Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads

Anja Lambrecht and Catherine Tucker

Research Question

What may make an ad serving algorithm appear biased?

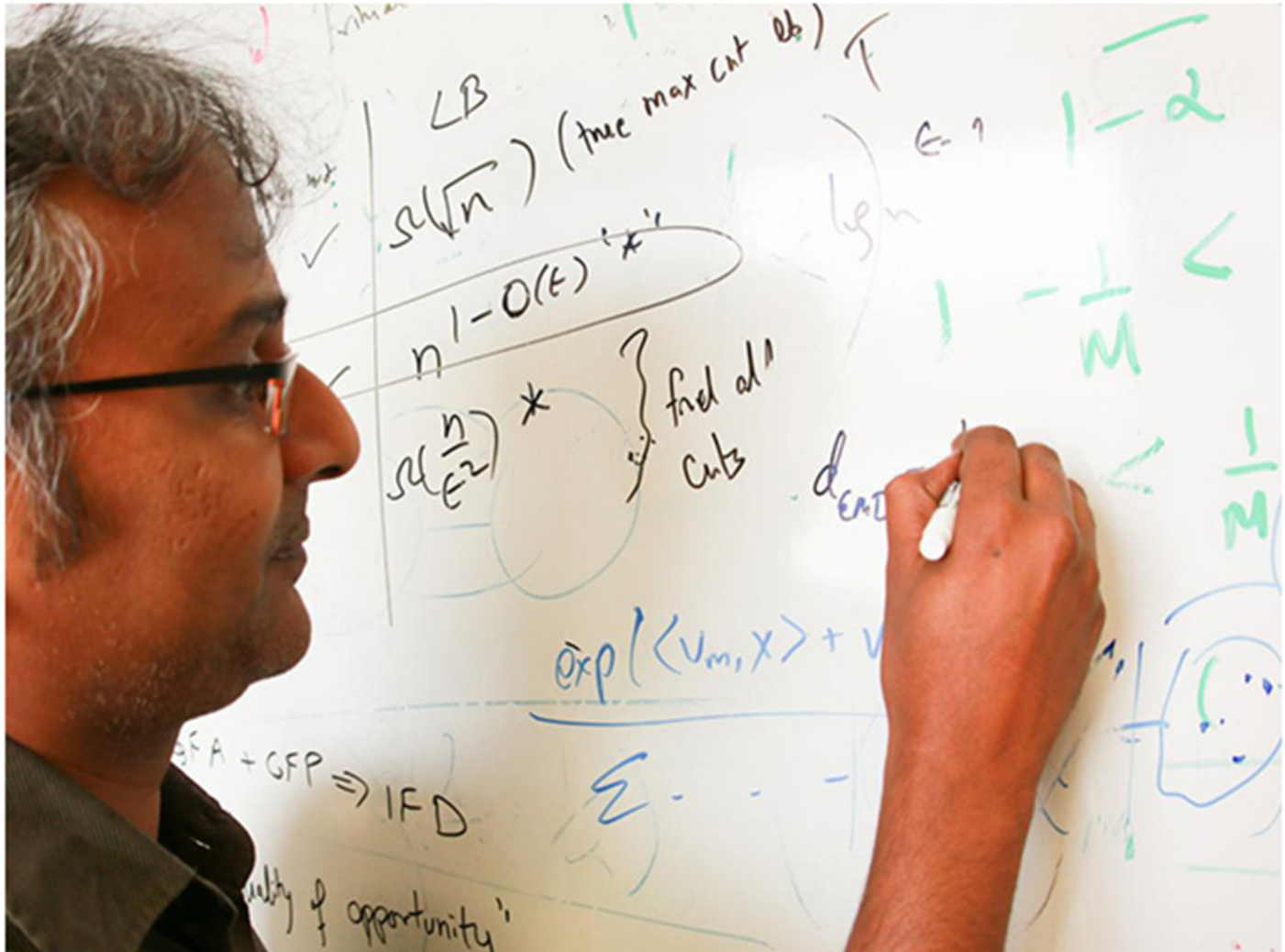
Motivation

- Papers in Computer Science have documented empirical pattern of apparently discriminatory ad serving behavior (Sweeney, 2013; Datta et al., 2015)

Computer Scientists Find Bias in Algorithms

By Lauren J. Young

Posted 21 Aug 2015 | 20:00 GMT




stephanie williams
My Network
Jobs
Interests
2,513 results for stephanie williams
Did you mean stephen williams?
Some search results have been filtered to improve relevance.
Show all results

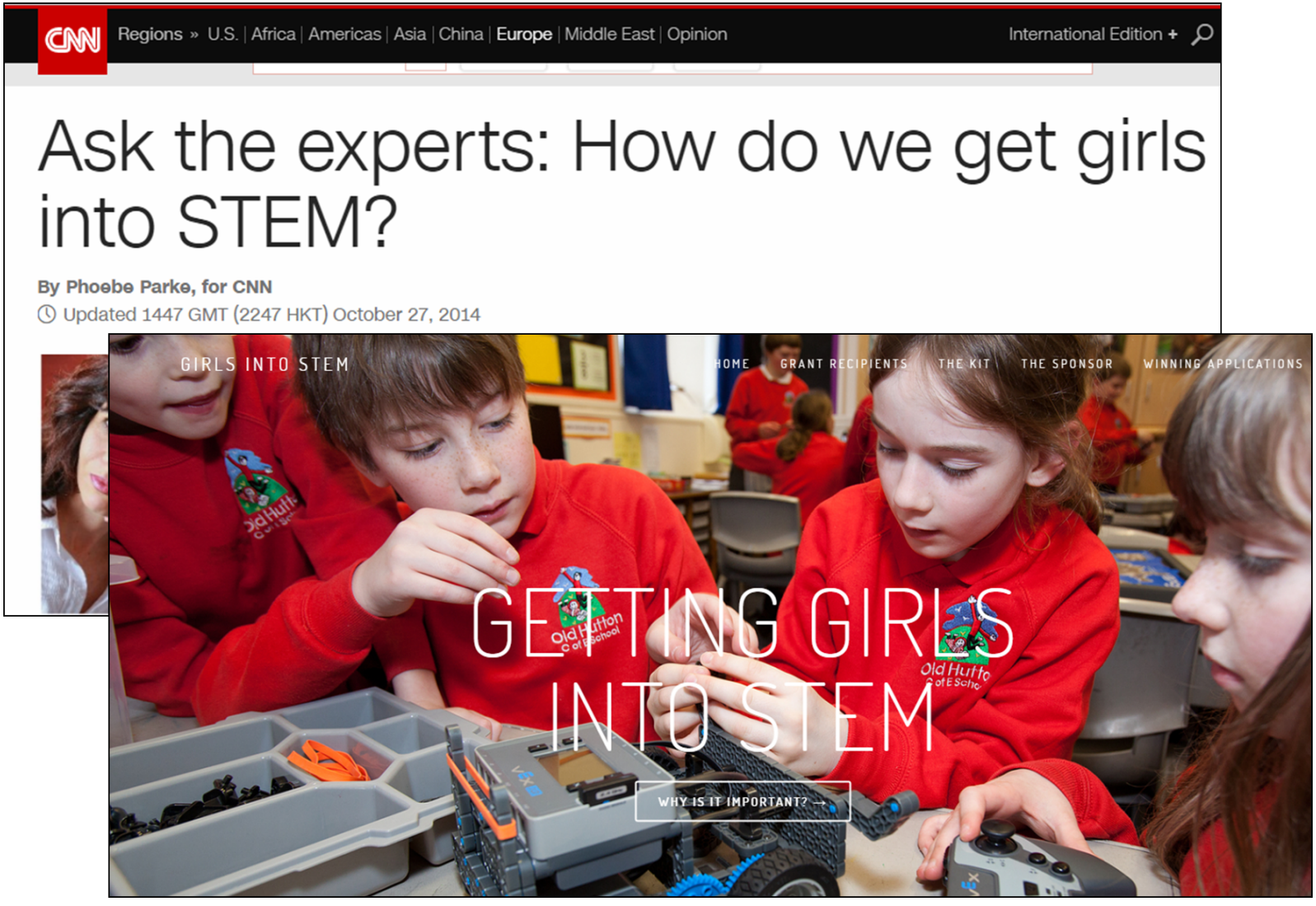
Why do we observe algorithmic bias?

- Tech's white male problem
- Algorithm learns discriminatory behavior from data
- Algorithm learns discriminatory behavior from people using it
- Maybe economics can help us understand why (apparent) algorithmic bias happens?

What we do

- Field Test data on STEM ad across 190 countries
  - Set up as gender neutral
  - But shown to men more than women

STEM ads are something we might worry about being distorted away from women



Why apparent algorithmic bias happens

- Not because of
  - Click propensity
  - Media usage
  - Underlying sexism
- Evidence that young women are valuable demographic and other advertiser bids crowd out intentionally gender neutral advertisers

Why does this matter?

- First paper to explore the why of apparent algorithmic-bias
- We find that apparent algorithmic bias may not be intentional but instead the result of other actors' economic behavior
- Fight urge to argue that it may be efficiency not 'discrimination' - that is our point.

Why this matters for policy

Home / Software

The FTC is worried about algorithmic transparency, and you should be too

Internationally algorithmic transparency gaining support

Merkel: murky internet giants distort perception of reality

DPA/The Local
news@thelocal.de

26 October 2016
11:51 CEST+02:00

merkel

facebook

google

Share this article

Angela Merkel. Photo: DPA.

Chancellor Angela Merkel called on Tuesday for internet giants to make public their closely-guarded algorithms, claiming that they are not giving people diverse enough information.

Merkel said at a media conference in Munich that "the algorithms must be made public, so

Policy implications

- Not much support in our findings for ‘Algorithmic Transparency’ as a policy solution
- Shows the need for understanding the underlying economics

Outline

Field Test

Data

Empirical Evidence

Results
- Do men indeed see more STEM ads than women?

Replication on Other Platforms
- Google Adwords
- Instagram
- Twitter

Implications

Sample ad

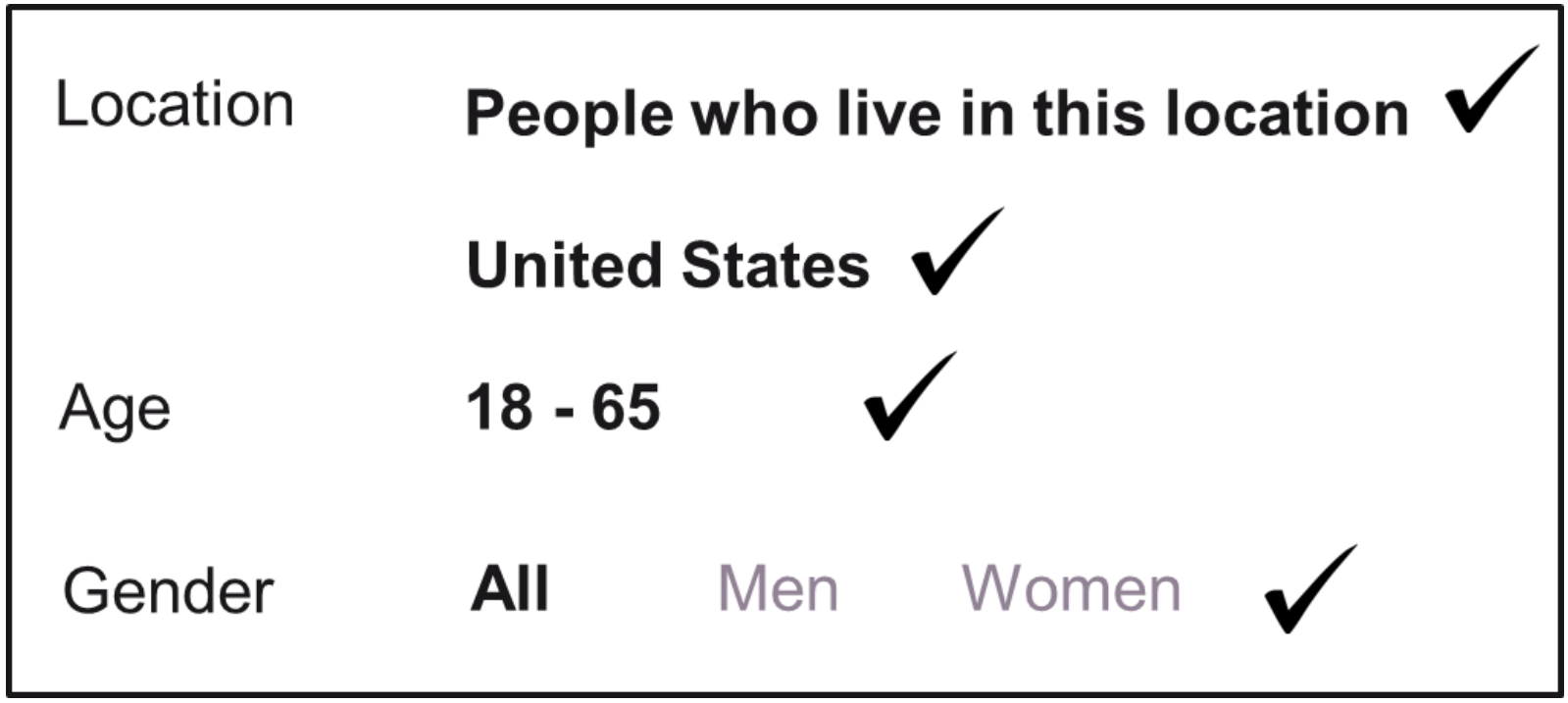
STEM Careers

Information about STEM Careers

This was a very straightforward field test

- All that varied was the country it was targeted at
- 191 countries
- Ensured that in each country the ad was shown at least to 5000 people

Ad targeting settings - intended to be shown to both men and women aged 18-65



To remind people of the obvious

- Ads are shown to eyeballs based on an ad-auction
- Therefore we had to ‘win’ the auction for that pair of eyeballs for our ad to be shown
- Our winning the auction is a function of the price we bid, our ‘quality score’, other bidders

Outline

- Field Test
- Data
- Empirical Evidence
- Results
  - Do men indeed see more STEM ads than women?
- Replication on Other Platforms
  - Google Adwords
  - Instagram
  - Twitter
- Implications

	Mean	Std Dev	Min	Max
Impressions	1911.8	2321.4	0	24980
Clicks	3.00	4.52	0	42
Unique Clicks	2.78	4.15	0	40
CPC	0.085	0.090	0	0.66
Reach	615.6	850.7	0	13436
Frequency	4.38	4.32	1	53

Table: Summary statistics at Campaign-Segment level

Outline

Really, this paper doesn’t need any complex analysis

Table: Raw Data reported

Age Group	Male Impr.	Female Impr.	Male Click Rate	Female Click Rate
Age18-24	746719	649590	0.0015	0.0018
Age25-34	662996	495996	0.0013	0.0015
Age35-44	412457	283596	0.0012	0.0017
Age45-54	307701	224809	0.0013	0.0018
Age55-64	209608	176454	0.0015	0.0021
Age 65+	192317	153470	0.0016	0.0021

Three obvious patterns in the data

- Men see more impressions of the ad than women
- Particularly in younger ad cohorts
- Women do not click less often

Outline

Do men indeed see more STEM ads than women?

For campaign i and demographic group j in country k on day t, the number of times an ad is displayed is modeled as a function of:

$$
\begin{aligned}
AdDisplay_{ijkt} = & \\
& + \quad \beta_1 Female_j \\
& + \quad \beta_2 Age_j \\
& + \quad \beta_3 Female_j \times Age_j \\
& + \quad \alpha_k + \epsilon_{jk} \qquad (1)
\end{aligned}
$$

Women are shown fewer ads than men

	(1) Impressions	(2) Impressions	(3) Reach	(4) Reach
Female	-479.3***	-209.7***	-228.1***	-98.97***
Female × Age18-24		-298.8		-234.3**
Female × Age25-34		-664.6***		-302.2***
Female × Age35-44		-464.9***		-159.9***
Female × Age45-54		-224.2**		-97.25***
Female × Age55-64		36.16		18.93
Age18-24	2753.6***	2902.6***	909.5***	1026.5***
Age25-34	2132.4***	2464.3***	561.4***	712.3***
Age35-44	920.5***	1152.6***	197.4***	277.2***
Age45-54	492.4***	604.1***	99.08**	147.5***
Age55-64	109.0*	90.53^{+}	16.56	6.911
Country Controls	Yes	Yes	Yes	Yes
Observations	2291	2291	2291	2291
R-Squared	0.485	0.488	0.442	0.446

Ordinary Least Squares Estimates. Dependent variable as shown. Omitted demographic groups are those aged 65+ and men. Robust standard errors. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Do our results directly reflect human behavior that the algorithm learns?

If they see the ad, women are more likely to click than men

	(1) Clicks	(2) Unique Clicks	(3) Clicks	(4) Unique Clicks
Female	0.221***	0.303***	0.264**	0.399***
	(0.0271)	(0.0290)	(0.0932)	(0.0875)
Female × Age18-24			-0.137	-0.166+
			(0.0975)	(0.0956)
Female × Age25-34			-0.0899	-0.135
			(0.113)	(0.109)
Female × Age35-44			0.0822	-0.0289
			(0.113)	(0.109)
Female × Age45-54			0.0633	0.000689
			(0.119)	(0.117)
Female × Age55-64			0.0465	-0.0573
			(0.136)	(0.129)
Age18-24	-0.175**	-0.214***	-0.105	-0.129+
	(0.0576)	(0.0557)	(0.0731)	(0.0704)
Age25-34	-0.375***	-0.460***	-0.332***	-0.394***
	(0.0593)	(0.0572)	(0.0823)	(0.0785)
Age35-44	-0.341***	-0.409***	-0.379***	-0.392***
	(0.0712)	(0.0657)	(0.0902)	(0.0839)
Age45-54	-0.190**	-0.222***	-0.220*	-0.220**
	(0.0613)	(0.0605)	(0.0865)	(0.0843)
Age55-64	-0.0186	-0.0199	-0.0426	0.00913
	(0.0682)	(0.0666)	(0.0955)	(0.0879)
Country Controls	Yes	Yes	Yes	Yes
Observations	4515014	1453890	4515014	1453890

Aggregate Logit Estimates. Age controls not shown. Omitted demographic groups are those aged 65+ and men. Robust standard errors. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Do our results reflect some capacity constraint faced by the algorithm?

- Do women spend less time on social media?
- No.
- At least every piece of recorded data says no.

Do our results reflect that the algorithm tries to distribute clicks equally across demographic groups?

- Click rates mean that number of clicks by men and women in an age group is broadly similar
- But little evidence of balancing of clicks across age segments - indeed appears quite distorted

Do our results reflect cultural prejudice or labor market conditions for women that the algorithm has learned?

- Collected data from World bank for variables which capture potential labor market prejudice against women

Algorithm does not appear to 'learn' from bias in country

	(1) Reach	(2) Reach	(3) Reach	(4) Reach	(5) Reach
Female	-208.5***	-183.0***	-249.8***	-225.3***	-237.8***
Female × High % Female Labor Part=1	-59.40				
Female × High % Female Primary=1		-139.0			
Female × High % Female Secondary=1			69.07		
Female × High Female Equality Index=1				-20.82	
Female × High GDP=1					32.22
					(60.94)
Country Controls	Yes	Yes	Yes	Yes	Yes
Observations	2291	2291	2291	2291	2291
R-Squared	0.442	0.443	0.442	0.442	0.442

Ordinary Least Squares Estimates. Age controls not shown. Dependent variable is whether someone is exposed to an ad. Omitted demographic groups are those aged 65+ and men. Robust standard errors. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Do our results simply reflect competitive spillovers?

Does price matter?

- Remaining explanation could be other actors' bidding behavior
- Across all campaigns, the average cost per click was nearly identical for men and women ($0.09)
- But maybe we just were not bidding high enough to reach women. So we collected separate data on suggested bids by demographic segment

Women are more expensive to advertise to on social media

	(1) Avg Suggested Bid	(2) Avg Suggested Bid	(3) Avg Suggested Bid
Female	0.0534*	0.0525*	-0.0464
Female × Age18-24			0.0648+
Female × Age25-34			0.174+
Female × Age35-44			0.150***
Female × Age45-54			0.0751
Female × Age55+			0.129**
Country Controls	No	Yes	Yes
Observations	2096	2096	2096
R-Squared	0.00443	0.569	0.571

Ordinary Least Squares Estimates. Age controls not shown. Omitted demographic groups are those aged between 13-17 and those of the male gender. Robust standard errors. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We are not the first to notice this

resolution media

MEN ARE CHEAP

KENSHOO social

View full report @ Kenshoo.com/MenAreCheap or ResolutionMedia.com/white-papers/MenAreCheap

Men are also less expensive to reach with Facebook ads.

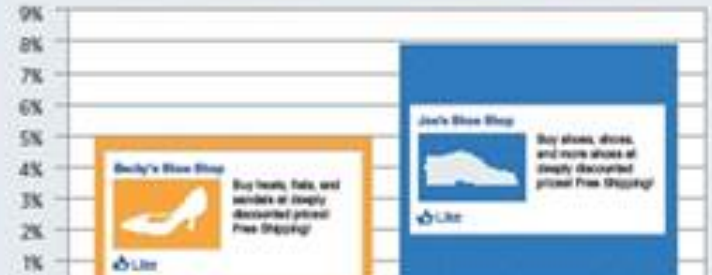
Cost-Per-Thousand (CPM)Impressions for Facebook Ads

Cost per Click (CPC) Impressions for Facebook Ads

Therefore, you can reach more of a unique audience when targeting men.

And you can afford to show them the same ad more often.
Sometimes men just need to be reminded :)

Exposure Rate for Facebook Ads



Frequency of Facebook Ads

Are women of higher value to advertisers?

- Use separate data set on the purchasing of consumer items as a result of a social media campaign
- Find that conditional on clicking, women are more likely to purchase than men - suggesting that women may offer higher ROI
- Thus, advertisers could benefit from bidding more on women

Young women may be a valuable demographic as they appear more likely to convert

	Clicks Out of Impressions (1)	Add-to-cart Out of Clicks (2)	Add-to-cart Out of Impressions (3)
Female	-0.0522***	-0.0231	-0.0979
Age Group 18-24	-0.795***	-0.528	-1.392**
Age Group 25-35	-0.533***	-0.149	-0.742***
Age Group 35-44	-0.244***	-0.168	-0.430**
Female $\times$ Age Group 18-24	0.408***	1.078*	1.553***
Female $\times$ Age Group 25-35	-0.0602**	0.701**	0.709**
Female $\times$ Age Group 35-44	-0.000403	0.509*	0.508*
Week Controls	Yes	Yes	Yes
Day of week controls	Yes	Yes	Yes
Product Controls	Yes	Yes	Yes
Observations	127617816	67501	127605845

Aggregate logit estimates. Dependent variable as listed. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Omitted demographic groups are men and those aged 45+.

Women are prized – because they make purchase decisions

- ”Women Make Up 85% of All Consumer Purchases” (Bloomberg)
- ”Women drive 70-80% of all consumer purchasing” (Forbes)
- ”Sorry, Young Man, You’re Not the Most Important Demographic in Tech” (The Atlantic)

Replication on Google Adwords

Table: Results of test on Google Display Network

Gender	Impressions	Clickrate	Cost Per Click
Female	26,817	1.71%	0.20
Male	38,000	.97%	0.19

Replication on Google Adwords

- Contextual Display advertising network
- Targeting Criteria ‘Science Jobs’ ‘Engineering Career’
- We used a manual bid strategy where we bid 50 cents per click. We spent $181 for the campaign.

Replication on Instagram

Table: Results of test on Instagram

Gender	Impressions	Clickrate	Cost Per Click
Female	1,560	0.27%	$1.74
Male	9,595	0.59%	0.95

Replication on Instagram

- Very Skewed. Perhaps because men clicked more on the ad.
- Budget $100

Replication on Twitter

Table: Results of test on Twitter

Gender	Impressions	Total Spend
Female	52,363	$31
Male	66,243	$46.84

Replication on Twitter

- We spent $100 total on the campaign.
- No results for clicks by gender, but we can at least see replication of impressions

Outline

- Field Test
- Data
- Empirical Evidence
- Results
  - Do men indeed see more STEM ads than women?
- Replication on Other Platforms
  - Google Adwords
  - Instagram
  - Twitter
- Implications

Limitations

- Descriptive paper
- Just look at gender
- Big (non-economist) questions are not tackled - Should we think of this as bias? Should we think of this as discrimination?
- What is the counterfactual?

Punchline

- ‘Gender-neutral’ STEM ad shown to more women than men
- Not because algorithm responds to click behavior or local prejudice, or because of a ’capacity constraint’ when aiming to target women
- Women’s desirability as a demographic means that an algorithm trained to be cost effective shows relatively less ads to them
- Apparent algorithmic bias may be an unintentional consequence of other actors’ economic behavior

Implications for practice

- Managers can’t assume an algorithm will neutrally deliver ads - even if a campaign is explicitly designed as gender-neutral
- In our case, can be easily solved by managing two separate campaigns for men and women and paying more for women
- But what about cases where the algorithm does not neutrally distribute ads with respect to harder-to-address factors such as economic marginalization or race?

Implications for policy

- Not clear how algorithmic transparency would help here
- Emphasizes the need for nuance in algorithmic auditing policy

Thank you!

alambrecht@london.edu cetucker@mit.edu

Datta, A., M. C. Tschantz, and A. Datta (2015). Automated experiments on ad privacy settings. *Proceedings on Privacy Enhancing Technologies 2015*(1), 92–112.

Sweeney, L. (2013). Discrimination in online ad delivery. *ACMQueue 11*(3), 10.