Network Science (2026), **14**, e10, 1–27
doi:10.1017/nws.2026.10028

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Enhancing joint ideal point estimation strategies with Twitter social networks and text data

Yicheng Shen[1], Christopher Bail[2] and Alexander Volfovsky[3]

[1]Biostatistics, Johns Hopkins University, USA, [2]Sociology, Duke University, USA, and [3]Statistical Science, Duke University, USA
Corresponding author: Alexander Volfovsky; Email: alexander.volfovsky@duke.edu

Abstract
Analyses of online social interactions on major platforms have become central to evaluating the ideological leanings of users. Social scientists often represent individuals' policy preferences as ideal points in a low-dimensional space. We first demonstrate that standard ideal-point estimation approaches can be temporally inconsistent by examining the evolution of ideological positions for political elites and non-elite Twitter users from 2017 to 2022. To address this instability and to improve overall measurement quality, we augment the social-network data conventionally used for ideal-point estimation with unstructured text data widely available on social media. Our results indicate that this data-augmented approach improves identification at the extreme ends of the ideological spectrum and yields robust estimates consistent with those from more complex word-embedding techniques, while offering greater computational efficiency and interpretability. This framework can enhance future studies that measure political ideology using behavioral data and inform broader discussions about integrating textual and network information.

Keywords: social network; text data; ideal point estimation; correspondence analysis; political polarization

1. Introduction

Social media platforms have become essential tools for political communication, mobilization, and engagement in recent decades (Tucker et al., 2018). With vast user communities generating real-time data, platforms such as Twitter,[1] Facebook, Reddit, and YouTube provide rich sources of behavioral information, including social networks and text, for analyzing political discourse around the world (Storck, 2011; Grimmer and Stewart, 2013; Effing et al., 2011; De Francisci Morales et al., 2021). One important application of these data is the measurement of political ideology among elites as well as the broader online public (Golbeck and Hansen, 2014; Barberá and Rivero, 2015; Mosleh and Rand, 2022).

Most approaches to studying ideological leanings rely on a single mode of data. While methods vary (see Section 2), they generally follow a common recipe: take a social network or a text corpus at a single point in time, embed the data, and project the embedding onto one or two latent dimensions (e.g., Figure 1). These approaches have performed well on massive volumes of network and text data (Himelboim et al., 2017; Vafa et al., 2020). A prominent example is Barberá et al. (2015), who introduced a computationally tractable network-based algorithm that has since become standard. However, although such methods typically estimate static positions, their temporal consistency and their ability to distinguish extremists from moderates remain unclear.

Meanwhile, high-profile political accounts attract large numbers of followers from across the spectrum, not necessarily because users agree with their views, but because they wish to monitor

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.
https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

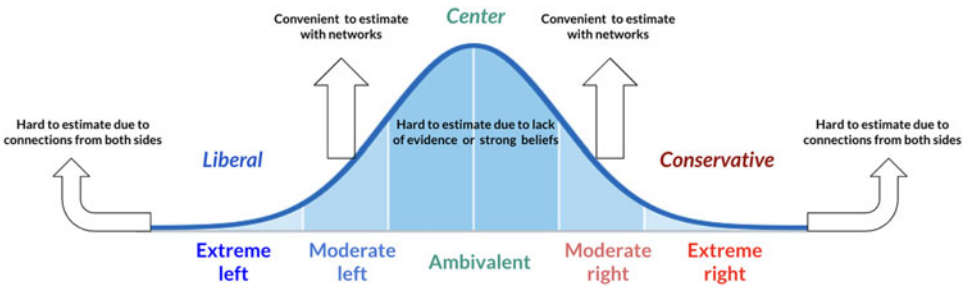


Figure 1. The ideological spectrum often serves as a tool for understanding and categorizing political positions. Usually, it is more difficult to accurately estimate individuals who either have sparse data and are ambivalent about their political beliefs or, conversely, have popular connections from both sides and hold extreme political beliefs.

those messages. Similarly, controversial politicians may exhibit apparent shifts in estimated ideal points as their audiences change, even when their views remain stable. Consequently, estimating ideal points using only network information can be inconsistent for highly visible elites and for users with cross-cutting connections on social media (Figure 1).

Inspired by the work of Bail et al. (2018) on exposure to opposing views and polarization, we connect the network-based approach of Barberá et al. (2015) with textual analysis to estimate ideology for both political elites and ordinary users over time. We propose a data-augmentation strategy that jointly leverages network and text data on a latent scale to improve reliability in practice. Our data include networks and posts from the Twitter accounts of 933 political elites and 1,176 experimental participants with self-reported party alignments, observed over five years. We argue that although networks may change for many reasons, language provides a stable source of ideological information and, as we show in simulations, facilitates ideal-point estimation.

Our results first verify that while network-based approaches provide reasonable estimates of ideological positions within a moderate left–right spectrum at a single point in time, these estimates can become inconsistent as the composition of elites' and users' networks evolves over time. Incorporating text information from users' tweet content substantially improves the estimation of ideal points, particularly by better identifying political elites at the ideological extremes. Furthermore, when comparing across estimators, we find that the text-only approach based on word embeddings captures meaningful linguistic variation but lacks sufficient structural information to accurately distinguish moderate users. By contrast, our data-augmented (DA) estimator, which integrates both network and text data, achieves the most distinct and stable separation of ideological groups, outperforming single-source approaches.

The remainder of this paper proceeds as follows. First, we review background on the ideal-point estimation problem, with a particular focus on correspondence analysis (CA) and its limitations in practical applications. We then describe our motivating data, introduce the proposed joint modeling framework, and present supporting simulation evidence in Section 3. In Section 4, we apply our data-augmented estimation method and discuss ideal-point results for both American political elites and ordinary Twitter users. Finally, we conclude by summarizing our contributions and outlining directions for future work.

2. Estimating political ideology with behavioral data

Social media platforms serve as a rich and dynamic source of data for studying political opinions, attitudes, and behaviors in democratic polities (Garton et al., 1997; Kavanaugh et al., 2011). With rapid advances in technology and Internet accessibility, people around the world increasingly turn to social media to read news, express ideas, build connections, and engage in discussions about political and social issues (Wike et al., 2022). This widespread engagement provides substantial

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

opportunities for researchers across all disciplines to study political behavior using large-scale observational data.

In late 2017, Bail et al. (2018) conducted a large-scale experiment involving more than 1,600 Twitter users who identified as either Democrats or Republicans. The study sought to understand how exposure to different political perspectives on social media might influence participants' political ideologies and, ultimately, patterns of political polarization in the United States. Participants were surveyed at the beginning and end of the experiment to assess treatment effects, and the study documented notable backfire effects, whereby exposure to opposing political views intensified partisan polarization. Nevertheless, these posttreatment ideological shifts are difficult to quantify formally and objectively. The lack of a standard measurement introduces both practical challenges and substantive concerns in analyzing outcomes. Our study is motivated by the need for a reliable and consistent approach to estimating ideological positions.

2.1 Understanding ideology through social media

Over the past several decades, the scholarly study of ideal-point estimation has gained significant traction, drawing contributions from multiple disciplines. These efforts have helped formalize ideal points as quantitative representations of latent preferences within a policy or issue space (Martin and Quinn, 2002). This conceptualization is particularly salient because it aligns with how political dimensions, such as the left–right spectrum, are commonly understood in spatial terms (Shor and McCarty, 2011; Laver, 2014).

Advances in statistical methodology have enabled researchers to model ideal points in multidimensional spaces using a variety of approaches, including logistic regression (Wasserman and Pattison, 1996), covariate-based mixed-effects models (Bailey, 2001), item response theory (Clinton et al., 2004; Martin and Quinn, 2002), multinomial models (Goplerud, 2019), and discriminant analysis (Takane et al., 1987). At the same time, concerns regarding the strict assumptions and complexity of traditional models have also motivated the development of nonparametric alternatives (Tahk, 2018).

With increasingly sophisticated strategies for estimating ideal points from public records and known network ties, researchers have gained deeper insight into the ideological spectrum of political elites. While congressional polarization has become a stylized fact in the United States (Hill and Tausanovitch, 2015), the ideal points of the general public have attracted growing interest, especially with the proliferation of large-scale data on political opinions (Imai et al., 2016; Barberá et al., 2019). More recent techniques, including social network analysis (Himelboim et al., 2017), topic modeling via latent Dirichlet allocation (Greene and Cross, 2017; Kherwa and Bansal, 2019), variational EM algorithms (Imai et al., 2016; Slapin and Proksch, 2008), and machine learning–based sentiment analysis (Hasan et al., 2018), have expanded the toolkit for estimating ideological positions among broad populations of social media users. A foundational contribution in this area is Barberá (2015), who developed a Bayesian spatial-following model to estimate ideal points for large samples of both elite and regular Twitter users in the United States and several European countries.

Building on prior work using latent-space models for networks and related positional estimates (Hoff et al., 2002; Shortreed et al., 2006), Barberá (2015) argue that ideology can be represented as a continuous latent variable along a multidimensional ideological spectrum. Individual point estimates are obtained by locating users within a low-dimensional ideological space. Their method employs a logit model that characterizes the probability that a user follows a given political account. The formulation relies on the idea that users tend to follow elites with similar ideological positions and engage with their content (Enelow and Hinich, 1984). In this setting, political elites refer to influential Twitter accounts whose followership patterns provide strong discriminatory power for predicting users' ideological orientations.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

2.2 Correspondence analysis

While the Bayesian spatial-modeling approach to ideal-point estimation is valued for its flexibility in incorporating prior information and quantifying uncertainty (Bafumi et al., 2005), its reliance on Markov chain Monte Carlo (MCMC) sampling makes it less practical for large-scale social network applications (Barberá, 2015). To facilitate the analysis of large user networks while preserving interpretability and comparability, Barberá et al. (2015) proposed using CA to project users onto a latent ideological space as a scalable alternative to a full Bayesian estimation framework.

Correspondence analysis has since been widely used to quantify and visualize the political ideology of social media users (Lozza et al., 2013; Lai et al., 2022). Broadly speaking, CA is a statistical technique for uncovering associations between categorical variables, including political attitudes and preferences (Pack and Jolliffe, 1992; Kroonenberg and Greenacre, 2004; Nenadic and Greenacre, 2007). The CA framework builds upon the same log-linear latent-space model that underlies network-based ideal-point estimation. Specifically, users' connections on Twitter are treated as indicators of similarity and used to construct a correspondence matrix (Barberá et al., 2015). Applying singular value decomposition (SVD) to the matrix of standardized residuals yields a low-dimensional ideological subspace that captures the strongest systematic associations in the data. Projecting users onto this subspace produces coordinates representing their estimated ideological positions. Empirical evidence suggests that these CA-based latent positions are highly correlated with those obtained from Bayesian MCMC methods, offering a conceptually aligned yet more interpretable and computationally efficient alternative (Lowe, 2008; Barberá et al., 2015).

2.3 Toward a comprehensive approach: joint network- and text-based estimation

Over time, with the rapid proliferation of social network site data, Barberá (2015) has become one of the most widely adopted methods in the field. It has been replicated and extended across a range of substantive areas, including protest survey analysis (Barrie and Frey, 2021), environmental discourse (Chang et al., 2022), and public health misinformation (Borghouts et al., 2023). Examining the political landscape of the lower house of the Polish legislature, Ecker (2017) found empirical support for the relationship between Twitter network–based policy position estimates and politicians' individual voting records. Similarly, Souza et al. (2017) applied the approach to the 2014 Brazilian elections, showing that network-based estimates for federal deputies were broadly consistent with roll-call–based measures, although interpreting ideological distances within a highly fragmented legislature remained challenging.

To further evaluate reliability, Barberá (2015) and Jost et al. (2018) demonstrated that follower-based ideal-point estimates for both elites and ordinary users correlate strongly with conventional measures of ideology, such as DW-NOMINATE scores and party registration records (Heller and Mershon, 2009). Ramaciotti Morales and Muñoz Zolotoochin (2022) also found that CA-based estimates largely align with classifications generated by language models using users' Twitter bios, though their study did not integrate network and text information into a single modeling framework.

A key gap in the existing literature concerns the temporal stability of the CA estimates. Social media networks are highly dynamic and evolve rapidly, yet most applications estimate ideal points at a single static time point. As a result, little is known about the consistency of applying a CA model constructed from elites' networks at one time to the networks of ordinary users observed at another. This lack of temporal robustness motivates the need for methods that jointly leverage multiple behavioral signals and better accommodate the evolving nature of online social networks.

Moreover, the accuracy of network-based approaches faces limitations in settings where social connections alone are insufficient to reveal true ideological positions. For example, a well-known

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

and socially active political figure located at an ideological extreme may be estimated near the center because they attract followers across the entire spectrum. In such cases, the complexity of social networks can obscure meaningful signals, whereas the textual content of users' tweets may provide a more direct indicator of ideological orientation. The text-based ideal-point model of Vafa et al. (2020) exemplifies this approach, using an unsupervised probabilistic topic model to identify latent topics and estimate ideal points based on political texts. In parallel, advances in automated text analysis have opened new avenues for recovering latent traits from political discourse, with increasing attention to measurement validity and theoretical grounding (Schoonvelde et al., 2019). Researchers have now applied a range of natural language processing techniques to extract ideological cues from tweets, including keyword usage, sentiment, topic structure, and linguistic features (Spell et al., 2020; Wu et al., 2023; Chao et al., 2022). These developments highlight the promise of text as a complementary source of information for ideal-point estimation, a direction we aim to extend in this paper.

Despite substantial progress, existing research has not fully leveraged the combined value of text and network data from online platforms. Most studies analyze either textual content or interaction patterns in isolation rather than integrating both sources. A closely related contribution is Kim et al. (2018), who proposed a supervised joint model of text and voting data to estimate legislators' positions, linking roll-call votes and speech text through a generalized method of moments estimator in a shared latent space. However, this approach relies on the availability of ground-truth votes and does not recover latent preferences directly from behaviorally generated relational and textual signals.

Given the complexity and richness of social media discourse, integrating both the content of communication and the structure of user connections has the potential to generate more accurate and comprehensive measures of political ideology. These gaps in the literature present an important opportunity for our research to capture a more complete picture of political sentiment and positioning on social media.

3. Data & ideal point estimation method

3.1 Twitter data

The data used in this study originate from Bail et al. (2018) and include Twitter network and text data (tweet timelines) spanning from 2017 through 2022. For our analysis, we focus on a subset of 933 political elites and 1,176 experimental participants with complete follow-ups. The original study used survey responses to infer participants' ideological positions; here, we estimate ideal points as an alternative continuous measure of ideology and validate these estimates against elites' known political affiliations and respondents' self-reported party alignments. In addition to party identification, we retain covariates such as age, gender, frequency of Twitter use, and interest in political news for downstream analyses, allowing further examination of the determinants of ideological positioning and change.

To complement the original dataset, we collected updated Twitter information for the same participants and elites in 2022. Using the Twitter Application Programming Interface (API), we retrieved (1) follower and following relationships between participants and elites and (2) up to the most recent 2,000 tweets per account to comply with standard API rate limits. Data collection was conducted between October and November 2022, ensuring cross-sectional consistency and avoiding artifacts from repeated querying. All publicly available tweets within the API limit were included, allowing for consistent comparison across accounts and facilitating the construction of user-level word-occurrence matrices and updated network graphs for ideal-point estimation.

Due to privacy settings, account deletions, and suspensions, not all individuals' data were retrievable through the API. These missing cases are discussed in more detail in Section 4.1 and Appendix A.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

3.2 Data-augmented estimation framework

This subsection introduces a unified framework for the joint estimation of users' ideal points from both network and text information. We first formalize the latent factor model in scalar notation consistent with standard ideal-point frameworks, then describe its spectral estimation procedure and algorithmic implementation.

3.2.1 Observed data and notation

Let there be N users (indexed by i), J political elites (indexed by j), and a vocabulary of K words (indexed by k). For each user, we observe two behavioral signals:

- **Network follows:** $y_{ij} \in \{0, 1\}$ indicates whether user i follows elite j.
- **Word usage:** $w_{ik} \in \{0, 1\}$ indicates whether user i uses word k at least once in their tweet history (extensions to word frequency or counts are also compatible and straightforward).

3.2.2 Latent structure and observation models

Each user i has an unobserved one-dimensional ideal point $u_i \in \mathbb{R}$ representing ideological preference. Each elite j and word k has a corresponding discrimination parameter $v_j^{(Y)}$ and $v_k^{(W)}$, measuring how strongly it differentiates along the ideological dimension. Intercepts α and β capture baseline tendencies in following or word use.

Network model. For each user–elite pair (i, j),

$$\Pr\left(y_{ij} = 1 \mid u_i, v_j^{(Y)}, \alpha_j^{(Y)}, \beta_i^{(Y)}\right) = \sigma\left(\alpha_j^{(Y)} + \beta_i^{(Y)} + u_i\, v_j^{(Y)}\right),$$

where $\sigma(\cdot)$ denotes a logistic or probit link. Users are more likely to follow elites whose estimated positions $v_j^{(Y)}$ align with user ideological preference u_i.

Text model. For each user–word pair (i, k),

$$\Pr\left(w_{ik} = 1 \mid u_i, v_k^{(W)}, \alpha_k^{(W)}, \beta_i^{(W)}\right) = \sigma\left(\alpha_k^{(W)} + \beta_i^{(W)} + u_i\, v_k^{(W)}\right),$$

so that ideologically distinctive words have loadings $v_k^{(W)}$ whose sign and magnitude correspond to their political orientation and salience.

Conditional independence. Conditional on u_i and the item parameters, $\{y_{ij}\}$ and $\{w_{ik}\}$ are independent. This single-trait structure follows the convention of ideal-point and item-response models: both modalities serve as noisy but complementary indicators of the same latent ideology u_i (Bafumi et al., 2005).

Identification and interpretation. The model is invariant to sign and scale transformations of u_i and $v^{(\cdot)}$. We orient the latent scale so that right-leaning (Republican) elites have positive loadings on average and standardize estimated user scores $\hat{\gamma}_i$ to have mean 0 and variance 1. The resulting dimension is interpreted as a left–right ideological continuum.

3.2.3 Spectral surrogate estimation and implementation

We adopt a spectral surrogate approach that approximates the latent structure by leveraging the leading eigenvectors of user–user similarity matrices. This method is mathematically equivalent to analyzing the eigenvectors of the normalized graph Laplacian constructed from the network and text similarity graphs, a standard technique in spectral graph theory that recovers low-dimensional embeddings of nodes based on shared connectivity patterns (Binkiewicz et al., 2017).

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

Under standard latent factor assumptions, this spectral representation consistently recovers the principal latent dimension corresponding to users' ideological orientation.

Conceptual overview. The key intuition is that two users are ideologically similar if they (i) follow similar sets of political elites, and (ii) use similar language in their tweets. These relationships are expressed in two matrices:

- The **network/correspondence matrix**, $\mathbf{Y} \in \{0,1\}^{N \times J}$, where each element $y_{ij} = 1$ indicates that user i follows elite j.
- The **word/text matrix**, $\mathbf{W} \in \{0,1\}^{N \times K}$, where each element $w_{ik} = 1$ indicates that user i has used word k at least once in their tweet history.

From these, we derive user–user similarity matrices based on co-following and co-word usage patterns:

$$\mathbf{S}_Y = \mathbf{Y}\mathbf{Y}^\top, \qquad \mathbf{S}_W = \mathbf{W}\mathbf{W}^\top.$$

Each entry (i, i') in $\mathbf{S}_Y$ or $\mathbf{S}_W$ represents the degree of similarity between users i and i' in their following or language behavior.

Normalization and aggregation. Because $\mathbf{S}_Y$ and $\mathbf{S}_W$ may differ substantially in scale and degree, we apply symmetric degree normalization to each:

$$\tilde{\mathbf{S}}_Y = \mathbf{D}_Y^{-1/2}\,\mathbf{S}_Y\,\mathbf{D}_Y^{-1/2}, \qquad \tilde{\mathbf{S}}_W = \mathbf{D}_W^{-1/2}\,\mathbf{S}_W\,\mathbf{D}_W^{-1/2},$$

where $\mathbf{D}_Y$ and $\mathbf{D}_W$ are diagonal matrices of row sums of $\mathbf{S}_Y$ and $\mathbf{S}_W$, respectively. This degree normalization improves comparability across modalities and mitigates the influence of high-degree users or very common words. The two normalized similarities are then aggregated:

$$\mathbf{X} = \tilde{\mathbf{S}}_Y + \tilde{\mathbf{S}}_W,$$

placing equal weight on behavioral and textual similarity. More generally, users may introduce a tuning parameter λ and consider $\mathbf{X} = \tilde{\mathbf{S}}_Y + \lambda\,\tilde{\mathbf{S}}_W$ to up- or down-weight the text modality.

Spectral extraction. The eigenstructure of $\mathbf{X}$ captures the principal dimensions of latent variation across users. The first eigenvector primarily reflects overall activity or popularity, while the second eigenvector captures the dominant non-trivial axis of variation, interpreted as the ideological dimension. Standardizing the second eigenvector yields the estimated ideal points:

$$\hat{\gamma} = \text{Standardize}\big(v_2(\mathbf{X})\big),$$

where $v_2(\mathbf{X})$ denotes the second eigenvector of $\mathbf{X}$. Sign orientation is fixed such that right-leaning elites have positive mean scores.

This spectral approximation offers an efficient and interpretable procedure for large-scale data integration, capturing the joint structure of users' ideological behavior across both network and linguistic dimensions. The complete estimation pipeline is summarized below.

3.2.4 Relation to correspondence analysis

CA applied to a user–elite matrix represents the network-only special case of the proposed model with $\lambda = 0$. The CA method embeds users and elites into a low-dimensional space by decomposing standardized residuals of follows. Our data-augmented approach generalizes this by incorporating an additional, conditionally independent source of information, the linguistic content of users' tweets, into the same latent space. Both modalities are modeled as functions of

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

Algorithm 1. Data-Augmented (DA) Ideal-Point Estimation

Input: User–elite matrix $\mathbf{Y} \in \{0,1\}^{N \times J}$, user–word matrix $\mathbf{W} \in \{0,1\}^{N \times K}$, and optional weight $\lambda \geq 0$ (default 1).

Preprocessing: Convert text to lowercase; remove URLs, mentions, emojis, and infrequent words.

Step 1: Compute $\mathbf{S}_Y = \mathbf{Y}\mathbf{Y}^\top$ (co-follow similarity)

Step 2: Compute $\mathbf{S}_W = \mathbf{W}\mathbf{W}^\top$ (co-word similarity)

Step 3: Degree-normalize and rescale each matrix so that $\lambda_{\max} = 1$: $\tilde{\mathbf{S}}_Y = \text{Normalize}(\mathbf{S}_Y)$, $\tilde{\mathbf{S}}_W = \text{Normalize}(\mathbf{S}_W)$.

Step 4: Form the aggregated matrix $\mathbf{X} = \tilde{\mathbf{S}}_Y + \lambda\, \tilde{\mathbf{S}}_W$.

Step 5: Extract the second eigenvector $v_2(\mathbf{X})$.

Step 6: Standardize $v_2(\mathbf{X})$ to mean 0 and variance 1; orient the sign such that right-leaning elites have positive mean.

Output: Estimated ideal points $\hat{\gamma}_1, \ldots, \hat{\gamma}_N$.

the same underlying ideological dimension, and their integration aims to enhance identification, particularly at the ideological extremes.

3.3 Empirical evidence of the joint modeling approach

Before applying the proposed joint model to the collected data, we present a simulation study to evaluate its ability to recover latent ideological positions from multiple data sources. The simulation demonstrates both the consistency of the data-augmented estimator under the model assumptions and its improved identification of ideological extremes compared to single-source estimators. For simplicity, ideology is again assumed to be a one-dimensional unobserved trait.

Data generation process. We first simulated $N = 100$ users with latent ideological positions drawn independently from a uniform distribution:

$$u_i \overset{\text{i.i.d.}}{\sim} \text{Unif}(-5, 5), \quad i = 1, \ldots, N.$$

The vector $\mathbf{U} = (u_1, \ldots, u_n)^\top$ was sorted in ascending order to facilitate examination of extreme positions. For each of S potential data sources (representing different types of behavioral signals), we generated:

$$D_s \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1), \quad \mathbf{V}_s = (v_{s1}, \ldots, v_{sn})^\top \overset{\text{i.i.d.}}{\sim} \text{Unif}(-1, 1), \quad s = 1, \ldots, S.$$

Each binary observation matrix $\mathbf{A}_s$ was generated via a probit model with Gaussian noise:

$$\mathbf{A}_s^{\text{raw}} = \mathbf{U} D_s \mathbf{V}_s^\top + \Sigma_s, \quad [\Sigma_s]_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1),$$
$$\mathbf{A}_s = \mathbb{I}(\mathbf{A}_s^{\text{raw}} > 0),$$

where $\mathbb{I}(\cdot)$ is the indicator function. This setup produced S distinct binary matrices that share the same latent ideological dimension $\mathbf{U}$.

Illustration of real data structure. In the real-data analysis, two such matrices are observed: (1) the adjacency matrix $\mathbf{Y}$, describing user–elite follows; and (2) the word-occurrence matrix $\mathbf{W}$, describing users' language use. To illustrate their structure from our Twitter data:

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

Adjacency matrix	@JoeBiden	@BarackObama	@realDonaldTrump	...
@user1	1	1	0	...
@user2	1	0	1	...
@user3	0	0	1	...
⋮				

(1)

Word occurrence matrix	american	civilrights	government	justice	...
@user1	0	1	0	0	...
@user2	1	0	1	1	...
@user3	1	0	1	1	...
⋮					

Each entry of **Y** indicates whether a user follows a given elite, and each entry of **W** indicates whether a user employs a specific word in their tweets. Both matrices are therefore sparse, high-dimensional representations of observable online behavior reflecting underlying ideological preferences.

Estimators of latent traits. Assuming that all data sources measure the same latent ideology, the ideal estimator would aggregate across sources to reconstruct **U**. We could then approximate this through the spectral aggregation of cross-product similarity matrices:

$$\hat{\gamma}_S = v_2\left(\sum_{s=1}^{S} \mathbf{A}_s \mathbf{A}_s^\top\right),$$

where $v_2(\cdot)$ denotes the second eigenvector of the summed similarity matrix. The first eigenvector captures overall activity or popularity, while the second captures the dominant ideological dimension. Following Krzakala et al. (2013), the second eigenvector of a normalized similarity matrix is maximally correlated with the true latent structure in sparse settings.

In our study, this corresponds to using both **Y** (network ties) and **W** (text features) to estimate ideal points. Reliable estimation is achieved even with moderate sample sizes ($N = 100$), and aggregation across modalities substantially improves the recovery of ideological extremes.

The simulation results shown in Figure 2 and Appendix B demonstrate that:

- The second eigenvector derived from a single source ($\hat{\gamma}_1$) provides a noisy but informative approximation of the latent ideology **U**.
- Aggregating information from multiple sources markedly enhances estimation accuracy and stability, especially in identifying extreme ideological positions.

This simulation parallels our empirical setting, where **Y** and **W** represent distinct yet complementary manifestations of users' political behavior. By aggregating information from both, the data-augmented estimator yields more precise and interpretable estimates of the latent ideological dimension, particularly at the extremes of the political spectrum.

3.4 Comparison with word–embedding approaches

We also explored an alternative variant of the textual input using TF-IDF–SVD embeddings (Ramos, 2003; Kadhim et al., 2017), described in detail in Section 4.3 and

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

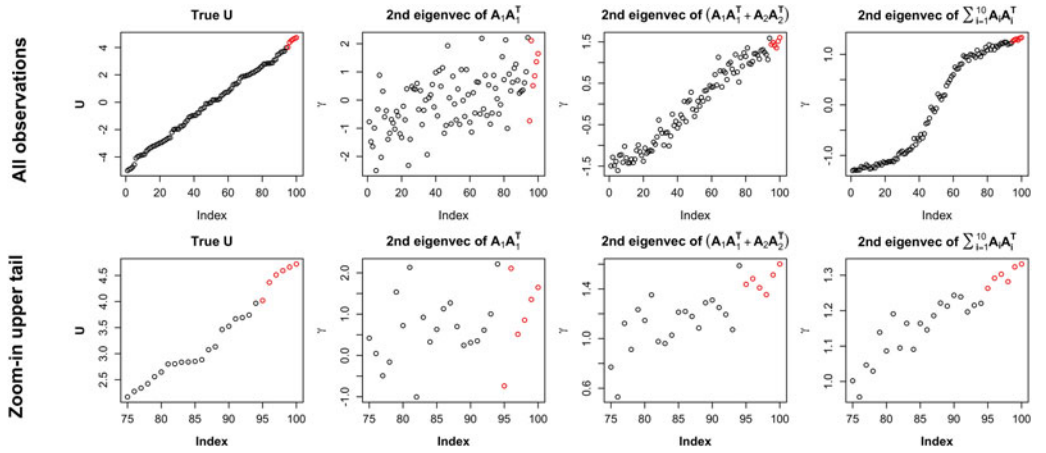


Figure 2. Simulation evidence supporting the data-augmented estimator. The **top row** shows all users, illustrating how adding more sources, from $\mathbf{A}_1\mathbf{A}_1^\top$ to $\sum_{s=1}^{10} \mathbf{A}_s\mathbf{A}_s^\top$, improves alignment with the true latent ideal points **U**. The **bottom row** focuses on the upper tail (the 25 most extreme users), showing that the aggregated estimator better preserves the ordering of ideological extremes (highlighted in red).

Appendix D. This analysis was conducted in parallel to confirm the robustness of our original word-occurrence-based approach and provide additional visualization of ideological structure.

To further benchmark our proposed approach, we conducted an additional robustness analysis using a standard neural word-embedding model. Specifically, we trained a skip-gram `word2vec` model (Mikolov et al., 2013) on the corpus of tweets posted by the same set of accounts (detailed implementation steps are provided in Appendix E). Each user's tweet text was represented as a bag-of-words distribution, and we derived their text-based embeddings by taking Smooth Inverse Frequency (SIF), the weighted averages of the learned word vectors (Arora et al., 2017). The resulting user-level embeddings were used to compute cosine similarity between users, followed by degree normalization to form a Laplacian matrix. The second eigenvector of this matrix provides text-only ideal point estimates, analogous to the CA and DA methods. This procedure serves as a contemporary text-encoding benchmark that captures semantic regularities in language and allows for a comparison with our data-augmented estimator.

4. Results

4.1 Data availability

The availability of longitudinal Twitter user data has continued to decline over recent years due to platform policy changes and user behavior shifts, making it increasingly difficult to track participants' current information. While Bail et al. (2018) originally recruited 1,176 participants, 335 accounts either restricted data access or were deactivated by 2022. Another 137 participants did not follow sufficient elite accounts for reliable ideal point estimation using network data. This attrition results in a final sample of 704 users with valid ideal point estimates.

Figures 9 and 10 (Appendix A) compare excluded versus analyzed participants. Three patterns of missingness emerge here: First, excluded users show significantly lower baseline Twitter activity in initial surveys. Second, while party affiliation remains balanced among excluded participants, the analyzed sample skews Democratic (62% Democratic vs. 38% Republican). Lastly, the experimental groups maintain balance within parties, comprising 439 Democrats (175 control, 264 treatment) and 265 Republicans (107 control, 158 treatment).

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

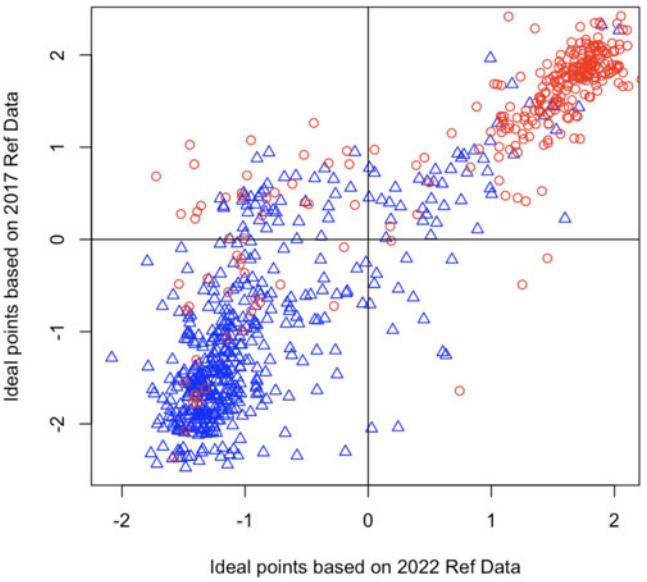


Figure 3. Point estimations of 704 participants' ideal points using the CA method with different reference data. Points are colored according to users' self-reported party affiliations (blue = Democrat, red = Republican).

4.2 Elite ideology over time

One primary concern with the previous correspondence analysis method for estimating ideal points is its consistency in projecting users onto the ideological plane over time. For instance, the U.S. political landscape and the key elites shaping it had evolved considerably from 2017 to 2022. The original CA method relied on a reference set of 1,186 elites whose Twitter data were collected in 2015, and subsequent analyses, as well as related applied work, have continued to use this same reference set (Barrie and Frey, 2021; Chang et al., 2022). This reliance on outdated elite ideal points as a reference risks introducing inaccuracies and inconsistencies in the analysis.

To address this issue, we compiled an updated list of 933 elites and collected their Twitter network and text data by the end of 2022 to serve as the new reference for estimating the ideal points. In our method, we define "elites" more broadly than the *tweetscore* package, which focuses primarily on elected officials such as congressional representatives. Our dataset includes both elected political elites (e.g., members of Congress) and non-elected elites, such as prominent journalists, think tank scholars, and political commentators. These accounts are typically verified, have large follower bases, and are frequently cited or retweeted in political discourse.

For identification purposes, positive ideal points are again assigned to right-leaning and conservative ideological traits, while negative values correspond to liberal traits. Figure 3 shows that updating the reference data leads to a significant change in the estimation of ideal points: a number of Democrats who were previously misestimated (located in the upper left quadrant) now have negative ideal points, indicating improved alignment of the ideal point estimates with self-identified ideology. This improvement supports using an up-to-date reference dataset that corresponds to the target estimations. In Figure 12 (Appendix C), we observe that the new estimates provide a greater separation in the distributions of the estimated ideal points between left and right-leaning individuals. Nevertheless, some discrepancies remain between the estimations and users' self-reported party affiliations. Notably, some users who self-identify as Democrats continue to be estimated with relatively high positive values across both methods. This finding highlights the inherent challenges of relying solely on network data for ideal point estimation, particularly for users who follow a diverse array of elites.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

Table 1. Selected elites' twitter accounts with their ideal point estimates and number of followers taken based on network data collected in 2015 and 2022

Twitter account	$\hat{y}$ in 2015	$\hat{y}$ in 2022	Followers in 2015	Followers in 2022	Followers increase ratio
@JoeBiden	−1.2524	−0.6864	998,737	36,903,062	36.95
@SpeakerPelosi	−1.1682	−0.6459	604,664	8,123,329	13.43
@sensanders	−1.4355	−0.6767	349,149	12,494,541	35.79
@HillaryClinton	−1.2127	−0.6078	3,612,817	31,580,745	8.74
@FoxNews	0.7601	0.1514	5,235,703	23,711,845	4.53
@TuckerCarlson	1.2155	1.6177	2,060,971	5,567,201	2.70

While pinpointing the exact causes of inconsistencies in ordinary users' ideal points over time is challenging without more detailed personal data, the ideal points of prominent political figures are generally expected to remain stable and consistent with public perception. Nevertheless, we observe clear shifts in the CA-estimated ideal points of elite accounts that do not align with their established ideological positions. For instance, Table 1 shows that the ideal points for @JoeBiden, @sensanders, and @HillaryClinton shifted approximately 50% towards zero between the network data collected in 2015 and 2022. This change does not necessarily reflect a true ideological shift among these Democrat political elites; rather, it may result from the substantial changes in the composition of their follower base, which now encompasses a broader range of ideological backgrounds (Eady et al., 2019).

4.3 Highlighting ideological extremes through network and text data

Here, we compare three estimation approaches for elite ideal points: (1) the proposed DA estimator that integrates both network and text information; (2) the baseline CA estimator that relies solely on network data; and (3) the word-embedding–based estimator that directly uses textual information learned from tweet corpora via a `word2vec` skip-gram model. Figure 4 shows the distributions of the estimated ideal points across Democratic, Republican, and Independent elites for each method. The figure illustrates clear differences in the ideological separation produced by each estimator. The CA method (blue) yields relatively compressed distributions, reflecting its limited ability to distinguish elites who share overlapping followers. The word-embedding method (yellow) incorporates linguistic signals from tweet contents and shows improvements over the CA, revealing more nuanced variations in word usages especially among moderates. Our DA estimator (pink), which jointly models network and text data, shows the most distinct ideological polarization, with clear separation between Democratic and Republican distributions and a centered spread for Independents. These results confirm that combining the two modalities provides the most consistent and interpretable estimates of elites' ideological positions as the DA framework effectively leverages complementary information from networks and texts to refine the latent ideological space.

Following established approaches for quantifying partisan lexical asymmetries (Monroe et al., 2008; Gentzkow et al., 2019), we also computed a lexical polarization index (LPI), defined as the difference in average word usage rates between the most extreme (top 10%) and the moderate (bottom 50%) elites. The highest-scoring "extremist" words include *FBI, Donald, FOX, podcast, GOP, Trumps, Georgia, Democrat, Hunter*, and *documents* – terms tied to partisan media, political scandals, and identity-laden topics. In contrast, words most associated with moderates include *communities, veterans, bipartisan, businesses, ensure, colleagues, legislation, resources*, and *infrastructure*, reflecting institutional or policy-oriented language (more detailed results

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

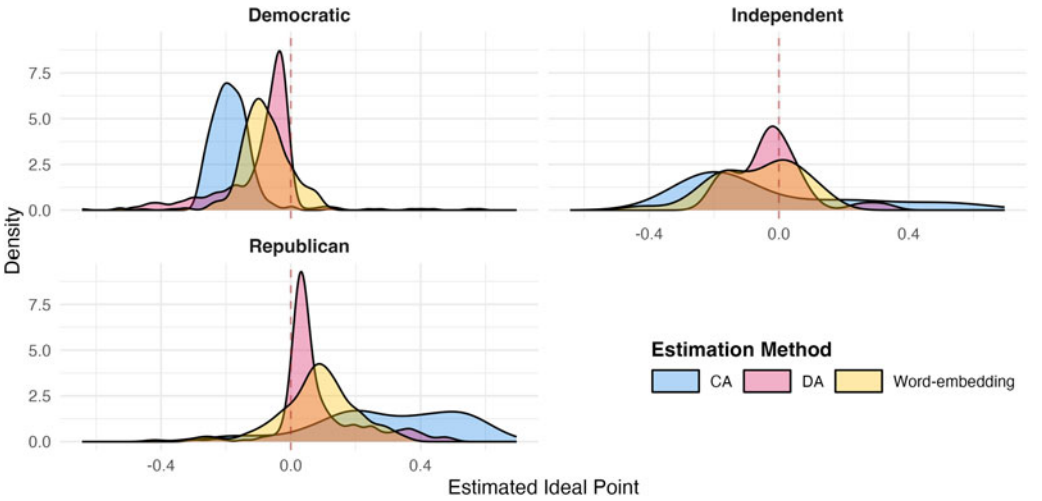


Figure 4. Density plots of estimated ideal points for 933 elites' Twitter accounts, grouped by their party affiliation. The CA estimator (blue) uses only network information; the word-embedding estimator (yellow) uses `word2vec`-derived textual embeddings only; and the DA estimator (pink) integrates network and text data. The DA method produces the clearest separation between Democratic and Republican elites, indicating improved identification of ideological extremes.

described in Appendix F). This contrast suggests that extremist discourse is more event- and identity-focused, whereas moderate discourse centers on governance and service (Figure 14).

Across all elites, the most frequent and cross-partisan words – *America, attack, business, children, court, power, president, rights*, and *Washington* – suggest a common civic vocabulary. We further compare the alignment of text and network structures in Appendix G, suggesting that moderates show higher text–network coupling than extremists.

Figure 5 compares the estimated ideal points of elites before and after incorporating text data. Our proposed DA method demonstrates a superior ability to identify and highlight elites with strongly liberal or conservative ideologies compared to the standard CA method. While the CA approach tends to cluster elites more centrally, our DA method more effectively adjusts the ideal points of moderate elites, independents, and third-party members toward the ideological center, reflecting their positions more accurately. Notably, prominent political figures and media outlets, such as Hillary Clinton, Tucker Carlson, and CNN, which were estimated to have moderate within-party ideal points, are now clearly differentiated using our DA approach. This differentiation underscores the value of leveraging text data, as it reveals distinct patterns of word usage in Twitter rhetoric that correlate with the ideological spectrum. These findings thus highlight the importance of incorporating rich text data, particularly for capturing the ideological extremes.

Moreover, we can also verify the overall DA estimation performance with the known ideological orientations of these elite accounts. For politicians and influencers, we distinguish them by their publicly known party affiliations. For news and media accounts, we use the media bias scores to classify their liberal and conservative leanings. These classifications are well reflected in Figure 5, further showing that our ideal point estimation strategy provides an accurate depiction of the true ideological distribution.

4.4 Revisiting Bail et al. (2018)

As a further and wider application of our method, we analyzed and compared the ideal points of real ordinary users from the experiment in Bail et al. (2018) through the CA method and our proposed DA approach. We first applied the CA method using the default reference data from 2015 to the experimental data (Barberá, 2021). As shown in Figure 6, the densities of estimated ideal

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

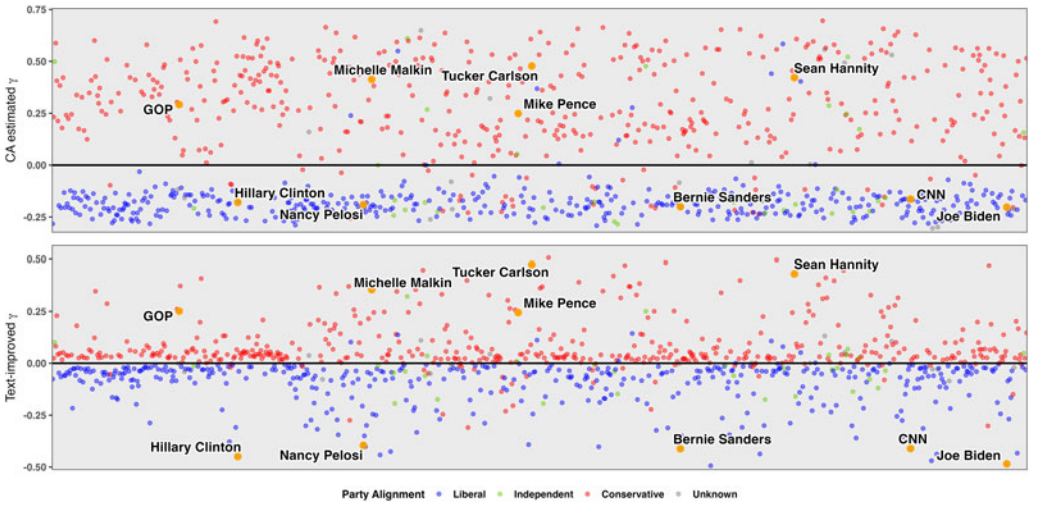


Figure 5. Estimated ideal points of political elites before and after adjusting for text data, with their true political alignment marked in colors. Some well-known elites with significant changes to their estimated ideal points after applying the new method are named and highlighted. Those with extreme ideologies are more distinct under the new estimation method while CA estimation does not clearly identify true extremes.

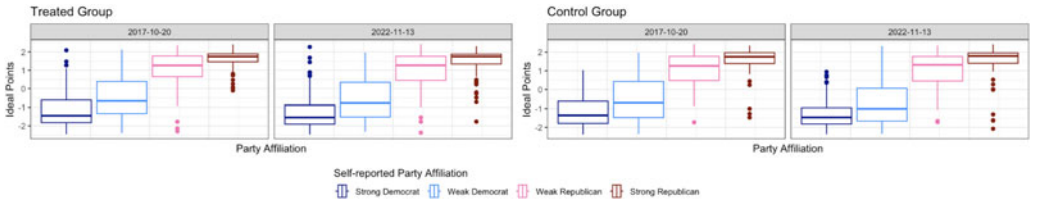


Figure 6. Estimated participants' ideal points by the CA method. Ideal points are generally consistent with party alignments reported by users during the experiment.

points align with the party affiliations self-reported by participants during the survey. Individuals with CA-estimated ideal points that contradict their affiliations predominantly self-reported moderate or weak party alignments, positioning them near the center of the estimated ideal point distribution. There are stronger discrepancies between ideal points and party affiliations among Democrats (18%) than Republicans (11%). Since the CA method relies solely on social network data, it reflects the phenomenon whereby a greater proportion of Democrat users tend to follow political elites whose views oppose their own, relative to their Republican counterparts. This finding contributes additional evidence to an ongoing debate in the literature, as empirical studies have offered conflicting perspectives on whether liberals are more inclined than conservatives to follow media and political accounts from the opposite sides (Barberá et al., 2015; Eady et al., 2019; Wojcieszak et al., 2022).

Figure 7 presents results obtained using our DA method with an updated elite list from 2022 to estimate ideal points longitudinally. Consistent with findings from the CA method, these estimated ideal points mostly align with users' self-reported party affiliations. However, unlike the relatively static CA-based ideal points, the DA method reveals more substantial shifts over the five-year period following the original experiment. Specifically, the overall sample average ideal point decreased from −0.036 in 2017 to −0.139 in 2022 (median changed from −0.094 to −0.139). Among Democratic users, the average ideal point changed slightly from −0.915 to −0.917, whereas Republican users exhibited a more pronounced shift, decreasing from 1.362

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

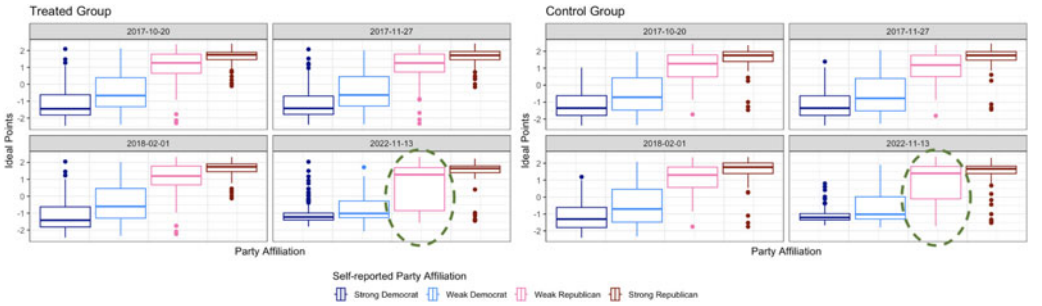


Figure 7. Estimated participants' ideal points by the DA method at various time points. Distributions remained stable during 2017, but experienced notable changes in 2022, as circled boxplots highlight that ideal points of weakly aligned Republicans spread across both positive and negative values.

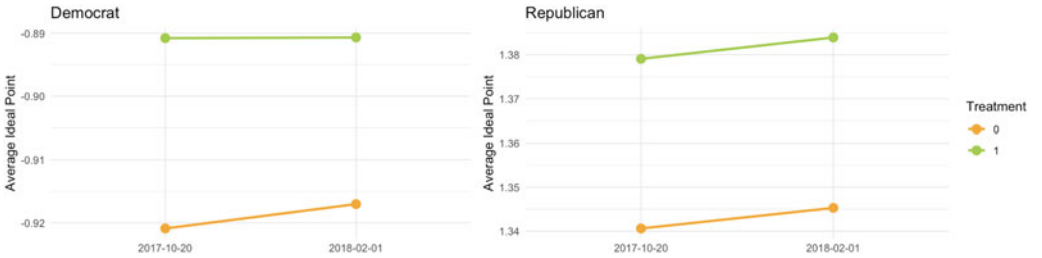


Figure 8. Estimations of changes in ideal points for Democrats and Republicans over time, in contrast to theoretical, no-treatment-effect changes if the parallel trend assumption holds.

to 1.123. These results suggest that the distribution of estimated ideal points among Republican users became more centered over time, potentially reflecting increased heterogeneity or moderation in online behavior among a subset of users. These distributional shifts are consistent with emerging literature documenting ideological heterogeneity and moderation among segments of the Republican electorate in recent years, particularly post-2016 and during the COVID-19 era (Pew Research Center, 2017; Ekins, 2017; Iyengar et al., 2019).

Finally, we revisited the analysis of Bail et al. (2018) to explore whether the exposure to opposing political views had any observable effect on users' actual online behavior. Unlike the original study, which focused on changes in self-reported policy attitudes, our analysis and method examine changes in users' Twitter behaviors, specifically their network connections and posted texts, as inferred through estimated ideal points. To do this, we estimated ideal points for participants just before and immediately after the study, assuming that the short intervention window did not lead to interactions among participants. As shown in Figure 8, we found no evidence of meaningful changes in the estimated ideal points from network composition and texts of participants between the two time points. This aligns with prior work suggesting that observable behavior, such as who users follow or how they communicate online, is relatively resistant to change following brief interventions (Valente, 2012).

5. Discussion

In this work, we advance ideal point estimation methodologies by proposing a joint model for network and text data that enhances the reliability and accuracy of estimating latent political preferences among users. In particular, our results reveal ideological extremity among political elites with greater precision, a task at which previous methods often faltered. This approach normalizes

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

the network and text data to a common scale, but this is not the only possible solution for scaling differences between covariates. It is, for example, possible to normalize the word occurrence matrix directly and introduce a tuning parameter when adding the network and text information. While this step may provide some additional gains if the tuning parameter is chosen correctly, setting such parameters is notoriously hard and frequently estimators behave consistently for a very wide range of parameter settings (e.g. Binkiewicz et al., 2017). Moreover, as embedding-based methods are increasingly popular in the literature and as a further robustness check, we also compare our joint data-augmented estimator with a text-only ideal point model using `word2vec` embeddings (Mikolov et al., 2013). Although the embedding-based approach captures semantically meaningful ideological signals from language alone, it still lags behind the DA estimator in distinguishing ideological extremes and overall consistency with known party affiliations. This reinforces that neither networks nor text data alone are sufficient, and that their combination offers the most comprehensive and stable measurement of ideology.

Beyond its empirical performance, a key advantage of the proposed framework lies in its interpretability relative to unsupervised embedding-based approaches. In our model, the estimated user positions $\hat{\gamma}_i$ correspond to coordinates on a latent left–right ideological continuum, which can be directly validated against known party affiliations or policy stances (Clinton et al., 2004; Bafumi et al., 2005; Barberá, 2015). The elite- and word-level parameters, $\nu_j^{(Y)}$ and $\nu_k^{(W)}$, represent discrimination weights indicating how strongly each elite or term contributes to ideological differentiation – positive values aligning with conservative orientation and negative values with liberal orientation (Clinton et al., 2004). These interpretable quantities link model output to observable political behaviors (e.g., following specific elites or using partisan language), allowing substantive validation and transparent comparison across studies (Barberá et al., 2015; Barberá, 2015). By contrast, unsupervised embedding models, such as `word2vec` or transformer-based text embeddings capture latent semantic regularities in high-dimensional spaces but lack an explicit, identifiable ideological axis (Mikolov et al., 2013; Devlin et al., 2019; Rodríguez et al., 2022). Thus, our approach offers both conceptual clarity and empirical transparency, facilitating theoretical interpretation and reproducibility in studies of online political behavior.

It is important to also acknowledge certain inherent challenges and potential extensions of using our estimated ideal points in downstream analyses. While our analysis uses data originally collected by Bail et al. (2018), we treat this dataset as a motivating real-world application for evaluating the utility of our data-augmented ideal point estimation method, rather than to replicate or test the causal hypotheses of the original study. Due to the significant attrition over time, where many participant accounts were deactivated, restricted, or lacked sufficient activity for estimation, the original experimental design's assumptions no longer hold. Moreover, network structures evolve dynamically over time, which violates the static network assumptions underlying the causal framework of Bail et al. (2018). That said, the extended analysis does yield substantive insight into the long-term evolution of ideological alignment on Twitter. For instance, our updated estimates reveal asymmetric patterns and variability of ideological drift among weakly aligned Republicans between 2017 and 2022, with many shifting closer to the ideological center. These results highlight how the joint estimation method may detect trends in political positioning that pure network-based approaches may miss.

Additionally, our simulation framework assumed a Uniform distribution for latent ideology. This simplifying choice does not capture the clustering and skewness seen in empirical ideology distributions (e.g., bimodal left–right polarization). Future research could adopt empirically grounded priors or mixture models to more closely approximate real-world ideological structures. This would allow testing whether the joint model maintains robustness under more complex ideological landscapes.

Although our empirical analyses estimate ideal points across time through re-estimating the ideal points based on users' networks and textual content available at each observed time

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

point, this design does not yet impose any temporal structure on these estimates. In contrast, time-varying ideal point models, such as those proposed by Martin and Quinn (2002), Carroll et al. (2009), and more recently Kubinec (2025), explicitly model latent trajectories that evolve dynamically over time. Extending our joint modeling framework to a time-varying context (e.g., via dynamic latent factor or state-space models that combine network and text data streams) represents a promising direction for future work.

While Twitter discontinued free access to its API in 2023, our methodological framework is broadly applicable to settings where network data are accompanied by text information. For example, there are large corpora of historical Twitter data that can be reanalyzed to understand how stable ideal points are. More platforms, such as Facebook, Reddit, and TikTok, continue to offer varying levels of API accessibility, which allow researchers to analyze current online users. Beyond these, researchers have routinely collected (symmetric and asymmetric) network data and accompanying open-ended survey questions that can be analyzed to identify positioning within these studies (Singh and Moody, 2022). Building on our work, future research could explore more sophisticated text analysis models and expand the joint modeling framework to incorporate additional data sources, such as multimedia content or user engagement metrics. Finally, using large language models to identify specific linguistic patterns or word choices that distinguish users with extreme ideological dispositions could offer deeper insights into the relationship between language and political behavior (Wu et al., 2023). These advancements would not only refine our understanding of ideological positioning but also enhance the applicability of the method across diverse digital platforms and contexts.

Acknowledgements. We are grateful to the members of the Polarization Lab at Duke for providing important feedback on early versions of this work.

Competing interests. No competing interest is declared.

Data availability statement. The code for this work is available in https://github.com/sheny2/Ideal-point-estimation, including a large user-elite network data and word occurrence matrix. Network and text information for the participants in the original study is not shareable publicly due to being personally identifying information.

Funding statement. This work was partially supported by DMS-CAREER 2046880 and the Templeton Foundation.

Note

1 Although Twitter was renamed "X" in July 2023, we refer to it by its former, more familiar name throughout.

References

Arora, S., Liang, Y., & Ma, T. (2017). A simple but tough-to-beat baseline for sentence embeddings. In *International Conference on Learning Representations*.

Bafumi, J., Gelman, A., Park, D. K., & Kaplan, N. (2005). Practical issues in implementing and understanding bayesian ideal point estimation. *Political Analysis*, *13*(2), 171–187.

Bail, C. A. (2016). Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media. *Proceedings of the National Academy of Sciences*, *113*(42), 11823–11828.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221.

Bailey, M. (2001). Ideal point estimation with a small number of votes: a random-effects approach. *Political Analysis*, *9*(3), 192–210.

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

Barberá, P. (2021). R package. GitHub. https://github.com/pablobarbera/twitter_ideology/tree/master/pkg/tweetscores.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

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

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

Barberá, P., & Rivero, G. (2015). Understanding the political representativeness of twitter users. *Social Science Computer Review*, *33*(6), 712–729.

Barrie, C., & Frey, A. (2021). Faces in the crowd: Twitter as alternative to protest surveys. *PLoS One*, *16*(11), e0259972.

Binkiewicz, N., Vogelstein, J. T., & Rohe, K. (2017). Covariate-assisted spectral clustering. *Biometrika*, *104*(2), 361–377.

Borghouts, J., Huang, Y., Gibbs, S., Hopfer, S., Li, C., & Mark, G. (2023). Understanding underlying moral values and language use of covid-19 vaccine attitudes on twitter. *PNAS Nexus*, *2*(3), pgad013.

Carroll, R., Lewis, J. B., Lo, J., Poole, K. T., & Rosenthal, H. (2009). Measuring bias and uncertainty in dw-nominate ideal point estimates via the parametric bootstrap. *Political Analysis*, *17*(3), 261–275.

Chang, C. H., Armsworth, P. R., & Masuda, Y. J. (2022). Environmental discourse exhibits consistency and variation across spatial scales on twitter. *Bioscience*, *72*(8), 789–797.

Chao, Z., Molitor, D., Needell, D., & Porter, M. A. (2022). Inference of media bias and content quality using natural-language processing. *arXiv e-prints*, arXiv:2212.00237.

Clinton, J., Jackman, S., & Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, *98*(2), 355–370.

De Francisci Morales, G., Monti, C., & Starnini, M. (2021). No echo in the chambers of political interactions on reddit. *Scientific Reports*, *11*(1), 2818.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT* (pp.4171–4186).

Eady, G., Nagler, J., Guess, A., Zilinsky, J., & Tucker, J. A. (2019). How many people live in political bubbles on social media? Evidence from linked survey and twitter data. *Sage Open*, *9*(1), 2158244019832705.

Ecker, A. (2017). Estimating policy positions using social network data: cross-validating position estimates of political parties and individual legislators in the polish parliament. *Social Science Computer Review*, *35*(1), 53–67.

Effing, R., Van Hillegersberg, J., & Huibers, T. (2011). Social media and political participation: Are facebook, twitter and youtube democratizing our political systems?. In *Electronic Participation* (pp. 25–35). Springer.

Ekins, E. (2017). The five types of trump voters. Democracy Fund Voter Study Group. https://www.voterstudygroup.org/reports/2016-elections/the-five-types-trump-voters. Accessed June 21, 2017.

Enelow, J. M., & Hinich, M. J. (1984). *The spatial theory of voting: An introduction*. CUP Archive.

Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying online social networks. *Journal of Computer-Mediated Communication*, *3*(1), JCMC313.

Gentzkow, M., Shapiro, J. M., & Taddy, M. (2019). Measuring group differences in high-dimensional choices: method and application to congressional speech. *Econometrica*, *87*(4), 1307–1340.

Golbeck, J., & Hansen, D. (2014). A method for computing political preference among twitter followers. *Social Networks*, *36*, 177–184.

Goplerud, M. (2019). A multinomial framework for ideal point estimation. *Political Analysis*, *27*(1), 69–89.

Greene, D., & Cross, J. P. (2017). Exploring the political agenda of the european parliament using a dynamic topic modeling approach. *Political Analysis*, *25*(1), 77–94.

Grimmer, J., & Stewart, B. M. (2013). Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297.

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications*, *23*(1), 11.

Heller, W., & Mershon, C. (2009). *Political parties and legislative party switching*. Springer.

Hill, S. J., & Tausanovitch, C. (2015). A disconnect in representation? Comparison of trends in congressional and public polarization. *The Journal of Politics*, *77*(4), 1058–1075.

Himelboim, I., Smith, M. A., Rainie, L., Shneiderman, B., & Espina, C. (2017). Classifying twitter topic-networks using social network analysis. *Social Media+ Society*, *3*(1), 2056305117691545.

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, *97*(460), 1090–1098.

Imai, K., Lo, J., & Olmsted, J. (2016). Fast estimation of ideal points with massive data. *American Political Science Review*, *110*(4), 631–656.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the united states. *Annual Review of Political Science*, *22*(1), 129–146.

Jost, J. T., Barberá, P., Bonneau, R., Langer, M., Metzger, M., Nagler, J., Sterling, J. and Tucker, J. A. (2018). How social media facilitates political protest: information, motivation, and social networks. *Political Psychology*, *39*, 85–118.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

Kadhim, A. I., Cheah, Y.-N., Hieder, I. A., & Ali, R. A. (2017). Improving tf-idf with singular value decomposition (SVD) for feature extraction on twitter. In *3rd International Engineering Conference on Developments in Civil and Computer Engineering Applications*.

Kavanaugh, A., Fox, E. A., Sheetz, S., Yang, S., Li, L. T., Whalen, T., Shoemaker, D., Natsev, P., & Xie, L. (2011). Social media use by government: from the routine to the critical. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (pp. 121–130).

Kherwa, P., & Bansal, P. (2019). Topic modeling: a comprehensive review. *EAI Endorsed Transactions on Scalable Information Systems*, *7*(24), e2.

Kim, I. S., Londregan, J., & Ratkovic, M. (2018). Estimating spatial preferences from votes and text. *Political Analysis*, *26*(2), 210–229.

Kroonenberg, P. M., & Greenacre, M. J. (2004). Correspondence analysis. *Encyclopedia of Statistical Sciences*, *2*, 613–619.

Krzakala, F., Moore, C., Mossel, E., Neeman, J., Sly, A., Zdeborová, L. & Zhang, P. (2013). Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences*, *110*(52), 20935–20940.

Kubinec, R. (2025). Generalized ideal point models for noisy dynamic measures in the social sciences. *Sociological Methods & Research* (forthcoming). https://doi.org/10.31219/osf.io/8j2bt_v3.

Lai, A., Brown, M. A., Bisbee, J., Bonneau, R., Tucker, J. A., & Nagler, J. (2022). Estimating the ideology of political youtube videos. Available at SSRN.

Laver, M. (2014). Measuring policy positions in political space. *Annual Review of Political Science*, *17*(1), 207–223.

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

Lozza, E., Kastlunger, B., Tagliabue, S., & Kirchler, E. (2013). The relationship between political ideology and attitudes toward tax compliance: the case of Italian taxpayers. *Journal of Social and Political Psychology*, *1*(1), 51–73.

Maaten, L.V D, & Hinton, G. (2008). Visualizing data using t-sne. *Journal of Machine Learning Research*, *9*, 2579–2605.

Martin, A. D., & Quinn, K. M. (2002). Dynamic ideal point estimation via markov chain monte carlo for the us supreme court, 1953-1999. *Political Analysis*, *10*(2), 134–153.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.

Monroe, B. L., Colaresi, M. P., & Quinn, K. M. (2008). Fightin'words: lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, *16*(4), 372–403.

Mosleh, M., & Rand, D. G. (2022). Measuring exposure to misinformation from political elites on twitter. *Nature Communications*, *13*(1), 7144.

Nenadic, O., & Greenacre, M. (2007). Correspondence analysis in r, with two-and three-dimensional graphics: the ca package. *Journal of Statistical Software*, *20*, 1–13.

Pack, P., & Jolliffe, I. T. (1992). Influence in correspondence analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *41*(2), 365–380.

Center, P. R. (2017). *Political typology reveals deep fissures on the right and left*. Pew Research Center-US Politics & Policy.

Ramaciotti Morales, P., & Muñoz Zolotoochin, G. (2022). Measuring the accuracy of social network ideological embeddings using language models. In *Information Technology and Systems: Proceedings of ICITS 2022* (pp. 267–276). Springer.

Ramos, J. (2003). Using TF-IDF to determine word relevance in document queries. *Proceedings of the First Instructional Conference on Machine Learning, New Jersey, USA* (vol. *242*, pp. 29–48).

Rodríguez, P., Cuadros, M., & Rigau, G. (2022). On the stability and interpretability of word embeddings. *Natural Language Engineering*, *28*(5), 648–672.

Rodriguez, P. L., & Spirling, A. (2022). Word embeddings: What works, what doesn't, and how to tell the difference for applied research. *The Journal of Politics*, *84*(1), 101–115.

Schoonvelde, M., Schumacher, G., & Bakker, B. N. (2019). Friends with text as data benefits: assessing and extending the use of automated text analysis in political science and political psychology. *Journal of Social and Political Psychology*, *7*(1), 124–143.

Shor, B., & McCarty, N. (2011). The ideological mapping of american legislatures. *American Political Science Review*, *105*(3), 530–551.

Shortreed, S., Handcock, M. S., & Hoff, P. (2006). Positional estimation within a latent space model for networks. *Methodology*, *2*(1), 24–33.

Singh, M. K., & Moody, J. (2022). Do social capital and networks facilitate community participation? *International Journal of Sociology and Social Policy*, *5*(6), 385–398.

Slapin, J. B., & Proksch, S.-O. (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science*, *52*(3), 705–722.

Souza, R. M.D, Graça, L. F. G.D, & Silva, R.D S. (2017). Politics on the web: using twitter to estimate the ideological positions of Brazilian representatives. *Brazilian Political Science Review*, *11*, e0003.

Spell, G., Guay, B., Hillygus, S., & Carin, L. (2020). An embedding model for estimating legislative preferences from the frequency and sentiment of tweets.In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 627–641).

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

Storck, M. (2011). *The role of social media in political mobilisation: a case study of the January 2011 Egyptian uprising*. Scotland: University of St Andrews.
Tahk, A. (2018). Nonparametric ideal-point estimation and inference. *Political Analysis*, *26*(2), 131–146.
Takane, Y., Bozdogan, H., & Shibayama, T. (1987). Ideal point discriminant analysis. *Psychometrika*, *52*(3), 371–392.
Tucker, J., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., Nyhan, B. (2018). Social media, political polarization, and political disinformation: a review of the scientific literature. *Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. http://dx.doi.org/10.2139/ssrn.3144139.
Vafa, K., Naidu, S., & Blei, D. M. (2020). Text-based ideal points. arXiv preprint arXiv:2005.04232.
Valente, T. W. (2012). Network interventions. *Science*, *337*(6090), 49–53.
Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to markov graphs and p. *Psychometrika*, *61*(3), 401–425.
Wike, R., Silver, L., Fetterolf, J., Huang, C., Austin, S., Clancy, L. & Gubbala, S. (2022). Social media seen as mostly good for democracy across many nations, but US is a major outlier. *Pew Research Center's Global Attitudes Project*. Retrieved January, 27: 2023.
Wojcieszak, M., Casas, A., Yu, X., Nagler, J., & Tucker, J. A. (2022). Most users do not follow political elites on twitter; those who do show overwhelming preferences for ideological congruity. *Science Advances*, *8*(39), 1–10.
Wu, P. Y., Nagler, J., Tucker, J. A., & Messing, S. (2023). Large language models can be used to estimate the latent positions of politicians. arXiv preprint arXiv:2303.12057.

Appendix A. Missing data analysis

To assess the extent and pattern of missing data across time, we plotted the user-level presence or absence of tweets and network data from 2017 to 2022. Our analysis first reveals considerable attrition: many users active in 2017 are no longer present in the 2022 dataset. Importantly, the data are not missing completely at random (MCAR); rather, it reflects informative dropout. Users may have deactivated accounts, been suspended, or changed behavior in ways correlated with ideological position or treatment group in the original experiment. As a result, reusing this dataset for causal inference without accounting for informative missingness would lead to biased conclusions. In our work, we treat the updated 2022 snapshot as an independent dataset and focus on evaluating the methodological robustness of joint ideal point estimation.

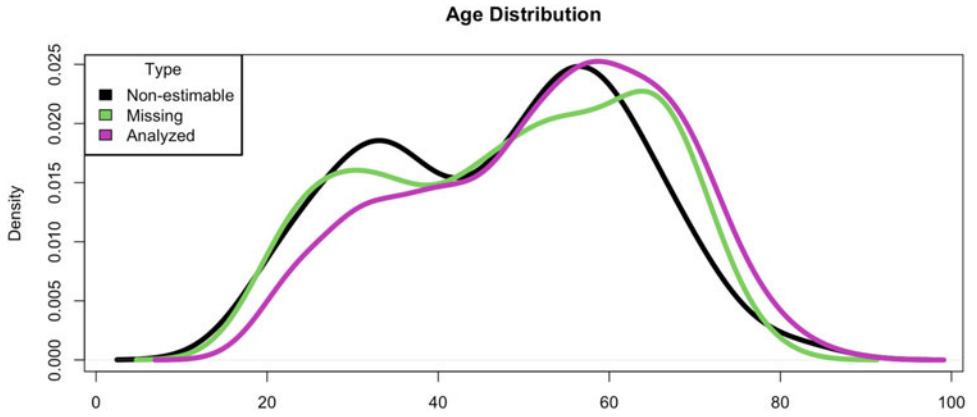


Figure 9. Age distribution across non-estimable, missing, and analyzed participants.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

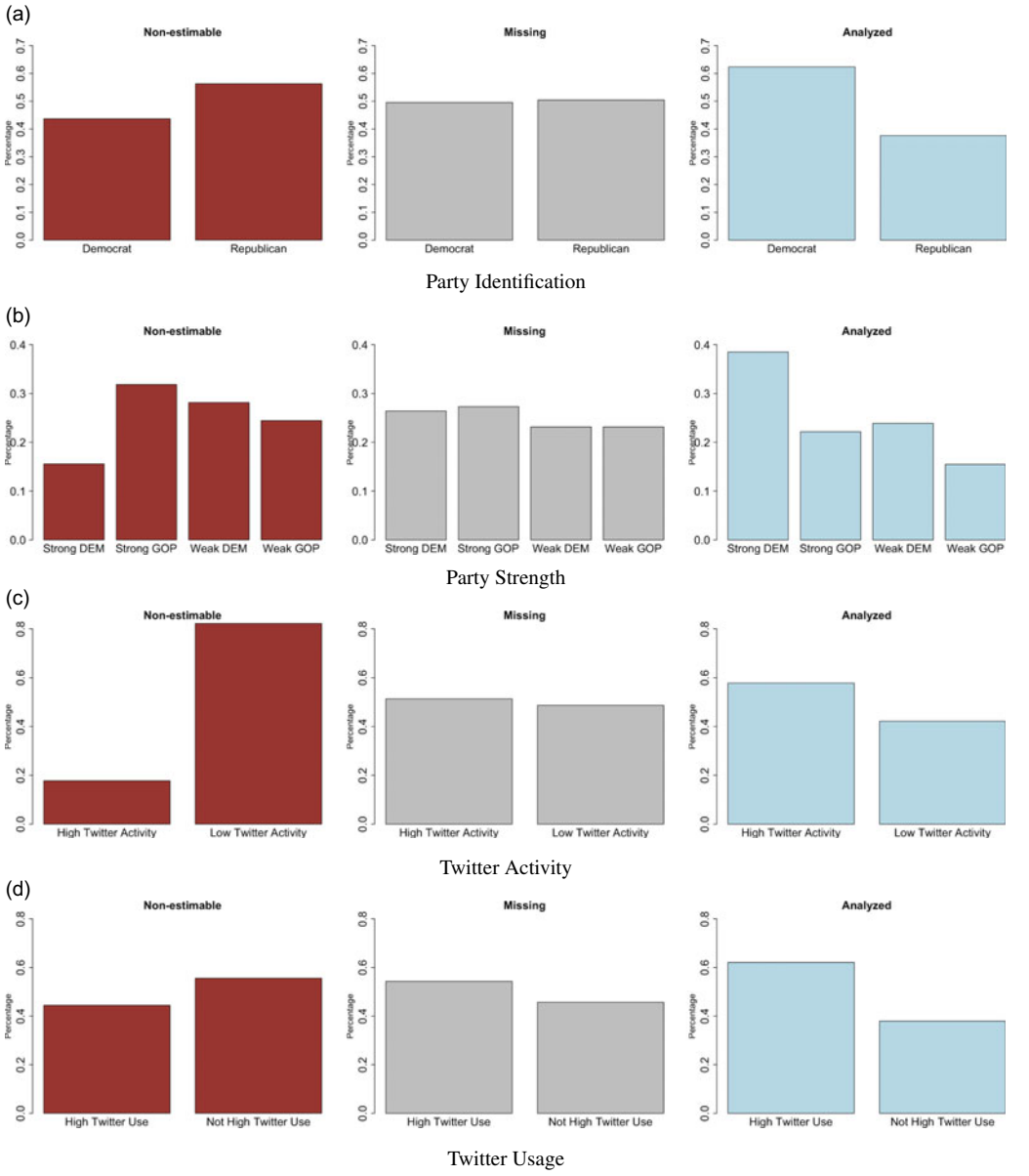


Figure 10. Illustration of non-estimable, missing, and analyzed participants. (a) Party identification, (b) Party strength, (c) Twitter activity, (d) Twitter usage.

Appendix B. Simulation results

Figure 11 visualizes the association between the top three eigenvectors of the network-text Laplacian and the estimated latent ideological positions. While the second eigenvector clearly aligns with the one-dimensional left–right ideological axis and is consistent with our theoretical justification from the spectral graph theory (Krzakala et al., 2013), the first and third eigenvectors exhibit substantially less ideological structure. The first eigenvector is known to capture trivial global properties, such as node degree or popularity, rather than the community structure. The third eigenvector may capture secondary or noisy structural signals but does not strongly align

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

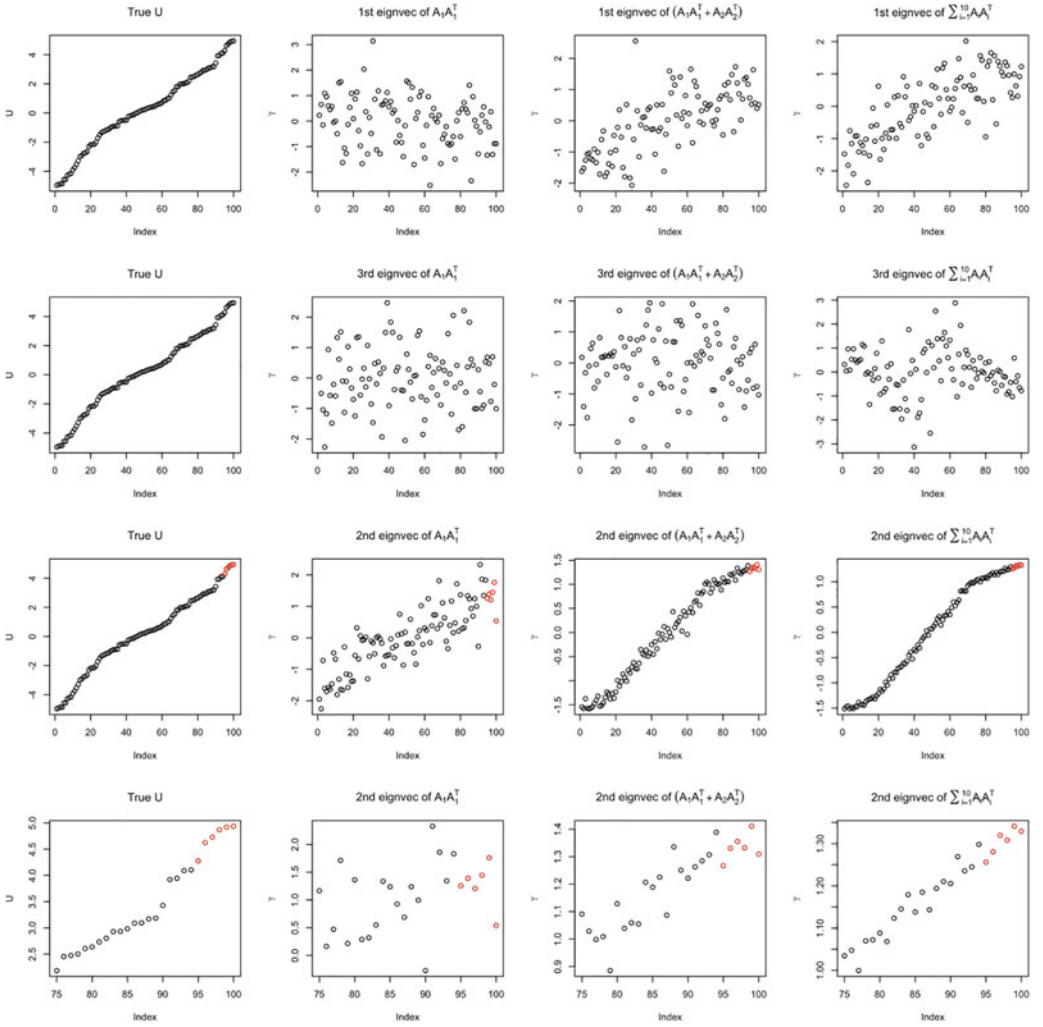


Figure 11. Full simulation results examining the use of different eigenvectors when extracting multiple sources of characteristics.

with partisan affiliation in this case. These findings reinforce our choice to focus on the second eigenvector for ideal point estimation and underscore that the latent ideological space is predominantly one-dimensional in our context.

Appendix C. Sensitivity to elite reference data

To examine the sensitivity of our estimation to the definition of the elite reference set, we re-estimated ideal points using different subsets of elite users in the network anchoring procedure. Figure 12 compares the resulting distributions across specifications. We observe that while the overall shape of the partisan distributions remain unimodal and approximately continuous, the location and scale shift depending on which elite subset is used for alignment. This underscores the importance of choosing a stable and interpretable reference group when interpreting the positions on the latent ideological scale. Nevertheless, relative ordering among users is preserved, and

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

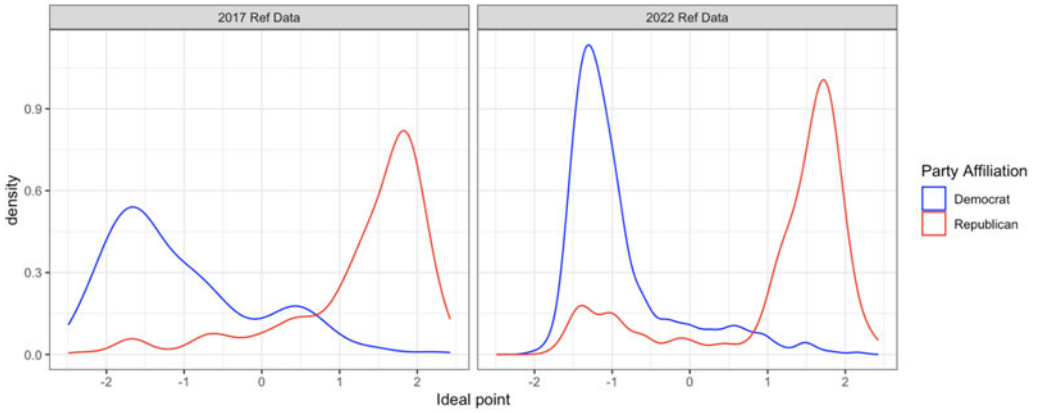


Figure 12. Density plots of all 704 participants' ideal points estimated using the correspondence analysis method with different reference data of elites taken in 2017 and 2022.

correlations between versions remain high, indicating that the qualitative ideological structure is relatively robust to the choice of elite anchors.

Appendix D. TF-IDF word embeddings approach

To validate and strengthen our modeling choice for the textual component, we implemented a TF-IDF transformation of the user-word matrix and derived a low-dimensional embedding via truncated SVD (Ramos, 2003; Kadhim et al., 2017). Specifically, we first applied TF-IDF reweighting to the word occurrence matrix using the `text2vec` package. We then extracted a 30-dimensional embedding. The resulting word embedding matrix preserves both lexical structure and document-level variation. Importantly, this TF-IDF–SVD variant produces highly consistent results compared to our original model that relied on simple word occurrence counts. This suggests that the key ideological signals are robust to alternative word representations, thereby supporting the validity of our original design. These findings are consistent with recent work on the interpretability and stability of word embeddings in applied settings (Rodriguez and Spirling, 2022), which emphasize that embedding-based representations can yield valid insights when carefully evaluated and aligned with research goals.

To further explore the ideological information captured by the text-based user embeddings, we applied t-distributed stochastic neighbor embedding to the 30-dimensional matrix obtained from the TF-IDF–SVD pipeline (Maaten and Hinton, 2008). This non-linear dimensionality reduction method visualizes high-dimensional data while preserving local structure in the embedding space. We set the perplexity to 30 and used a fixed random seed for reproducibility.

Figure 13 shows the resulting two-dimensional projection, where each point represents a user and is colored by their self-reported party affiliation. The axes are abstract dimensions created by t-SNE and do not have direct interpretable meaning like principal components. However, their spatial configuration reveals that users form coherent clusters largely corresponding to their political alignment. Liberal users (blue) and conservative users (red) are well-separated, suggesting that their patterns of word usage, as encoded through TF-IDF and compressed into the low-dimensional embedding, reflect meaningful ideological divisions. This supports our assumption that text features contain relevant information about latent political positions and validates the utility of our original, simpler word-occurrence approach as a strong baseline.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

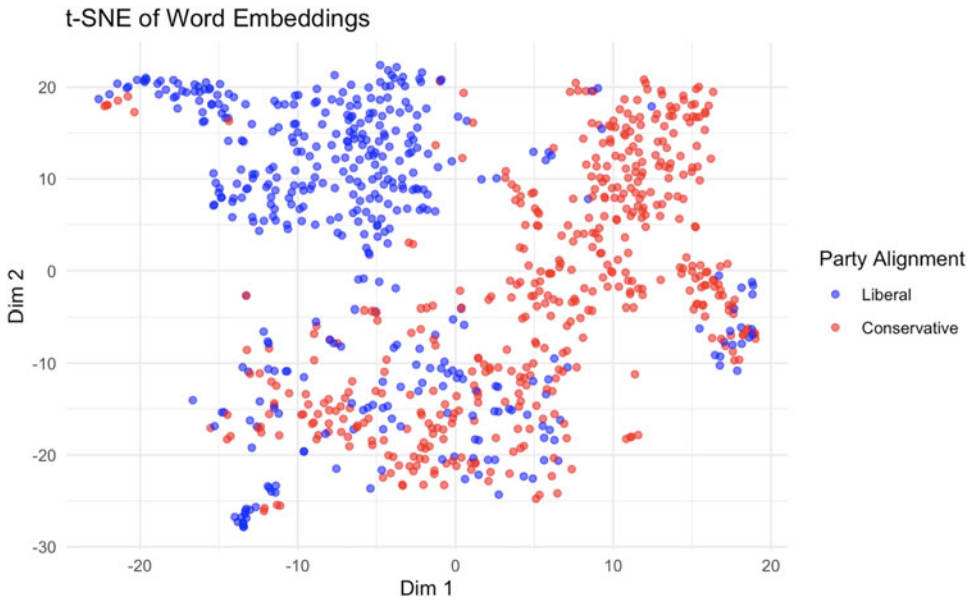


Figure 13. Two-dimensional t-SNE projection of user-level TF-IDF–SVD embeddings. The projection reveals clear clustering by political alignment, indicating that textual features learned from users' tweet content encode meaningful ideological signals.

Appendix E. Word-embedding approach using word2vec

To complement our data-augmented estimator and to benchmark it against more contemporary text-encoding techniques, we additionally implemented a word-embedding–based ideal point estimator using the skip-gram variant of the `word2vec` model (Mikolov et al., 2013). This model learns vector representations of words by predicting surrounding tokens within a fixed context window, thereby capturing both syntactic and semantic regularities in language. The learned embeddings are subsequently aggregated to the elite level to infer text-only ideological positions.

Data and preprocessing. We used the same corpus of tweets collected for the elite accounts described in Section 3.1. Each elite's tweet history was tokenized, lowercased, and cleaned by removing URLs, mentions, emojis, and non-alphabetic symbols. The tweets for each elite were concatenated into a single pseudo-document, producing a corpus of 933 documents (one per elite). This ensures consistent preprocessing with the text data used in our DA method.

Model training. The skip-gram `word2vec` model was trained on this corpus to learn d-dimensional word embeddings. The model parameters were set as follows: embedding dimension $d = 200$, context window size $= 5$, minimum word frequency $= 5$, number of negative samples $= 10$, and training iterations $= 10$. The optimization was performed using negative sampling. The resulting embedding matrix $\mathbf{E_W} \in \mathbb{R}^{V \times d}$ represents each word w among $V = 18,946$ total words as a continuous vector $\mathbf{e}_w$ in a latent semantic space of dimension d.

Embeddings and estimation. To obtain text-based representations for each elite, we applied the SIF weighting (Arora et al., 2017). Let c_{wj} denote the count of word w in elite j's tweets and $p(w)$ the empirical frequency of w across all elites. The SIF-weighted elite embedding is defined as

$$\mathbf{X}_j = \frac{\sum_{w \in \mathcal{V}_j} \frac{a}{a+p(w)}\, c_{wj}\, \mathbf{e}_w}{\sum_{w \in \mathcal{V}_j} \frac{a}{a+p(w)}\, c_{wj}},$$

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

where $a = 10^{-3}$ is a smoothing constant and $\mathcal{V}_j$ denotes the set of words appearing in elite j's tweets. Each resulting vector $\mathbf{X}_j$ is then L_2-normalized to yield the elite-level text embedding.

Using the normalized elite embeddings, we constructed the elite–elite cosine similarity matrix

$$\mathbf{S_W} = \mathbf{Z}\,\mathbf{Z}^\top,$$

where Z stacks all normalized elite embeddings as rows. From this, we computed the degree-normalized Laplacian

$$\mathbf{L_W} = \mathbf{D}^{-\frac{1}{2}}\mathbf{S_W}\mathbf{D}^{-\frac{1}{2}}, \qquad \mathbf{D} = \mathrm{diag}(\mathbf{S_W}\mathbf{1}),$$

and extracted its second eigenvector, $\boldsymbol{\gamma}_W$, which serves as the word-embedding ideal point estimator. Values of $\boldsymbol{\gamma}_W$ were standardized to have mean 0 and unit variance for interpretability. This procedure parallels the spectral framework of our CA and data-augmented estimators, thereby allowing a direct comparison across all three methods.

Interpretation. The `word2vec`-based estimator identifies a one-dimensional ideological structure that emerges purely from the linguistic content of elites' tweets. As shown in Figure 4, the resulting distribution of estimated ideal points reveals clear partisan clustering that is consistent with known political alignments, but remains less sharply separated than the DA estimates.

Appendix F. Lexical polarization and discourse differences

To further investigate how linguistic signals relate to ideological differentiation, we conducted a lexical analysis comparing the language of extreme versus moderate elites. Specifically, we computed the LPI, defined as the difference in average word usage rates between extreme and moderate elites (Monroe et al., 2008; Gentzkow et al., 2019). The resulting contrasts reveal clear thematic distinctions in discourse.

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

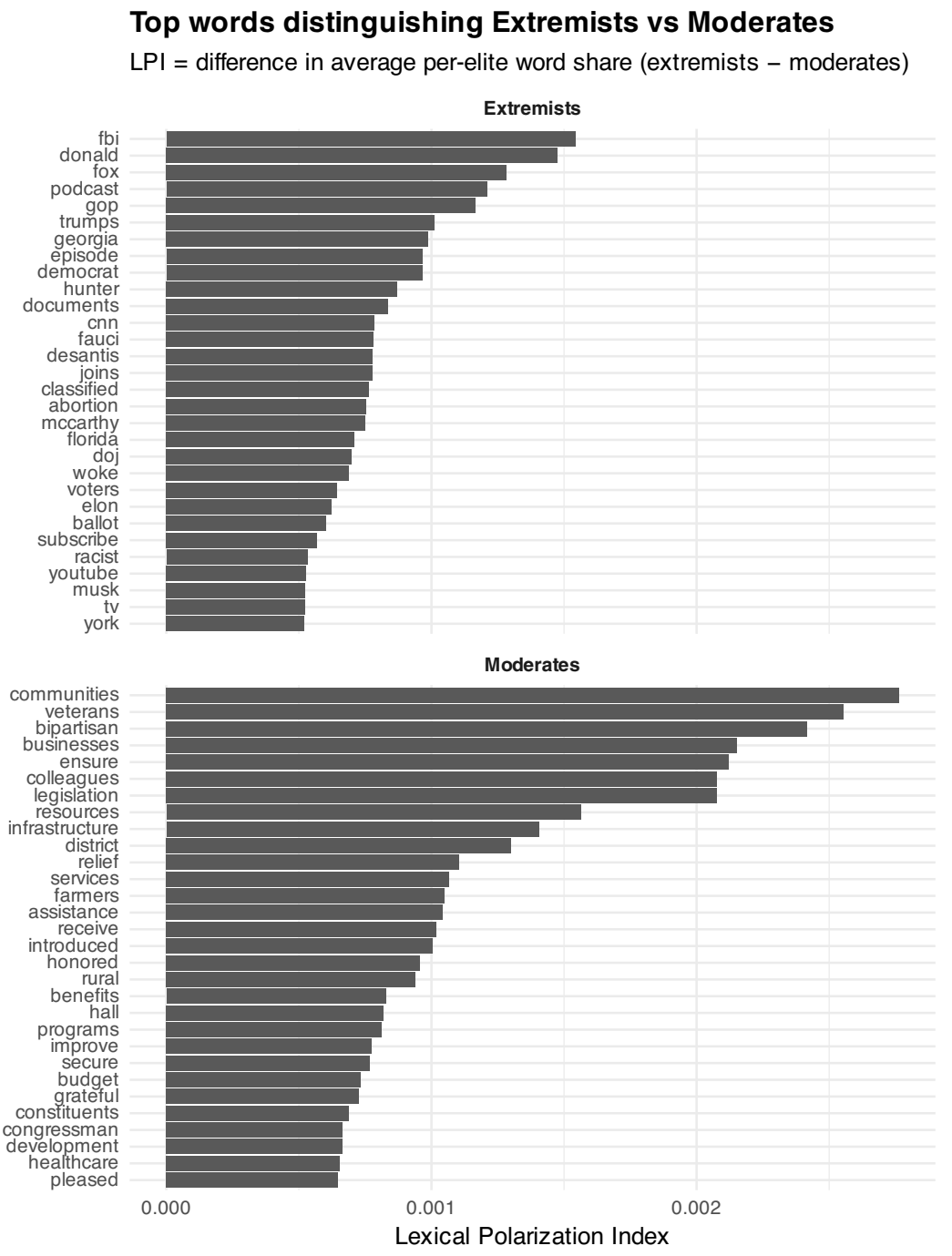


Figure 14. Lexical polarization by extremity. Bars show the top words (by magnitude of the lexical polarization index) that distinguish extremists (top 10%) from moderates (bottom 50%). Extremist discourse concentrates on partisan media and identity-laden topics (e.g., *FBI, Donald, FOX, GOP*), while moderate discourse emphasizes institutional and policy language (e.g., *communities, veterans, bipartisan, legislation, infrastructure*).

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press

Appendix G. Text–network coupling and ideological extremity

Following previous work that integrates network and text data to study polarization and echo-chamber effects (Bail, 2016; Barberá et al., 2015), we compared the alignment between elites' text-based and network-based similarity structures to assess how discourse and social connectivity jointly shape ideological differentiation. For the network dimension, we constructed a symmetric co-occurrence similarity matrix S_{net} from a bipartite follower (or retweet) incidence matrix Y (rows = users, columns = elites). The network similarity was computed as:

$$S_{\text{net}} = \mathbf{D}^{-\frac{1}{2}}(\mathbf{Y}^{\mathsf{T}}\mathbf{Y})\mathbf{D}^{-\frac{1}{2}},$$

where $\mathbf{D}$ is a diagonal matrix of column degrees $\text{diag}(Y^{\mathsf{T}}Y)$, following the standard normalized Laplacian transformation used in spectral graph analysis. Each entry $S_{\text{net},ij}$ represents the normalized co-occurrence between elites i and j (e.g., overlap in follower or retweet audiences).

We then compared this structure to the text-based similarity matrix S_{text}, derived from cosine similarity of elites' embedding representations. Two complementary measures were computed: (i) a global correlation between all off-diagonal elements of S_{text} and S_{net}, and (ii) a per-elite correlation

$$\rho_i = \text{cor}(S_{\text{text},i\cdot}, S_{\text{net},i\cdot}),$$

which captures how well an elite's text-similar neighbors coincide with their network neighbors.

Figure 15 summarizes the per-elite correlations by ideological extremity group (0–50% = moderates, 50–90% = high, 90–100% = extremes). We observe that moderates exhibit higher text–network coupling on average, while the correspondence weakens among more extreme elites. This indicates that moderates' discourse tends to align more closely with the overall network structure as their language overlaps with multiple communities, whereas extremists' linguistic patterns diverge even from their own network clusters. In other words, the separation of extremists reflects not only structural segregation but also the distinctiveness of their rhetorical choices, which are less well explained by networks alone.

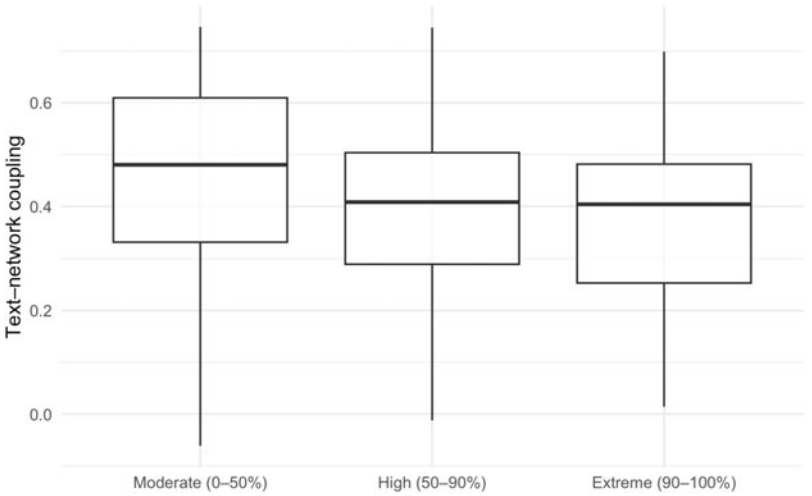


Figure 15. Text–network coupling by ideological extremity. Each box shows the distribution of per-elite correlations between text-based and network-based similarity matrices. Higher values indicate stronger alignment between linguistic and structural neighborhoods. Moderates display the highest coupling, suggesting that their language reflects broader network connections, whereas extremists show weaker coupling, indicating discourse that diverges from network ties and contributes to greater ideological differentiation.

Cite this article: Shen Y., Bail C. and Volfovsky A. (2026). Enhancing joint ideal point estimation strategies with Twitter social networks and text data. *Network Science* **14**, 1–27. https://doi.org/10.1017/nws.2026.10028

https://doi.org/10.1017/nws.2026.10028 Published online by Cambridge University Press