Measuring the Ideology of Audiences for Web Links and Domains Using Differentially Private Engagement Data

Cody Buntain[1], **Richard Bonneau**[2], **Jonathan Nagler**[2], **Josua A. Tucker**[2]
[1] iSchool, University of Maryland, College Park
[2] Center for Social Media and Politics, New York University
cbuntain@umd.edu, rb133@nyu.edu, jonathan.nagler@nyu.edu, joshua.tucker@nyu.edu

Abstract

This paper demonstrates the use of differentially private hyperlink-level engagement data for measuring ideologies of audiences for web domains, individual links, or aggregations thereof. We examine a simple metric for measuring this ideological position and assess the conditions under which the metric is robust to injected, privacy-preserving noise. This assessment provides insight into and constraints on the level of activity one should observe when applying this metric to privacy-protected data. Grounding this work is a massive dataset of social media engagement activity, provided by Facebook and the Social Science One (SS1) consortium, where privacy-preserving noise has been injected into the data prior to release. We validate our ideology measures in this dataset by comparing to similar work on sharing-based, homophily- and content-oriented measures, where we show consistently high correlation (> 0.87). We then apply this metric to individual links from six popular news domains and construct link-level distributions of audience ideology. We further show this estimator is robust to engagement types besides sharing, where domain-level audience-ideology assessments based on views and likes show no significant difference compared to sharing-based estimates. Estimates of partisanship, however, suggest the viewing audience is more moderate than the audiences who share and like these domains. Beyond providing thresholds on sufficient activity for measuring audience ideology and comparing three types of engagement, this analysis provides a blueprint for ensuring robustness of future work to differential privacy protections.

Introduction

Datasets of large-scale online behavior and digital traces are growing more sensitive as privacy expectations and regulations mature. To address such concerns, data providers are turning to differential privacy to balance large-scale data releases with maintaining privacy guarantees for individuals whose data may be included in these releases. Differential privacy techniques operate by injecting noise into observations to prevent identification of individuals in these datasets (see Wood et al. (2020) for an introduction to these methods). These protections come at a cost, however, as standard analyses may produce biased or erroneous results if they do not account for such protections (Evans et al. 2019). This issue is particularly evident in the release of the "Facebook Privacy Protected Full URLs Dataset," referred to as the "Condor" dataset, where Facebook and the SS1 consortium have released a massive collection of 63.5 million links shared on the Facebook platform along with differential-privacy-protected engagement data on age, gender, location, and political preference (Messing et al. 2020). Condor is the largest dataset of link-level engagement released to date and holds marked potential for studying large-scale online behaviors, but researchers lack guidance in and examples of methods that account for differential privacy protections.

This paper provides this guidance by 1) examining a simple weighted-average metric for calculating *ideological positions of audiences* for web domains[1] based on link-sharing in Facebook in the presence of differential-privacy protections, 2) showing how differential privacy impacts this metric, 3) establishing bounds on how this metric should be used, and 4) validating this metric against similar domain-level, sharing-based measures. While similar metrics have been proposed, those efforts rely on highly sensitive data, such as internal Facebook data (as in Bakshy, Messing, and Adamic (2015)) or Twitter profiles aligned with sensitive "voter-file" information (as in Robertson et al. (2018)); in contrast, this paper's metric can be calculated solely from this differentially private, public dataset.

After establishing constraints for our differential-privacy-resilient metric, we use it to extract novel insights about individual hyperlinks, where sparsity issues have forced previous approaches to use domain-level measures. We then assess how different types of engagement–views and likes–impact our measures. For individual hyperlinks, we estimate distributions of link-level ideology measures for several thousand individual links across six popular domains, including YouTube.com, providing insight into long-standing questions about partisan audiences on that platform. For different types of engagement, we measure differences in domains' audience ideologies using link-sharing, viewing, and liking behaviors, also answering open questions about consistency in measurement across engagement types; results show no significant deviation in domain-level estimates across these activities–though a domain's viewing audience is on average more moderate than its sharing or liking audi-

[1]As distinct from the ideological position taken by the domain.

ences. Given the commercial value of viewing data in online platforms, this result is particularly encouraging for the generalizability of share-based studies and for future efforts that leverage protected versions of this sensitive data.

This work's core contributions are:

- A demonstration of how a simple metric for estimating ideology of a domain's audience can be made robust to differential privacy protections;
- An examination of link-level distributions of ideology across six major news sources; and
- New insight into how varied forms of engagement (shares, views, and likes) impact audience ideology estimates.

Related Work

This paper engages with two main communities: First, a large body of research exists on inferring ideology of audiences in studies of media bias and polarization, especially in online spaces (Gentzkow and Shapiro 2010; Bakshy, Messing, and Adamic 2015; Budak, Goel, and Rao 2016; Robertson et al. 2018), which both motivates this work and provides sources for validation. Second, much of the data that could be useful for similar studies is often restricted and sensitive; recent work has explored methods for providing data protections of such sensitive data while still enabling inference on this data (D'Orazio, Honaker, and King 2015), which directly informs our work. Before describing contributions to these communities, we first provide an overview of differential privacy to situate this work.

A Brief Primer on Differential Privacy

At its core, *differential privacy* is an approach to collecting and disseminating aggregate statistics in a way that guarantees some level of privacy for individuals whose data is used to generate these statistics. While Wood et al. (2020) provides an overview of differential privacy for non-technical audiences, such protections generally provide a form of plausible deniability for individuals whose data is included in these statistics. This deniability comes from the property that a third-party cannot learn anything about a single individual whose data contributes to these statistics that could not be learned if that individual's data were excluded. Hence, an individual could claim their data was never included in the released statistics at all, allowing them to deny potential allegations derived from the data. Digital trace data can therefore benefit from applications of differential privacy, as large-scale, aggregate datasets can be released in privacy-protected forms that reduce potential harm to the populations from which the data is collected D'Orazio, Honaker, and King (2015). These protections are also consistent with calls for and regulations on enhanced protections of digital consumer data, and groups like the US Census Bureau are using similar ways to protect sensitive data.

These protections are generally applied by adding noise into the computations of aggregate statistics. Researchers can tune characteristics of this noise to quantify potential privacy loss, and by tracking this loss over subsequent dataset releases, those creating these datasets can maintain privacy guarantees. Characteristics of this noise can be shared as part of the release process without risking these privacy guarantees, so researchers can account for this noise without identifying individuals within the dataset. In the context of the Condor dataset, Facebook bounds privacy loss by adding noise to the aggregated engagement statistics prior to releasing the dataset to external researchers. That is, if a hyperlink has been shared X times in a given month, Facebook adds noise ϵ drawn from a Gaussian distribution to this value, and external researchers only ever see $X + \epsilon$.

Measuring Political Ideology

Methods for estimating ideology, partisan lean, media slant, or similar aspects of information sources (e.g., newspapers, websites, or communities) are well-studied and generally fall into one of two categories: content- or homophily-based approaches. Content-based approaches generally analyze language, while homophily-based methods propagate individuals' ideological preferences to the information sources these individuals share or consume. Content-based analyses like the study of media slant in Gentzkow and Shapiro (2010) or media bias in Budak, Goel, and Rao (2016) are powerful but require content analysis and either manual assessment (as in Budak, Goel, and Rao 2016) or information about political preferences of people sharing that content to learn mappings of particular language to political preference (as in Gentzkow and Shapiro 2010). In contrast, while homophily-based methods need information about political preferences, they do not require analysis of actual content and instead rely on interactions among nodes in a network. Through these interactions, one can propagate political preferences to neighboring nodes, making these methods particularly amenable to algorithmic assessment. In online social networks, such interactions are often computationally cheap to collect through APIs or found data, making these approaches popular in research. Homophily-based methods for inferring political ideology have been used to measure online/offline ideological segregation (Gentzkow and Shapiro 2011), diversity in online news (Bakshy, Messing, and Adamic 2015), ideological biases in search engines (Robertson et al. 2018), and even political lean of disinformation agents (Golovchenko et al. 2020).

Despite the clear utility and popularity of such homophily-based approaches, when these methods use social media data to measure ideology of a news source's or web domain's audience – as in Robertson et al. (2018), Golovchenko et al. (2020), Eady et al. (2020) and others – they commonly rely on easily collectable *sharing* behavior (e.g., an individual shares a tweet with a link to a domain). While sharing behavior is easy to collect from sources like Twitter and Reddit, prior research on social media spaces and online communities shows that the vast majority of users on the platform do not actively share or produce content (Nonnecke and Preece 2000; Preece, Nonnecke, and Andrews 2004; Gong, Lim, and Zhu 2015) – Benevenuto et al. (2009) in particular suggests that 92% of all behavior in one social network was comprised of content viewing alone, which does not produce collectable artifacts in many public

APIs. Data about these viewing behaviors, however, is commercially sensitive, and most social media platforms do not make this data publicly available. For studies of viewing behavior in Facebook, for example, up to the release of SS1's Condor dataset, one has had to rely on researchers employed by Facebook, as in Bakshy, Messing, and Adamic (2015), or partner with researchers at Facebook.

This reliance on content production and sharing leads to a problematic implication: A media source's ideological slant is significantly affected by the source's audience, and as share-based metrics omit activity from a significant portion of the viewing audience, measures of the media source's audience may differ significantly from the true distribution Gentzkow and Shapiro (2010). A unique aspect of the Condor dataset provided by SS1, however, is that it provides engagement data across both sharing and viewing behaviors, binned across several political-preference buckets. Hence, the work in this paper can shed some needed light on the differences in estimates based on shares versus views. That is, by first comparing results from our share-based audience ideology metric to existing work in this area, we can validate our metric despite the privacy-protecting noise injected into Condor observations. Then, by comparing results from our share-based estimator to estimators based on views – and indeed other behavior, such as Facebook's "Like" reaction, which is similar to Twitter's "Favorite" affordance – we can evaluate whether differences in share- and view-based estimates differ significantly.

Inference and Protected Data

While the above context on measures of media bias and audience ideology show a clear need for understanding the impacts of share-versus-view-based metrics, as mentioned, view data is both commercially sensitive and highly private. Facebook has endeavored to help researchers in this need by releasing the Condor dataset and protecting it with differentially private noise, as described in Messing et al. (2020). While works such as D'Orazio, Honaker, and King (2015) and Evans and King (2022) outline how differential privacy can support inference in social sciences, how these protections impact on researcher utility remains an open question. Evans and King (2022) even shows ignoring differential privacy can lead to unpredictable biases in results, including biasing estimated effects towards zero, or in some particularly problematic cases, inverting the sign of the estimates. Despite these risks, Evans and King (2022) shows corrections are feasible in certain scenarios, as the noise added to data for establishing differential privacy guarantees is equivalent to increasing standard measurement error, and for linear systems (e.g., linear regression models), one can correct for noise if details of the noise-generating distribution are known. For non-linear systems like the weighted-average metric we present, however, analytically based corrections are not readily available.

We instead build on Evans and King (2021), which outlines the context in which bias in a ratio metric can be bounded. Evans and King (2021) claims that, if the noise introduced is generally much smaller than true observations, bias in the noisy metric is minimal. While this result is valuable, no guidance is provided regarding how large observations should be relative to noise nor how to evaluate whether one is in this regime. Hence, this paper provides this much-needed guidance for using these noisy observations in the Condor dataset to study media and ideology in a robust manner. We further validate these methods against extant results where such privacy protections are not in place.

Condor: The Facebook Privacy Protected Full URLs Dataset

As a brief overview, the "Privacy Protected Full URLs" dataset, provided to academic researchers by Facebook and the SS1 consortium, is a large-scale collection of URLs and associated engagement data for 63,574,836 hyperlinks shared on the Facebook platform. This dataset exists to provide researchers new insight into how individuals engage with hyperlinks on the Facebook platform while simultaneously maintaining strong privacy guarantees for Facebook users. For a URL to be included in this dataset, it must have been publicly shared by approximately 100 unique individuals (see Messing et al. (2020) for more details). As of this writing, the dataset is on its ninth iteration and contains monthly engagement metrics for all months between 1 January 2017 and 31 December 2021.

For each of these URLs, the dataset contains counts for 11 actions one can take on the Facebook platform (sharing, viewing, liking, commenting, clicking, etc.), broken down by month and audience demographics. These demographics cover an individual's country, age, and gender, from one of 45 countries, seven age groups, and three gender groups (Messing et al. 2020). In the US, Condor further decomposes these counts across six bins representing individual-level political preference, using a "political page affinity" (PPA) metric, a homophily-based measure defined by Barberá et al. (2015) and described in Messing et al. (2020). PPA measures an individual's political ideology on a scale $b \in \{-2, -1, 0, +1, +2\}$ (-2 is very liberal, and +2 is very conservative), with an additional bin for individuals whose PPA is unknown–we exclude this sixth group from our analyses. In this manner, the Condor dataset contains makes available highly sensitive but valuable engagement data for large volumes of online information sharing and consumption behavior.

Given the sensitivity of this data and to protect users of the Facebook platform from potential de-identification, researchers using the Condor dataset do not have direct access to the raw monthly counts of these activities. Instead, researchers can only observe counts of these activities after Facebook has added zero-centered Gaussian noise to them in accordance with zero-Concentrated Differential Privacy (Bun and Steinke 2016). By controlling the amount of noise relative to the amount of engagement across these demographic bins, Condor provides privacy guarantees about the probability of an individual person's single action (e.g., share, view, like, etc.) being attributed to that person. That is, more noise can be injected into counts of views compared to counts of shares or clicks, while noise added across a single action comes from the same normal distribution.

Parameter	Description
ℓ	A single hyperlink
D	A domain D, with multiple associated hyperlinks $\ell \in D$
b	A political page affinity (PPA) bin, such that $b \in \{-2, -1, 0, 1, 2\}$
$s_b(\ell)$	Actual shares of the associated hyperlink ℓ in the PPA bin b
$\epsilon_{s,b}(\ell)$	Differentially private noise added to shares of ℓ in PPA bin b
$\widehat{s_b}(\ell)$	Observed shares of a given hyperlink ℓ in PPA bin b, after noise has been added, equal to $s_b(\ell) + \epsilon_{s,b}(\ell)$
$s_b(D)$	Actual shares across all links in domain D and PPA bin b, summed over $s_b(\ell)$ over all links $\ell \in D$
$\widehat{s_b}(D)$	Observed shares across all hyperlinks in domain D, after noise has been added, equal to $\sum_{\ell \in D} s_b(\ell) + \epsilon_{s,b}(\ell)$
$\widehat{s}(D)$	Observed shares across all PPA bins and hyperlinks in domain D, after noise has been added
ζ_D	Actual audience-level ideology of a domain D
$\widehat{\zeta_D}$	Estimated audience-level ideology of a domain D, using privacy-protected shares

Table 1: Summary of notation in audience-level ideology estimates. $\widehat{x}$ denotes estimates of x using differentially private data.

For this work, we focus on URLs shared primarily in the US. All access to this data is allowed through the SS1 approval process and is conducted on the Facebook Open Research and Transparency platform.

A Robust Metric for Audience Ideology

We now turn to a metric for estimating the ideology of a domain's audience from this dataset. Prior work on media slant has shown audiences' political preferences have a marked relationship with the message, topic selection, and framing of news sources (Budak, Goel, and Rao 2016; Gentzkow and Shapiro 2010; Bakshy, Messing, and Adamic 2015). In this context, Bakshy, Messing, and Adamic (2015) propose a homophily-based metric of the degree to which a news article is aligned with a partisan audience "by averaging the ideological affiliation of each user who shared the article." In the differential-privacy-protected Condor dataset, we can replicate this metric at the web domain level by measuring the average political ideology of the individuals who share content from this domain. Table 1 summarizes the notation we use in defining our version of this metric. While Condor includes engagement metrics, we focus on sharing for consistent comparison with other work on ideology estimation.

As Condor provides engagement data for each PPA bin, we can interpret these counts as the frequencies for which an individual who is very liberal (PPA$= -2$), liberal (PPA$= -1$), etc. has engaged with this content. To estimate a domain's audience ideology ζ_D from this PPA metric, we calculate the weighted average across these five PPA bins, omitting the sixth PPA bin that contains audience engagement with unknown ideological affinity. Eq. 1 shows this metric as the product of each PPA value with the count of individuals who have shared that domain and have that PPA value. In Eq. 1, $s_b(D)$ represents the number of individuals with the PPA value b who have shared the domain D. In Condor, however, engagement frequencies are at the URL/hyperlink ℓ level, not the domain level, so we must first aggregate $s_b(D)$ over all links ℓ in domain D, as shown in Eq. 2.

$$\zeta_D = \frac{1}{\sum_b s_b(D)} \sum_{b \in [-2..+2]} b \cdot s_b(D) \tag{1}$$

$$s_b(D) = \sum_{\ell \in D} s_b(\ell) \tag{2}$$

In the absence of differentially private data protections, Eq. 1 is fundamentally the same metric as in Bakshy, Messing, and Adamic (2015) and is similar to Robertson et al. (2018). With the introduction of zero-centered Gaussian noise, however, these metrics are ill-behaved and can result in discontinuities, as we explain below.

How Noise Impacts this Metric Analytically

While the metric ζ_D in Eq. 1 is a straightforward calculation, differential privacy protections preclude observing the actual number of shares for a given PPA value directly. Instead, we observe a noised version of this value, shown in Eq. 3, where $\epsilon_{s,b}$ is drawn from a zero-centered Gaussian distribution with standard deviation σ. This σ is constant for a single action (e.g., sharing) and reported in the Condor codebook (Messing et al. 2020). Hence, when calculating the number of individuals in PPA bin b who have shared a domain D, we can only construct a noisy estimate of this quantity $\widehat{s_b}(D)$ (Eq. 4). Substituting this value into Eq. 1 yields a noisy estimate of domain-level ideology $\widehat{\zeta_D}$, as shown in Eq. 7.

$$\begin{aligned} \widehat{s_b}(\ell) &= s_b(\ell) + \epsilon_{s,b} & (3)\\ \widehat{s_b}(D) &= \textstyle\sum_{\ell \in D} \widehat{s_b}(\ell) & (4)\\ &= \textstyle\sum_{\ell \in D} (s_b(\ell) + \epsilon_{s,b}(\ell)) & (5)\\ &= s_b(D) + \epsilon_{s,b}(D) & (6) \end{aligned}$$

$$\widehat{\zeta_D} = \frac{1}{\sum_b \widehat{s_b}(D)} \sum_{b \in [-2..+2]} b \cdot \widehat{s_b}(D) \tag{7}$$

Critically, the ratio in Eq. 7 is ill-behaved when the magnitudes of actual shares $s_b(D)$ and the noise $\epsilon_{s,b}(D)$ are similar. In such cases, because the Gaussian noise is zero-centered and can be negative, the denominator can approach zero, which inflates the metric (examples of this behavior are shown in the section below on link-level estimates). This scenario can also lead to pathological cases in which denominator is exactly zero (i.e., the noise exactly cancels the number of shares), resulting in discontinuities in the ideology estimate. Given the number of URLs in the dataset, these rare cases occur sufficiently often as to be problematic.

Analytically, Eq. 7 can be viewed as a ratio of Gaussian distributions, but these pathological cases results in this ratio having a Cauchy distribution, which has an undefined expected value. It is consequently difficult to isolate and correct for bias introduced by differentially private noise analytically. Fortunately, other work has examined this bias, and we rely on Hayya, Armstrong, and Gressis (1975), Evans et al. (2019), and Evans and King (2021) for their discussion of weighted averages in the face of noise.

In particular, if counts are normally distributed, we could treat this instance as a ratio of correlated, non-central normal distributions and use the result from Hayya, Armstrong, and Gressis (1975), to find the expected value of this ratio. In that case, as long as mean of the denominator's distribution is sufficiently large compared to the mean of numerator, bias in this expectation goes to zero. While we cannot assume normally distributed counts in this dataset (see Papakyriakopoulos, Serrano, and Hegelich (2020) for a discussion of log-normal distributions in social media engagement data), accounting for bias when the denominator is sufficiently large is supported by Evans and King (2021). Evans and King (2021) relies on Taylor approximation to expand a ratio of noised observations, leading to an upper bound on potential bias in this estimate, shown in Eq. 8, following from Eq. 2 of Evans and King (2021) where we replace K with the number of PPA bins. Specifically, as long as the number of shares $s_b(D)$ is sufficiently greater than the variance of the noise added, Eq. 8 goes to zero. Restated, as long as $s_b(D) >> \sigma_s$, or equivalently, $s_b(D)/\sigma_s >> 1$, bias in this metric should be negligible. Borrowing from signal processing, we refer to this ratio of engagement to noise as the signal-to-noise ratio (SNR), defined in 9.

$$bias < 4 \cdot \frac{\sigma^2}{(\sum_b s_b(D))^2} \tag{8}$$

$$SNR = \frac{(\sum_b s_b(D))^2}{\sigma^2} \tag{9}$$

Impacts of Noise via Simulation

As we show above, for a sufficiently high SNR, bias in our metric is negligible, but that analysis does not tell us what a sufficient-SNR regime might be. We thus turn to simulation to evaluate potential bias in the environment specific to the Condor dataset and construct two experiments: First, we evaluate whether the environment observed for popular domains in the Condor dataset are sufficient for our metric to be unbiased. Second, we examine the relationship between SNR and bias to get a sense for what levels of SNR and observed sharing are necessary to produce tight estimates of political ideology.

Estimating Bias for Popular Domains In the first simulation experiment, we test the hypothesis that $\zeta_D - \widehat{\zeta_D} = 0$, or whether the noise in Condor drives a significant difference in our estimates of ideology. We perform this analysis after observing engagement data for the top 1% most shared domains in the Condor dataset, which we select under the expectation that these popular domains achieve the necessary share-to-noise ratio. In the alternative case, i.e., that these domains do not have sufficient shares to be in the high-SNR regime, the noise added to this dataset may overwhelm any useful signal.

At a high level, each run in the simulation starts by drawing ideology scores ζ_D for a given number of domain observations n_{obs}. For each domain, we sample a count of the links to this domain u_D from a log-normal distribution and sample link-level ideology estimates for each link ζ_ℓ, according to a normal distribution centered at ζ_D for this domain. We then sample the number of shares for each link $s(\ell)$ in this domain, also from a log-normal distribution, and distribute these shares across the five PPA bins according to ζ_ℓ. This process yields a collection of domains with associated links and shares across PPA bins for each link, mirroring the Facebook collection prior to noise injection.

To simulate the noise-generate process, we add noise to each link's simulated shares $\widehat{s_b}(\ell)$ using the exact process outlined in the Condor codebook. We aggregate these link-level shares up to the domain level, and estimate ideology $\widehat{\zeta_D}$ from these noisy observations. Comparing this estimate from noisy sharing counts to the actual ideology yields an estimate of the bias added by the differentially private noise. Parameters for this simulation come from qualitative assessment of the Condor dataset and are shown in Table 2.

We then run the simulation with $n_{sim} = 100$ iterations, sampling $n_{obs} = 100$ domains per iteration, and calculate the mean bias $\zeta_D - \widehat{\zeta_D}$ and Monte-Carlo standard error for each iteration. Simulation results produce an estimated bias of -4.006×10^{-5} with a Monte-Carlo standard error of 5.4940×10^{-5}. Variance across simulation iterations is also small, at 3.0485×10^{-7}. These results shows the proposed estimator's bias is neither statistically significant nor is this difference practically significant on the $[-2, 2]$ scale of PPA.

A Note on Aggregation One may be tempted to first estimate ideology by calculating ζ_ℓ at the link level and taking the mean across all links in domain D to calculate ζ_D, which is more consistent with the metric provided in Bakshy, Messing, and Adamic (2015). This approach produces a much higher MC standard error in this case, however, as the sharing signal in a given link is generally much lower compared to the additive noise than in the aggregate. We provide guidance on when such link-level estimates are reasonable in a later section.

Relationships between SNR and Bias In the second simulation experiment, we examine the relationship between SNR and variance in our ideology estimator. This experiment fixes the number of links a domain has and varies the number of shares necessary to achieve a target SNR, defined in Eq. 9. We run this experiment with two fixed values for the number of links in a domain, first setting $u_D = 1024$ to evaluate SNR for domain-level aggregates and then setting $u_D = 1$ for cases where researchers want to study a single link. Varying SNR on the interval $[1, 1024]$ shows that an $SNR \geq 16$ results in tight estimates on ideology, regardless of whether we aggregate over many or few links. These metrics derive from $n_{sim} = 500$ runs of $n_{obs} = 100,000$ domains with uniformly distributed ideologies for each SNR.

Parameter	Description
ζ_D	Audience-level ideology of a domain D, drawn from a three-component Gaussian mixture model.
ζ_ℓ	The mean political page affinity for a given link $\ell \in D$, drawn from $N(\zeta_D, 0.25)$, with $\sigma = 0.5$ to provide separation between bins.
u_D	The number of hyperlinks to domain D for which we have sharing data, drawn from a log-normal distribution $LN(9, 1)$, as estimated from the Condor dataset.
$s(\ell)$	The number of shares for link ℓ, drawn from a log-normal distribution $LN(7, 1)$, as observed within the Condor dataset.
$s_b(\ell)$	The number of shares in political page affinity bin b for link ℓ, which we allocate from $s(\ell)$ by drawing 100 samples from $N(\zeta_\ell, 0.5625)$ and scaling up. $\sigma = 0.75$ is chosen so most mass is ± 1.5 from the mean.
$\epsilon_b(D)$	Noise added in to sharing in political page affinity bin b for domain D, drawn from $N(0, \sigma^2 \cdot 3 \cdot 7 \cdot 36)$ (σ is taken from the Condor codebook, and the multiplicative factors account for the aggregations along demographic and temporal bins; i.e., gender, age, and month).

Table 2: Simulation parameters, drawn from observation of the Condor dataset. Distributional values in this table are based on qualitative assessments of distributions within v7 of the Condor dataset, which covered 36 months.

A remaining question concerns the level of *observed* sharing needed for tight estimates. To this end, we have explored the maximal number of observed sharing (i.e., noised share counts), averaged over several link-sharing counts and privacy-protecting noise levels, necessary for tight bounds. This exploration suggests the relationship between sufficient observed sharing and aggregate noise (i.e., noise accumulated from differential privacy protections and aggregations over multiple links, demographic bins, and temporal timeframes) is linear in log-log space for a fixed SNR. Using this framework, we then estimate the relationship between this noise and a target observed sharing volume needed for tight estimates at $SNR = 16$ (sufficiently high to ensure tight estimates). This model is shown in Eq. 10, where 16 is the SNR; 5, 3, and 7 are number of PPA bins, gender bins, and age bins respectively; m is the number of months the aggregation covers; u_D is the number of links over which one is aggregating; and σ_{dp} is the noise added for differential privacy. This equation for $\widehat{s}(D)$, or the amount of observed sharing, allows us to set a lower bound on the volume of observed sharing necessary for stable estimates. This model also allows us to vary injected noise σ_{dp}, meaning we can estimate the minimum quantity of engagement one should observe for other types of activity in Condor as well.

$$\widehat{s}(D) \gtrsim 1.578 \cdot \sqrt{(16 \cdot 5 \cdot 3 \cdot 7 \cdot m \cdot u_D \cdot \sigma^2_{dp})} \qquad (10)$$

Audience Ideology for Popular Domains

We now use this metric and bounds on sharing to examine the distribution of audience ideologies among Facebook's top 1% most popular US domains among politically engaged Facebook users. This analysis covers 2,629 domains out of a possible 2,644 as 15 domains were excluded for having insufficient activity to achieve the target SNR relative to the number of links u_D.These excluded domains include `SoundCloud.com` and `ReverbNation.com`, which both have high numbers of unique links in Condor, leading to high differences between observed and needed sharing, making their estimates suspect.

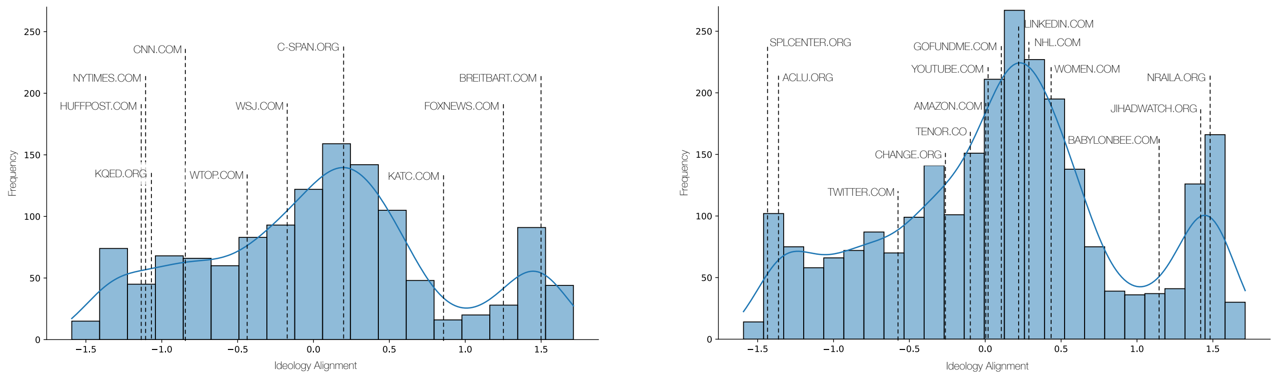
Figure 1 presents distributions of our estimated domain-level ideologies, with a selection of domain annotations, divided into news (1a) and non-news domains (1b). For news domains, we select domains that have ratings from the NewsGuard[2] trust rating service and exist in our top-1% set, resulting in 1,279 news-oriented domains. This collection shows a tri-modal distribution, with traditionally liberal news sources (e.g., CNN, the New York Times, and Huffington Post) on the left, conservative news on the right (e.g., Fox News, Breitbart), and more centrist reporting such as C-SPAN around the center. For non-news sites, we see many centrally oriented domains are primarily shopping, social networking, crowd-funding, and sports sites, whereas non-news domains in the ideological extremes are primarily activist organizations (e.g., the Southern Poverty Law Center splcenter.org or the National Rifle Association's Institute of Legislative Action nraila.org).

As our metric captures ideological leans of a domain's audience, in the context of news sources, Gentzkow and Shapiro (2010) suggests this measure should be highly correlated with the "slant" of these sources (which we indeed see in Figure 2d in the following section). For national news like the New York Times, Breitbart, etc., these sources are consistent with traditionally accepted partisan placement (e.g., as in Media Bias Fact Check). At the local level, we find local-affiliate news stations (e.g., WTOP in Washington, DC or KATC in Lafayette, Louisiana) are aligned with more ideologically moderate audiences, with KQED in Berkeley, CA having the most liberal and partisan audience of the local affiliates; the national media outlets, on the other hand, are more well-separated. This alignment among local affiliates is consistent with the literature (Bakshy, Messing, and Adamic 2015; Gentzkow and Shapiro 2010), as Berkeley, CA leaned heavily liberal in the 2016 presidential election, and Lafayette, LA leaned heavily conservative (Dottle 2019). The data also suggests, as seen in other work (Jurkowitz et al. 2020), a wider gap between the moderate and conservative components compared to the moderate and liberal components, suggesting the conservative media sites are more insulated from mainstream media.

Comparisons to Other Ideology Measures

To validate the audience ideology metric calculated from differentially private sharing data, we compare our results

[2]https://www.newsguardtech.com/



(a) News Domains ($N = 1,279$). Local affiliates are bounded between KQED.org and KATC.com.

(b) All Domains ($N = 2,629$). Many moderate domains are apolitical sources, e.g., social media, shopping, or sports.

Figure 1: Density plot of domain-level ideologies, for all (a) and news-only (b) domains (-2 is strongly liberal, and +2 is strongly conservative).

from these top-1% most popular domains to four similar domain-level measures. This first comparison is with measures of "partisan audience bias score" from Robertson et al. (2018), which uses ratios of shares from registered Republicans' and Democrats' Twitter accounts; we find a significant Pearson's correlation $\rho = 0.9295$ here for 1,675 domains (see Figure 2a). Our second comparison is with a homophily-based measure introduced by Eady et al. (2020) (Figure 2b), where we achieve a strong correlation $\rho = 0.9386$ across 154 domains. Our third comparison is with the similarly defined ideology scores introduced in Bakshy, Messing, and Adamic (2015), where we find the highest correlation ($\rho = 0.9522$) for 112 domains. Lastly, we compare against a content-based media slant estimate for 16 newspapers analyzed in Budak, Goel, and Rao (2016), where we find our lowest but still strong correlation ($\rho = 0.8675$). Despite the complexity introduced by differential privacy protections in Condor, our estimates are strongly correlated across all four of these comparison.

Beyond the Most Popular Domains

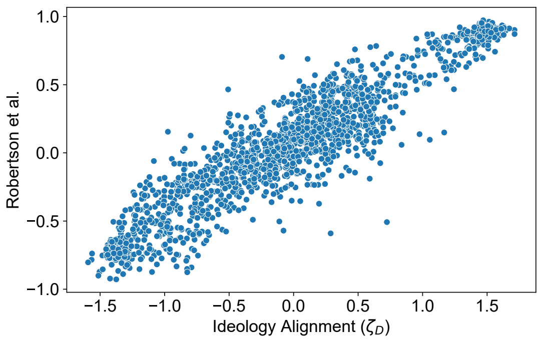
Above, we focus on the top-%1 of domains, as these domains are more likely to exceed the threshold established in Eq. 10 and because these domains are well-captured in other works on audience ideology. Our method is not restricted to only popular domains, however, as domains that are shared less often may still have sufficient signal to exceed our threshold: e.g., using Eq. 10, if we observe a domain with a single link and use only one month of data, that domain need only have about 906 observed shares (i.e., shares with added noise) to provide stable estimates. As the most recent iteration of the Condor dataset contains 363,738 domains over five years, one may then ask how many of these similarly exceed the thresholds we establish for this stability. This quantity is also important for the creators of the Condor dataset, as it can shed light on the tradeoff between differential privacy protections and data utility in downstream analysis.

To answer this question, we randomly sample 1,024 domains from the Condor dataset and measure the proportion that exceed the threshold of observed shares in Eq. 10. For all domains, we use the same $\sigma^2_{dp} = 14$ value and set m and u_D to the number of months for which the domain has data in Condor and the number of unique hyperlinks to that domain, respectively. Of these 1,024 domains, 43 exceed this threshold, accounting for the top 4.2% of the domains in this set. In comparison, over 99.4% of the top-1% of domains exceed this threshold. While this proportion is low, that still leaves in excess of 15 thousand domains that will produce stable ideology estimates using the method described above.

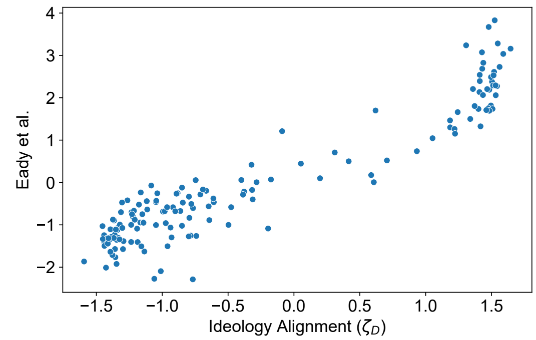
This result also has an important implication for the Condor dataset's construction more generally. While we note a couple of ways one might increase this proportion through relaxing constraints or focusing on link-level estimates (as we do in the following section), it is also true that the application of differentially private noise to the Condor dataset is done with limited insight into the downstream impact this noise has on analyses. Hence, this finding motivates a call to Facebook to revisit its privacy budget and investigate the balance between adding noise and reducing utility of this large dataset.

Link-Level Audience Ideology Estimates

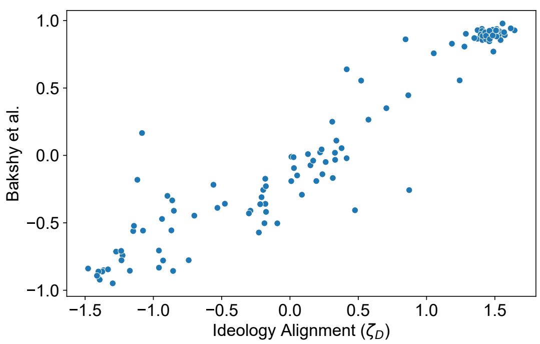
In the preceding section, we have focused on demonstrating the validity of the audience-ideology metric by showing the bounds in SNR for which ideology estimates are tight and comparing our domain-level metrics to several extant, non-differential-privacy-protected datasets. This metric is not specific to domain-level metrics, however, and is equally amenable to estimating audience ideology at the individual link level, as in Bakshy, Messing, and Adamic (2015). This link-level analysis is a major advantage of the Condor dataset as the scale at which it provides these engagement metrics alleviates sparsity issues, which are a major barrier to link-level analyses in other sources. That is, the Condor dataset provides a much larger volume of link-level data than academic researchers are generally able to access, allowing for novel insights into ideological distribu-



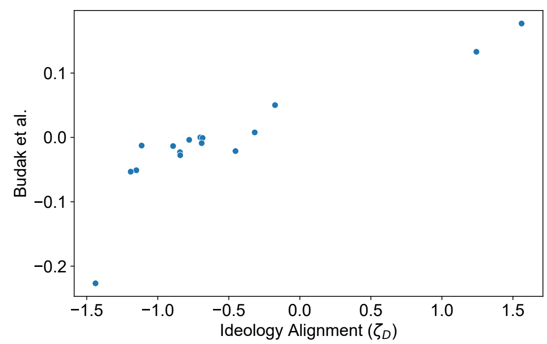
(a) Versus Robertson et al. ($N = 1,675, \rho = 0.9295$).



(b) Versus Eady et al. ($N = 154, \rho = 0.9386$).



(c) Versus Bakshy et al. ($N = 112, \rho = 0.9522$).



(d) Versus Budak et al. ($N = 16, \rho = 0.8765$).

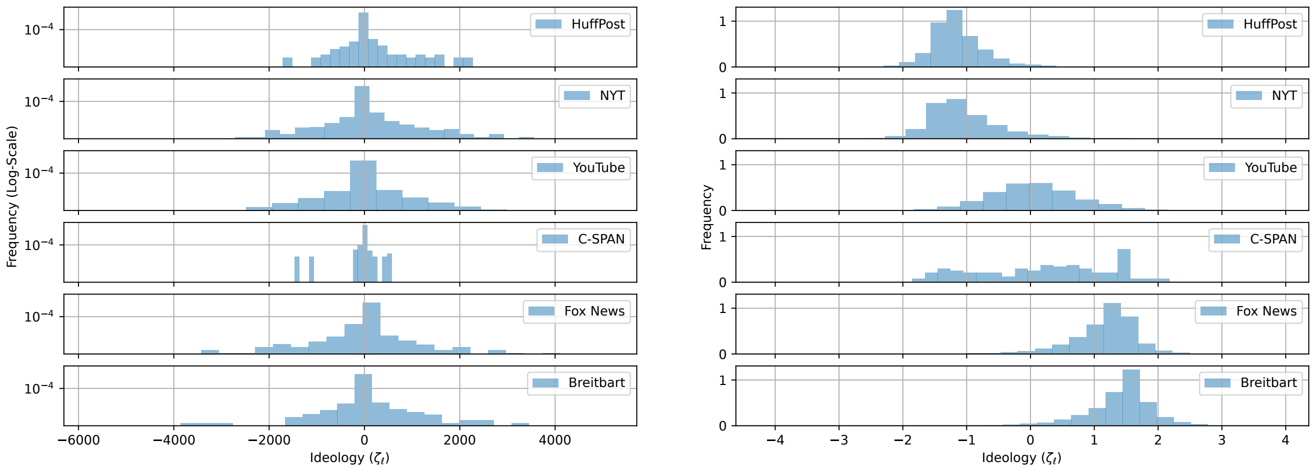
Figure 2: Comparisons of the proposed domain-level ideology metric and extant domain-level measures. Results show strong correlations between our proposed metric and extant literature. Of these four methods, Budak, Goel, and Rao (2016) provides a content-based measure of media slant, whereas the remaining three provide homophily-based measures.

tions across individual links rather than domain-level aggregates. That said, noise added to engagement metrics for individual links may be relatively high compared to domain-level aggregates, as many individual links likely do not receive sufficient engagement alone to support ideology estimates using our proposed metric.

To illustrate this point, we estimate audience ideologies for individual links to six popular domains across the ideological spectrum. Of these domains, five are news sources, which we order from left to right–liberal to conservative: Huff Post, the New York Times, C-SPAN, Fox News, and Breitbart. Traditionally, Huff Post and the New York Times are considered left-leaning sources, whereas Fox News and Breitbart are right- and far-right sources; in contrast, C-SPAN is a non-profit, non-partisan source that primarily covers the US House of Representatives. We also include YouTube as our sixth domain, given its substantial role in the online news ecosystem. For each of these domains, we calculate audience estimates using the ninth iteration of the Condor dataset, covering 2017-2021 and show link-level distributions for all links in each domain (Figure 3a) and only for those links with sufficient engagement (Figure 3b) as estimated by Eq. 10; i.e., $s(\ell) > 7,014$. Tables 3a and 3b show summaries of these figures as well. Distributions of audience ideology when we use all links from each domain consistently show extreme variation, with YouTube showing the widest range, from -5,760.0 to 5,167. In contrast, focusing only on links with with SNR $>$ 16 yield more stable averages and more informative distributions. Interestingly, links to YouTube videos appear symmetrically distributed in their audience ideologies, and C-SPAN shows wider variation (though still constrained to between $[-1.9, 2.2]$). The remaining four domains, traditionally considered partisan-leaning, exhibit the expected partisan distributions, with the majority of links falling on one side of the ideological spectrum. Some links from these domains do cross the ideological divide though; e.g., 1,169 of the links in Figure 3b from the New York Times have an audience ideology measure $\zeta_\ell > 0$. One article in particular, "I Wanted to Be a Good Mom. So I Got a Gun," was shared by a solidly right-leaning audience ($\zeta_\ell = 2.2109$).[3] Likewise, Breitbart has 255 links shared among left-leaning audiences ($\zeta_\ell < 0$), with an article concerning Harvey Weinstein seeing the most liberal audience.[4]

[3]https://www.nytimes.com/2018/03/05/opinion/mom-gun-safety-intruder.html

[4]http://www.breitbart.com/big-hollywood/2017/10/12/bombshell-weinstein-board-knew-about-harveys-payoffs-in-



(a) Link-level ideology distributions for all links in each domain. (b) Ideology distributions for only links with more than 7,014 shares.

Figure 3: Histograms of link-level ideologies for six major web domains, ordered from left to right as most liberal to most conservative–Huff Post, New York Times, YouTube, C-SPAN, Fox News, and Breitbart–using all links in Condor (a) and only links with sufficient observed sharing (b). Distributions using all links show extreme variation, whereas links with sufficient sharing produce reasonable distributions in the range $[-4, 4]$.

Domain	**Mean ζ_ℓ**	σ	N
HuffPost	-0.6434	34.83	31,303
NYTimes	-0.3288	42.67	122,972
YouTube	0.0009	44.11	10,584,364
C-SPAN	-1.4369	55.43	1,458
FoxNews	0.6920	44.33	119,010
Breitbart	1.1475	45.27	82,536

(a) Using all links in the dataset. Ideology measures exhibit high variance, with averages towards zero.

Domain	**Mean ζ_ℓ**	σ	N
HuffPost	-1.1511	0.4197	5,380
NYTimes	-1.0318	0.5988	18,387
YouTube	0.0462	0.6749	189,436
C-SPAN	0.2183	1.0472	117
FoxNews	1.1632	0.5152	20,576
Breitbart	1.4050	0.5250	13,447

(b) Using only links with more than 7,014 observed shares. Variation in ideology is more constrained at the cost of smaller N.

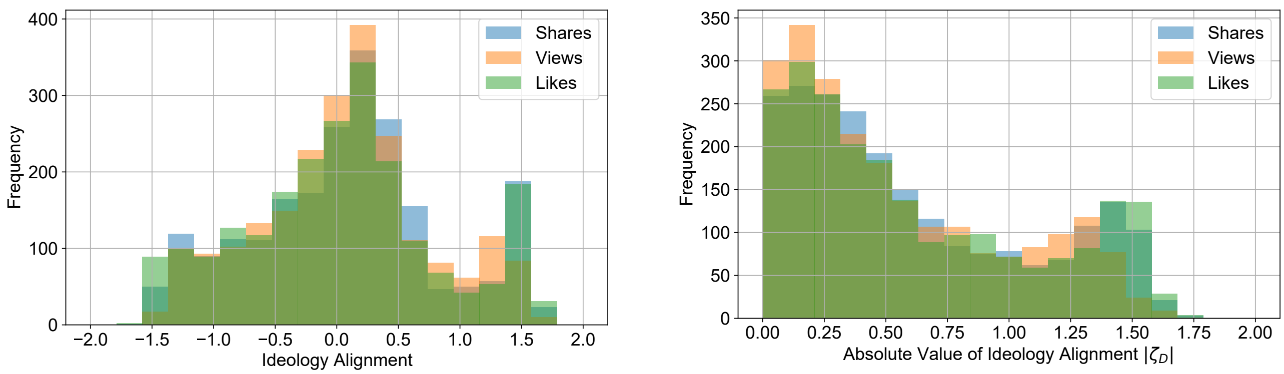
Table 3: Summary statistics for link-level audience ideology for six major domains. Consistent with Figure 3, ideology measures using all links (a) exhibits high variance, potentially masking useful structure, which emerges when we constrain links to those that are sufficiently popular (b).

Comparing Shares, Likes, and Views

Prior sections use sharing as the primary mode of engagement, so we can compare against similar methods on Facebook (e.g., Bakshy, Messing, and Adamic 2015) and Twitter (e.g., Robertson et al. 2018; Eady et al. 2020). Concerns with these measures include 1) sharing as a proxy for viewership and 2) counter-attitudinal sharing, wherein an individual shares a particular article to criticize it. Often, *exposure* to information is more important than who is sharing information, but sharing activity is more readily available, so it is used in place of exposure. Similarly, while prior work shows criticism is one of the primary motivations to share content in political discourse (Kim, Jones-Jang, and Kenski 2020), counter-attitudinal sharing is relatively rare (An, Quercia, and Crowcroft 2014), so it is generally ignored. The Condor dataset is not limited to sharing though, as it contains measures of likes, views, and other activity, though the added noise varies for these actions (e.g., $\sigma = 10$ and $\sigma = 2,228$ for likes and views, respectively). We can therefore compare whether the population *sharing* a particular domain is significantly different from the population who *likes* or *views* this content, potentially mitigating concerns around share-based measures. We thus compute audience ideology metrics for 2,227 domains across three engagement types–sharing, viewing, and liking– and compare them in Figure 4.

This figure demonstrates limited statistically significant differences in ideological distributions exist among share-, view-, and like-based measures – supported by a one-way ANOVA test ($F(2, 2225) = 1.899, p = 0.1498$). That is, in comparing the distributions of inferred domain-level ideology metrics using the three activity types, we see no significant differences in these audience-ideology values. Correlation among all three metrics is also very strong (> 0.97),

2015-what-about-disney/



(a) Ideology Measure Using Views, Shares, and Likes.

(b) Distribution of Absolute Values in Ideology Measures.

Figure 4: Comparison of Ideology Metrics Based on Shares, Views, and Likes. Overall distributions (a) show no significant deviation; distributions of absolute values (b) suggests domains are viewed by more moderate audiences than those who share and like them.

suggesting differences based on engagement are less impactful than the cross-method analysis shown in Figure 2. Hence, despite potential concerns around share-based measures as proxies for exposure and counter-attitudinal sharing, the ideological alignment of audiences sharing, viewing, and liking these domains are statistically indistinguishable.

Separate from these concerns, one may expect differences in the extremes of these ideology distributions, as political sharing on Facebook is relatively rare (Bakshy, Messing, and Adamic 2015). We therefore test an alternative measure by taking the absolute value of the ideology metric $|\zeta_D|$, to measure potential partisanship rather than liberal/conservative ideology. Comparing this partisanship metric based on shares, views, and likes (Figure 4b) shows a more significant difference among these three distributions ($F(2, 2225) = 17.95, p << 0.001$). A post-hoc Tukey test shows a moderating effect in viewership in that a domain's viewing audience appears significantly more moderate than its sharing and liking audiences ($p < 0.001$).

A Note on Views We note that the interpretation of "views" in Condor does not directly capture the volume of individuals who have visited and viewed a given URL outside of the Facebook platform. Instead, the Condor codebook defines "views" as the "number of users who viewed a post containing the URL" (Messing et al. 2020). That is, an individual may view a link outside the Facebook platform, and this view would not be captured in Condor's "view" count. Conversely, an individual may view a post in Facebook containing a specific link without visiting the link, and this interaction would be captured in Condor's "view" count–actually visiting the link is captured by the "click" count. While this interpretation omits off-platform engagement, Condor's operationalization of "views" still captures important network-driven exposure aspects that other work largely is forced to omit, given the commercial sensitivity of this measure. Hence, the result above should be interpreted as: The audience exposed to a particular domain within Facebook is significantly more moderate the audience that shares and likes that domain.

Threats to Validity

Though the Condor dataset is a milestone in the availability and transparency of social media data, concerns remain around how such data is collected and protected. First, the process by which the Condor dataset is constructed is opaque to researchers in that parties outside of Facebook are not allowed to inspect the code used to select URLs or calculate the metrics like PPA. As a result, researchers are forced to trust that Facebook's URL selection process is correct. Likewise, researchers are given limited insight into how much data is omitted from the Condor dataset because the links do not meet the 100-unique-user threshold on public shares. While internal Facebook developers have made some data available about this threshold's relation to the distribution of on-platform links, external review of the data pre- and post-application of the privacy-protecting noise remains unavailable. This latter issue is of particular concern as, in the fall of 2021, external researchers identified a flaw in the Condor dataset that significantly undercounted engagement in the US (Alba 2021). This flaw led to Condor's omission of engagement from US users whose political preferences (i.e., PPA bin) could not be identified; that is, while other demographic bins could be null to capture shares from, say, individuals with an unknown gender, no data existed in the dataset for the many users who did not follow sufficient political pages to have an identifiable PPA value. Though this paper was unaffected by this error (as we ignore shares from null-PPA users), and Facebook has since corrected this issue, the lack of transparency around Condor's creation and population remains a problem.

Second, while the privacy protections applied to the Condor dataset serve a crucial purpose, how these protections impact research methods remains an open question. Our proposed metric requires a sufficient level of activity to overcome additive noise, which means many important but rare phenomena may be masked. Consequently, domains and links shared among extreme partisan audiences may be included in the dataset but have insufficient signal for useful analyses. More worryingly, this masking could be asymmetric and result in ideological bias in what links are included. To examine this possibility, we have examined 500 domain-level ideology scores from Bakshy, Messing, and Adamic (2015) and assess the overlap between that work and our set. Using a logistic regression model to assess whether a domain's audience ideology scores predict its inclusion in our dataset, we find no statistically significant relationship between the two factors. For less popular phenomena, however, this question of bias remains open, but a fundamental tension exists between these rare instances and the differential privacy protections, as these protections add more noise to these rare instances to prevent identification. How these two factors interact needs to be a subject of future work.

Third, we stress our ideology metrics measure the *audience* of a domain/link, not the actual content of the domain or link itself. While much of the related work in this area makes similar assessments (e.g., Bakshy, Messing, and Adamic (2015), Robertson et al. (2018)), it is important to note that content in some of these sites may not be overtly partisan but are more attractive or known to partisan audiences. As noted in Gentzkow and Shapiro (2010), the ideological slant of a media outlet's audience does affect choices of what that outlet covers and how it does so, but this distinction between content and audience is important for interpreting this work.

Ethics and Competing Interests

This work's intent is to provide a broader audience with an example for working with social media and digital trace data that has been protected with differential privacy techniques. Though this work is focused on audience ideology, the methods are equally applicable to aggregations across other demographic bins or activities. Similarly, our focus on ideology results in a US-oriented analysis, as the Condor dataset only provides ideology-relevant PPA assessments for US users. While a clear limitation of this work, it does hint at the need for broader perspectives on how such left/right scales can be generalized to other national contexts, as discussed in Lo, Proksch, and Gschwend (2014). Ultimately though, teams internal to Facebook would have to extend the Condor dataset to include page-affinity scores for non-US audiences.

Regarding ethics in research, this work and the Condor dataset more generally has some considerations worth noting. Condor's privacy protections provide value in preventing identification of individual users' actions on the platform but at the cost of obfuscating rare phenomena. Vulnerable and minority groups who might be over-represented in these rare instances are potentially disproportionately impacted by these protections, as researchers balance preserving privacy with studying how behaviors on the platform may impact these groups. More work is needed to assess how platforms like Facebook interact with these populations and how we might study these interactions while still providing a reasonable level of protection for these users.

While we claim no conflicts of interest, for transparency, we note that one of the authors of this work has received funding from Facebook related to the Social Science One initiative. This funding was not for this work, and while Facebook has had the opportunity to review this work prior to publication as part of the Social Science One agreement, they do not have authority to prevent publication. Finally, this work was reviewed by university internal review boards as a prerequisite for gaining access to the Condor dataset.

Conclusions

Through the above assessments of our proposed audience-ideology measure, based on a simple weighted average of online behavior across ideologically grouped audiences, this paper presents three core contributions: First, this measure and its assessment provide guidance for researchers seeking to use differential privacy-protected digital trace data in analyses of online political behaviors, which we make more compelling by demonstrating agreement with other published measures that do not have these protections. Second, we extend this work on domain-level analyses to demonstrate how our proposed metric can provide insights at the individual link level, which is often made difficult by concerns of sparsity in other datasets. Third, we contribute to studies of media slant and online political engagement by extending this analysis to other online types of online activity beyond just sharing – i.e., views and likes. As Condor is the largest dataset of its kind and the primary mode of access to Facebook data for researchers unaffiliated with Facebook, the endogenous metric for audience ideology we provide – along with the related insights for SS1 researchers looking to leverage this unique dataset – may accelerate research in this space.

Data Availability

Access to the Social Science One dataset used in this analysis is governed by the Research Data Agreement made available as a joint effort between Facebook and the Social Science One Consortium: https://socialscience.one/research-data-agreement.

References

Alba, D. 2021. Facebook sent flawed data to misinformation researchers. *The New York Times*. https://www.nytimes.com/live/2020/2020-election-misinformation-distortions/facebook-sent-flawed-data-to-misinformation-researchers.

An, J.; Quercia, D.; and Crowcroft, J. 2014. Partisan sharing: Facebook evidence and societal consequences. *COSN 2014 - Proceedings of the 2014 ACM Conference on Online Social Networks*, (Figure 1): 13–23.

Bakshy, E.; Messing, S.; and Adamic, L. A. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239): 1130–1132.

Barberá, P.; Jost, J. T.; Nagler, J.; Tucker, J. A.; and Bonneau, R. 2015. Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10): 1531–1542.

Benevenuto, F.; Rodrigues, T.; Cha, M.; and Almeida, V. 2009. Characterizing user behavior in online social networks. *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, 49–62.

Budak, C.; Goel, S.; and Rao, J. M. 2016. Fair and Balanced? Quantifying Media Bias through Crowdsourced Content Analysis. *Public Opinion Quarterly*, 80: 250–271.

Bun, M.; and Steinke, T. 2016. Concentrated Differential Privacy: Simplifications, Extensions, and Lower Bounds. In Hirt, M.; and Smith, A., eds., *Theory of Cryptography*, 635–658. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN 978-3-662-53641-4.

D'Orazio, V.; Honaker, J.; and King, G. 2015. Differential Privacy for Social Science Inference. *SSRN Electronic Journal*, 02138: 1–44.

Dottle, R. 2019. FiveThirtyEight.com: Where Democrats And Republicans Live In Your City. https://projects.fivethirtyeight.com/republicans-democrats-cities/. Accessed: 2023-04-17.

Eady, G.; Bonneau, R.; Tucker, J. A.; and Nagler, J. 2020. News Sharing on Social Media: Mapping the Ideology of News Media Content, Citizens, and Politicians. https://osf.io/ch8gj/. Accessed: 2023-04-17.

Evans, G.; and King, G. 2021. Statistically Valid Inferences from Differentially Private Data Releases, II: Extensions to Nonlinear Transformations. https://gking.harvard.edu/files/gking/files/wt.pdf. Accessed: 2023-04-17.

Evans, G.; and King, G. 2022. Statistically Valid Inferences from Differentially Private Data Releases, with Application to the Facebook URLs Dataset. *Political Analysis*, 1–21.

Evans, G.; Schwenzfeier, M.; King, G.; Schwenzfeier, M.; and Thakurta, A. 2019. Statistically Valid Inferences from Privacy Protected Data. https://gking.harvard.edu/files/gking/files/udp.pdf. Accessed: 2023-04-17.

Gentzkow, M.; and Shapiro, J. 2010. What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1): 35–71.

Gentzkow, M.; and Shapiro, J. M. 2011. Ideological segregation online and offline. *Quarterly Journal of Economics*, 126(4): 1799–1839.

Golovchenko, Y.; Buntain, C.; Eady, G.; Brown, M. A.; and Tucker, J. A. 2020. Cross-Platform State Propaganda: Russian Trolls on Twitter and YouTube during the 2016 U.S. Presidential Election. *The International Journal of Press/Politics*, 0(0): 1940161220912682.

Gong, W.; Lim, E. P.; and Zhu, F. 2015. Characterizing silent users in social media communities. *Proceedings of the 9th International Conference on Web and Social Media, ICWSM 2015*, 140–149.

Hayya, J.; Armstrong, D.; and Gressis, N. 1975. A Note on the Ratio of Two Normally Distributed Variables. *Management Science*, 21(11): 1338–1341.

Jurkowitz, M.; Mitchell, A.; Shearer, E.; and Walker, M. 2020. U.S. Media Polarization and the 2020 Election: A Nation Divided. Technical report, Pew Research Centre. https://www.journalism.org/2020/01/24/americans-are-divided-by-party-in-the-sources-they-turn-to-for-political-news/.

Kim, D. H.; Jones-Jang, S. M.; and Kenski, K. 2020. Why Do People Share Political Information on Social Media? *Digital Journalism*, 0(0): 1–18.

Lo, J.; Proksch, S.-O.; and Gschwend, T. 2014. A Common Left-Right Scale for Voters and Parties in Europe. *Political Analysis*, 22(2): 205–223.

Messing, S.; Degregorio, C.; Hillenbrand, B.; King, G.; Mahanti, S.; Mukerjee, Z.; Nayak, C.; Persily, N.; State, B.; and Wilkins, A. 2020. Facebook Privacy-Protected Full URLs Data Set. Technical report, Social Science One, New York City.

Nonnecke, B.; and Preece, J. 2000. Lurker demographics: Counting the silent. *Conference on Human Factors in Computing Systems - Proceedings*, 2(1): 73–80.

Papakyriakopoulos, O.; Serrano, J. C. M.; and Hegelich, S. 2020. Political communication on social media: A tale of hyperactive users and bias in recommender systems. *Online Social Networks and Media*, 15.

Preece, J.; Nonnecke, B.; and Andrews, D. 2004. The top five reasons for lurking: Improving community experiences for everyone. *Computers in Human Behavior*, 20(2): 201–223.

Robertson, R. E.; Jiang, S.; Joseph, K.; Friedland, L.; Lazer, D.; and Wilson, C. 2018. Auditing partisan audience bias within Google search. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW).

Wood, A.; Altman, M.; Bembenek, A.; Bun, M.; Gaboardi, M.; Honaker, J.; Nissim, K.; O'Brien, D.; Steinke, T.; and Vadhan, S. 2020. Differential Privacy: A Primer for a Non-Technical Audience. *Vanderbilt Journal of Entertainment & Technology Law*.