I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17749

Narratives of Migration and Political Polarization: Private Preferences, Public Preferences and Social Media

Eugenio Levi
Michael Bayerlein
Gianluca Grimalda
Tommaso Reggiani

MARCH 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17749

Narratives of Migration and Political Polarization: Private Preferences, Public Preferences and Social Media

Eugenio Levi
Link University and Masaryk University

Michael Bayerlein
German Institute for International and Security Affairs

Gianluca Grimalda
University of Passau and South East Technological University of Waterford

Tommaso Reggiani
Cardiff University, Masaryk University and IZA

MARCH 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Narratives of Migration and Political Polarization: Private Preferences, Public Preferences and Social Media*

We study how preferences for migration-related narratives differ between private and public contexts and how social media fuel opinion polarization. Using a German representative sample (n=1,226), we found that individuals, especially from the left and center, avoided publicly endorsing anti-migration narratives. In an experiment on Twitter (n=19,989) we created four Twitter profiles, each endorsing one of the narratives. Far-right users exhibited markedly different engagement patterns. While initial public endorsements, measured by follow-back rates, aligned with private preferences, social media interactions amplified support for the most hostile and polarizing narrative. We conclude that social media significantly distort private preferences and amplify polarization.

JEL Classification: D72, D91, C93

Keywords: immigration, narratives, political polarization, economic reciprocity, survey experiment, field experiment, group identity, social media, Twitter

Corresponding author:
Eugenio Levi
Link University
Via Casale San Pio V 44
00165 Rome
Italy

E-mail: e.levi@unilink.it

* We thank all participants at the 2021 Applied Micro Workshop at the Free University of Bozen-Bolzano, MUES seminar at Masaryk University, M-BEES workshop at Maastricht University, SIEPI Conference at the University of Naples Federico II, IMEBESS 2023, 2023 CEPR Conference on the Political Economy of Populism, seminar at Science Po, for their useful comments and suggestions. We also thank James Pennebaker, Tabea Meier and Linda Bos for dictionary resources, and Andrea Guido and Shaun Hargreaves Heap for helpful suggestions. Any remaining errors are our own. Eugenio Levi gratefully acknowledges the financial support from the Czech Ministry of Education and Masaryk University through the program MSCAfellow3@MUNI (reg. no.: CZ.02.2.69/0.0/0.0/19_074/0012727). Tommaso Reggiani gratefully acknowledges the support of the European Regional Development Fund project Foreign Interference in The Context of Geopolitical and Technological Change (reg. no.: CZ.02.01.01/00/23_025/0008692). Pre-registration: OSF https://doi.org/10.17605/OSF.IO/Z6SGN for survey 1, https://doi.org/10.17605/OSF.IO/2T46H for the social media experiment, and https://doi.org/10.17605/OSF.IO/Q6FEV for survey 2. IRB approval: Central Ethics Committee of the University of Kiel #19-21 and #20-21.

1 Introduction

Since social media achieved mass diffusion, their role in fostering political polarization has been widely debated. On the one hand, social media are seen by many as distorting political preferences, thus spurring polarization. This may be the case because the functioning of social media makes it more likely that extreme views become amplified through the creation of "echo chambers" or the diffusion of "fake news" (Sunstein, 2018; Cinelli et al., 2021; Lazer et al., 2018). Moreover, individuals with extreme views may be more active on social media (Schumann et al., 2021), or they may be prompted to express different opinions in public as opposed to private settings (Lelkes, Sood, and Iyengar, 2017; Vosoughi, Roy, and Aral, 2018; G. Levy and Razin, 2019; Zhuravskaya, Petrova, and Enikolopov, 2020; Alesina, Miano, and Stantcheva, 2020). On the other hand, social media may simply reflect actual political preferences, some of which may have previously been underrepresented in the public sphere (Pellert et al., 2022; Groseclose and Milyo, 2005), possibly because of pluralistic ignorance (Prentice and D. T. Miller, 1996; Bicchieri, 2016; Bursztyn, González, and Yanagizawa-Drott, 2020; D. Miller, 2023). Our knowledge of the extent to which social media distort or reflect privately held opinions is scant, given the obvious difficulty of capturing opinions both privately and publicly.

Several studies gauged the extent to which fake news become widespread on social media and how people can be educated not to fall for them (Henry, Zhuravskaya, and Guriev, 2022; Allcott and Gentzkow, 2017; Pennycook and Rand, 2019; Barrera et al., 2020; Costello, Pennycook, and Rand, 2024). Our approach is different. We hold facts constant and study how the *narrative* of such facts get amplified or stifled when moving from a more private to a public context[1]. Our focus is on narratives of immigration, one of the most divisive issues in Western politics (Alesina and M. Tabellini, 2024)[2]. Participants were asked to rank six proposed narratives (3 in acceptance of migration and 3 hostile to migration) according to how much they agreed with them: we define the outcome of this ranking as a participant's 'preferences' over the narratives[3]. We observe how these preferences over narratives differ in three different settings at the time of the 2021 general elections in Germany: (1) A survey on a representative sample of the German population in which preferences over narratives are revealed to the experimenter under anonymity but are not made public. This is our 'private' setting. We screen participants according to their voting intentions in the upcoming elections, oversampling voters at the extreme of the political spectrum. (2) A decision, included in the same survey as (1), as to whether individuals want to publicly endorse one of the narratives on an internet website, as in Romano et al. (2021). Participants are free to opt-out so that we can infer whether a possible change in the manifestation of preferences is due to the intensive margin, i.e., people manifest different opinions in the two contexts, or to the extensive margin, i.e., some people decline to enter the public sphere. (3) A field experiment on the popular social media platform Twitter (now X). We created four artificial profiles, which were identical in terms of layout but differed in their pinned tweet. This tweet reproduced one of the four most preferred narratives in the private context. After observing

[1] By narratives we mean the stories that are interwoven around facts, in order to make sense, interpret, explain, and explore the ramifications of such facts, possibly triggering concerns or emotions (Shiller, 2017; G. Akerlof and Snower, 2016).

[2] Stories that fuel hostility against foreigners or minorities have always been powerful. For example, the Protocols of Elders of Zion were forged by czarist propaganda in Russia in 1903 to induce antisemitism. It was a fictional story, but people took it for real, and even though by 1935 the story of the forgery had come out and was officially confirmed, the Protocols are still published and widely distributed even today.

[3] While narratives are different from social preferences or consumer goods, we follow Eliaz and Spiegler (2020) and Schwartzstein and Sunderam (2021) in assuming that individuals choose their preferred narratives based on some utility function (see section 2).

Twitter users' activities on the platform, we classified n=19,989 of them according to their political orientation. This is our "subject pool". Then, as in Mosleh et al. (2021) our experimental "stimulus" consisted of having the four profiles following one randomly selected Twitter user from our pool. Our primary observed outcome was the frequency of the follow-back by the "treated" subjects to our four profiles. We also monitored the total number of engagements, i.e., visualizations, replies, likes and retweets of our four pinned tweets.

We find that narratives matter to people. Even within the acceptance/hostility dimension, people's private views on which narrative they agree the most differ substantially both within and across political orientation. Interestingly, we find some general patterns that hold across all political orientations. Secondly, when participants had the option to publicly support one of the narratives on a website, we found a substantial modification of the manifestation of individual attitudes. As a result, pro-immigration narratives become more visible in the public compared to the private setting. This change is driven by people from the centre-left and the centre-right of the political spectrum who tend to shy away from the public setting when they hold hostile narratives privately. We also show that narratives over COVID-19 are less divisive than those concerning migration.

In the Twitter field experiment, we find that our primary observational variable, i.e., the follow-back rate, mirrors rather closely the private preference over narratives. However, as interactions grow, the most hostile and divisive narrative - according to a polarization index we develop - is the one that gets overwhelming traction by a factor of at least ten to one. This is the case across all four measures of engagement we use. This result might be due to the fact that far-right users are disproportionately more active than other users. This is the case both in following-back - by a factor of three to one - and in spreading the hostile divisive narrative. We cannot rule out that the functioning of the algorithm also plays a role in this result. Overall, the transition from private considerations to social media appears to distort the public portrayal of migration preferences by amplifying divisive, anti-immigration narratives, thereby intensifying polarization.

We contribute to several strands of literature. Firstly, our study is relevant to the quickly growing literature on social media, populism, and polarization. Some experimental studies investigated the effect of exposing social media users to news from the political side in opposition to their own. R. Levy (2021) shows that this manipulation decreases polarization on Facebook while C. A. Bail et al. (2018) find the opposite on Twitter. In general, the evidence on whether social media lead to more polarization is mixed. For instance, Fujiwara, Müller, and Schwarz (2024) found no positive effect of Twitter use on the Republican vote share in the 2016 and 2020 presidential elections, while Boxell, Gentzkow, and Shapiro (2017) found a greater increase in polarization in age groups that are less likely to use social media. In contrast, other studies focusing on the US show that social media may be a force behind political polarization (Lelkes, Sood, and Iyengar, 2017; Allcott, Braghieri, et al., 2020). In the same vein, Manacorda, G. Tabellini, and Tesei (2022) provide evidence of a positive effect of mobile internet access on the support for right-wing populist parties in Europe, particularly in economically deprived areas, where communitarian policy views, intolerance of strangers, and nationalistic tendencies were more likely to spread. This reasoning is based on earlier work by G. Tabellini (2008) about an in-group (communitarian) versus out-group (universalistic) cleavage structuring contemporary politics. Messages of in-group love and especially out-group hate are, given their emotionally charged content, particularly likely to spread over social media (see also Crockett, 2017; Rathje, Van Bavel, and Van Der Linden, 2021)[4], even though we know much less about what happens when the out-group consists

[4]The effect is well-established on social media. Its effect in proper political campaigning is much more ambiguous

of immigrants. Our study differs from those above by combining a field experiment on social media with survey evidence, thus enabling us to understand how polarization evolves from private/public settings to social media and its underlying determinants.

Secondly, migration issues have become a key concern for studies analyzing public opinion and electoral competition. Previous studies focused on changes in attitudes toward migration (see e.g., Facchini and Mayda, 2009; Facchini and Mayda, 2012; David Card, Dustmann, and Preston, 2012), the rise of radical right-wing parties (see e.g., Dinas et al., 2019; Dustmann, Vasiljeva, and Piil Damm, 2019; Edo et al., 2019; Levi, Mariani, and Patriarca, 2020), and on the demand for and supply of (anti-)migration policy positions, especially in the wake of the increased influx of migrants and asylum seekers to Europe in 2015 and 2016 (see e.g., Marx and Naumann, 2018; Brug and Harteveld, 2021; Steinmayr, 2021). These papers usually focus on the geographical distribution of immigration flows to ascertain the impact of migrants while we look at an alternative (and potentially more relevant) channel, i.e. popular narratives about migration and their diffusion in social media.

Thirdly, we contribute to the growing literature on narratives in their role of shaping attitudes, preferences, and decision-making (Shiller, 2017; Shiller, 2020; G. Akerlof and Snower, 2016; Eliaz and Spiegler, 2020; Schwartzstein and Sunderam, 2021; Antinyan et al., 2024). Galasso, Morelli, et al. (2022) finds that emotionally charged narratives aiming at discrediting populist politicians are more effective than objective information in reducing support for the populist agenda. Contributions addressing the role of migration narratives have found that narratives of assimilation and symbolic representation of national identity reduce opposition to immigration (Kaufmann, 2019; Wright and Citrin, 2011). In general, research has shown that migration narratives can affect hostility toward immigrants when based on misinformation (Barrera et al., 2020; Hameleers and Van der Meer, 2020; Schäfer and Schadauer, 2018), if focusing on immigrants' crimes (Manzoni et al., 2024; Keita, Renault, and Valette, 2024) or if using slant language (Djourelova, 2023). Some observational evidence already points to politicians increasingly recurring to harsh rhetoric on migrants (Dallas Card et al., 2022), as demonstrated in Donald Trump's presidency (see Figure J1).

In line with this research, Gehring et al. (2022) analyze the migration narratives in German newspapers in a large dataset of articles between 2000 and 2019. They find cultural narratives are more frequent than economic ones, with religious differences generally portrayed negatively. Gehring et al. believe this is instrumental to fostering anti-migration attitudes in Germany and the rise of the far-right Alternative für Deutschland (AfD henceforth) party. Moreover, it has also been shown that offering some rationale for out-group hate increases the likelihood of sharing anti-migrant messages on social media (Bursztyn, Haaland, et al., 2020) or for justifying joining an anti-immigrant organization (Bursztyn, Egorov, et al., 2022). Compared to this literature, we show how even minor manipulations in narratives based on established theories can matter. Moreover, our analysis of narratives extends to multiple settings, allowing us to assess if and how the "medium makes the message".

Moreover, our study is related to the extensive literature on so-called pluralistic ignorance. This has been defined as a situation in which the plurality of a group is ignorant of, or misperceives, its own beliefs, perceptions, and practices (D. Miller, 2023). In other words, the group experiences second-order misperceptions of first-order beliefs (Bicchieri, 2016).[5] Pluralistic ignorance has been ascertained in several contexts (Bursztyn, González, and Yanagizawa-Drott, 2020; Andre et al., 2024) and has been

(Galasso, Nannicini, and Nunnari, 2023). The literature on "negative campaigning" has found mixed results.

[5]While pluralistic ignorance is a group-level phenomenon, the false consensus effect, i.e., the belief that the majority holds attitudes in line with the individual's attitudes (Ross, Greene, and House, 1977; Marks and N. Miller, 1987; Engelmann and Strobel, 2000), holds at the individual level.

associated with the rise of populist parties in Western countries (Barrera et al., 2020). In our study, we find a significant modification of opinions in a private and public context. This is compatible with pluralistic ignorance, although we show that alternative mechanisms could explain it. We also find that change in preferences occurred with respect to migration narratives but not for those concerning anti COVID-19 vaccination, another divisive topic at the time of our study.

Even if the narratives in our design report factually correct information, our study is also relevant to the literature focusing on misinformation and social media interaction (Allcott and Gentzkow, 2017; Grinberg et al., 2019; Schäfer and Schadauer, 2018). A relevant role for social media algorithms in feeding information consistent with the user's profiled ideology has been ascertained (C. Bail, 2021; Bowen, Dmitriev, and Galperti, 2021; Garz, Sörensen, and Stone, 2020; R. Levy, 2021; Wojcieszak and Garrett, 2018). This results in the formation of "echo chambers" or "filter bubbles", in which users only process the news in line with their views and exclusively engage with like-minded individuals (Acemoglu, Ozdaglar, and Siderius, 2021; Cinelli et al., 2021; Mosleh et al., 2021; Pariser, 2011; Sunstein, 2018; Stein, Keuschnigg, and Rijt, 2023). In our field experiment on social media we qualify these findings, showing a widespread tendency by far-right individuals to engage with users from different political views from their own, although in the context of a generalized attitude of engaging more with tweets expressing more polarized views[6].

We focus on Germany both for its international political and economic relevance and because it has attracted large numbers of immigrants after the Syrian civil war. AfD reached 10.3% of the votes in the 2021 national election and was the second most voted party, polling at 16% of the votes, in the 2024 European elections. We exploited a specific time frame in German politics for fielding our study: the run-up to the 2021 general election. This was very important to make the narratives more salient, and to increase the perceived impact of the publicly endorsed narratives on German politics.

The paper is structured as follows: Section 2 discusses the theoretical background of our research together with our hypotheses, while Section 3 introduces the narratives, which are the cornerstone of our experimental design. Section 4 discusses the experimental design and the results related to private preferences, while Section 5 presents our results related to public preferences on an internet website. Section 6 provides the experimental design and the results of the social media field experiment. Section 7 explains the mechanisms and Section 8 concludes the paper.

2 Theoretical background

We consider narratives that vary across two dimensions: (a) acceptance or hostility toward immigrants; and (b) the group on which the narrative focuses, distinguishing between a focus on the in-group (i.e., co-nationals), on the out-group (i.e., immigrants), or on reciprocity between the two groups (Tajfel et al., 1979; Brewer, 1999; Rand et al., 2009; Columbus et al., 2023). We posit that an agent's preferences for narratives are determined by two basic components. First, an agent has a *private* preference for narratives. We define a narrative as an element $\upsilon \in N$ where N is the set of possible narratives.

[6] Lastly, we contribute to the strand of literature on the rise of the AfD party. Recent analyses have focused on explaining the determinants of support for a right-wing populist party in Germany (Arzheimer, 2015; Berbuir, Lewandowsky, and Siri, 2015). Despite focusing on nativism and hostility towards migration as specific policy issues (Arzheimer and Berning, 2019; Bieber, Roßteutscher, and Scherer, 2018; Hambauer and Mays, 2018; Pesthy, Mader, and Schoen, 2021), other contributions have analyzed the role of narratives and misinformation, especially on social media (Sängerlaub, M. Meier, and Rühl, 2020; Serrano et al., 2019; Weisskircher, 2020). Consequently, descriptive research on social media and AfD support draws a similar conclusion as previous research by finding the most homogeneous networks among AfD supporters (Gärtner and Wuttke, 2019) and an increased social media engagement by AfD supporters (Schumann et al., 2021).

We assume that an agent has her own ranking over such elements $\upsilon \in N$. We assume that such an individual ranking may be represented by a utility function $U_i(\nu, ., .)$. Let us then define the following:

$$\bar{\upsilon}_i = \underset{\upsilon \in N}{argmax} U_i(\nu, ., .) \tag{1}$$

$\bar{\upsilon}_i$ is the narrative that maximizes the utility function with respect to the private component exclusively. We have no prior on the relative popularity of the reciprocity narratives. On the basis of existing evidence, we hypothesize that out-group narratives will be chosen by most agents as their preferred ones on both sides of the political spectrum (e.g. Rathje, Van Bavel, and Van Der Linden, 2021), albeit in relation to opposing policy positions on migration.

H1: Narratives with an out-group focus will be preferred to the in-group/reciprocity ones. However, far-left and left-wing individuals will prefer narratives in acceptance of migration more than hostile ones, while the opposite holds for far-right and right-wing individuals.

Second, we posit that an agent has a *public* preference over narratives. By public preference, we mean that an agent has an ordering of narratives that depends on the agent's social image in supporting a certain narrative rather than another one. This is based on the idea that individuals derive utility from their self-esteem, which is in turn dependent on the individual's second-order belief in others' esteem of self (G. A. Akerlof and Kranton, 2000; Bénabou and Tirole, 2006; Bénabou, Falk, and Tirole, 2018). In the present setting, we assume that public preferences are mediated by inter-group relationships and beliefs (Tajfel et al., 1979). We assume that an agent gains utility by publicly endorsing narratives that are the most *distinctive* for the political group they support. This idea is drawn from 'optimal group distinctiveness' theory (Leonardelli, Pickett, and Brewer, 2010) or self-categorization theory (Turner et al., 1987) in social psychology. The main idea of these theories is that individuals strive to identify themselves with their own group and differentiate themselves from other groups.

We adopt the following specification. Let us assign each i to a group $\bar{I}_i$ to which i feels to belong. In line with G. A. Akerlof and Kranton (2000), $\bar{I}_i$ represents the "reference group" from which i derives her self-esteem. We assume that the four political groups that we considered provide a partition of the space of 'belongingness'. In other words, each i belongs to one and only one of the four $\bar{I}_k, k = \{FL, L, R, FR\}$, where the meaning of the acronyms k is self-explanatory. Given $\bar{I}_i$, that is, i's ingroup, we also identify $-\bar{I}_i$ as the set of agents to which i does not belong, i.e., her outgroup.

To define distinctiveness, we consider the first- and second-order beliefs over the 'popularity' of a narrative in $\bar{I}_i$ and $-\bar{I}_i$. By 'popularity', we mean the share of people in the relevant group who prefer that narrative among the six. Hence, the sum of the beliefs over popularity must sum up to one. The first component of distinctiveness is, simply stated, the desire to please others in one's group. This is determined by the first-order belief over the popularity of a narrative in $\bar{I}_i$. The second component of distinctiveness is the desire to displease others in the outgroup. This is determined by what i believes that others in her ingroup believe is unpopular in $-\bar{I}_i$. That is, this is the second-order belief over the popularity of narratives in the outgroup. In formulas:

$$Dist_{i\upsilon} = \frac{\pi_i^1(P_{\bar{I}\upsilon})}{\pi_i^2(P_{-\bar{I}\upsilon})} \tag{2}$$

where $\pi_i^1(P_{\bar{I}\upsilon})$ is i's first-order belief over the popularity of narrative υ in group $\bar{I}_i$ and $\pi_i^2(P_{-\bar{I}\upsilon})$ is i's second-order belief over the popularity of narrative υ in group $-\bar{I}_i$. $Dist_{I\upsilon}$ close to 1 means that the narrative υ was perceived as equally preferred by supporters of one's own and other parties. The higher the ratio, the higher the perception that a specific narrative will be liked by members of the same party as the subject and disliked by other parties' members.

Finally, we identify the most distinctive narrative for each political group as the maximum of (2):

$$\hat{\nu}_I = \max_{\upsilon \in N}(Dist_{I\nu}) \tag{3}$$

To be sure, public preferences may depend on other motivations concerning the public sphere. Agents may be motivated by just one of the two components of distinctiveness. Or they could follow expressive rationality, that is, acting in the public sphere as if they would be pivotal in public decision-making (Kahan, 2013). Or they may shy away from supporting a narrative perceived as racist out of pluralistic ignorance (D. Miller, 2023), as suggested in the introduction. We abstract away from these motivations in the utility function, but we will give a cursory discussion of some of them in section 5.2.4 and in Appendix I.

We assume that private and public preferences are combined linearly in agents' utility function:

$$U(\nu, \hat{\nu}_i) = (\nu - \bar{\upsilon}_i)^2 + \alpha (\nu - \hat{\nu}_i)^2 \tag{4}$$

$\alpha \geq 0$ is the weight associated with the public "social image" components compared to private preferences. Upon maximization of (4), the overall optimal action for the agent will then be the following:

$$\upsilon^* = \underset{\upsilon \in N}{argmax} U(\upsilon, \hat{\upsilon}_i) = \frac{\bar{\upsilon} + \alpha\hat{\upsilon}}{1+\alpha} \tag{5}$$

It is then possible that $\upsilon^* \neq \bar{\upsilon}$. We can in general hypothesize the following:

H2: Preferred narratives may differ between a private and a public setting, with choices being closer to the most distinctive narrative for one's group in the public setting.

To evaluate changes in polarization across settings, we develop an index of overall polarization. In this case, we only consider the component of others' reactance always following "optimal distinctiveness theory". We operationalize this notion considering both an individual's most preferred narrative, as expressed by her answers or behavior, and the most disliked ones by other parties' members. The latter is the narrative that ranks the lowest when individuals are asked to rank the statements at the beginning of Survey 1. Intuitively, suppose an agent chooses a narrative that is most disliked by a high (low) number of other parties' members. In that case, the polarization score for this individual will be high (low). We then aggregate such individual-level scores of polarization across individuals belonging to the same parties and then across parties, weighing each aggregation by the relevant number of members of each party. Formally, the polarization index for each party p in setting s is:

$$PolIndex_{ps} = \frac{\Sigma_{i_p} ShDis_{i_p s}}{n_{ps}} \tag{6}$$

where:

$$ShDis_{i_p s} = D_{ijs} \frac{\Sigma_{-p}\Sigma j \frac{Dis_j}{n_{-ps}}}{3} \tag{7}$$

and:

$$D_{ijs} = \begin{cases} 1 & if\ Pref^*_{is} = Dis^*_j \\ 0 & otherwise \end{cases} \tag{8}$$

$ShDis_{i_p s}$ in equation (7) is the share of other parties' supporters who disliked the preferred narrative of participant i_p the most. n_{ps} in (6) is the total number of observations from party p in setting s, while n_{-ps} in (7) is the number of observations from a party different from p in setting s. D_{ijs} in (8) is an indicator function taking value 1 if the preferred narrative by participant i in setting s is the most disliked one by participant j supporting another party: $Pref^*_{is} = Dis^*_j$.

The interpretation of the index is straightforward. The more the preferred narratives by voters of a party are disliked by voters of other parties, the higher the polarization index for that party, and *vice versa.* A $PolIndex_p$ equal to zero for party p means that no supporter of other parties disliked the most the preferred narratives within party p. A $PolIndex_p$ of 100 means that voters of party p all preferred the same narrative, and this narrative was the most disliked by all voters of all other parties.

3 Narratives construction

Central to our design is the use of various narratives in all of our settings. Even if the so-called "Grand Koalition"[7] that governed Germany during the influx of Middle Eastern asylum seekers in 2015/16 was substantially in favor of immigration,[8] a language analysis of their communication on Twitter reveals substantial differences between its supporting parties[9]. To avoid any effect related to the messenger of the narratives, we build artificial narratives that have minor variations based on established theories but that are similar to those found in the actual debate on immigration. All narratives present the same fact, that is, the number of people who migrated to Germany from 2015 to 2020[10]. Our main experimental manipulation was then to construct different and opposing ramifications of this fact in

[7] This was an alliance formed by the most prominent political parties over the period 2013–2021. It was formed by the Christian Democratic Union (CDU), the Christian Social Union (CSU) and the Social Democratic Party (SPD). The CDU and CSU are center-right parties who have also supported migration historically but have also been critical as to the risks of migration. The SPD is a party of the centre-left firmly rooted in social democracy with historical support of migration. Additionally, "Die Grüne" ("The Greens") also took a pro-migration stance, demanding an even more generous asylum regime.

[8] Support for immigration was justified on both humanitarian and economic grounds. However, this was not unconditional, as ex-chancellor Angela Merkel was famously at odds with the then minister president Horst Seehofer who demanded a quantitative limit to the amount of immigrants Germany would take in 2015.

[9] We applied the Linguistic Inquiry and Word Count (LIWC) dictionary (Tausczik and Pennebaker, 2010; T. Meier et al., 2019), a psychometric dictionary commonly used in psychology, and Moral Foundation Theory (MFT) dictionary (Bos and Minihold, 2022) to tweets by German politicians from 2016 to 2021. See Appendix A for technical details about this analysis. The center-right party CDU used words referring to group identity about 25%, 50%, and twice as much as the center-left SPD, the Greens, and the AfD, respectively. The SPD and the Greens tended to use words referring to fairness twice as often as the CDU and the AfD. The CDU attributed positive moral traits to native Germans about 40% of the times more than the SPD and the Greens, and nearly three times as much as the AfD. Moreover, words such as "family" or "home country" were used by all parties, but the AfD almost never used words like "shared" and "together". In addition, the CDU further avoided using words with the root "equal'.

[10] Based on the Federal Statistical Office of Germany, this number is approximately 10 million (https://www.destatis.de/EN/Press/2022/06/PE22_268_12411.html). We considered inflows and not net migration, as we wanted the key fact around which narratives were built to attract participants' attention as much as possible. We acknowledge that this number may have primed people to take a stand against migration. However, this choice cannot affect our results because narratives are held constant across the various settings. In fact, we do not find a higher hostility to migrants in our study than what is usually found in surveys. For example, based on Eurobarometer data from 2017 (https://migrant-integration.ec.europa.eu/sites/default/files/2021-05/ebs_469_infographics_en.pdf), 38% of people in Germany think that immigration is more of a problem than an opportunity compared with 35% who think the opposite; we found 48% vs. 52% in our data.

the cultural and economic domain[11]. We needed short narratives, both because of the Twitter limit on the number of characters and because we did not want participants in the survey to get inattentive. Therefore, narratives included only two sentences, the first one referring to cultural values, the second to economic aspects[12]. The first dimension is the general stance on the migration issue. We restricted to two opposing stances, one of acceptance of migration (A), one of hostility (H). The second dimension concerned the arguments in support of this general stance. The narratives emphasized the in-group (I), the out-group (O) or reciprocity motives (R). This approach generated a 3×2 design in terms of the narratives' policy positions and focus. More specifically, the two narratives focusing on the out-group were the following:

Hostility + Out-group focus (H/O): From 2015 to 2020, almost 10 million migrants arrived in Germany. The unacceptable values and practices of many of these immigrants are incompatible with our cultural lives. Furthermore, immigrants also have job skills and work attitudes that threaten to harm our economy permanently.

Acceptance + Out-group focus (A/O): From 2015 to 2020, almost 10 million migrants arrived in Germany. The values and practices of many of these immigrants can enrich our cultural lives. Furthermore, immigrants also carry the job skills and work attitudes that are needed for our economy.

The first sentence provides the central fact of migration, which is common across narratives. The two narratives focus on the same aspect of migration—who immigrants are, that is, the out-group—but build two opposing policy positions depending on how migrants are described vis-a-vis Germans both in cultural terms ("incompatible with our cultural life" or "can enrich our cultural life") and in economic ones ("menace to permanently harm our economy" or "are needed for our economy"). This type of narrative can be found in the actual political debate from Germany's main parties.[13]

In turn, we used the following narratives for the in-group focus:

Hostility + In-group focus (H/I): From 2015 to 2020, almost 10 million migrants arrived in Germany. The German values and practices we hold so dear have to be preserved from migrants. Furthermore, Germans have all that they need to sustain a strong economy, even without immigrants.

Acceptance + In-group focus (A/I): From 2015 to 2020, almost 10 million migrants arrived in Germany. The German values and practices we hold so dear can be relied upon to live peacefully with migrants. Furthermore, Germans have all that they need to sustain a strong economy, even together with immigrants.

The focus here is on how Germans can benefit or be harmed by migrants rather than on how migrants can benefit or harm Germany. "German values and practices" and the Germans' "strong economy" are the subjects of these narratives. The content of these narratives in terms of meaning is exactly the same as in the previous ones, with the same ramifications in cultural and economic

[11] We decided to go for both cultural and economic domains because these are the main competing explanations in the scientific literature to hostility towards immigration (Alesina and M. Tabellini, 2024).

[12] The political economy literature puts forward a lot of possible explanations for opposition to migration. Among them, competition in the labor market (Mayda, 2006; Edo et al., 2019; Moriconi, Peri, and Turati, 2022; Mayda, Peri, and Steingress, 2022), for welfare expenses (Facchini and Mayda, 2009; Alesina, Miano, and Stantcheva, 2023), cultural distance (Harmon, 2018; Brunner and Kuhn, 2018), integration costs (Levi, Mariani, and Patriarca, 2020), etc. See Levi, Mariani, and Patriarca (2023) or Alesina and M. Tabellini (2024) for reviews on this topic. In creating the narratives, we referred to these arguments in general terms. For example, the reference to the "job skills" of migrants in negative terms is derived from theories on competition in the labor market among low-skilled workers.

[13] An example of the H/O narrative can be found in the AfD campaign (https://www.afd.de/wahlprogramm-asyl-einwanderung/). The A/O narrative can be identified in the the CDU campaign (https://www.gruene-niedersachsen.de/fluechtlingspolitik-migration-und-teilhabe/).

domains. Still, the subject and object are inverted, leading to a fundamental shift in the focus from the out-group to the in-group. Again, these narratives are similar to those used in the German political debate.[14]

Finally, the two narratives focusing on reciprocity are:

Hostility + Reciprocity (H/R): From 2015 to 2020, almost 10 million migrants arrived in Germany. Germans and immigrants have different values and practices, as well as different job skills and work attitudes. The integration of immigrants into our society represents too costly an investment, and the costs to integrate them will never be compensated for in the future.

Acceptance + Reciprocity focus (A/R): From 2015 to 2020, almost 10 million migrants arrived in Germany. Germans and immigrants have different values and practices, as well as distinct job skills and work attitudes. The integration of immigrants into our society represents a profitable investment, and the costs to integrate them will be more than compensated for in the future.

These narratives strike a neutral tone when it comes to evaluating immigrants vis-a-vis Germans, stating just that there are different attributes between migrants and Germans. The focus here is on whether the integration of migrants is valuable as an investment: in the anti-immigration narrative it is "too costly", and the costs "will never be compensated for", while in the pro-immigration narrative it is "profitable" and the costs "will be more than compensated for". Similar narratives were also used in the electoral campaign.[15]

4 Study 1: Measurement of private preferences

4.1 Design of the survey on private preferences

An online sample of 1,226 participants was recruited by the international polling company Kantar for our first survey. We conducted this survey between the 10th and 17th of September 2021 among German residents who declared that they were registered to vote. Our recruitment strategy included a screening stage in which participants were asked to express their voting intentions in the upcoming national elections (see question 6 from the survey in Appendix K for the exact question). Participants were then classified into four groups according to their political orientation: far-left, left-wing, right-wing, and far-right. We assigned participants to one of these four groups based on the classification offered by the Parliaments and Governments Database[16]. The far-left group was mostly made up of "The Left" ("Die Linke") prospective voters, the left-wing of the SPD or the Greens, the right-wing of the CDU/CSU or the liberals of the Free Democratic Party (FDP), and the far-right by AfD voters. At the time of the survey, opinion polls gave roughly the same votes to the CDU-CSU and SPD plus Greens, while the AfD and The Left were expected to receive about 11% and 6% of the votes, respectively[17]. Given our specific interest in polarization, we oversampled groups at the extremes of the political spectrum. Those unwilling to disclose their voting intentions were screened out of the survey. In total, we turned down 1,324 participants who were foreign citizens, did not expect to vote

[14] An example of the H/I narrative can be found in the AfD campaign (https://www.afd.de/staatsvolk/) and the A/I narrative can be found in the Greens'(https://www.cdu.de/thema/integration).

[15] An example of the H/R narrative can be found in the declarations by AfD members (https://afd-fraktion-bw.de/pressemitteilung/integration-ist-gescheitert/) and the A/R narrative in declarations by the SPD (https://www.spd.de/aktuelles/einwanderungsgesetz).

[16] This is available at http://www.parlgov.org/explore/deu/party/

[17] See https://www.politico.eu/europe-poll-of-polls/germany/, accessed on 7.9.21.

in the upcoming elections, or were unsure about their voting behaviour. The resulting attrition rate is 45%. In addition to party quotas, we also applied quotas to gender, aiming for a 50% split, and age, according to the following groups: [18-30]; [31-60]; [61-99]. Our age quotas corresponded to the real-life age distribution based on a poll by the Institut für neue soziale Antworten (INSA). The final sample included 162 far-left, 498 left-wing, 340 right-wing and 226 far-right supporters, which is in line with our pre-registration targets (see Table J1 in Appendix J for the other characteristics of the sample).

In the first part of the survey, we showed participants the six narratives in random order and asked them to rank the narratives from the one they agreed with the most to the one they agreed with the least (question 9). As a robustness check, in the ensuing screen, we also asked each narrative to be agreed on a Likert scale from 1 to 10 (question 10). Immediately after eliciting preferences, we elicited beliefs over the most preferred and second most preferred narrative by other participants (questions 11 and 12). The reward for a correct answer was 50 cents on top of the participation fee. The survey ended with a final set of demographic and socio-economic questions (see Appendix K for the full survey).

4.2 Results of the private preferences survey

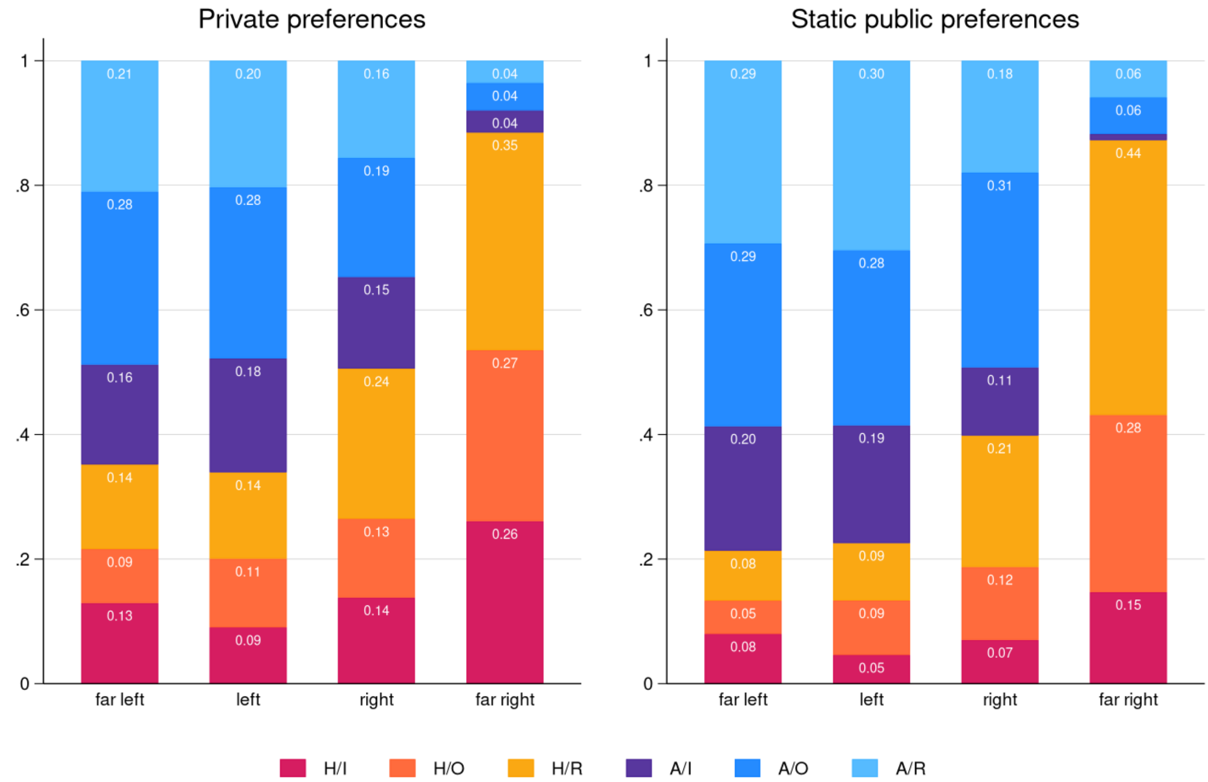
We use chi-squared tests to assess if participants preferred pro-immigration narratives over anti-immigration ones, and Mann-Whitney tests to compare private preferences across political groups. More than half of far-left and left-wing supporters ranked narratives accepting immigration first (64.8% and 66%, respectively). The null that the mean of either of the two distributions is equal to 0.5 is rejected at $p < 0.0001$, while the null of equality of the two distributions is not rejected ($p = 0.771$). Right-wing perspective voters were almost equally split between accepting and hostile narratives (50.6% vs. 49.4%, with the equal split $p = 0.828$; compared to the two left-wing parties, $p < 0.01$). 88% of far-right supporters preferred an anti-immigration narrative (the null hypotheses that the mean differs from the equal split and that the distributions are the same for each pair are rejected at $p < 0.0001$).[18]

Result 1: Far-left and left-wing voters preferred acceptance narratives over hostile ones, while the opposite holds for far-right voters. Right-wing voters were equally split between acceptance and hostile narratives.

Within the domains of acceptance and hostility of immigrants, participants clearly differentiated their level of agreement with the proposed narratives. We use Kolmogorov-Smirnov (KS) tests and regressions to evaluate these differences formally. Among hostile narratives, H/R was ranked ahead of both H/O and H/I at the strong level of significance (H/R Vs. H/O: $p = 0.0002$; H/R Vs. H/I: $p = 0.0001$; see Figure 1, left panel). Remarkably, H/R was the most preferred hostile narrative in all political groups. This was the case more strongly for right-wing voters (H/R Vs. H/O: $p = 0.001$; H/R Vs. H/I: $p = 0.003$), and, albeit in one comparison only, for left-wing voters, too (H/R Vs. H/O: $p = 0.243$; H/R Vs. H/I: $p = 0.031$). The test did not reach conventional significance levels for far-right participants (H/R Vs. H/O: $p = 0.153$; H/R Vs. H/I: $p = 0.105$) and was far from significant for far-left ones (H/R Vs. H/O: $p = 0.188$; H/R Vs. H/I: $p = 1.000$). As for acceptance narratives, A/O was ranked at the top significantly more often than both A/R ($p = 0.005$) and A/I ($p = 0.0001$). Once again, this was the case for each of the four groups. The result held more strongly for left-wing supporters (A/O Vs. A/I: $p = 0.003$; A/O Vs. A/R: $p = 0.020$), and partly so for far-left individuals (A/O Vs. A/I: $p = 0.032$; A/O Vs. A/R: $p = 0.260$), while they did not hold at conventional levels

[18] The results on beliefs from Survey 1 and 2 can be found in Appendixes E and I.

Figure 1: Private and static public preferences over narratives

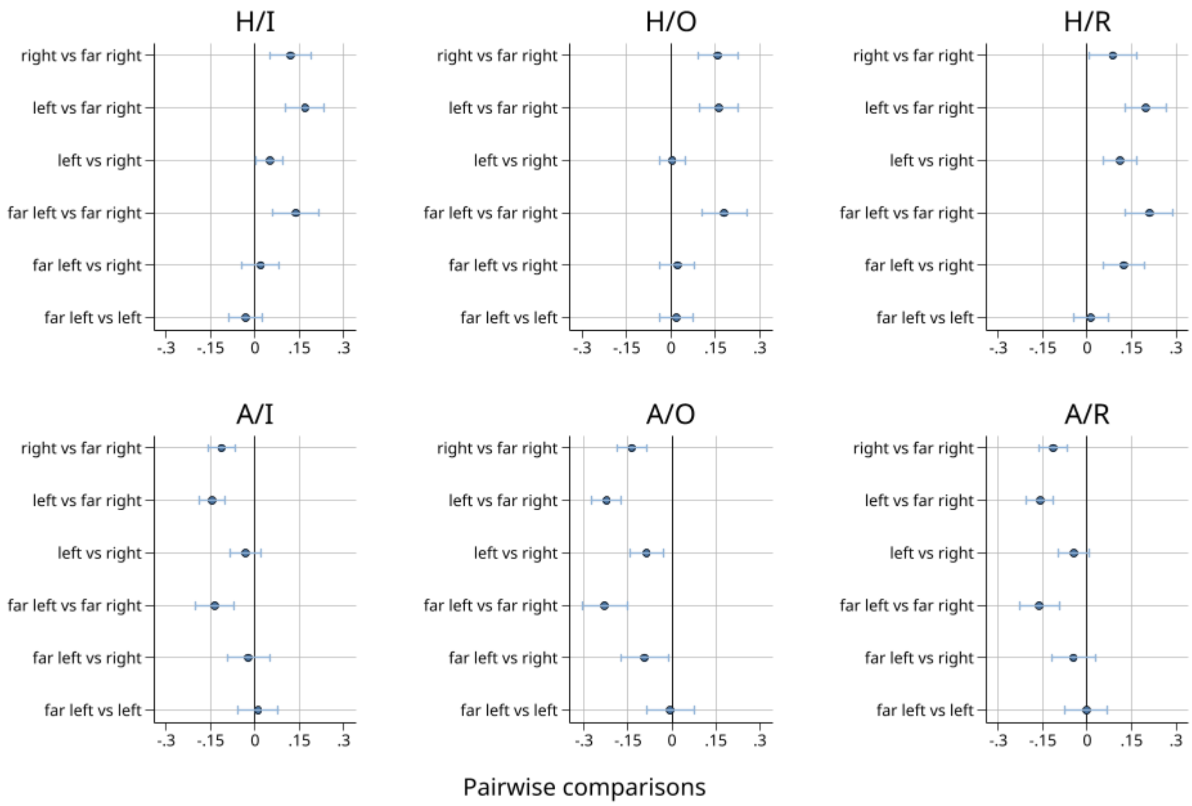


Notes: On the left, frequency of 1st ranked narratives by political orientation from Survey 1. The total sample is 1226 participants (162 far-left, 498 left-wing, 340 right-wing and 226 far-right). On the right, the frequency of endorsed narratives by political orientation from Survey 1. The total sample is 522 participants (75 far-left, 217 left-wing, 128 right-wing and 102 far-right).

for right-wing (A/O Vs. A/I: $p = 0.163$; A/O Vs. A/R: $p = 0.311$) and far-right supporters (A/O Vs. A/I: $p = 0.815$; A/O Vs. A/R: $p = 0.815$).

To assess whether this evidence was robust to controlling for individual characteristics, we fitted a multinomial logit regression on private preferences by party groups controlling for individual demographic and socio-economic characteristics. Figure 2 plots pairwise comparisons of relevant coefficient pairs. We find that left-wing voters do not show any difference in preferences from far-left ones. Right-wing voters prefer H/R more and A/O less than far-left and left-wing voters. In contrast, far-right voters' preferences significantly differ from those of any other group. Econometric analysis confirms that far-right voters prefer any anti-immigration narrative to any pro-immigration ones.[19]

Figure 2: Differences in private preferences over narratives between groups



Pairwise comparisons between coefficients from a multinomial logit regression on private preferences. The corresponding regression table can be found in Table B1 in Appendix B. Controls include age groups (young, middle age, old), female, immigrant status, education (4 groups), income (3 groups), employment status (3 groups), occupation (3 groups), religion (4 groups). Reported confidence intervals are at 95% level.

Result 2: Far-left and left-wing voters, when evaluating acceptance narratives, prefer a focus on the out-group to in-group/reciprocity ones. In partial opposition to H1, right-wing and far-right voters, while evaluating hostile narratives, prefer a focus on reciprocity to the out-group/in-group ones.

[19] In Appendix B, we explore heterogeneity in individual demographic characteristics. We find minor differences in the probability of supporting a specific narrative by characteristic and replicate the well-established result that younger, more educated, and more religious individuals tend to be more in favor of immigration (Scheve and Slaughter, 2001). Interestingly, the only individual characteristic that remains significant after controlling for political affiliation is being Protestant, which leads to a significantly higher probability of preferring Acceptance narratives (Appendix B, Table B1, column 8).

These results are robust to expunging from the analysis participants who gave inconsistent answers between the ranking and the Likert scale questions. 67% of participants were consistent in their preferred narratives across the two questions, while 51% of them were consistent in both the preferred and least preferred narrative. The latter percentage does not vary much by political orientation, ranging from 45% for right-wing to 56.6% for far-right voters. We still found the same qualitative results when restricting to this smaller sample. If anything, Result 1 is stronger, as 66.7% and 72.3% of consistent far-left and left-wing supporters were in acceptance of migration. Right-wing supporters were again equally split between acceptance and hostility (50.3% vs 49.7%) and an astonishing 96.4% of far-right supporters were hostile to migration. As shown in Figure J3, we also confirmed Result 2 with H/R and A/O being by far the preferred narratives in this sub-sample, too. We also found the same results by including in our sample only participants born in Germany, thus excluding migrants and those who refused to state their country of birth (see Figure J4). At the end of Survey 2, which was ran a couple of weeks later than the present study, we elicited again private preferences over the same narratives. Reassuringly, both Result 1 and 2 hold (see Table I1).

5 Study 2: Measurement of public preferences in a static setting

5.1 Design of the survey on public preferences

The second part of Survey 1 (see section 4.1) was devoted to studying how preferences evolve from a private to a public setting within a within-subject design. All participants in Survey 1 were given the opportunity to publicly endorse a narrative on the freely accessible website labeled "Who-supports-what". The text of the question was:

We would now like to give you the possibility to express your support for one of the statements you saw in the previous section on the publicly accessible webpage "https://Who-supports-what.com." Your support will be made public, along with the support of all other participants to this survey who decide to do so. The website "https://Who-supports-what.com" will be active on the 20th of September and will be deleted on the 25th of September. It will never be reactivated again. You can check the website to see which statement has received the most endorsements from the participants in this research.

We created and maintained this website for the purpose of this study. The week in which it was active coincided with the elections week. We showed a picture of the website (see Figure J2 in the Appendix) and informed participants that, in case they chose to go public, they would be asked an alias name at the end of the survey under which their support would be shown.[20] Romano et al. (2021) used a similar design to make participants' experimental choices public. Public endorsement was voluntary and participants could pick up a narrative for public endorsement different from the one at the top of their private ranking. We first asked participants to choose the narrative to be publicly endorsed (question 13 in Appendix K) and then their willingness to endorse it publicly (question 14).

[20] We would have preferred to leave participants free to give their real name or an alias name on the website. However, the polling company refused the option of participants giving their real names on grounds of protecting participants' privacy.

5.2 Results of the public preferences survey

5.2.1 Comparing private and static public preferences for narratives

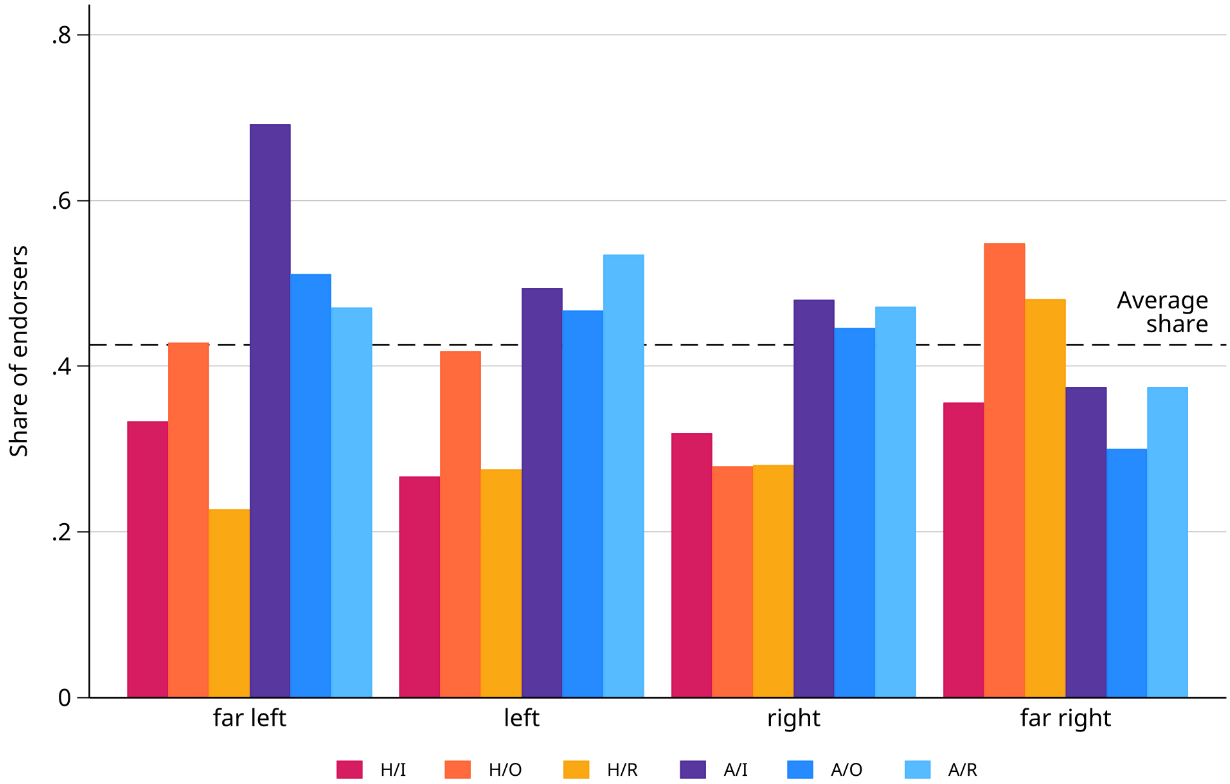
First, we compare preferences by the whole sample of participants when involved in the private survey with those expressed by participants who opted to endorse a narrative publicly (see panels in Figure 1 for the distribution of such preferences). In this way, we can compare how preferences "evolve" from a private setting (when the choice is only observed by the researcher) to the public setting (when the choice is observed both by the researcher and internet users accessing the website). We found significant changes between these two settings with KS tests. The null of equality of distributions between the private and public setting is strongly rejected in the overall sample ($p = 0.001$) and is also rejected for the non-extreme political groups ($p = 0.041$ for the left-wing, $p = 0.040$ for the right-wing, $p = 0.279$ for the far-left, $p = 0.321$ for the far-right).

We then used Mann-Whitney ranksum tests between private and public preferences on acceptance/hostile narratives to assess specific differences. Overall, we find a tendency to observe more support for pro-immigration narratives in the public vis-a-vis the private setting among all groups ($p = 0.0003$). We observe a 13.9%, 11.3%, and 10.8% increase in support for pro-immigration policies over the total among far-left, left-wing, and right-wing perspective voters ($p = 0.032$ for far-left, $p = 0.002$ for left-wing, $p = 0.038$ for right-wing supporters). Support for anti-immigration narratives by far-right supporters remained nearly unchanged - from 88.5% in the private setting to 87.3% in the public setting ($p = 0.748$)-, and remained the highest between political groups. In particular, differences get close to significance for far-left supporters for some narratives ($p = 0.264$ for H/I, $p = 0.217$ for H/R, and $p = 0.160$ for A/R). We also observed significantly higher support for A/R (30.41% vs 20.28%, $p = 0.003$) and significantly lower support for both H/I (4.61% vs 9.22%, $p = 0.041$) and H/R (9.22% vs 13.86%, $p = 0.084$) in the public vis-a-vis the private setting for left-wing prospective voters. We observed significantly higher support for A/O in the public (31.3% of total support) than the private (19.1% of total support, $p = 0.005$) setting for right-wingers. This went mainly at the expense of H/I ($p = 0.044$), whose support dropped from 13.8% in the private setting to 7% in the public setting. Far-right participants showed less support for H/I in public (14.71% of total support) than in private (26.11% of support, $p = 0.022$), and this seems to be accounted for by an increase in H/R (34.96% vs 44.12%, $p = 0.114$).

Result 3: Consistently with H2, we observe significant changes in the distribution of preferences manifested in the public and the private settings. Pro-immigration narratives are more frequently supported in the public by far-left, left-wing and right-wing voters, while far-right voters maintain their support for hostile narrative while shifting away from the hostile narrative with a focus on the in-group. We also observe a shift towards out-group narratives for the right-wing voters while left-wing ones move more towards A/R.

As shown in Figure J5, these results are robust to expunging from the analysis observations by participants who were inconsistent in their stated preferences, both with respect to the tendency to observe more support for pro-immigration policies in the public than the private setting and with respect to the specific narratives being supported. The same holds if we include in our sample only participants born in Germany, as shown in Figure J6 in Appendix J.

Figure 3: Share of endorsers by political party and by narrative



The sample consists of 1226 participants. The bars represent the share of participants by party - conditional on their private preferences - who are willing to publicly endorse a narrative. The dashed line represents the average share across all the political parties.

5.2.2 The choice of endorsing a narrative

The result of greater overall support for pro-immigration narratives in the static public vs. the private setting could be due to changes at the extensive margin (people with anti-immigration preferences shy away from the public setting), or to changes in the intensive margin (people tend to switch their preferences from anti-immigration to pro-immigration when in the public vis-a-vis the private setting), or to both. In this section, we show that Result 3 is mainly driven by the first of these mechanisms.

522 participants out of 1,226, namely, slightly above 40% of our sample, decided to go public with a narrative. Hence, although anonymous, the individuals seemed to think of this choice as meaningful. The willingness to go public was 46.3% for far-left, 43.5% for left-wing, 37.7% for right-wing and 45.1% for far-right supporters. Overall, prospective voters of extreme parties were more inclined to publicly endorse, but by Mann-Whitney tests, differences are statistically significant only for right-wing voters and only at weak levels of statistical significance ($p = 0.076$ with respect to far-right, $p = 0.065$ with respect to far-left supporters). Figure 3 shows the share of endorsers for each initial private preference and for each party. Anti-immigration far-left, left-wing and right-wing participants shied away from going public. Using KS tests, the distributions of private preferences between endorsers and non-endorsers were significantly different for each party group with the only exception of the far-right (far-left: $p = 0.009$, left-wing: $p < 0.0001$, right-wing: $p = 0.0001$, far-right: $p = 0.148$). Evidence in this direction also came from OLS regressions in Table 1 where we looked at the probability of becoming endorsers by initial private preferences and by individual characteristics. The probability of becoming an endorser significantly increases by 12.3% if the private preference is in favour of immigration. When we control for the supported party in column (2), the same result still holds. When we add in column (3) the interactions between the private preference and the party group, we find that the interaction between far-right voters and acceptance narratives is negative; furthermore, by a post-estimation Wald test, we cannot reject the null that the sum of the interaction with the baseline coefficients is equal to zero, $p = 0.754$. In Appendix B, we further discuss the role of individual characteristics on the probability of becoming an endorser, showing that men, people older than 65, and people on high incomes are significantly more likely to endorse narratives than others.

Result 4: Far-left, left-wing and right-wing voters publicly endorse a narrative less if their preferred private narrative is hostile to migration. Far-right voters' endorsement decision does not depend on their private preferences.

5.2.3 Changes in preferences of the endorsers

Did the endorsers change their preferences when going public? In Table C1 in Appendix C, we show the full matrices of changes by party. The number of those who changed their preference was higher than that of those who kept the same preference in the aggregate and for the main political groups (311 vs 211, from chi-squared tests, $p < 0.001$ in the aggregate, $p = 0.203$ for far-left, $p = 0.025$ for left-wing, $p < 0.001$ for right-wing and $p = 0.843$ for far-right voters). Importantly, changes were mostly limited to public preferences in support of the same policy position. According to a Fisher chi-test, we fail to reject the null that distributions of acceptance/hostility narratives are different between public and private preferences for the endorsers ($p = 0.389$). Moreover, differences in the distributions between private and public preferences tested with KS tests reach statistical significance neither overall nor for each party (overall: $p = 0.396$, far-left: $p = 0.970$, left-wing: $p = 0.894$, right-wing: $p = 0.910$, far-right: $p = 0.593$), suggesting that these changes balance out in the aggregate. In Appendix C, we

Table 1: **The probability of endorsing a narrative by private preferences and party groups**

	Endorser		
	(1)	(2)	(3)
Acceptance private preference	0.123***	0.127***	0.204**
	(0.028)	(0.035)	(0.082)
Party group			
left		-0.025	0.012
		(0.045)	(0.075)
right		-0.062	-0.022
		(0.048)	(0.075)
far right		0.078	0.152**
		(0.054)	(0.074)
Acceptance pref. # Party group			
Yes # left			-0.060
			(0.093)
Yes # right			-0.060
			(0.097)
Yes # far right			-0.320**
			(0.130)
Acceptance narrative in public			
Yes		0.040	0.036
		(0.035)	(0.035)
R-squared	0.044	0.052	0.058
Number of observations	1226	1226	1226

*** p<.01, ** p<.05, * p<.1

OLS regressions on the choice of becoming an endorser. The main variable of interest is a dummy variable taking a value of 1 if the participants has a private preference in acceptance of migration (A/I, A/O or A/R). Controls (not reported) include age group (young, middle age, old), female, immigrant status, education (4 groups), income (3 groups), employment status (3 groups), occupation (3 groups), religion (4 groups).

provide further statistical evidence in this direction from a multinomial logit.

Result 5: Endorsers' public preferences were, on aggregate, the same as their private ones. Moreover, at the individual level, they were more likely to be preferences expressing the same policy position.

5.2.4 Interpreting the change in preferences

As laid out in section 2, we postulated that social image concerns and the desire to publicly endorse the most distinctive narrative for their own political group were additional motivating factors to endorsing the privately preferred narrative. The key idea of a distinctive narrative is that it optimizes the desire to express support for one's own political group most preferred narrative and the desire to differentiate from the most preferred narratives by other groups, as modelled in equations (2) and section (3) in section 2.

Since eliciting all first- and second-order beliefs to accurately measure distinctiveness according to expression (3) would have been unfeasible in Survey 1,[21] we constructed such measures from Survey 2 (see Section 7.4.1 and questions 9 and 13 in Appendix L). As laid out in our pre-registration, we computed the most distinctive narrative for each party in Survey 2 and assumed that all individuals

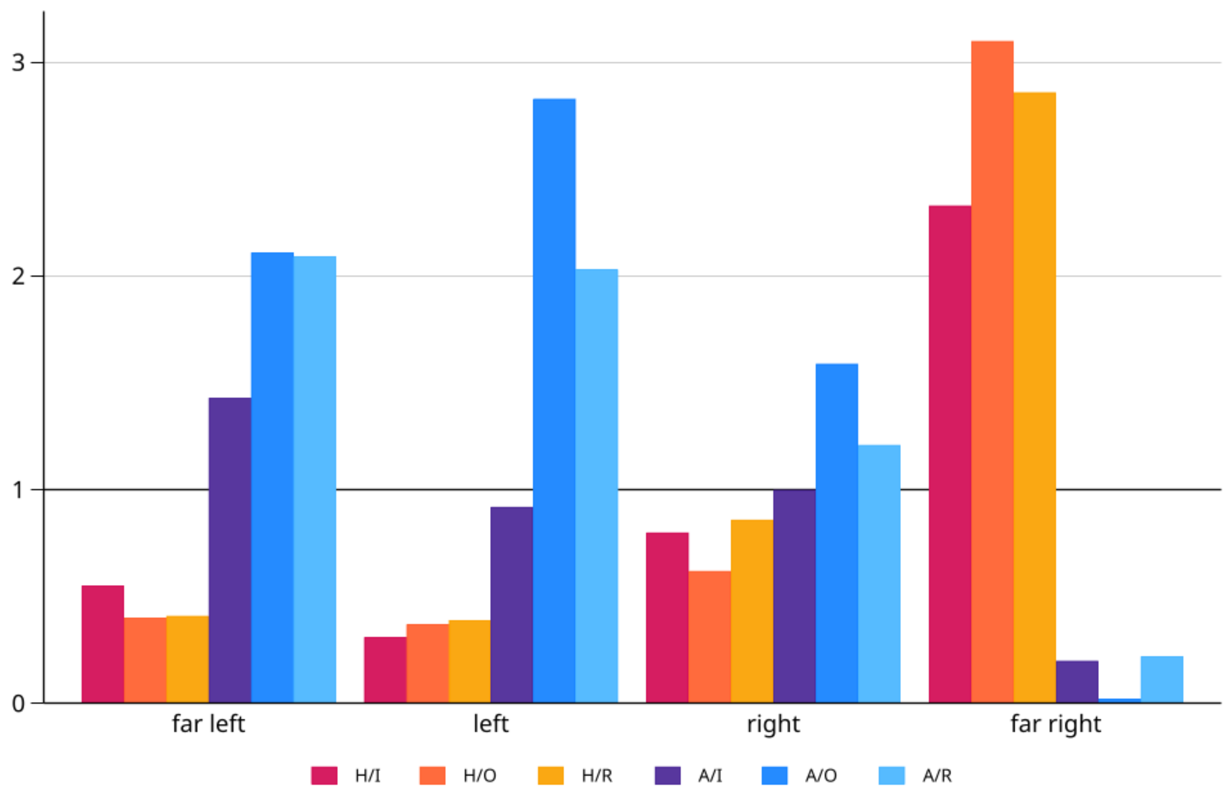
[21] At the theoretical level, the elicitation of this measure of Survey 1 would have had to come after the public endorsement decision. We then thought that beliefs over the popularity of the narrative for one's own and other political groups would have been anchored to the previous decision, thus offering an unreliable and "noisy" indicator of our targeted construct. Logistically, this elicitation would have come toward the end of the survey, with arguably low attention rates. For all these reasons, we preferred to elicit this measure in Survey 2 and proceed as described in the pre-registration plan.

in Survey 1 belonging to the same political group held the same distinctive narrative as determined in Survey 2 for the same political group. That is, we proxied (3) with the following measure:

$$Dist_{Iv} = \frac{E(Sh_{\bar{I}v})}{E(Sh_{-\bar{I}v})} \tag{9}$$

where $E(Sh_{\bar{I}v})$ is the percentage of voters in one's own party who declared that the preferred narrative $\bar{I}_i$ is v and $E(Sh_{-\bar{I}v})$ is the percentage of own party voters who declared the preferred narrative in the other parties $\bar{I_{-i}}$ is v.

Figure 4: Distinctivness ranking by party



Ratio of how many supporters of a party believed the narrative was preferred by their own party supporters over the supporters of other parties from Survey 2. The total sample over which these calculations are made is 771 participants.

In Appendix I, we show that what can be construed as false consensus (Ross, Greene, and House, 1977) is sizeable, as 54% of participants predicted that fellow party members follow the same narrative as themselves, but were wrong in 59% of such cases. Interestingly, right-wingers are those most likely to be wrong when predicting the narrative preferred by others, both in their own party and in other parties. Conversely, far-right supporters are those least imprecise in their predictions.

Figure 4 displays the degree of distinctiveness for each narrative by political group. It emerges that the most distinctive narrative is A/O for all political groups except for the far-right, whose supporters instead have H/O as their distinctive narrative. We then use Mann-Withney tests to assess if the static public preferences of the endorsers have higher distinctiveness than private preferences in the overall sample. In other words, we associate to each preference (either private or public) the level of distinctivness such as measured in Survey 2. We then ran a test on the null hypothesis that the distribution of choices in terms of distinctiveness was the same in the private and public static context. We find that the null hypothesis is rejected ($p = 0.0892$), as participants choose their own party's most distinctive narrative $\hat{\nu}_I$ more frequently in public than in private context. This result is mostly driven by right-wing voters ($p = 0.0051$).

In Appendix D, we show that the distinctive narrative determined for each group is highly significant in predicting which narrative has been endorsed in the public static decision. We also test for alternative explanations for the change between private and public preferences. We test for the predictive power of the two components of the distinctive narrative, that is, the most popular narrative in one's own political group and in the other political groups. Finally, we also deploy answers to a question asking to state which narrative was the "most politically correct" (see question 8 in the Survey 2 questionnaire). Although all of these measures have significant predictive power, the coefficient for distinctiveness in predicting the endorsed narrative is the highest.

In principle, pluralistic ignorance may also be a motivating factor for the change of preferences between the private and static public settings. Explicit testing for pluralistic ignorance was beyond the scope of this study. Nevertheless, an essential component of pluralistic ignorance is the belief that one's own preferred opinion is not upheld by the majority of people in a society. In our surveys, however, the narrative that most survey participants expected to be the most preferred was the anti-immigration narrative (see Appendix I: Table I2). This was the case for all groups except for the far-right. Since all groups, except for the far-right, turned away from the anti-immigration narratives in the static public setting compared to the private setting, this pattern of behavior appears to be incompatible with pluralistic ignorance. The one group for which pluralistic ignorance may have hindered their public endorsement of anti-immigration narratives was the far-right. However, their behavior did not change significantly in the public vs. the private setting. Therefore, pluralistic ignorance was definitely not a motivating factor for them.

In sum, our available evidence points to relevance for social image concerns, in line with our hypotheses. Although other related mechanisms may be at play, distinctiveness seems to have significant power in predicting the publicly endorsed narrative. On the other hand, the pattern of beliefs over the most preferred narratives by others seems to be incompatible with pluralistic ignorance.

5.2.5 Private and static public preferences over vaccination narratives

In order to appreciate the extent to which the patterns and polarization of private and static public preferences in Study 1 and 2 were specific to immigration we ran a "false experiment" in which we replicated the above design with respect to narratives related to the COVID-19 pandemic.

After the sections described above, Survey 1 had an additional section on COVID-19 vaccination. The structure of this additional section was the same as the one used for migration narratives. That is, we constructed six narratives in which unvaccinated people replaced immigrants as the main subject of the narratives. We manipulated such narratives on both the hostility/acceptance dimension and on the in-group/out-group/reciprocity focus (see Appendix F for the text of the narratives).

The vast majority of respondents was in favor of vaccination. This was definitely the case for far-left, left-wing and right-wing voters, of whom 79.01%, 80.12% and 77.07% were accepting of supporters of vaccination. On the other hand, far-right individuals were evenly split, with 49.56% in favor of vaccination (see Figure 5). A chi-squared test rejects the null of half split at $p < 0.0001$ for all parties except for the far right ($p = 0.894$). We use KS tests to evaluate differences across narratives. We find no differences across hostile ones ($p > 0.25$ for all narratives in each party group except for $p = 0.0505$

for H/O > H/R for far-right supporters). Among the acceptance ones, we found a marked prevalence of agreement with A/I ($p < 0.0001$ for left-wing and right-wing and $p = 0.0116$ for far-left with A/R, $p = 0.0024$ for left-wing, $p = 0.0368$ for right-wing and $p = 0.3634$ for far-left with A/O), while A/O was generally preferred to A/R (far-left: $p = 0.1302$, left-wing: $p = 0.0004$, right-wing: $p = 0.0072$). Only among far-righters the differences across acceptance narratives were not significant ($p > 0.20$). This evidence shows a substantial change from immigration to vaccination: migration is revealed to be in our survey a more controversial topic than vaccination, with more people against it across all party groups. Moreover, the A/I narrative now became the most preferred one, which was not the case on migration where A/O and H/R were more popular.

We also consider here if respondents' behavior in a public setting with opt-out possibility and no interactions was the same on vaccination as on migration. 500 subjects over 1226 decided to go public on vaccination and these form our sample on public preferences, and 451 of these subjects were the same that went public on migration (86.40% of those who went public on immigration made the same choice on vaccination). Contrary to the immigration narratives, a KS test does not reject the null of equality of distribution between the private and public setting neither in the overall sample ($p = 0.586$), nor for each party group (far-left: $p = 0.938$, left-wing: $p = 0.832$, right-wing: $p = 0.967$, far-right: $p = 0.768$). We observe, nevertheless, a moderate increase in support for pro-vaccination narratives, from 73.49% to 77.60%, however the difference by Mann-Whitney tests is only weakly significant in the overall sample ($p = 0.0750$) and not significant for any party (far-left: $p = 0.4582$, left-wing: $p = 0.2703$, right-wing: $p = 0.2296$, far-right: $p = 0.6079$). When we looked at the specific focus of the narratives, no change in preferences came out as significant at conventional levels.

6 Study 3: Preferences over narratives endorsed in social media

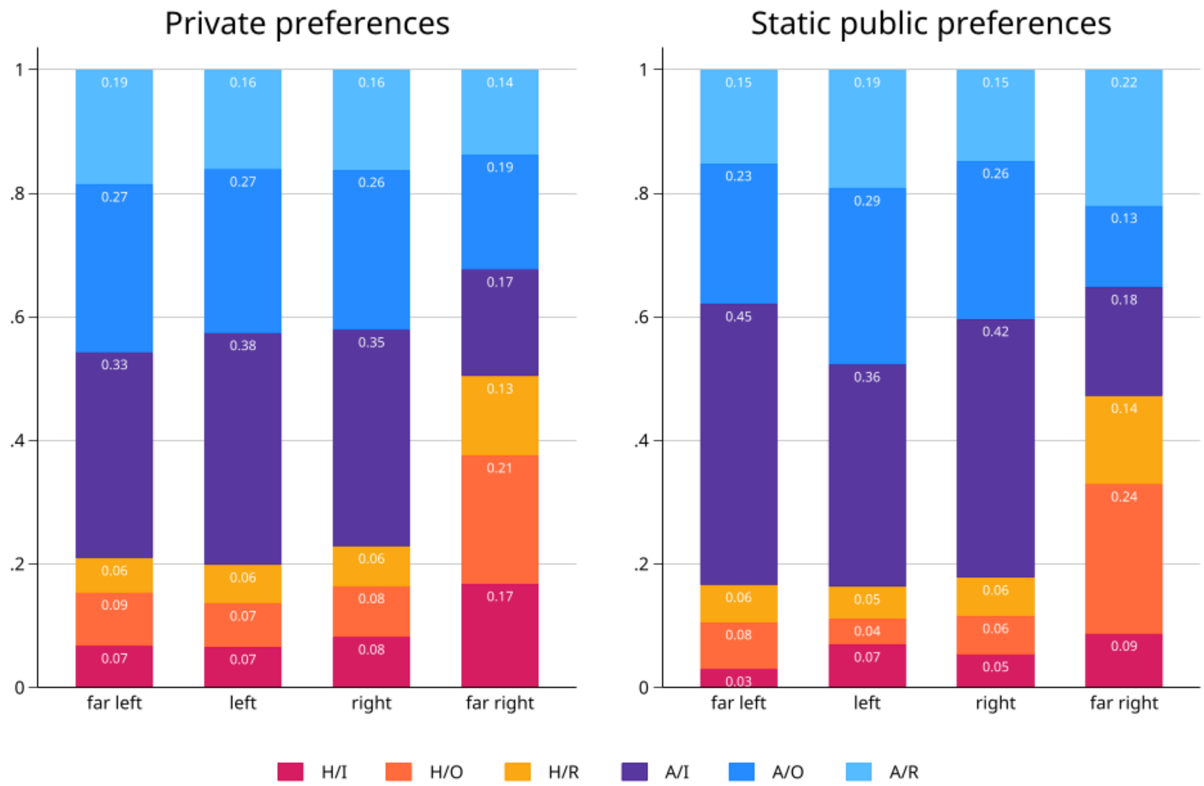
6.1 Design of the field experiment

Preferences studied in Study 2 (see section 5) were public but without social interaction. That is why we call them "static" public preferences. Moreover, participants were aware that their choice would have been observed by the experimenter. In our social media experiment, both features were modified. We used the four narratives with the highest number of endorsements in Survey 1 (see section 5.2.2), thus omitting the two in-group narratives in the hostility and acceptance domains. We excluded two narratives because we conjectured that, with six narratives, the available number of observations in this experiment would not have given us enough power to avoid Type-1 errors. We implemented our social media experiment on Twitter (currently X), the most popular social media platform dealing with political issues at the time of the study[22]. The narratives used on Twitter had to be slightly shortened from the version used in the survey (see section 4.1) due to the characters limits on Twitter (see Appendix G for the revised narratives). Hypotheses were pre-registered at OSF (https://doi.org/10.17605/OSF.IO/2T46H), and we got ethical approval from the Central Ethics Committee of the University of Kiel under the conditions listed below.

Our experimental manipulation consisted of assigning one narrative to the "pinned" tweet of four fictitious Twitter profiles that we set up. Pinned tweets are tweets a user can decide to fix at the top of the timeline on their personal Twitter page. A pinned tweet is then clearly visible whenever another

[22]See https://www.pewresearch.org/short-reads/2014/11/12/facebook-and-twitter-as-political-forums-two-different-dynamics/ by the Pew Research Centre. As Twitter became X, several changes have been introduced over time. For instance, likes are now hidden from the timelines, and the tweets have been renamed "posts".

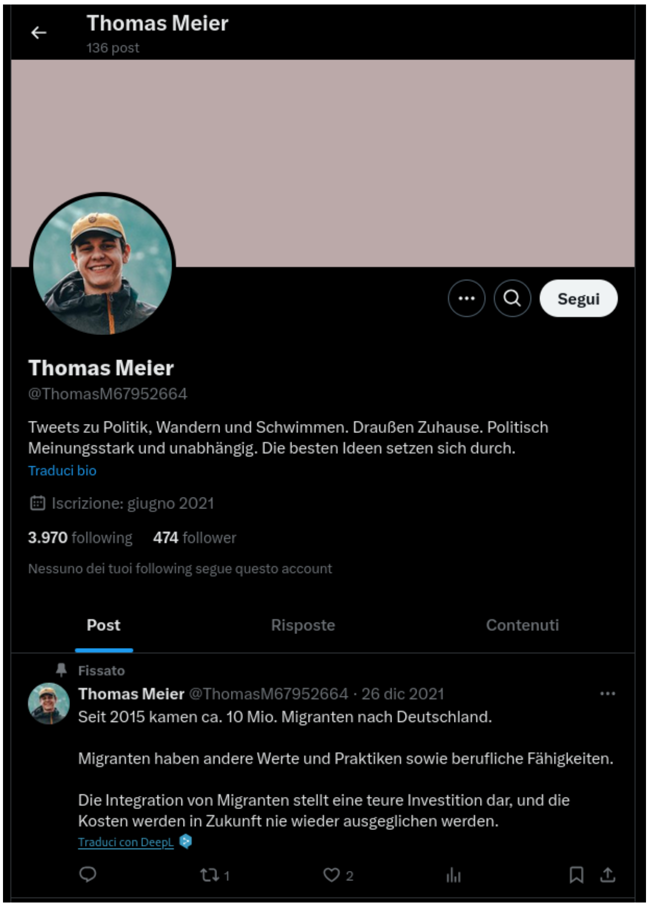
Figure 5: Private and static public preferences over vaccination narratives



On the left, frequences of 1st ranked narratives by political orientation from Survey 1. The total sample is 1226 participants (162 far-left, 498 left-wing, 340 right-wing and 226 far-right). On the right, the frequency of endorsed narratives by political orientation from Survey 1. The total sample is 500 participants (66 far-left, 214 left-wing, 129 right-wing and 91 far-right).

user lands on the profile page. We administered our 'stimulus' by following a randomly selected number of Twitter users from one and only one of our four profiles. In other words, our design can be seen as following a between-subjects approach. Based on their previous activities on Twitter, which used to be publicly accessible on Twitter through APIs, we classified Twitter users into the four groups of political orientation used in the online survey, whenever their political orientation was clear. We randomly assigned each user from this pool of *targets* to one of our four profiles. The order in which they would be targeted over the course of the experiment was also randomized. More precisely, our methods — piloted twice a few months before the actual experiment was fielded — were as follows:

Figure 6: Thomas Meier



Screenshot of one of the Thomas Meier profiles we created on Twitter in June 2021. All profiles looked the same except for the pinned tweet that stayed fixed at the top of the timeline. An example of a pinned tweet is at the bottom of the screenshot.

1. In June 2021, we opened four Twitter accounts, all under the name of 'Thomas Meier'. The accounts were initially identical (see Figure 6 for one example). In particular, the photo profile and the description were the same for all four accounts.[23] We strove to represent an "average" young German male with a common name and surname. The account's activity up to the experiment's starting date was limited to posting politically neutral daily news from mainstream German online outlets. The experiment started on September 22nd. From that date, each profile started having a different "pinned" tweet on immigration corresponding to one of our four narratives.

[23]The profile picture was taken from the website Unsplash, whose photos can be used freely for all purposes with no need for permission request.

2. We selected four different pools of German-speaking Twitter users. Each pool was composed of users aligned to a specific political orientation (right-wing: CDU/CSU + FDP, left-wing: SPD+Greens, far-right: AfD, far-left: Die Linke). We selected users into the pools based on their past Twitter activity: if they retweeted in the recent past—starting from January 2021—a tweet by one of the parties listed above, then they were categorized as being supporters of the corresponding party and selected into the pool. Drawing on an algorithm, we then screened out accounts that we suspected were either bots, journalists or institutions[24]. We were then left with 19,989 Twitter users, which we randomly assigned to four different target experimental groups. We used a blocked randomization process with the weight of each party within the block determined by data from the INSA poll from the 17th of September.

3. Each experimental Twitter profile started "following" the assigned target users on the 22nd of September and finished the pool on the 96th day of the experiment[25]. Given the technical constraints imposed by Twitter, we were able to follow approximately 200 accounts per day per single experimental profile. We had to end our experiment when our profiles followed more than 5,000 users because of Twitter's limitations on the followers/following ratio.

4. As requested by the Kiel University Ethics Board, at the end of the experiment we replaced the pinned tweet with the following disclaimer in German (here translated) in the four accounts: *"This profile was created as part of a research study on political opinions. It will be removed in a few days. Please contact this email address with your Twitter ID if you would like to receive further information: ThomasMeierResearch@mail.com."*. The disclaimer was shown for one week, after which the four accounts were deleted. No user contacted the above email address asking for information or placed comments afterwards. Three users placed a like or retweeted the disclaimer.

As stated in the pre-registration, our key variable of interest is the follow-back rate, the same variable used in Mosleh et al. (2021). In some cases, a user followed back at first but then "unfollowed" before the end of the experiment. As a robustness check, we consider both variables attributing "success" to back-followers who were still followers at the end of the experiment and the broader set of back-followers who unfollowed before the end of the experiment.

We also report results concerning the number of engagements with the pinned tweets (i.e., the number of visualizations, comments, likes, and retweets)[26], as they clearly reveal which narrative spread the most over the platform. Since we expected a much lower level of engagement for our pinned tweets than what turned out to be the case, we did not include engagement variables in our pre-registration. The pinned tweets were removed and reinstated almost on a daily basis to provide enough independent observations.

It is apparent that such engagement variables involve different levels of interest, support and publicity of support. The follow-back rate is an expression of interest and is arguably the mildest form of

[24] More specifically, we dropped users who retweeted tweets of more than one party in their Twitter timeline, those who signed up after 2019 (because of their higher probability of being bots created to influence voting), national or local branches of the parties, and users who had less than 100 tweets in their timeline (again because of the higher probability that they were bots). See Appendix G for more details on this procedure.

[25] Our starting date was before the election date of the 26th of September because we wanted to ensure that users did not behave differently before and after the elections. Although the sample pool before the election was too small to make any statistical inference, the results were qualitatively the same.

[26] Visualizations are the number of users who see a tweet. Comments are posts that are placed in the thread following a post. We did not evaluate whether comments were positive or negative. Likes are expressions of interest or support by a user on a post. Retweets are relaunch of a post published by another user on one's own profile page.

support, as it does not appear on any user's timeline. A user may verify whether a user is following another user by searching through their list of followed accounts. However, this verification may be laborious when a user follows many other users. Moreover, it is rather common that a user follows other adversarial users, possibly as a way to publicly counter their claims.[27] Likes and retweets are, unambiguously, stronger manifestations of support than follow-backs. Both the number of likes and retweets are visible under the tweet and users can easily verify who liked and retweeted. Arguably, retweets are an even stronger manifestation of support than likes, because retweets appear in the retweeter's own profile page.

Visualizations cannot be seen *prima facie* as expressions of support, because they are determined by the social media matching algorithm. However, they may be seen as a comprehensive measure of how large was the diffusion of a certain tweet. Generally speaking, social media algorithms select posts from other users that are deemed as interesting to a certain user (on the basis of the profiling of a user carried out by the algorithm) and make them visible on the user's homepage. Algorithms work using quasi-random selection mechanisms, where posts that start receiving a large number of likes and retweets are more likely to appear on the homepage of other users. Therefore, by liking and retweeting a certain post, a user is both expressing their support for this post and making it more likely that it will appear on some other user's homepage (see e.g. Narayanan 2023; Pariser 2011).

6.2 Results of the field experiment

6.2.1 The follow-back rate

A total of 942 users followed back our profiles. The follow-back rate was 2.97% for far-left users, 2.68% for left-wing ones, 3.07% for right-wing ones and 10.3% for far-right supporters. It is noticeable that far-right supporters followed back more than the supporters of other parties by a factor of roughly 3 to 1 (by Whitney-Mann tests, $p < 0.0001$ for all pairwise comparisons between far-right supporters and each of the other groups; pairwise comparisons between the other parties have $p > 0.20$)[28].

The proportion of follow-backs across the four narratives is similar to both private preference rankings and public-endorsed narratives in Survey 1 (see sections 4.2 and 5; see Figure 7 for stacked bars showing relative preferences in the three studies). We use KS tests to evaluate differences across distributions. Private preferences and follow-backs are not significantly different ($p = 0.672$). This is the case for all political areas ($p = 0.960$ for the far-left, $p = 0.976$ for the left-wing, $p = 0.643$ for the right-wing) except for far-right users ($p < 0.0001$). Likewise, there are no significant differences between follow-backs and endorsed narratives in Study 2. However, it is noticeable that, in this case, the p-value is considerably closer to the significance region in the aggregate ($p = 0.162$), as well as for individual parties (far-left: $p = 0.112$; left-wing: $p = 0.085$; right-wing: $p = 0.432$), while the null is strongly rejected for far-right users ($p < 0.0001$). Conversely and consistently with Result 3, when comparing private preferences and endorsed narratives from Study 2 limitedly to the four most popular narratives used in Study 3, the null is now rejected ($p = 0.041$). Hence, eyeballing p-values from such KS tests, it seems as if the distribution of preferences in the Twitter experiment lies between the private and the endorsed preferences distribution, being "closer" to the former than the latter.

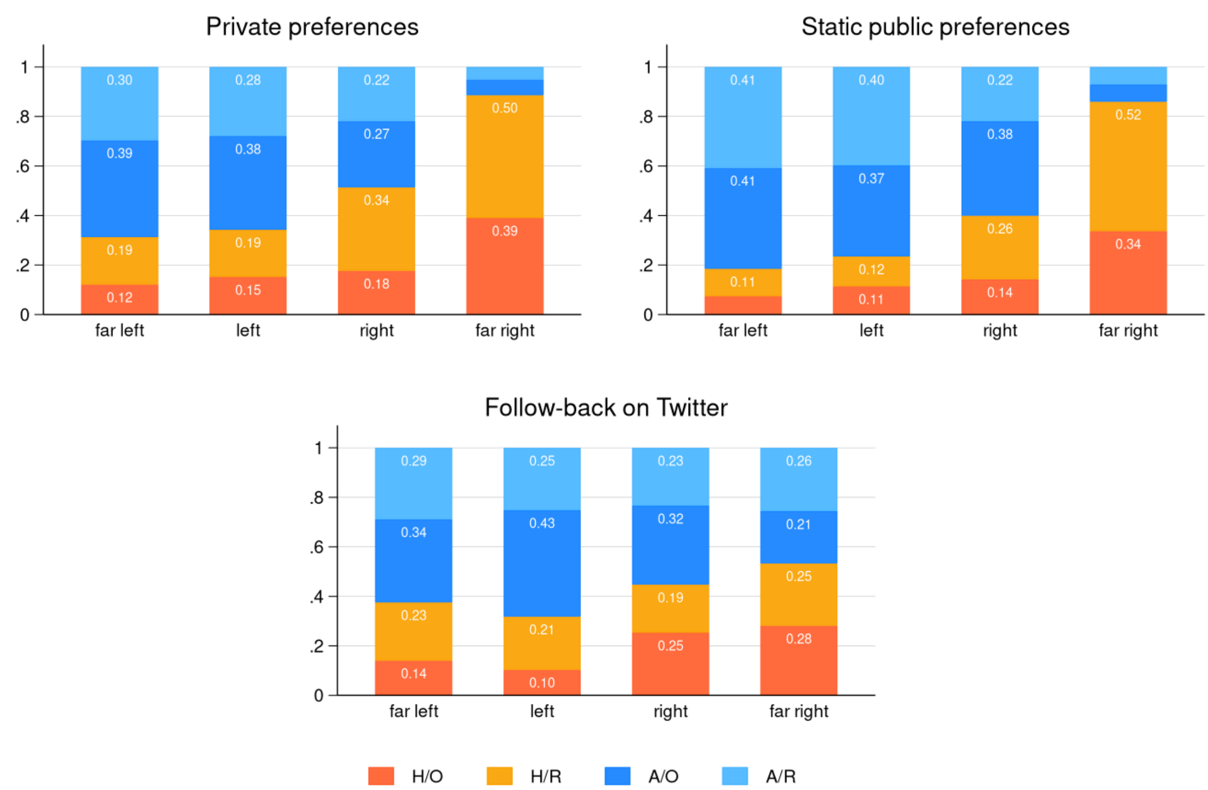
As in the survey, left-wing and far-left users followed back more the two pro-immigration profiles

[27] For instance, the pinned tweet of the account of popular journalist Ian Bremmer reads "If you're not following some people you dislike, you're doing it wrong."

[28] The total number we followed is 5009 for far-left, 5040 for left-wing, 5023 for right-wing and 4917 for far-right supporters.

than the anti-immigration profiles (68.2% for the left, 62.4% for the far-left), while right-wing users followed back acceptance and hostility profiles almost equally (55.2%). Again, for these three groups we do not detect appreciable differences in distribution compared to private preferences in the survey (p = 1.000 for far-left supporters, N=149 on Twitter and N=162 in the survey; p = 0.960 for left-wing ones, N=135 on Twitter and N=498 in the survey; p = 0.801 for right-wing ones, N=154 on Twitter and N=340 in the survey)[29]. Conversely, far-right users displayed a much more balanced preference over pro- and anti-immigration narrative than in the survey. The ratio between follow-back of anti-immigration and pro-immigration was 53.4% on social media, while it was 88% in the survey experiment, the difference in distributions with the survey being strongly statistically significant ($p < 0.0001$, N=504 on Twitter and N=226 in the survey). If we took this result at face value, it would seem as if far-right users became more pro-immigration on Twitter. More likely, these users followed back with a considerably higher probability regardless of the content of the pinned tweet, maybe following some strategic behavior. We will further explore possible mechanisms behind this behavior in section 7.1.

Figure 7: Comparison of preferences across settings



We report relative frequencies by political orientation for the 1st ranked narrative in Survey 1 (panel on the top left), for the endorsed narrative in Survey 1 (panel on the top right) and for the followed-back narratives in the Twitter experiment -relative to all narratives that were followed-back (panel at the bottom). For private and static public preferences, we computed relative frequencies using only the four narratives used in the Twitter experiment.

Result 6: The distribution of follow-backs by narrative and political orientation on Twitter is

[29]From Figure 7, we can also appreciate some other minor differences associated with the focus of the narratives. More specifically, it looks that for far-left and left-wing individuals H/R is now more strongly preferred to H/O and not that less popular than A/R.

not different from the distribution of private preferences and public static preferences from the survey, except for far-right supporters who follow-back more also our pro-immigration artificial profiles. In general, far-right supporters tend to follow-back more than the other parties regardless of the narrative.

For robustness, in Figure J7 we considered the follow-back rate measured at any point in time regardless of the date of the end of the experiment. That is, we also included as follow-backs users who later unfollowed our profiles. The follow-back rate so calculated more imprecisely reflects preferences because it conflates true preferences with users who followed-back out of confusion or of an initial instinct of pure reciprocation. It is reassuring that results still go in the same direction, even if they are slightly noisier. In Figure J8 and J9 we considered other related outcomes, such as unfollowing our profile after having followed it back at first, and blocking it after our profile has followed the user. Results are qualitatively similar to those on follow-backs, however numbers are smaller and the evidence is then noisier.

6.2.2 Engagement over the pinned tweets and diffusion on Twitter

In the case of engagements with the pinned tweets, we were unable to identify the political orientation of all Twitter users who engaged with the tweets in question. It is possible that users who would not be classifiable into the four political groups could engage with the tweets[30]. The analysis was then conducted at the aggregate level. Pinned tweets were removed and reposted by us approximately once every two days. To be precise, we had 52 iterations of pinned tweets with H/O, 48 with H/R, 50 with A/O, 50 with A/R, which constitute independent observations. The clear result from this analysis is that the H/O narrative triggered much more engagement than any other narratives and became much more diffused according to all of the four measures we use (see Figure 8). On average, the H/O pinned tweet triggered 16 times as many likes than the other narratives. Results are even clearer for retweets, where H/O got retweeted on average 45 times more than the other narratives. H/O got also more visualizations and replies by an average factor of 5 to 1. The differences involving H/O were statistically significant, with Mann-Withney ranksum tests, in almost every comparison (see Table 2). When it came to the comparisons involving H/O, Cohen's d effect sizes were between 0.38 and 0.47, suggesting a medium size effect (Cohen, 1988). We also found some evidence that A/O was more supported and diffused than A/R and H/R because the number of likes and visualizations was statistically significantly higher (see Table 2). In Figure J10, we show the coefficients from OLS regressions where we additionally controlled for date fixed effects and cluster errors at the level of our artificial profiles. Not only were the no parametric results on H/O confirmed on every dimension, but we found that A/O attracted significantly more engagement and was more diffused than A/R on every dimension except for retweets. The above results are all the more striking, considering that H/O was not even the most followed-back narrative in our experiment. In fact, the A/O narrative was the most followed-back, exceeding H/O by a factor of 1.22 to 1.

Result 7: Out-group narratives drove more engagement over the pinned tweets and became much more diffused on Twitter than reciprocity ones. The corresponding hostility one, in particular, was the only narrative going "viral".

[30]See section 7.2 for an attempt at identifying the political orientation of these users and the diffusion process of the pinned tweets.

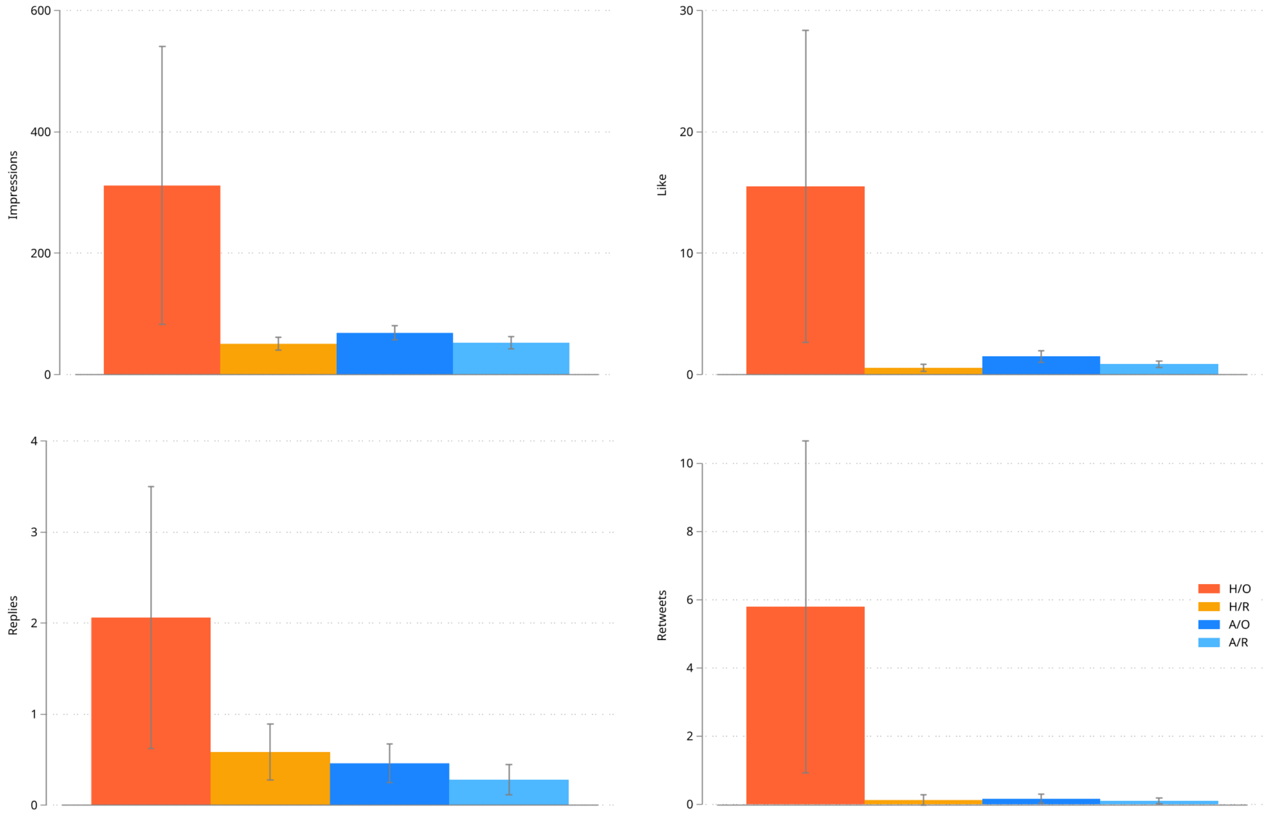


Figure 8: Pinned tweets by profile

Mean engagement associated with the four pinned tweets. The total sample includes 200 pinned tweets. Reported confidence intervals are at 95% level.

Table 2: Mann-Whitney ranksum tests and Cohen's d on H/O

	H/O vs H/R	H/O vs A/O	H/O vs A/R	A/O vs H/R	A/O vs A/R	H/R vs A/R
Visualizations						
Cohen's d	0.4311	0.4053	0.4325	-0.4449	0.4132	-0.0479
Mann-Whitney p-value	0.0004	0.1320	0.0014	0.0028	0.0075	0.5989
Likes						
Cohen's d	0.4407	0.4170	0.4363	-0.6458	0.4566	-0.3009
Mann-Whitney p-value	0.0000	0.1101	0.0070	0.0009	0.1165	0.0239
Replies						
Cohen's d	0.3811	0.4211	0.4703	0.1318	0.2613	0.3461
Mann-Whitney p-value	0.2816	0.2316	0.0137	0.9118	0.1433	0.1320
Retweets						
Cohen's d	0.4414	0.4432	0.4480	-0.0700	0.1522	0.0581
Mann-Whitney p-value	0.0000	0.0000	0.0000	0.3564	0.7030	0.5492

Cohen's d statistics and p-values from Mann-Whitney ranksum tests for treatments' comparisons over different measures of engagements with the pinned tweets. Observations are 52 iterations of pinned tweets with H/O, 48 with H/R, 50 with A/O, 50 with A/R.

7 Analyzing possible mechanisms

7.1 Users' activity on Twitter

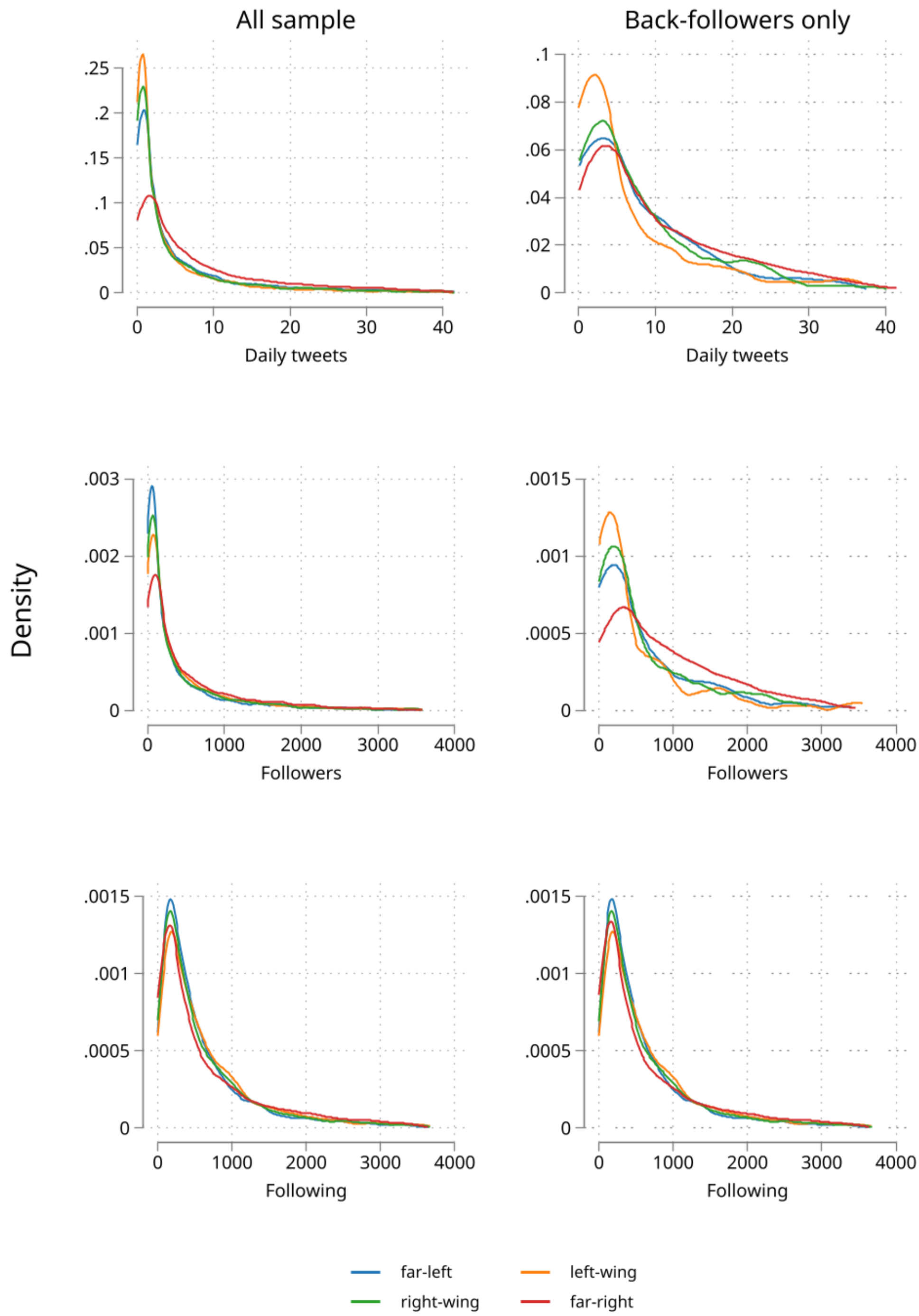


Figure 9: Twitter activity of our sample, overall and restricted to those who follow-back Thomas

Kernel densities with Epanechnikov kernel trimming the distributions at the 95 percentile over the number of daily tweets, of user's followers and of users followed. Panels on the left consider the whole sample (5009 for far-left, 5040 for left-wing, 5023 for right-wing, 4917 for far-right users), those on the right only the back-followers (149 for far-left, 154 for left-wing, 135 for right-wing and 504 for far-right users).

An interesting piece of evidence from our Twitter experiment was that far-right users followed-back more than users from other parties regardless of the narrative in the pinned tweets. This may have underlain the lower variability in their follow-back behavior on Twitter compared to their private preferences in Survey 1 (Result 6); they may have simply been following back our profiles more regardless of their narrative. To investigate this issue, we exploited the information we had on several dimensions of Twitter behavior from the users in our sample, specifically the number of tweets per day, the number of users who followed them, and the number of users whom they followed. Figure 9 summarizes this information with kernel densities both for the whole sample and for those who followed-back our profiles. The question we were interested here was: were far-right users in our sample more active on Twitter than the other ones? We use Mann–Whitney ranksum tests to answer this question.

The short answer is yes. Far-right users tended to tweet more than far-left and right-wing users, who, in turn, tweeted more than left-wingers. All these differences were significant ($p < 0.0001$ for all comparisons except left-wing vs. right-wing, $p = 0.008$). The distribution of tweeting was skewed to the right, with approximately 25% of the users tweeting more than the mean value. Far-right users also tended to have more followers. The mean values masked this result, but a comparison between distributions by parties shows that far-right users were the most followed. Then came left-wing users, followed by right-wing, and finally far-left users ($p < 0.0001$ for all comparisons except left-wing vs. right-wing, $p = 0.001$). The evidence was qualitatively the same if we considered the whole sample or if we restricted it to those who followed-back our profiles.

The "following" behavior presented a similar story, but only with respect to the subgroup of far-right users who ended up following our profiles. Far-right users who followed back our profiles followed other users more and the difference in distributions was significant with left-wing and right-wing users (left-wing $p = 0.001$, right-wing, $p = 0.031$, far-left, $p = 0.167$). However, on the whole sample the difference in distribution with right-wing and far-left users was not significant (right-wing $p = 0.894$, far-left $p = 0.440$) while the difference with left-wing ones was in the opposite direction and also statistically significant ($p = 0.0002$). To sum up, for those who followed our profiles back, general behavior may have mattered more than the narratives of our profiles.

7.2 Understanding the popularity of out-group hostility on social media

In this section, our aim is to understand the popularity of H/O through likes and retweets associated with the pinned tweets (Result 7). In fact, while we were unable to accurately identify the political orientation of users who engage with our pinned tweets, we managed to recover some information on them for a sub-sample of the pinned tweets[31]. More specifically, we retrieved the usernames of 25.8% (25.6%) of the total likes (retweets) associated with the pinned tweets. Interestingly, 65.13% (50.7%) of these users were not in our sample, meaning that they must have seen the tweet either in their timeline as a suggested tweet or from another user's retweet, and for them we do not have any information on their political orientation. This left us with 76 (31) users belonging to our sample[32]. Of these 76 (31), 78.9% (80.6%) liked (retweeted) H/O, and 93.3% (96%) of them were far-right users (see Table 3 for

[31] For unknown reasons, Twitter eliminated some of our pinned tweets from their database. We managed to recover likes (retweets) only for 22% (42%) of the tweets that had a positive number of likes (retweets).

[32] It is also interesting from a methodological perspective that some users were followed by one of our accounts, but then liked or retweeted the pinned tweet of another one of our accounts. Concerning H/O, only 10.9% (12%) of the likers (retweeters) are original viewers of H/O. For the other narratives instead, almost only original viewers become likers or retweeters. As with likers and retweeters not in our sample, the non-original viewers must have seen the tweet somewhere else, so either in their timeline as a suggested tweet or from another user's retweet. We observe the same behavior on following, as we have a limited group of users (138 over 1720 subjects) that follows more than one of our profiles.

the full decomposition by narrative and political orientation).

Table 3: **Likes and retweets in a subsample**

Party	Likes (retweets)				
	H/O	H/R	A/O	A/R	Total
far-left	0 (0)	0 (0)	2 (0)	1 (1)	3 (1)
left-wing	0 (0)	0 (0)	4 (2)	0 (0)	4 (2)
right-wing	4 (1)	1 (0)	2 (1)	2 (0)	9 (2)
far-right	56 (24)	3 (1)	0 (1)	1 (0)	60 (26)
Total	60 (25)	4 (1)	8 (4)	4 (1)	76 (31)

Statistics associated to retweets in parenthesis. The sample for these statistics comes from a subsample of pinned tweets we were able to retrieve and focuses on usernames we were able to associate to our original sample (see text for more details on this).

Not only almost all likes and retweets for H/O were from far-right users, but these users also tended to like and retweet H/O much more than H/R (56 vs 3 for likes, 24 vs 1 for retweets). Indeed, the numbers associated with the other narratives were in general very small, however they still followed a partisan pattern, as H/R was also liked or retweeted mostly by far-right individuals, A/O by left-wing ones and A/R seemed to have a mixed consensus. Combined with the previous evidence in section 7.1, this evidence suggests that the engagement in Twitter over H/O was driven by its popularity among far-right users rather than by its heightened capacity to overcome partisan barriers or spark debates. Furthermore, far-right users disproportionately chose H/O over the other hostile narrative H/R, meaning that they were mostly attracted by the most divisive one. Last but not least, H/O became popular well beyond our sample, suggesting that hostile polarizing narratives have the potential to spread on Twitter beyond the close network of followers even when coming from non-popular users like our artificial profiles.

7.3 Polarization across settings

The index proposed in section 2 can be applied to all our studies to gauge the overall level of polarization in a certain setting. The index can be straightforwardly applied to our survey. However, in the Twitter experiment, we lack information on Dis_j, p, and n_p with respect to engagements other than the follow-backs (see equation 6). We circumvent this issue assuming that the distribution of conditional probabilities that a 'liker' or 'retweeter' came from a certain party given a certain narrative was the same in the Twitter as in the private preferences in the survey. This assumption seems reasonable as the distribution of private preferences is close to that of the follow-backs on Twitter (see Result 6 and section 6.2.1).[33] If we took as reference the likes and retweets of section 7.2 instead, one could interpret our estimates below as very conservative. The values of the index so calculated for each party and each setting can be found in Table 4.

[33]For example, considering an H/O private preference in the survey, there was a 14.4% probability that it came from a far-left supporter, 18.5% from a left-wing one, 21.2% from a right-wing one and 46% from a far-right one (see Figure 7 with the probabilities computed across parties keeping the narrative constant).The same H/O narrative in the pinned tweets had been liked 806 times on Twitter. We used the probabilities just calculated on the private preferences to impute this 806 to the parties: this left us with 116.5 far-left users, 149 left-wing ones, 170.6 right-wing ones, and 370 far-right ones. Summing up across the narratives for each party gave us an estimate of n_{ps} for Twitter's likes and retweets.

The index so constructed ranges between 15.67 and 32.65 at the party level. It is noticeable that, in all of the seven settings we consider, far-right voters are always those choosing the most polarizing narratives. Most of all, while private preferences, static public preferences and follow-backs in Twitter yield similar levels of polarization, dynamic engagements on Twitter - with likes and retweets - bring about an increase of about half in polarization (from approximately 17 for the former settings to approximately 26 in the latter). Finally, it is also noticeable that narratives over COVID-19 vaccination induced less polarization than immigration narratives, as the polarization index decreases by about 20% for private preferences (from 17.2 to 14.3) and by about 29% for public static preferences (from 17.7 to 13.5).

Table 4: Polarization index by party and by setting

	Party group				
	far left	left	right	far right	Total
Private preferences	16.47	16.60	15.75	19.91	17.18
Static public preferences	16.88	17.18	15.67	20.87	17.65
Twitter – Follow-backs	17.18	16.66	17.79	18.46	17.53
Twitter – Likes	22.54	22.80	25.23	32.16	25.68
Twitter – Retweets	23.78	23.85	26.55	32.65	26.71
Private preferences on vacc.	13.27	13.28	13.50	17.23	14.32
Static public preferences on vacc.	12.09	13.03	12.55	16.25	13.48

The polarization index for each party captures the average share of other parties' voters who disliked the liked narratives of own party's voters. See the main text for details on how the polarization index is calculated, especially in the case of Likes and Retweets on Twitter.

7.4 Study 4: Underlying mechanisms and additional results

7.4.1 Design of the second survey

In an additional survey conducted one week after the previous one, we explored the possible mechanisms behind the behavior observed in the previous studies[34]. We recruited 771 participants through Kantar between the 24th and 29th of September 2021 using the same quotas for party affiliation, age, and gender as in Survey 1. We recruited 105 far-left, 314 left-wing, 214 right-wing and 138 far-right voters. The samples from Surveys 1 and 2 are not statistically significantly different in any observed characteristic for any party group, except for 3 rejections of the null at the 10% level out of 36 tests that would reduce to zero if errors were corrected for multiple hypotheses testing (see Table J1 and J2 for the descriptive statistics, see Table J3 for the KS tests). This entails that we can assume that preferences and attitudes elicited in the two surveys are generated by the same distribution.

[34]Similar to Study 1, we added some questions on the topic of COVID-19 vaccination at the end.

The first goal of this study was to build an individual-level distinctiveness ranking of narratives, as outlined in sections 2 and 5.2.4. For this purpose, we elicited normative and empirical beliefs over the most preferred narratives for one's own party and other parties (see section 5.2.4 for more details and questions 9-23 in Appendix L). Eliciting beliefs separately from preferences should provide a cleaner measurement of beliefs. Beliefs were monetarily incentivized, as a correct guess earned participants an additional €0.50 on top of the participation fee. Normative expectations are analyzed in Appendix I.

The second goal was to elicit the emotions triggered by the narratives and political correctness (see questions 8), as these can provide insights on the psychological mechanisms used by individuals to prefer or endorse certain narratives. For each narrative, we used 7-point Likert scales to elicit two negative emotions (fear and anger) and a positive emotion (happiness) taken from the six basic emotions proposed by Ekman, Friesen, and Ellsworth (2013). Always with a 7-point Likert scale, we also elicited the participants' opinions on whether the narrative was, to the best of their judgment, politically correct. The full text of the survey can be found in Appendix L.

7.4.2 Emotions and political correctness associated with the narratives

Emotions could be an important determinant of which narrative to prefer privately and endorse publicly. Indeed, it has been shown that political communication by populist parties is often charged with negative emotions and that these - especially anger - influence voters' attitudes (Rico, Guinjoan, and Anduiza, 2017; Caiani and Di Cocco, 2023; Widmann, 2021). We first note that far-right supporters declared on average higher emotional involvement than all other party groups (Mann-Whitney tests: $p < 0.0001$ in every pairwise comparison). At the other extreme, left-wing supporters were those stating, on average, the lowest emotional involvement. However, the pairwise differences are significant only with respect to right-wing supporters (Mann-Whitney test: $p = 0.010$).

In Appendix H we show descriptive statistics associated with the emotions (see Figure H1) and analyze differences across emotions and by political position for each political group. Two aspects emerged that are related to our results from Survey 1. First, negative emotions were associated with hostile narratives (both in comparison with positive emotions and with acceptance narratives) for all party groups except for far-right supporters. This may have underlain the decision of individuals not to endorse their hostile preference in the public vs. the private setting. Second, H/I was associated with more positive emotions than the other hostile narratives for all party groups. This could well have defined its status as the weakest narrative in terms of distinctiveness for far-right supporters.

We now try to directly account for the role of emotions in explaining private preferences, the choice of becoming an endorser, static public preferences and follow-backs on Twitter. For this goal, we first associated each private or public preference in Survey 1 and the Twitter profile in the social media experiment with the corresponding average values of emotions at party level. We then ran regressions where the variables of interest were the emotions associated with each narrative and their interactions with the party groups. In the models on the chosen private and public preferences, we associated 6 observations to each participant, each corresponding to a narrative, and the outcome variable was a dummy identifying the chosen one. Although we cannot claim causality, these associations are suggestive of possible mechanisms.

Results in Table 5 suggest that the chosen narrative in the survey was mostly associated with negative emotions, both the private (column (1)) and the public (column (5)) preference, and the choice of becoming an endorser (column (3)). Respondents shied away from narratives associated with anger and embraced those associated with fear, while the coefficient of happiness was smaller in size and almost never significant. These results masked an interesting heterogeneity by party group, as left-wing voters chose narratives associated with more happiness and anger rather than fear. For the choice of endorsing (column (4)), the interaction estimates were noisier with all coefficients losing significance except for far-right endorsing more when the private preference was associated with fear. As for follow-backs on Twitter (columns (7) and (8)), all emotions were positively associated with the follow-back rate. However, contrary to the survey, the relation with happiness on social media was stronger than that with anger and fear (by a Wald post-estimation test, $p = 0.072$ and $p = 0.002$ respectively), and in turn anger was more associated with the follow-back than fear ($p = 0.067$). These estimates, in the interacted model, were significant only for the far-right, and the sign of fear even switched to negative.

Table 5: **Preferences and follow-backs by emotions**

	Chosen narrative in private		Endorser in public setting		Chosen narrative in public		Follow-back	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Happiness	0.070***		0.034		0.046*		0.066***	
	(0.017)		(0.050)		(0.025)		(0.008)	
Party group # Happiness								
far left		0.065		0.207		0.063		0.017
		(0.041)		(0.133)		(0.060)		(0.018)
left		0.115***		0.034		0.136***		0.013
		(0.031)		(0.100)		(0.046)		(0.013)
right		0.034		0.102		0.070		0.013
		(0.033)		(0.106)		(0.049)		(0.016)
far right		0.092***		-0.023		0.042		0.106**
		(0.026)		(0.089)		(0.039)		(0.050)
Anger	-0.108***		-0.186**		-0.235***		0.056***	
	(0.027)		(0.088)		(0.041)		(0.011)	
Party group # Anger								
far left		-0.099		0.261		-0.248**		0.018
		(0.085)		(0.269)		(0.123)		(0.034)
left		0.185***		-0.113		0.119		0.021
		(0.066)		(0.207)		(0.098)		(0.028)
right		-0.374**		-0.097		-0.911***		0.092
		(0.146)		(0.494)		(0.234)		(0.084)
far right		-0.181***		-0.196		-0.250***		0.079*
		(0.042)		(0.141)		(0.063)		(0.047)
Fear	0.146***		0.155***		0.244***		0.024***	
	(0.020)		(0.059)		(0.029)		(0.007)	
Party group # Fear								
far left		0.167		-0.472		0.331*		-0.026
		(0.117)		(0.372)		(0.169)		(0.047)
left		-0.185**		0.064		-0.132		-0.028
		(0.074)		(0.215)		(0.110)		(0.038)
right		0.461***		-0.021		0.976***		-0.096
		(0.146)		(0.491)		(0.235)		(0.083)
far right		0.224***		0.197**		0.330***		-0.115**
		(0.026)		(0.096)		(0.039)		(0.058)
Number of observations	7356	7356	1226	1226	3132	3132	19989	19989
R-squared	0.026	0.034	0.045	0.058	0.054	0.067	0.022	0.029

*** $p<.01$, ** $p<.05$, * $p<.1$

OLS on the chosen narrative as private preferences in columns (1) and (2), on the choice of becoming an endorser in columns (3) and (4), on the chosen narrative as public preference for the subsample of the endorsers in columns (5) and (6), on the choice of following-back in columns (7) and (8). See text for more details on how emotions are associated to the narratives. Controls in column (3) and (4) include age (3 categories), immigration status, education (4 categories), income (4 categories), employment status (3 categories), occupation (3 categories), and religion (4 categories). Controls in columns (7) and (8) include the number of users followed, the number of users whom they follow, the number of tweets posted, the year in which users signed up.

We now turn to political correctness. Descriptive statistics by narrative and political orientation are shown in Figure H2. We conjectured that the perception of the degree to which a narrative was "politically correct" could have been a driver of individual preferences, both in the survey private choice but also in the change of preferences between private and public survey. This may support a social image explanation to how individuals changed their preferences when going from private to public and to how they behaved on Twitter. As we did with emotions, we tried to directly account for the role of political correctness in explaining private preferences, the choice of becoming an endorser, public preferences and follow-backs on Twitter. We followed the same steps and show the results in Table H1 in Appendix H. Higher political correctness was strongly associated with the chosen narratives both in private and in public, to the choice of endorsing a narrative and of following-back on Twitter. Interestingly, this held for every party group with comparable intensities.

8 Discussion

Our research design aimed to combine a "controlled" survey with a natural experiment, and we must acknowledge some limitations of our approach. Plenty methodological studies have discovered comprehensive distortions of private preferences in survey measurements, including framing and anchoring effects, social desirability biases, and cognitive dissonance (Manski, 1990; Bertrand and Mullainathan, 2001; De Quidt, Haushofer, and Roth, 2018; Galizzi and Navarro-Martinez, 2019; Falco and Zaccagni, 2021). Although the existence of such biases is undeniable, we do not believe it hampers the validity of our results. First of all, the elicitation of public static preferences is exposed to the same type of conditions, i.e., observation of the participant's decisions by the experimenter, the use of an online survey as a medium for the interview, confidentiality guaranteed by the opinion poll company, as the elicitation of private preferences. As such, the two measures are truly comparable. Additionally, recent studies convincingly point out that measurement of individual preferences through surveys, either having an experimental nature or not, have both internal and external validity and have predictive power for real-life behavior (Jahedi and Méndez, 2014; Hainmueller, Hangartner, and Yamamoto, 2015; Kistler, Thöni, and Welzel, 2017; Snowberg and Yariv, 2021; Kaiser and Oswald, 2022).

Another possible limitation of our design is that the focus on either the in-group or the out-group may have been imperfect because inevitably mentioning the in-group highlights the existence of an out-group and *vice versa*. This is the reason why we preferred to talk about a "'focus" on a group or on reciprocity rather than a full-fledged group. Even if it could be questionable whether we can interpret narratives that we labelled as in-group-focused (out-group-focused) as truly representing the construct they refer to, we believe that the essential thing is that participants did react significantly to the use of one or the other narrative (see Section 4.2). The fact that some patterns of behavior appear to hold regardless of political affiliation (see Result 1) adds, we believe, validity to our findings. We also have to accept that the narratives used in the survey needed to be shortened to be compatible with the Twitter length restrictions. Even if this is the case, we do not believe that shortening the narrative affected their content in any meaningful way. This should be evident by a comparison between the main narratives and the Twitter ones, as the only changes are stylistic and not about content (see Section 3 and Appendix G). This is also confirmed by Result 6, i.e. the fact that the initial manifestation of preference over narratives in the Twitter experiment, that is, the follow-back, was remarkably similar to the private preferences measurement in the survey.

Similar considerations apply to the fact that we reduced the number of narratives studied on Twitter by two compared to the survey experiment. On the one hand, this adjustment was necessary given the relatively small size of the sample of Twitter users that could be classified according to their political affiliation. In fact, our search algorithm to include Twitter users in our subject pool was increasingly failing to find eligible subjects as our sampling was progressing. We suspect that our subject pool of n=19,989 is close to the total number of users eligible for our research requirements (listed in Appendix G) in the whole population of German Twitter users. Since the follow-back rate was very low, having six "treatments" in our natural experiment would have considerably increased the chances of Type I error. On the other hand, running a between-subject survey, in which each participant is exposed to only one narrative, was impractical for similar reasons, as we went close to exhausting the entire sample pool by Kantar for participants with extreme political orientation.

We also must acknowledge that, as in every natural experiment, some considerable loss of control cannot be prevented. Even if our design for the Twitter experiment aimed to be "between-subject", we later found that some subjects had been exposed to more than one of our "Thomas-accounts". Most likely, this is due to the fact that the algorithm proposing matches to Twitter users noted some similarity between our "Thomas-accounts", thus proposing to a user who previously followed back one of our "Thomas-accounts" the possibility to follow another "Thomas-account" (see section 7.2 for more details). Given the overwhelming diffusion of the out-group hostility narrative compared to the other narratives (Result 7), though, we believe that this shortcoming only had a marginal impact on the results while the related evidence provides an interesting insight on how Twitter algorithms work.

Another cause of concern would be if Result 7 on H/O going "viral" was entirely driven by Twitter algorithms. Since Twitter algorithms spread more the posts that are more likely to create interest and engagement by Twitter users, a post that would receive a high level of engagement on one day would then be likely to be disseminated even the next day. If, by pure chance, the first pinned tweet that grew in traction was the one that included the H/O narrative, then the algorithm may have simply continued to select that same pinned tweet. The spread of the H/O narrative may then be a complete fluke. The data completely reject this possibility. If the above conjecture were true, we would be able to identify a clear monotonic trend over time in how the H/O narrative became widespread. When analyzing each of the engagement measures, however, we could not detect any time trends both in terms of general patterns and through statistical tests.[35] To be sure, it is still possible that Twitter algorithms identified the H/O narrative as the one having greater potential to spur engagement, possibly on the basis of its wording and content. For this reason, it then repeatedly became widespread. In fact, we believe that this account is quite plausible and adds to the greater propensity to engage by far-right users to bring about Result 7. Unfortunately, we are not able to quantify the relative weight of the Twitter algorithm's way of operation and of far-right users' patterns of engagement on social media.

We believe that the main limitation of our study is that we were unable to track the behavior of the survey participants on social media and Twitter in particular. Personal identification, even with prior consent by participants, was denied by the opinion poll company. Even asking survey participants the extent to which they were active on social media in a follow-up survey was not approved by the company. Therefore, we are unable to measure the extent to which the magnification of some narrative preferences that we observed in the public static in comparison with the private setting would likely carry over on social media. Even if this piece of information would have been important in interpreting our results, we believe that the additional information we were able to gather from

[35]We ran OLS regressions of visualizations, likes, replies and retweets over a time trend measured at the minute level. The coefficient was never significant in several specifications, with p-values ranging from 0.71 to 0.95. Results are available upon request.

the Twitter environment (see Section 7.1 and 7.2) suffices to give a reasonably accurate picture of the mechanisms at play.

9 Conclusions

The main goal of this paper was to investigate the extent to which preferences over narratives significantly differ in private and public contexts and whether they become more polarized in public vis-à-vis private settings. We focused on narratives of migration, one of the most divisive issues in the political discourse. To do this, we developed six different narratives, which kept constant one fact - namely, the number of migrants having entered Germany from 2015 to 2020, and varied (a) their acceptance or hostility of migrants, and (b) their focus on the in-group, the out-group, or reciprocity. Our samples were stratified across different dimensions including participants' political orientation.

Our first result is that narratives do seem to matter to people. While, as expected, people leaning to the right (left) of the political spectrum preferred hostility (acceptance) narratives to others, people seemed to like some narratives significantly more than others in our survey eliciting private preferences. Interestingly, the ranking of the narratives showed striking similarities across all four political groups. For instance, the reciprocity hostile narrative was ranked above the other hostile narratives across all political parties, and the same happened with the outgroup acceptance one.

Our second key result is that the manifestation of narratives differed significantly in the private vis-à-vis public setting that we labelled "static". When individuals were asked to publicly endorse one of the six narratives on a publicly accessible website, people from the left or the center of the political spectrum who had supported hostility narratives in the private setting opted out from endorsing a narrative publicly. On the other hand, far-righters tended to endorse narratives publicly in line with their private preferences. The overall result is that narratives seemed to be less hostile in the static public than in the private context. By additional analyses, this "distortion" in the manifestation of preferences over narratives is more consistent with a willingness to support the narrative that is most distinctive for one's own political group than with pluralistic ignorance.

We then linked one of the four most preferred narratives in the private setting to one "pinned tweet" of four artificial profiles created on Twitter. The other characteristics of the profile were kept as constant as possible. Our experimental stimulus consisted of having each of the four profiles follow a random set of Twitter users every day. Before the experiment, we had classified n=19,989 German Twitter users according to their political orientation. Our first measure of preference manifestation in this setting was, perhaps surprisingly, in line with private preferences. The follow-back rate of narratives by Twitter users mirrored quite closely the preferences expressed privately - even more than those expressed in the static public setting. The notable exceptions were far-righters who appeared to follow the four profiles in a similar proportion and who appeared to follow-back more frequently. Nonetheless, as soon as Twitter users were left free to interact over the narratives, the one narrative that appeared to receive widespread attention was the hostility out-group narrative, in spite of not being the most preferred narrative either privately or static-publicly. This was the case across all possible measures of dynamic engagement, i.e., visualizations, likes, replies, and retweets. Our supplementary investigation ascertained that this result is at least partly due to far-righters being conspicuously more active than Twitter users from the other groups in spreading hostile outgroup narratives. However, it is also plausible that the algorithm selecting Twitter posts for diffusion played a considerable role in this result.

Our index of polarization signals a significantly higher level of polarization on Twitter for all engagements different from the follow-back than in all other cases. In a control experiment conducted in the private and static public setting only, we also ascertained a significantly lower level of polarization in narratives applied to COVID-19 vaccines than in those applied to migration. Moreover, we did not find any bias in the static public setting in comparison with the private setting over-vaccination narratives.

An open question of our study is whether this result is caused by the functioning of algorithms, self-selection into Twitter, or different patterns of engagement by users from different political affiliations. Even if our analysis on this cannot be more than speculative, the fact that the initial form of engagement with narratives on social media mirrors private preferences seems to go against the hypothesis of self-selection. This leaves the other two mechanisms as those more likely to account for our results.

Most of the literature has thus far focused on the impact of fake news in producing polarization or echo-chamber effects. The main result of this paper is, arguably, that distorting facts is not necessary to create polarization. The mere use of one rather than another narrative may suffice to induce significantly higher polarization on social media than in private or static public contexts. Although policy implications cannot be directly derived from our study, we would advocate social media companies to be more transparent on how social media algorithms actually work, as they appear to considerably distort preferences as expressed privately, fostering polarization.

References

Acemoglu, Daron, Asuman Ozdaglar, and James Siderius (June 2021). *A Model of Online Misinformation*. Working Paper 28884. National Bureau of Economic Research. DOI: `10.3386/w28884`. URL: `http://www.nber.org/papers/w28884`.

Akerlof, George and Dennis J Snower (2016). "Bread and bullets". In: *Journal of Economic Behavior & Organization* 126, pp. 58–71.

Akerlof, George A and Rachel E Kranton (2000). "Economics and identity". In: *The quarterly journal of economics* 115.3, pp. 715–753.

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva (2020). "The polarization of reality". In: *AEA Papers and Proceedings*. Vol. 110. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 324–328.

— (2023). "Immigration and redistribution". In: *The Review of Economic Studies* 90.1, pp. 1–39.

Alesina, Alberto and Marco Tabellini (2024). "The political effects of immigration: Culture or economics?" In: *Journal of Economic Literature* 62.1, pp. 5–46.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow (2020). "The welfare effects of social media". In: *American economic review* 110.3, pp. 629–676.

Allcott, Hunt and Matthew Gentzkow (2017). "Social media and fake news in the 2016 election". In: *Journal of economic perspectives* 31.2, pp. 211–236.

Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk (2024). "Globally representative evidence on the actual and perceived support for climate action". In: *Nature Climate Change* 14.3, pp. 253–259.

Antinyan, Armenak, Thomas Bassetti, Luca Corazzini, and Filippo Pavesi (2024). "Narratives and opinion polarization: a survey experiment". In: *Scientific reports* 14.1, p. 19732.

Arzheimer, Kai (2015). "The AfD: finally a successful right-wing populist Eurosceptic party for Germany?" In: *West European Politics* 38.3, pp. 535–556.

Arzheimer, Kai and Carl C Berning (2019). "How the Alternative for Germany (AfD) and their voters veered to the radical right, 2013–2017". In: *Electoral Studies* 60, p. 102040.

Bail, Chris (2021). *Breaking the social media prism*. Princeton University Press.

Bail, Christopher A, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky (2018). "Exposure to opposing views on social media can increase political polarization". In: *Proceedings of the National Academy of Sciences* 115.37, pp. 9216–9221.

Barrera, Oscar, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya (2020). "Facts, alternative facts, and fact checking in times of post-truth politics". In: *Journal of Public Economics* 182, p. 104123.

Bénabou, Roland, Armin Falk, and Jean Tirole (2018). *Narratives, imperatives, and moral reasoning*. Tech. rep. National Bureau of Economic Research.

Bénabou, Roland and Jean Tirole (2006). "Incentives and prosocial behavior". In: *American economic review* 96.5, pp. 1652–1678.

Berbuir, Nicole, Marcel Lewandowsky, and Jasmin Siri (2015). "The AfD and its sympathisers: Finally a right-wing populist movement in Germany?" In: *German Politics* 24.2, pp. 154–178.

Bertrand, Marianne and Sendhil Mullainathan (2001). "Do people mean what they say? Implications for subjective survey data". In: *American Economic Review* 91.2, pp. 67–72.

Bicchieri, Cristina (2016). *Norms in the wild: How to diagnose, measure, and change social norms.* Oxford University Press.

Bieber, Ina, Sigrid Roßteutscher, and Philipp Scherer (2018). "Die Metamorphosen der AfD-Wählerschaft: Von einer euroskeptischen Protestpartei zu einer (r) echten Alternative?" In: *Politische Vierteljahresschrift* 59.3, pp. 433–461.

Bos, Linda and Sophie Minihold (2022). "The Ideological Predictors of Moral Appeals by European Political Elites; An Exploration of the Use of Moral Rhetoric in Multiparty Systems". In: *Political Psychology* 43.1, pp. 45–63.

Bowen, Renee, Danil Dmitriev, and Simone Galperti (2021). *Learning from shared news: when abundant information leads to belief polarization.* Tech. rep. National Bureau of Economic Research.

Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro (2017). "Greater Internet use is not associated with faster growth in political polarization among US demographic groups". In: *Proceedings of the National Academy of Sciences* 114.40, pp. 10612–10617.

Brewer, Marilynn B (1999). "The psychology of prejudice: Ingroup love and outgroup hate?" In: *Journal of social issues* 55.3, pp. 429–444.

Brug, Wouter van der and Eelco Harteveld (2021). "The conditional effects of the refugee crisis on immigration attitudes and nationalism". In: *European Union Politics* 22.2, pp. 227–247.

Brunner, Beatrice and Andreas Kuhn (2018). "Immigration, cultural distance and natives' attitudes towards immigrants: Evidence from Swiss voting results". In: *Kyklos* 71.1, pp. 28–58.

Bursztyn, Leonardo, Georgy Egorov, Ingar Haaland, Aakaash Rao, and Christopher Roth (2022). "Scapegoating during crises". In: *AEA Papers and Proceedings.* Vol. 112, pp. 151–55.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott (2020). "Misperceived social norms: Women working outside the home in Saudi Arabia". In: *American economic review* 110.10, pp. 2997–3029.

Bursztyn, Leonardo, Ingar K Haaland, Aakaash Rao, and Christopher P Roth (2020). *Disguising prejudice: Popular rationales as excuses for intolerant expression.* Tech. rep. National Bureau of Economic Research.

Caiani, Manuela and Jessica Di Cocco (2023). "Populism and emotions: a comparative study using Machine Learning". In: *Italian Political Science Review/Rivista Italiana di Scienza Politica* 53.3, pp. 351–366.

Card, Dallas, Serina Chang, Chris Becker, Julia Mendelsohn, Rob Voigt, Leah Boustan, Ran Abramitzky, and Dan Jurafsky (2022). "Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration". In: *Proceedings of the National Academy of Sciences* 119.31, e2120510119.

Card, David, Christian Dustmann, and Ian Preston (2012). "Immigration, wages, and compositional amenities". In: *Journal of the European Economic Association* 10.1, pp. 78–119.

Cinelli, Matteo, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini (2021). "The echo chamber effect on social media". In: *Proceedings of the National Academy of Sciences* 118.9.

Cohen, Jacob (1988). *Statistical power analysis for the behavioral sciences.* Routledge.

Columbus, Simon, Isabel Thielmann, Ingo Zettler, and Robert Böhm (2023). "Parochial reciprocity". In: *Evolution and Human Behavior* 44.2, pp. 131–139.

Costello, Thomas H, Gordon Pennycook, and David G Rand (2024). "Durably reducing conspiracy beliefs through dialogues with AI". In: *Science* 385.6714, eadq1814.

Crockett, Molly J (2017). “Moral outrage in the digital age”. In: *Nature human behaviour* 1.11, pp. 769–771.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth (2018). “Measuring and bounding experimenter demand”. In: *American Economic Review* 108.11, pp. 3266–3302.

Dinas, Elias, Konstantinos Matakos, Dimitrios Xefteris, and Dominik Hangartner (2019). “Waking up the golden dawn: does exposure to the refugee crisis increase support for extreme-right parties?” In: *Political analysis* 27.2, pp. 244–254.

Djourelova, Milena (2023). “Persuasion through slanted language: Evidence from the media coverage of immigration”. In: *American economic review* 113.3, pp. 800–835.

Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm (2019). “Refugee migration and electoral outcomes”. In: *The Review of Economic Studies* 86.5, pp. 2035–2091.

Edo, Anthony, Yvonne Giesing, Jonathan Öztunc, and Panu Poutvaara (2019). “Immigration and electoral support for the far-left and the far-right”. In: *European Economic Review* 115, pp. 99–143.

Ekman, Paul, Wallace V Friesen, and Phoebe Ellsworth (2013). *Emotion in the human face: Guidelines for research and an integration of findings*. Vol. 11. Elsevier.

Eliaz, Kfir and Ran Spiegler (2020). “A model of competing narratives”. In: *American Economic Review* 110.12, pp. 3786–3816.

Engelmann, Dirk and Martin Strobel (2000). “The false consensus effect disappears if representative information and monetary incentives are given”. In: *Experimental Economics* 3, pp. 241–260.

Facchini, Giovanni and Anna Maria Mayda (2009). “Does the welfare state affect individual attitudes toward immigrants? Evidence across countries”. In: *The review of economics and statistics* 91.2, pp. 295–314.

— (2012). “Individual attitudes towards skilled migration: An empirical analysis across countries”. In: *The World Economy* 35.2, pp. 183–196.

Falco, Paolo and Sarah Zaccagni (2021). “Promoting social distancing in a pandemic: Beyond good intentions”. In: *PloS one* 16.12, e0260457.

Fujiwara, Thomas, Karsten Müller, and Carlo Schwarz (2024). “The effect of social media on elections: Evidence from the united states”. In: *Journal of the European Economic Association* 22.3, pp. 1495–1539.

Galasso, Vincenzo, Massimo Morelli, Tommaso Nannicini, and Piero Stanig (2022). “Fighting populism on its own turf: Experimental evidence”. In.

Galasso, Vincenzo, Tommaso Nannicini, and Salvatore Nunnari (2023). “Positive Spillovers from Negative Campaigning”. In: *American Journal of Political Science* 67.1, pp. 5–21. DOI: `https://doi.org/10.1111/ajps.12610`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12610`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12610`.

Galizzi, Matteo M and Daniel Navarro-Martinez (2019). “On the external validity of social preference games: a systematic lab-field study”. In: *Management Science* 65.3, pp. 976–1002.

Gärtner, Lea and Alexander Wuttke (2019). “Eingeschlossen in der Filter Bubble? Politische Kommunikationsnetzwerke im Wahlkampf”. In: *Zwischen Polarisierung und Beharrung: Die Bundestagswahl 2017*, pp. 97–110.

Garz, Marcel, Jil Sörensen, and Daniel F Stone (2020). “Partisan selective engagement: Evidence from Facebook”. In: *Journal of Economic Behavior & Organization* 177, pp. 91–108.

Gehring, Kai, Joop Age Harm Adema, and Panu Poutvaara (2022). "Immigrant Narratives". In: *CESifo Working Paper* 10026. URL: https://www.econstor.eu/bitstream/10419/267259/1/cesifo1_wp10026.pdf.

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer (2019). "Fake news on Twitter during the 2016 US presidential election". In: *Science* 363.6425, pp. 374–378.

Groseclose, Tim and Jeffrey Milyo (2005). "A measure of media bias". In: *The quarterly journal of economics* 120.4, pp. 1191–1237.

Haidt, Jonathan (2012). *The righteous mind: Why good people are divided by politics and religion.* Vintage.

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto (2015). "Validating vignette and conjoint survey experiments against real-world behavior". In: *Proceedings of the National Academy of Sciences* 112.8, pp. 2395–2400.

Hambauer, Verena and Anja Mays (2018). "Wer wählt die AfD?–ein vergleich der sozialstruktur, politischen einstellungen und einstellungen zu flüchtlingen zwischen AfD-WählerInnen und der WählerInnen der anderen Parteien". In: *Zeitschrift für vergleichende Politikwissenschaft* 12.1, pp. 133–154.

Hameleers, Michael and Toni GLA Van der Meer (2020). "Misinformation and polarization in a high-choice media environment: How effective are political fact-checkers?" In: *Communication Research* 47.2, pp. 227–250.

Harmon, Nikolaj A (2018). "Immigration, ethnic diversity, and political outcomes: Evidence from Denmark". In: *The Scandinavian Journal of Economics* 120.4, pp. 1043–1074.

Henry, Emeric, Ekaterina Zhuravskaya, and Sergei Guriev (2022). "Checking and sharing alt-facts". In: *American Economic Journal: Economic Policy* 14.3, pp. 55–86.

Jahedi, Salar and Fabio Méndez (2014). "On the advantages and disadvantages of subjective measures". In: *Journal of Economic Behavior & Organization* 98, pp. 97–114.

Kahan, Dan M (2013). "Ideology, motivated reasoning, and cognitive reflection". In: *Judgment and Decision making* 8.4, pp. 407–424.

Kaiser, Caspar and Andrew J Oswald (2022). "The scientific value of numerical measures of human feelings". In: *Proceedings of the National Academy of Sciences* 119.42, e2210412119.

Kaufmann, Eric (2019). "Can narratives of white identity reduce opposition to immigration and support for hard Brexit? A survey experiment". In: *Political Studies* 67.1, pp. 31–46.

Keita, Sekou, Thomas Renault, and Jérôme Valette (2024). "The usual suspects: Offender origin, media reporting and natives' attitudes towards immigration". In: *The Economic Journal* 134.657, pp. 322–362.

Kistler, Deborah, Christian Thöni, and Christian Welzel (2017). "Survey response and observed behavior: Emancipative and secular values predict prosocial behaviors". In: *Journal of Cross-Cultural Psychology* 48.4, pp. 461–489.

Lazer, David MJ, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. (2018). "The science of fake news". In: *Science* 359.6380, pp. 1094–1096.

Lehmann, Pola, Tobias Burst, Theres Matthieß, Sven Regel, Andrea Volkens, Bernhard Weßels, and Lisa Zehnter (2022). *The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR).*

Version 2022a. Berlin. DOI: 10.25522/manifesto.mpds.2022a. URL: https://doi.org/10.25522/manifesto.mpds.2022a.

Lelkes, Yphtach, Gaurav Sood, and Shanto Iyengar (2017). "The hostile audience: The effect of access to broadband internet on partisan affect". In: *American Journal of Political Science* 61.1, pp. 5–20.

Leonardelli, Geoffrey J, Cynthia L Pickett, and Marilynn B Brewer (2010). "Optimal distinctiveness theory: A framework for social identity, social cognition, and intergroup relations". In: *Advances in experimental social psychology*. Vol. 43. Elsevier, pp. 63–113.

Levi, Eugenio, Rama Dasi Mariani, and Fabrizio Patriarca (2020). "Hate at first sight? Dynamic aspects of the electoral impact of migration: the case of Ukip". In: *Journal of Population Economics* 33.1, pp. 1–32.

— (2023). "Immigration and Voting". In: *Handbook of Labor, Human Resources and Population Economics*, pp. 1–16.

Levy, Gilat and Ronny Razin (2019). "Echo chambers and their effects on economic and political outcomes". In: *Annual Review of Economics* 11.1, pp. 303–328.

Levy, Ro'ee (2021). "Social media, news consumption, and polarization: Evidence from a field experiment". In: *American economic review* 111.3, pp. 831–70.

Manacorda, Marco, Guido Tabellini, and Andrea Tesei (2022). "Mobile Internet and the Rise of Political Tribalism in Europe". In: URL: https://www.cesifo.org/DocDL/cesifo1_wp9955.pdf.

Manski, Charles F (1990). "The use of intentions data to predict behavior: A best-case analysis". In: *Journal of the American Statistical Association* 85.412, pp. 934–940.

Manzoni, Elena, Elie Murard, Simone Quercia, and Sara Tonini (2024). "News, emotions, and policy views on immigration". In.

Marks, Gary and Norman Miller (1987). "Ten years of research on the false-consensus effect: An empirical and theoretical review." In: *Psychological bulletin* 102.1, p. 72.

Marx, Paul and Elias Naumann (2018). "Do right-wing parties foster welfare chauvinistic attitudes? A longitudinal study of the 2015 'refugee crisis' in Germany". In: *Electoral Studies* 52, pp. 111–116.

Mayda, Anna Maria (2006). "Who is against immigration? A cross-country investigation of individual attitudes toward immigrants". In: *The review of Economics and Statistics* 88.3, pp. 510–530.

Mayda, Anna Maria, Giovanni Peri, and Walter Steingress (2022). "The political impact of immigration: Evidence from the United States". In: *American Economic Journal: Applied Economics* 14.1, pp. 358–389.

Meier, Tabea, Ryan L Boyd, James W Pennebaker, Matthias R Mehl, Mike Martin, Markus Wolf, and Andrea B Horn (2019). ""LIWC auf Deutsch": The development, psychometrics, and introduction of DE-LIWC2015". In: *PsyArXiv* a.

Miller, Dale (2023). "A century of pluralistic ignorance: what we have learned about its origins, forms, and consequences". In: *Frontiers in Social Psychology* 1, p. 1260896.

Moriconi, Simone, Giovanni Peri, and Riccardo Turati (2022). "Skill of the immigrants and vote of the natives: Immigration and nationalism in European elections 2007–2016". In: *European Economic Review* 141, p. 103986.

Mosleh, Mohsen, Cameron Martel, Dean Eckles, and David G Rand (2021). "Shared partisanship dramatically increases social tie formation in a Twitter field experiment". In: *Proceedings of the National Academy of Sciences* 118.7.

Narayanan, Arvind (2023). *Understanding social media recommendation algorithms.* Working Paper. Knight First Amendment Institute: Algorithmic Amplification and Society. DOI: https://doi.org/10.7916/khdk-m460. URL: https://academiccommons.columbia.edu/doi/10.7916/khdk-m460.

Pariser, Eli (2011). *The filter bubble: What the Internet is hiding from you.* penguin UK.

Pellert, Max, Hannah Metzler, Michael Matzenberger, and David Garcia (2022). "Validating daily social media macroscopes of emotions". In: *Scientific reports* 12.1, p. 11236.

Pennycook, Gordon and David G Rand (2019). "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning". In: *Cognition* 188, pp. 39–50.

Pesthy, Maria, Matthias Mader, and Harald Schoen (2021). "Why Is the AfD so Successful in Eastern Germany? An Analysis of the Ideational Foundations of the AfD Vote in the 2017 Federal Election". In: *Politische Vierteljahresschrift* 62.1, pp. 69–91.

Prentice, Deborah A and Dale T Miller (1996). "Pluralistic ignorance and the perpetuation of social norms by unwitting actors". In: *Advances in experimental social psychology.* Vol. 28. Elsevier, pp. 161–209.

Rand, David G, Thomas Pfeiffer, Anna Dreber, Rachel W Sheketoff, Nils C Wernerfelt, and Yochai Benkler (2009). "Dynamic remodeling of in-group bias during the 2008 presidential election". In: *Proceedings of the National Academy of Sciences* 106.15, pp. 6187–6191.

Rathje, Steve, Jay J Van Bavel, and Sander Van Der Linden (2021). "Out-group animosity drives engagement on social media". In: *Proceedings of the National Academy of Sciences* 118.26, e2024292118.

Rico, Guillem, Marc Guinjoan, and Eva Anduiza (2017). "The emotional underpinnings of populism: How anger and fear affect populist attitudes". In: *Swiss Political Science Review* 23.4, pp. 444–461.

Romano, Angelo, Matthias Sutter, James H Liu, Toshio Yamagishi, and Daniel Balliet (2021). "National parochialism is ubiquitous across 42 nations around the world". In: *Nature communications* 12.1, p. 4456.

Ross, Lee, David Greene, and Pamela House (1977). "The "false consensus effect": An egocentric bias in social perception and attribution processes". In: *Journal of experimental social psychology* 13.3, pp. 279–301.

Sängerlaub, Alexander, Miriam Meier, and Wolf-Dieter Rühl (2020). "Islamische Grabschparty in Schorndorf!": Die Bundestagswahl 2017 und das Phänomen» Fake News". In: *Fake News und Desinformation.* Nomos Verlagsgesellschaft mbH & Co. KG, pp. 219–244.

Schäfer, Claudia and Andreas Schadauer (2018). "Online fake news, hateful posts against refugees, and a surge in xenophobia and hate crimes in Austria". In: *Refugee News, Refugee Politics.* Routledge, pp. 109–116.

Scheve, Kenneth F and Matthew J Slaughter (2001). "Labor market competition and individual preferences over immigration policy". In: *Review of Economics and Statistics* 83.1, pp. 133–145.

Schumann, Sandy, Diana Boer, Katja Hanke, and James Liu (2021). "Social media use and support for populist radical right parties: assessing exposure and selection effects in a two-wave panel study". In: *Information, communication & society* 24.7, pp. 921–940.

Schwartzstein, Joshua and Adi Sunderam (2021). "Using models to persuade". In: *American Economic Review* 111.1, pp. 276–323.

Serrano, Juan Carlos Medina, Morteza Shahrezaye, Orestis Papakyriakopoulos, and Simon Hegelich (2019). "The rise of Germany's AfD: A social media analysis". In: *Proceedings of the 10th international conference on social media and society*, pp. 214–223.

Shiller, Robert J (2017). "Narrative economics". In: *American economic review* 107.4, pp. 967–1004.

Shiller, Robert J (2020). *Narrative economics: How stories go viral and drive major economic events.* Princeton University Press.

Snowberg, Erik and Leeat Yariv (2021). "Testing the waters: Behavior across participant pools". In: *American Economic Review* 111.2, pp. 687–719.

Stein, Jonas, Marc Keuschnigg, and Arnout van de Rijt (2023). "Network segregation and the propagation of misinformation". In: *Scientific Reports* 13.1, p. 917.

Steinmayr, Andreas (2021). "Contact versus exposure: Refugee presence and voting for the far right". In: *Review of Economics and Statistics* 103.2, pp. 310–327.

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

Tabellini, Guido (2008). "The scope of cooperation: Values and incentives". In: *The Quarterly Journal of Economics* 123.3, pp. 905–950.

Tajfel, Henri, John C Turner, William G Austin, and Stephen Worchel (1979). "An integrative theory of intergroup conflict". In: *Organizational identity: A reader* 56.65, pp. 9780203505984–16.

Tausczik, Yla R and James W Pennebaker (2010). "The psychological meaning of words: LIWC and computerized text analysis methods". In: *Journal of language and social psychology* 29.1, pp. 24–54.

Turner, John C, Michael A Hogg, Penelope J Oakes, Stephen D Reicher, and Margaret S Wetherell (1987). *Rediscovering the social group: A self-categorization theory.* basil Blackwell.

Vosoughi, Soroush, Deb Roy, and Sinan Aral (2018). "The spread of true and false news online". In: *science* 359.6380, pp. 1146–1151.

Weisskircher, Manès (2020). "Neue Wahrheiten von rechts außen? Alternative Nachrichten und der „Rechtspopulismus "in Deutschland". In: *Forschungsjournal Soziale Bewegungen* 33.2, pp. 474–490.

Widmann, Tobias (2021). "How emotional are populists really? Factors explaining emotional appeals in the communication of political parties". In: *Political Psychology* 42.1, pp. 163–181.

Wojcieszak, Magdalena and R Kelly Garrett (2018). "Social identity, selective exposure, and affective polarization: How priming national identity shapes attitudes toward immigrants via news selection". In: *Human communication research* 44.3, pp. 247–273.

Wright, Matthew and Jack Citrin (2011). "Saved by the stars and stripes? Images of protest, salience of threat, and immigration attitudes". In: *American Politics Research* 39.2, pp. 323–343.

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov (2020). "Political effects of the internet and social media". In: *Annual review of economics* 12.1, pp. 415–438.

A Analysing German politicians' tweets

Using Twitter APIs and a home-made algorithm to avoid the Twitter limit for scraping of 3,200 tweets per user, we collected all tweets by the German leaders of the main political parties (during their mandate) and by the official Twitter account of those parties from 2016 to 2021; then, we classified them based on the topic, pre-processed the text, and performed a dictionary-based text analysis to classify tweets by party according to the categories in the Linguistic Inquiry and Word Count (LIWC) and in the Moral Foundation Theory (MFT) dictionaries. The LIWC is a dictionary that aims at connecting important psychosocial constructs and theories with words, phrases, and other linguistic constructions. LIWC has many categories, but for this analysis we only considered the category 'affiliation' related to group identity, which includes 384 words such as 'we', 'our', 'us', 'help', etc. MFT is a psychological framework that suggests human morality is shaped by innate moral foundations: Care/Harm, Fairness/Cheating, Loyalty/Betrayal, Authority/Subversion, Sanctity/Degradation, and Liberty/Oppression (e.g. Haidt 2012). Its related dictionary includes words able to characterize those foundations, and we specifically consider the Loyalty foundation as better classifying closeness to the in-group. Of these dictionaries, we used the German version as provided to us by T. Meier et al. (2019) and by Bos and Minihold (2022). To classify by topic, we considered the Manifesto Project, which manually codes sentences of political parties' Manifestos by topic (Lehmann et al., 2022). We specifically considered words in sentences related to migration, cleaned this list of the words of common use and used the remaining words to identify tweets related to migration. Finally, to apply the psychological dictionaries, we pre-processed the text of the tweets with the usual natural language processing procedure of tokenizing, removing 1-character words, stemming the words, and removing stopwords. We ultimately associated tweets to categories based on a correspondence between the words in the tweets and in the dictionaries. The full tables with the results of this classification are available from the authors on request.

B The role of individual characteristics

All individual characteristics played a very limited role in the choice of the specific individual private preference. We report marginal effects from the multinomial logit on private preferences in the first four columns of Table B1. Being female decreased the probability of choosing A/I and being protestant increased it while high income individuals tended to choose H/O over H/R. All coefficients related to age, immigration status, education, employment status and occupation were statistically not significant or only weakly so. In the last two columns of Table B1 we perform a logit regression over the political position in the chosen narrative (in acceptance vs hostile to migration). When we do not include the political party variable, we find that adults between 30 and 65 are less likely to be in favor of migration, while the opposite holds for more educated individuals and for protestant compared to individuals of other or no religion. When we do include the party affiliation variables, all coefficients, except for Protestant religious denomination, lose their statistical significance, meaning that these characteristics translate well into support for political parties. We also looked at individuals characteristics by political orientation by interacting the individual characteristics with the party groups over the probability of choosing each narrative in OLS regressions. Some of the interaction coefficients were significant and in directions that were consistent with the expected popularity of narratives for specific subgroups of people. However, we did not find any systematic pattern that we could relate to existing economic and cultural explanations of hostility towards migration.

Table B1: **Private preferences by party group and individual characteristics**

	Private preference						Acceptance	
	H/I	H/O	H/R	A/I	A/O	A/R		
Party group								
left	-0.031	0.018	0.013	0.010	-0.006	-0.003		0.014
	(0.029)	(0.028)	(0.030)	(0.035)	(0.041)	(0.037)		(0.197)
right	0.019	0.021	0.123***	-0.023	-0.093*	-0.047		-0.672***
	(0.032)	(0.030)	(0.036)	(0.037)	(0.042)	(0.037)		(0.205)
far right	0.138***	0.179***	0.210***	-0.136***	-0.230***	-0.162***		-2.623***
	(0.040)	(0.039)	(0.041)	(0.033)	(0.039)	(0.034)		(0.271)
Age								
between 30 and 65	0.013	-0.003	0.051	-0.012	-0.030	-0.018	-0.429**	-0.278
	(0.031)	(0.030)	(0.035)	(0.032)	(0.035)	(0.030)	(0.181)	(0.196)
above 65	0.005	-0.041	-0.004	-0.018	0.005	0.053	0.072	0.200
	(0.038)	(0.037)	(0.042)	(0.039)	(0.046)	(0.041)	(0.228)	(0.249)
Female	0.002	0.003	0.013	-0.054***	0.033	0.003	0.095	-0.087
	(0.021)	(0.020)	(0.024)	(0.021)	(0.024)	(0.021)	(0.123)	(0.134)
Born in Germany								
immigrant	-0.040	0.053	-0.056	0.106	-0.020	-0.043	-0.028	0.139
	(0.050)	(0.056)	(0.057)	(0.072)	(0.062)	(0.051)	(0.328)	(0.357)
Education								
high school	0.002	0.004	-0.061	0.001	0.011	0.044	0.477***	0.270
	(0.025)	(0.025)	(0.030)	(0.025)	(0.030)	(0.027)	(0.151)	(0.165)
university or higher	0.016	0.010	-0.066	-0.002	0.000	0.042	0.334*	0.192
	(0.031)	(0.030)	(0.035)	(0.030)	(0.036)	(0.032)	(0.181)	(0.197)
advanced vocational	0.027	0.006	-0.039	0.055	-0.049	0.001	0.106	0.041
	(0.029)	(0.029)	(0.034)	(0.031)	(0.033)	(0.029)	(0.173)	(0.189)
High income								
between 1500 and 3200 euros	-0.022	0.025	-0.036	0.037	0.010	-0.014	0.035	0.182
	(0.026)	(0.024)	(0.031)	(0.023)	(0.028)	(0.027)	(0.149)	(0.162)
above 3200 euros	-0.057	0.102**	-0.107**	-0.020	0.071	0.010	0.188	0.295
	(0.036)	(0.041)	(0.041)	(0.033)	(0.046)	(0.042)	(0.228)	(0.246)
Employment status								
retired	0.018	-0.063*	0.034	0.019	0.034	-0.042	0.196	0.033
	(0.031)	(0.029)	(0.036)	(0.030)	(0.035)	(0.029)	(0.178)	(0.194)
unemployed, student and other	0.007	-0.059*	0.022	0.013	0.044	-0.027	0.147	0.156
	(0.030)	(0.027)	(0.035)	(0.031)	(0.036)	(0.032)	(0.180)	(0.195)
Occupation								
low-skilled white collars	0.039	-0.021	-0.037	0.036	-0.011	-0.007	0.062	0.083
	(0.024)	(0.023)	(0.027)	(0.024)	(0.027)	(0.025)	(0.141)	(0.152)
blue collars	0.014	0.023	-0.037	0.036	-0.023	-0.013	-0.066	0.022
	(0.028)	(0.030)	(0.032)	(0.030)	(0.033)	(0.031)	(0.172)	(0.188)
Religion								
protestant	-0.028	0.021	-0.054	0.070**	-0.035	0.026	0.419***	0.298*
	(0.024)	(0.025)	(0.028)	(0.026)	(0.027)	(0.027)	(0.148)	(0.161)
catholic	0.001	0.033	-0.041	0.039	0.008	-0.039	0.038	0.043
	(0.026)	(0.025)	(0.029)	(0.026)	(0.031)	(0.025)	(0.151)	(0.165)
other	-0.046	0.091	-0.054	-0.038	0.013	0.034	0.203	0.020
	(0.043)	(0.051)	(0.051)	(0.034)	(0.052)	(0.050)	(0.267)	(0.282)

*** p<.01, ** p<.05, * p<.1

For private preferences, marginal effects from a multinomial logit regression over the 1st ranked narratives. For Acceptance, logit regressions. The number of observations is 1226 and the R-squared 0.08. The baseline categories are under 35 for Age, natives for Born in Germany, vocational training for Education, employed individuals for Employment status, managers and professional for Occupation, not belonging to any religion for Religion, and the far left for Party. Standard errors are corrected with the Bonferroni correction to account for multiple hypotheses testing.

As for the choice of endorsing a narrative, we find in regressions that participants with higher income, older than 65, and men were more likely to publicly endorse a narrative (Table B2).

Table B2: **The probability of endorsing a narrative by private preferences and party groups**

	Endorser		
	(1)	(2)	(3)
Acceptance narrative	0.123***	0.127***	0.204**
	(0.028)	(0.035)	(0.082)
Party group			
left		-0.025	0.012
		(0.045)	(0.075)
right		-0.062	-0.022
		(0.048)	(0.075)
far right		0.078	0.152**
		(0.054)	(0.074)
Acceptance narr. # Party group			
Yes # left			-0.060
			(0.093)
Yes # right			-0.060
			(0.097)
Yes # far right			-0.320**
			(0.130)
Acceptance narrative in public			
Yes		0.040	0.036
		(0.035)	(0.035)
Age			
between 30 and 65	0.030	0.028	0.023
	(0.043)	(0.044)	(0.043)
above 65	0.125**	0.125**	0.119**
	(0.054)	(0.055)	(0.055)
Female	-0.093***	-0.083***	-0.080***
	(0.029)	(0.030)	(0.030)
Born in Germany			
immigrant	-0.021	-0.020	-0.028
	(0.079)	(0.079)	(0.079)
Education			
high school	-0.002	0.005	-0.001
	(0.036)	(0.036)	(0.036)
university or higher	0.027	0.037	0.030
	(0.044)	(0.044)	(0.044)
advanced vocational	0.053	0.052	0.046
	(0.041)	(0.041)	(0.041)
High income			
between 1500 and 3200 euros	0.046	0.048	0.049
	(0.036)	(0.036)	(0.036)
above 3200 euros	0.118**	0.123**	0.124**
	(0.054)	(0.055)	(0.055)
Employment status			
retired	-0.028	-0.026	-0.029
	(0.043)	(0.043)	(0.043)
unemployed, student and other	0.033	0.028	0.028
	(0.043)	(0.043)	(0.043)
Occupation			
low-skilled white collars	0.027	0.026	0.029
	(0.034)	(0.034)	(0.034)
blue collars	0.028	0.024	0.024
	(0.041)	(0.041)	(0.041)
Religion			
protestant	0.041	0.047	0.050
	(0.035)	(0.035)	(0.035)
catholic	-0.015	-0.004	-0.002
	(0.036)	(0.036)	(0.036)
other	0.100	0.114*	0.114*
	(0.064)	(0.064)	(0.064)
R-squared	0.044	0.052	0.058
Number of observations	1226	1226	1226

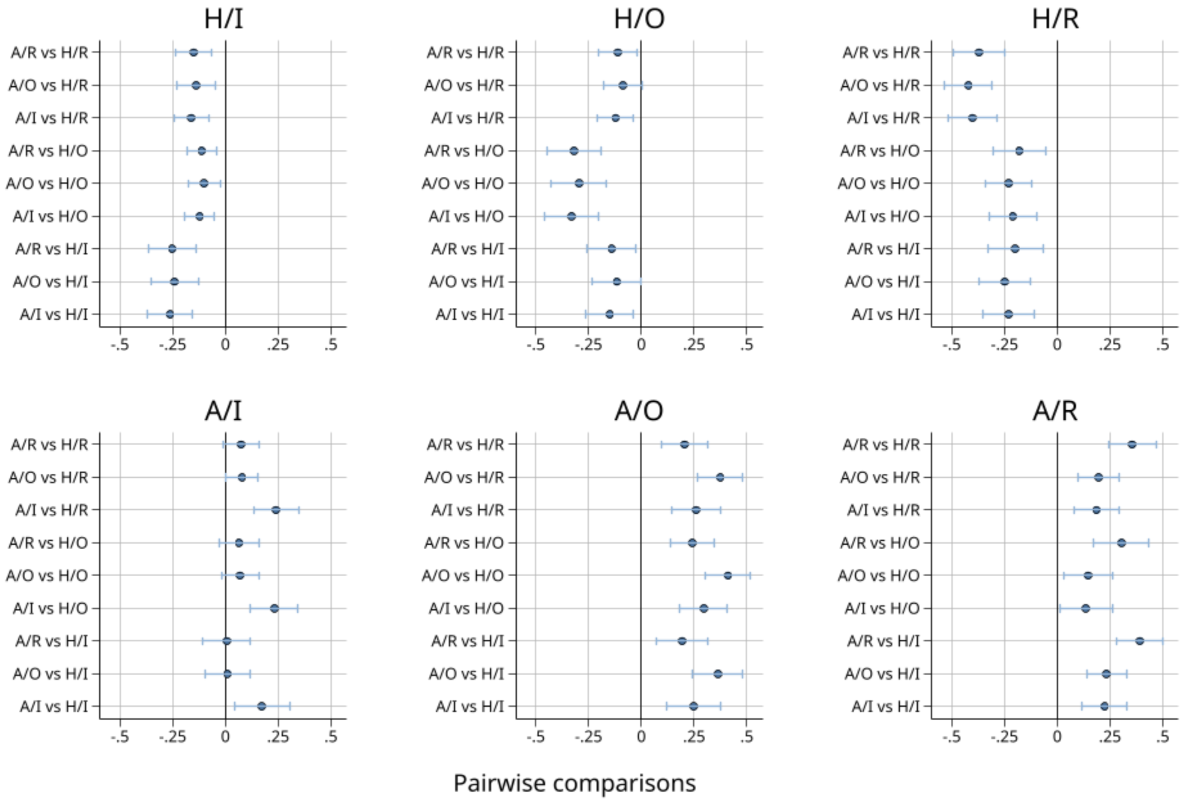
*** p<.01, ** p<.05, * p<.1

OLS regressions on the choice of becoming an endorser. The baseline categories are under 35 for Age, natives for Born in Germany, vocational training for Education, employed individuals for Employment status, managers and professional for Occupation, not belonging to any religion for Religion, and the far-left for Party.

C Changes in preferences for the endorsers

To provide some additional statistical evidence on Result 5, we ran a multinomial logit regression on public preferences by the privates ones and by individual characteristics for the subsample of endorsers[36]. The marginal effects from this regression can be found in Table C2, while we show in Figure C1 pairwise comparisons tests of coefficients associated with the private preferences. In each panel, the probability of publicly endorsing a narrative was between 10% and 40% higher if the private preference was supporting the same policy position and the comparisons were stronger in magnitude for the narrative with the same focus.

Figure C1: Change in preferences between private and public for endorsers



Pairwise comparisons between coefficients from a multinomial logit on public preferences for the endorsers. Errors are corrected with the Bonferroni correction to account for multiple hypothesis testing. The regression table can be found in Table C2 in Appendix C. Controls include age group (young, middle age, old), female, immigrant status, education (4 groups), income (3 groups), employment status (3 groups), occupation (3 groups), religion (4 groups), prejudices on migrants and contact with immigrants. Reported confidence intervals are at 95% level.

[36]Since we did not find any relevant pattern by political orientation in the raw data, and given that the sample size is here smaller than in the main survey, in this regression we did not interact private preferences with party groups to gain more power for our statistical analysis .

Table C2: **Change from private to public preference by the endorsers**

	Public preference					
	H/I	H/O	H/R	A/I	A/O	A/R
Private preference						
H/O	-0.140	0.180	-0.020	-0.059	-0.048	0.088
	(0.065)	(0.081)	(0.078)	(0.056)	(0.048)	(0.050)
H/R	-0.102	-0.029	0.172	-0.067	-0.011	0.037
	(0.068)	(0.066)	(0.078)	(0.052)	(0.050)	(0.039)
A/I	-0.263***	-0.150**	-0.230***	0.172**	0.248***	0.223***
	(0.055)	(0.057)	(0.063)	(0.067)	(0.065)	(0.053)
A/O	-0.241***	-0.116	-0.250***	0.010	0.363***	0.233***
	(0.057)	(0.059)	(0.061)	(0.056)	(0.061)	(0.048)
A/R	-0.252***	-0.140*	-0.199**	0.006	0.194***	0.392***
	(0.056)	(0.059)	(0.067)	(0.058)	(0.061)	(0.056)
Contact with imm	0.011	-0.008	0.006	-0.012	-0.005	0.008
	(0.006)	(0.008)	(0.009)	(0.009)	(0.011)	(0.010)
Prejudice against migrants	0.009	-0.047	0.074	-0.047	0.029	-0.017
	(0.038)	(0.054)	(0.056)	(0.065)	(0.079)	(0.075)
Age						
between 30 and 65	-0.061	0.014	0.061	0.012	0.033	-0.059
	(0.049)	(0.049)	(0.051)	(0.045)	(0.054)	(0.057)
above 65	-0.077	0.000	0.080	0.000	0.016	-0.020
	(0.054)	(0.056)	(0.061)	(0.054)	(0.065)	(0.068)
Female	-0.044*	-0.041	0.027	0.011	0.058	-0.011
	(0.024)	(0.032)	(0.034)	(0.030)	(0.036)	(0.035)
Born in Germany						
immigrant	0.023	-0.062	0.185	-0.038	-0.120	0.013
	(0.068)	(0.066)	(0.101)	(0.069)	(0.085)	(0.101)
Education						
high school	0.009	-0.043	-0.002	0.004	-0.028	0.060
	(0.027)	(0.036)	(0.039)	(0.039)	(0.047)	(0.045)
university or higher	0.008	-0.007	0.055	0.055	-0.052	-0.059
	(0.031)	(0.044)	(0.048)	(0.051)	(0.054)	(0.045)
advanced vocational	0.050	-0.008	0.032	-0.058	-0.109*	0.092
	(0.038)	(0.042)	(0.045)	(0.036)	(0.049)	(0.053)
Income						
between 1500 and 3200 euros	0.016	-0.029	0.040	-0.051	0.041	-0.017
	(0.028)	(0.040)	(0.039)	(0.043)	(0.044)	(0.045)
above 3200 euros	0.009	-0.061	0.045	-0.093	0.020	0.081
	(0.043)	(0.052)	(0.060)	(0.054)	(0.064)	(0.066)
Employment status						
retired	-0.057**	0.038	-0.022	0.028	-0.010	0.023
	(0.024)	(0.046)	(0.044)	(0.048)	(0.052)	(0.051)
unemployed, student and other	0.029	-0.056	-0.009	-0.024	0.107	-0.047
	(0.045)	(0.037)	(0.054)	(0.040)	(0.057)	(0.050)
Occupation						
low-skilled white collars	-0.003	0.009	-0.040	0.036	0.018	-0.020
	(0.027)	(0.035)	(0.037)	(0.035)	(0.042)	(0.042)
blue collars	0.042	0.027	0.009	0.067	0.005	-0.151***
	(0.038)	(0.044)	(0.051)	(0.046)	(0.052)	(0.043)
Religion						
protestant	-0.000	0.033	-0.048	0.013	0.023	-0.021
	(0.029)	(0.038)	(0.037)	(0.035)	(0.043)	(0.042)
catholic	-0.016	-0.027	0.032	0.024	0.048	-0.061
	(0.026)	(0.034)	(0.041)	(0.040)	(0.047)	(0.043)
other	0.038	-0.021	0.069	0.026	-0.044	-0.069
	(0.055)	(0.057)	(0.078)	(0.065)	(0.072)	(0.069)

*** p<.01, ** p<.05, * p<.1 Marginal effects from a multinomial logit regression over the publicly endorsed narratives for the endorsers. The number of observations is 522 and the R-squared 0.23. The baseline categories are under 30 for Age, natives for Born in Germany, vocational training for Education, employed individuals for Employment status, income below 1500 euros for Income, managers and professional for Occupation, not belonging to any religion for Religion. Prejudice on migrants is the difference between the estimated percentage of illegal migrant workers over the total working population and the estimated percentage of illegal native workers over the total working population; the actual value, based on estimates, ranges between 10% and 30%. Standard errors are corrected with the Bonferroni correction to account for multiple hypotheses testing.

Table C1: **Table with the frequency of public and private preferences**

	Public preference						
	H/I	H/O	H/R	A/I	A/O	A/R	Total
far right							
H/I	5	3	12	0	1	0	21
H/O	3	19	8	0	2	2	34
H/R	7	7	22	0	0	2	38
A/I	0	0	0	1	1	1	3
A/O	0	0	1	0	2	0	3
A/R	0	0	2	0	0	1	3
Total	15	29	45	1	6	6	102
right-wing							
H/I	3	4	5	2	0	1	15
H/O	3	1	5	0	1	2	12
H/R	2	3	10	2	4	2	23
A/I	0	2	4	4	10	4	24
A/O	0	3	2	3	13	8	29
A/R	1	2	1	3	12	6	25
Total	9	15	27	14	40	23	128
left-wing							
H/I	3	1	1	3	3	1	12
H/O	4	6	7	3	0	3	23
H/R	2	4	7	3	2	1	19
A/I	0	2	2	16	12	13	45
A/O	1	4	1	10	30	18	64
A/R	0	2	2	6	14	30	54
Total	10	19	20	41	61	66	217
far left							
H/I	3	2	0	1	1	0	7
H/O	2	0	2	1	0	1	6
H/R	0	1	2	0	1	1	5
A/I	0	0	0	7	7	4	18
A/O	1	1	0	3	11	7	23
A/R	0	0	2	3	2	9	16
Total	6	4	6	15	22	22	75

Matrixes with Private and Public preferences for the subsample of endorsers by party. In the rows there are the Private preferences, in columns the Public ones.

D Interpreting the change in preferences

We explored various mechanisms to interpret the changes between private and static public preferences. We considered the four following properties of a narrative, to which we refer as "attributes": (a) Distinctiveness (see sections 2 and 5.2.4 in the main paper); Beliefs over the most popular narrative (b) within one's own political group; and (c) within the other political groups; (d) Political correctness. (b) and (c) are the two components of (a). All these beliefs were elicited in Survey 2. In section 5.2.4, we also discuss the potential role of pluralistic ignorance, concluding that some of its constitutive elements do not seem to occur in our survey.

For our analyses, we consider the following econometric model:

$$\overline{S_A}\left(v_i^*\right) = \alpha + \beta PUBLIC_i + \gamma POL_i + \delta PUBLIC\_X\_POL_i + \eta D_i + \varepsilon_i \tag{10}$$

where $\overline{S_A}\left(v_i^*\right)$ is the average score for the attribute A for a narrative v_i^* preferred by agent i. A is one of the four attributes we considered, that is, distinctiveness, beliefs over popularity within one's own political group, beliefs over popularity in others' political groups, and political correctness. The average is obtained from Survey 2 and refers to agent i's relevant group, that is one's own group for distinctiveness, popularity in own party and political correctness, and the other group for popularity in other party group. We standardized the scores S_A separately for each party group to make them comparable. For each endorser in our sample, v_i^* includes both decisions in the private and the static public setting. For non-endorsers, v_i^* only includes the decision in the private setting. α is a constant. $PUBLIC_i$ is the key variable in this analysis, as it is a dummy variable that identifies whether the decision was made in the private or public static setting. POL_i is i's political group. $PUBLIC\_X\_POL_i$ is an interaction term between the latter two variables. D_i is a vector of demographic characteristics (not reported). Finally, ε_i is a normally distributed error term.

All mechanisms had some significant role in explaining the change from the private to the static public preferences (see Table D1). The narratives chosen in static public settings had a significantly higher score than those chosen in the private setting for all the four attributes. This means that, in the public setting, endorsers chose narratives that were (a) more distinctive, (b) believed to be more popular in their group, (c) believed to be less popular in other party groups, and (d) more politically correct, compared to the narratives chosen in the private setting. Given the high correlation between these variables, it is not surprising that the coefficient for $PUBLIC_i$ is statistically significantly different from 0 for all four attributes, albeit only weakly so for beliefs over popularity in other groups. Overall, the coefficient for $PUBLIC_i$ was larger, in absolute value, for distinctiveness than the other three attributes. It is also noteworthy that the role of the four attributes seemed to be different for far-right supporters compared to all others in the private setting. Moreover, far-righters attached significantly lesser importance to politically correctness in the static public than the private setting, while this was not the case for others (Table D1, column 8).

Apart from this evidence in a regression framework, there are other (minor) differences between the private and the static public setting that are related to specific narratives and that are consistent with an explanation to the change in preferences in terms of distinctiveness. These differences are the lower distinctiveness of H/I compared to H/O and H/R for far-right supporters and the higher one of A/R compared to A/O for far-left supporters (see Figure 4). In Section 5.2.1, we exactly find that far-right supporters decrease their preferences for H/I in public compared to the private setting in favour of

H/R, and also that far-left individuals tend to endorse more A/R in the public vs the private setting.

Table D1: **Social image concerns for the endorsers and alternative mechanisms**

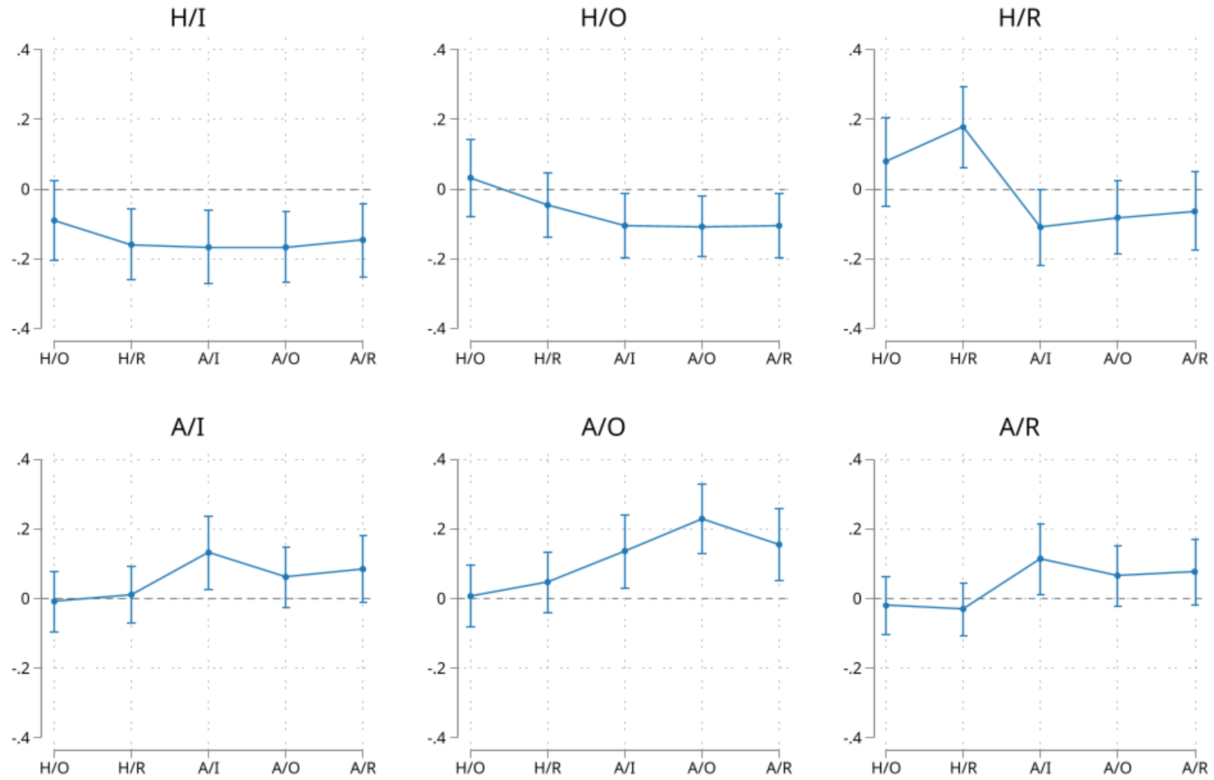
	Distinctivness		Beliefs on most popular in my party		Beliefs on most popular in other parties		Political correctness	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Static public preference	0.189***	0.250**	0.162***	0.177	-0.085*	-0.168	0.146***	0.234**
	(0.047)	(0.124)	(0.048)	(0.125)	(0.047)	(0.122)	(0.043)	(0.111)
Party group								
left		0.044		0.041		-0.058		0.041
		(0.081)		(0.082)		(0.080)		(0.073)
right		-0.223***		-0.122		0.218**		-0.217***
		(0.086)		(0.087)		(0.085)		(0.077)
far right		0.453***		0.412***		-0.523***		0.421***
		(0.092)		(0.094)		(0.091)		(0.082)
Party group # Static publ. pref.								
left		-0.081		-0.023		0.043		-0.051
		(0.143)		(0.145)		(0.141)		(0.128)
right		0.030		0.018		0.110		-0.068
		(0.154)		(0.156)		(0.152)		(0.138)
far right		-0.238		-0.100		0.266*		-0.320**
		(0.162)		(0.165)		(0.160)		(0.145)
R-squared	0.027	0.074	0.024	0.056	0.015	0.079	0.028	0.081
Number of observations	1748	1748	1748	1748	1748	1748	1748	1748

*** p<.01, ** p<.05, * p<.1

OLS regressions on the distinctiveness, the popularity within own party group, the popularity within other parties' groups, and the political correctness of the private (overall sample) and public (subsample of endorsers) preferences. Observations associated to endorsers are duplicated to account for their public preferences (see text for more details). Static public preference is a dummy variable taking the value of 1 if the preference is a public preference by an endorser. The control variables are age group (young, middle age, old), female, immigrant status, education (4 groups), income (3 groups), employment status (3 groups), occupation (3 groups), religion (4 groups).

E The role of beliefs: an alternative explanation

Figure E1: Beliefs on narratives



Multinomial logit on beliefs on the private preferences of others. Baseline is H/I. Standard errors are adjusted with the Bonferroni correction. Controls include age (3 categories), immigration status, education (4 categories), income (4 categories), employment status (3 categories), occupation (3 categories), and religion (4 categories).

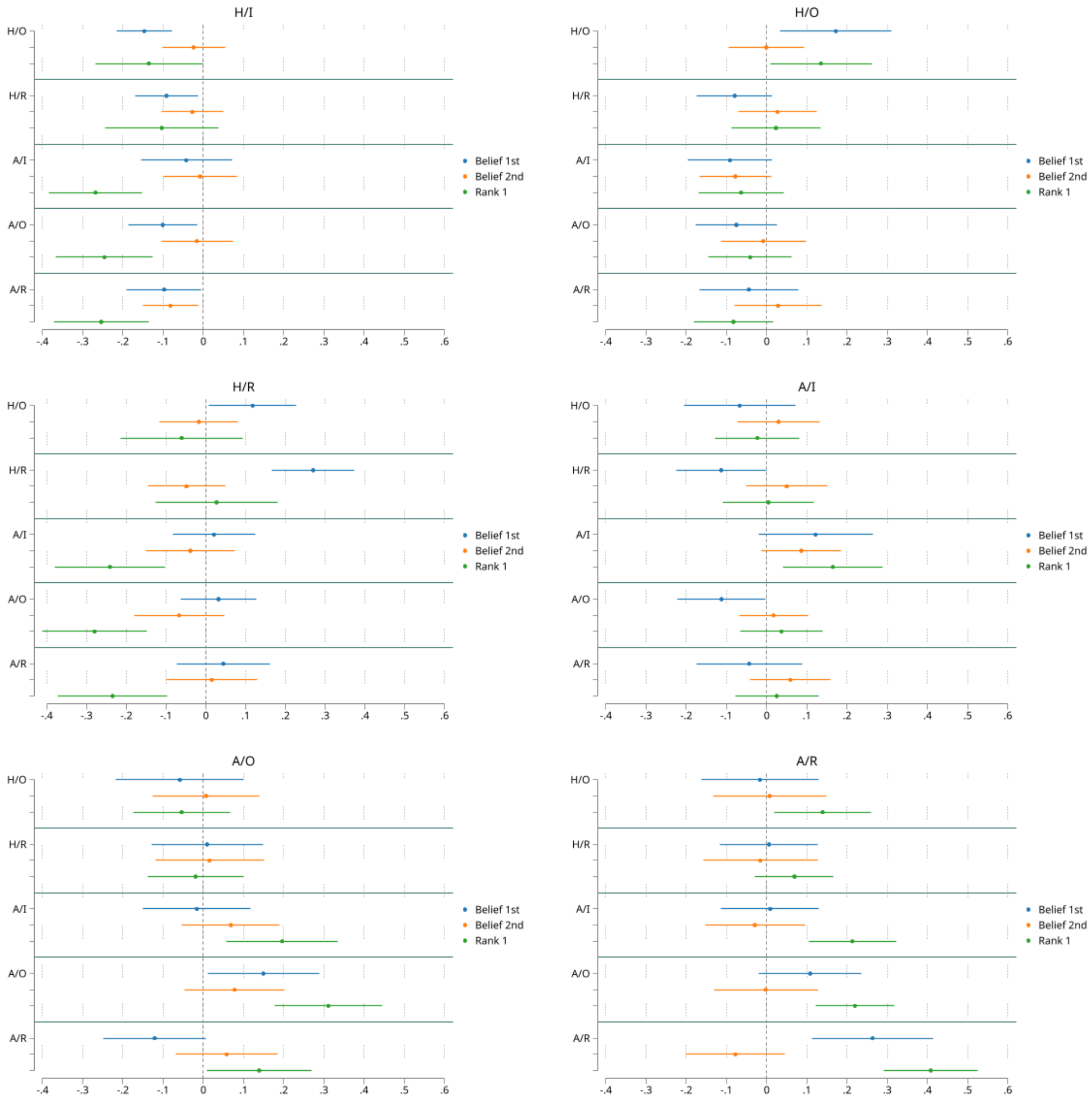
To explain the change between private and public preferences, we also took into account the beliefs about the most preferred narrative (see questions 11 and 12 in Survey 1 in Appendix K). This provided some evidence of alternative motives, such as conformism, when deciding which narrative to publicly endorse. First, we examined whether participants overestimated the number of people at the "society level" agreeing with their own private preferences. Given that our sample is representative of the German population, we estimate preferences at the "society level" with the aggregate results from our survey. Subsequently, we examined how beliefs entered the decision process on public preferences.

To assess whether the individuals overestimated the percentage of others agreeing with the same narrative as themselves, we performed a multinomial logit regression over beliefs on the most preferred narrative. These beliefs were regressed over the preferred narratives while controlling for individual characteristics. Based on that regression, Figure E1 shows the results of the marginal effects of each preferred narrative—with H/I as the baseline category—on each possible belief.

The coefficients associated with pro-immigration narratives were always negative in the panels on anti-immigration narratives' beliefs and positive in the panels on pro-immigration narratives' beliefs. Hence, the popularity of narratives against migration was overestimated by individuals who preferred anti-immigration narratives and vice versa. To be more specific, let us consider each panel in Figure E1 separately. On beliefs on H/I, all coefficients were negative, and all except H/O were significantly so at the 5% level, suggesting that H/I, compared with the other ones, was indeed overestimated by individuals who found it to be the one that they agreed with the most. The same pattern held for the panels on H/R and A/I: the coefficients were significantly different from zero only for the corresponding narrative. Out-group narratives, both hostile and accepting, and A/R seemed to display a different pattern. In these cases, it did not seem to make a difference in the participant's beliefs if they expressed a specific preference for that narrative, as long as the political position expressed by the narrative—anti-or proimmigration—was the same.

We restricted the sample to those who were willing to share a narrative publicly. In this way, we could disentangle the influence of beliefs from that of preferences on the choice of endorsing one specific narrative rather than another. We ran a multinomial logit regression where the dependent variable was the endorsed narrative and the three independent variables of interest were the beliefs on each narrative and the highest ranked narratives. Then, the average marginal effects were computed with the Bonferroni correction to account for multiple hypothesis testing, which have been reported in Figure E2 as graphs.

Figure E2: Determinants of endorsing a narratives



Average marginal effects from multinomial logit regression on the endorsed narrative. The sample for this regression is the subgroup of participants who are willing to publicly endorse a narrative. Baseline category is H/I. Standard errors are adjusted with the Bonferroni correction. Controls include age (3 categories), immigration status, education (4 categories), income (4 categories), employment status (3 categories), occupation (3 categories), and religion (4 categories).

First, the beliefs on the second preferred narrative were never significant, with just one exception. This type of beliefs was inserted in the regression as a placebo. Then, to understand how much beliefs weighted compared to preferences on the choice of which narrative to support, we considered the coefficients of beliefs on the most popular narrative vs. the individual preference. Both seemed to matter because the coefficients of the beliefs and preferences corresponding to each narrative were significantly different from zero for most of the narratives. The weight varied depending on the narrative under consideration. For H/I, both had the same weight: not having an individual preference for proimmigration narratives seemed to matter more than specifically liking the H/I one. In the H/O panel, those coefficients associated with the corresponding belief and preference were the only significant ones, both with the same order of magnitude. To endorse H/R, two things mattered: having the corresponding belief and disagreeing with proimmigration narratives. For A/I, only the corresponding preference mattered. A/O and A/R displayed a pattern in which what mattered was agreeing with a proimmigration narrative and believing that the correspondent one was the one individuals agreed with the most.

F Narratives on vaccination

H/I: From June to August, almost 20 million people got at least a dose of vaccine. The way unvaccinated people question scientific results and still experimental vaccines is a rational way of reacting to the pandemic. Furthermore, unvaccinated people are contributing to reach the point to which we are all immune against Covid-19 fast without at the same time putting themselves at a great risk.

H/O: From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated people mostly show blind faith in science and in still experimental vaccines, which is an irrational way of reacting to the pandemic. Furthermore, vaccinated people are delaying the point in time to which we are all immune against Covid-19 without at the same time really protecting themselves.

H/R: From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated and unvaccinated people show opposing levels of faith in science and in still experimental vaccines and have conflicting views on the rational way of reacting to the pandemic. Every attempt to make vaccinated and unvaccinated cooperate to reach the point we defeat Covid-19 and protect ourselves will always prove to be too costly and should not be pursued.

A/I: From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated people mostly show faith in science and in still experimental vaccines, which is a rational way of reacting to the pandemic. Furthermore, vaccinated people are contributing to reach the point to which we defeat Covid-19 while at the same time protecting themselves.

A/O: From June to August, almost 20 million people got at least a dose of vaccine. The way unvaccinated people question scientific results and still experimental vaccines is just an irrational way of reacting to the pandemic. Furthermore, unvaccinated people are delaying the point in time to which we defeat Covid-19 while at the same time putting themselves at a great risk.

A/R: From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated and unvaccinated people show opposing levels of faith in science and in still experimental vaccines and have conflicting views on the rational way of reacting to the pandemic. Every attempt to make vaccinated and unvaccinated cooperate to reach the point we defeat Covid-19 and protect ourselves will always prove to be profitable and should be pursued.

G Details on the Twitter experiment

G.1 Revised narratives

We had to adapt our narratives for Twitter by slightly reducing the number of words. The modifications did not impinge on the fundamentals of our narratives. Each narrative still encompassed the fact, policy position, and emphasis both on the cultural and economic domain.

H/O: Since 2015, approximately 10 million migrants have come to Germany. The unacceptable values and practices of many migrants are incompatible with our cultural lives. In addition, many migrants have work attitudes threatening to permanently harm their economy.

H/R: Since 2015, approximately 10 million migrants have come to Germany. Migrants have different values and practices, as well as professional skills. The integration of migrants is an expensive investment, and the costs will never be recouped in the future.

A/O: Since 2015, approximately 10 million migrants have come to Germany. The values and practices of many migrants can enrich our cultural lives. In addition, migrants bring professional qualifications that are necessary for our economy.

A/R: Since 2015, about 10 million migrants have come to Germany. Migrants have different values and practices, as well as professional skills. The integration of migrants is a worthwhile investment, and the costs will be more than offset in the future.

G.2 Sample

To build our sample, we performed several operations to obtain stratification by party and ensure we were not collecting undecided people or bots:

1. We collected the retweets of the main German parties (AfD, CDU, CSU, FDP, Greens, SPD, Die Linke) from the 17th of September backwards. On the first time we performed this operation, we collected a total number of 158,286 retweets, which provided us with a list of 39,152 unique Twitter users. Then, we did this five more times with more limited numbers to enlarge the sample, taking caution that there was no overlap in the users.

2. For each of these users, we collected information (photo profile, description, location, etc.) and a list of their 100 more recent tweets.

3. The users' timelines were used to spot those who retweeted from parties of different political orientations among far-left, left-wing, right-wing, and far-right ones. We dropped those users for which we found some overlap in the retweeting. We only found 0.009% of the users engaging in this type of online behavior.

4. We further dropped the following types of users: 1) those who signed up in 2020 or afterwards because of their higher chance of being bots or fake accounts, 2) the users who represented national or local branches of the parties, and 3) those users who tweeted less than 100 tweets in total. We strove to form a sample roughly balanced in the size of the four political orientations of approximately 5,000 users for each group (AfD, SPD+Greens, FDP+CDU+CSU, Die Linke). Within each group, we sought to balance users according to the parties that make up a group based on the last poll available on the 17th of Setpember.[37] For calculating the ratio between CDU and CSU, we used the ratio in terms of the total number of votes in the German Federal elections of 2017.

G.3 Procedures

Our research protocol involved performing several tasks every day before starting the experiment on September 22nd. Some of these tasks were continued on each day of the experiment.

[37] This was an INCA poll with data from the 13th to 15th of September.

Firstly, starting in July and throughout the course of the experiment, we randomly posted articles with politically neutral content taken from online newspapers and magazines. The tweets were about sports, art, history, culture, and business events. These tweets almost never received any interaction and only worked to keep our Thomas Meier "real". This was mostly needed because, during our pilots, we realized that, with no activity, our experimental profiles following other users did not trigger any notification to the users who were followed.

Secondly, we built our sample (see the previous subsection for details). This was done six times in total because our first recruitment attempt was not enough to complete the experiment.

Thirdly, three days before the starting date of the experiment, we bought fake followers on https://famousfollowe We started on the planned date with approximately 250 followers per Thomas: this was needed because otherwise—as we realized during our pilots— no notification of our following another user would be produced for the followed user. Of course, these fake users were not considered part of the sample, so they were left out of our results completely.

Fourthly, each day, we started following approximately 220 users for each of our "Thomas-accounts", stratified with our sample by political orientation (i.e., 55 per party group). Because of Twitter's technical constraints, we could not follow more than 1,000 users per day with our Twitter API account. This procedure was always enacted around 11 a.m. to 12 a.m., a time in which we assumed that Twitter users were most active (e.g., https://statusbrew.com/insights/best-times-to-post-on-social-media/). It took us around 20 minutes to complete this operation with Twitter APIs. Every operation with its exit code was recorded through APIs. On the 96th day of the experiment, we reached a limit of 5,000 users, the maximum for Twitter based on an undisclosed following/followed ratio.

H Emotions and political correctness

Figure H1 represents the average value of each emotion by narrative and by party group (answers for this analysis come from question 8 in Survey 2, shown in Appendix L). To test differences between emotions or across narratives, we use Wilcoxon signrank tests over our sample. On average, anger was the most common emotion associated with the narratives at the individual level, scoring higher than fear, with happiness scoring the lowest ($p < 0.0001$ in all pairwise comparisons using the whole sample). This still held when focusing on hostile narratives at the aggregate level ($p < 0.0001$ in all comparisons). Interestingly, fear and anger scored higher than happiness regardless of political orientation ($p < 0.0001$ for each party for each of the two comparisons), anger scores higher than fear in far-left and left-wing voters (far-left: $p = 0.009$, left-wing: $p < 0.0001$) while the opposite held for far-right ones ($p = 0.029$). The difference between anger and fear was not significant for right-wing participants instead ($p = 0.970$). At the aggregate level, acceptance narratives do not seem to trigger a specific emotion ($p > 0.10$ in every comparison). This masks relevant heterogeneity by political orientation. Acceptance narratives triggered more happiness than anger and fear for far-left and left-wing voters (far-left: $p = 0.028$ with anger, $p = 0.115$ with fear, left-wing: $p < 0.0001$ in both comparisons), while the opposite held for far-right voters ($p < 0.0001$ in both comparisons). Furthermore, in far-right voters anger prevailed over fear ($p = 0.0003$), while we failed to find significant differences in the other party groups (far-left: $p = 0.272$, left-wing: $p = 0.985$). In addition, no comparison between emotions on acceptance narratives was significant for right-wing individuals (happiness vs anger: $p = 0.731$, happiness vs fear: $p = 0.664$, fear vs anger: $p = 0.590$). When comparing emotions across narratives, we find that on aggregate hostility narratives are associated with more anger and fear and less happiness

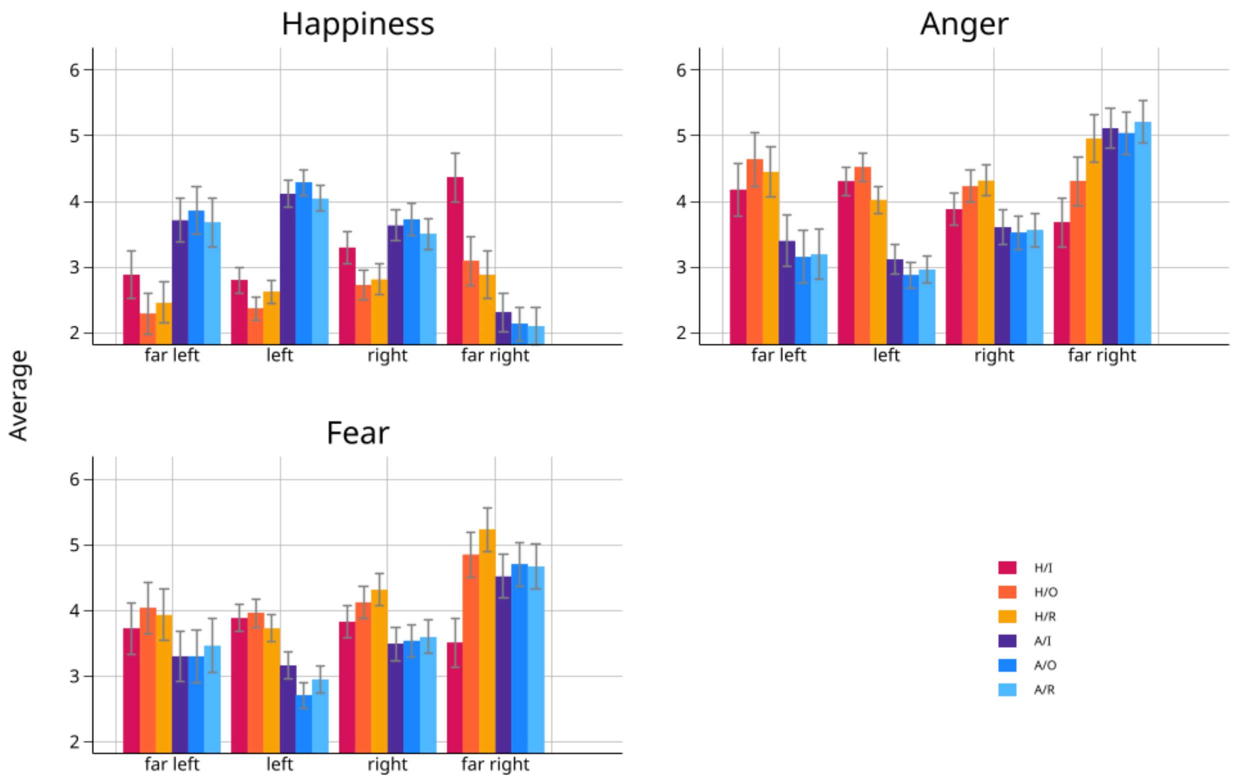


Figure H1: Emotions by party and by narrative

Mean values of emotions by political orientation from Survey 2. The total sample is 771 participants. Reported confidence intervals are at 95% level.

than acceptance ones ($p < 0.0001$ for each emotion). However, this held for far-left, left-wing and far-right voters only (on happiness: $p < 0.0001$ for far-left and left-wing supporters, $p = 0.0002$ for right-wing ones; on anger: $p < 0.0001$ for far-left and left-wing supporters, $p = 0.0007$ for right-wing ones, on fear: $p < 0.0001$ for left-wing and right-wing supporters, $p = 0.002$ for far-left ones). Far-right supporters were happier and less angry with hostility narratives instead ($p < 0.0001$), while we did not find significant differences on fear ($p = 0.784$).

To complete our analysis, we then looked at emotions by the specific focus of the narratives. With Mann-Withney ranksum tests we did not find differences in almost any emotion between A/I, A/O, and A/R (always $p > 0.10$ except for A/O vs A/R on happiness $p = 0.044$). The same between H/O and H/R (happiness: $p = 0.144$, anger: $p = 0.318$, fear: $p = 0.927$). However, H/I was associated with more happiness and less anger and fear compared to H/O and H/R (always $p < 0.0001$ except $p = 0.007$ in comparison with H/R on anger). This result was driven by far-right individuals ($p < 0.0001$ on every comparison except $p = 0.018$ for H/I vs H/O on anger), although most differences were significant also in the other groups[38].

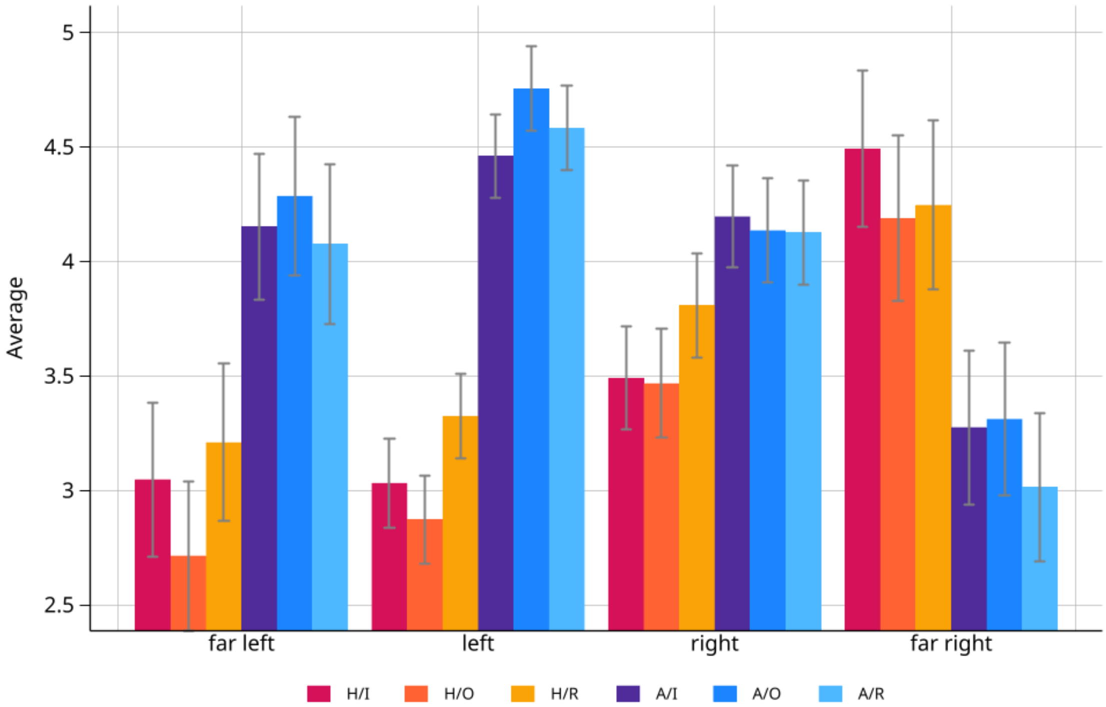


Figure H2: Political correctness by party and by narrative

Mean values of political correctness by political orientation from Survey 2. The total sample is 771 participants. Reported confidence intervals are at 95% level.

Figure H2 shows the average level of political correctness by political orientation and by narrative (always from question 8 in Survey 2, shown in Appendix L). From the Figure, it emerges a stark difference when judging the relative appropriateness of the policy positions (more than of the specific

[38]Compared to H/O, the tests gave the following results. On happiness: $p = 0.018$ for far-left voters, $p = 0.004$ for left-wing ones, $p = 0.0009$ for right-wing ones; on anger: $p = 0.100$ for far-left ones, $p = 0.123$ for left-wing ones, $p = 0.043$ for right-wing ones; on fear: $p = 0.265$ for far-left ones, $p = 0.574$ for left-wing ones, $p = 0.101$ for right-wing ones. Compared to H/R, the tests gave the following results. On happiness: $p = 0.147$ for far-left voters, $p = 0.388$ for left-wing ones, $p = 0.003$ for right-wing ones; on anger: $p = 0.367$ for far-left ones, $p = 0.065$ for left-wing ones, $p = 0.013$ for right-wing ones; on fear: $p = 0.457$ for far-left ones, $p = 0.317$ for left-wing ones, $p = 0.008$ for right-wing ones.

focus) on migration. Therefore, we took the average value of political correctness between H/I, H/O and H/R and between A/I, A/O and A/R for each individual and, for each party group, we used signrank tests to test 1) if the policy position is judged politically correct (by testing differences with the average neutral value of 3.5), 2) differences in political correctness between anti-immigration and pro-immigration narratives. For far-left, left-wing and right-wing voters, pro-immigration narratives were deemed politically correct ($p < 0.0001$). For the far-right, they were not and were even judged to be incorrect ($p = 0.037$). Maybe not surprisingly, we got the opposite evidence for anti-immigration narratives (for far-left, left-wing and far-right supporters, $p < 0.0001$), except for right-wing voters who were more split over the judgment ($p = 0.260$). It is then not surprising that far-left, left-wing and right-wing supporters deemed narratives in acceptance of migration as more politically correct compared to the hostile ones while the opposite held for far-right supporters ($p < 0.0001$ for all parties except $p = 0.002$ for right-wing voters). Interestingly, this also held for the right-wing voters, whom we saw as more split when looking at private preferences in Survey 1. If political correctness was related to social image concerns, then the decision not to publicly endorse anti-immigration narratives for this political orientation may be at least partly explained by these concerns (see also Appendix D).

When we looked at the focus, H/R was deemed as more politically correct than H/O and H/I (by Mann-Withney ranksum tests, $p = 0.0001$ with H/O and $p = 0.039$ with H/I)[39]. These differences were driven by far-left, left-wing and right-wing supporters (far-left: $p = 0.039$ with H/O and $p = 0.568$ with H/I, left-wing: $p = 0.0005$ with H/O and $p = 0.019$ with H/I, right-wing: $p = 0.040$ with H/O and $p = 0.048$ with H/I), while they were not significant when focusing on far-right supporters ($p = 0.819$ with H/O and $p = 0.403$ with H/I). We then failed to find differences between the acceptance narratives (A/I vs A/O: $p = 0.121$, A/I vs A/R: $p = 0.913$, A/O vs A/R: $p = 0.102$).

Table H1: **Preferences and follow-backs by political correctness**

	Chosen narrative		Endorser in public setting		Chosen narrative in public		Follow-back	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Political correctness	0.093***		0.108***		0.134***		0.009***	
	(0.007)		(0.024)		(0.010)		(0.002)	
Party group # Political correctness								
far left		0.101***		0.126***		0.146***		0.009***
		(0.008)		(0.029)		(0.012)		(0.003)
left		0.094***		0.109***		0.135***		0.008***
		(0.007)		(0.025)		(0.010)		(0.002)
right		0.093***		0.099***		0.135***		0.009***
		(0.007)		(0.027)		(0.011)		(0.003)
far right		0.099***		0.110***		0.140***		0.027***
		(0.007)		(0.026)		(0.011)		(0.003)
Number of observations	7356	7356	1226	1226	3132	3132	19989	19989
R-squared	0.024	0.025	0.044	0.048	0.052	0.053	0.009	0.028

*** p<.01, ** p<.05, * p<.1

OLS on the chosen narrative as private preferences in columns (1) and (2), on the choice of becoming an endorser in columns (3) and (4), on the chosen narrative as public preference for the subsample of the endorsers in columns (5) and (6), on the choice of following-back in columns (7) and (8). See text for more details on how political correctness is associated to the narratives. Controls in column (3) and (4) include age (3 categories), immigration status, education (4 categories), income (4 categories), employment status (3 categories), occupation (3 categories), and religion (4 categories). Controls in columns (7) and (8) include the number of users followed, the number of users whom they follow, the number of tweets posted, the year in which users signed up.

[39]Weakly, also H/I was deemed more correct than H/O ($p = 0.065$).

I Beliefs over own and other parties preferred narratives and moral disapproval

In this section, we examine in greater detail beliefs over narratives by political orientation related to own and others' parties. For this purpose, we exploit the multiple elicitation of beliefs in Survey 2. We asked participants to state which narrative was most preferred by people supporting their own party and people supporting other parties (see questions 9-23 in Appendix L). Correct answers were rewarded with 0.50 Euro. To estimate the relevance of false consensus effect, we match participants' beliefs with information on their own most preferred narrative and compute forecast errors.

We find a strong tendency for participants to believe that fellow party members preferred the same narrative as themselves. 54% of the participants made this prediction, with only small fluctuations between political groups (Figure I1, panel a). Consistent with a false consensus effect (Ross, Greene, and House, 1977), however, they were wrong in 59% of the cases. Analyzing forecast errors by political group, we find that right-wingers were the most prone to a false consensus effect, as their prediction that party members follow the same narrative as themselves was wrong in 73% of the cases, followed by left-wingers and far-left voters with error rates of 57% and 55%, respectively. The far-right supporters were those least prone to false consensus, as they are wrong in 43% of the cases. A Kruskall-Wallis test rejects the null that the distribution of forecast errors, conditional on participants predicting that others follow the same narrative as themselves, is the same for the four political groups ($p = 0.0007, n = 417$). In particular, using Kolmogorov-Smirnov tests of equality of distributions, we can reject the two null hypotheses that observations for right-wingers are the same as those for the far-right (p=0.001, n=187) and left-wingers (p=0.088, n=292). The same patterns can also be detected among those who believe that fellow party members do not prefer their same narrative. Even in this subgroup, right-wingers are those most likely to make wrong predictions, with an error rate of 77%, followed by far-left (74%), far-right (71%) and left-wingers (70%). Forecast errors are here much more similar across groups and the null of equality of distributions is not rejected either across all four political groups ($p = 0.61, n = 354$) and in all pairwise comparisons.

When predicting the preferred narrative by other party members, the share of respondents who believe that non-party members prefer the same narrative as themselves drops to 24% , with an error rate of 76%. Since in this case false consensus seems relevant, we analyze the whole sample. Right-wingers are again the group with the highest error rate (84%, very close to pure random guessing), followed by far-left and left with error rates of 83% and 74%, respectively. Even in this case, far-right supporters are those committing the lowest number of errors (71%). Again, the null of equality of distributions across the four political groups is rejected in a Kruskall-Wallis test ($p = 0.0030$). However, Kolmogorov-Smirnov tests on equality of distributions in pairs of political groups fail to reject the null of equality of distributions – although significance is close to the 10% level for right-wing vs. far-right ($p = 0.11$, n=352) and right-wing vs. left ($p = 0.14$, n=528). Overall, it then seems that right-wing people tend to be particularly imprecise in their forecasts both on fellow members and other party members.

We also asked participants to estimate the exact *share* of other participants from their own party and from other parties who preferred the narrative they indicated as the most preferred for the relative group. Even in this case, each correct answer was monetarily incentivized. While there is a large variability in predictions that encompasses both overestimation and underestimation, overestimation largely prevails (Figure I1). This holds across parties, narratives, and beliefs over own and other party supporters. The forecasting error is higher in beliefs over own party group than over the other ones (signrank test: $p = 0.0310$). The forecast errors tend in this case to be higher for the far-right supporters than for other groups (ranksum test in pairwise comparisons of far-right with other parties for beliefs over one's own political group: $p = 0.1636$ vs. far-left; $p = 0.0224$ vs. left; $p = 0.0581$ vs. right-wing; as for beliefs over other political group: $p = 0.2819$ vs. far-left, $p = 0.0350$ vs. left-wing; $p = 0.0073$ vs. right-wing). In contrast, we do not find significant differences between other political groups[40]. The forecasting errors by far-righters is higher in the belief that supporters of other parties hold the same belief as their own (Figure I1, Panel c). Overestimation occurs even when the narrative expected to be the most preferred by others is different from the participant's own preferred narrative (Figure I1, Panel b and d). This suggests that people tend to be over-confident in their ability to predict others' behavior. This component is likely to affect also beliefs that others' preferred narrative is the same as the individual preference. This suggests that the false consensus effect may be partly driven by over-confidence in one's own predictive ability.

In sum, while right-wingers are the most imprecise in the straight forecast of the narrative preferred by others and far-right supporters the least imprecise, the latter are those most imprecise in overstating the *share* of people following a certain narrative.

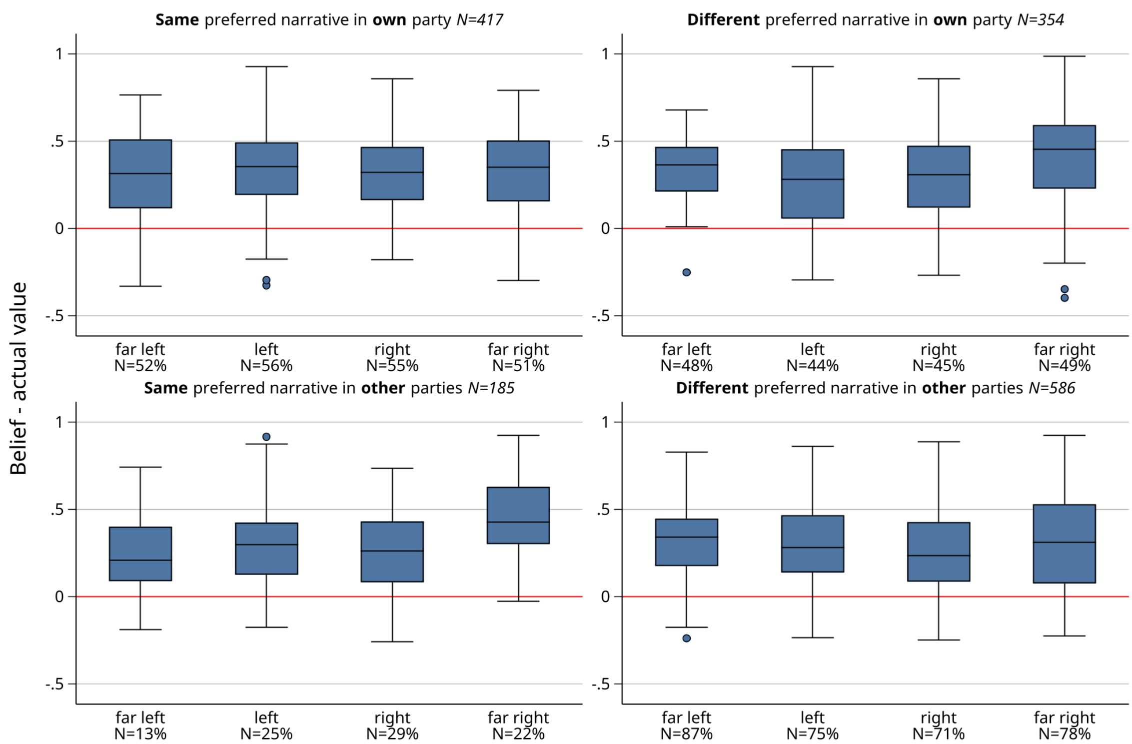


Figure I1: Forecast errors for beliefs on own and other parties' preferred narratives
Error forecast over beliefs on the share of participants who supported the most preferred narrative in own - on the top - and other - on the bottom - parties by party group when that narrative corresponds to the own preferred one - to the left - or to a different one - to the right. The center line in each box represents the 50th percentile (median) of the error forecast in each category. The bottom (top) of each box represents the 25th (75th) percentile. The bottom (upper) whisker below (above) the box represents the lower (upper) adjacent value. The circles lying above or under the hinges identify outside values.

[40] Ranksum test over one's own political group: far-left vs left $p = 0.5752$, far-left vs right $p = 0.6456$, left vs right $p = 0.9139$; over the other political group: far-left vs left $p = 0.3969$, far-left vs right $p = 0.0748$, left vs right $p = 0.1943$.

Table I1: Actual vs predicted supporters of own preferred narrative

		Party affiliation			
		far-right	right-wing	left-wing	far-left
Predicted:	H/I	66.86	44.06	53.83	42.50
Actual:	H/I	21.01	14.49	7.32	8.57
Predicted:	H/O	70.71	41.67	53.00	65.00
Actual:	H/O	24.64	8.41	5.73	11.43
Predicted:	H/R	66.65	45.61	52.70	37.33
Actual:	H/R	44.93	23.83	14.33	12.38
Predicted:	A/I	50.00	61.07	55.11	41.50
Actual:	A/I	1.45	10.75	15.92	13.33
Predicted:	A/O		53.31	61.18	56.76
Actual:	A/O	2.90	28.04	34.71	35.24
Predicted:	A/R	50.00	62.47	60.67	66.55
Actual:	A/R	5.07	14.49	21.97	19.05

Mean values of the beliefs on the share of supporters of own party who supported the own preferred narrative. These values are contrasted in the bottom rows with the actual share of supporters by party group from Survey 2.

We now look at the perceived level of disagreement with supporters of other parties. Figure I2 presents a measure of this disagreement by showing, for each party, the percentage of voters who believed the majority of participants of other parties preferred the opposite type of narrative—anti-immigration for pro-immigration participants and vice versa. The perceived level of disagreement was higher for extreme parties than for mainstream ones: using a Wilcoxon rank sum test, the difference was statistically significant at the 1% level. Moreover, it was not significant when confronting far-right and far-left ($p > 0.15$) and was only weakly so when confronting left-wing and right-wing ($p > 0.05$).

These analyses suggests there might be some overestimation of the level of disagreement across different parties' supporters, but it is not conclusive; that analysis does not tell us anything about the discrepancy between the perceived level and actual one. We then performed an analysis on overestimation at the aggregate level by confronting beliefs on the other parties' preferred narratives with the actual distribution from Survey 1. Table I2 shows the distribution of beliefs on the preferred narrative by the other parties. This can be confronted with numbers in Figure 4 in the main text to determine the discrepancy. We performed this comparison with Kolmogorov–Smirnov tests among the distributions. Over 50% of far-right individuals correctly predicted that a pro-immigration narrative was the preferred one, and 29% of them also correctly guessed that A/O was the preferred one. However, the aggregate distribution of beliefs does not reflect well the actual distribution of preferences of the other parties' supporters ($p < 0.01$), because it overestimates how many people of other parties are in favor of migration. Right-wing individuals tended to believe that most people from other parties were against migration, while the opposite was actually true ($p < 0.01$). Moving on to left-wing individuals, 54%

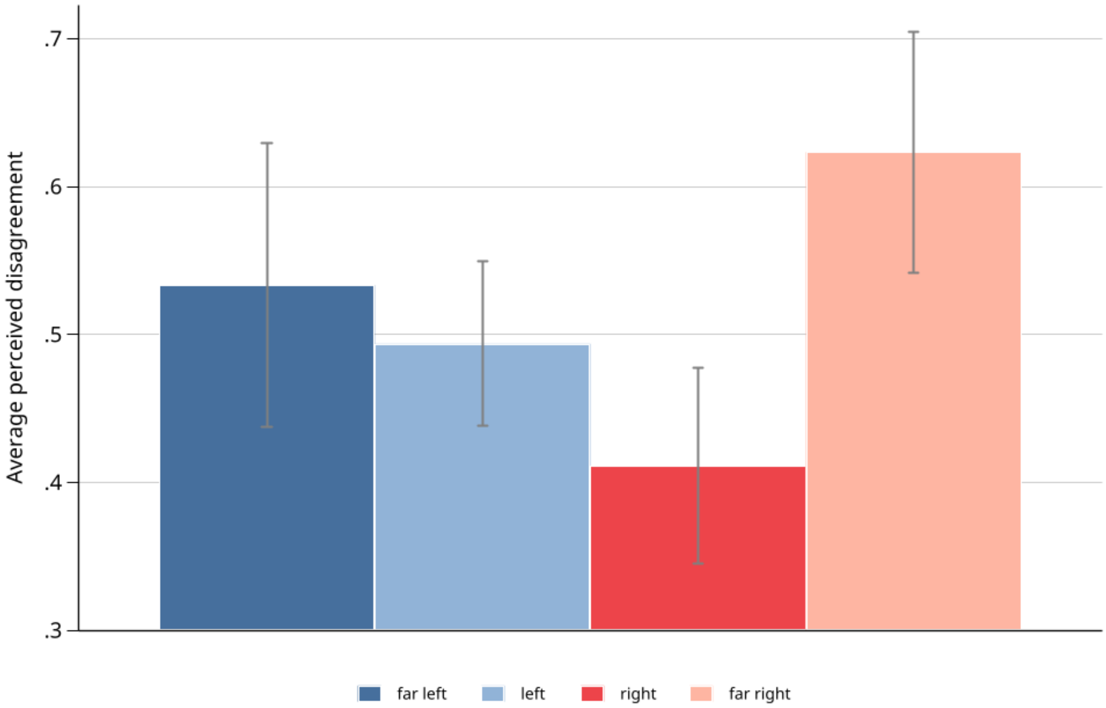


Figure I2: Perceived level of disagreement
The share of voters from Survey 2 by party who believed the majority of participants of other parties preferred the opposite type of narrative—anti-immigration for pro-immigration participants and vice versa

of them correctly predicted that anti-immigration narratives were the most preferred by other parties, and 26% of them correctly guessed that H/R attracted more favor. However, they exaggerated how many people preferred A/I over A/O, so that the distributions of beliefs and actual preferences still come out as significantly different ($p < 0.01$). More or less, the same held for far-left individuals. Here, 59% of them were right that anti-immigration narratives were preferred, and 28% of them guessed the preferred one was H/R. However, they exaggerated how many participants preferred H/I over H/O so that the aggregate distributions are significantly different ($p < 0.01$).

Finally, Figure I3 shows moral disapproval, that is, the discrepancy between empirical and normative expectations. The level of moral disapproval was around 35%, meaning that more or less one-third of the participants experienced dissonance between the narrative they thought should be endorsed and the one they expected most people to endorse. This was weaker for far-left individuals. Only 25% of them experienced such discrepancy, and the difference with left-wing individuals by a Wilcoxon rank sum test was statistically significant ($p < 0.05$). This difference was only close to significance instead when comparing far-left individuals with right-wing ones ($p > 0.05$) and far-right ones ($p > 0.10$).

Table I2: Beliefs on the most preferred narrative by other parties

	Party affiliation				
	far-right	right-wing	left-wing	far-left	Total
Beliefs on most preferred of other parties					
H/I	15	35	48	22	120
	10.87%	16.36%	15.29%	20.95%	15.56%
H/O	10	29	41	10	90
	7.25%	13.55%	13.06%	9.52%	11.67%
H/R	21	57	82	30	190
	15.22%	26.64%	26.11%	28.57%	24.64%
A/I	25	31	64	14	134
	18.12%	14.49%	20.38%	13.33%	17.38%
A/O	40	34	41	18	133
	28.99%	15.89%	13.06%	17.14%	17.25%
A/R	27	28	38	11	104
	19.57%	13.08%	12.10%	10.48%	13.49%

No. of supporters (and percentages in the bottom rows) by party who believe a narrative is the most preferred by supporters of the other party groups.

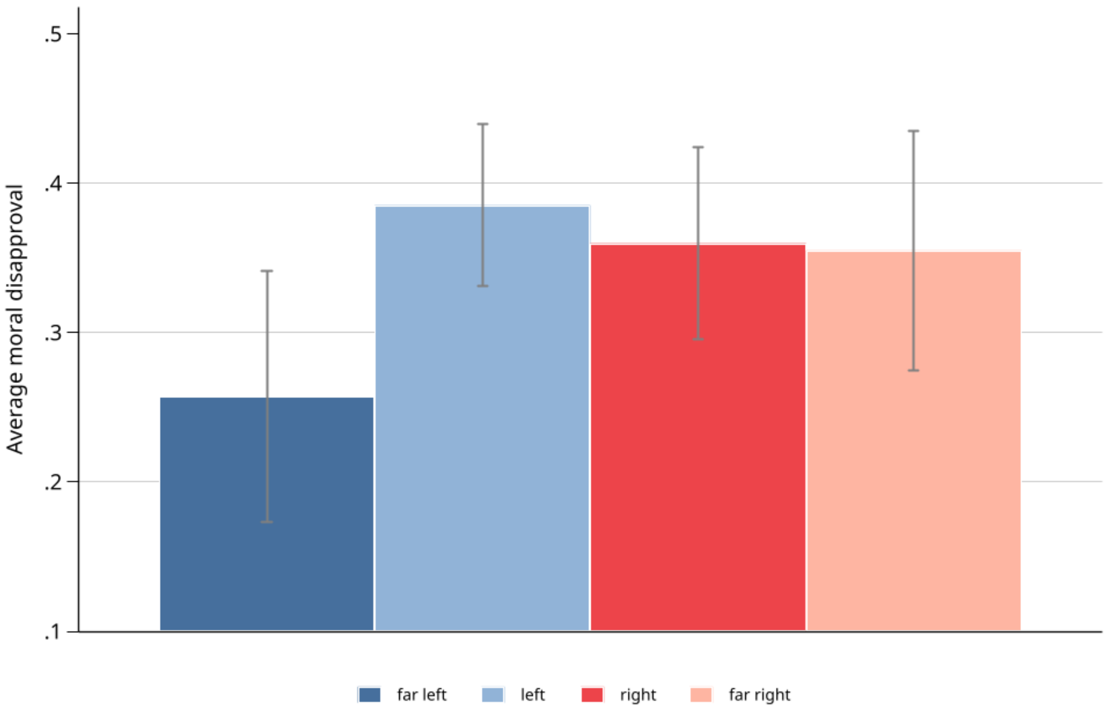


Figure I3: Moral disapproval by party
The share of individuals expressing moral disapproval by party group. Moral disapproval is a dummy variable taking a value of 1 if there is a discrepancy between the belief over the most preferred narrative in the overall sample and the belief over the narrative that respondents should agree with.

J Additional tables

Table J1: Descriptive statistics on our sample for Survey 1

	Party				
	far right	right	left	far left	Total
Age					
Under 30	23	47	100	18	188
	10.18%	13.82%	20.08%	11.11%	15.33%
Between 30 and 60	143	182	228	84	637
	63.27%	53.53%	45.78%	51.85%	51.96%
Over 60	60	111	170	60	401
	26.55%	32.65%	34.14%	37.04%	32.71%
Gender					
Female	86	177	280	70	613
	38.05%	52.06%	56.22%	43.21%	50.00%
Men	140	163	218	92	613
	61.95%	47.94%	43.78%	56.79%	50.00%
Born in Germany					
No	9	15	15	3	42
	4.02%	4.44%	3.05%	1.85%	3.46%
Yes	215	323	476	159	1,173
	95.98%	95.56%	96.95%	98.15%	96.54%
Mother born in Germany					
Yes	209	305	455	151	1,120
	93.30%	90.50%	91.92%	93.79%	92.03%
No	15	31	35	8	89
	6.70%	9.20%	7.07%	4.97%	7.31%
Do not know	0	1	5	2	8
	0.00%	0.30%	1.01%	1.24%	0.66%
Education					
No degree	2	1	3	1	7
	0.88%	0.29%	0.60%	0.62%	0.57%
Hauptschule	29	16	44	11	100
	12.83%	4.71%	8.84%	6.79%	8.16%
Realschule	65	84	112	28	289
	28.76%	24.71%	22.49%	17.28%	23.57%
High school diploma	30	60	99	26	215
	13.27%	17.65%	19.88%	16.05%	17.54%
Other high school diploma	20	54	76	33	183
	8.85%	15.88%	15.26%	20.37%	14.93%
University degree	27	48	53	21	149
	11.95%	14.12%	10.64%	12.96%	12.15%
Degree from a university of applied sciences	1	2	6	3	12
	0.44%	0.59%	1.20%	1.85%	0.98%
Ph.D.	3	7	6	3	19
	1.33%	2.06%	1.20%	1.85%	1.55%
Dual vocational training	37	32	53	19	141
	16.37%	9.41%	10.64%	11.73%	11.50%
Master's degree	3	21	12	9	45
	1.33%	6.18%	2.41%	5.56%	3.67%
Other professional degree	9	15	34	8	66
	3.98%	4.41%	6.83%	4.94%	5.38%
Income					
under 900 Euro	18	10	34	20	82
	8.04%	2.96%	6.88%	12.42%	6.74%
900-1300 Euro	15	20	58	21	114
	6.70%	5.92%	11.74%	13.04%	9.37%
1301-1500 Euro	17	24	40	14	95
	7.59%	7.10%	8.10%	8.70%	7.81%
1501-2000 Euro	37	40	67	23	167
	16.52%	11.83%	13.56%	14.29%	13.72%
2001-2600 Euro	42	56	79	27	204
	18.75%	16.57%	15.99%	16.77%	16.76%
2601-3200 Euro	39	61	74	23	197
	17.41%	18.05%	14.98%	14.29%	16.19%
3201-4500 Euro	34	81	87	22	224
	15.18%	23.96%	17.61%	13.66%	18.41%
4501-6000 Euro	16	27	43	11	97
	7.14%	7.99%	8.70%	6.83%	7.97%
more than 6001 Euro	6	19	12	0	37
	2.68%	5.62%	2.43%	0.00%	3.04%
Employment status					
Employed	112	158	199	62	531
	49.56%	47.02%	40.12%	38.51%	43.56%
Self-employed	14	31	25	10	80
	6.19%	9.23%	5.04%	6.21%	6.56%
450 Euro employment	10	10	21	4	45
	4.42%	2.98%	4.23%	2.48%	3.69%
Working without registration	3	3	6	1	13
	1.33%	0.89%	1.21%	0.62%	1.07%
Currently not employed and not looking for work	5	10	23	5	43
	2.21%	2.98%	4.64%	3.11%	3.53%
Looking for work but currently unemployed	15	7	14	13	49
	6.64%	2.08%	2.82%	8.07%	4.02%
Student	2	11	32	7	52
	0.88%	3.27%	6.45%	4.35%	4.27%
Retired	50	89	158	50	347
	22.12%	26.49%	31.85%	31.06%	28.47%
Apprentice and trainee	3	5	9	3	20
	1.33%	1.49%	1.81%	1.86%	1.64%
Other	12	12	9	6	39
	5.31%	3.57%	1.81%	3.73%	3.20%
Occupation					
Managers	25	53	37	18	133
	11.21%	15.87%	7.61%	11.32%	11.06%
Professionals	72	124	187	65	448
	32.29%	37.13%	38.48%	40.88%	37.27%
Technicians and Associate Professionals	14	19	24	8	65
	6.28%	5.69%	4.94%	5.03%	5.41%
Clerical Support Workers	43	70	117	33	263
	19.28%	20.96%	24.07%	20.75%	21.88%
Service and Sale Workers	19	18	42	5	84
	8.52%	5.39%	8.64%	3.14%	6.99%
Skilled Agricultural Forestry and Fishery Workers	2	9	4	3	18
	0.90%	2.69%	0.82%	1.89%	1.50%
Craft and Related Trades Workers	24	12	25	9	70
	10.76%	3.59%	5.14%	5.66%	5.82%
Plant and Machines Operators and Assemblers	6	6	8	2	22
	2.69%	1.80%	1.65%	1.26%	1.83%
Unskilled Labor	15	16	41	13	85
	6.73%	4.79%	8.44%	8.18%	7.07%
Armed Forces Occupations	3	7	1	3	14
	1.35%	2.10%	0.21%	1.89%	1.16%
Religion					
I do not belong to any religious community	129	145	214	103	591
	57.08%	42.90%	43.32%	63.98%	48.48%
Protestant	39	75	150	30	294
	17.26%	22.19%	30.36%	18.63%	24.12%
Catholic	51	98	101	21	271
	22.57%	28.99%	20.45%	13.04%	22.23%
Christian Orthodox	3	4	5	2	14
	1.33%	1.18%	1.01%	1.24%	1.15%
Islamic	0	5	11	1	17
	0.00%	1.48%	2.23%	0.62%	1.39%
Jewish	0	1	1	0	2
	0.00%	0.30%	0.20%	0.00%	0.16%
Other	4	10	12	4	30
	1.77%	2.96%	2.43%	2.48%	2.46%

Table J2: Descriptive statistics on our sample for Survey 2

	Party				
	far right	right	left	far left	Total
Age					
Under 30	13	28	50	16	107
	9.42%	13.08%	15.92%	15.24%	13.88%
Between 30 and 60	75	117	163	42	397
	54.35%	54.67%	51.91%	40.00%	51.49%
Over 60	50	69	101	47	267
	36.23%	32.24%	32.17%	44.76%	34.63%
Gender					
Female	66	109	160	42	377
	47.83%	50.93%	50.96%	40.00%	48.90%
Men	72	105	154	63	394
	52.17%	49.07%	49.04%	60.00%	51.10%
Born in Germany					
No	8	5	14	2	29
	5.97%	2.37%	4.47%	2.00%	3.83%
Yes	126	206	299	98	729
	94.03%	97.63%	95.53%	98.00%	96.17%
Mother born in Germany					
Yes	127	193	285	97	702
	93.38%	91.04%	91.64%	93.27%	92.01%
No	8	19	25	5	57
	5.88%	8.96%	8.04%	4.81%	7.47%
Do not know	1	0	1	2	4
	0.74%	0.00%	0.32%	1.92%	0.52%
Education					
No degree	1	0	3	1	5
	0.72%	0.00%	0.96%	0.95%	0.65%
Hauptschule	19	12	23	5	59
	13.77%	5.61%	7.32%	4.76%	7.65%
Realschule	46	55	67	25	193
	33.33%	25.70%	21.34%	23.81%	25.03%
High school diploma	16	41	65	23	145
	11.59%	19.16%	20.70%	21.90%	18.81%
Other high school diploma	11	21	49	17	98
	7.97%	9.81%	15.61%	16.19%	12.71%
University degree	18	32	37	12	99
	13.04%	14.95%	11.78%	11.43%	12.84%
Degree from a university of applied sciences	1	3	2	0	6
	0.72%	1.40%	0.64%	0.00%	0.78%
Ph.D.	0	7	3	0	10
	0.00%	3.27%	0.96%	0.00%	1.30%
Dual vocational training	20	26	44	14	104
	14.49%	12.15%	14.01%	13.33%	13.49%
Master's degree	3	9	9	3	24
	2.17%	4.21%	2.87%	2.86%	3.11%
Other professional degree	3	8	12	5	28
	2.17%	3.74%	3.82%	4.76%	3.63%
Income					
under 900 Euro	13	7	16	15	51
	9.56%	3.29%	5.10%	14.56%	6.66%
900-1300 Euro	14	13	39	19	85
	10.29%	6.10%	12.42%	18.45%	11.10%
1301-1500 Euro	13	12	29	12	66
	9.56%	5.63%	9.24%	11.65%	8.62%
1501-2000 Euro	17	26	41	14	98
	12.50%	12.21%	13.06%	13.59%	12.79%
2001-2600 Euro	17	27	40	17	101
	12.50%	12.68%	12.74%	16.50%	13.19%
2601-3200 Euro	24	34	42	10	110
	17.65%	15.96%	13.38%	9.71%	14.36%
3201-4500 Euro	20	46	46	10	122
	14.71%	21.60%	14.65%	9.71%	15.93%
4501-6000 Euro	7	18	35	2	62
	5.15%	8.45%	11.15%	1.94%	8.09%
more than 6001 Euro	3	16	6	0	25
	2.21%	7.51%	1.91%	0.00%	3.26%
10	8	14	20	4	46
	5.88%	6.57%	6.37%	3.88%	6.01%
Employment status					
Employed	68	93	134	32	327
	49.64%	43.66%	42.68%	31.07%	42.63%
Self-employed	4	23	19	4	50
	2.92%	10.80%	6.05%	3.88%	6.52%
450 Euro employment	6	9	8	1	24
	4.38%	4.23%	2.55%	0.97%	3.13%
Working without registration	0	1	0	3	4
	0.00%	0.47%	0.00%	2.91%	0.52%
Currently not employed and not looking for work	5	5	12	6	28
	3.65%	2.35%	3.82%	5.83%	3.65%
Looking for work but currently unemployed	6	3	12	5	26
	4.38%	1.41%	3.82%	4.85%	3.39%
Student	2	9	19	5	35
	1.46%	4.23%	6.05%	4.85%	4.56%
Retired	42	58	91	43	234
	30.66%	27.23%	28.98%	41.75%	30.51%
Apprentice and trainee	1	5	8	0	14
	0.73%	2.35%	2.55%	0.00%	1.83%
Other	3	7	11	4	25
	2.19%	3.29%	3.50%	3.88%	3.26%
Occupation					
Managers	11	30	31	7	79
	8.09%	14.22%	9.94%	6.86%	10.38%
Professionals	51	83	121	29	284
	37.50%	39.34%	38.78%	28.43%	37.32%
Technicians and Associate Professionals	7	8	20	6	41
	5.15%	3.79%	6.41%	5.88%	5.39%
Clerical Support Workers	35	49	65	20	169
	25.74%	23.22%	20.83%	19.61%	22.21%
Service and Sale Workers	6	18	18	15	57
	4.41%	8.53%	5.77%	14.71%	7.49%
Skilled Agricultural Forestry and Fishery Workers	0	2	4	3	9
	0.00%	0.95%	1.28%	2.94%	1.18%
Craft and Related Trades Workers	9	8	16	6	39
	6.62%	3.79%	5.13%	5.88%	5.12%
Plant and Machines Operators and Assemblers	7	5	2	6	20
	5.15%	2.37%	0.64%	5.88%	2.63%
Unskilled Labor	10	7	34	10	61
	7.35%	3.32%	10.90%	9.80%	8.02%
Armed Forces Occupations	0	1	1	0	2
	0.00%	0.47%	0.32%	0.00%	0.26%
Religion					
I do not belong to any religious community	84	86	160	65	395
	61.31%	40.38%	51.12%	63.11%	51.57%
Protestant	25	57	78	22	182
	18.25%	26.76%	24.92%	21.36%	23.76%
Catholic	20	60	50	10	140
	14.60%	28.17%	15.97%	9.71%	18.28%
Christian Orthodox	3	2	5	1	11
	2.19%	0.94%	1.60%	0.97%	1.44%
Islamic	1	0	8	1	10
	0.73%	0.00%	2.56%	0.97%	1.31%
Jewish	0	1	2	1	4
	0.00%	0.47%	0.64%	0.97%	0.52%
Other	2	4	6	0	12
	1.46%	1.88%	1.92%	0.00%	1.57%

Table J3: Balance of individual characteristics between Survey 1 and 2

	Party			
	far left	left	right	far right
Age	0.842	0.893	1.000	0.398
Gender	0.055	0.274	1.000	0.066
Born in Germany	1.000	1.000	1.000	1.000
Mother born in Germany	1.000	1.000	1.000	1.000
Education	0.461	0.995	1.000	0.967
Income	0.466	0.219	0.311	0.758
Employment status	0.432	0.950	1.000	0.965
Occupation	0.083	0.647	0.814	0.598
Religion	1.000	0.156	1.000	0.974

Each cell has the p-value of a Kolmogorov-Smirnov test on equality of distributions on the characteristics listed in rows by Survey 1 (N=1226) or 2 (N=771). In Survey 2, there are 162 far-left, 498 left-wing, 340 right-wing and 226 far-right respondents. In Survey 2, there are 105 far-left, 314 left-wing, 214 right-wing and 138 far-right respondents.

Figure J1: wordcloud of Trump tweets on migration during his presidency
The wordcloud is based on our own calculations from tweets by Donald Trump during his presidency. We acquired the tweets using an algorithm of our own making before he was banned.

Figure J2: Who-supports-what website

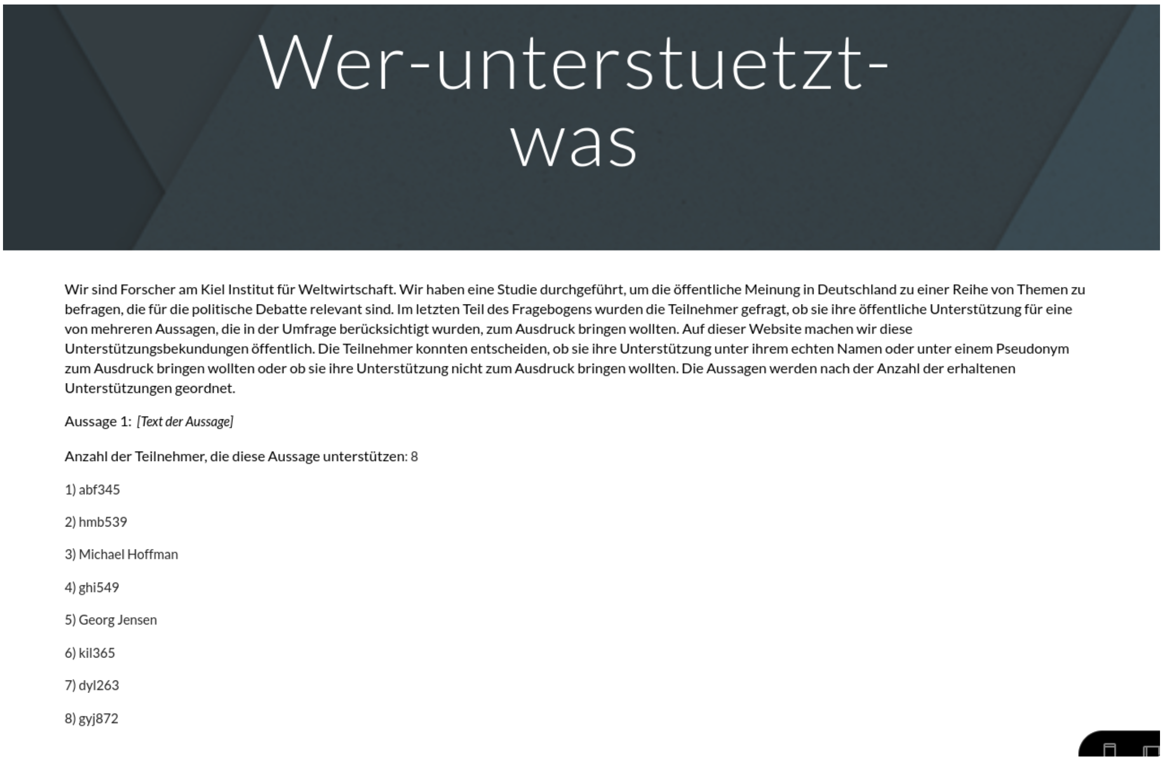
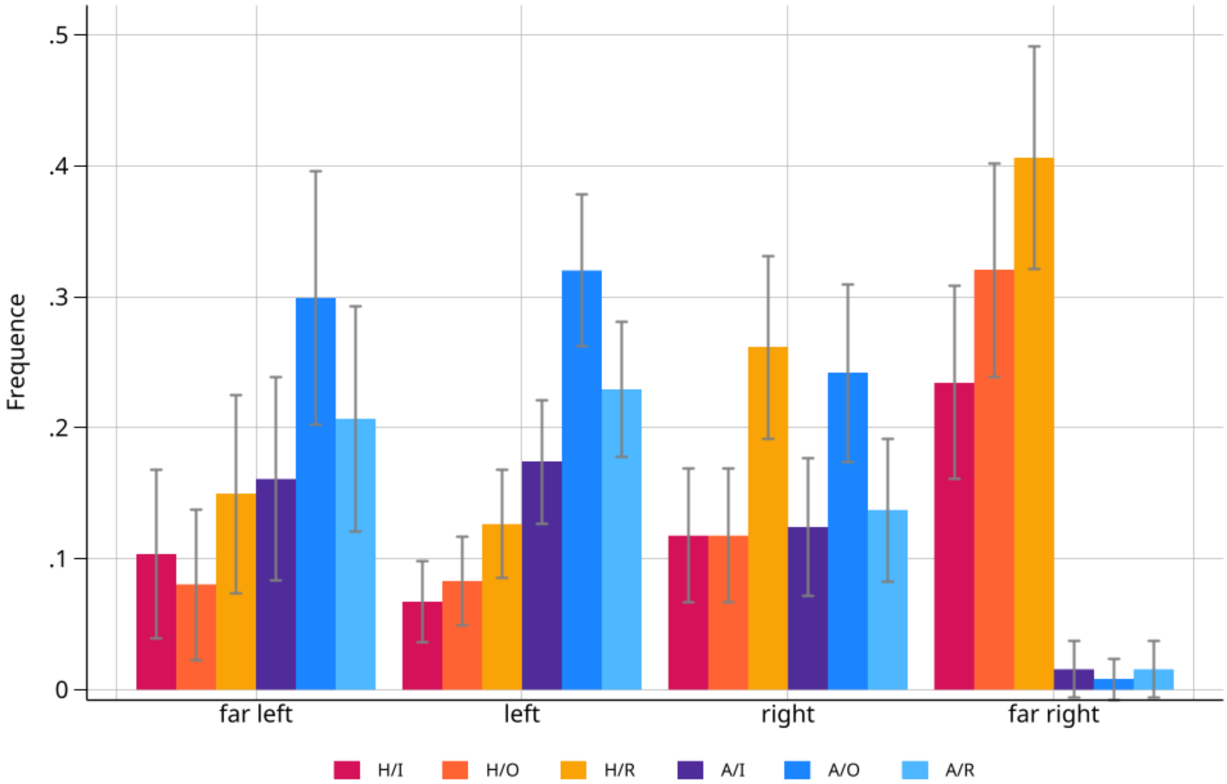


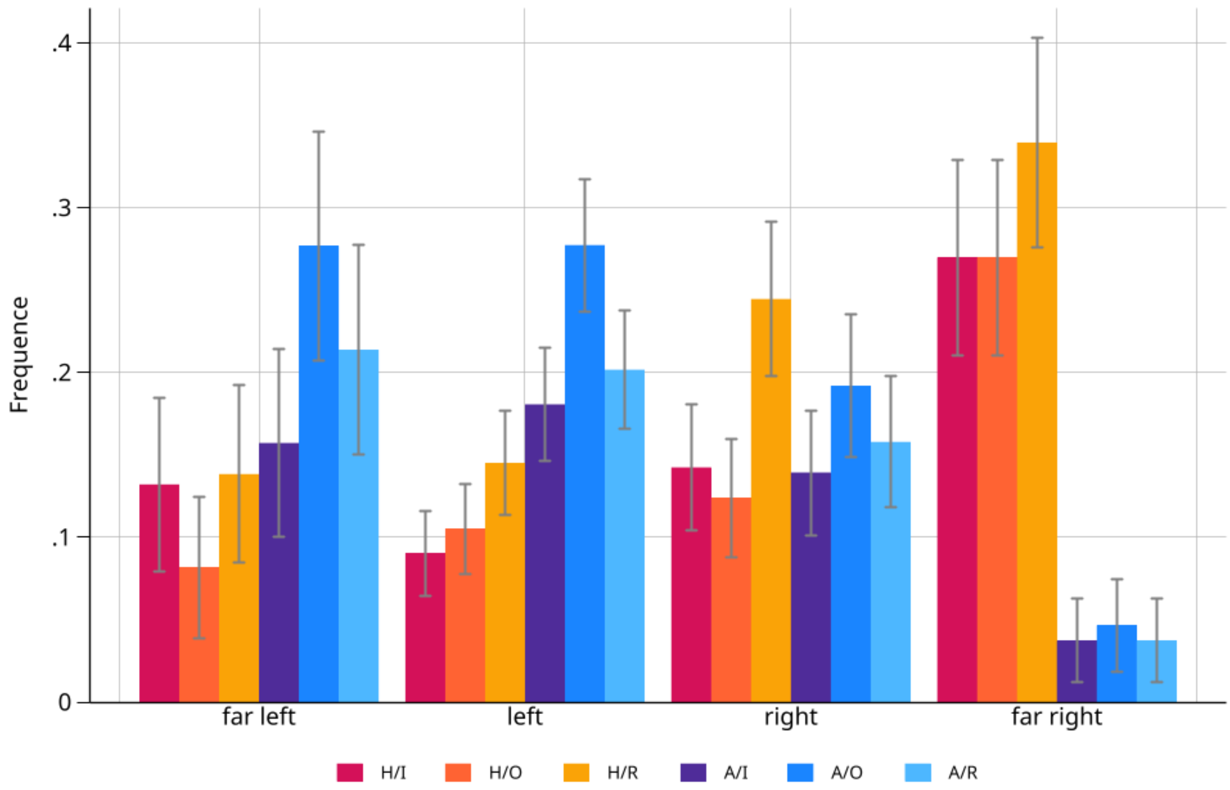
Image that participants in Survey 1 saw when deciding if to publicly endorse a narrative on the *Who-supports-what* website.

Figure J3: Private preferences over narratives - excluding low quality



Private preferences by party group. From the overall sample N=1226, we excluded N=605 participants who reported inconsistent answers between the elicitation through a ranking and through Likert scales.

Figure J4: Private preferences over narratives - Germans only



Private preferences by party group. From the overall sample N=1226, we excluded N=53 participants who were either immigrants or who did not report their country of birth.

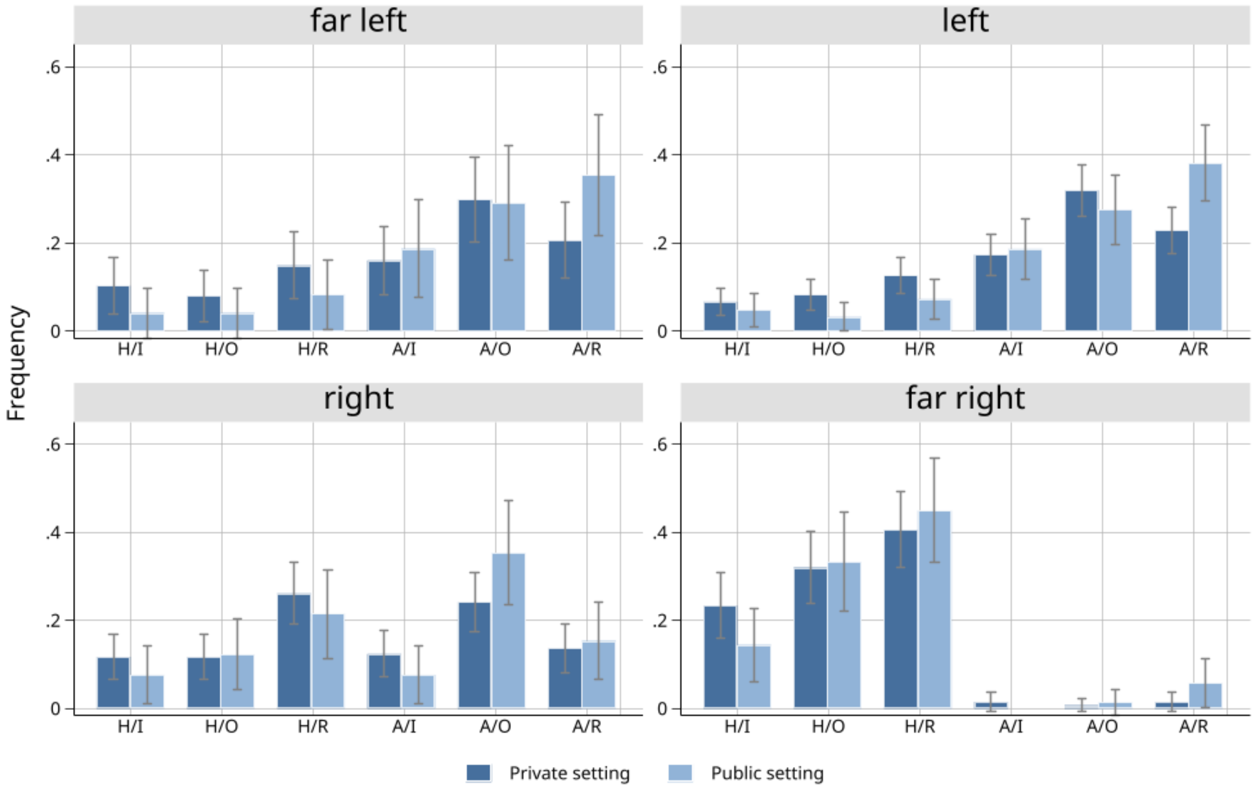


Figure J5: Comparison between private and public preferences - excluding low quality
Private and Public preferences by party group. From the overall sample N=1226 (N=522 for the Public preferences), we excluded N=605 (N=217) participants who reported inconsistent answers between the elicitation through a ranking and through Likert scales.

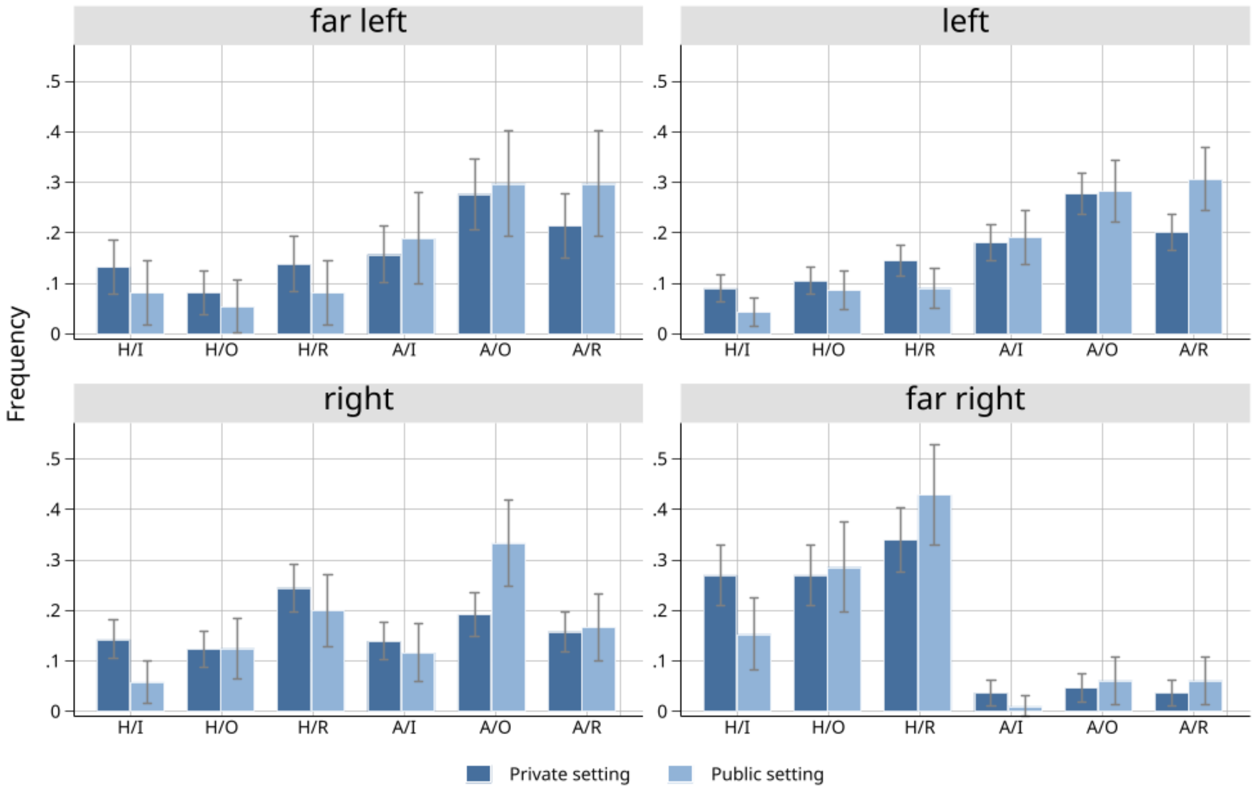


Figure J6: Comparison between private and public preferences - Germans only
Private and Public preferences by party group. From the overall sample N=1226 (N=522 for the Public preferences), we excluded N=53 (N=21) participants who were either immigrants or who did not report their country of birth.

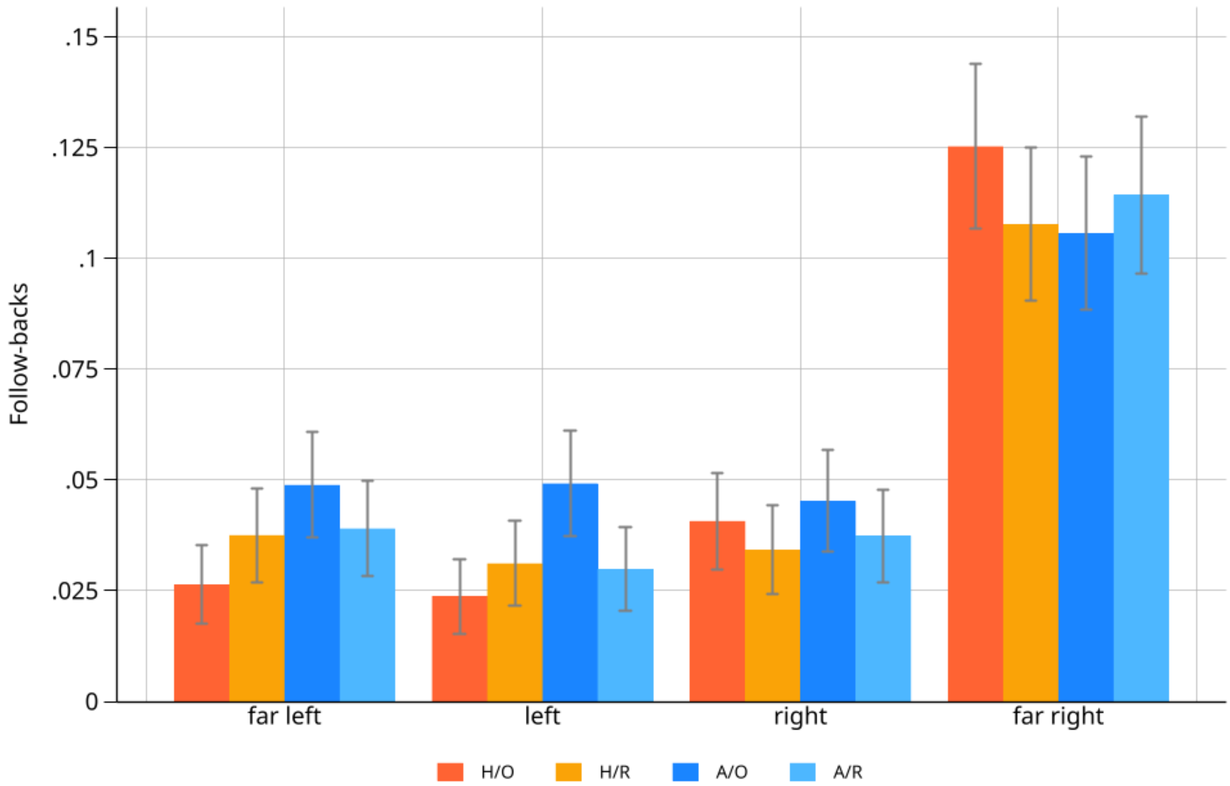
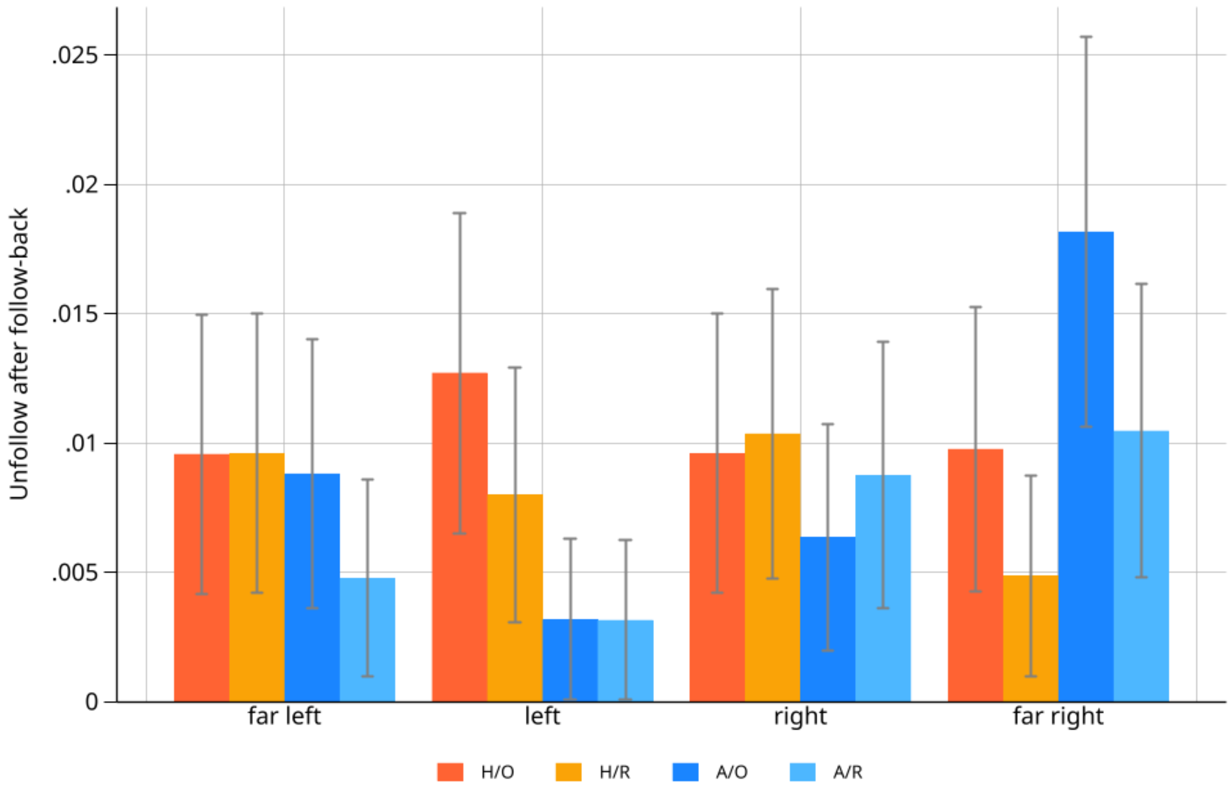


Figure J7: Follow-back by pinned tweet and by party - including those that later unfollow

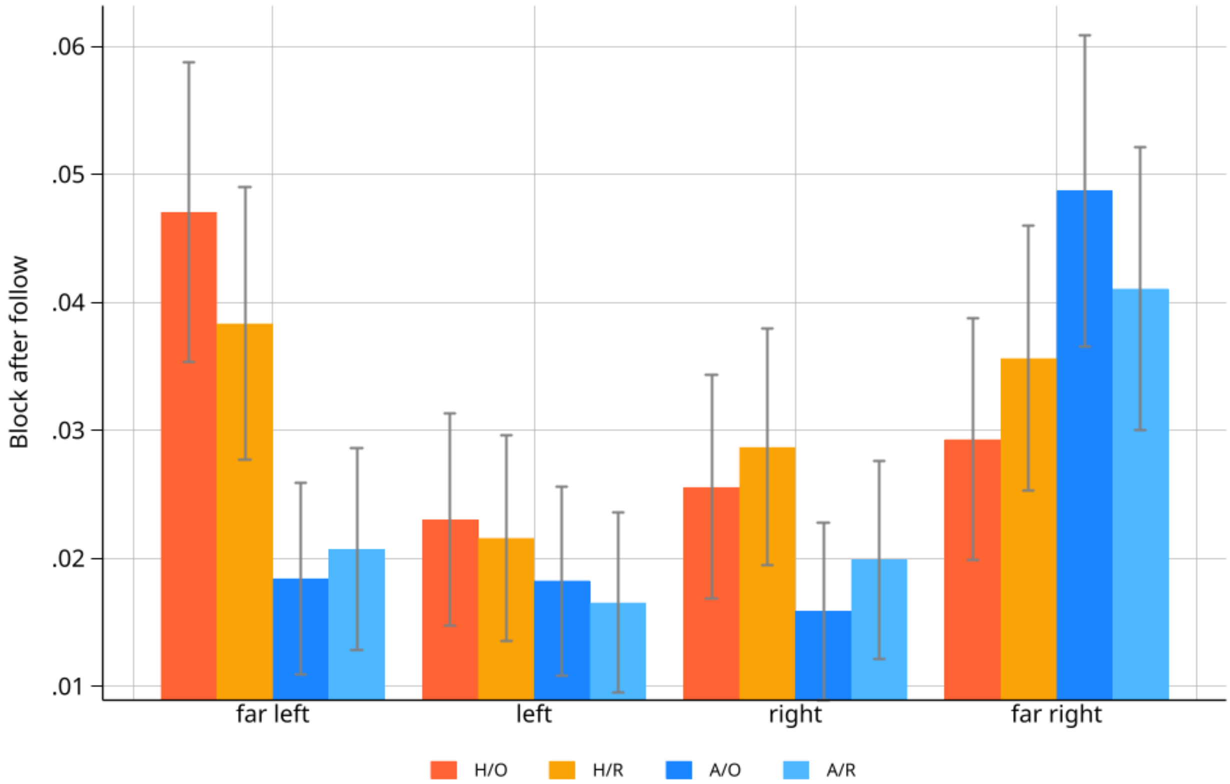
Frequency of users who followed-back by political orientation from the Twitter experiment regardless if they later unfollow. The total sample includes 19,989 users. Reported confidence intervals are at 95% level.

Figure J8: Unfollow after follow-back by pinned tweet and by party group



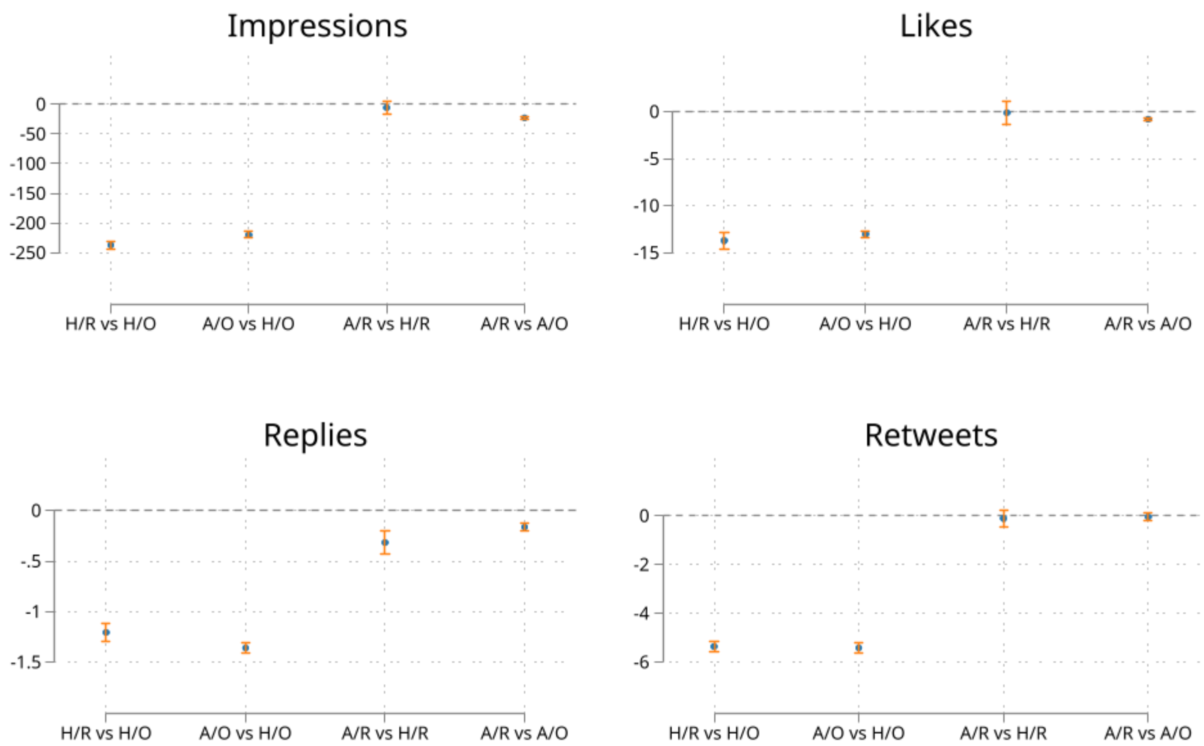
Frequency of users who unfollowed after having followed-back by political orientation from the Twitter experiment. The total sample includes 19,989 users. Reported confidence intervals are at 95% level.

Figure J9: Blocking by pinned tweet and by party group



Frequency of users who block the experimental profile after the profile followed them by political orientation from the Twitter experiment. The total sample includes 19,989 users. Reported confidence intervals are at 95% level.

Figure J10: Pairwise comparisons from regressions on pinned tweets



OLS regression on impressions, likes, replies and retweets associated with each tweeting of the pinned tweet. Errors are clustered at artificial profile level. Included controls are fixed effects associated with the date of the tweeting. Each coefficient represent the result of a pairwise comparison of engagements over the pinned tweets.

K Instructions for Survey 1

--- Page 1 ---

The items are numbered just for reference. The questions were not numbered in the survey,

Dear participant, our names are Gianluca Grimalda and Michael Bayerlein. We are researchers at the Kiel Institute for the World Economy.

The goal of the study is to survey public opinion in Germany over a range of topics relevant for our society – in particular immigration and COVID-19. If you decide to participate, you will be asked to complete a research survey about your views and attitudes over such topics. The survey will take approximately 10-15 minutes and you will be paid according to the standard rate of compensation for this time. In addition, you will be asked some basic demographic questions, including your political orientation and voting intention in the next general elections. In the last part of the questionnaire you will be asked whether you want to express public support for some of the statements that will form part of the survey, at your discretion. Occasionally you will also be asked some questions just to check your attention.

Your participation in this study is purely voluntary and your data will be treated confidentially by our research team. The data we receive from the company that contacted you are fully anonymized. Except for the initial questions, you are free not to answer some of the questions, but we would really grateful if you answered all the questions. Your opinion matters to us and remember, the final version of the questionnaire is fully anonymous. You are free to opt out of the survey at any time you wish, but in this case you will forfeit your payment.The results of the study may be published or presented at professional meetings, but only group characteristics will be discussed. All payments are managed through Kantar and we will never contact you directly.

We will be happy to answer any questions you have about the study. You may contact us at phone: XXXXX or email: XXXXXX. Thank you for your consideration. If you would like to participate, please click on the button below. When you are done, simply click "finish" to complete the survey.

With kind regards,

Dr.Gianluca Grimalda and Mr. Michael Bayerlein

--- Page 2 ---

Thank you for agreeing to participate in this survey. It is very important for the success of our research that you read the questions carefully and answer thoughtfully. Please answer all of the questions to the best of your knowledge, keeping in mind that your responses are anonymous.

The completion time is about 10-15 minutes.

--- Page 3 ---

1) How old are you?

2) Please indicate your gender.

- Woman
- Male
- Diverse

3) Please indicate your nationality.

- German nationality

- German and other nationality
- What other nationality do you have?
- No German citizenship

4) What is your highest educational qualification?

- No degree
- Hauptschulabschluss
- Realschulabschluss
- High school diploma
- Other high school diploma
- University degree
- Degree from a university of applied sciences
- Doctorate (Dr.)
- Dual vocational training
- Master's degree
- Other professional degree

5) Are you eligible to vote in the upcoming federal election?

- Yes
- No

6) Which party will you vote for in the upcoming federal election?

- CDU/CSU
- SPD
- Greens
- AfD
- The Left Party
- FDP
- Other
- Which party?
- I will not vote
- No answer

7) Many people use the terms 'left' and 'right' when referring to different political attitudes. Thinking about your own political views, where would you rank those views on this scale?

--- Page 4 ---

8) Please briefly describe your opinion on the situation of this country in relation to immigration (max 100 words.):

--- Page 5 ---

9) Immigration is a hot topic in the current political debate. We would like you to read the six statements below and tell us how much you agree with them. Such statements reflect the views recently expressed by political parties in Germany.

Please rank the following statements from 1 to 6 based on how much you agree with them, where 1 corresponds to the statement you agree the most, 2 corresponds to the second statement with which you agree the most, and so on until 6, which corresponds to the statement you agree the least. If you are indifferent between two statements, please do your best to think which one you prefer, and give it a better ranking, even if slightly. Please do not rank two statements with the same number.

From 2015 to 2020, almost 10 million migrants arrived to Germany. Germans and immigrants have different values and practices as well as different job skills and work attitudes. The integration of immigrants into our society represents a profitable investment and the costs to integrate them will be more than compensated in the future.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The values and practices of many of these immigrants can enrich our cultural life. Furthermore, immigrants also carry job skills and work attitudes that are needed for our economy.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The German values and practices we hold so dear can be relied upon to live peacefully with migrants' values. Furthermore, Germans have all that it needs to sustain a strong economy even together with immigrants.

From 2015 to 2020, almost 10 million migrants arrived to Germany. Germans and immigrants have different values and practices as well as different job skills and work attitudes. The integration of immigrants into our society represents too costly an investment and the costs to integrate them will never be compensated in the future.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The unacceptable values and practices of many of these immigrants are incompatible with our cultural life. Furthermore, immigrants also carry job skills and work attitudes that menace to permanently harm our economy.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The German values and practices we hold so dear have to be preserved from migrants' values. Furthermore, Germans have all that it needs to sustain a strong economy even without immigrants.

--- Page 6 ---

10) Consider again the six statements seen before. They are now ordered based on the ranking you chose in the previous screen. Please state how much you agree with each one of them on a scale from 1 (not at all) to 10 (very much).

--- Page 7 ---

11) Consider again the six statements seen before. Which statement do you think has been selected by most people as the one with which they agreed the most? At the end of this research, we will count how many people selected a certain statement as the one they most agreed with, and will determine which statement was selected by most people as the one they most agreed with. You will receive 50 Lifepoints on the top of your base earnings if your guess is correct.

12) Which statement do you think is the second most selected as the one with which they agreed the most? You will receive 50 Lifepoints on the top of your base earnings if your guess is correct.

--- Page 8 ---

We would now like to give you the possibility to express your support for one of the statements you saw in the previous section on the publicly accessible webpage "https://Who-supports-what.com". Your support will be made public along with the support of all other participants to this survey who decide to do so. The website "https://Who-supports-what.com" will be active from the 20th of September and will be deleted on the 25th of September. It will never be reactivated again. You can check the website to see which statement has received most endorsement by participants in this research.

You have the following options:

- Decline to express support;
- Express support under an alias name;
- Express support under your real name.

If you choose to express support under an alias name, this will have a format like: "abc123". You will be asked to indicate three letters and three numbers at the end of this research, and that will be your pseudonym.

Please see the layout of the webpage here:

We are researchers at the Kiel Institute for the World Economy. We conducted a study to survey public opinion in Germany on a range of topics relevant for the political debate. In the last part of the questionnaire, participants were asked whether they wanted to express public support for one among several statements that were considered in the survey. In this website we make these expressions of support public. Participants could decide to express support under their real name, under an alias name, or not express support, The statements are ranked according to the number of preferences they received.

Statement 1:

Number of participants who support this statement:

13) Please select which statement you would like to support publicly. (Please indicate your preference even if you later decline to express support publicly):

14)

- Would you like to express support for this statement publicly?
- No, I decline to express support publicly
- Yes, I express support under an alias name on the website https://Who-supports-what.com

--- Page 8.1 ---

15) You have decided to express support for a statement under an alias name. Please select three letters from a to z and three numbers from 0 to 9. This will be your alias name:

--- Page 9 ---

16) Please briefly describe your opinion on the situation of this country in relation to vaccination (max 100 words.):

--- Page 10 ---

17) Vaccination is another hot topic in the current political debate. We would like you to read the six statements below and tell us how much you agree with them. Such statements reflect the views recently expressed by political parties in Germany.

Please rank the following statements from 1 to 6 based on how much you agree with them, where 1 corresponds to the statement you agree the most, 2 corresponds to the second statement with which you agree the most, and so on until 6, which corresponds to the statement you agree the least. Please do not rank two statements with the same number.

From June to August, almost 20 million people got at least a dose of vaccine. The way unvaccinated people question scientific results and still experimental vaccines is a rational way of reacting to the pandemic. Furthermore, unvaccinated people are contributing to reach the point to which we are all immune against Covid-19 fast without at the same time putting themselves at a great risk.

From June to August, almost 20 million people got at least a dose of vaccine. The way unvaccinated people question scientific results and still experimental vaccines is just an irrational way of reacting to the pandemic. Furthermore, unvaccinated people are delaying the point in time to which we defeat Covid-19 while at the same time putting themselves at a great risk.

From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated people mostly show faith in science and in still experimental vaccines, which is a rational way of reacting to the pandemic. Furthermore, vaccinated people are contributing to reach the point to which we defeat Covid-19 while at the same time protecting themselves.

From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated and unvaccinated people show opposing levels of faith in science and in still experimental vaccines and have conflicting views on the rational way of reacting to the pandemic. Every attempt to make vaccinated and unvaccinated cooperate to reach the point we defeat Covid-19 and protect ourselves will always prove to be too costly and should not be pursued.

From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated and unvaccinated people show opposing levels of faith in science and in still experimental vaccines and have conflicting views on the rational way of reacting to the pandemic. Every attempt to make vaccinated and unvaccinated cooperate to reach the point we defeat Covid-19 and protect ourselves will always prove to be profitable and should be pursued.

From June to August, almost 20 million people got at least a dose of vaccine. Vaccinated people mostly show blind faith in science and in still experimental vaccines, which is an irrational way of reacting to the pandemic. Furthermore, vaccinated people are delaying the point in time to which we are all immune against Covid-19 without at the same time really protecting themselves.

--- Page 11 ---

18) Consider again the six statements seen before. They are now ordered based on the ranking you chose in the previous screen. Please state how much you agree with each one of them on a scale from 1 (not at all) to 10 (very much).

--- Page 12 ---

19) Consider again the six statements seen before. Which statement do you think has been selected by most people as the one with which they agreed the most? At the end of this research, we will count how many people selected a certain statement as the one they most agreed with, and will determine which statement was selected by most people as the one they most agreed with. You will receive 50 Lifepoints on the top of your base earnings if your guess is correct.

20) Which statement do you think is the second most selected as the one with which they agreed the most? You will receive 50 Lifepoints on the top of your base earnings if your guess is correct.

--- Page 13 ---

SEE ABOVE PAGE 8

--- Page 14 ---

23) To your knowledge, what is the percentage [between 0 and 100%] of Germans who have been without a legal job in relation to the total German population in 2019?

24) To your knowledge, what is the percentage [between 0 and 100%] of immigrants who have been without a legal job in relation to the total immigrant population in 2019?

25) To your knowledge, what is the percentage of vaccinated people [between 0 and 100%] who are in intensive care over unvaccinated ones among the total population as of September 2021?

--- Page 15 ---

Final Questionnaire

Before ending the questionnaire, we would like to ask you some final questions. Please answer all questions honestly and accurately. Your answers will stay anonymous.

26) Please indicate the postal code of your current place of residence:

27) Were you born in Germany?

- Yes

- No
- Please indicate in which country/region you were born

28) Was your mother born in Germany?

- Yes
- No
- Please indicate in which country/region your mother was born
- Don't know

29) Was your father born in Germany?

- Yes
- No
- Please indicate in which country/region your father was born
- Do not know

30) Were your grandparents born in Germany?

- Yes
- No
- Please indicate the countries/regions where your grandparents were born
- Partly
- Don't know

31) What is your monthly net household income? That is, the total income of all family members living in the household after deducting taxes and contributions per month.

- under 900 Euro
- 900-1300 Euro
- 1301-1500 Euro
- 1501-2000 Euro
- 2001-2600 Euro
- 2601-3200 Euro
- 3201-4500 Euro
- 4501-6000 Euro
- more than 6001 Euro

32) What is your current employment situation? (If you have more than one job, please indicate only your main job):

- Employed (more than 450 euros subject to social security contributions)
- Self-employed
- 450 Euro employment
- Working without registration (i.e. without social security)
- Currently not employed and not looking for work
- Looking for work, but currently unemployed
- Student
- Retired
- Apprentice and trainee
- Other

33) Take a look at the categories below. Which category most closely matches your current or most recent job?

- Manager

- Professional
- Technician and associate professional
- Office worker and support staff
- Service and sales worker
- Skilled worker in agriculture, forestry and fisheries
- Craftsman and related occupation
- Plant and machine operators and assemblers
- Unskilled laborer
- Military

34) Do you belong to a religious community? If yes, which one?

- I do not belong to any religious community.
- Protestant
- Catholic
- Christian Orthodox
- Islamic
- Judaism
- Other

35) How often do you meet with migrants? Meant personal meetings and conversations, not mere greetings (e.g., at work or in the neighborhood):

- Daily
- Several times a week
- Once a week
- Once a month
- Less often than once a month
- Never

36) How often do you wear a face mask?

- Daily
- Several times a week
- Once a week
- Once a month
- Less often than once a month
- Never

--- Page 16 ---

Thank you. We are really grateful for your participation in this survey. You can check the website "https://Who-supports-what.com" to see the results of this survey from the 20th to the 25th of September.

Finish the survey

L Instructions for Survey 2

--- Page 1 ---

The items are numbered just for reference. The questions were not numbered in the survey,

Dear participant, our names are Gianluca Grimalda and Michael Bayerlein. We are researchers at the Kiel Institute for the World Economy.

The goal of the study is to survey public opinion in Germany over a range of topics relevant for our society – in particular immigration and COVID-19. If you decide to participate, you will be asked to complete a research survey about your views and attitudes over such topics. The survey will take approximately 10-15 minutes and you will be paid according to the standard rate of compensation for this time. In addition, you will be asked some basic demographic questions, including your political orientation and voting intention in the next general elections. In the last part of the questionnaire you will be asked whether you want to express public support for some of the statements that will form part of the survey, at your discretion. Occasionally you will also be asked some questions just to check your attention.

Your participation in this study is purely voluntary and your data will be treated confidentially by our research team. The data we receive from the company that contacted you are fully anonymized. Except for the initial questions, you are free not to answer some of the questions, but we would really grateful if you answered all the questions. Your opinion matters to us and remember, the final version of the questionnaire is fully anonymous. You are free to opt out of the survey at any time you wish, but in this case you will forfeit your payment.The results of the study may be published or presented at professional meetings, but only group characteristics will be discussed. All payments are managed through Kantar and we will never contact you directly.

We will be happy to answer any questions you have about the study. You may contact us at phone: XXX or email: XXXX. Thank you for your consideration. If you would like to participate, please click on the button below. When you are done, simply click "finish" to complete the survey.

With kind regards,

Dr.Gianluca Grimalda and Mr. Michael Bayerlein

--- Page 2 ---

Thank you for agreeing to participate in this survey. It is very important for the success of our research that you read the questions carefully and answer thoughtfully. Please answer all of the questions to the best of your knowledge, keeping in mind that your responses are anonymous.

The completion time is about 10-15 minutes.

--- Page 3 ---

1) How old are you?

2) Please indicate your gender.

- Woman
- Male
- Diverse

3) Please indicate your nationality.

- German nationality
- German and other nationality
- What other nationality do you have?
- No German citizenship

4) What is your highest educational qualification?

- No degree
- Hauptschulabschluss

- Realschulabschluss
- High school diploma
- Other high school diploma
- University degree
- Degree from a university of applied sciences
- Doctorate (Dr.)
- Dual vocational training
- Master's degree
- Other professional degree

5) Are you eligible to vote in the upcoming federal election?

- Yes
- No

6) Which party will you vote for in the upcoming federal election?

- CDU/CSU
- SPD
- Greens
- AfD
- The Left Party
- FDP
- Other
- Which party?
- I will not vote
- No answer

7) Many people use the terms 'left' and 'right' when referring to different political attitudes. Thinking about your own political views, where would you rank those views on this scale?

--- Page 4 ---

Immigration is a hot topic in the current political debate. We would like you to read six statements and tell us your attitudes towards them. Such statements reflect the views recently expressed by political parties in Germany.

Please note: This survey is a follow-up of another survey on the same topic. You are welcome to participate in this second survey even if you took part in the previous one.

8) Please tell us how you feel about each statement you will see in the next pages:

{Randomize narratives}

From 2015 to 2020, almost 10 million migrants arrived to Germany. Germans and immigrants have different values and practices as well as different job skills and work attitudes. The integration of immigrants into our society represents a profitable investment and the costs to integrate them will be more than compensated in the future.

{Randomize emotions and keep politically correct at the bottom}

- This statement makes me happy.
- This statement makes me angry.
- This statement makes me fearful.
- This statement is politically correct.

{Repeat for other narratives}.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The values and practices of many of these immigrants can enrich our cultural life. Furthermore, immigrants also carry job skills and work attitudes that are needed for our economy.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The German values and practices we hold so dear can be relied upon to live peacefully with migrants' values. Furthermore, Germans have all that it needs to sustain a strong economy even together with immigrants.

From 2015 to 2020, almost 10 million migrants arrived to Germany. Germans and immigrants have different values and practices as well as different job skills and work attitudes. The integration of immigrants into our society represents too costly an investment and the costs to integrate them will never be compensated in the future.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The unacceptable values and practices of many of these immigrants are incompatible with our cultural life. Furthermore, immigrants also carry job skills and work attitudes that menace to permanently harm our economy.

From 2015 to 2020, almost 10 million migrants arrived to Germany. The German values and practices we hold so dear have to be preserved from migrants' values. Furthermore, Germans have all that it needs to sustain a strong economy even without immigrants.

--- Page 5 ---

In previous research, we asked all participants to rank the six statements you saw before according to how much they agreed with them. So, each participant indicated which statement they agreed the most. We are now going to ask you to guess how other participants answered these questions. We interviewed participants from the whole range of political orientation in Germany, aiming to achieve a representative sample of the German population with respect to political orientation.

You will receive **50 Lifepoints** on top of your base earnings for each of your guesses that is correct.

In the initial page, you told us that you plan to vote for Party {*Insert name of the party the participant is planning to vote for*} in the next election.

We would now like to ask you something about other participants who, like you, plan to vote for Party {*Name of participant's party*}.

9) Which statement do you think has been selected by most people planning to vote for Party {*Name of participant's party*} as the one with which they agreed the most?

10) Out of 100 participants planning to vote for Party {*Name of participant's party*}, how many do you think have selected the statement you have just indicated as the one they agreed the most with? Your answer may be any integer varying from 0 (no participant has selected that statement) to 100 (all participants selected that statement).

11) Which statement do you think has been the second most selected by people planning to vote for Party {*Name of participant's party*} as the one with which they agreed the most?

12) Out of 100 participants planning to vote for Party {*Name of participant's party*}, how many do you think have selected the statement you have just indicated as the second one they agreed the most with?

--- Page 6 ---

We would now like to ask you something about other participants who do not plan to vote for Party {*Name of participant's party*}

You will receive **50 Lifepoints** on top of your base earnings for each of your guesses that is correct.

13) Which statement do you think has been selected by most people planning to vote for a party

different from {*Name of participant's party*} as the one with which they agreed the most?

14) Out of 100 participants not planning to vote for Party {*Name of participant's party*}, how many do you think have selected the statement you have just indicated as the one with which they agreed the most?

15) Which statement do you think has been the second most selected by people not planning to vote for Party {*Name of participant's party*} as the one with which they agreed the most?

16) Out of 100 participants not planning to vote for Party {*Name of participant's party*}, how many do you think have selected the statement you have just indicated as the second one they agreed the most with?

--- Page 7 ---

Please consider again the six statements.

17) Which statement do you agree with the most?

We are now going to ask you to guess how other participants answered these questions. You will receive **50 Lifepoints** on the top of your base earnings for each one of your guesses that is correct.

18) Which statement do you think other participants agree with the most?

19) Out of 100 participants, how many do you think have selected the statement you have just indicated?

20) Which statement do you think other participants should agree with the most?

21) Out of 100 participants, how many do you think have selected the statement you have just indicated?

22) Which statement do you think other participants thought that others should agree with the most? (In other words, what is the most selected statement in the question above?)

23) Out of 100 participants, how many do you think have selected the statement you have just indicated?

--- Page 8 to 14 ---

REPEAT FOR VACCINATION

--- Page 15 ---

Before ending the questionnaire, we would like to ask you some final questions.

24) Please indicate the postal code of your current place of residence:

25) Were you born in Germany?

- Yes
- No
- Please indicate in which country/region you were born

26) Was your mother born in Germany?

- Yes
- No
- Please indicate in which country/region your mother was born

-Don't know

27) Was your father born in Germany?

- Yes
- No
- Please indicate in which country/region your father was born
- Do not know

28) Were your grandparents born in Germany?

- Yes
- No
- Please indicate the countries/regions where your grandparents were born
- Partly
- Don't know

29) What is your monthly net household income? That is, the total income of all family members living in the household after deducting taxes and contributions per month.

- under 900 Euro
- 900-1300 Euro
- 1301-1500 Euro
- 1501-2000 Euro
- 2001-2600 Euro
- 2601-3200 Euro
- 3201-4500 Euro
- 4501-6000 Euro
- more than 6001 Euro

30) What is your current employment situation? (If you have more than one job, please indicate only your main job):

- Employed (more than 450 euros subject to social security contributions)
- Self-employed
- 450 Euro employment
- Working without registration (i.e. without social security)
- Currently not employed and not looking for work
- Looking for work, but currently unemployed
- Student
- Retired
- Apprentice and trainee
- Other

31) Take a look at the categories below. Which category most closely matches your current or most recent job?

- Manager
- Professional
- Technician and associate professional
- Office worker and support staff
- Service and sales worker
- Skilled worker in agriculture, forestry and fisheries
- Craftsman and related occupation
- Plant and machine operators and assemblers
- Unskilled laborer
- Military

32) Do you belong to a religious community? If yes, which one?

- I do not belong to any religious community.
- Protestant
- Catholic

- Christian Orthodox
- Islamic
- Judaism
- Other

33) How often do you meet with migrants? Meant personal meetings and conversations, not mere greetings (e.g., at work or in the neighborhood):

- Daily
- Several times a week
- Once a week
- Once a month
- Less often than once a month
- Never

34) How often do you wear a face mask?

- Daily
- Several times a week
- Once a week
- Once a month
- Less often than once a month
- Never

- 35) Did you consult the website "Who supports what?"
- 36) Did you take part in a survey inquiring about your opinion on the same statements as the present survey last week?

37) Do you think that the researcher had any preference on what you should answer to this survey? If so, please state which preference

– – – – – –

--- Page 16 ---

Thank you. We are really grateful for your participation in this survey.

Finish the survey