



Unmasking Social Media Accountability

An Exploration of Technical Anonymity, Partisan Media Engagement, and Sentiment During COVID-19

Aishwarya Balaji, Katrina Liu, Shraavasti Bhat

About our project

We hope to understand the correlation between user's **anonymity**, **sentiment**, and engagement with **partisan** media during the COVID pandemic using Twitter Data (Pre-X).

Background

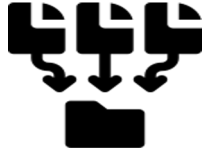
- Social Identity model of Deindividuation Effects (**SIDE**)
 - Anonymity on digital platforms allows individuals to feel less accountable for their actions (Zimbardo et al, 1969 & Alonzo et al, 2004)
- Commonly-used anonymity classification frameworks are **single factor** (Peddinti et al, 2014)
- Exposure to **partisan** media leads to extreme **political attitudes** (Earle et al, 2022)
- Media outlets can be classified based on their **partisanship**
 - Left, left-center, center, right-center, and right (Faris et al, 2017)

Our Hypothesis

H1: Higher **anonymity** facilitates more extreme **sentiment** (:: SIDE)

H2: More extreme **partisan** media engagement correlates to more extreme **sentiment** (:: Earle et al)

Research Methods



**Data Extraction +
Wrangling**



News Tweets about
COVID



User Profiles



User Replies



**User and Outlets
Classification**



Anonymity Levels
(Scores)



Political Spectrum



Sentiment Analysis
Using **VADER**



**Data
Analysis**



Partisanship
Score



Anonymity Scores

Data Extraction

98 News Outlets

- Removed suspended accounts, repeats, private accounts

30 COVID-related threads per news outlet

- All **replies** + **profiles**
- During the first 100 days of the covid outbreak



combined_replies.csv

~13000 replies

id	text	conversation_id	author_id	...
12482519 87470188 546	@dcexaminer https://t.co/ki4YjBLOEU	12482511581473341 52	300423 184	...



combined_profiles.csv

~9800 profiles

username	name	location	url	...
DevLilAidRy	Andrea Allen	United States	null	...

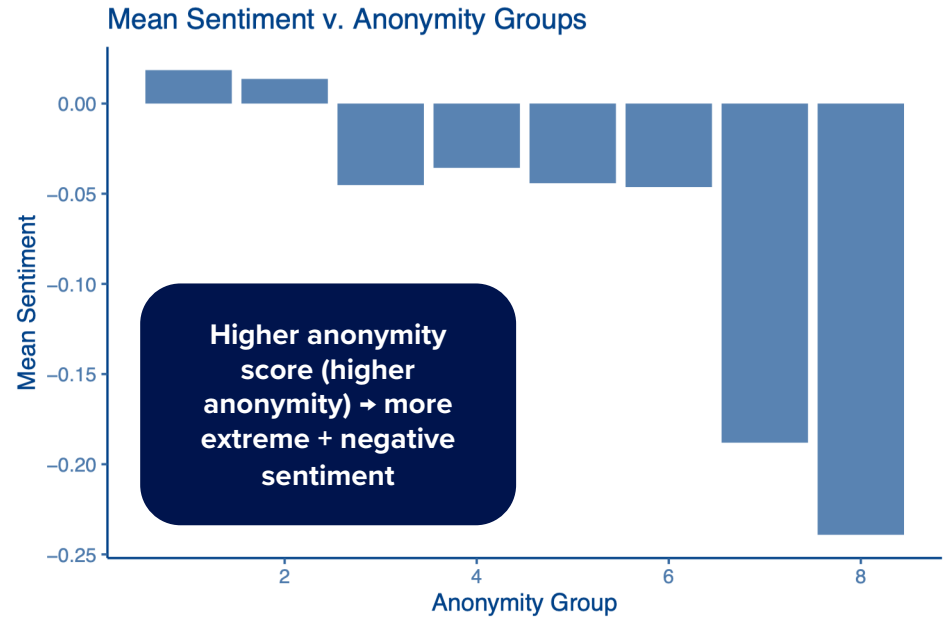
Anonymity Classification

- Created a multi-factor **anonymity scoring framework** based on Twitter profile features
 - e.g. username, bio
- Verified by manual evaluation on a small sample, will further verify using:
 - Amazon Mechanical Turk
 - Peddinti et al (2014) & Peddinti, Ross et al (2017)
 - Few-shot GPT-3.5 Turbo
 - Brown et al (2020)

Results

H1: Higher **anonymity** facilitates more extreme **sentiment**

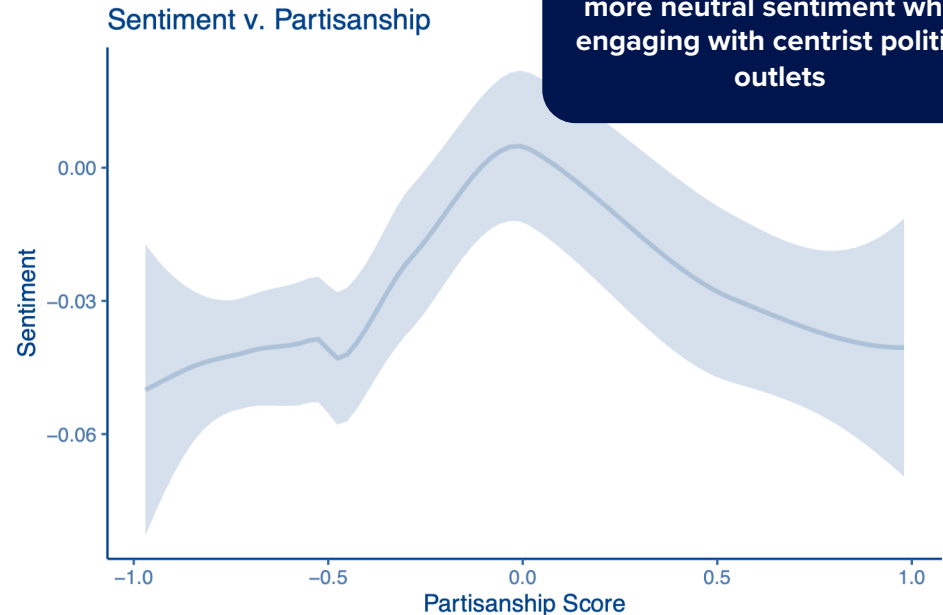
- 1 Grouped the profiles into 8 anonymity groups per profile feature
- 2 Plotted the mean sentiment of each group v. anonymity groups
- 3 Performed t-test ($p > 0.05$) on the least and most anonymous groups
 - Therefore, cannot reject the null hypothesis



Results

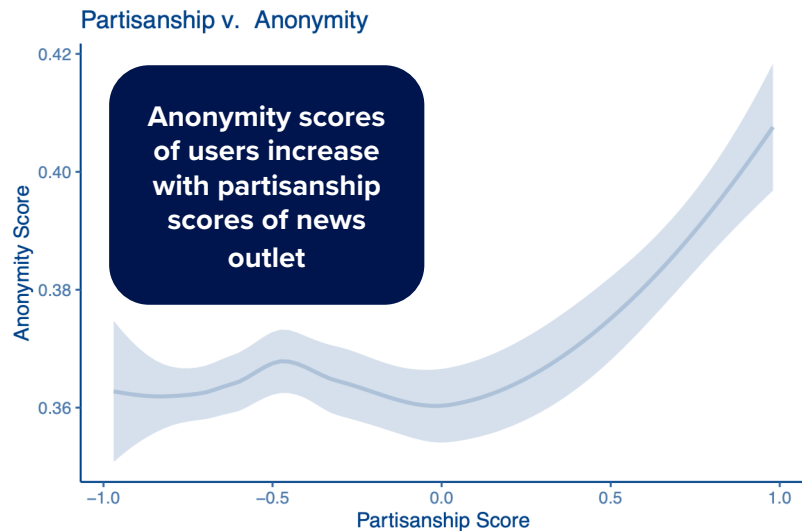
H2: More extreme **partisan media engagement** correlates to more extreme **sentiment**

- 1 Plotted sentiment score v. partisanship of media outlets
- 2 Grouped the outlets into 5 groups
 - Left, left-center, center, right-center, and right (Faris et al, 2017)
- 3 Performed t-tests ($p < 0.05$) between polar and neutral groups
 - Displayed statistical significance

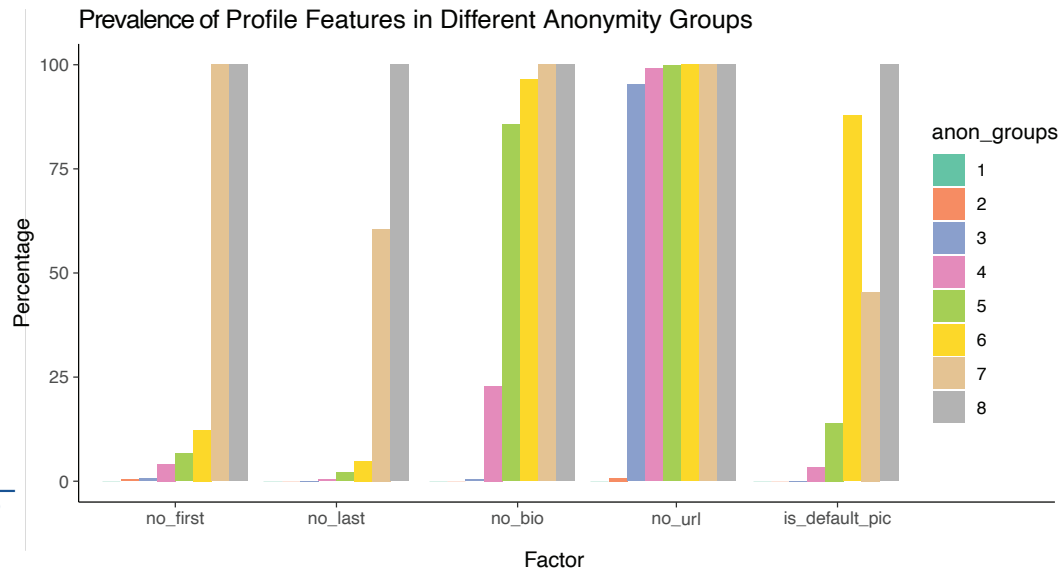


Other Findings

Anonymous accounts tend to interact with media displaying **right-leaning partisanship**



Only the **least** anonymous group tend to include **external URLs** in profile



Conclusion

- H1: Higher **anonymity** trending towards more extreme **sentiment**
- H2: More extreme **partisan** media engagement correlates to more extreme **sentiment**
- Other findings:
 - **Anonymous** accounts tend to interact with media displaying higher **partisanship**
 - Only the least **anonymous** group tend to include **external URLs** in profile

Significance

- Improve social media **platform design**
 - Inherently hold users more accountable by increasing mandatory profile features, reducing the effects of SIDE
 - Increase online psychological safety by reducing profile features
 - e.g. Sidechat
- Determine user credibility (anonymity score)
 - Determine **stock** perception based on more credible (less anonymous) profiles
 - Perform **market research** on credible profiles to accurately determine product interest
 - Suggest **medical** benefits to qualified (less anonymous, more real) users



Acknowledgement(s):

Preliminary Study Supervisor: Professor Metaxa

Preliminary Study Collaborators: Siddharth, Caia



Appendix

Multi-factor Anonymity Scoring Framework

- Anonymity Factors:
 - Legitimacy of the (1) username and (2) name
 - Presence of a (3) Bio, (4) External URL, (5) Location, and (6) Profile Photo
 - (7) User Creation Date
 - (8) Verification Status



combined_profiles.csv

~9800 profiles

username	name	location	url	...
DevLilAidRy	Andrea Allen	United States	null	...