

Lecture 19: Social media and social ads

Matthew J. Salganik

Sociology 204: Social Networks
Princeton University



Social media:

- ▶ Lecture 17: Social media and individuals
- ▶ [Lecture 18: Social media and society](#)
- ▶ Lecture 19: Social ads in social media
- ▶ Lecture 20: Fixing social media
- ▶ Lecture 21: Facebook Files

Social media:

- ▶ Lecture 17: Social media and individuals
- ▶ Lecture 18: Social media and society
- ▶ [Lecture 19: Social ads in social media](#)
- ▶ Lecture 20: Fixing social media
- ▶ Lecture 21: Facebook Files

Community minute

1. Duhigg, C. (2012). How companies learn your secrets. *New York Times*.
2. Goel, S. and Goldstein D.G. (2014). Predicting Individual Behavior with Social Networks. *Marketing Science*.
3. Bakshy, E. et al. (2012). Social influence in social advertising: Evidence from field experiments. *EC*



<https://www.youtube.com/watch?v=n2H8wx1aBiQ>

Social media companies want to “help connect everyone around the world and to bring the world closer together”

Social media companies want to “help connect everyone around the world and to bring the world closer together” and sell ads

Social media companies want to “help connect everyone around the world and to bring the world closer together” and sell ads that make lots and lots of money

Facebook posts a 33 percent increase in revenue and a 53 percent jump in profit.



By **Mike Isaac**

Jan. 27, 2021

Facebook on Wednesday reported surging profits and revenue driven by soaring ad sales, but cautioned that it might face “headwinds” in the future from regulation and technology changes.

The social network’s revenue in the fourth quarter grew to \$28 billion, up 33 percent from a year earlier and beating Wall Street estimates. Profits totaled \$11.2 billion, up 53 percent.

Copyrighted Material

THE AGE OF SURVEILLANCE CAPITALISM

THE FIGHT FOR A
HUMAN FUTURE
AT THE NEW
FRONTIER OF POWER

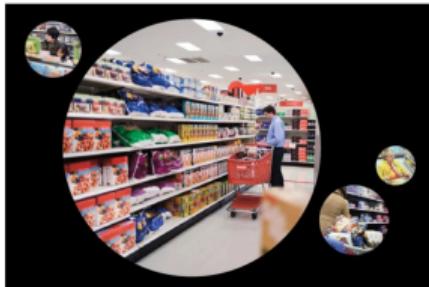
SHOSHANA
ZUBOFF

Copyrighted Material

Commodification of personal data for profit:

https://en.wikipedia.org/wiki/Surveillance_capitalism

How Companies Learn Your Secrets



Antonio Bolfo/Reportage for The New York Times

By Charles Duhigg

Feb. 16, 2012

- ▶ Pregnancy story illustrates surveillance capitalism, but Duhigg doesn't talk about social networks at all

How Companies Learn Your Secrets



Antonio Bolfo/Reportage for The New York Times

By Charles Duhigg

Feb. 16, 2012

- ▶ Pregnancy story illustrates surveillance capitalism, but Duhigg doesn't talk about social networks at all
- ▶ Discussion of habits may be quite related to what you are experiencing now during your self-experiment

Bakshy et al. “We regard social advertising as any advertising methods that uses information about consumers’ social networks to target ads and/or provide personalized social signals.”

- ▶ Goel and Goldstein don't care about causality so they use observational data
- ▶ Bakshy et al. care about causality a lot so they run online field experiments
- ▶ Both paper use the data from millions of people

Predicting Individual Behavior with Social Networks

Sharad Goel, Daniel G. Goldstein

Microsoft Research, New York, New York 10011 {sharadg@microsoft.com, dgg@microsoft.com}

- ▶ Focus on managerially relevant question of assessing the worth of social network data for targeting and prediction

Predicting Individual Behavior with Social Networks

Sharad Goel, Daniel G. Goldstein

Microsoft Research, New York, New York 10011 {sharadg@microsoft.com, dgg@microsoft.com}

- ▶ Focus on managerially relevant question of assessing the worth of social network data for targeting and prediction
 - ▶ Does social network data help even in the presence of other data that companies already have (e.g., demographics and previous behavior)?
 - ▶ How many targets can network data help find?

Predicting Individual Behavior with Social Networks

Sharad Goel, Daniel G. Goldstein

Microsoft Research, New York, New York 10011 {sharadg@microsoft.com, dgg@microsoft.com}

- ▶ Focus on managerially relevant question of assessing the worth of social network data for targeting and prediction
 - ▶ Does social network data help even in the presence of other data that companies already have (e.g., demographics and previous behavior)?
 - ▶ How many targets can network data help find?
- ▶ Given their focus, they don't care about causality

- ▶ Social data comes from Yahoo! communications network.
- ▶ Edge between people who mutually exchanged email or instant messages during a fixed two-month period.
- ▶ 132 million people and 719 million edges, with a mean of 11 contacts per individual.

 [Sign in](#)[Mail](#)[Mail](#)[COVID-19](#)[News](#)[Finance](#)[Sports](#)[Sportsbook LIVE](#)[Entertainment](#)[Life](#)[Shopping](#)[Yahoo Plus](#)[More...](#)

New Jersey Homeowners Are Getting Huge Payday

If you owe less than \$625,000 on your home and have no missed payments in 3 months, you better read this.

[Learn More](#)[Close Ad](#) |

Ad • LendGo ::



Is this the best pump fake in NCAA tournament history?

Houston and Baylor tipped off the Final Four on Saturday and delivered a baffling highlight right after the game began.

[Watch the play unfold »](#)

Trending Now

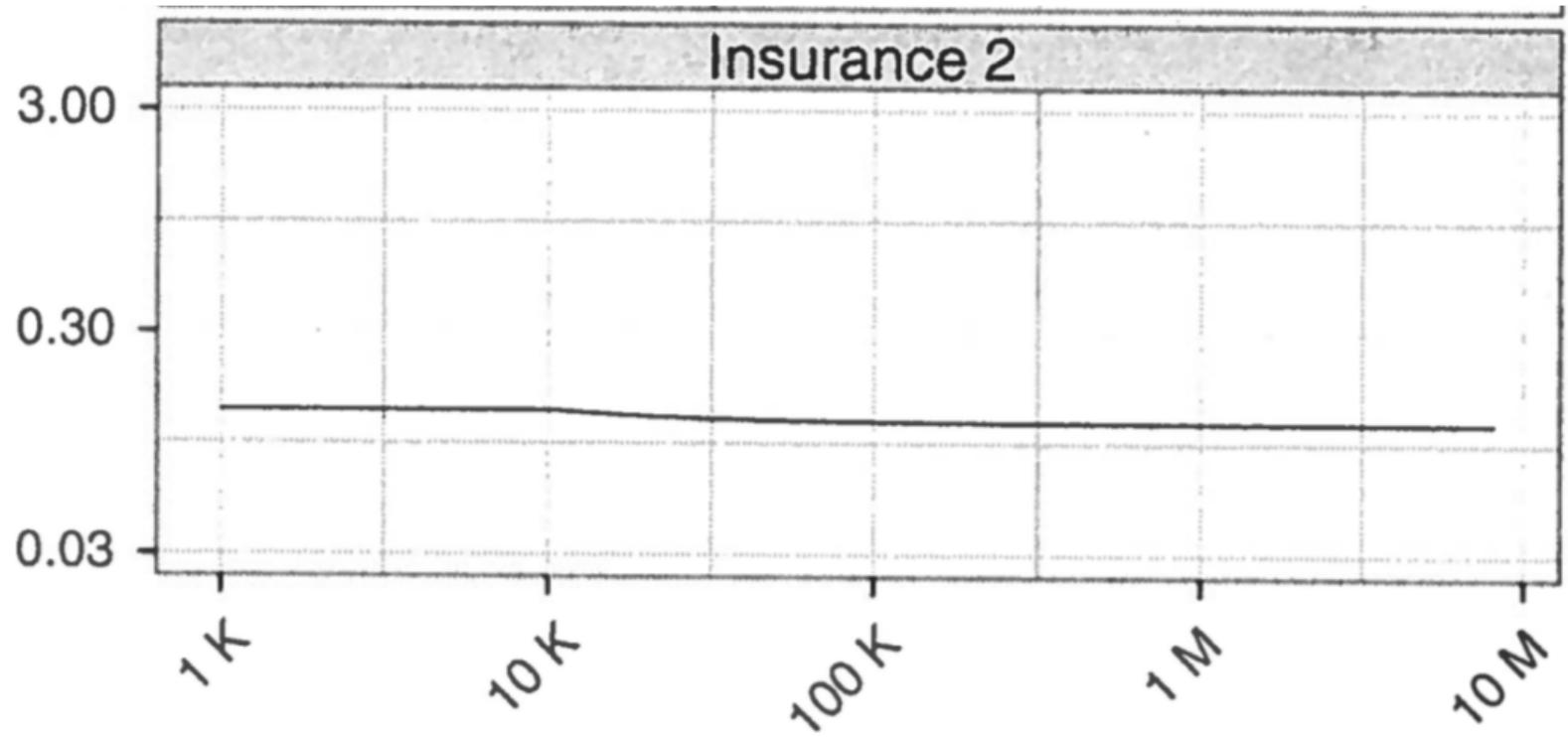
1. Mossimo Giannulli
2. Dmx Overdose
3. David Schwimmer
4. Milwaukee Bucks
5. Simple Gantt Chart
6. Noah Green
7. Volkswagen Passat Off...
8. Window Shades Online
9. Chelsea FC
10. Matt Gaetz

Domain	Click rates for individuals without contacts who clicked (%)	Click rates for individuals with contacts who clicked (%)	
Movie 1	0.038	0.47	
Government	0.209	0.46	
Movie 2	0.225	0.44	
TV	0.260	0.50	
Transportation	0.155	0.25	
Insurance 1	0.124	0.19	
Apparel	1.723	2.43	
Household	0.205	0.27	
Insurance 2	0.118	0.13	
Movie 3	1.185	1.30	

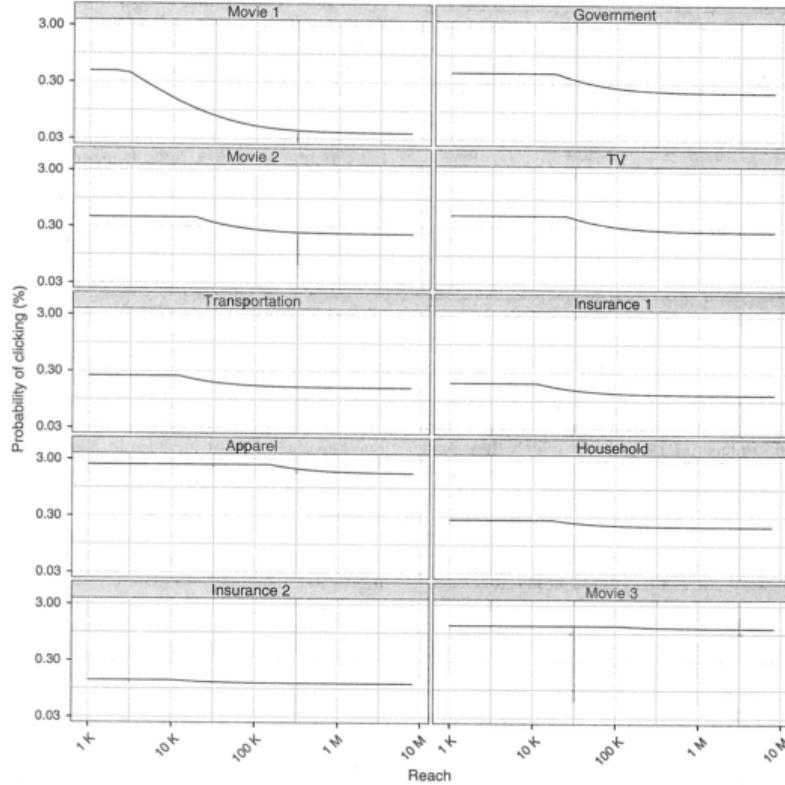
- ▶ People with contacts who clicked on the ad are much more likely to click on the ad (probably not causal). Seems like a good chance to make some money.

Domain	Click rates for individuals without contacts who clicked (%)	Click rates for individuals with contacts who clicked (%)	Percentage of individuals with contacts who clicked
Movie 1	0.038	0.47	0.036
Government	0.209	0.46	0.225
Movie 2	0.225	0.44	0.239
TV	0.260	0.50	0.303
Transportation	0.155	0.25	0.160
Insurance 1	0.124	0.19	0.138
Apparel	1.723	2.43	1.881
Household	0.205	0.27	0.222
Insurance 2	0.118	0.13	0.129
Movie 3	1.185	1.30	1.335

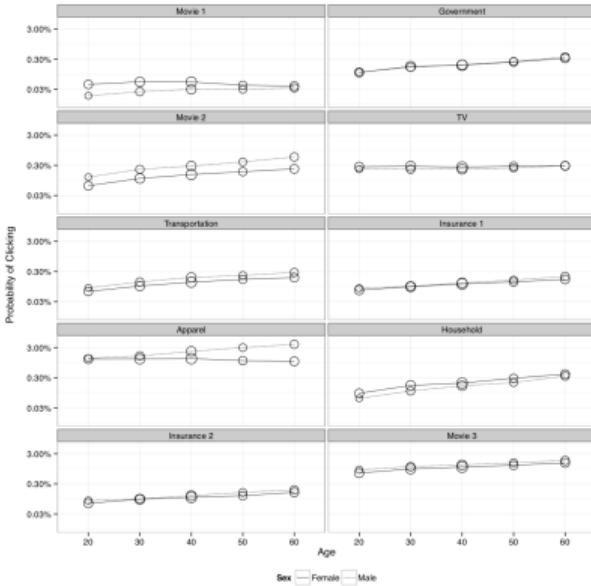
- ▶ People with contacts who clicked on the ad are much more likely to click on the ad (probably not causal). Seems like a good chance to make some money.
- ▶ But. . . . very few people have contacts who clicked on the ads.



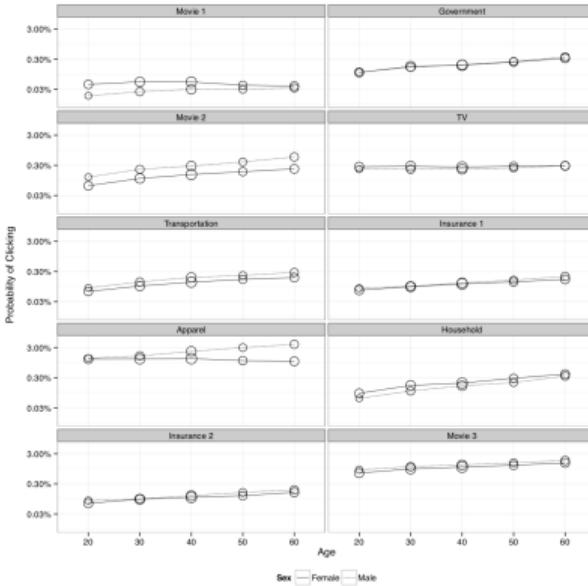
Top-k analysis: order people by the predicted probability of clicking, calculate predicted probability for different sized buckets of people



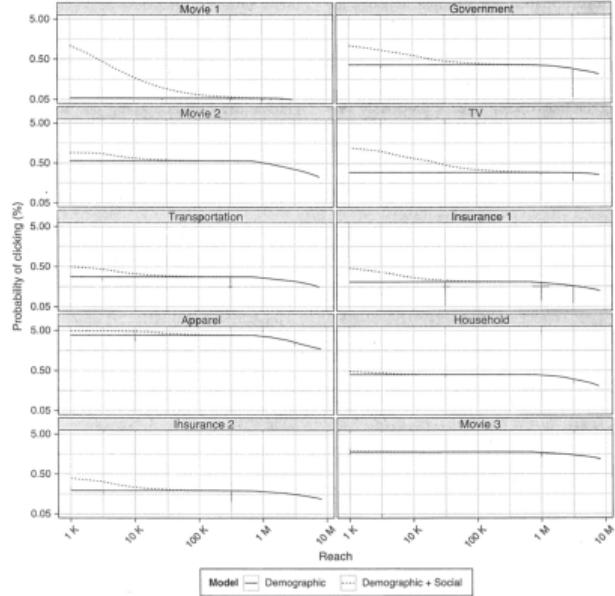
- ▶ social signal allows you to construct pools of between 10,000 and 100,000 candidates who are more likely to click



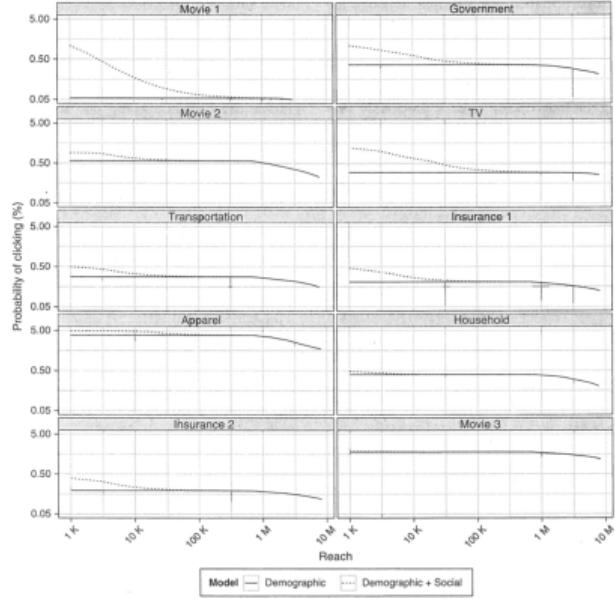
- ▶ But there are important demographic differences as well.



- ▶ But there are important demographic differences as well.
- ▶ Does social add value even if you already have demographic data?

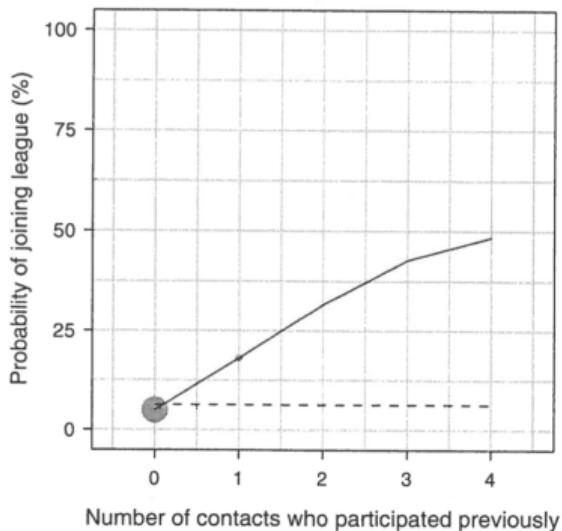


- ▶ Demographic data alone can find large numbers of people who are more likely than average to click



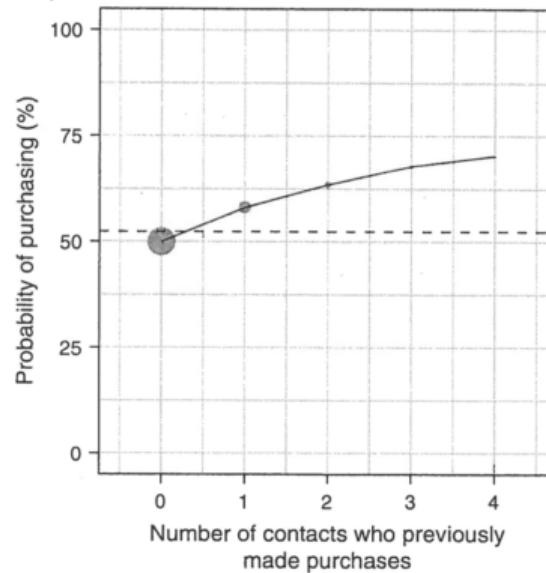
- ▶ Demographic data alone can find large numbers of people who are more likely than average to click
- ▶ Even given that, social data since finds smaller groups more likely to click. But this group is small enough that it might not matter much.

(a) Probability of joining the fantasy football league in 2009 related to the number of social contacts one has who joined in the previous year



(a) Fantasy football

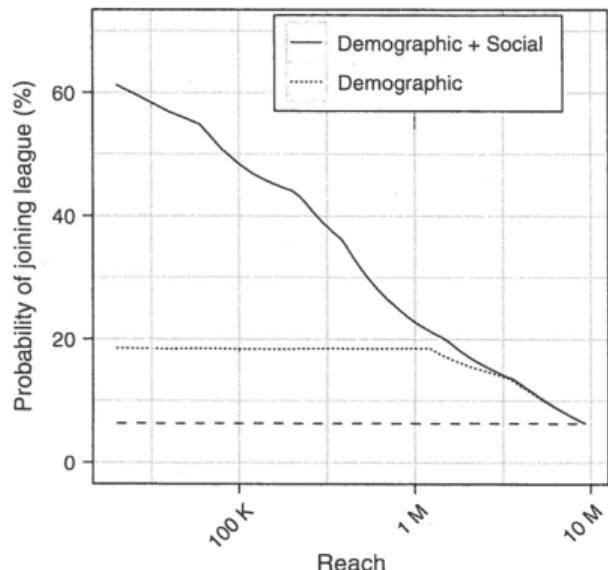
(a) Probability of purchasing at the department store in the second period related to the number of social contacts one has who shopped at the store in the first period



(b) Retail purchase

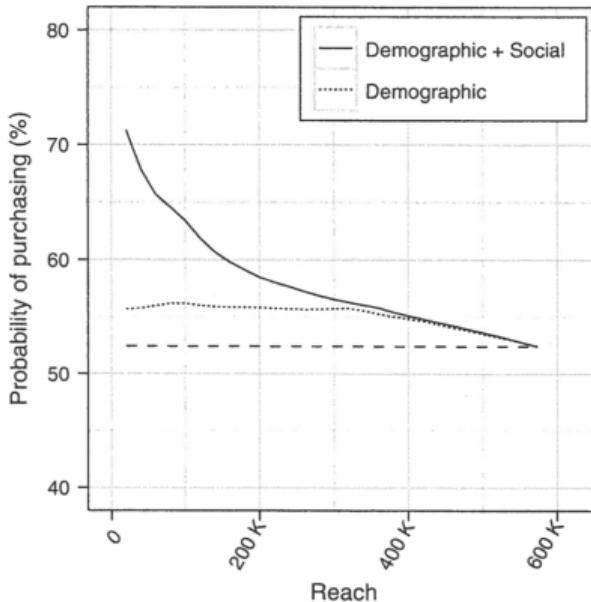
- ▶ Probability of doing activity goes up with number of friends doing it but most people don't have friends going it

Figure 4 Probability of Joining the Fantasy Football League for Varying Numbers of High-Scoring Individuals Under a Demographic Model and a Model That Includes Both Demographic and Social Attributes



(a) Fantasy football

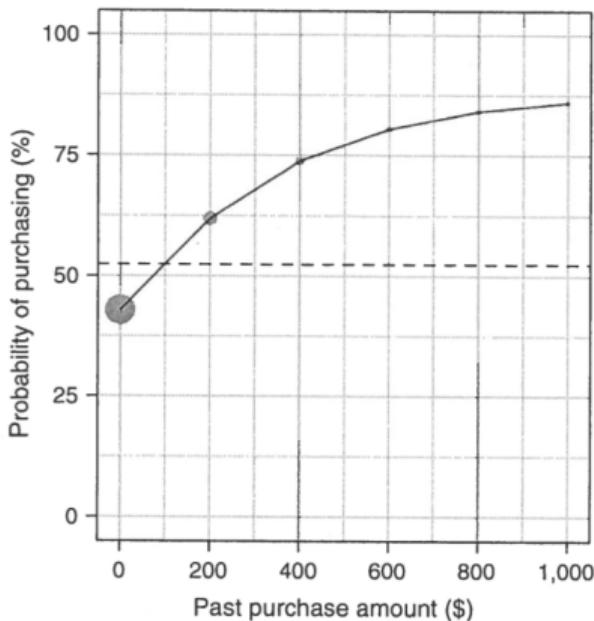
Figure 7 Purchase Rates for Varying Numbers of High-Scoring Individuals Under a Demographic Model and a Model That Includes Both Demographic and Social Attributes for the Shopping Domain



(b) Retail purchase

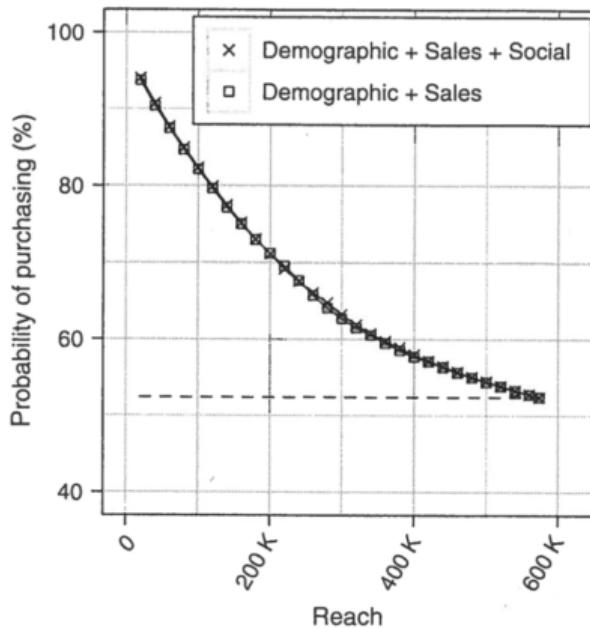
- ▶ Demographics beats baseline
- ▶ Social + demographics beats demographics alone

(a) Probability of purchase in the second period as a function of the amount spent at the store in the first period



- ▶ People who spent more money in the past are more likely to purchase in next time period

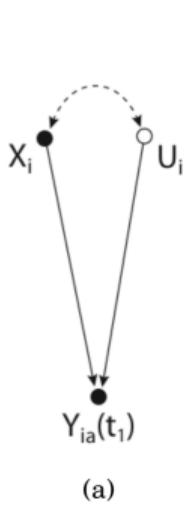
(b) Purchase rates for varying numbers of high-scoring individuals under a model that incorporates demographics and past sales vs. a model that adds the additional feature of the number of contacts who previously made purchases



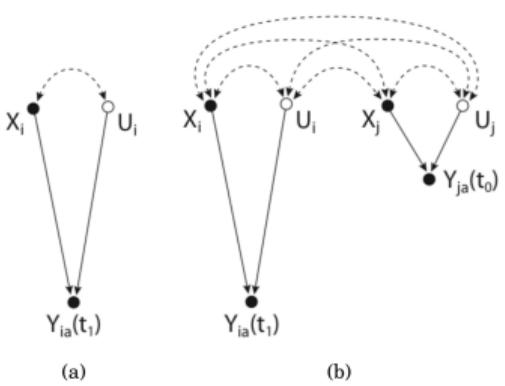
- ▶ Additional value of social is zero once model already has demographics and sales data

What about other ways that social data could be used?

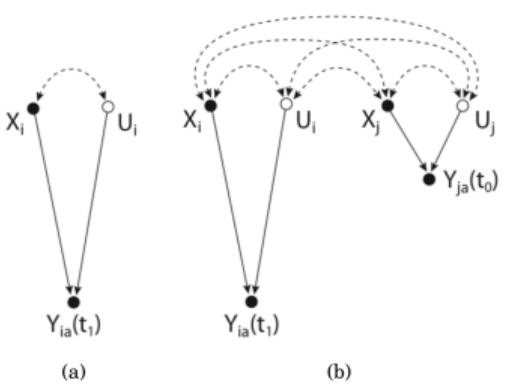
Bakshy et al. “We regard social advertising as any advertising methods that uses information about consumers’ social networks to target ads and/or provide personalized social signals.”



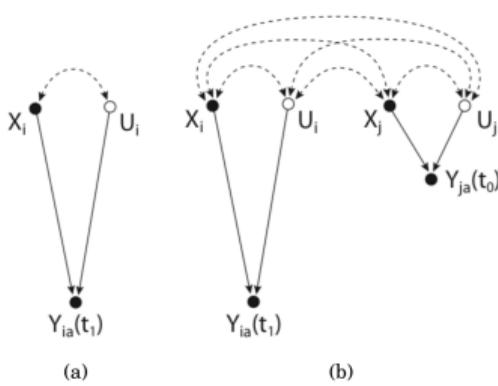
- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .
- ▶ Measured characteristics X_i and unmeasured characteristics U_i might be correlated in unknown ways.



- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .

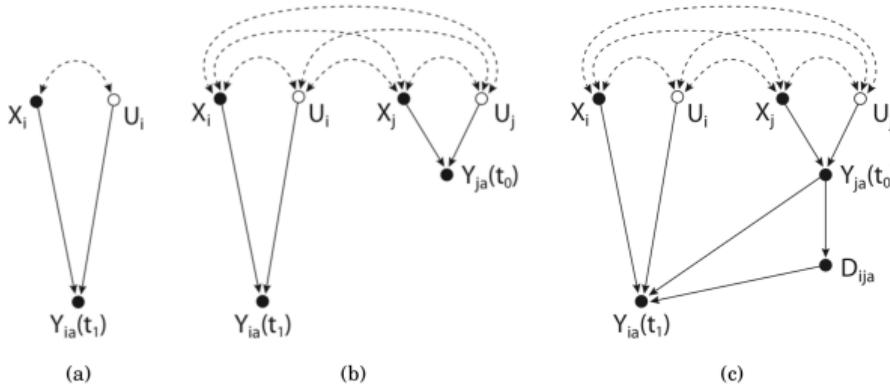


- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .
- ▶ $Y_{ja}(t_0)$ (person j clicking on ad a at time 0) is caused by measured characteristics X_j and unmeasured characteristics U_j .

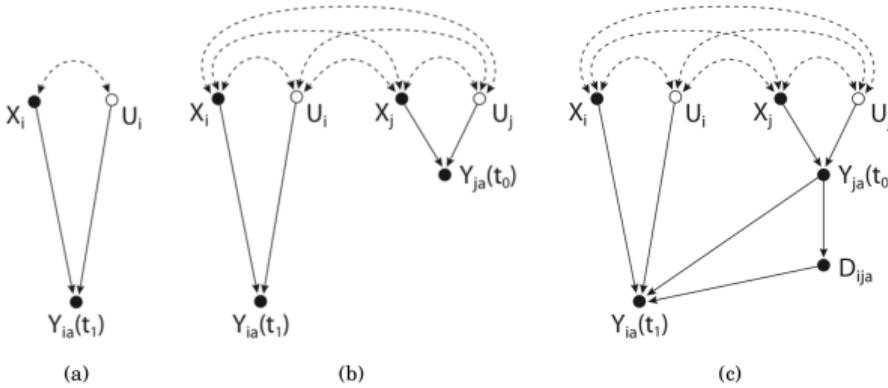


- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .
 - ▶ $Y_{ja}(t_0)$ (person j clicking on ad a at time 0) is caused by measured characteristics X_j and unmeasured characteristics U_j .
 - ▶ Measured characteristics X_i and X_j and unmeasured characteristics U_i and U_j might be correlated in unknown ways.

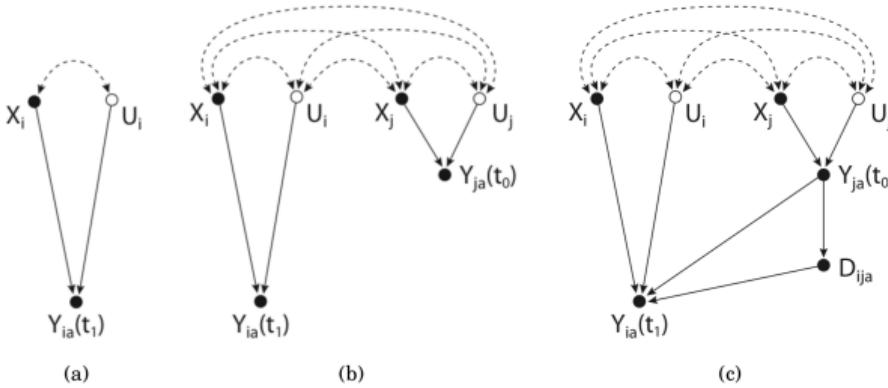
When there is homophily j 's behavior at time 0 predicts i 's behavior at time 1.



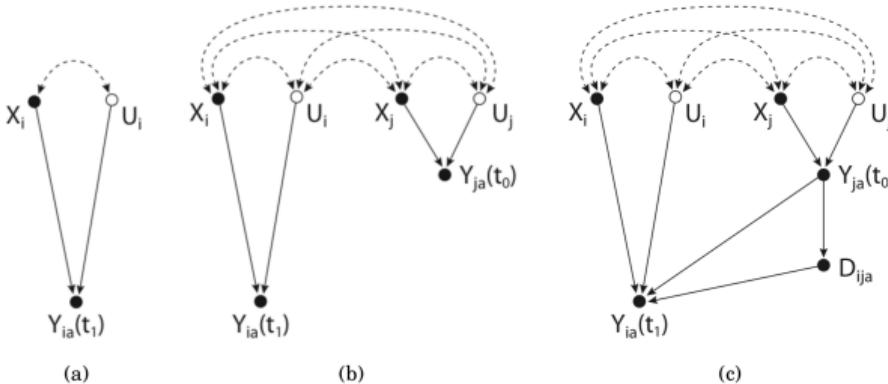
- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .



- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .
- ▶ $Y_{ja}(t_0)$ (person j clicking on ad a at time 0) is caused by measured characteristics X_j and unmeasured characteristics U_j .



- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .
- ▶ $Y_{ja}(t_0)$ (person j clicking on ad a at time 0) is caused by measured characteristics X_j and unmeasured characteristics U_j .
- ▶ Measured characteristics X_i and X_j and unmeasured characteristics U_i and U_j might be correlated in unknown ways.



- ▶ $Y_{ia}(t_1)$ (person i clicking on ad a at time 1) is caused by measured characteristics X_i and unmeasured characteristics U_i .
- ▶ $Y_{ja}(t_0)$ (person j clicking on ad a at time 0) is caused by measured characteristics X_j and unmeasured characteristics U_j .
- ▶ Measured characteristics X_i and X_j and unmeasured characteristics U_i and U_j might be correlated in unknown ways.
- ▶ Most important new addition: social signal D_{ija} might impact $Y_{ia}(t_1)$.

If you see correlated behavior what is causing that? Social influence or correlation of unmeasured characteristics?



Clifton likes Tough Mudder.



Tough Mudder
Like

(a)



Clifton and Joey like
Tough Mudder.



Tough Mudder
Like

(b)



Clifton, Joey and
Maciej like Tough Mudder.



Tough Mudder
Like

(c)



Facebook exec says we're moving from the “wisdom of crowds” to the “wisdom of friends”

Dean Takahashi

January 24, 2011 5:21 AM

[f](#) [t](#) [in](#)

Rose said that when you see an ad with your friend's name in it, the recall of that ad goes up by 60 percent. Facebook has been working closely with Nielsen, which measures ad effectiveness, for a couple of years to make Facebook's ads effective by connecting them to your social circle of friends.

“Word of mouth marketing is the best advertising,” he said. “At Facebook, we are implementing it at scale.”

TECHNOLOGY

Facebook Reaches Settlement in Sponsored Stories Lawsuit

A judge in California today approved a \$20 million settlement in a class action lawsuit against Facebook, over their use of users' likenesses (and 'likes') in paid advertisements.

BRIAN FELDMAN AUGUST 27, 2013

[https://www.theatlantic.com/technology/archive/2013/08/
facebook-reaches-settlement-sponsored-stories/311753/](https://www.theatlantic.com/technology/archive/2013/08/facebook-reaches-settlement-sponsored-stories/311753/)



Clifton likes Tough Mudder.



Tough Mudder
Like

(a)



Clifton and Joey like
Tough Mudder.



Tough Mudder
Like

(b)



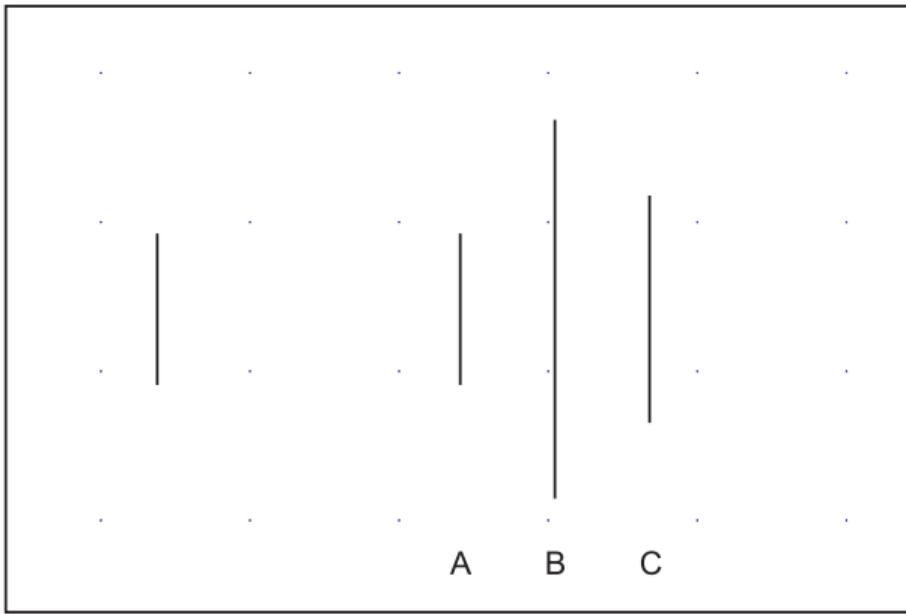
Clifton, Joey and
Maciej like Tough Mudder.

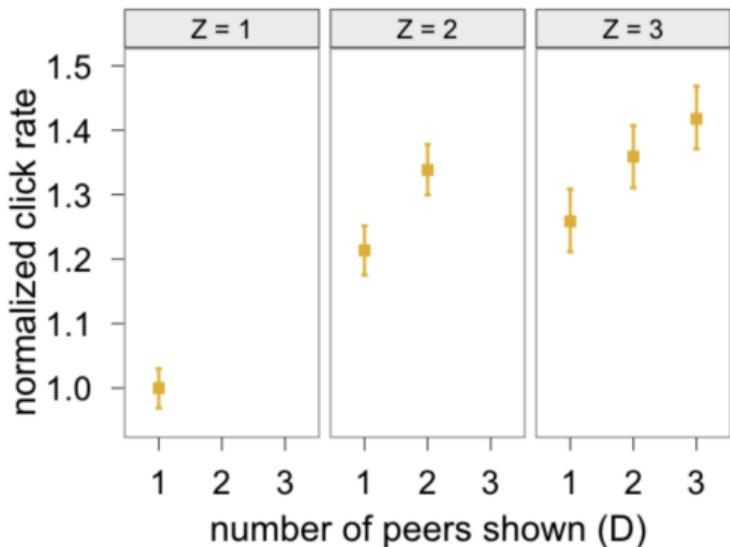


Tough Mudder
Like

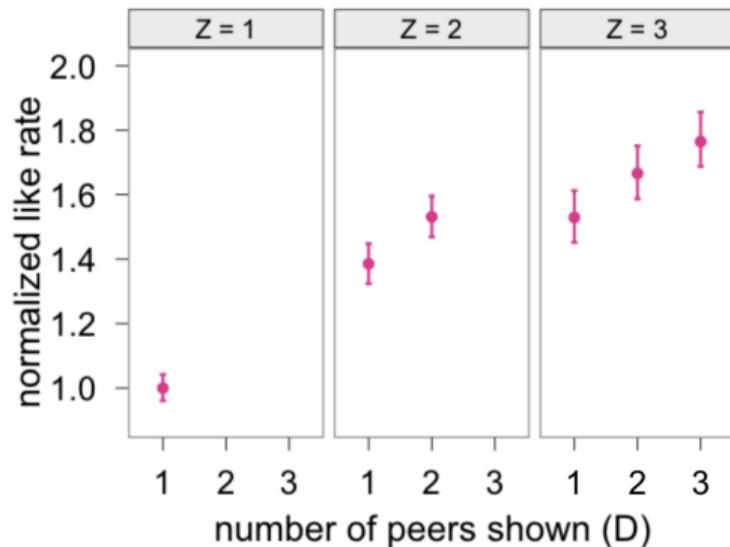
(c)

7.1





(a)



(b)

- ▶ For each value of Z , more peers shown more clicks and likes
- ▶ Difference between $D = 1$ for different values of Z suggests homophily on unobservable characteristics (but with caution)
- ▶ They don't publish raw rates, and papers often go through business review before sending them to a journal

Experiment 2:

History



"Like" HISTORY and enter to win a free trip to New York City and \$5,000.

Like · 357,462 people like this.

(a)

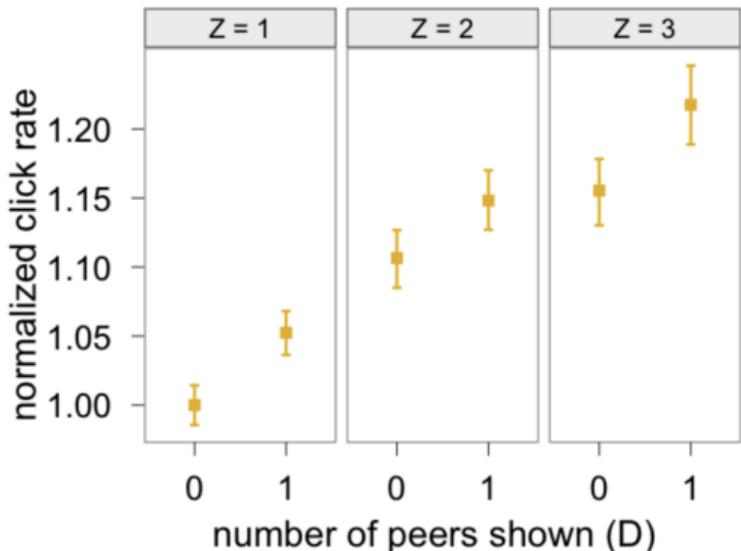
History



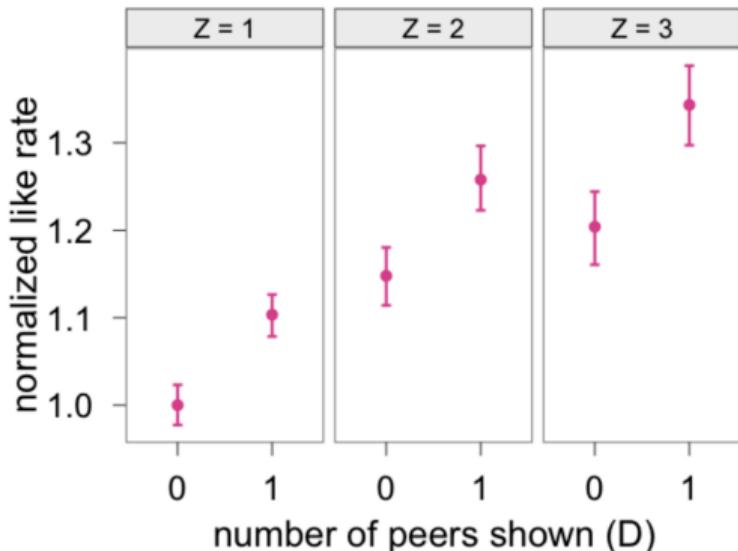
"Like" HISTORY and enter to win a free trip to New York City and \$5,000.

Like · Jina [redacted] likes this.

(b)



(a)

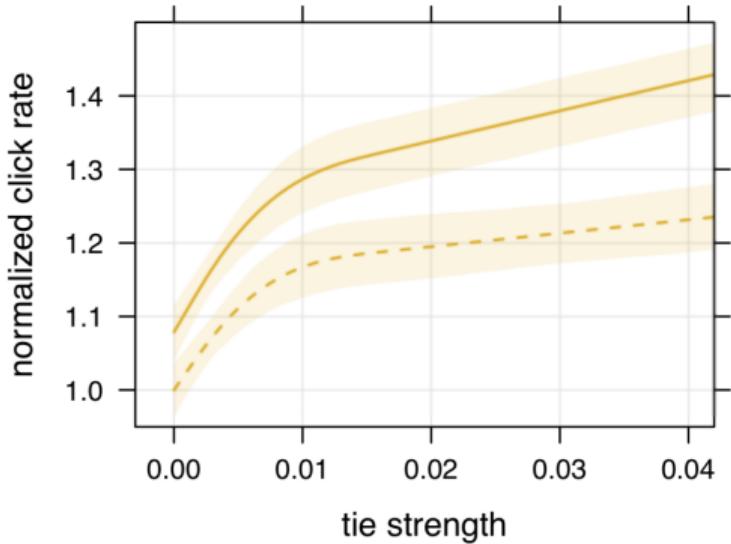


(b)

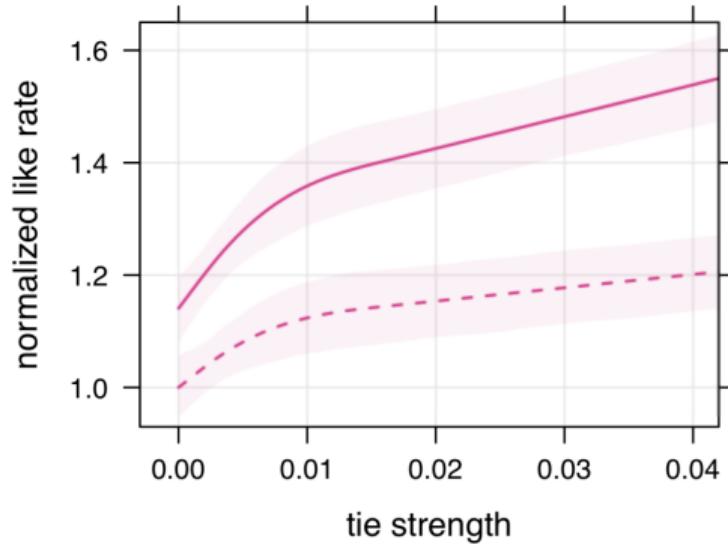
- ▶ For each value of Z , showing specific peer leads to more clicks and likes
- ▶ Difference between $D = 0$ for different values of Z suggests homophily on unobservable characteristics (but with caution)

What about tie strength? Who is the best person to mention to maximize click-through rates?

Tie strength between i and j is measured the fraction of i 's communication that is directed at j or posts by j .



(a)



(b)

- As tie strength increases, click and like rate increase whether the signal is present (solid) or absent (dashed). The risk ratio increased as tie strength increased (Fig 8).

You can bet that this information was used to decide which friend to show in the ad



<http://pubsonline.informs.org/journal/mksc>

MARKETING SCIENCE

Vol. 39, No. 6, November–December 2020, pp. 1142–1165

ISSN 0732-2399 (print), ISSN 1526-548X (online)

Social Advertising Effectiveness Across Products: A Large-Scale Field Experiment

Shan Huang,^a Sinan Aral,^b Yu Jeffrey Hu,^c Erik Brynjolfsson^b

^a Michael G. Foster School of Business, University of Washington, Seattle, Washington 98195; ^b Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142; ^c Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30332

- ▶ Replicate and extend Baskhy et al. on WeChat
- ▶ 71 products, 25 categories, 37 million users



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前

推广



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前



成成

推广



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前



成成, adrianli德苑, xLi, 浩君

- ▶ 3 treatments: no social cue, one like, organic likes



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前

推广



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前



成成

推广



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前

推广



成成, adrianli德苑, xLi, 浩君

- ▶ 3 treatments: no social cue, one like, organic likes
- ▶ main comparison is between no social cue and one like



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前

推广



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

10分钟前



成成



宝马中国

超模Gigi Hadid 到底上了谁的车?



查看详情 ↗

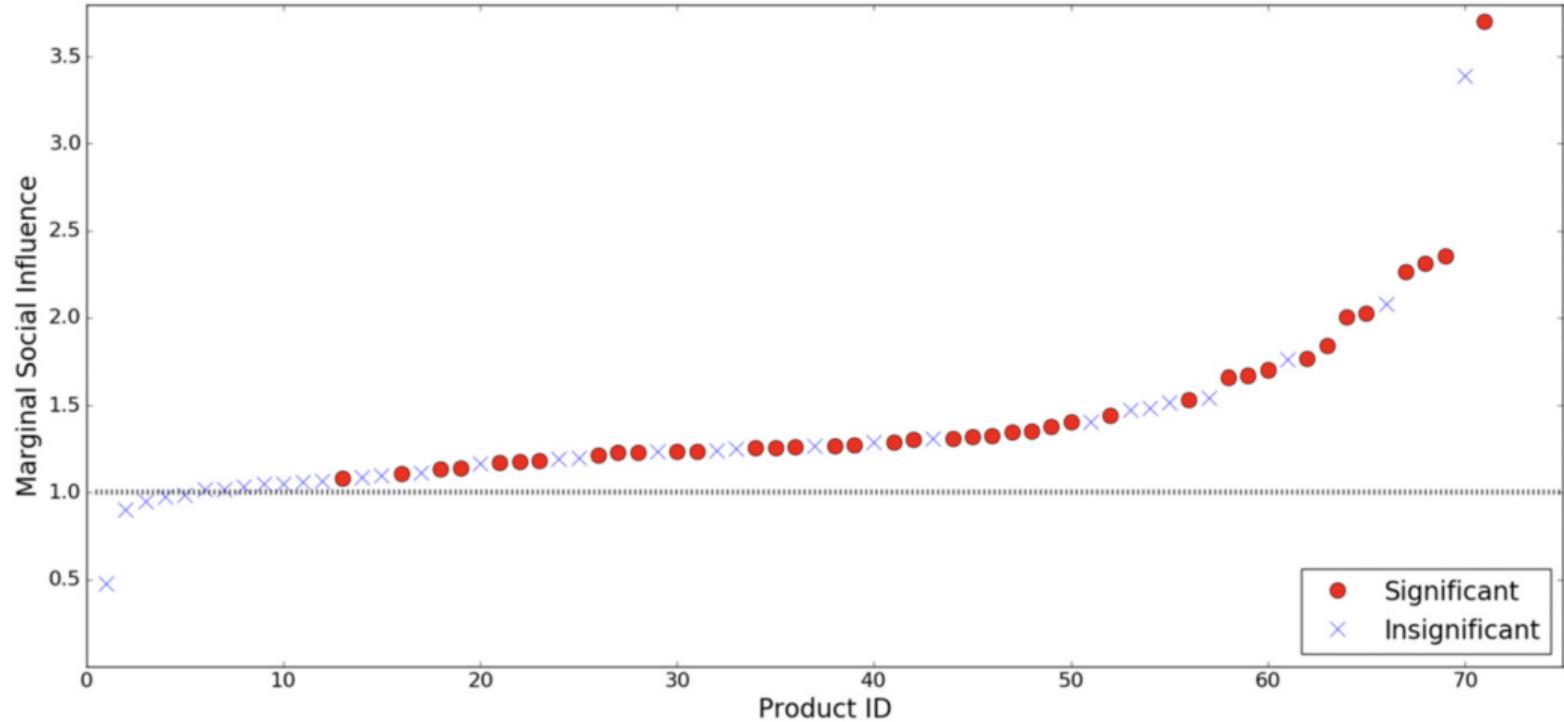
10分钟前

推广

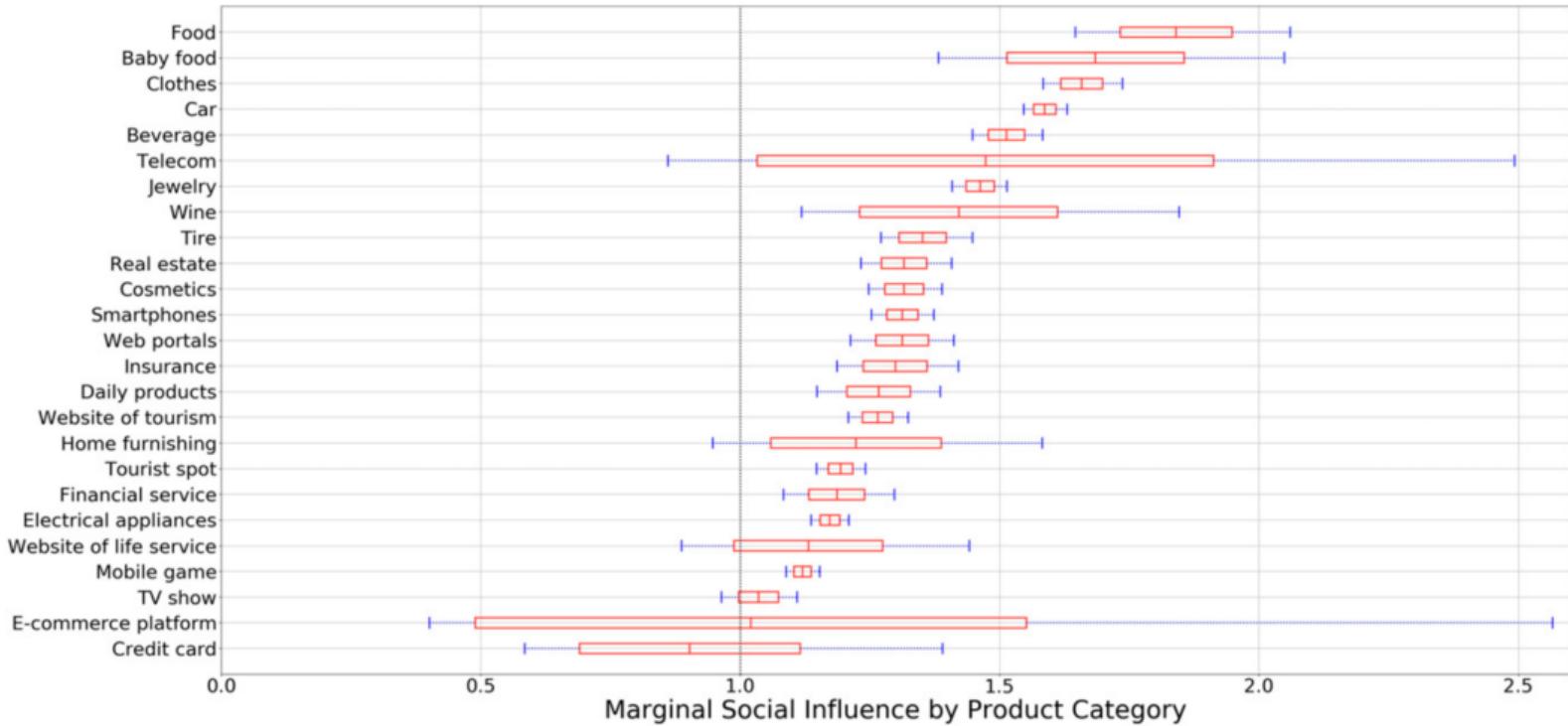


成成, adrianli德苑, xLi, 浩君

- ▶ 3 treatments: no social cue, one like, organic likes
- ▶ main comparison is between no social cue and one like
- ▶ displaying a social cue make users 33.75% more likely to click on an ad



- ▶ Effect was positive for virtually all 71 products



- ▶ Effect was biggest for food, baby food, clothes, and cars.
- ▶ Effect was smallest for credit card, e-commerce platform, and a TV show

Stepping back:

- ▶ social media enable social advertising

Stepping back:

- ▶ social media enable social advertising
- ▶ network data can be used to target ads. The value of adding network data depends on the type of information the advertiser already has. In some cases, social helped in other cases it does not.

Stepping back:

- ▶ social media enable social advertising
- ▶ network data can be used to target ads. The value of adding network data depends on the type of information the advertiser already has. In some cases, social helped in other cases it does not.
- ▶ network data can also be used as part of the ad itself, and social signals of different kinds can increase engagement (click-through-rate, likes)

Fixing social media:

- ▶ Frank, R.H. (2021). The economic case for regulating social media. *New York Times*.
- ▶ Pennycook, G. and Rand, D. (2020). The right way to fight fake news. *New York Times*.
- ▶ Bail, C.A. et al. (2018). Exposure to opposing views on social media can increase political polarization. *PNAS*.
- ▶ Kaiser, B., Mayer, J. and Mattias, J.M. (2021). Warnings that work: Combating misinformation without deplatforming. *Lawfare*.
- ▶ Kaiser, B. et al. (2021). Adapting security warnings to counter online disinformation. *30th USENIX Security Symposium (USENIX Security 21)*.