

Discussion of Athey et al. (2014) and Breznitz & Palermo (2014)

Pinar Yildirim
Assistant Professor of Marketing

June 6, 2014
5th Annual Conference on Internet Search and Innovation
Searle Center on Law, Regulation, and Economic Growth



PART I:
Social Media and News Consumption

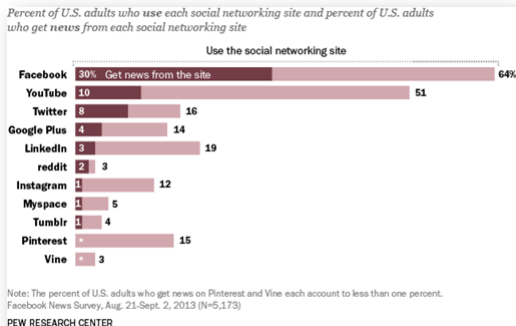
by Athey, Mobius, and Pal (2014)

Motivation

(**Disclaimer:** I only received the slides.)

Motivation

(Disclaimer: I only received the slides.)



- An increasing number of consumers receive news from social media - interesting demand and supply side questions
- How do social media affect the composition of news consumption?
 - Focus on **topics** & on **polarization**

The Paper

- Strategy
 - **Impressive** data collection: toolbar data, Wikipedia data, article data, others
 - Text mine Wikipedia, text mine news articles, determine important Wikipedia topics, assign articles based on similarity
 - Use **Mechanical Turk** & crowdsource to generate topic and article classifications, then compare articles of direct navigation and social media

- Strategy
 - **Impressive** data collection: toolbar data, Wikipedia data, article data, others
 - Text mine Wikipedia, text mine news articles, determine important Wikipedia topics, assign articles based on similarity
 - Use **Mechanical Turk** & crowdsource to generate topic and article classifications, then compare articles of direct navigation and social media
- Results
 - Some differences in article characteristics (length, opinion pieces, extreme views, more conservative, etc.) and super topics (less sports, more individual people, more crime)
 - Overall, DN and SM news consumption are **similar** (contrast to Gentzkow and Shapiro 2011 - ideological segregation of online news consumption is low but higher than the segregation of most offline news consumption)

Comments

Social & Traditional News Media

- **Aim:** Capture the characteristics of news consumed on / via social media

Comments

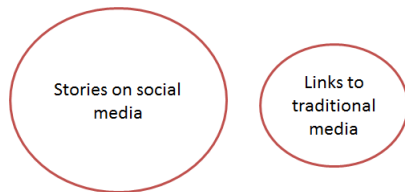
Social & Traditional News Media

- **Aim:** Capture the characteristics of news consumed on / via social media
- **Strategy:** collect news urls, investigate SM referrals to news sites to infer the topics consumed / shared on social media

Comments

Social & Traditional News Media

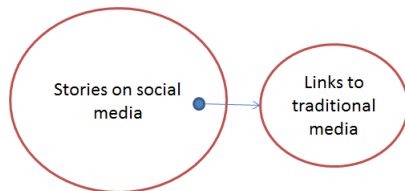
- **Aim:** Capture the characteristics of news consumed on / via social media
- **Strategy:** collect news urls, investigate SM referrals to news sites to infer the topics consumed / shared on social media
- How much do navigation data from social media capture what is being talked about on social media?



Comments

Social & Traditional News Media

- **Aim:** Capture the characteristics of news consumed on / via social media
- **Strategy:** collect news urls, investigate SM referrals to news sites to infer the topics consumed / shared on social media
- How much do navigation data from social media capture what is being talked about on social media?



Comments

Links on Social Media

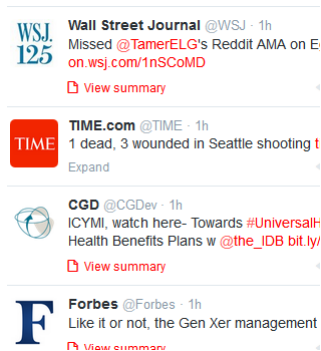
- How do news links appear on social media in the first place?

- How do news links appear on social media in the first place?
 - Take Twitter- A portion of the tweets contain links organically, but a high number of links are posted by the news sources
 - **Promoted** - highly *salient* and endogenously selected to go viral
 - What is the process through which these stories are promoted?

Comments

Links on Social Media

- How do news links appear on social media in the first place?
 - Take Twitter- A portion of the tweets contain links organically, but a high number of links are posted by the news sources
 - **Promoted** - highly *salient* and endogenously selected to go viral
 - What is the process through which these stories are promoted?



The screenshot displays a vertical list of four tweets from news organizations. Each tweet includes a profile picture, the account name and handle, the time since posted, and the text of the tweet. A 'View summary' link is visible below each tweet.

- WSJ. 125** Wall Street Journal @WSJ · 1h
Missed @TamerELG's Reddit AMA on E...
on.wsj.com/1nSCoMD
[View summary](#)
- TIME** TIME.com @TIME · 1h
1 dead, 3 wounded in Seattle shooting t...
Expand
- CGD** @CGDev · 1h
ICYMI, watch here- Towards #UniversalH...
Health Benefits Plans w @the_IDB bit.ly/
[View summary](#)
- F** Forbes @Forbes · 1h
Like it or not, the Gen Xer management
[View summary](#)

Broader Questions

Social & Traditional News Media

- What is the role of social media in news consumption?
 - A **substitute** for traditional news media?
 - **Complement** traditional media? Spread stories that are not reported or less represented?

Broader Questions

Social & Traditional News Media

- What is the role of social media in news consumption?
 - A **substitute** for traditional news media?
 - **Complement** traditional media? Spread stories that are not reported or less represented?
- **Informativeness of missing links** - Let's try the same data exercise in Turkey, Russia, China...
 - Many news stories spread on Twitter with almost no links or referral to traditional media outlets
 - Finding differences is important - but this cannot be done by analyzing data from navigated links (they won't exist!)

Broader Questions

Who is making the news?

- Changing news media landscape: who supplies the news? Consumers? News media?
 - News consumers spread information, bloggers raise issues, others post pictures and videos (Enikolopov, Petrova and Sonin 2013)
 - News media may respond to the issues that matter to consumers, or directly report / or try to deny news produced by citizens
- Is a comparison of content from social media and traditional media assuming an **independent** consumption and production decision using publicly available data meaningful?

Minor Comments

Crowdsourcing & Measurement

- A novel aspect of the paper is using crowdsourcing for classification of large scale data

Minor Comments

Crowdsourcing & Measurement

- A novel aspect of the paper is using crowdsourcing for classification of large scale data
 - Survey issues - responses are not exclusive & subjective, etc.

Is the topic of the articles of general interest or is it interesting only to a special group of people?

- The topic is of general interest
- The topic is interesting only to a special group of people

Do you think the articles about this topic can be politically biased?

- Yes
- No

Do you think the articles about this topic might make people feel strong emotions?

- Yes
- No

What do you think is the broad topic of the articles?

- Crime, terrorism, war
- Government, politics, political scandals
- Disasters, accidents
- Stories about individual people and their experiences
- Celebrities / Entertainment (including stories about crimes involving celebrities)
- Business / economy / real estate
- Science, technology, health
- Sports
- Other

- Can bias be measured objectively via crowdsourcing?
- Have moved away from judgment based methods in political bias measurement
 - Gentzkow & Shapiro (2010): text mine congressional speeches, identify Republican and Democratic keywords, search for them in news articles
 - Using Wikipedia & text mining to match articles based on commonalities - what about bias? Recency effect? (Greenstein and Zhu 2012)

PART II:

**Life Is But an Online Shopping Journey?
Exploring the Dynamic Interactions Between Targeted and
Paid Search Advertisement Mix**

by Breznitz & Palermo

- Motivation
 - Understanding the mechanisms that affect advertisers' performance after they invested in both behavioral targeting (BT) and paid search (PS) advertising
 - Investigate the interactions and **spillovers** between different modes of advertising
- Strategy
 - Use data from a network on firms that advertise in both BT and PS modes, tracked for two years
 - Lagged performance measures for instruments
- Stated Results
 - BT has higher return than PS
 - Users that click on an ad and land on the advertised webpage are less likely to conclude a transaction in the same week

The Question

- PS and BT are leading consumer acquisition tools

The Question

- PS and BT are leading consumer acquisition tools
- Key challenges faced by empirical studies of paid search advertising are in part due to the nature of widely available data
 - Google and other major search engines provide data **aggregated** at keyword level to their advertisers
 - Online shoppers who use a keyword are **heterogeneous** in their preferences and responses to search tools
 - There is an unobserved auction process for PS, which generates the **endogenous position** of the ad

The Question

- PS and BT are leading consumer acquisition tools
- Key challenges faced by empirical studies of paid search advertising are in part due to the nature of widely available data
 - Google and other major search engines provide data **aggregated** at keyword level to their advertisers
 - Online shoppers who use a keyword are **heterogeneous** in their preferences and responses to search tools
 - There is an unobserved auction process for PS, which generates the **endogenous position** of the ad
- Several papers tried to uncover heterogeneous preference to resolve position effect from aggregate data (Song and Chintagunta 2006, Chen and Yang 2007, Rutz and Trusov 2010) and study spillovers from generic and branded keywords perspective (Rutz and Bucklin 2011)

- Authors have a large and more refined dataset...

- Authors have a large and more refined dataset...
 - Consumer preferences are unknown, or are assumed to be **homogeneous** in the model

- Authors have a large and more refined dataset...
 - Consumer preferences are unknown, or are assumed to be **homogeneous** in the model
 - Heterogeneity among advertisers is unknown
 - Data are obtained from 15 different industries and 75 different advertisers, but data show advertising gaps & differences

- Authors have a large and more refined dataset...
 - Consumer preferences are unknown, or are assumed to be **homogeneous** in the model
 - Heterogeneity among advertisers is unknown
 - Data are obtained from 15 different industries and 75 different advertisers, but data show advertising gaps & differences
 - Heterogeneity in ads is narrowed down to a classification of BT and PS
 - Missing **textual properties** and ad position information of PS
 - **Generic** vs. branded keywords in PS
 - Product properties for ads are unknown

- Authors follow advertiser's classification of BT and PS, but details of the ad allocation process is unclear
 - How is PS differentiated from BT?
 - Is the BT really “behavioral targeting” or is it just “targeting”? Is there re- targeting?
 - Do they always advertise on both BT & PS, and if so why? Do they switch the amount of ads? Do they respond to changing ad prices? Competitor response?

- Authors provide separate estimations for CTR and CR - treating them as independent decisions
 - Model whether a consumer clicks or not and, conditional on click-through, use an independent model of conversion
 - There is **no correlation** in consumer preferences across the two decisions and no correction for selection bias
 - Possible correlation between consumers' preferences across click and conversion needs to be taken in consideration
- These correlations can be useful in deriving counterfactuals

- For PS, there is an endogeneity concern with respect to position
 - Position is determined based on the expected performance of the ad
(simultaneity bias)

- For PS, there is an endogeneity concern with respect to position
 - Position is determined based on the expected performance of the ad (**simultaneity bias**)
 - Outcome of an auction and influenced by competition, which is unobserve (**omitted variables bias**)
 - Authors model competition by a count of other firms advertising via the same network in a given week - level of advertising and advertising on other networks are unknown

- For PS, there is an endogeneity concern with respect to position
 - Position is determined based on the expected performance of the ad (**simultaneity bias**)
 - Outcome of an auction and influenced by competition, which is unobserve (**omitted variables bias**)
 - Authors model competition by a count of other firms advertising via the same network in a given week - level of advertising and advertising on other networks are unknown
 - How is position / CPC reported? If position is only reported on average then this may lead to further bias (**errors-in-variables**).

Thank you