

Online Social Networks and Media

Fairness, Diversity

Fairness, Non-discrimination

To **discriminate** is to treat someone differently

(Unfair) discrimination is based on *group membership*,
not individual merit

Some attributes should be irrelevant, called **protected**,
or **sensitive**

What is the cause?

Data

- **Correctness and completeness** Garbage in, garbage out (GIGO)
 - Poorly selected
 - Incomplete
 - Incorrect
 - Outdated
 - Selected with bias
- **Data as a social mirror:** perpetuating and promoting historical biases
- **Sample size disparity**
 - learn on majority (Errors concentrated in the minority class)

What is the cause?

Processing

- Algorithms as **black boxes**
- Output models that are **hard to understand**
- Unrealistic **assumptions**
- Algorithms that **do not compensate** for input data problems
- Output **presentation** that is faulty (biased, unfair)
- Personalization and recommendation services that **narrow** instead of expand user options
- Decision making systems that assume **correlation** implies **causation**
- **BIAS REINFORCEMENT CYCLE**

Disparate treatment and impact

Disparate treatment: Treatment depends on class membership (protected attribute directly used in the decision)

Disparate impact: Outcome depends on class membership (Even if (apparently) people are treated the same way)

Doctrine solidified in the US after [Griggs v. Duke Power Co. 1971] where a high school diploma was required for unskilled work, excluding black applicants

Fairness through blindness

Ignore all irrelevant/protected attributes

Useful to avoid formal disparate treatment

Fairness: definition

Classification

- Classification/prediction for people with similar non-protected attributes should be similar
- Differences should be mostly explainable by non-protected attributes

Individual fairness

General principle: Similar people should be treated similarly

What does “similar” people mean?

Let V be a set of individuals

A task-specific distance metric $d: V \times V \rightarrow \mathbb{R}$

- Expresses *ground truth* (or, best available approximation)
- Public
- Open to discussion and refinement
 - Externally imposed, e.g., by a regulatory body, or externally proposed, e.g., by a civil rights organization

Group fairness

Three basic types of group fairness, based on

- Base rates
- Group-conditioned accuracy
- Calibration

Base rate (statistical parity)

Compare

$$P[\hat{Y} = \text{yes} | S = 1]$$

*Probability of favorable outcome
for privilege group*

with

$$P[\hat{Y} = \text{yes} | S \neq 1]$$

*Probability of favorable outcome
for minority group*

Catalog of evils

Self-fulfilling prophecy:

Deliberately *choosing the “wrong” members of S* in order *to build a bad “track record”* for S

A less malicious vendor simply selects *random members of S* rather than qualified members
(problem with parity)

Reverse tokenism:

Goal is to create convincing refutations

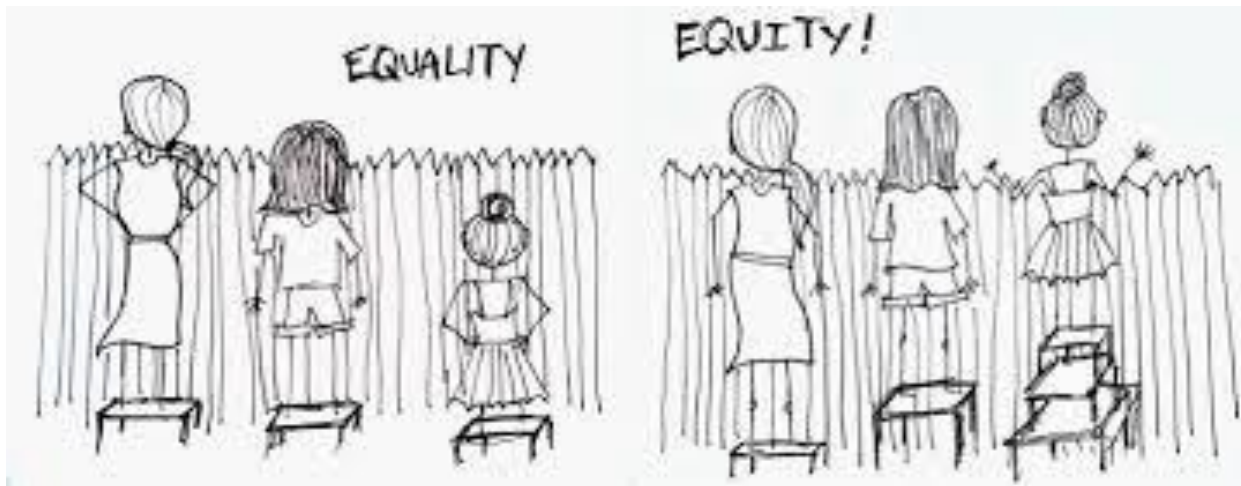
Deny access to a qualified member of S^c

c is a token rejectee

Discussion

Individual fairness

Statistical parity



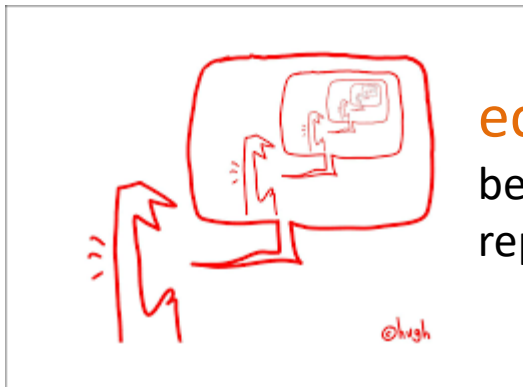
Diversity: filter bubbles

personalized searches and recommendations

filter bubble a state of intellectual isolation where users become separated from information that disagrees with their viewpoints,

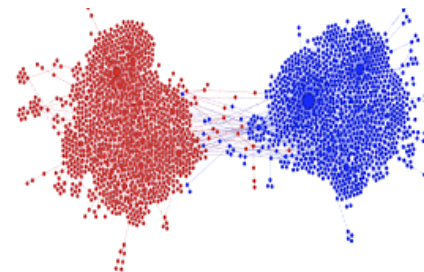


Social media has become the main source of news online with more than **2.4 billion** internet users, nearly **64.5%** receive breaking news **from social media instead of traditional media**



echo chambers: a situation in which information, ideas, or beliefs are amplified or reinforced by communication and repetition inside a defined system

polarity



Diversity

- *No useful information is missed*: results that cover all user intents
- *Better user experience*: less boring, more interesting, human desire for discovery, variety, change
- *Personal growth*: limited, incomplete knowledge, a self-reinforcing cycle of opinion

Better (Fair? Responsible?) decisions

Network Diversity

Improve awareness

Blue Feed, Red Feed site -- See Liberal Facebook and Conservative Facebook, Side by Side

<http://graphics.wsj.com/blue-feed-red-feed/>

Is your news feed a bubble? -- PolitEcho shows you the political biases of your Facebook friends and news feed.

<http://politecho.org/>

Link recommendation algorithms

Content recommendation algorithms (e.g., feed selection algorithms)

Filter Bubble – Eco Chambers: an experiment

Created two Facebook accounts

“Rusty Smith”, *right-wing avatar*, liked a variety of conservative news sources, organizations, and personalities, from the Wall Street Journal and The Hoover Institution to Breitbart News and Bill O’Reilly.

“Natasha Smith”, *left-wing avatar*, liked The New York Times, Mother Jones, Democracy Now and Think Progress.

Ten US voters – five conservative and five liberal – liberals were given log-ins to the conservative feed, and vice versa

<https://www.theguardian.com/us-news/2016/nov/16/facebook-bias-bubble-us-election-conservative-liberal-news-feed>

Coverage

Assuming different topics (e.g., concepts, categories, aspects, intents, interpretations, perspectives, opinions, etc)

Find items that **cover** all (most) of the topics

For example,

Rakesh Agrawal, Sreenivas Gollapudi, Alan Halverson, Samuel Jeong: *Diversifying search results*. WSDM 2009

Content Dissimilarity

Assuming (multi-dimensional, multi-attribute) items + a *distance measure (metric)* between the items
Find the most *different/distant/dissimilar* items

- Distance depends on the items and the problem
- Diversity ordering of the attributes

Defining distance/dissimilarity is key

For example, Sreenivas Gollapudi, Aneesh Sharma: *An axiomatic approach for result diversification*. WWW 2009

Novelty

Assuming the *history* of items seen in the past
Find the items that are the *most diverse (coverage, distance)* with respect to what a user (or, a community) *has seen in the past*

- *Marginal relevance*
- *Cascade (evaluation) models*: users are assumed to *scan result lists* from the top down, *eventually stopping* because either their information need is satisfied or their patience is exhausted

Novelty

Relevant concept: **serendipity**

represents the “unusualness” or “surprise”

(some notion of semantics – the guitar vs the animal)

For example, Charles L. A. Clarke, Maheedhar Kolla, Gordon V. Cormack, Olga Vechtomova, Azin Ashkan, Stefan Büttcher, Ian MacKinnon: *Novelty and diversity in information retrieval evaluation*. SIGIR 2008

Yuan Cao Zhang, Diarmuid Ó Séaghdha, Daniele Quercia, Tamas Jambor: *Auralist: introducing serendipity into music recommendation*. WSDM 2012

Homophily

“Ὅμοιος ὁμοίῳ αἰεὶ πελάζει” (Plato)

“Birds of a feather flock together”

Caused by two related social forces

- *Selection*: People seek out similar people to interact with
- *Social influence*: People become similar to those they interact with

Both processes contribute to homophily and lack of diversity, but

- Social influence leads to community-wide homogeneity
- Selection leads to fragmentation of the community

Opinion Formation

Complex process: many models

Commonly-used opinion-formation model (of Friedkin and Johnsen, 1990) (opinion – real number)

- Each individual i has an innate and an expressed opinion.
- At each step updates her expressed opinion
 - adheres to her innate opinion with a certain weight a_i and
 - is *socially influenced* by its neighbors with a weight $1-a_i$

Opinion Formation

An opinion formation process is **polarizing** if it results in increased divergence of opinions.

Empirical studies have shown that homophily results in polarization.

Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. *Exposure to Ideologically Diverse News and Opinion on Facebook*. Science 348:1130–1132, 2014

Stages in Facebook Exposure Process

1. *Friends network*: ideological homophily
2. *News feed*: more or less diverse content with algorithmically ranked News Feed
3. *Users' choices*: click through to ideologically discordant content.

Stages in Facebook Exposure Process

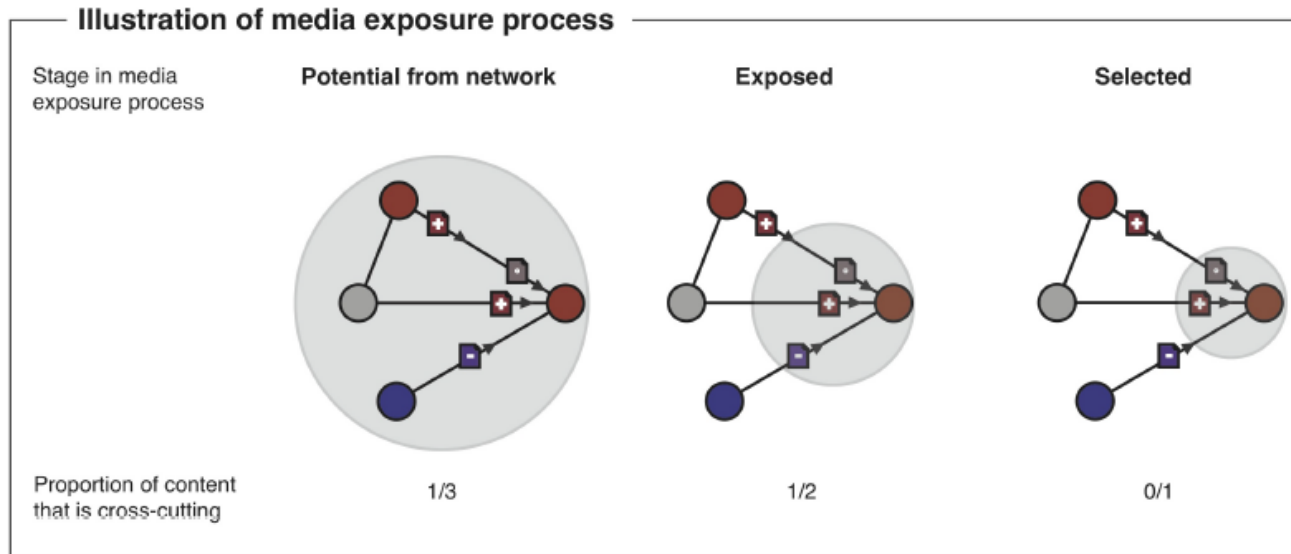


Illustration of how the exposure process consists of three phases: (1) the news your friends share (Potential from network), (2) ranking and the time that individuals take to scroll governs what they see in their News Feeds (Exposed), (3) clicking through to actual article (Selected).

- (1) what your **friends** share
- (2) what appears and in which position in the **News Feed**
- (3) what you choose to **click**

News Feed Ranking

“The *order* in which users see stories in the News Feed depends on *many factors*, including

- how often the viewer visits Facebook,
- how much they interact with certain friends, and
- how often users have clicked on links to certain websites in News Feed in the past.”

Dataset: users

10.1 million *active* U.S. users *who self-report* their ideological affiliation

All Facebook users can self-report their political affiliation, **9%** of U.S. over 18

Dataset: content

7 million distinct Web links (URLs) shared by U.S. users over a 6-month period between 7 July 2014 and 7 January 2015

Classified stories as

- **Hard content** (such as national news, politics, or world affairs) or
- **Soft content** (such as sports, entertainment, or travel)

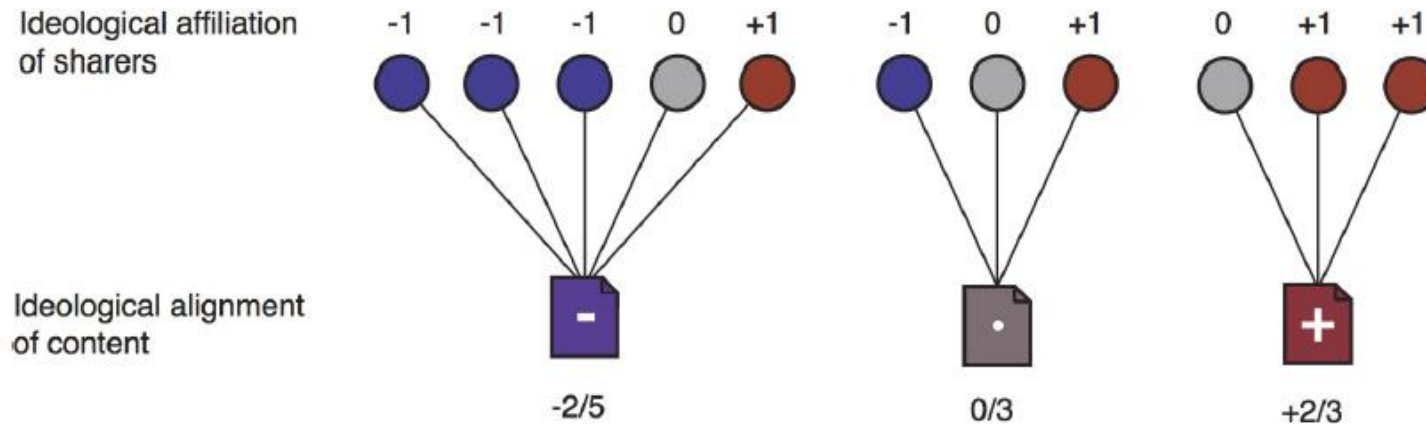
by training a *support vector machine* on unigram, bigram, and trigram text features

Approximately 13% hard content.

226,000 distinct hard-content URLs shared by at least 20 users who volunteered their ideological affiliation in their profile

Labeling stories (content alignment)

measure *content alignment (A)* for each hard story:
average of the ideological affiliation of each user who shared the article.



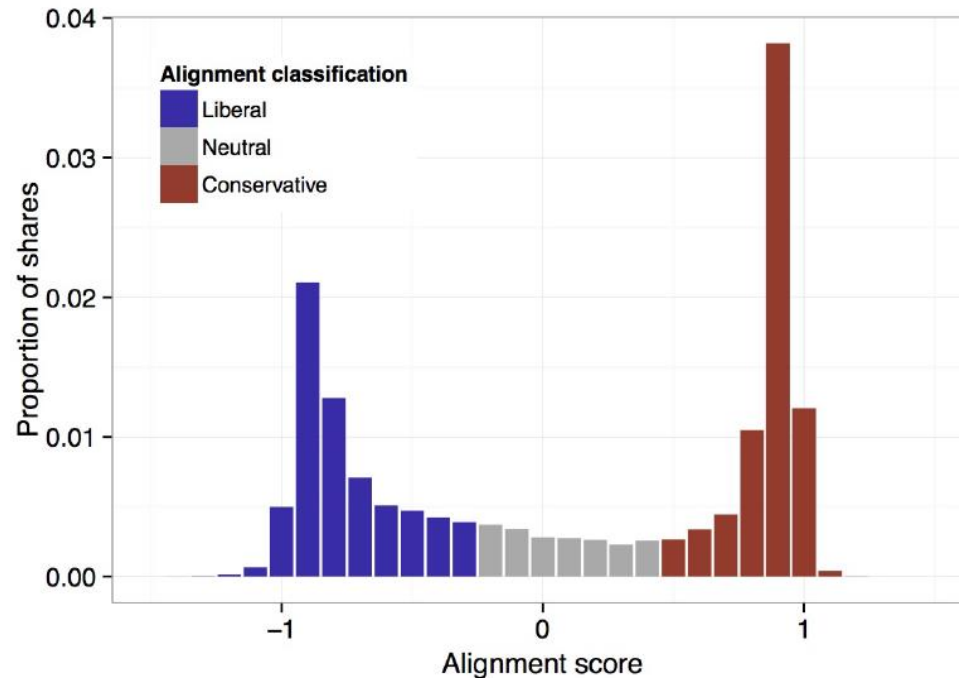
- measure of the *ideological alignment of the audience* who shares an article, *not a measure of political bias or slant of the article*

Labeling stories (content alignment)

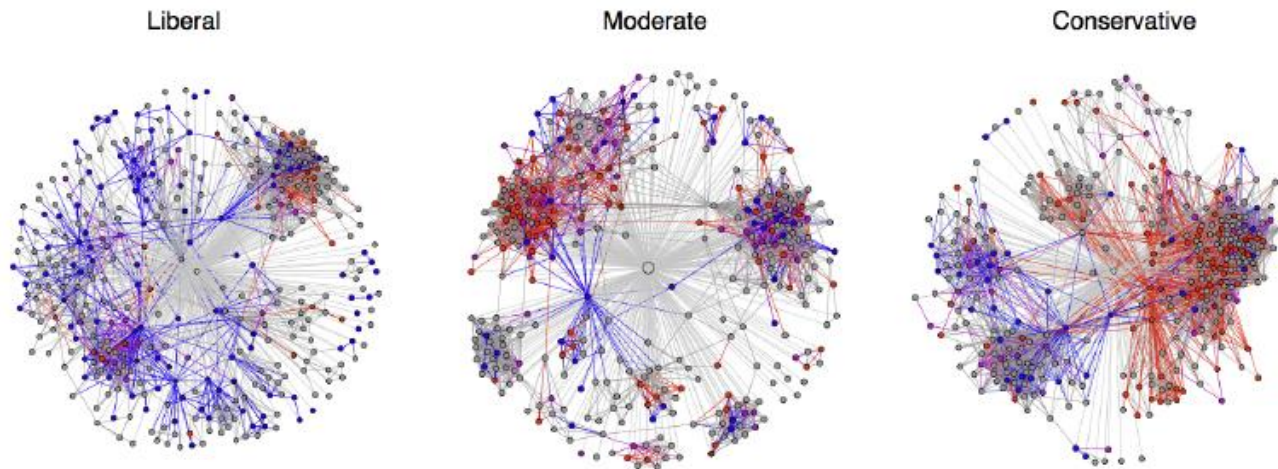
FoxNews.com is aligned with conservatives (As = +.80)

HuffingtonPost.com is aligned with liberals (As = -.65)

Substantial polarization

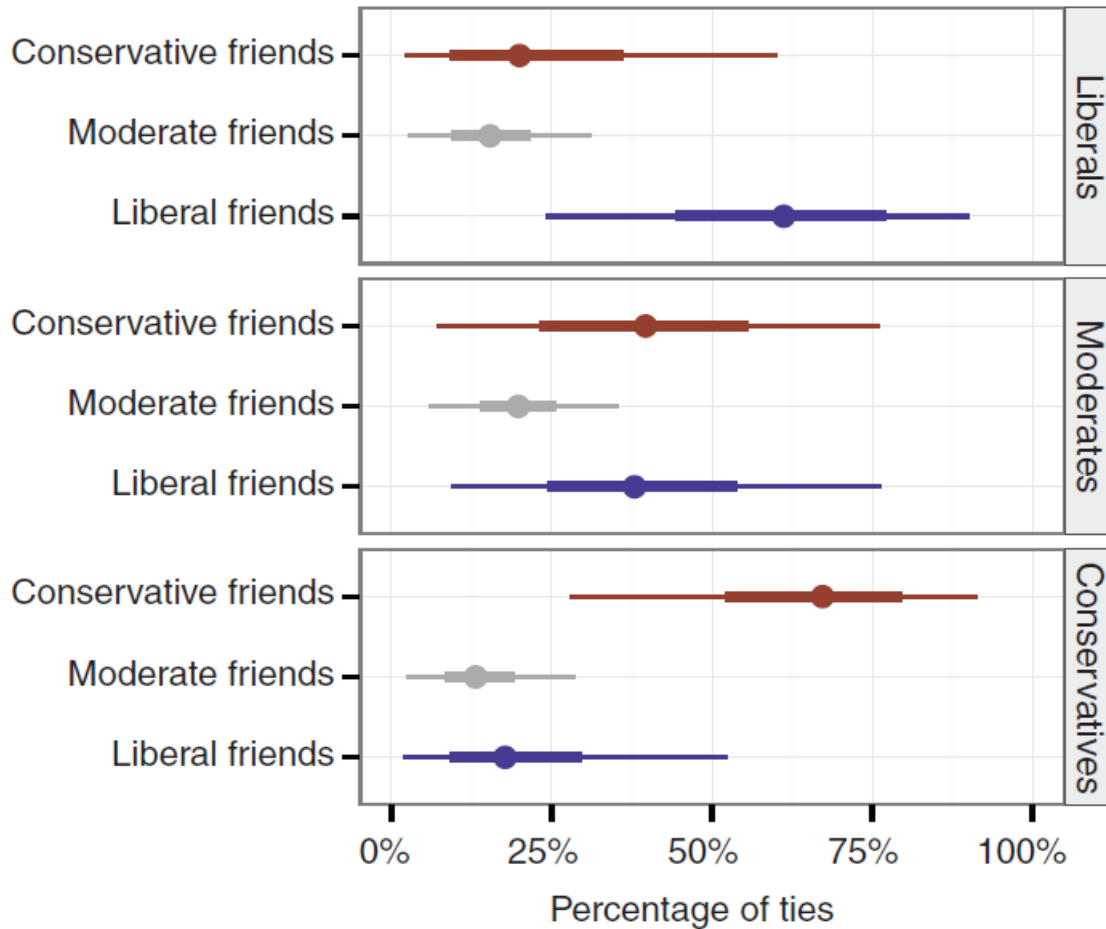


Homophily in the Friends Network



Example social networks for a liberal, a moderate, and a conservative. Points are individuals' friends, and lines designate friendships between them.

Homophily in the Friends Network



Median proportion of friendships

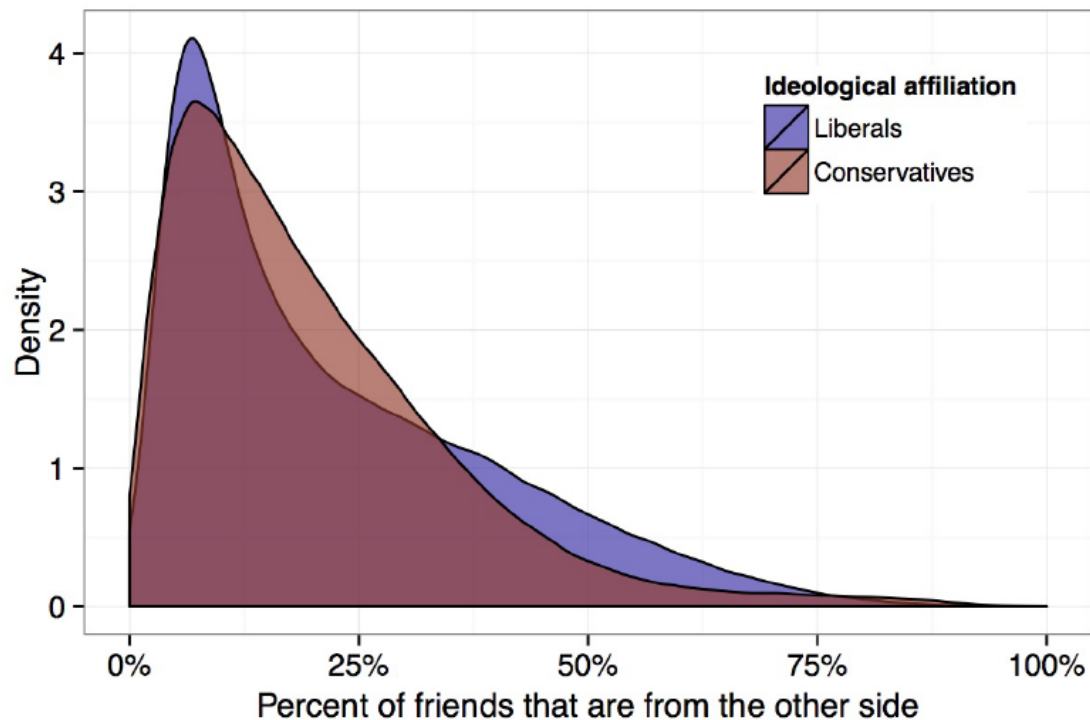
- of liberals with conservatives **0.20**,
- of conservatives maintain with liberals **0.18**

Homophily in the Friends Network

On average, about 23% of their friends report an affiliation on the opposite side

A wide range of network diversity

- 50% between 9 and 33 percent,
- 25% less than 9 percent
- 25% more than 33 percent



Content shared by friends

If from **random** others,
~45% cross-cutting for liberals
~40% for conservatives

If from **friends**,
~24% crosscutting for liberals
~35% crosscutting for conservatives

News Feed

After ranking, there is on average *slightly less crosscutting*

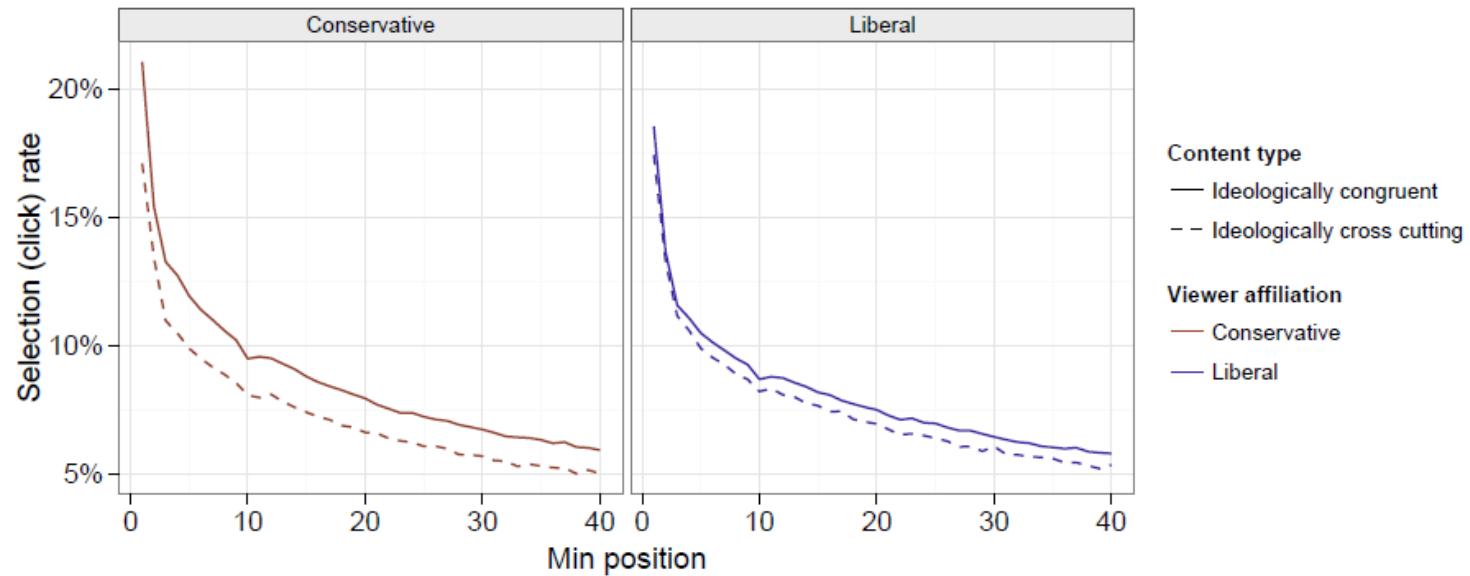
risk ratio of x percent:

people were x percent less likely to **see crosscutting** articles that have been shared by friends, compared to the likelihood of **seeing ideologically consistent** articles that have been shared by friends.

risk ratio

- 5% for conservatives
- 8% for liberals

Clicked



the click rate on a link is negatively correlated with its position in the News Feed

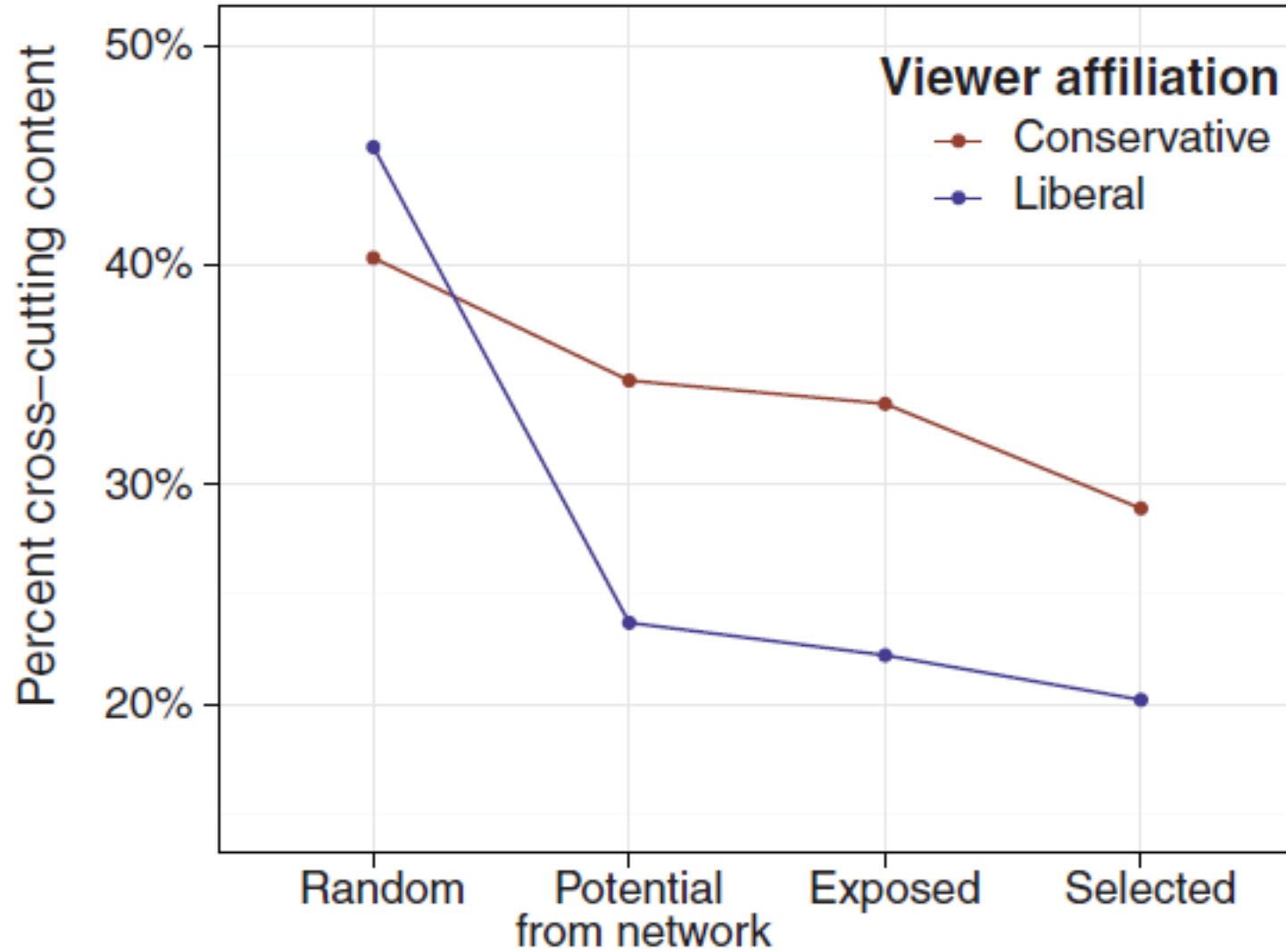
Clicked

Risk ratio

17% for conservatives

6% for liberals,

On average, viewers clicked on 7% of hard content available in their feeds



Questions?