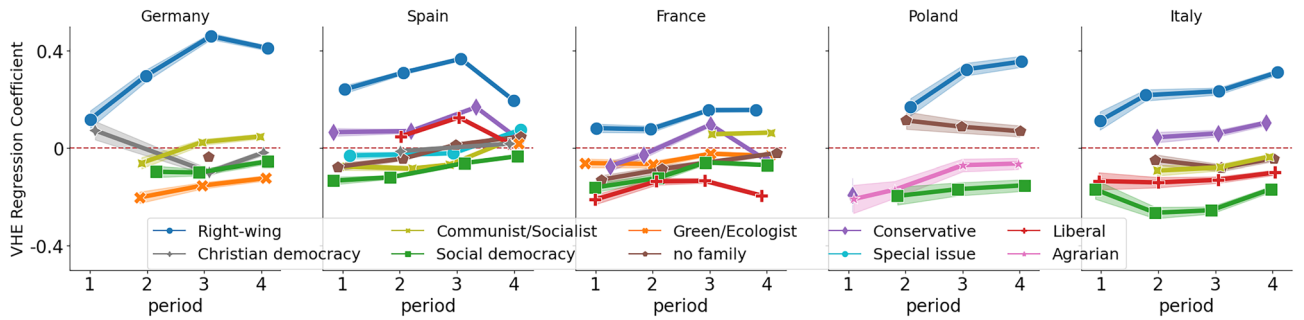
Figure 3. Significant OLS coefficients (at $p < 0.01$ with Bonferroni correction) modeling the users' VHE score using user interest in political parties, grouped in families using ParlGoV. Shaded areas indicate the 99% confidence intervals. Showing countries having sufficient model fit over 4 time periods. For the most part, we find the effect sign of a political party family to remain consistent over time.

that the sign of the coefficients rarely flips. The coefficients for the Right-wing family of parties remain positive, and that for Social Democracy remains negative. However, we find country-specific peculiarities: Liberal parties are associated positively with VHE score in Spain (and not in France or Italy), and the Green parties are more negatively in Germany.

Finally, we investigate the relationship between the VHE score and the four party characteristics as defined by ParlGov, namely Left vs. Right, Liberty vs. Authority, Anti vs. Pro-EU, and State vs. Market. Figure 4 shows the distribution of coefficients for parties in different quintiles of characteristics, accompanying bootstrapped confidence intervals, and horizontal brackets indicating statistical comparison using the one-sided Mann-Whitney U test. For Left vs. Right, Liberty vs. Authority, and Anti vs. Pro-EU dimensions, we find statistically robust differences between parties in the first and last quintiles, and sometimes with the middle quintile as well. Users following Left-leaning politicians are less likely to endorse vaccine-hesitant content, whereas those following Right-wing politicians are much more. Similarly, those following parties closer to the Liberty characteristics – those promoting expanded personal freedoms such as abortion, same-sex marriage, or greater democratic participation – are less likely to endorse vaccine-hesitant content, and the opposite for the parties closer to the Authority side – those promoting authoritarian ideals of order, tradition, and stability. Interestingly, even the party's stance on European Union correlates with the VHE score of users following them: those opposing European Union are more likely to have a higher VHE score. On the other hand, the trend for the economic dimension of State (State-controlled economy) vs. Market (free-market economy) displays a peak in association with the VHE score in the third quintile.

RQ2 Is a user's politicization related to their endorsement of hesitant content?

We address this question by defining two measures of politicization for each user: *political interest* (the proportion of all accounts a user follows that are politicians) and *political focus* (the share of politicians in the user's
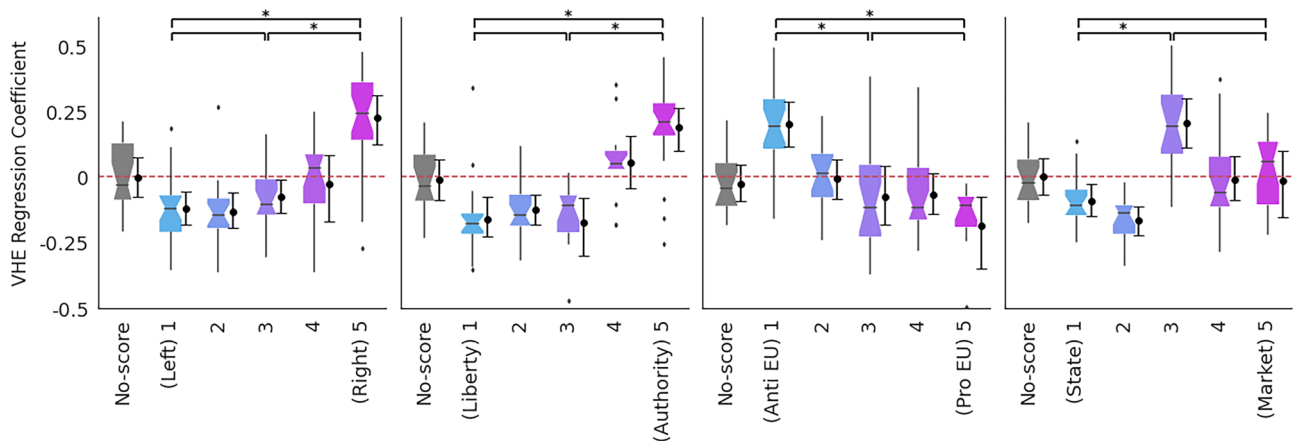
Figure 4. Distribution (boxplots) of OLS coefficients modeling users' VHE score by their interest in parties having one of four dimensions defined by ParlGov, grouped in quintiles. Accompanying points and whiskers indicate a 99% bootstrapped confidence interval. Horizontal brackets on top indicate the comparisons among quintiles 1, 3 and 5, * signifies whether one distribution is statistically greater than the other (one-sided Mann-Whitney U test at $p < 0.01$ with Bonferroni correction).
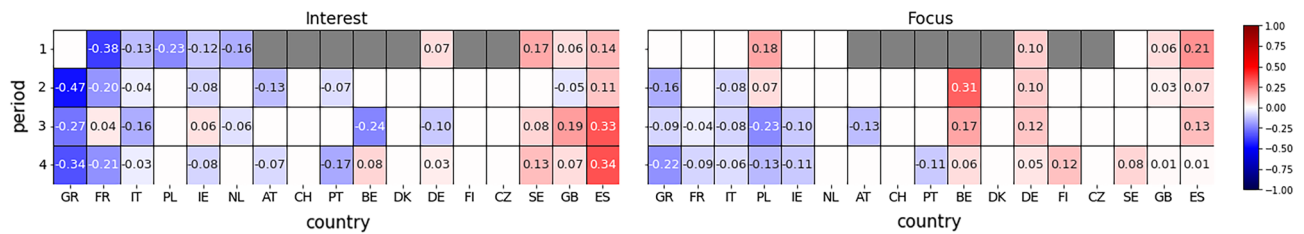
Figure 5. Spearman correlation between political interest and VHE score (left) and political focus and VHE score (right) (see RQ2 Methods). Grey cells indicate country/periods having fewer than 300 users. White cells indicate a non-significant correlation.

most followed party). Figure 5 shows the Spearman rank correlation coefficient between these two measures and the VHE score, for each country/period (those that have fewer than 300 users are greyed out). First, note that the two measures tend to produce similar results for the same countries. Second, the relationship may be different for different countries. For some, there is a positive relationship between both measures and VHE, especially Spain, indicating that (in the case of interest) the greater the share of politicians that a user follows, the more likely they are to endorse vaccine-hesitant content in their timeline. On the other hand, other countries display a negative correlation, especially Greece, suggesting that a greater political interest corresponds to less vaccine-hesitant content. However, this relationship tends to remain constant over time, though can sometimes change, as in the case of Poland when considering political focus. Thus we conclude that the relationship between politicization and VHE is not straightforward, but specific to a particular country and its political situation (echoing our findings in RQ1).

RQ3 Are political actors more influential in the vaccination debate compared to non-political users?

Finally, we compare the engagement with content posted by political actors —the Twitter accounts of prominent politicians— to other users in that country. We match politicians with other users on the number of followers, followees, and daily posting rate in order to achieve a fair comparison. We then consider country/period networks in which there are at least 10 politicians involved in the retweet network, resulting in 44 experiments. First, when comparing the number of retweets these users have received, we find that 40 out of 44 comparisons (91%) do not show a significant difference (using one-sided Wilcoxon test, with Bonferroni correction); only in 4 cases political accounts had more retweets. The result is the same if we consider the number of unique retweeters. Second, measuring the PageRank centrality of both groups, we find only 2 cases when they are significantly different (politicians having higher centrality). Third, considering mentions, we find that in 10 cases (less than a quarter of all cases) politicians are mentioned significantly more than others, but there are no cases of the reverse. In summary, we do not find evidence of political Twitter users being much more influential than other users who have a similar social profile, using the metrics above.

Discussion

A connection between far-right and anti-vaccination movements has been seen on the streets of many European countries, with documented "anti-vax" protests of white supremacists in Britain[29] and arrests of Italy's far-right New Force during COVID-19 riots[30]. However, our study shows that the connection between interest in right-wing political actors and vaccination-hesitant content is not limited to the extreme cases on the streets, but can be found on one of the largest social media websites. We notice that, according to ParlGov, during the study period no right-wing government was in power in the considered countries (with the exception of North League from February 2021 in Italy), thus they were in an opposition role, which may have shaped their audience.

Our findings are further supported by survey evidence. A large survey conducted from December 2020 to January 2022 in Spain has shown that far-right supporters were almost twice as likely to be vaccine-hesitant than the overall population – a consistent trend with a brief lessening around October 2021[31]. A survey of UK residents in April-May, 2020 found that those believing in conspiracy theories around the "authorities" exaggerating COVID-19 deaths and pushing vaccination, among other conspiratorial beliefs, corresponded with a lack of engagement in health-protective behaviours[32]. Elsewhere, a survey of the residents of Norway showed that the refusal to vaccinate is associated with right-wing ideological constraint, with the authors concluding that "vaccine refusal is partly an act of political protest and defiance which attaches itself to consistent right-wing attitudes"[33] (a similar association between right-wing political attitude and refusal to vaccinate was found in a German survey[34]). However, a survey of French respondents found that those both on the right and left extremes were more likely to refuse vaccination[16]. A more refined look at the personal right-wing ideology was used in a survey of respondents in Germany, Poland, and the United Kingdom. When separated into two distinct right-wing dimensions, high "social dominance orientation" (SDO) is associated with higher vaccine hesitancy, while high "right-wing authoritarianism" (RWA) is associated with lower hesitancy, pointing to additional nuance within political ideologies[35]. As we show, political situations and their association with vaccination hesitancy vary greatly amongst countries, more work is necessary to understand the local political incentives around public health communication.

These findings, however, also present an opportunity. The fact that those who are likely to find and endorse vaccine-hesitant information on social media are likely to follow particular political actors points to a possible direct way to communicate with them. If such actors can be persuaded to team up with the public health authorities to promote scientifically-grounded information, their audience would receive such messages from sources they already trust. For instance, the partisan divide in terms of COVID vaccination in the UK has been shown to be less drastic than in the US, likely to the greater pro-vaccination position of the Conservative UK government[36]. Additionally, public health messaging and interventions could be tailored to specific political affiliations or beliefs to better address vaccine hesitancy.

Note that, even though we did not explicitly label content for misinformation, labelers have encountered many instances of possible misleading or erroneous information (which has been found to be posted even by accounts identified as "Business/NGO/government"[37]). Further, we acknowledge that our study does not include bot detection, as bots may play an important role in information propagation[38]. A recent study has shown there is a global proliferation of low-quality information within and across national borders[10]. Whether this information is propagated by political actors is an important question in terms of accountability and responsibility of those representing a public office. For instance, in early 2023, the UK Conservatives suspend a lawmaker for posting vaccine misinformation on Twitter[39].

This study has certain notable limitations. Although the focus of this paper is Europe, by far not all countries were included, either due to data sparsity (despite the multi-lingual queries) or the lack of external resources (such as ParlGov). Twitter users are not representative of the larger populations and tend to be more political[40], potentially overestimating political engagement in the whole country. Although we find some relationships which are stable over time, both public health and political spheres are highly dynamic, limiting our findings to the unique (and unprecedented) time of the rollout of COVID-19 vaccines. Further, keyword-based data collection is always bound to miss content that is phrased differently from the query, with additional challenges in a multilingual setting. However, our decision to keep the keyword set consistent across countries and time was to conserve comparability between countries and to avoid topic drift (for instance, discussion around vaccine-related "green passes" during the vaccination period). The methodology used in this study also favors users who are active and engage in retweeting, excluding those who engage with the topic rarely (previous work shows it is difficult to ascertain the opinions of such users even for human annotators[8]). Observational studies also fail to capture the impact of merely seeing information, that is, whether encountering vaccine-hesitant content actually changes people's minds, although previous studies have found a link between social media use and COVID-19 vaccine hesitancy[41,42]. Complementary methodologies including surveys may be needed to further understand cognitive and psychological nuances of the interaction between social media use, political interest, and vaccine stance deliberation.

Finally, as this study deals with healthcare decisions that may be highly personal in nature, privacy is an important consideration of this work. Because the Streaming API was used to collect this data, it is possible that some of the content was removed either by the user or by Twitter by the time of the analysis. However, besides annotating select tweets, throughout the paper we deal with the structural properties of the RT network and followership of popular political accounts, and present our findings in an aggregated fashion. Further, our methodology identifies pockets of users agreeing on a topic, which could be misused to target or harass people because of their opinions. Any interventions, therefore, possibly based on this analysis must comply with the ethical standards of public communication, and further research possibly dealing directly with individuals must follow the human research guidelines of their institution[43].

Methods

Below, we describe the initial data collection using the Twitter Streaming API, the selection of country-specific content using user geo-location, and the construction of retweet networks that represent one user's endorsement of another's content. We then enrich this data from two sides: by estimating the vaccination stance of the tweets users are likely to encounter in these networks, and by estimating the political interest of these users. Finally, we compare these two sides to answer our research questions.

Data collection

First, in 2019 we compiled a catalog of terms related to vaccines that are translated into 18 different languages (using 2-letter ISO: *bg, cz, fi, de, el, en, es, fi, fr, hu, it, nl, pl, pt, ro, ru, sv, tr*) starting from the list created for the research of tracking vaccination discussion on Twitter[8]. Using Twitter search, we iteratively added keywords to the list and searched again until no new keywords could be found. Then, these keywords were translated by native speakers with the task of including several common grammatical variations or relevant local keywords. This resulted in a collection of 459 keywords, including *vaccine*, *novax*, *measles*, *MMR*, *vaccinated*, and others (for full listing, see Data availability). Once the collection began, we did not modify the list to incorporate new vaccine-related words to keep the selection process consistent and the data comparable across time. This methodological choice likely led to some exclusion of the tweets mentioning new vaccines that did not also use the above keywords (see sec. Limitations).

The final collection, spanning 2019/10/01 - 2021/3/31, is divided into four three-month periods: (1) pre-COVID (October 2019 - December 2019), (2) pre-vaccine (July 2020 - September 2020), (3) vaccine development (October 2020 - December 2020), and (4) vaccine rollout (January 2021 - March 2021) periods (see Supplementary Fig. S1 for a volume graph). The time periods were chosen based on the international news about the vaccine development, though we acknowledge that they fit loosely the vaccine rollout schedule in each individual country. For instance, in period 2 the Sputnik V vaccine was announced (2020/08/11), in period 3 Pfizer-BioNTech (2020/11/09) and Moderna (2020/12/18) vaccines were announced, and at the beginning of period 4, the first AstraZeneca vaccine was administered (2021/01/04). Note that we exclude the period from January to June 2020, when the COVID pandemic first begins. This results in a dataset of 319M of tweets. The distribution across various languages exhibits a marked heterogeneity, with English accounting for 49.0% of the tweets, followed by Spanish at 24.8%, Portuguese at 13.7%, and French at 4.7%. Other languages, including Czech, Finnish, Danish, Romanian, Bulgarian, and Hungarian, each represent a negligible portion of the traffic, constituting less than 0.1% of the total (note that they were queried separately, so the large volume of popular languages should not have affected their collection).

Geo-localization
Next, we assigned to each tweet a location using the self-reported location field provided by users in their description, by matching it to GeoNames[44], a large geographical database of locations. We manually eliminated over 500 words most commonly associated with non-locations for false matches. Using this approach, we were able to geolocate more than 49% of the users. Since some users may provide a fake position or matching may fail (e.g. in the case of homonym places), to reinforce proper geo-localisation, for each country we filtered out tweets written in a language other than the official ones of their country. We tested this approach using tweets that come with geographic coordinates in their metadata as the gold standard, resulting in 93.7% (95% CI [83.2, 100.]) accuracy (which is on average more than twice as accurate as localizing users writing in other languages). See Supplementary Table S1 for accuracy estimates by country by language. Finally, we constrained our analysis to the European countries appearing in the ParlGov dataset (more on this below in the Political Analysis section).

Debate as networks
 Following previous literature[8–10], we represent the vaccine debate in each country using directed, weighted graphs that capture retweet interactions among users. It has been widely shown (e.g. Ref.[45]) that retweet networks have high homophily concerning intrinsic characteristics of the nodes, such that users who retweet each other are likely to share the same opinions and are exposed to the same type of content, thus such networks are often called *endorsement networks*[46]. For each country and period, we built a retweet (RT) network where the nodes are the users in the dataset and the weight on edges from node i to j is the number of times user i retweeted user j's tweets. For each network, we kept the nodes in the biggest *weakly connected component* (WCC), which on average corresponds to 93% of the total (while the size of the second largest WCC is usually less than 1%.). In the following analysis, we considered RT networks with at least 300 nodes from the following 17 countries: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Poland, Portugal, Spain, Sweden, Switzerland, and United Kingdom. This resulted in 61 networks (17 countries × 4 periods), as 7 networks were removed from period 1 due to the size threshold. In the Supplementary Table S2 the sizes of the WCCs in each country by period are listed.

1: $G \leftarrow G(V,E)$ ▷ RT-Network G
2: **for** $i \leftarrow 1$ to 100 **do**
3:      $G' \leftarrow$ Perturb the network G
4:      $\{c\} \leftarrow$ Community detection on G'
5:      $\gamma_{VH}^{(c)} \leftarrow$ VH Exposure in each community c
6: **end for**
7: **for** $u \in V$ **do**   ▷ V nodes of G
8:      $VHE_u \leftarrow$ Average of $\gamma_{VH}^{(c)}$ where c's are the communities to which u was assigned.
9: **end for**

Algorithm 1. Assigning users' VHE score

Vaccination stance labeling

Next, we turn to the labeling of the tweets in terms of vaccination stance (and measuring users' exposure to these stances). To this end, we create a dataset of manually annotated samples of the content shared on the RT networks. To avoid the over-representation of some portions of the social network, we performed a stratified sampling of the tweets first dividing the networks into 15 communities and then selecting 6 tweets per community with the largest difference between the fraction of internal and external retweeters (i.e. the fraction of users from inside a community retweeting the post, minus a fraction of users from outside retweeting it outside). The chosen number of communities and tweets per community are a trade-off between an accurate representation and our capability of manual annotation. Notably, in this step, community detection serves to perform a stratification, this is why the number of communities is the same for all networks. This strategy allowed us to identify tweets that are at the same time popular and representative of a network portion. In this way, we were able to achieve a

high coverage ranging between 37.5% (1st period) and 4.5% (4th period) of the total number of tweets produced by users in the RT network and an average fraction of users covered (i.e. for whom we noted at least one of their posts) between 61.8% to 38.5%.

The selected tweets were then labeled by 13 expert annotators with a background in the vaccine debate and as far as possible proficient in the original language of the tweets. Nevertheless, we translated all tweets into English using Google Translate for cross-checking. The classification task consisted in deciding whether a tweet was *pro-vaccine*, *vaccine-hesitant*, or *other*. In particular, vaccine-hesitant tweets included those stating directly the user will not vaccinate, questioning their efficacy or safety, or espousing conspiratorial views around their creation or distribution. The annotators were encouraged to take into account all information available (images, videos, URLs, etc.) if the tweet was still available online. Annotator agreement, measured using *Cohen's kappa* of $k = 0.47$ showed the task to be moderately complex. However, when the label *other* is excluded, the agreement is $k = 0.78$, with an overlap in labels of 92.3%. We make the labeled dataset of 5667 tweets in 13 languages available to the research community (see Data availability).

With a partial annotation of the tweets at hand, we aim at assigning to each user a score, which we dub *Vaccine Hesitancy Exposure* (VHE) score, capturing the stance of the content they might be exposed to but not necessarily their own stance on vaccination. To accomplish this task, we begin by propagating the labels of the manually annotated data to other users who have retweeted (but not "quoted") them. Next, we apply a procedure summarized in Algorithm 1. In simple words, we randomize the network, perform community detection and assign a score in [0,1] to each individual to be exposed to vaccine-hesitant content, given its class affiliation. The process is repeated 100 times, and the VHE score is obtained by averaging the result over all trials. We now describe in detail the steps in the scheme of the algorithm.

Perturb the network
We generate 100 versions of the network using perturbation methods described in[47]. In this way, we attempt to mitigate weaker clustering signals that may be an artifact of the particular sample of the RT network. We sample randomly 15% of the retweets and change their target, selecting the new one according to the *weighted-in-degree* distribution of the nodes, such that the account popularity information is preserved. We empirically choose a fraction of 15% as the maximum amount of noise we can introduce before the smallest networks' community structure is unrecoverable, while still allowing us to have significant randomization effects on the denser networks of the largest countries.

Community detection
Both for the stratified sampling and for the VHE definition, we perform community detection with the spectral clustering algorithm for weighted networks introduced in Ref.[48], applied to a symmetrized version of our graph. We adopt the "spin-glass" version of the regularized Laplacian matrix and extend its use to more than two communities as per[49]. We choose this algorithm for its speed, its efficiency on weighted, and sparse graphs, but also because it is one of the few known methods to estimate the number of communities in graphs. For interpretation purposes, we split the graph into $k = 15$ communities, whenever the estimated number exceeded this threshold. As a robustness check, we compared the VHE scores created using the Louvain algorithm[50], as well as the results for RQs 1 & 2, and obtained generally consistent results to those using spectral clustering (for more on VHE score robustness checks, see Supplementary Section S4 and additional results in Supplementary Figs. S3–S6).

Stance endorsement within communities
In each community c detected above, we define a measure $\gamma_{VH}^{(c)}$ that captures the extent to which its members are likely to endorse vaccine-hesitant content, defined as the difference between the fraction of hesitant and pro tweets:

$$\gamma_{VH}^{(c)} = \frac{1}{2}\left[\frac{N_{VH} - N_{Pro}}{N_{VH} + N_{Pro} + N_{other}} + 1\right]$$

where N_{VH}, N_{Pro}, N_{other} is the number of vaccine-hesitant, pro-vaccine, and other tweets in the community. The value $\gamma_{VH}^{(c)}$ spans between 0 and 1 and captures how much likely it is that a user in the community c retweets a hesitant tweet instead of a pro-vaccine one.

From community-level to user-level scores
Finally, for each user u we compute their Vaccine Hesitance Endorsement VHE_u score as the mean on the set of 100 values of $\gamma_{VH}^{(c)}$ where the c's are the communities to which u was assigned. The process results in a score for each user from 0 to 1, where 0 means that the user is more likely to retweet a pro-vax tweet, and 1 more likely to retweet a vaccine-hesitant one. Intuitively, having a value greater than 0.5 means that on average the user was sorted into communities where more hesitant tweets were shared than pro-vaccine. As a validation of this approach, we consider the 4th time period (the one having most activity) and annotate a sample of 5 tweets for a sample of users stratified by VHE score in Italy, France, Spain, and United Kingdom, each country having between 225-275 tweets. The Spearman correlation between the the difference between anti- and pro-vax tweets for a user and their VHE score ranged from 0.38 to 0.57, and out of all anti-vax tweets, 50-83% were posted by the users assigned VHE score in the highest tercile (for details, see Supplementary Section S5).

Political analysis

In order to study the connections between VHE score and politics, we identified and characterized the Twitter accounts in our dataset linked to politicians from each respective country.

Identifying politicians and parties

We began by building an extensive list of politicians (parliamentarians and parties' leaders) who are associated with a personal Twitter account using *Politicians on Social Media*[51] and *Twitter Parliament*[52] datasets. Then, we annotated them with the party affiliations in the target period (October 2019 – March 2021) by first matching them to *WikiData*[53] and then manually annotating those not found in WikiData using all available resources. Since parties may be dynamic (dissolution, merging, renaming, etc.) we took as gold standard those parties which are present in the national Parliament during the last government session up to December 2021 provided by the *ParlGov* dataset (the *Parliamentary Governments Database*[28] is an extensive and widely-utilized resource providing detailed information about political parties, elections, cabinets, and governments in parliamentary democracies worldwide). For those parties not in *ParlGov* but which were active in the target period, we either labeled them with a larger party in *ParlGov* that they have joined, or as *Other* if a larger party could not be found. Indeed, it is important to take smaller parties into account when analyzing social media data, since their supporters can be vocal on Twitter[54]. A link to the list of all political users annotated is provided in the Data availability section.

Comparison of user interest in political parties to the VHE score (RQ1)

To compare the user's political interests to their VHE score, we employ an OLS regression model that predicts a user's VHE score using the fraction of politicians followed by them in each party, along with a set of confounding variables. These confounding variables include the number of followers and followees, daily posting rate, weighted in-degree and weighted out-degree (number of retweets they had, and number of retweets they made in the vaccine debate, respectively), and the proportion of followed users who are politicians (political interest, defined above). These variables were selected using the Variance Inflation Factor (VIF), to make sure they do not introduce multicollinearity. We then standardized all features within each country, including the target variable. For each country/period network, we run OLS and note the model fit (Adjusted R^2) and the coefficients of the politically-related variables. We compute 61 such models, one for each network except for the 7 networks that have fewer than 300 users in the WCC. Further, we only consider the models which have a fit of Adjusted $R^2 > 0.1$[55]. To alleviate the multiple hypothesis testing problems, we apply the Bonferroni correction to the p-values of the variable coefficients, selecting those significant at $p < 0.01$, with the correction. Finally, we report aggregated results in terms of the proportion of significant coefficients, their direction (positive vs. negative), and the magnitude and stability of coefficients for select parties.

We perform a similar analysis by aggregating the parties by classification assigned by ParlGov in terms of the political family (conservative, social democrats, liberal, green, etc.), and four dimensions: left/right, State/market (economic policy), liberty/authority (personal freedom), pro/anti-EU. For instance, when aggregating per political family, we run a model for each network that models the VHE score by considering the proportion of politicians followed by the user in a particular political family, along with a set of confounding variables. In the case of the four dimensions, each dimension is run as a separate model, with the numerical score of the dimension binned into quintiles (as well as a "none" score when a party does not have a score in that dimension). For each dimension, the quintile bins were computed on the whole dataset, before being applied to the parties in each country. Similar Adjusted R^2 and Bonferroni-corrected p−value filters were applied to these models. When comparing the average coefficients of the variables in different groups, we compute a confidence interval (CI) using bootstrapping ($n = 1000$). See Supplementary Material Section S7 for details on the features used in and performance of the above regressions.

Comparison of user politicization to the VHE score (RQ2)

To assess the politicization of the users in our data, we collected the followers of the politician identified above using the Twitter Followers API. The share of users who follow at least one politician varies on average between 54% and 88% except for Portugal at only 22%. Note that, as the API does not provide historical knowledge of the follower relationship, here we assume that the followership does not drastically change over time (between the start of our data on October 2019 and the followership collection in December 2022). We then define two measures of user politicization by each user: *political interest* is the proportion of all accounts a user follows that are politicians, and *political focus* is the share of politicians in the user's most followed party (for example, if a user follows 5 politicians in party A, 3 in party B, and 2 in party C, the political focus is 5/10 = 0.5). To make sure the metrics are computed on enough information, we constrain our consideration to users who follow at least 5 politicians when computing political focus (excluding on average 44% (99% CI [20, 93]) users), and to users who follow at least 100 accounts (not necessarily politicians) when computing political interest (excluding on average 9% (99% CI [4, 18]) users). Finally, we compute the Spearman correlation between each of these two measures and the VHE score.

Comparison of politicians to others (RQ3)

Finally, to ascertain the importance of politicians in the vaccination debates in each country, we compare the politicians and their posts to a matched set of other users in terms of retweets and mentions, and their PageRank in the RT network (using a one-sided Wilcoxon signed-rank test). We also considered the unique sets of retweeters and those mentioning the politicians, getting the same results. To perform a fair comparison, we match each political account with another on the number of followers, followees, and the daily posting rate using the

Euclidean distance with the standardization of all variables. The distributions of these variables with matches were checked with paired t-tests to make sure the political accounts were indeed close to the matched baseline.

Data availability

References

1. Centers for Disease Control and Prevention (CDC). Impact of vaccines universally recommended for children-United States, 1990–1998. *Morb. Mortal. Wkly. Rep.* **48**, 243–248 (1999).
2. Flash Eurobarometer 505. Attitudes on vaccination against covid-19 - february 2022. https://www.quotidianosanita.it/allegati/allegato1650373320.pdf (2022).
3. World Health Organization. Ten threats to global health 2019. https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 (2019).
4. Kalimeri, K. *et al.* Human values and attitudes towards vaccination in social media. In *Companion Proceedings of The 2019 World Wide Web Conference*, 248–254 (2019).
5. Browne, M., Thomson, P., Rockloff, M. J. & Pennycook, G. Going against the herd: Psychological and cultural factors underlying the vaccination confidence gap. *PLoS ONE* **10**, e0132562 (2015).
6. Jolley, D. & Douglas, K. M. The effects of anti-vaccine conspiracy theories on vaccination intentions. *PLoS ONE* **9**, e89177 (2014).
7. Dubé, E. *et al.* Vaccine hesitancy: An overview. *Hum. Vaccines Immunother.* **9**, 1763–1773 (2013).
8. Cossard, A. *et al.* Falling into the echo chamber: The Italian vaccination debate on twitter. *Proc. Int. AAAI Conf. Web Soc. Media* **14**, 130–140 (2020).
9. Crupi, G., Mejova, Y., Tizzani, M., Paoletti, D. & Panisson, A. Echoes through time: Evolution of the Italian covid-19 vaccination debate. Proc. *Int. AAAI Conf. Web Soc. Media* **16**, 102–113 (2022).
10. Lenti, J. *et al.* Global misinformation spillovers in the online vaccination debate before and during covid-19. *JMIR Infodemiol.* (2023).
11. Mønsted, B. & Lehmann, S. Characterizing polarization in online vaccine discourse-a large-scale study. *PLoS ONE* **17**, e0263746 (2022).
12. Murphy, J. *et al.* Psychological characteristics associated with covid-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. *Nat. Commun.* **12**, 29 (2021).
13. Jennings, W. *et al.* Lack of trust, conspiracy beliefs, and social media use predict covid-19 vaccine hesitancy. *Vaccines* **9**, 593 (2021).
14. Esser, F., Stepińska, A. & Hopmann, D. N. 28. populism and the media. cross-national findings and perspectives. *T. Aalberg, F. Esser, C. Reinemann, J. Strömbäck & C. Vreese (Eds), Populist political communication in Europe* 365–380 (2017).
15. Guerrero-Solé, F., Suárez-Gonzalo, S., Rovira, C. & Codina, L. Social media, context collapse and the future of data-driven populism. *Profesional de la información* **29** (2020).
16. Peretti-Watel, P. *et al.* A future vaccination campaign against Covid-19 at risk of vaccine hesitancy and politicisation. *Lancet Infect. Dis* **20**, 769–770 (2020).
17. Kreps, S. E. & Kriner, D. L. Resistance to covid-19 vaccination and the social contract: Evidence from Italy. *npj Vaccines* **8**, 60 (2023).
18. Wanat, Z. Poland's vaccine skeptics create a political headache. Politico. https://www.politico.eu/article/poland-vaccine-skeptic-vax-hesitancy-political-trouble-polish-coronavirus-covid-19/ (2021).
19. Funk, Y. C., Tyson, A., Pasquini, G. & Spencer, A. Pew Research. Americans Reflect on Nation's COVID-19 Response. https://www.pewresearch.org/science/2022/07/07/americans-reflect-on-nations-covid-19-response/ (2022).
20. Recio-Román, A., Recio-Menéndez, M. & Román-González, M. V. Vaccine hesitancy and political populism. An invariant cross-European perspective. *Int. J. Environ. Res. Public Health* **18**, 12953 (2021).
21. Stoeckel, F., Carter, C., Lyons, B. A. & Reifler, J. The politics of vaccine hesitancy in Europe. *Eur. J. Pub. Health* **32**, 636–642 (2022).
22. Clark, S. E., Bledsoe, M. C. & Harrison, C. J. The role of social media in promoting vaccine hesitancy. *Curr. Opin. Pediatr.* **34**, 156–162 (2022).
23. Kobayashi, R. *et al.* Evolution of public opinion on covid-19 vaccination in Japan: Large-scale twitter data analysis. *J. Med. Internet Res.* **24**, e41928 (2022).
24. Baron, R. J. & Emanuel, E. J. Politicians should not be deciding what constitutes good medicine. https://www.statnews.com/2022/03/07/politicians-should-not-be-deciding-what-constitutes-good-medicine/ (2022).
25. Garimella, K., Morales, G. D. F., Gionis, A. & Mathioudakis, M. Quantifying controversy on social media. *ACM Trans. Social Comput.* **1**, 1–27 (2018).
26. Ansari, M. T. J. & Khan, N. A. Worldwide covid-19 vaccines sentiment analysis through twitter content. *Electron. J. Gen. Med.* **18** (2021).
27. Reshi, A. A. *et al.* Covid-19 vaccination-related sentiments analysis: A case study using worldwide twitter dataset. In *Healthcare* Vol. 10 (ed. Reshi, A. A.) 411 (MDPI, 2022).
28. Döring, H., Huber, C. & Manow, P. ParlGov 2022 Release, https://doi.org/10.7910/DVN/UKILBE (2022).
29. Saphore, S. White supremacist and far right ideology underpin anti-vax movements. https://theconversation.com/white-supremacist-and-far-right-ideology-underpin-anti-vax-movements-172289 (2021).
30. Broderick, R. Italy's anti-vaccination movement is militant and dangerous. https://foreignpolicy.com/2021/11/13/italy-anti-vaccination-movement-militant-dangerous/ (2021).
31. Serrano-Alarcón, M., Wang, Y., Kentikelenis, A., Mckee, M. & Stuckler, D. The far-right and anti-vaccine attitudes: Lessons from Spain's mass covid-19 vaccine roll-out. *Eur. J. Pub. Health* **33**, 215–221 (2023).
32. Allington, D., Duffy, B., Wessely, S., Dhavan, N. & Rubin, J. Health-protective behaviour, social media usage and conspiracy belief during the covid-19 public health emergency. *Psychol. Med.* **51**, 1763–1769 (2021).
33. Wollebæk, D., Fladmoe, A., Steen-Johnsen, K. & Ihlen, Ø. Right-wing ideological constraint and vaccine refusal: The case of the covid-19 vaccine in Norway. *Scand. Polit. Stud.* **45**, 253–278 (2022).
34. Fischer, H., Huff, M., Anders, G. & Said, N. Metacognition, public health compliance, and vaccination willingness. *Proc. Natl. Acad. Sci.* **120**, e2105425120 (2023).

35. Bilewicz, M. & Soral, W. The politics of vaccine hesitancy: An ideological dual-process approach. *Soc. Psychol. Personal. Sci.* **13**, 1080–1089 (2022).
36. Klymak, M. & Vlandas, T. Partisanship and covid-19 vaccination in the UK. *Sci. Rep.* **12**, 19785 (2022).
37. Kouzy, R. *et al.* Coronavirus goes viral: quantifying the covid-19 misinformation epidemic on twitter. *Cureus* **12** (2020).
38. Gallotti, R., Valle, F., Castaldo, N., Sacco, P. & De Domenico, M. Assessing the risks of 'infodemics' in response to Covid-19 epidemics. *Nat. Hum. Behav.* **4**, 1285–1293 (2020).
39. AP News. Uk conservatives suspend lawmaker for vaccine misinformation. https://apnews.com/article/british-politics-health-united-kingdom-government-f463bd4fdb343a6efb9953ba50b0dfa5 (2023).
40. Bestvater, S., Shah, S., Rivero, G. & Smith, A. Politics on twitter: One-third of tweets from u.s. adults are political. Pew Research. https://www.pewresearch.org/politics/2022/06/16/politics-on-twitter-one-third-of-tweets-from-u-s-adults-are-political/ (2022).
41. Jennings, W. *et al.* Lack of trust, conspiracy beliefs, and social media use predict covid-19 vaccine hesitancy. *Vaccines* **9** (2021).
42. Sallam, M. *et al.* High rates of covid-19 vaccine hesitancy and its association with conspiracy beliefs: A study in Jordan and Kuwait among other Arab countries. *Vaccines* **9**, 42 (2021).
43. CITI Program. Human subjects research (hsr). https://about.citiprogram.org/series/human-subjects-research-hsr/, how="[Accessed on April 20, 2023]" (2023).
44. Geonames. http://www.geonames.org/.
45. Conover, M. *et al.* Political polarization on twitter. *Proc. Int. AAAI Conf. Web Soc. Media* **5**, 89–96 (2011).
46. Garimella, K., De Francisci Morales, G., Gionis, A. & Mathioudakis, M. The effect of collective attention on controversial debates on social media. In *Proceedings of the 2017 ACM on Web Science Conference*, 43–52 (2017).
47. Casas-Roma, J., Herrera-Joancomartí, J. & Torra, V. A survey of graph-modification techniques for privacy-preserving on networks. *Artif. Intell. Rev.* **47**, 341–366 (2017).
48. Dall'Amico, L., Couillet, R. & Tremblay, N. Nishimori meets bethe: A spectral method for node classification in sparse weighted graphs. *J. Stat. Mech. Theory Exp.* **2021**, 093405 (2021).
49. Dall'Amico, L., Couillet, R. & Tremblay, N. A unified framework for spectral clustering in sparse graphs. *J. Mach. Learn. Res.* **22**, 9859–9914 (2021).
50. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
51. Haman, M. & Školník, M. Politicians on social media. the online database of members of national parliaments on twitter. *Profesional de la información* **30** (2021).
52. van Vliet, L., Törnberg, P. & Uitermark, J. The twitter parliamentarian database: Analyzing twitter politics across 26 countries. *PLoS ONE* **15**, e0237073 (2020).
53. Vrandečić, D. & Krötzsch, M. Wikidata: A free collaborative knowledgebase. *Commun. ACM* **57**, 78–85 (2014).
54. Jungherr, A., Jürgens, P. & Schoen, H. Why the pirate party won the German election of 2009 or the trouble with predictions: A response to Tumasjan, a., Sprenger, to, sander, pg, & welpe, im predicting elections with twitter: What 140 characters reveal about political sentiment. *Soc. Sci. Comput. Rev.* **30**, 229–234 (2012).
55. Ozili, P. K. The acceptable r-square in empirical modelling for social science research. In *Social Research Methodology and Publishing Results: A Guide to Non-Native English Speakers* (ed. Ozili, P. K.) 134–143 (IGI Global, 2023).
56. Lenti, J. Global misinformation spillovers in the online vaccination debate before and during COVID-19, https://doi.org/10.5281/zenodo.7716817 (2023).

Acknowledgements

Author contributions

M.T., Y.M., J.L., and G.P. collected the data, G.P. conducted the experiments, G.P., Y.M., M.S., K.K., M.T., L.D. wrote the manuscript. All authors participated in the experimental design, interpretation, and review of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-54863-7.

Correspondence and requests for materials should be addressed to Y.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.