

Networks and Social Media

CS 224C Spring 2024

Instructor:

Mobility Networks

(Nilforoshan et al., 2023, "Human mobility networks reveal increased segregation in large cities")

Interactions among individuals



Human mobility networks reveal
increased segregation in large cities



Socioeconomic Segregation



Large cosmopolitan areas
(# of Residents)

Human mobility networks reveal increased segregation in large cities

Why this research is important?

“Extreme Segregation is Costly”

“Decrease Economic Mobility”

“Foster Health Problems and Political Polarization”

Why haven't this been solved?

Reform Design to Rise Urbanization → Cosmopolitan Mixing Hypothesis:

“In **large c**
constrain
diverse in
reducing

Not Tested → No data to measure each of the
important variables

ersity,
ill bring
er,

Why can we test it now? What is “New”?

Measurement for: **Human mobility networks**

SafeGraph – de-identified GPS location pings from mobile phones

1,570,782,460 Edges (interaction)

9,567,559 Nodes (individuals)

382 Metropolitan Statistical Areas

2,829 Counties in the U.S.



- Where people go;
- When they go there;
- Whom they have encountered.

Why can we test it now? What is “New”?

Measurement for: **Human mobility networks**

Edge between a **pair of nodes** signifies that the **two individuals crossed paths** with and encountered each other

1. **Time**
2. **Location Coordinates**
(Latitude, Longitude)

Why can we test it now? What is “New”?

Measurement for: **Socioeconomic Segregation**

SafeGraph, CoreLogic, Zillow → Socioeconomic Status for Individual

- SafeGraph: Identify coordinates of the Nighttime Fixed Location (with a set timeframe & radius for movements)
- CoreLogic: Identify properties near the coordinates
- Zillow: Identify the rent estimates for the property

Coordinates → Properties → Rent Estimates → SES for Individual

Why can we test it now? What is “New”?

Measurement for: **Exposure Segregation**

Exposure: two devices being within D metres (50m) of each other within T minutes (5min)

Segregation of Each Geographical Region:

Pearson Correlation between a person's SES and the mean SES of everyone to whom they are exposed through a path crossing
(0: Perfect Integration → 1: Complete Segregation)

Why can we test it now? What is “New”?

Measurement for: **Exposure Segregation**

New Approach

- capturing the diversity of person-to-person exposures localized in space and time.
- Measured by correlation between a person’s SES and the **mean SES of everyone to whom they are exposed through a pass crossing**

Conventional Approach

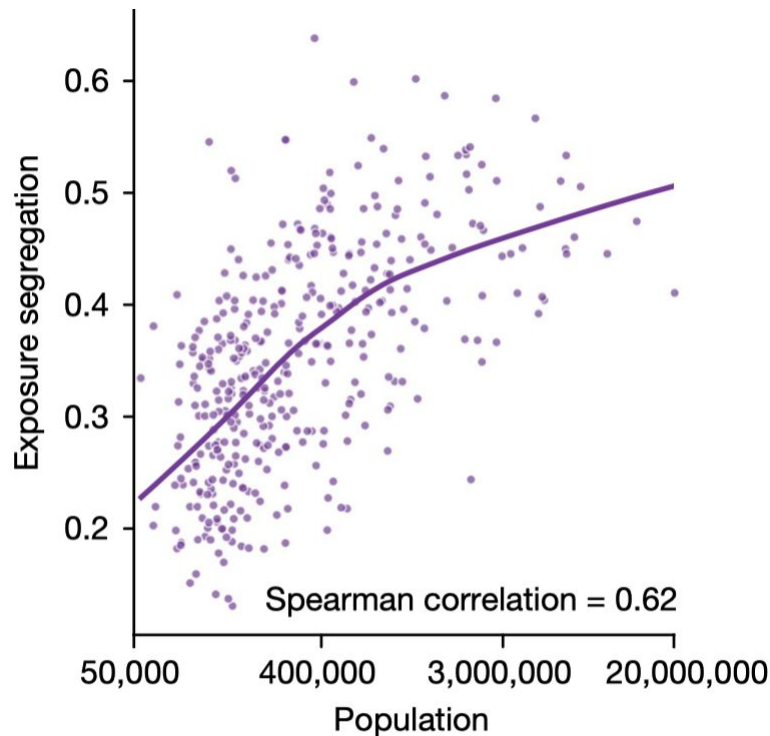
- the neighbourhood sorting index
- **assumes that individuals are exposed to other residents only within their home census tract** (occur uniformly among co-residents of the same home tract)

38% Differences in Median ($P < 10^{-4}$, 95% CI = 37–41%)

What can we conclude?

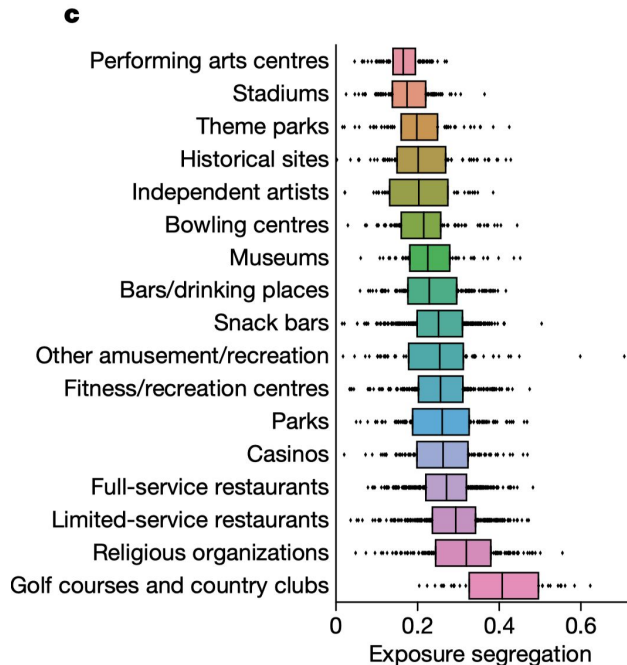
- Exposure Segregation is **Higher** in Large MSAs
- 10 Largest MSAs by population size are 67% more segregated than small MSAs with fewer than 100,000 residents

~~Cosmopolitan Mixing Hypothesis~~



Potential Explanation of the Results

Measurement for: **Exposure Segregation across different space**

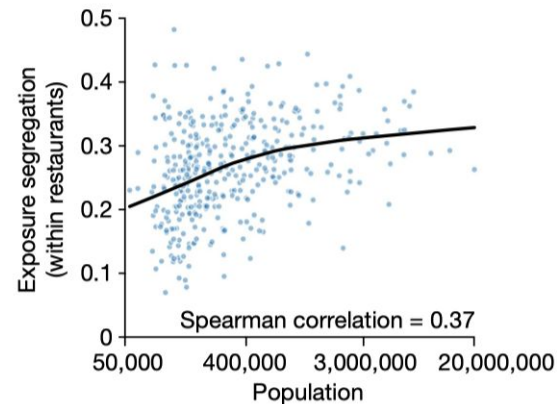
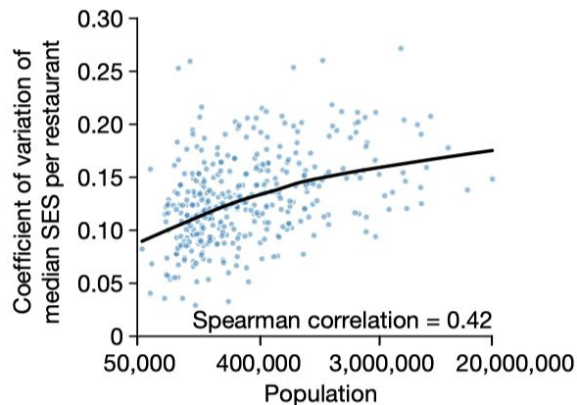
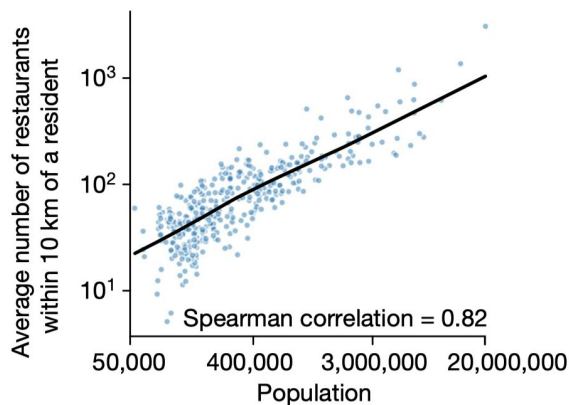


Variability in Point of Interest in Exposure Segregation because:

Serve Small & Socioeconomically Homogeneous Communities

Potential Explanation of the Results

Measurement for: **Exposure Segregation across different space**



Restaurant

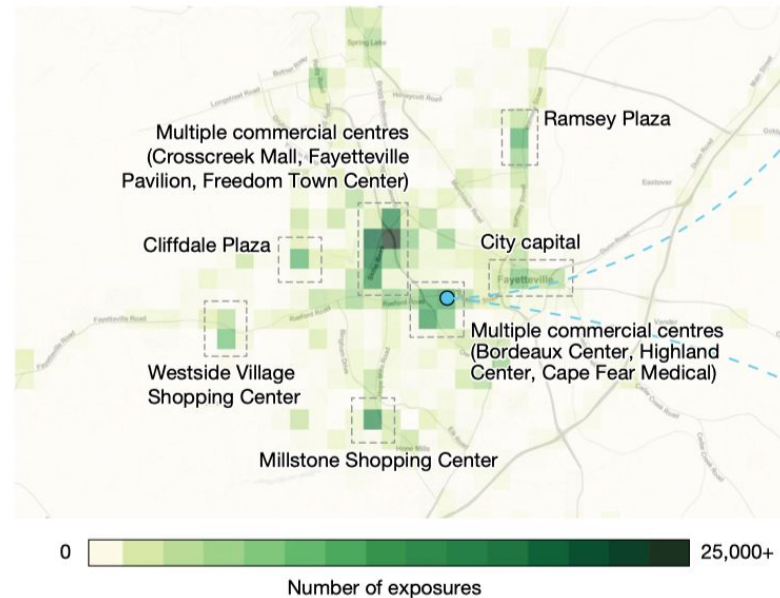
What Can We Do?

Segregation could be mitigated when “hubs” are positioned in Close Proximity to Diverse Neighbourhoods → **Bridging Index**

Diversity within Clusters → Clusters of individuals by the nearest hub to their home

the majority (56.9%, 95% CI = 56.9–56.9%) of exposures across all 382 MSAs occur in close proximity (within 1 km) to a **commercial centre**

(With controls for race, population size, economic inequality and many other variables)



What Can We Do?

Segregation could be mitigated when “hubs” are positioned in Close Proximity to Diverse Neighbourhoods → **Bridging Index**

Cities can **mitigate** this **segregative effect** by **having hubs that are located in bridging zones** drawing in people from diverse neighbourhoods

Peer Reviewer:

Mobility Networks

(Nilforoshan et al., 2023, "Human mobility networks reveal increased segregation in large cities")

Peer Review

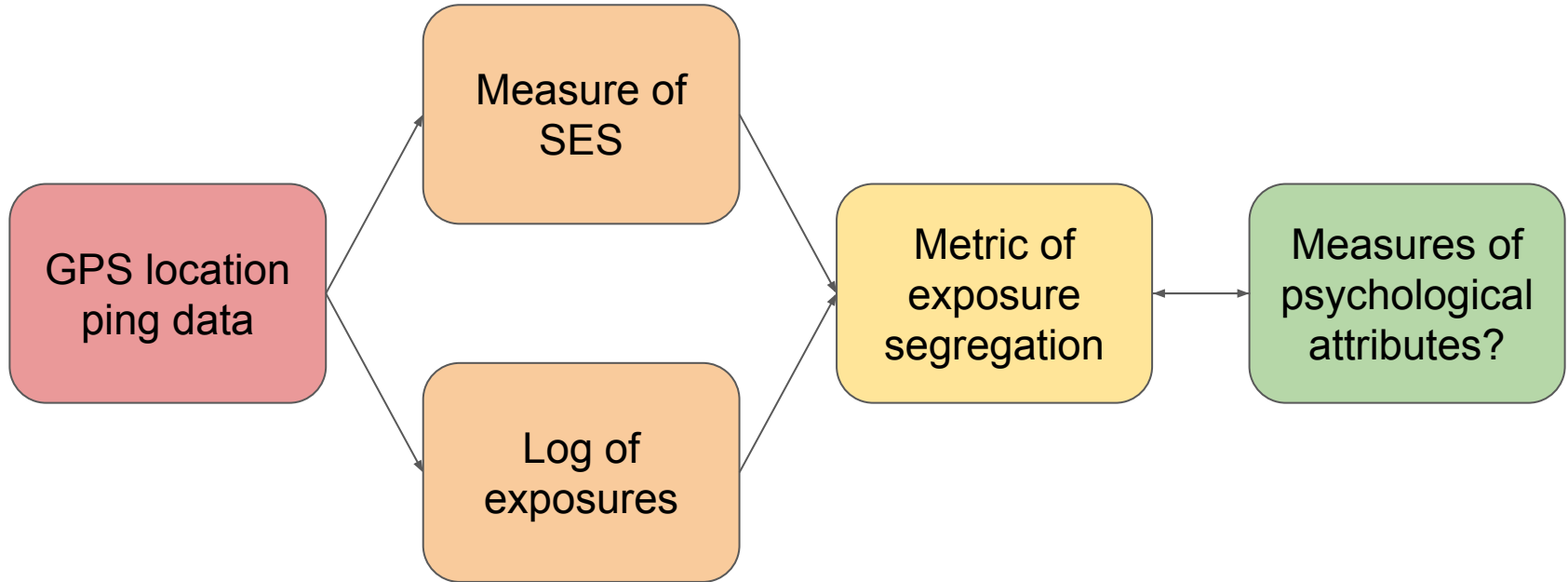
- **Problem:** Could it be that greater variation in SES in large cities (rent estimates) drive the result? This could in fact drive also more detailed results when these groups are more different in their consumption and leisure behavior
- **Solution:** Robustness check with rescaled measure.
- **Problem:** Raj Chetty's social capital atlas is using Facebook connections and draws a different conclusion: <https://www.socialcapital.org/>.
Solution: Compare and reason about the differences.
- **Problem:** Reverse Causality: Segregation can also be a cause of separate points of interest, not just a consequence.
- **Problem:** Hubs are not exogenous but determine the characteristics of the neighborhood (incl. rent prices)
Solution: The finding on bridging points would be more convincing when matching comparable cases where SES and neighbourhood composition are similar but the hub is placed differently.

Academic Researcher:

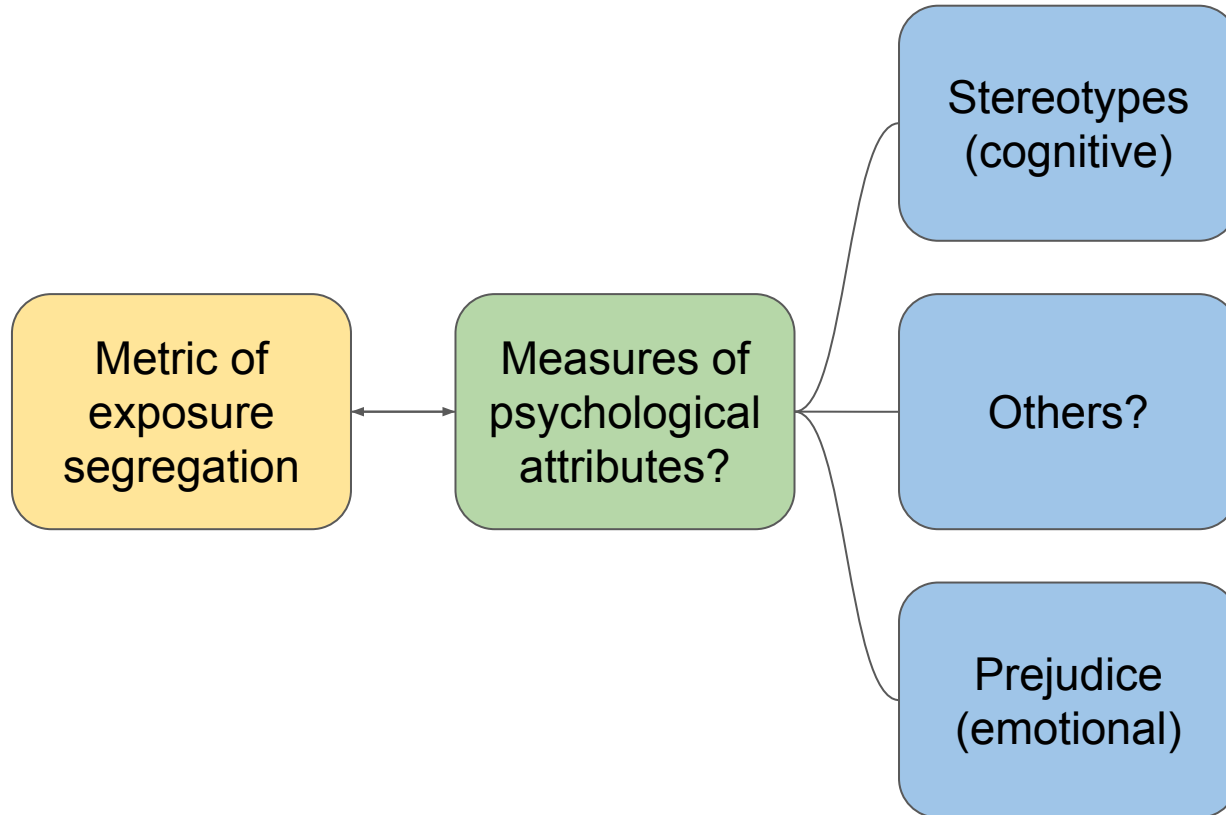
Mobility Networks

(Nilforoshan et al., 2023, "Human mobility networks reveal increased segregation in large cities")

Incorporating Psychological Measures



Incorporating Psychological Measures



Discuss with one or two people around you:

1. What is a psychological feature that would be interesting to relate to exposure segregation?
2. How might we measure/operationalize this psychological feature?
3. What relationship would you expect to see between this feature and exposure segregation? Why?

Limitations of Using Psychological Variables

- Much smaller sample size
 - Potential solution 1: smartphone-based survey pings
 - Potential solution 2: acquire psychological data through apps for which users have already consented to the use of their data for research purposes
 - This is how the GPS data were acquired
- Greater ethical/privacy concerns
 - Location + psychological attributes pose a concern for deanonymization

Social Impact Assessor: **Mobility Networks**

(Nilforoshan et al., 2023, "Human mobility networks reveal increased segregation in large cities")

Impact of economic segregation

In the United States, economic segregation is very high, with income affecting where one lives⁹, who one marries¹⁰, and who one meets and befriends¹¹. This extreme segregation is costly. It reduces economic mobility¹²⁻¹⁵, fosters a wide range of health problems¹⁶⁻¹⁸ and increases political polarization¹⁹⁻²². Although there are all manner

Self-assessment of positive impact

Improves detection of economic segregation, revealing a more severe problem than originally imagined.

are located in bridging zones that can draw in people from diverse neighbourhoods. We were able to detect these pockets of homophily (and the counteracting effects of bridging hubs) because we have developed a dynamic measure of economic segregation that captures everyday socioeconomic mixing at home, work and leisure.

The dynamic measure is "more realistic" and better reflects behavior and preference patterns.

Self-assessment of positive impact

Justifies urban design advice meant to reduce economic segregation.

Mitigating segregation through urban design

Our results suggest that segregation could be mitigated when frequently visited POIs, which we refer to as 'hubs', are positioned in close proximity to diverse neighbourhoods. These hubs would serve as bridges between residents of nearby high-SES and low-SES neighbourhoods, enabling them to easily visit the hubs⁴⁴⁻⁴⁶ and encounter one another (Fig. 3c). We developed the bridging index (see the 'Bridging index' section of the Methods) to measure whether hubs are located in such bridging positions. Our index measures the economic diversity of

Devises methods to implement this advice (like the bridging index, which would help in such planning).

Self-assessment of positive impact

This is all to suggest that dynamic segregation data are rich enough to overcome many seeming limitations. The dynamic approach that we have taken here could further be extended to examine cross-population differences in the sources of segregation and to develop a more complete toolkit of approaches for reducing segregation and improving urban design.

The good left out

The positive impacts of having more accurate perceptions of where economic segregation actually occurs:

- Better direction of efforts to mitigate segregation (for example: housing and cost-of-living subsidies encouraged in highly segregated areas);
- Further studies regarding these efforts or the contexts where they are needed can now collect better data, and make better use of it
- Opportunities for community-focused and data-driven activism by aid and advocacy groups, which can act more effectively with greater knowledge of their targets
- Greater awareness in community members of the degree to which they are polarized by status, and – as a result of good planning – potentially a sincere drive in these community members to reach out and make contact with those they are behaviorally and habitually divided from.

The unmentioned bad

Rent value of nighttime location may not be an accurate representation of individual social or economic status (though the paper acknowledges this):

time. Finally, our measure of SES relies on housing consumption, an indicator that does not exhaust the concept of SES. It is again reas-

Its own conclusions are robust to different measures, but the reliance of the exported SES measure on housing consumption may not be appropriate for all contexts where researchers may be inclined to use it.

The unmentioned bad

Visual exposure is not connection.

Creating an economic segregation measure based on visual exposure means that any solution based on this measure will ensure that people frequently **see** those of different economic backgrounds.

Sight does not guarantee change, empathy, or help.



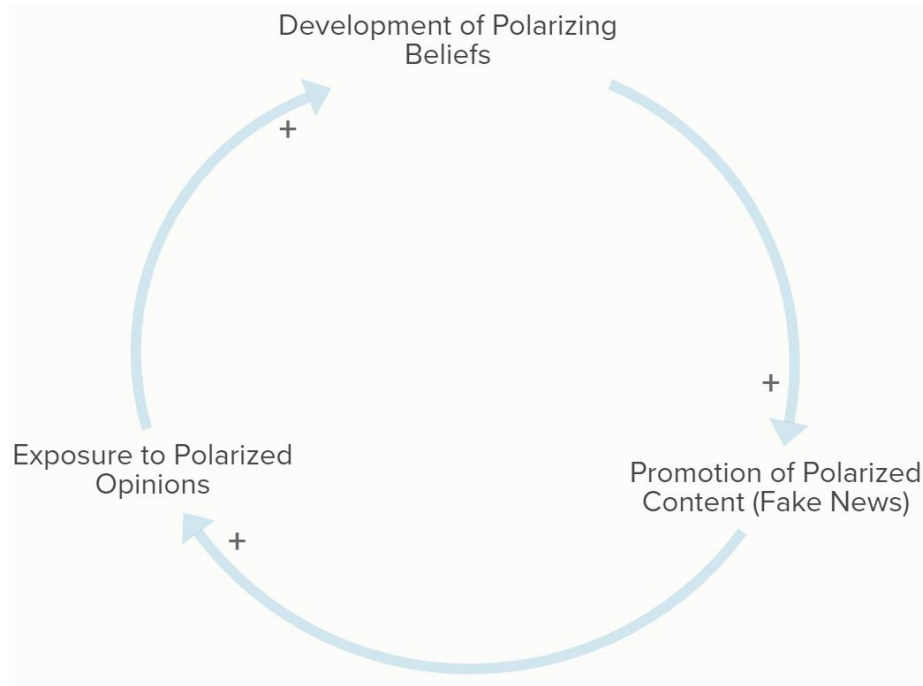
Instructor:

Echo Chambers

(Cinelli et al., 2021, "The echo chamber effect on social media")

What is an echo chamber?

The Vicious Cycle of Echo Chambers



This is intensified by:

- Selective Exposure (algorithms only recommend one-sided content)
- Confirmation Bias (people purposefully seek out evidence to support their beliefs)

What was Studied?

- Hypothesis: "Social media may limit exposure to diverse perspectives and favor the formation of groups of like-minded users framing and reinforcing a shared narrative, i.e., echo chambers."
- Novelty: Among the first comparative studies on social media, especially concerning news consumption
- Authors perform a comparative analysis on >100 million pieces of content concerning controversial topics (abortion, vaccines, politics)
- Data comes from Gab, Facebook, Reddit, and Twitter, four uniquely structured communities.



Methods

Two variables to measure echo chambers:

- Inference of the user's leaning for a specific topic as the average leaning $[-1, 1]$ of their posts
- Structure of the user's interactions on the platform, i.e. their social network's directed nodes

Two dimensions of analysis:

- Homophily in the interaction networks, i.e. the distribution of user leanings into clusters
- Bias in diffusion toward like-minded peers, i.e. the reach of network influencers



Methods (Cont.)

- Variables are applied to unique data on different topics from each site, measuring how users interacted with politically-charged content.
- Homophily is compared between discussions regarding abortion on Twitter, vaccines on Facebook, r/Politics on Reddit, and all of Gab.
 - (Each site's results were qualitatively similar across topics)
- Bias across networks is measured with an adaptation of the epidemiological SIR model.
- Finally, a direct comparison is carried out between Facebook and Reddit, looking only at news consumption.

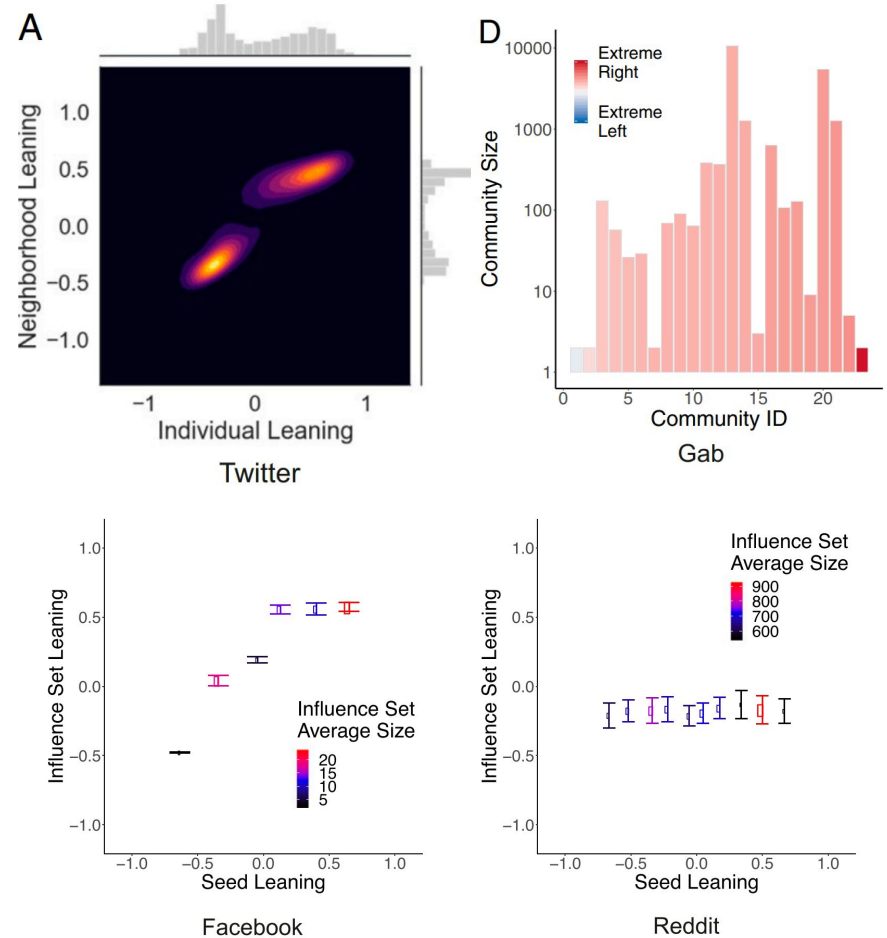
Table 1. Dataset details

Media	Dataset	T_0	T	C	N	n_c
Twitter	Gun control	June 2016	14 d	19 million	3,963	0.93
	Obamacare	June 2016	7 d	39 million	8,703	0.90
	Abortion	June 2016	7 d	34 million	7,401	0.95
Facebook	Sci/Cons	January 2010	5 y	75,172	183,378	1.00
	Vaccines	January 2010	7 y	94,776	221,758	1.00
	News	January 2010	6 y	15,540	38,663	1.00
Reddit	Politics	January 2017	1 y	353,864	240,455	0.15
	the_donald	January 2017	1 y	1.234 million	138,617	0.16
	News	January 2017	1 y	723,235	179,549	0.20
Gab	Gab	November 2017	1 y	13 million	165,162	0.13

For each dataset, we report the starting date of collection T_0 , time span T expressed in days (d) or years (y), number of unique contents C , number of users N , and coverage n_c (fraction of users with classified leaning). For Twitter, T represents the window to sample active users, of which we retrieve all of the tweets related to the topic via the Application Programming Interface (API) (more information in [SI Appendix](#)). Sci/Cons, Scientific and Conspiracy content.

Results

- Users' aggregation in homophilic clusters dominates online dynamics, especially on Facebook and Twitter.
 - Gab and Reddit formed single clusters biased to the right and left respectively.
 - This can be seen in spatial clustering and community leaning distributions.
- Whereas Facebook and Twitter users experienced biased diffusion, Reddit and Gab users received similarly biased info across clusters.
- A direct comparison of news consumption on Facebook and Reddit shows higher segregation on Facebook and an evenly-distributed left-wing bias on Reddit.



Peer Reviewer:

Echo Chambers

(Cinelli et al., 2021, "The echo chamber effect on social media")

Peer Review

- **Problem:** The measure of political leaning seems to pick up little signal on Reddit, but seems to be extremely sensitive to the Facebook data. Compositional effect?
Solution: Can we make this measure more comparable across platforms? Is there some way to verify that this measure does capture real political leaning?
- **Problem:** Measurement error in social ties can severely bias the results since only few links are necessary to bridge communities.
Solution: Use same measurement strategy for all four platforms (co-commenting). Yet, commenting behavior may be different. Thus, use network completion methods as a robustness check (for example to achieve similar density).
- **Problem:** News consumption measure assumes that every user is equally likely to produce content. This may be more likely for reddit than for facebook. When content is over proportionally produced by extreme actors that are connected to unique sets of users, news may be consumed as in an echo chamber while the network does not display it.
Solution: Measure who is over proportionally producing content.

Academic Researcher:

Echo Chambers

(Cinelli et al., 2021, "The echo chamber effect on social media")

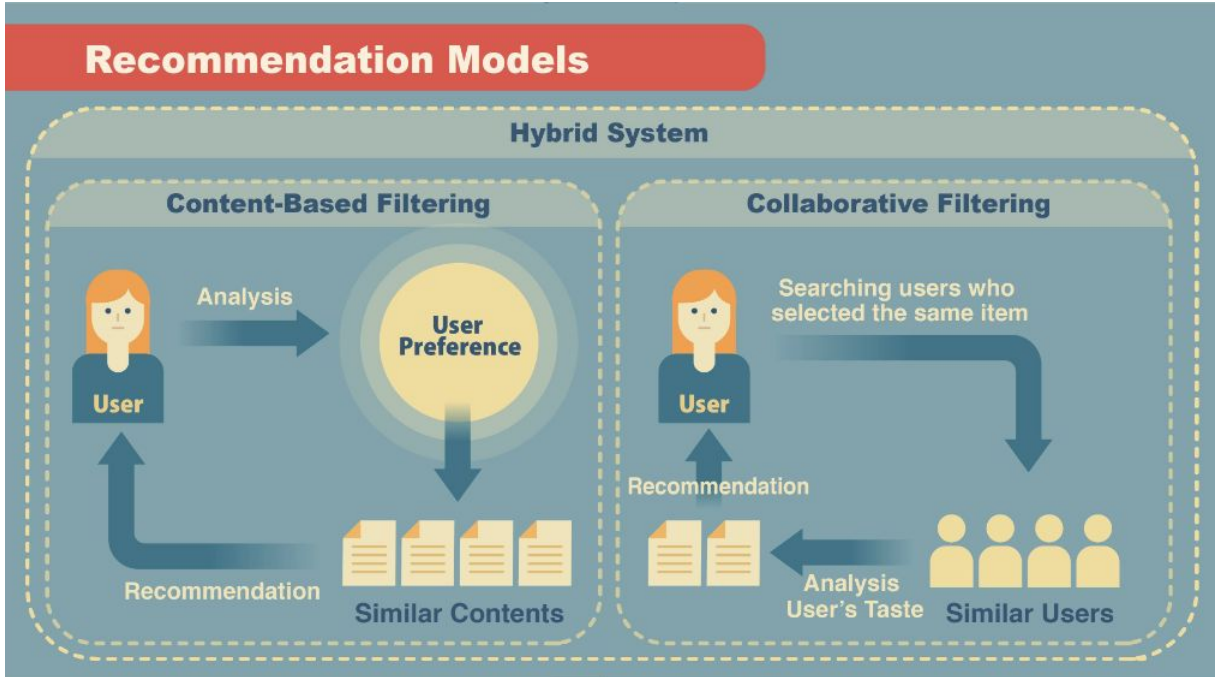
Development of Network Homophily Over Time

“The next envisioned step addresses the **temporal dimension** of echo chambers, to understand better how different social **feedback mechanisms**, specific to distinct platforms, can impact their **formation**.”

Development of Network Homophily Over Time

Theory: As a content recommendation system learns more about a user's preferences, that user will become more deeply involved with homophilic networks.

Recommendation Models



Development of Network Homophily Over Time

Theory: As a content recommendation system learns more about a user's preferences, that user will become more deeply involved with homophilic networks.

Method: Collect longitudinal data from new users joining a platform. Vary “strength” of recommendation system between users (if possible). Otherwise, compare across different platforms. Examine the association between degree of feed personalization and homophily (as defined in Cinelli et al. (2021)) over time.

Social Impact Assessor: **Echo Chambers**

(Cinelli et al., 2021, "The echo chamber effect on social media")

Impact of social media and echo chambers

Social media platforms provide direct access to an unprecedented amount of content. Platforms originally designed for user entertainment changed the way information spread. Indeed, feed algorithms mediate and influence the content promotion accounting for users' preferences and attitudes. Such a paradigm shift affected the construction of social perceptions and the framing of narratives; it may influence policy making, political communication, and the evolution of public debate, especially on polarizing topics. Indeed, users online tend to prefer information adhering to their worldviews, ignore dissenting information, and form polarized groups around shared narratives. Furthermore, when polarization is high, misinformation quickly proliferates.

Self-assessment of positive impact

Contributes to the debate around the effects and "very existence" of echo chambers,

29–31). However, recently, the effects and the very existence of echo chambers have been questioned (2, 27, 32). This issue is also fueled by the scarcity of comparative studies on social media, especially concerning news consumption (33). In this

especially because it is a comparative study:

or not. As recently shown in the case of selective exposure to news outlets, studies considering multiple platforms can offer a fresh view on long-debated problems (34). Different plat-

Comparative studies help us see the problem in many different contexts. In this case, we get to see how the social dynamics of different media platforms affect echo chamber formation.

Self-assessment of positive impact

Helps us understand "social media's influence on information consumption and public opinion formation," (1)

Our work provides important insights into the understanding of social dynamics and information consumption on social media. The next envisioned step addresses the temporal dimension of echo chambers, to understand better how different social feedback mechanisms, specific to distinct platforms, can impact their formation.

– sets up ways to refine and improve upon its conclusions in the future.

Self-assessment of positive impact

Operationalizes definition of echo chambers,

ics (35). We introduce an operational definition of echo chambers to provide a common methodological ground to explore how different platforms influence their formation. In particular,

permitting their identification in non-obvious settings and comparison of their properties, effects and working principles across platforms.

The good left out

Characterizing social networks as more or less conducive to echo chamber formation helps us understand how echo chambers form.

This makes it easier for us to **prevent** their formation in social media settings, and to **identify** and **disassemble** them in other, less familiar contexts.

In a social media setting:

- Adapt algorithms to serve a diversity of opinions to the user

in the information diffusion toward like-minded users. **A clear-cut distinction emerges between social media having a feed algorithm tweakable by the users** (e.g., Reddit) and social media that don't provide such an option (e.g., Facebook and Twitter).

The good left out

In a social media setting:

- Help the platform not "**burst**" echo chamber effects – if we can find out what this means

These results indicate that information diffusion is biased toward individuals who share a similar leaning in some social media, namely Twitter and Facebook. In contrast, in others—Reddit and Gab in our analysis—this effect is absent. Such a latter configuration may depend upon two factors: 1) Gab and Reddit are not bursting the echo chamber effects, or 2) we are observing the dynamic inside a single echo chamber.

The unmentioned bad

- **Greater division:** social media-related policymaking based on such research may drive people to find other, less visible media exclusive to their community, where it can be defended and perpetuated.

Truth Social

Truth Social is America's "Big Tent" social media platform that encourages an open, free, and honest global conversation without discriminating on the basis of political ideology.

- **Harmful legislation:** *This is not a problem with Cinelli et al. specifically.*

The popularization of rhetoric around disinformation and polarization can provide governments with justification to curtail freedom of speech and of organization.

Turkey's strict new press and social media 'disinformation' law

By Reuters

October 14, 2022 10:09 AM PDT · Updated 2 years ago



The unmentioned bad

Florida lawmakers pass social media crackdown, governor expected to sign it

Encouragement of social media-related fears can lead governments to curtail personal freedoms.

TALLAHASSEE — With one lawmaker likening social media to a “dark alley,” the Florida House on Wednesday gave final approval to a bill that seeks to keep children off social-media platforms.

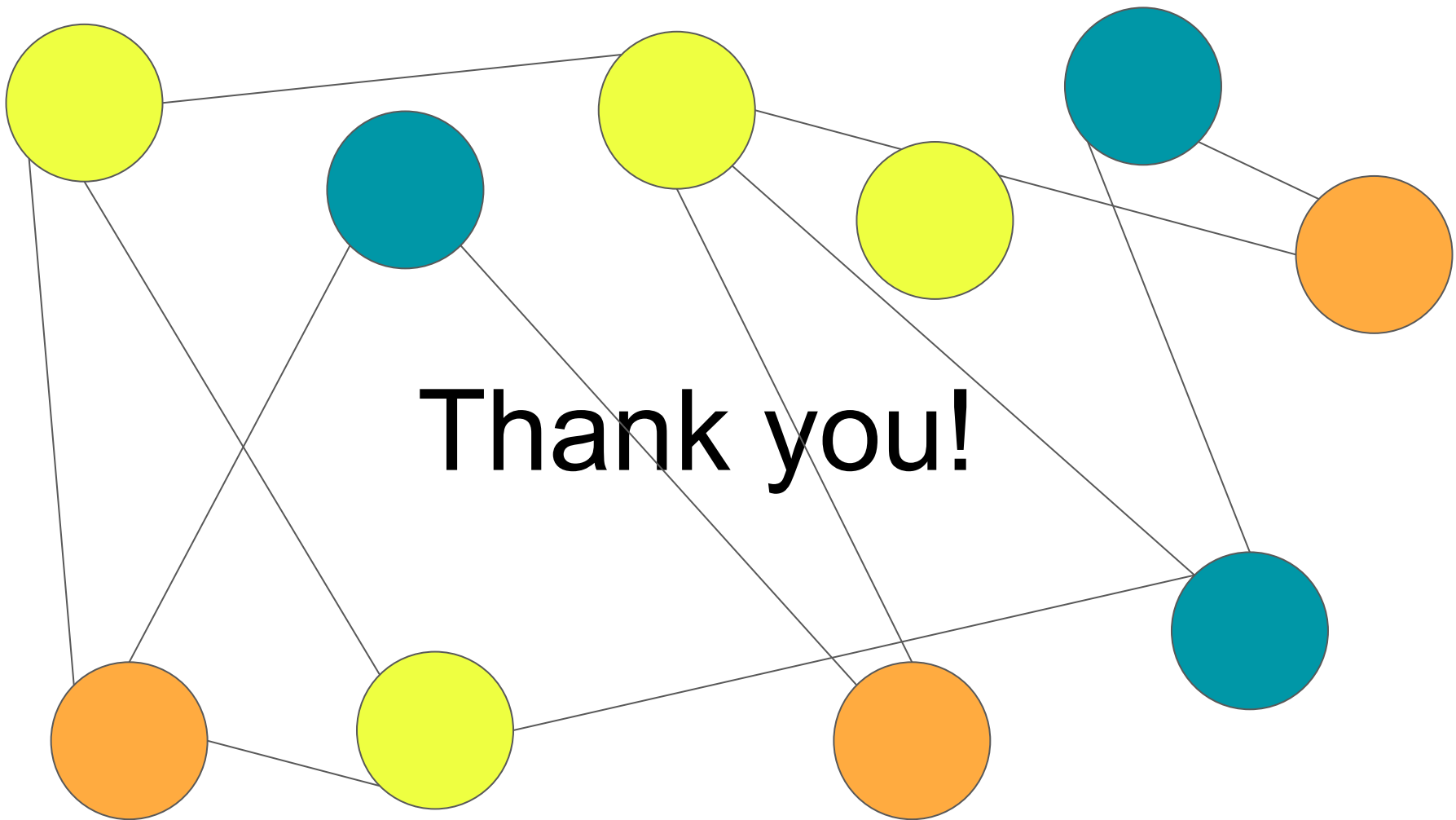
The House voted 109-4 to approve the bill (HB 3), which passed the Senate on Monday. It will go to Gov. Ron DeSantis, who is expected to sign it after vetoing an earlier version.

House Speaker Paul Renner, R-Palm Coast, made cracking down on social media perhaps the highest-profile issue of this year’s legislative session. He contends that social media use harms children’s mental health and can lead to sexual predators communicating with minors.

“This is something that I believe will save the current generation and generations to come if we’re successful,” Renner said after the bill passed.

Rep. Tyler Sirois, a Merritt Island Republican who helped sponsor the bill, said that if social media “is the new town square, then God help us.”

“For our children, social media is no town square,” Sirois said. “It is a dark alley.”



Thank you!