



OPEN

## Evaluating the relationship between news source sharing and political beliefs

Sofía M. del Pozo<sup>1,2,5</sup>, Sebastián Pinto<sup>1,2,5</sup>, Matteo Serafino<sup>3</sup>, Federico Moss<sup>1</sup>, Tomás Cicchini<sup>1,4</sup>, Hernán A. Makse<sup>3</sup> & Pablo Balenzuela<sup>1,2</sup>✉

In an era marked by an abundance of news sources, access to information significantly influences public opinion. Notably, the bias of news sources often serves as an indicator of individuals' political leanings. This study explores this hypothesis by examining the news sharing behavior of politically active social media users, whose political ideologies were identified in a previous study. Using correspondence analysis, we estimate the Media Sharing Index (MSI), a measure that captures bias in media outlets and user preferences within a hidden space. During Argentina's 2019 election on Twitter, we observed a predictable pattern: center-right individuals predominantly shared media from center-right biased outlets. However, it is noteworthy that those with center-left inclinations displayed a more diverse media consumption, which is a significant finding. Despite a noticeable polarization based on political affiliation observed in a retweet network analysis, center-left users showed more diverse media sharing preferences, particularly concerning the MSI. Although these findings are specific to Argentina, the developed methodology can be applied in other countries to assess the correlation between users' political leanings and the media they share.

**Keywords** News sharing, Social media, Correspondence analysis, Political leanings.

The massive and growing use of the internet and digital platforms has undoubtedly brought about changes in the way we live, including the resources at our disposal and the habits we have incorporated<sup>1</sup>. The political leaning of users plays a key role in both sharing and accessing specific news items<sup>2,3</sup>. Indeed, the media bias observed in shared news on Twitter has been employed as a proxy for users' political ideologies<sup>4,5</sup>.

News sharing behavior on Twitter within the context of political elections was analyzed by Weaver et al.<sup>6</sup>. The authors used a bipartite network of users and news articles and analyzed the emergence of communities in the projection onto the news layer and identified their main features that explain these communities. A similar approach was addressed in<sup>7</sup> for Argentinian media outlets, comparing electoral and non-electoral periods. Within this framework, the researchers observed that groups of users on Twitter form based on their preferences for specific media outlets.

Inferring political leaning from available information of users on social media presents a challenge that has been addressed in different ways. For instance one approach involves utilizing hashtag information from tweets, as in<sup>8</sup>, to quantify the level of support for the impeachment of the former Brazilian president. Another example is seen in<sup>9</sup>, where hashtags were utilized to train a machine learning model aimed at predicting electoral trends during the 2019 Argentinian elections. Also, Cinelli et al., deduce users' political leanings based on the media bias of the news outlets they share on Twitter<sup>4</sup>. However, directly inferring user ideology from shared media bias remains a hypothesis that is awaiting validation.

An alternative method for uncovering the political preferences of social media users was introduced by Barbera et al.<sup>10,11</sup>, where they developed a Bayesian model. This model treats ideology as a latent variable, which can be inferred from observed connections among users, assuming that these connections adhere to the principle of homophily. Specifically, the authors estimated latent parameters by utilizing correspondence analysis on the adjacency matrix of users who follow political accounts on Twitter. A recent application of correspondence analysis is found in Flamino et al. (2023)<sup>12</sup>, where it was utilized to examine political polarization during the

<sup>1</sup>Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina. <sup>2</sup>Instituto de Física Interdisciplinaria y Aplicada (INFINA), CONICET - Universidad de Buenos Aires, Buenos Aires, Argentina. <sup>3</sup>Levich Institute and Physics Department, City College of New York, New York 10031, USA. <sup>4</sup>CONICET - Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Instituto de Cálculo (IC), Buenos Aires, Argentina. <sup>5</sup>Sofía M. del Pozo and Sebastián Pinto have contributed equally to this work. ✉email: balen@df.uba.ar

2016 and 2020 US presidential elections. The authors estimated individual positions from the users-influencers adjacency matrix. Similarly, Falkenberg et al. employed a comparable method to estimate latent opinions within the online discourse surrounding measures to address climate change<sup>13</sup>.

In this work, correspondence analysis is employed not to infer users' ideology, but to quantify their media preferences based on the news articles they share. We compare these preferences with their political ideology, which was previously inferred through a machine learning model developed and validated in Zhou et al., based only on the hashtags supporting one of the two coalitions that contested the elections that year<sup>9</sup>. Specifically, we use users' connections to Argentine news articles to deduce preferences for media outlets within a latent space, achieved by conducting correspondence analysis of the user-media matrix. The comparison of both metrics allows us to measure how the political ideology of the users is related to the preference for the media they share. Additionally, we delve into the emergence of community structures within the retweet network (excluding retweets containing links to news articles) to discern whether user interactions are driven by political leaning or media outlet preferences.

This work aims to contribute to the quantitative analysis of the relationship between political ideology and media preferences, continuing the line of previous studies such as<sup>14</sup> and<sup>15</sup>.

## Background

### Argentinian context

In this section, given our focus on Argentina, we explore the country's political and media landscape during the 2019 presidential election campaign to provide essential contextualization.

Even though today Argentina is governed by Javier Milei of a libertarian party<sup>16,17</sup>, over the last decade, Argentina's political landscape has been dominated by two major coalitions: one, a center-left coalition (CL) led by Cristina Fernández de Kirchner, known as *Frente de Todos*, and the other, a center-right coalition (CR) led by Mauricio Macri, referred to as *Juntos por el Cambio*. Cristina Kirchner held the presidency in Argentina during the periods of 2007–2011 and 2011–2015, while Mauricio Macri served as president from 2015 to 2019<sup>18–20</sup>. During the 2019 elections, the center-left coalition presented Alberto Fernández and Cristina Fernández de Kirchner as their candidates. Meanwhile, the center-right coalition sought a second term for Mauricio Macri as president, with Miguel Ángel Pichetto as his vice-presidential candidate.

National elections in Argentina comprise two obligatory phases: the primary election, known as PASO (which stands for *Primarias, Abiertas, Simultáneas y Obligatorias* in Spanish, translating to Open, Simultaneous, and Obligatory Primaries in English), and the general election. In 2019, these events occurred on August 11th and October 27th, respectively. Additionally, if the results of the general election necessitate it, a third round, referred to as a *ballotage*, may also be conducted.

Regarding the media landscape, the digital media scene in Argentina is primarily characterized by three major players: *Infobae*, *Clarín* and *La Nación*, each boasting approximately 20 million unique users in 2020, as reported by Comscore data<sup>21</sup>. In particular, in 2019, *Infobae*, *Clarín*, *La Nación* and other media outlet, *Todo Noticias (TN)*, accounting for 80% of the total reported online visits<sup>22</sup>. These media outlets are relevant through this work. Following closely are a second tier of media outlets characterized by audience numbers ranging from 6 to 13 million unique visitors. Prominent among this group are *Página 12*, *Ámbito Financiero* and *El Destape Web*, among others.

In Argentina, a pronounced polarization has been reported through the distinct ideological orientations of the country's primary media outlets<sup>7,23</sup>. For instance, *Página 12* is recognized as a left-of-center broadsheet newspaper<sup>24</sup>, *Infobae* is considered a center-left outlet<sup>25</sup>, while *Clarín* is considered a center-right tabloid<sup>26</sup>, and *La Nación* is characterized as a center-right broadsheet newspaper<sup>24,27</sup>.

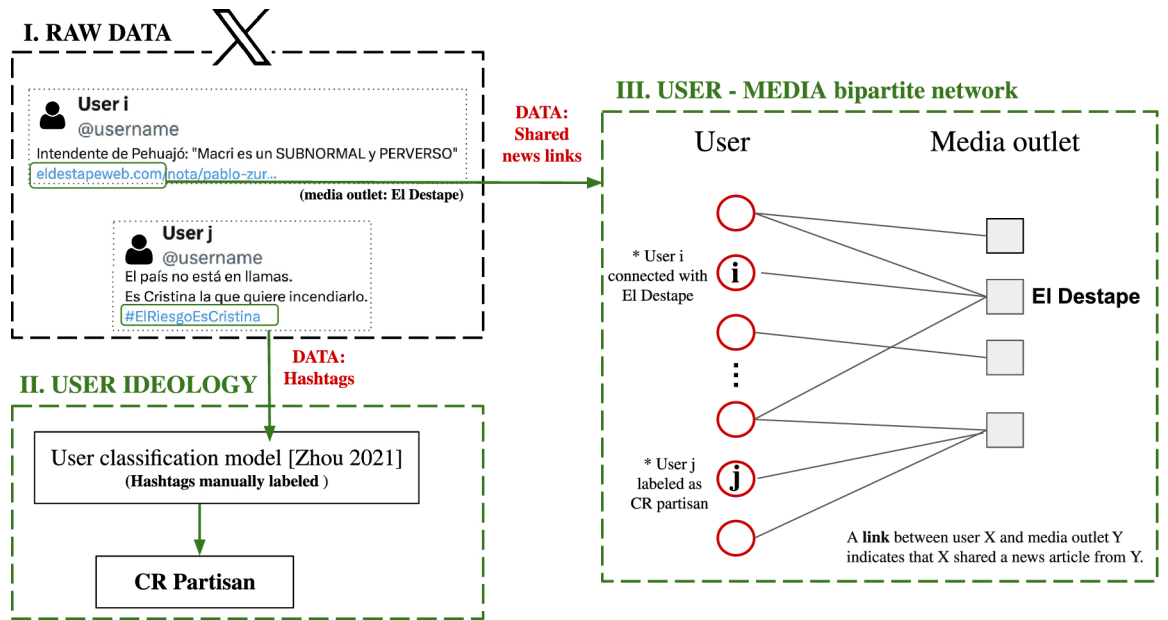
## Data and methods

In this section, we present the data description (Section “Data description”), which outlines the dataset utilized in our study, followed by the explanation about how we use this data and its study through correspondence analysis (Section “Methods”).

### Data description

In this research, we employed a pre-existing Twitter dataset<sup>9</sup> which comprised tweets collected between March 1, 2019, and August 27, 2019. The details of this dataset can be found in the “Appendix”. From this dataset, we filtered for all types of tweets, including tweets, retweets, and quotes, that exclusively contained external URLs linking to Argentinian media outlets listed in the ABYZ News Links Guide<sup>28</sup>. From this list, we selected 17 media outlets, ranking the outlets in descending order by the number of times they were shared on Twitter and selecting those with broad recognition and influence not only on this platform, but also across Argentine media as a whole, as reported in sources such as<sup>22</sup>. Given this filter, we first obtained a dataset encompassing the activity of 123,180 users, who collectively generated 1,039,281 tweets, sharing 66,982 unique news articles. Secondly, we incorporated data concerning individuals' voting intentions, which had been determined previously using a model detailed in<sup>9</sup>. In this paper, the authors developed a method to infer political preference of Twitter users by implementing a dynamic classification model based on the balance of tweets in favor of each of the contending coalitions. Such a model, described in detail in<sup>9</sup> provides a temporal label to a subset of 17,349 users as supporters of the center-left (CL) candidates (Fernández-Fernández) and 15,361 individuals as sympathizers of the center-right (CR) coalition (Macri-Pichetto). Supporters of the CL coalition shared 19,276 news articles, while those leaning towards the CR coalition shared 10,135.

Figure 1 depicts the methodology followed to organize the large set of tweets sharing news of politically tagged users into a bipartite network of users-media outlets.



**Fig. 1.** Methodology pipeline. (I). Raw data. Example of original data from Twitter (now X), with a tweet sharing a URL to a news article at the top and another one with a political hashtag at the bottom. (II) User ideology. Hashtags were used to train a logistic regression model to classify tweets as supportive of either candidate. Users are assigned to the candidate for whom they demonstrate the highest number of supportive tweets (further details in<sup>9</sup>). (III) User-media. The news URLs in the tweets are used to identify the media outlet. A bipartite network of users and media outlets is then created. For example, user *i* is linked to the media outlet *El Destape* because this user shared a news article from it.

### Methods

We organize the data in a user-media matrix, where each row is associated with a user and each column is associated with a media outlet. The components of the matrix represent the number of times a given user shares an article from a specific media outlet. By applying correspondence analysis (detailed below), we calculate what we refer to as the Media Sharing Index (MSI) in this context. This index positions users within a latent space reflecting their preferences for specific groups of media outlets. Simultaneously, it places media outlets within the same latent space, determined by the average preferences of their audience. Essentially, users closer in this space tend to share similar media outlets, indicating comparable preferences in media sharing. Similarly, media outlets situated closely in this latent space imply shared usage by a similar set of users. We compare the MSI with the political leaning of those users identified by<sup>9</sup> and their position within the interaction network.

#### Correspondence analysis

Following previous studies<sup>10,12</sup> that propose a methodology for inferring user coordinates in a latent space of social media based on correspondence analysis<sup>29</sup>, we begin by establishing a bipartite network denoted as  $G = (U, V, E)$ , where  $U$  represents the set of users,  $V$  denotes the news outlets, and  $E$  stands for the edges in the graph. The corresponding adjacency matrix associated with this network is denoted as  $Y$ . The element  $y_{ij}$  represents the number of times user  $i$ , with  $i \in U$  shares news from a media outlet  $j$ , with  $j \in V$ . The main difference between our implementation and the seminal proposal<sup>10</sup> is that the bipartite network here is based on the content of users' tweets, rather than explicit network connections (e.g., following-followers relations or retweets).

The matrix  $Y$  is converted into the correspondence matrix  $P$  by dividing each element by its grand total  $P = Y / \sum_{ij} y_{ij}$ . The element  $p_{ij}$  represents the probability of finding an event in which user  $i$  shares an article from media outlet  $j$ . From matrix  $P$ , the matrix of standardized residuals  $S$  is computed as:

$$S = D_r^{1/2} (P - r c^T) D_c^{1/2}$$

where vectors  $r$  and  $c$  are defined as  $r_i = \sum_j p_{ij}$  and  $c_j = \sum_i p_{ij}$ . The element  $r_i$  represents the likelihood that user  $i$  shares an article from any media outlet. Conversely,  $c_j$  represents the probability of media outlet  $j$  being shared by any user. The elements of outer product  $r c^T$  ( $r_i c_j$ ) can be interpreted as the probability of user  $i$  sharing a media outlet  $j$  given a null model where only the activity of user  $i$  and the frequency of with which media outlet  $j$  is shared matter. By defining diagonal matrices  $D_r = \text{diag}(r)$  and  $D_c = \text{diag}(c)$ , we can express the elements of  $S$  as follows:

$$s_{ij} = \frac{p_{ij} - r_i c_j}{\sqrt{r_i c_j}}$$

This expression can be interpreted as the deviation, measured in standard units, of  $p_{ij}$  from a null model where users and media outlets are independent.

In order to compute the MSI for each user, we first perform singular value decomposition on  $S$ , that is:

$$S = UD_{\alpha}V^T$$

where  $UU^T = V^TV = I$  and  $D_{\alpha}$  is a diagonal matrix with the singular values on its diagonal. The Media Sharing Index for the user  $i$ ,  $MSI_i$ , is then identified by the standard row coordinates by projecting only over the first singular component:

$$MSI_i = (D_r^{1/2}U)_i$$

and finally normalizing these values to have zero mean and a standard deviation equal to 1. As such, users with similar values of MSI imply that they share a similar set of media outlets. In particular, if user sharing behavior is driven by two distinct groups of media outlets, as shown in<sup>7</sup>, we would expect to observe a group of individuals with  $MSI_i > 0$  and another with  $MSI_i < 0$ .

Finally, we define the MSI for media outlets as the weighted average of the MSI of the users, weighted by the number of times user  $i$  shares media  $j$ :

$$MSI_j = \frac{\sum_i y_{ij}MSI_i}{\sum_i y_{ij}} \quad (1)$$

The interpretation of  $MSI_j$  is analogous to one provided for  $MSI_i$ : media outlets with  $MSI_j > 0$  will have a very different set of users who share its articles compared to media outlets with  $MSI_j < 0$ .

## Results

### Media sharing index

As described above, we construct the bipartite adjacency matrix  $Y$ , where each element  $y_{ij}$  represents the number of news articles from media outlet  $j$  shared by user  $i$ . We apply correspondence analysis to this matrix to calculate the Media Sharing Index (MSI), as discussed in Section “Methods”. For simplicity, we focus on the primary 12 media outlets, excluding those that are shared by only a few users and where the majority of the shared articles come from the outlets themselves. This reduces the dataset to 59,874 unique news articles (approximately 88% of the total unique news in the dataset) and 120,626 users (about 97% of the users in the original dataset). These users originate from a total of 1,015,380 tweets containing links to one of these 12 main outlets, which constitutes about 98% of the original tweet volume.

Figure 2 illustrates the probability density function of the MSI for users who share articles from at least one of the primary 12 media outlets. This figure reveals the emergence of a bimodal distribution in the Media Sharing Index. Unimodality is rejected with a p-value practically equal to zero ( $p < 0.001$ ), as determined by the Dip test<sup>30,31</sup>. This bimodal distribution reflects the preferences of users sharing content from two distinct groups of media outlets. Specifically, a majority of users share news from a group of outlets that includes Clarín, La Nación, Todo Noticias, among others, with an MSI close to +1. Conversely, a minority group prefers outlets such as El Destape, Página 12, and Minuto Uno, with an MSI close to -1.

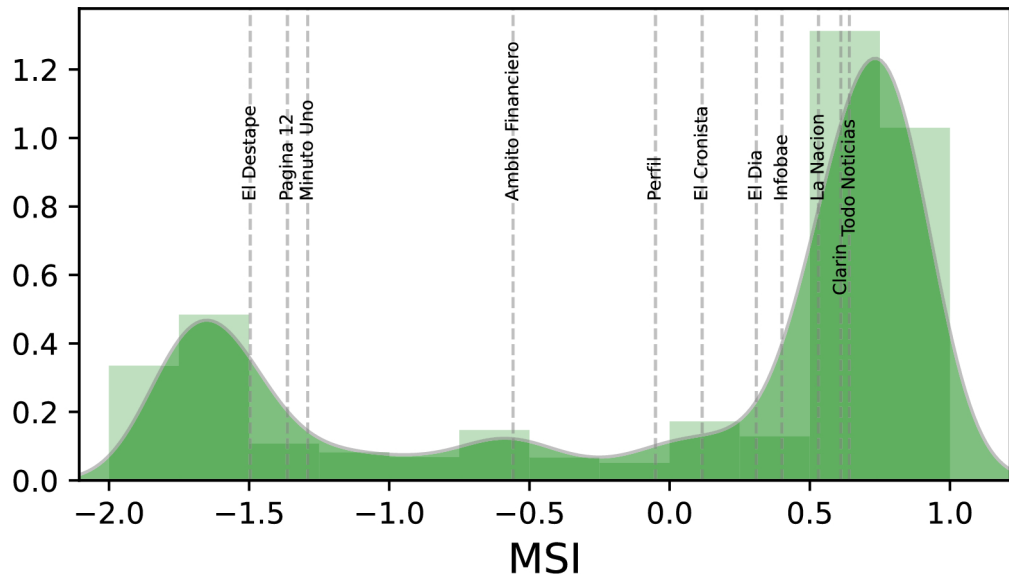
Given that Clarín and La Nación are considered center-right outlets<sup>26,27</sup>, and Página 12 is viewed as a left-of-center broadsheet<sup>24</sup>, we can associate the Media Sharing Index with media bias along the left-right political dimension. These results prompt the question: do left-leaning users predominantly share news from left-leaning newspapers, aligning with their beliefs? Or is news sharing behavior independent of their political preferences? In other words, can the media bias reflected in the news users share serve as a proxy for their political ideology? This is the question we aim to answer in this paper.

### Media sharing and political preferences

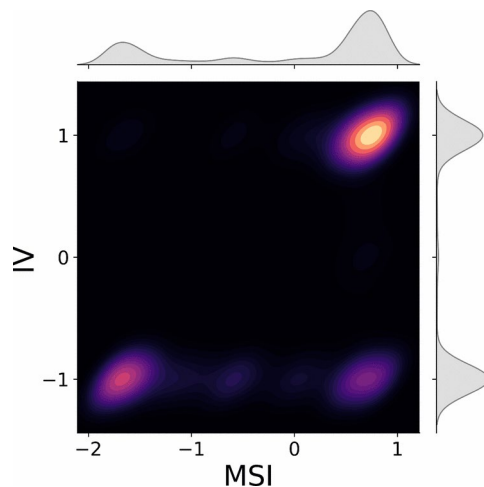
In this section, we explore the potential link between the sharing preferences of users, as observed in Fig. 2, and their underlying political polarization. As described in Section “Data description”, the dataset used in this analysis is the same one employed by Zhou et al.<sup>9</sup> to infer the political preferences of Twitter users based on their posts. The labels assigned to users in<sup>9</sup> are dynamic, allowing for a more nuanced description of a user’s ideology. Therefore, we define the ideology valence of a user  $i$  ( $IV_i$ ) as:

$$IV_i = \frac{\#CR_i - \#CL_i}{\#CR_i + \#CL_i} \quad (2)$$

Here,  $\#CR_i$  and  $\#CL_i$  denote the number of times user  $i$ , identified as a center-right or center-left partisan at that time, shared a news article, respectively. The sum  $\#CR_i + \#CL_i$  represents the total number of news articles shared by user  $i$ . By this definition,  $IV_i = 1$  indicates a user consistently labeled as center-right,



**Fig. 2.** Probability density of the Media Sharing Index (MSI). This graph shows the probability density of the MSI for users and the 12 main media outlets. Radio Mitre is excluded from this display as it is a positive MSI outlier. Histograms have been smoothed using a Gaussian kernel with a bandwidth of 0.15. The grey lines indicate the positions of the media outlets.



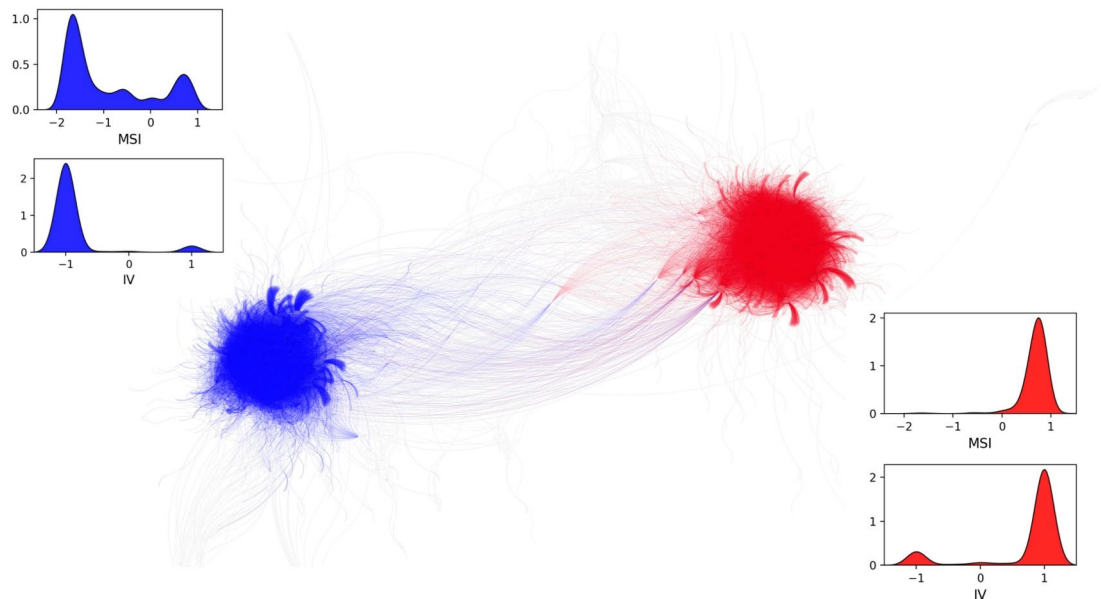
**Fig. 3.** Joint Probability Density of MSI and IV. This figure illustrates the relationship between the media sharing behavior of users and their ideological leaning.

representing a pure CR partisan. Conversely,  $IV_i = -1$  indicates a pure CL partisan.  $IV$  is only defined for users with a label assigned in<sup>9</sup>.

The relationship between the ideology valence ( $IV$ ) and the Media Sharing Index ( $MSI$ ) is illustrated in Fig. 3. As previously mentioned, the  $MSI$  can serve as an indicator of media bias along the left-right political dimension. However, Fig. 3 reveals that media-sharing behavior on social media cannot always be directly associated with users' ideological leanings. This figure demonstrates that while center-right users exhibit a distinct media sharing profile, with a preference for sharing outlets aligned with their political leanings, center-left users share news from media sources corresponding to both peaks observed in the probability density of the  $MSI$  in Fig. 2.

A possible interpretation of the asymmetric behavior observed between ideological groups is that media outlets with  $MSI > 0$ , such as Clarín, La Nación, and Infobae, are also predominant players in the Argentine media landscape<sup>21,22</sup> (see Section “Argentinian context”). These outlets' extensive reach and influence may contribute to the observed sharing patterns among center-left users at the aggregate level.





**Fig. 4.** Retweet network. The two largest communities detected by the Louvain algorithm are displayed (account for 75% of the entire network). Histograms illustrate the distribution of the MSI and IV for each community. Based on these distributions, the red community can be associated with a center-right political leaning, while the blue community can be associated with a center-left leaning.

### Retweet user networks

In this section, we analyze user interactions by constructing the retweet network. We specifically focus on the relationship between the emerging community structure, the Media Sharing Index (MSI), and the users' political preferences. By definition, the retweet network is directed and weighted. The direction of the links indicates the flow of information (i.e., arrows point from a retweeted user to the user who retweets), and the weights reflect the number of retweets between users. This network comprises 114,673 nodes, representing approximately 90% of the users described in Section “Data description”. The remaining 10% are users who did not retweet or were not retweeted by any other users during the analyzed period. Additionally, there are 12,993,644 edges, corresponding to the total number of retweets among users in this network. It is important to note that this network was constructed only considering retweets that do not include links to news articles, meaning it contains no data used in the calculation of the MSI.

In Fig. 4, the two main communities within the retweet network are shown, identified using the Louvain algorithm<sup>32</sup>. Although the algorithm detects 440 communities in total, the two largest communities account for 75% of the entire network, with nearly an equal number of nodes in each community. Figure 4 also reveals a highly modular structure of the retweet network, with a modularity score of approximately  $Q \sim 0.48$ .

Histograms in Fig. 4 reveal the profile of each community in relation to the Media Sharing Index and the Ideological Valence. Each community correlates strongly with a distinct ideological position: The red community in Fig. 4, representing 38% of the network, aligns with the center-right, while the blue community, representing 37% of the network, aligns with the center-left. The fact that the histograms for the Ideological Valence index show a clear peak for both communities suggests that the community structure emerging from the retweet network is a reliable proxy for the ideological positions of its members.

The association found between the community label and the ideological position of users supports the reproduction of the results discussed in Section “Media sharing and political preferences” at a community level: the center-right community exhibits media sharing behavior favoring center-right media outlets, as depicted in the blue histogram of the MSI. Conversely, the center-left community displays a less biased MSI distribution, indicating that this community shares content from both center-left media outlets and those biased towards the opposite ideological spectrum. As mentioned in Section “Media sharing and political preferences”, the observed diversity in sharing patterns among the CL group may be linked to the prominence of CR media outlets in the overall media ecosystem. This prominence is reflected in audience metrics, as we mentioned in Section “Argentinian context”.

### Discussion and conclusions

In this work, we describe the collective news sharing behavior of thousands of Twitter's users by a coordinate in a latent space which we named Media Sharing Index (MSI). This is obtained by performing correspondence analysis to the bipartite network of users - media outlets which emerge from the information of news articles shared by users.

The MSI metric enables us to set up a scale that describes the preferences of social media users in news sharing behavior. In this work, we specifically observe a bimodal distribution of the MSI, where the two clearly

defined peaks can readily be linked to two distinct groups of media outlets. In the case of Argentina, these two groups are exemplified by Clarín, La Nación, Infobae, and Todo Noticias, among others, on one side, and on the other side, Página 12 and El Destape. These six media outlets stand out as the most widely shared on social media. Given that Clarín and La Nación are considered as center-right outlets<sup>26,27</sup> and Página 12 as a left-of-center broadsheet newspaper<sup>24</sup>, we can associate the Media Sharing Index as a measure of media bias in the left-right political dimension.

The strength of the Media Sharing Index lies in its ability to quantify news sharing behavior among diverse media outlet groups. Coupled with the inference of users' political leaning derived from the machine learning model outlined in<sup>9</sup>, we can explore the correlations between these dimensions and address questions such as: *Can the bias of the shared media outlets be used as a proxy of user's ideology?*

Our analysis reveals distinct patterns: while the CR group predominantly shares news from media that align with their ideological stance, the CL group exhibits a broader range of media sources in their sharing behavior. This observed diversity could likely be influenced by the prominent role of CR media outlets within the broader media ecosystem in Argentina, as we stated in section 2.1 and can be observed in<sup>21,22</sup>. An ideological asymmetry in news preferences and exposure has been also documented in other contexts. In the US, studies have examined the asymmetry between right-wing and left-wing groups at the media ecosystem level<sup>33,34</sup> and the preference for news exposure in social media. However, the users' behavior is not similar to what we observe in this work. For instance, in<sup>35</sup>, the authors analyze the reinforcing effects of liberal and conservative media on political beliefs during the 2016 US election, finding that conservative beliefs contribute more to a conservative media echo chamber than liberal beliefs do to a liberal one. Similarly,<sup>36</sup> explores ideological segregation in political news on Facebook during the 2020 US election, revealing that conservatives consume a disproportionately large amount of news, while liberal sources are less prevalent. Research conducted in countries outside the US also suggests that ideology has a stronger influence on right-wing users compared to left-wing users when it comes to news-sharing behavior<sup>37</sup>.

The most significant limitation of this study is that it is restricted to Argentina during the 2019 electoral period. It is essential to examine how these results vary in other electoral periods in Argentina and what the results would be in different countries. Nonetheless, we have developed a methodology here that can be extrapolated to all these scenarios in future studies.

## Data availability

Twitter data analyzed in this work are provided according to its terms and are available at <https://osf.io/u29tk/>.

## Code availability

Analytical codes are available in <https://github.com/spinto88/MediaSharingIndex>.

## Appendix: Original dataset

As mentioned in Section “Data description”, we employed a pre-existing Twitter dataset<sup>9</sup> comprising tweets collected between March 1, 2019, and August 27, 2019. This dataset was obtained through the Twitter Academic API (at the time of acquiring the data, this API allowed full access to the complete history of the social media platform) by searching for the following queries: *Pichetto, Espert, victoria, Bregman, argentina, Voto, Kirchner, Propuesta Republicana, Moreno, Peron, Sola, kirchnerismo, Massa, UCR, PASO, Alternativa Federal, Justicialista, PRO, Movimiento Socialismo, Rossi, #ArgentinaVota, Union Civica Radical, fuerza, Votar, Vidal, Alfonsín, alferdez, kicillof, Scioli, vamos, Frente Izquierda, Donda, miguelpichetto, CFK, Vote, kicillofok, Alberto Fernandez, Urtubey, elecciones, sergiomassa, ganamos, Peronista, CFKArgentina, peronismo, #EleccionesArgentina, PJ, Lavagna, Partido Justicialista, mauriciomacri, Consenso 19, Cristina, Lousteau, Cambiemos, macrismo, mariuvidal, Capitanich, apoyo, Macri*. Exclusively, tweets written in Spanish were selected.

The full dataset analyzed comprised around 170 million tweets and about 3.5 million users. Although the dataset is massive, given the keywords used to obtain it, the dataset may include users with a certain bias towards tweeting about politics and, therefore, must be considered with caution when evaluating its representativeness of the Argentinian population. On the other hand, Zhou et al.<sup>9</sup> reported an average classification per day of 114,653 tweets as CR supporters and 96,576 as CL supporters (MP and FF, respectively, in<sup>9</sup>), which indicates a balanced classification.

All of these numbers are drastically reduced when considering only those tweets with URLs linking to Argentinian outlets, as stated in Section “Data description”.

Received: 28 May 2024; Accepted: 19 December 2024

Published online: 02 January 2025

## References

- Graham, M. and Dutton, W.H. *Society and the Internet: How Networks of Information and Communication are Changing Our Lives*. Oxford University Press, (2019).
- Osmundsen, M. et al. Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *Am. Political Sci. Rev.* **115**(3), 999–1015 (2021).
- Eady, G. et al. News sharing on social media: Mapping the ideology of news media content, citizens, and politicians. In: (2020).
- Cinelli, M. et al. The echo chamber effect on social media. *Proc. Natl. Acad. Sci.* **118**(9), e2023301118 (2021).
- An, J. et al. Sharing political news: The balancing act of intimacy and socialization in selective exposure. *EPJ Data Sci.* **3**, 1–21 (2014).
- Weaver, I. S. et al. Communities of online news exposure during the UK General Election 2015. *Online Soc. Netw. Media* **10**, 18–30 (2019).

7. Cicchini, T. et al. News sharing on twitter reveals emergent fragmentation of media agenda and persistent polarization. *EPJ Data Sci.* **11**(1), 48 (2022).
8. Cota, W. et al. Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Sci.* **8**(1), 35 (2019).
9. Zhou, Z. et al. Why polls fail to predict elections. *J. Big Data* **8**(1), 1–28 (2021).
10. Barberá, P. et al. Tweeting from left to right: Is online political communication more than an echo chamber?. *Psychol. Sci.* **26**(10), 1531–1542 (2015).
11. Barberá, P. Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Anal.* **23**(1), 76–91 (2015).
12. Flamino, J. et al. Political polarization of news media and influencers on Twitter in the 2016 and 2020 US presidential elections. In: *Nature Human Behaviour* (2023), pp. 1–13.
13. Falkenberg, M. et al. Growing polarization around climate change on social media. *Nat. Climate Change* **12**(12), 1114–1121 (2022).
14. Calvo, E. et al. Winning! election returns and engagement in social media. *Plos one* **18**(3), e0281475 (2023).
15. Aruguete, N., Calvo, E. & Ventura, T. News by popular demand: Ideological congruence, issue salience, and media reputation in news sharing. *Int. J. Press/Politics* **28**(3), 558–579 (2023).
16. *Presidencia de la Nación Argentina*. <https://www.argentina.gob.ar/presidencia> Accessed: 19th March 2024.
17. *Argentina Government and Media Profile*. <https://mediabiasfactcheck.com/argentina-government-and-media-profile/> Retrieved from: 19th March 2024.
18. Murillo, M. V. & Zarazaga, R. Argentina: Peronism returns. *J. Democr.* **31**, 125 (2020).
19. Benzra, K. Neoliberalism in Crisis. *J. Latin Am. Cult. Stud.* **29**(3), 461–467 (2020).
20. Cantamutto, F. Kirchnerism in Argentina: A populist dispute for hegemony. *Int. Crit. Thought* **62**, 227–244. <https://doi.org/10.1080/021598282.2016.1172325> (2016).
21. *Todo Medios (Comscore data)*. <https://www.totalmedios.com/nota/54090/infobae-lidero-en-un-agosto-marcado-por-las-paso-y-c-on-crecimiento-de-todos-los-sitios>. Accessed 14 Nov. 2023.
22. *Digital News Report*. <https://www.digitalnewsreport.org/survey/2019/argentina-2019/>. Retrieved from: 27th September 2024.
23. Newman, N. et al. Digital news report 2023. In: *RISJ: Reuters Institute for the Study of Journalism* (2023).
24. Bonner, M. D. Media and punitive populism in Argentina and Chile. *Bull. Latin Am. Res.* **37**(3), 275–290 (2018).
25. *Infobae Bias*. <https://mediabiasfactcheck.com/infobae/>. Accessed: 19th March 2024.
26. *Clarín Bias*. <https://mediabiasfactcheck.com/clarin-bias/>. Accessed: 19th March 2024.
27. *La Nación Argentina Bias*. <https://mediabiasfactcheck.com/la-nacion-argentina-bias/>. Accessed: 19th March 2024.
28. *ABYZ Web Links Inc.* Accessed 14 Nov. 2023. : <http://www.abyznewslinks.com/argen.htm>.
29. Greenacre, M. J. Correspondence analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2**(5), 613–619 (2010).
30. Hartigan, J. A. and Hartigan, P. M. The dip test of unimodality. In: *The Annals of Statistics*, pp. 70–84 (1985).
31. *Diptest Python Library*: : <https://pypi.org/project/diptest/>.
32. Blondel, V. D. et al. Fast unfolding of communities in large networks. *J. Stat. Mech.: Theory Exp.* **2008**(10), 10008 (2008).
33. Benkler, Y. & Faris, R. and Roberts, H. (Manipulation, disinformation, and radicalization in American politics. Oxford University Press, Network propaganda, 2018).
34. Jamieson, K. H. and Cappella, J. N.. *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. Oxford University Press, (2008).
35. Hmielowski, J. D., Hutchens, M. J. & Beam, M. A. Asymmetry of partisan media effects?: Examining the reinforcing process of conservative and liberal media with political beliefs. *Political Commun.* **37**(6), 852–868 (2020).
36. González-Bailón, S. et al. Asymmetric ideological segregation in exposure to political news on Facebook. *Science* **381**(6656), 392–398 (2023).
37. Aruguete, N., Calvo, E., and Ventura, T. *News by Popular Demand: Ideology, Reputation, and Issue Attention in Social Media News Sharing*. (2019).

## Author contributions

SMdP, SP, and MS (ORCID: 0000-0002-7907-1375) were responsible for collecting the raw data. SMdP and SP developed the computational code utilized consistently throughout the paper. TC and FM made contributions to the statistical analysis. PB and HAM conceptualized the research. All authors engaged in discussions about the results and collaborated on the development of the manuscript.

## Funding

HAM and MS were supported by NSF Grant No. 2214217. PB, SMdP and SP were supported by PICT-2020-SERIEA-00966.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to P.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024