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

Anja Lambrecht and Catherine Tucker*

March 9, 2018

Abstract

We explore data from a field test of how an algorithm delivered ads promoting job opportunities in the Science, Technology, Engineering and Math (STEM) fields. This ad was explicitly intended to be gender-neutral in its delivery. Empirically, however, fewer women saw the ad than men. This happened because younger women are a prized demographic and are more expensive to show ads to. An algorithm which simply optimizes cost-effectiveness in ad delivery will deliver ads that were intended to be gender-neutral in an apparently discriminatory way, due to crowding out. We show that this empirical regularity extends to other major digital platforms.

*Anja Lambrecht is Associate Professor of Marketing at London Business School. Catherine Tucker is the Sloan Distinguished Professor of Marketing at MIT Sloan School of Management, Cambridge, MA and Research Associate at the NBER. Thank you to NSF CAREER Award 6923256 for financial support. All errors are our own. We thank Ben Edelman for extremely useful comments. We also thank seminar and conference participants at Boston College, Oxford University, University of Colorado Boulder, Vrije Universiteit Amsterdam, Oxford University, FTC-Marketing Science conference, NBER, Economics of Intellectual Property, Software and the Internet Conference for their helpful comments.

1 Introduction

The increased use of algorithms to automate decision-making has sparked deep concern that such automated choices may produce discriminatory outcomes. In settings where ads are allocated by algorithm, research has documented instances where historically discriminated-against groups are more likely to be associated with undesirable ads (Sweeney, 2013) and less likely to see desirable ads (Datta et al., 2015). However, these papers do not attempt to understand *why* ad algorithms can produce apparently discriminatory outcomes.

We explore this question using data from a field test of an ad that was intended to promote job opportunities and training in STEM (Science, Technology, Engineering and Math).[1] Our empirical focus on information about STEM careers is motivated by the fact that policymakers in many countries are concerned about a shortage of graduates in the STEM sector,[2] particularly among women.[3] Providing information about STEM careers is an integral part of this policy challenge because of evidence suggesting that the shortage is not necessarily a result of hiring practices: Williams and Ceci (2015) document for an academic context that, conditional on applying, women are more likely than men to be hired into STEM jobs. Instead, distortions in perceptions about careers in STEM across genders may potentially explain why women do not apply (Diekman et al., 2010). Thus, disseminating information about STEM to women and encouraging women to enter this field is an important policy goal (Cheryan et al., 2011; Shapiro and Williams, 2012).

The targeting of the ad in our field test was intended to be gender-neutral, so the advertiser instructed the ad-serving algorithm to show the ad to both men and women. The ad

[1] This means we focus on disparity of access to information, rather than echoing the majority of the economics literature which has focused on disparities in wages (Oaxaca, 1973; Gunderson, 1989; Brown and Corcoran, 1997; Goldin, 2014; Altonji et al., 2015; Bertrand and Duflo, 2017).

[2] According to one estimate, the UK needs 100,000 new graduates in STEM subjects every year until 2020 just to maintain current employment numbers (`http://www.girlsintostem.co.uk/girlsintostem-1`).

[3] In the US, one in seven engineers are female, and in the UK, women make up only 6% of the engineering workforce. See `http://edition.cnn.com/2014/10/27/world/europe/how-to-get-girls/index.html` and `http://www.girlsintostem.co.uk/girlsintostem-1`.

was tested in 191 countries across the world. We show empirically that the ad was shown to over 20% more men than women. The difference is particularly pronounced for individuals in the age range 25-54 years. It is popular to suppose that such outcomes occur either because those who program the algorithm intend to discriminate or have unconscious biases, or because the algorithm itself will learn to be biased on the basis of the behavioral data that feeds it (O'Neil, 2016). Paralleling these popular assumptions, we explore three potential explanations for our result.

The first class of explanation is that the algorithm learned the apparently discriminatory behavior from actual consumer behavior: If women were less likely to click on the ad, an algorithm trying to maximize click probability might be more likely to show the ad to men than to women. However, we present evidence that if women were shown the ad, they were more likely to click on the ad than men, ruling out this explanation. A similar explanation could be that there were simply fewer women available on the social media platform, for example because they spend less time there than men, meaning they were less likely to see the ads. Again, we present evidence that that is not the case.

The second class of explanations is that the algorithm learned the behavior from other data sources that it was trained on, which in turn might reflect a pattern of discrimination against women in different countries. If that were the case, the ad-serving algorithm might simply reflect differences in underlying gender roles in the culture of the host country and the algorithm could have learned over time to present ads in a way which reflects that bias. We use country-specific data from the World Bank on levels of female education, female labor market participation or general gender inequality to reflect the likely level of institutional bias in that country. We show that these factors were not related to the result that the STEM ad was more likely to be shown to men than women.

The third class of explanations is that the algorithm's decision to display the STEM ad less often to women than to men was a reflection of the economics of ad delivery. In online advertising, multiple advertisers compete to display ads to the same set of eyeballs. This competition means that there can be spillovers from other advertisers' decisions even if they are advertising different products. We present evidence from a separate data collection effort that on average across the world, female 'eyeballs' are more expensive than male eyeballs. We find that the price premium an advertiser has to pay to show ads to women, relative to men, is particularly pronounced for the age group for which we observe the strongest negative effect for the display of the STEM ad. We provide evidence why this may be the case: A marketing literature suggests women largely control household purchases, making them potentially more valuable targets for advertisers. Using data from a separate online retailer, we then document that the higher prices paid by advertisers for female clicks may be profit-maximizing, as, conditional on clicking on the ad, women are more likely than men to purchase.

Last, we explore the generalizability of the finding that an advertising platform is more likely to display STEM ads to men than to women. We implemented a similar advertising campaign for information about STEM careers on three other online advertising platforms. On all platforms, we observed that men received more impressions of the ad than women, implying that our results are characteristic of the advertising ecosystem in general.

Our results suggest that advertiser behavior that is not intended to be discriminatory, such as implementing a campaign that does not discriminate by gender, can nevertheless lead to outcomes where people of one gender are more likely to be exposed to the ad. This occurs because in an advertising ecosystem there are spillovers from one economic actor's valuation of an eyeball to the distribution of ads by another.[4] The spillover across different industry sectors may be especially worrisome if there are societal reasons to care about who sees what kind of communication. For example, society may care about who sees apparently desirable

[4] As such, this paper also adds to research about interactions between different advertisers when bidding for impressions (Athey and Gans, 2010).

advertising which highlights beneficial employment, financial and housing opportunities, or about who sees potentially less desirable advertising, such as ads for predatory lending services.

These insights are important because of optimism among economists (Becker, 2010) that economic forces might limit discrimination. Becker focuses on the area of employment and highlights that firms that do not discriminate would be at a competitive advantage, as they would be able to employ more cheaply members of the groups which were discriminated against (Arrow, 1973; Becker, 1993). However, in that context, the group that was discriminated against was also 'less costly' to employ or engage with.

Our paper, by contrast, examines a case where the group that policymakers may worry about not receiving the same information as men – women – is also more costly to engage with. The key allocation mechanism that dictates the distribution of information is not a measure of the desirability of information dissemination, but instead is the return on investment on advertising across all industry sectors. Advertising allocation decisions by a retail sector selling household products therefore affect communication opportunities and costs in the sector offering job opportunities.

This paper contributes to three literatures.

The first literature is a nascent literature on apparent algorithmic discrimination in advertising. Sweeney (2013) shows that a background check service's ads were more likely to appear in a paid search ad displayed after a search for names that are traditionally associated with African-Americans. Datta et al. (2015) find that women were less likely to see ads for an executive coaching service in India. In general, this literature has focused on documenting empirical patterns consistent with algorithmic discrimination, rather than empirically examining underlying causes of the discriminatory outcomes. For example, Datta et al. (2015) states, 'We cannot determine who caused these findings due to our limited visibility into the ad ecosystem, which includes Google, advertisers, websites, and users.' Sweeney (2013) asks, 'Why is this discrimination occurring? Is this [the background check company's], Google, or society's fault?", but then says that "Answering these questions is beyond the scope of this writing.' Our paper intends to be a first step at uncovering why ad algorithms may lead, here unintentionally, to outcomes which appear to be discriminatory. We believe that answering the question of 'why' is of utmost importance to policy makers who need to think about how best to shape policy.

The second literature is in industrial organization, documenting discriminatory behavior in online markets. Scott Morton et al. (2003) investigate whether the internet leads to less discriminatory behavior in car buying. Edelman and Luca (2014) and Edelman et al. (2017) document racially discriminatory behavior in an online rental market. Relatedly, Pope and Sydnor (2011) find racial discrimination in peer-to-peer lending. Ge et al. (2016) explore discrimination by drivers of peer transportation companies and observe longer waiting times and more cancellations for customers with African-American names; there is also some evidence that drivers took female passengers for longer, more expensive, rides. While such research demonstrates how biases of individuals can lead to discrimination in the digital economy, one view is that when algorithms - not humans - make decisions, such biases should disappear. Our results demonstrate that even when decisions are made by algorithms and human biases are removed, the outcome may still disadvantage one group relative to another.

The third literature we contribute to is a more general discussion in economics about the potential use of algorithms or machine learning techniques to solve policy problems. Mullainathan and Spiess (2017) provide a good overview of this nascent literature, and highlight work such as Kleinberg et al. (2017) who show that using an algorithm to help guide decisions regarding bail, can help relative to a counterfactual where the judge's judgment could be clouded by the time of day or other external factors. Similarly, Cowgill (2017) shows that algorithmically-based hiring decisions may be less 'biased' than human decision making. As a counterpoint to this optimism about automated predictions improving the quality of policy-making, our paper emphasizes that economic forces may distort algorithmic decision-making in unexpected directions.

The research has two separate policy implications. First, our results emphasize the difficulty of regulating algorithms to prevent instances of apparent discrimination. One popular policy prescription has been a focus on algorithmic transparency where algorithmic codes are made public. Such policies are gaining increasing momentum - for example, the Federal Trade Commission (FTC) launched a new unit focused on algorithmic transparency, and in Germany, Chancellor Merkel asked Internet firms to make their algorithms public.[5] Our research suggests, however, that in the empirical context we study, algorithmic transparency would not have helped regulators to foresee uneven outcomes. The reason is that an examination of the algorithmic code would likely have revealed an algorithm focused on minimizing ad costs for advertisers. Without appropriate knowledge about the economic context and how such cost-minimization might affect the distribution of advertising, such 'transparency' would not have been particularly helpful.

Regulators also face the challenge that an apparently discriminatory outcome may not be informative about whether it was the intention of the advertiser or the algorithm to discriminate. Therefore, regulators need to understand potential economic forces before imputing discrimination to the platform or advertiser. We emphasize that our findings do not mean that algorithms may not be biased because of non-economic forces, but instead that economic forces may lead to apparently discriminatory outcomes. Further, any policy prescriptions need to reflect that there may be tradeoffs between the aim of reducing apparent bias and the aim of using economic mechanisms to allocate resources efficiently through algorithms.

[5]See `http://www.pcworld.com/article/2908372/the-ftc-is-worried-about-algorithmic-transparency-and-you-should-be-too.html`, `https://www.ftc.gov/news-events/blogs/techftc/2015/03/booting-new-research-office-ftc` and `https://www.thelocal.de/20161026/merkel-demands-transparency-from-internet-giants`

Second, our results suggest new policy challenges posed by algorithms in areas which are governed by legal protections designed to prevent discrimination. For example, in the US a firm's actions are restricted by federal employment discrimination law. Title VII of the Civil Rights Act of 1964 makes the following distinction with regards to employment discrimination. In recruitment, disparate treatment occurs when a firm treats a potential employee differently on a prohibited basis. Disparate impact occurs if there is a practice which on its face appears neutral and non-discriminatory, but which has a particularly negative impact on a certain group of applicants. Unlike disparate treatment, disparate impact does not require discriminatory intent on the part of the employer.

As of yet, however, the law is not settled about how targeted advertising falls within this employment discrimination framework, and it is not clear whether the law would apply if a firm tried to ensure that it used targeting so that employment opportunities were seen by more members of a protected class (Swire, 2014). Indeed, one implication of the current legal framework is that advertisers may be restricted from taking steps to ensure they can correct for any imbalance in advertising distribution that may result from advertising algorithms.

A superficially attractive solution for advertisers who are concerned about striking a gender balance is to manage two different campaigns that each target men and women separately, rather than relying on an algorithm to ensure an even distribution of impressions. Such an approach would allow an advertiser to ensure that the same number of men and women saw the ad, even if prices differ substantially. However, when we implemented this approach, the ad was automatically not 'approved' by the platform because targeting an employment ad towards only one gender is not in compliance with Federal law.[6] This insight highlights an unexpected tension in the application of Federal anti-discrimination law in a

[6] Facebook's policy is that advertisers can not 'use our [Facebook] audience selection tools to (a) wrongfully target specific groups of people for advertising... or (b) wrongfully exclude specific groups of people from seeing their ad', with particular reference to employment and housing ads. `https://www.facebook.com/policies/ads/prohibited_content/discriminatory_practices`T

digital ecosystem governed by algorithms. If algorithms lead to unbalanced outcomes in the distribution of information because advertisers can target by gender in unregulated sectors such as retail, then attempts to correct for an imbalance in access to information in protected sectors by using targeting methods may be restricted by legal concerns.

The tension we highlight illustrates an evident need for policy guidance in this area. One potential solution is for platforms to offer advertisers the option for a specific campaign of distributing ads equally across specified demographic groups. Such a solution may build on previously suggested policies for platforms on how to protect advertisers from spillovers from the bidding decisions of other advertisers (Ghosh et al., 2009).

2 Field test

For the field test, we partnered with a small website that gives information about careers in the STEM sector. We ran advertising campaigns that directed users who clicked on the ad to this website. We use the term 'field test' rather than 'field experiment' as there was no randomization in ad delivery. Instead, an ad was 'tested' in 191 countries. We use the word 'test' to reflect the fact that there was no strategy underpinning the selection of countries, ad format, or wording of the ad which could provide an alternative explanation of the results.

The field test was for an ad that promoted careers in STEM. The text of the ad was very simple; it said 'Information about STEM careers' accompanied by a picture that represented the different fields in STEM. Figure 1 displays a mock-up of the ad.

The field test was conducted on Facebook, currently the largest social media site in the world. On such social media platforms, advertisers specify the target audience along geography, demographics or interests and bid for display advertising impressions to their target audience by specifying a maximum price they are willing to pay per click. A separate

Figure 1: Sample ad

ad campaign was created with an identical ad for 191 countries spanning the world.[7] We use the cross-national variation later in the paper to explore whether the differences in ad allocation we observe can be ascribed to different economic and cultural conditions regarding the role of women in different nations.

In all cases the ad was targeted at both men and women over the age of 18 years. The only variation across the 191 ad campaigns was the country it was running in. Figure 2 displays the ad targeting settings for a typical ad.

When a user loads a webpage, the ad platform typically conducts an advertising auction in the background that determines which advertiser will show an ad to that user. The outcome of the auction is usually determined by the maximum bid an advertiser places, relative to the bids placed by other advertisers. In addition, the auction accounts for the 'quality score' of an ad. The quality score is the outcome of a predictive method that measures the likelihood a user will click on any particular ad (Athey and Nekipelov, 2010), thus adjusting for the relative merit of a bid for the advertising platform.

[7] According to the United Nations, there are 195 countries. According to the social media platform, there are 213 countries and regions it marks as territories, such as American Samoa or the Channel Islands. The missing countries in our dataset are ones where the social media platform did not reach. For example, North Korea attempts to ensure that its citizens do not browse the broader web, meaning that it is not part of our dataset (`http://www.businessinsider.com/the-six-countries-that-block-social-media-2015-4` Though Turkey is sometimes mentioned as a country that does block social media and has in the past banned Twitter, we were still able to collect advertising data on it.

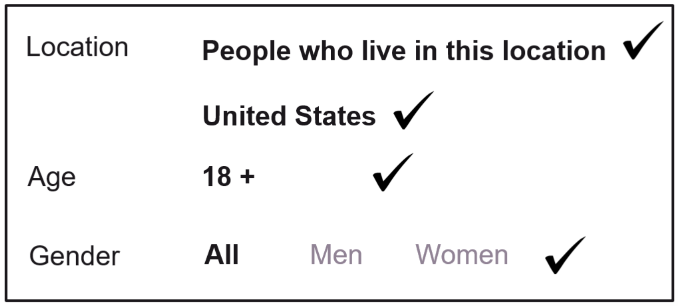

Figure 2: Ad targeting Settings - ad intended to be shown to both men and women aged 18+.

Facebook refers to their quality score as a 'relevance score,'[8] saying, "The more positive interactions we expect an ad to receive, the higher the ad's relevance score will be. (Positive indicators vary depending on the ad's objective, but may include video views, conversions, etc.)" Facebook also mentions that the relevance score "can lower the cost of reaching people. Put simply, the higher an ad's relevance score is, the less it will cost to be delivered." The actual calculation of the quality score and the bids of other advertisers that the advertising auction algorithm uses to allocate advertising is a black box to the advertiser and researcher.

The STEM website initially set a maximum bid per click of $0.20 for all countries. At the end of each day, the STEM website paused campaigns where the ad had been shown to more than 5,000 viewers. The delay in pausing the ad campaign meant that in some cases the ad was shown to up to 24,980 users in a country. If after a week that campaign had not been viewed by 5,000 unique users, the bid was raised to a higher amount that varied by country but was a maximum of $0.60. Bids were raised for 29 countries, or 15% of those in the study. These countries tend to be wealthier ones, such as the UK, the US and Switzerland, which had higher ad prices.

One concern is that the appearance of the ad itself might drive different responses across genders. To investigate this, we tested on Amazon's Mechanical Turk whether the ad ap-

[8] See https://www.facebook.com/business/news/relevance-score

pealed to both men and women. We asked 152 participants from the US (75 male, 77 female) to assume they viewed the ad when browsing the internet and to rate their own likelihood of clicking on the ad on a scale from 1 (very unlikely) to 5 (very likely). We find that the average stated likelihood of clicking on the ad does not differ significantly between men (mean 2.053, standard deviation 1.077) and women (mean 2.105, standard deviation 1.102; p=0.770).[9]

3 Data

For each of the campaigns in each of the 191 different countries, Facebook released extensive data on their performance. Table 1 summarizes the data. Our data is not on the level of individual consumers but groups all variables of interest (impressions, reach, clicks, unique clicks) by country, age and gender group. The age groups that were identified and reported on by the social media platform were 18-24, 25-34, 35-44, 45-54, 55-64 and 65+ years old. Therefore, an observation in Table 1 is at the demographic group (that is age group $\times$ gender)-country level which is the unit of observation we use in our regression analysis. On average, each age group and gender combination were shown 1911 'impressions' of the ad. 'Impressions' refers to the number of times a particular ad was shown. As some individuals saw more than one ad, the reach - which measures how many people saw an ad - was, on average, 616, that is, a campaign for a particular age and gender combination on average reached 616 distinct individuals. On average across all demographic-country groupings, a

[9] The share of participants who rated their probability with 3 or higher is not statistically different either (men: 0.293, women: 0.338, p=0.560). We also asked participants whether they thought the ad was targeted towards a male audience, a female audience or both male and female audiences. Of the participants, 78.29% felt the ad was targeted towards both audiences, 19.08% thought the ad was targeted towards a male audience and 2.63% thought the ad was targeted towards a female audience. The response to this question is likely to reflect inherent biases not about the ad's appearance but about the fact it is focused on STEM. For example, one survey found that '50% of teachers and 34% of parents perceive STEM subjects are more geared towards boys' (see `https://www.accenture.com/t20170905T101544Z__w__/ie-en/_acnmedia/PDF-60/Accenture-Girls-in-STEM-Research-Report-2017-online.pdf`).

	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 1: Summary statistics
Reported at the demographic group-country level.

campaign had 3 clicks and 2.78 unique clicks, indicating that occasionally users clicked more than once on an ad.

As shown in Table 1, the price paid for each click was low relative to other social media campaigns (Tucker, 2014b,a). Figure A1 in the appendix reflects the distribution of costs per click paid by the campaign.

4 Results

4.1 Model-free evidence

The main results of the field test are visible in the raw statistics supplied by the platform. Table 2 summarizes the total number of impressions, clicks and click-rates by demographic group (gender $\times$ age) across all countries in the study.

Table 2: Raw data

Age Group	**Male Impr.**	**Female Impr.**	**Male Clicks**	**Female Clicks**	**Male ClickRate**	**Female ClickRate**
Age 18-24	746719	649590	1156	1171	.0015	.0018
Age 25-34	662996	495996	873	758	.0013	.0015
Age 35-44	412457	283596	501	480	.0012	.0017
Age 45-54	307701	224809	413	414	.0013	.0018
Age 55-64	209608	176454	320	363	.0015	.0021
Age 65+	192317	153470	307	321	.0016	.0021

Reported at the aggregate level by gender $\times$ age group.

There are three obvious patterns in the data. First, men see more impressions of the ad than women. Second, younger women see fewer ads than younger men. Third, on average women are more likely to click on an ad if they see it. Across all campaigns, the average click-rate for men is 0.131 of a percent, and for women it is 0.167 of a percent (p <0.001), slightly higher than some reported in the literature such as Tucker (2014b).

The fact that women are exposed to fewer ad impressions than men is concerning. If women are not exposed to information on STEM careers, they may never apply for STEM jobs (Diekman et al., 2010). We next explore the robustness of these empirical regularities and provide suggestive evidence about why they occur.

4.2 Do men indeed see more STEM ads than women?

Though the empirical regularities may seem obvious in Table 2, we check that they are robust to a standard regression framework which allows us to control for country-specific characteristics.

For demographic group j in country k, the number of times an ad is displayed is modeled as a function of:

$$\begin{aligned} AdDisplay_{jk} = & \\ & + \beta_1 Female_j \\ & + \beta_2 Age_j \\ & + \beta_3 Female_j \times Age_j \\ & + \alpha_k + \epsilon_{jk} \end{aligned} \tag{1}$$

$Female_j$ is an indicator for whether or not the demographic group consisted of women. Age_j is a vector of fixed effects that capture the different age groups of the social media platform's users. We include a vector of country fixed effects α_k to capture variation in the number of impressions due to country size and other country characteristics, such as technological sophistication and social media usage.

Column (1) of Table 3 shows the results of a simple regression with no interactions. It suggests that women were indeed less likely to see the ad. Column (2) reports the full specification laid out in equation (1) and suggests that the disparity in impressions is driven by younger women seeing the ad less often then younger men.

Columns (3)-(4) replicate the results for the number of distinct individuals in a group that saw at least one impression ('reach'). It reflects the fact that in some groups some individuals may have seen more than one ad. Columns (5)-(6) explore the effects of gender on ad frequency, that is, the average number of ads any one individual saw. We find that conditional on seeing an ad, a woman is more likely to see it multiple times. This result suggests that in general our measure of impressions may understate the extent to which women were not shown our ad. Therefore, for the rest of the paper we focus on ad reach, that is the number of unique individuals who saw the ad, as the main dependent measure.

As an additional robustness test, we also check whether the results hold for countries where the maximum bid was kept at $0.20 throughout the study and for countries where the maximum bid was adjusted upwards. We find that the results generally hold, with the caveat that for countries where the bid was adjusted upward, interactions between gender and age are insignificant due to the small sample size.

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

One potential explanation for the fact that women saw the ad fewer times then men is learned behavior on the part of the algorithm. The divergence in impressions could reflect an accurate prediction by the algorithm that women are less likely to click on ads. Such an

Table 3: Women are shown fewer ads than men

	(1) Impressions	(2) Impressions	(3) Reach	(4) Reach	(5) Frequency	(6) Frequency
Female	-479.3***	-209.7***	-228.1***	-98.97***	0.729***	1.276***
	(97.09)	(44.26)	(35.45)	(20.44)	(0.150)	(0.305)
Female × Age18-24		-298.8		-234.3**		-0.523+
		(193.1)		(75.83)		(0.268)
Female × Age25-34		-664.6***		-302.2***		-0.630*
		(154.4)		(48.64)		(0.272)
Female × Age35-44		-464.9***		-159.9***		-0.900***
		(110.5)		(31.26)		(0.246)
Female × Age45-54		-224.2**		-97.25***		-0.903**
		(69.94)		(24.70)		(0.300)
Female × Age55-64		36.16		18.93		-0.326
		(39.58)		(14.33)		(0.412)
Age18-24	2753.6***	2902.6***	909.5***	1026.5***	-0.473*	-0.212
	(248.0)	(284.3)	(108.5)	(131.2)	(0.207)	(0.174)
Age25-34	2132.4***	2464.3***	561.4***	712.3***	-0.683***	-0.369*
	(204.4)	(236.5)	(67.32)	(83.38)	(0.163)	(0.143)
Age35-44	920.5***	1152.6***	197.4***	277.2***	-0.556***	-0.107
	(117.4)	(135.2)	(40.61)	(47.39)	(0.144)	(0.167)
Age45-54	492.4***	604.1***	99.08**	147.5***	-0.471***	-0.0198
	(84.60)	(85.93)	(31.03)	(35.27)	(0.108)	(0.167)
Age55-64	109.0*	90.53+	16.56	6.911	0.0107	0.173
	(51.37)	(52.72)	(18.93)	(19.70)	(0.182)	(0.147)
Country Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2291	2291	2291	2291	2291	2291
R-Squared	0.485	0.488	0.442	0.446	0.776	0.778

Ordinary least squares estimates. Dependent variable as shown. Omitted demographic groups are those aged 65+ and men. Robust standard errors. + $p < 0.1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

explanation for our results seems natural given that ad algorithms use quality scores which aim to reflect how likely a consumer is to click on an ad.

Our data consists of a number of successes (unique clicks) out of a number of trials (impressions) for each demographic segment-day. To pin down the likelihood of clicking across men and women, we first estimate an aggregate logit model using maximum likelihood (Flath and Leonard, 1979). We also estimate an OLS specification with a linear clickthrough rate, since a linear specification provides more straightforward interpretation of interactions.

Let F denote the logistic likelihood function. Due to the aggregate nature of the data the social media platform provides to advertisers that does not include user-level variables, all individuals i in demographic group j in country k have the same vector of x control variables. The likelihood of observing each observation of the sum of positive unique clicks as a function of the sum of reach for that campaign that day is:

$$F(\beta x)^s\{1 - F(\beta x)\}^{r-s} \tag{2}$$

where s is the number of unique clicks and r is the population of social media platform users exposed to the messages.

Table 4 reports the result of our investigation of clicks. Column (1) presents results of a simple specification for clicks as a function of impressions. It suggests that women are more likely to click on the ad. Column (2) repeats the analysis but instead of using impressions it uses reach, which is the number of unique users exposed to a message, as the measure of population. Again, it suggests women are more likely to click on the ad. Columns (3) and (4) show that our results replicate when using as dependent variable a linear clickthrough rate and estimate using an OLS specification. We repeat the analysis with the same age and gender interactions that we used in Table 3. As shown in Columns (5)-(6), these interactions are not significant, indicating that click propensity did not differ by age group and gender. However, we do observe across most columns that younger people are less likely to click. Columns (7)-(8) report the results for an OLS specification and suggest similar (if less precisely estimated) results.

Table 4: If They See The Ad, Women Are More Likely To Click Than Men

	(1) Clicks	(2) Unique Clicks	(3) Click Rate	(4) Reach Rate	(5) Clicks	(6) Unique Clicks	(7) Click Rate	(8) Reach Rate
Female	0.221***	0.303***	0.0362***	0.280***	0.264**	0.399***	0.0425+	0.366*
	(0.0271)	(0.0290)	(0.00713)	(0.0599)	(0.0932)	(0.0875)	(0.0233)	(0.177)
Female × Age18-24					-0.137	-0.166+	-0.0156	-0.107
					(0.0975)	(0.0956)	(0.0265)	(0.164)
Female × Age25-34					-0.0899	-0.135	-0.0254	-0.223
					(0.113)	(0.109)	(0.0283)	(0.209)
Female × Age35-44					0.0822	-0.0289	-0.0136	-0.244
					(0.113)	(0.109)	(0.0273)	(0.196)
Female × Age45-54					0.0633	0.000689	-0.00486	-0.180
					(0.119)	(0.117)	(0.0288)	(0.178)
Female × Age55-64					0.0465	-0.0573	0.0221	0.238
					(0.136)	(0.129)	(0.0308)	(0.221)
Age18-24	-0.175**	-0.214***	-0.0216	-0.117	-0.105	-0.129+	-0.0138	-0.0637
	(0.0576)	(0.0557)	(0.0139)	(0.0825)	(0.0731)	(0.0704)	(0.0152)	(0.0585)
Age25-34	-0.375***	-0.460***	-0.0500***	-0.271**	-0.332***	-0.394***	-0.0373*	-0.160*
	(0.0593)	(0.0572)	(0.0127)	(0.0850)	(0.0823)	(0.0785)	(0.0180)	(0.0680)
Age35-44	-0.341***	-0.409***	-0.0493***	-0.189*	-0.379***	-0.392***	-0.0425*	-0.0668
	(0.0712)	(0.0657)	(0.0133)	(0.0904)	(0.0902)	(0.0839)	(0.0174)	(0.112)
Age45-54	-0.190**	-0.222***	-0.0288*	-0.166+	-0.220*	-0.220**	-0.0264+	-0.0764
	(0.0613)	(0.0605)	(0.0123)	(0.0865)	(0.0865)	(0.0843)	(0.0158)	(0.0680)
Age55-64	-0.0186	-0.0199	-0.00190	0.149	-0.0426	0.00913	-0.0129	0.0296
	(0.0682)	(0.0666)	(0.0149)	(0.0912)	(0.0955)	(0.0879)	(0.0167)	(0.0863)
Country Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4515014	1453890	2291	2291	4515014	1453890	2291	2291
Log-Likelihood	-52298.6	-40388.3	1055.8	-3356.6	-52291.8	-40384.6	1058.5	-3349.3
R-Squared			0.173	0.314			0.175	0.318

Aggregate logit estimates in Columns (1)-(2) and (5)-(6). Ordinary least squares estimates in Columns (3)-(4) and (7)-(8). In Columns (2), (4), (6) and (8) the population variable is ad reach. In Columns (1), (3), (5) and (7) the population variable is ad impressions. The dependent variable is whether someone who was exposed to an ad clicked. Omitted demographic groups are those aged 65+ and men. Robust standard errors. + $p < 0.1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Though there is no evidence that Facebook has implemented such a policy, or that such a policy would even maximize revenue for Facebook, one possible explanation of our results could be that the algorithm aims to obtain the same number of clicks from each demographic group and that they showed the ad to fewer women in the expectation they would click more. We investigated whether there was such evidence of balancing across age groups. We found no evidence of it; indeed, the number of clicks was quite uneven across the different age groups in our data.

5.1 Do our results simply reflect the fact that women use Facebook less?

The ad algorithm could also echo that women spend less time on social media so have less opportunity to be shown the ad. If women spent less time on social media, it may simply be harder for advertisers to reach them. However, both Facebook's internal data and industry reports suggest that women are more likely to use social media platforms then men and also that they are more active on the sites, and consequently are more likely to be exposed to ads.

In the US, 54% of Facebook users are women, 46% are men. Intensity of usage by women also appears higher. Data from 2018 suggests that on average in the previous month, men had liked 18 posts and made 10 comments, while women liked 36 posts and made 29 comments. In the average month, men had liked 19 posts and made 19 comments, whereas women had liked 35 posts and made 24 comments.[10] On Facebook in general across the world, 59% of active users are women and 41% are men, though in terms of profiles 44% of Facebook profiles are for women and 56% are men.

Facebook's internal data is supported by industry reports and survey evidence. ComScore reports for Latin America, Europe, North America and Asia-Pacific that in each case the

[10] This data was derived from `https://www.facebook.com/ads/audience-insights`

average engagement with social media is higher by women (Shaw, 2012). Similarly, Vermeren (2015) reports that 76% of women and 66% of men use Facebook, while women have more than twice as many posts on their Facebook walls and have 8% more 'friends' than men.

Last, while the campaign we study was not targeted towards individuals who had indicated an interest in science or engineering, it is possible that Facebook viewed such individuals as more likely to click on the ad. If that were the case, we would want to ascertain that Facebook offers a sufficiently large pool of women who are interested in science or engineering and to whom the focal ad could be shown. We use a feature that allows an advertiser on Facebook to obtain data on the number of individuals in a specific target group.[11] We find that in the US, a campaign targeting an interest in science or engineering has a potential reach of 26 million men and 33 million women, suggesting that a shortage of women with an interest for the subject matter is unlikely to be the cause of fewer impressions being displayed to women than to men.

6 Do the results reflect cultural prejudice or labor market conditions that the algorithm has learned?

Another potential explanation for our results is that the underlying ad algorithm has learned the preferences of the host country and knows that in a particular country it is undesirable to show a specific type of ad, or employment ads in general, to women. Bias could result either if the algorithm was trained on a training dataset that reflected such bias, or if it had learned such bias in earlier campaigns run by different advertisers. In this case, the relative lack of impressions shown to women could simply reflect the fact that in most countries, women's labor market rights and careers lag behind men's.

To explore that possibility, we augmented our advertising data with data from the World

[11] See `https://www.facebook.com/ads/audience-insights`.

Bank pertaining to the status of women and the female labor force. We used some of the indicators from 'The Gender Data Portal,' which is the 'World Bank Group's comprehensive source for the latest sex-disaggregated data and gender statistics covering demography, education, health, access to economic opportunities, public life and decision-making, and agency.'[12] Much of the data is missing for many of the countries, so we focus on four measures where there is the most data available: The extent of female labor participation; the extent of female primary and secondary education; and an index constructed by the World Bank to capture a variety of measures of female equality (CPIA), assessing the extent to which the country has created institutions and programs to enforce equal access for men and women in education, health, the economy, and protection under law. A higher index implies more equality. These measures should all reflect the extent to which women in that country are likely to be able to obtain access to careers in the Science, Technology, Engineering and Math fields. Information on female labor market participation was available for 80 countries, information on female primary and secondary education was available for 90 countries and the female equality index was available for 42 countries. As with much World Bank data, the last year the data was collected varied by country, but in all cases it was reasonably recent.

Table 5 displays the results of this investigation. In each case, we estimate how the number of women who saw an ad campaign in a country was moderated by whether or not that country scored above the median by that measure of gender equality. If according to the World Bank, the data was missing for a country, it was treated as not being above the median. Column (1) demonstrates that the interaction of female labor market participation with how many women were reached is insignificant and we still find that the ads reach significantly fewer women than men. Columns (2) and (3) confirm the results when we instead interact with female primary and secondary education. Column (4) demonstrates the effect likewise holds when accounting for whether a country ranks high on the female equality index. The

[12] `http://data.worldbank.org/`

general lack of significance suggests that the particular cultural prejudices of the country towards women or women's social situation in the country cannot explain the fact that more ads are being shown to men than women. Table A1 in the appendix shows that these results (or at least the general lack of measured significant effects) hold for impressions as well.

In the final column of Table 5, we explore whether the phenomenon we observe is specific to poorer countries where, potentially, women have less access to careers in STEM and there might be inherent bias in how the social media platform allocates impressions across gender. Alternatively, it may be possible that our results are driven by the way that the ad algorithm allocates advertising to consumers in richer countries. We again interact the female indicator by whether the GDP of the country was above the median GDP in our data.[13] The results indicate that women are less likely to be exposed to the ads independently of whether they live in a poorer or a richer country.

[13] Again, we treat countries and territories for which GDP was not available as not having above median GDP.

Table 5: Women being exposed to fewer ads than men is not driven entirely by underlying gender disparity in labor market conditions in that country

	(1) Reach	(2) Reach	(3) Reach	(4) Reach	(5) Reach
Female	-208.5***	-183.0***	-249.8***	-225.3***	-237.8***
	(48.58)	(28.75)	(46.65)	(39.16)	(47.91)
Female × High % Female Labor Part=1	-59.40				
	(64.31)				
Female × High % Female Primary=1		-139.0			
		(94.51)			
Female × High % Female secondary=1			69.07		
			(66.96)		
Female × High Female Equality Index (CPIA)=1				-20.82	
				(87.22)	
Female × High GDP=1					32.22
					(60.94)
Age18-24	909.5***	909.5***	909.6***	909.5***	909.5***
	(108.5)	(108.5)	(108.5)	(108.5)	(108.5)
Age25-34	561.3***	561.3***	561.4***	561.3***	561.4***
	(67.34)	(67.33)	(67.34)	(67.34)	(67.34)
Age35-44	197.4***	197.4***	197.5***	197.4***	197.4***
	(40.62)	(40.61)	(40.62)	(40.62)	(40.62)
Age45-54	99.05**	99.01**	99.11**	99.07**	99.09**
	(31.04)	(31.02)	(31.04)	(31.04)	(31.04)
Age55-64	16.53	16.49	16.59	16.55	16.57
	(18.93)	(18.92)	(18.94)	(18.93)	(18.93)
Country Controls	Yes	Yes	Yes	Yes	Yes
Observations	2291	2291	2291	2291	2291
Log-Likelihood	-18053.6	-18051.1	-18053.4	-18054.1	-18054.0
R-Squared	0.442	0.443	0.442	0.442	0.442

Ordinary least squares estimates. Dependent variable is whether someone is exposed to an ad. Omitted demographic groups are those aged 65+ and men. Robust standard errors. + $p < 0.1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7 Do our results reflect competitive spillovers?

We now explore whether competitive spillovers and pricing pressure for certain demographic groups may explain our results.

In digital advertising markets, advertisers usually bid to pay a maximum price per click. Across all campaigns, the average price that was paid for a click was \$0.12086 for men and \$0.12087 for women. The similarity of these average prices provides little suggestion that the prices of advertising to different demographic groups caused our results. However, recall that the advertiser set the same maximum bid across men and women. As such, it is possible that there was a significantly larger share of individual auctions for female eyeballs (than for male eyeballs) that the advertiser did not win which could explain why prices paid are nearly identical.

We collected further data on how the social media platform advised advertisers to set their bids for each of the demographic groups in each of the countries we had targeted in order to have a good chance at 'winning' the advertising auction to show an ad. This suggested pricing data for different demographic groups was available to advertisers on the Facebook advertising platform.[14] The differences in prices across demographic groups would not necessarily be obvious to advertisers as Facebook only displays to them an average suggested bid across all consumers they intend to advertise to, not differences across (sub)segments of a target group. The advertiser would need to set out explicitly to collect more granular data.

Table 6 presents summary statistics for the average bid suggested by the social media platform as well as the minimum and maximum of the suggested bid range. It reports the unconstrained amount that the social media platform recommends that an advertiser should

[14]This data was available for any advertiser to view and collect at the time of our field test in 2016. Facebook has since changed the way it reports data on targeting and prices to advertisers so that this precise data can no longer be collected as easily.

	Mean	Std Dev	Min	Max
Avg Suggested Bid	0.45	0.66	0.010	15.7
Min Suggested Bid	0.19	0.31	0.010	4
Max Suggested Bid	0.77	1.32	0.017	43
Female	0.50	0.50	0	1

Table 6: Summary statistics

pay to reach a certain demographic group. Similar recommended bid data has been used in previous scholarship such as Goldfarb and Tucker (2011). Though such data has the disadvantage that researchers have no information about the precise 'black box' that is used to calculate the values, this is less of a concern in our study, as we are using it simply to proxy for the likely competitive bidding environment for a particular gender-age group within a country, rather than trying to precisely interpret the economic implications of a price.

The data on recommended bids also deviates from our original data in terms of the age cohorts we analyze. In general, to avoid the restrictions on advertising to children inherent under COPPA and other privacy regulations designed to protect children, the field test ad was not shown to anyone claiming to be under the age of 18. However, we were able to collect pricing data on this group and use them as a baseline for the analysis. Furthermore, because in some countries there were too few 65+-aged people in the data for us to be able to get separate estimates, we combine the 55-64 and 65+ cohorts in the analysis and use as our suggested bid the average across the two age groups.

7.1 Analysis of secondary pricing data

We estimate the relationship between demographic groups and suggested bidding prices. Columns (1) and (2) in Table 7 show that on average the platform suggests that advertisers bid about 5 cents more to advertise to women. In terms of age, those in the 25-44 year old age group are also more expensive to advertise to though this is not precisely estimated. Column (3) explores how the pattern changes when we include interactions between gender and age. It shows that women between 25 and 44 are more expensive to advertise to.

One interpretation of the results is that, rather than an ad algorithm itself discriminating actively against women, the fact that other advertisers prize the 'eyeballs' of young women, means that any ad algorithm designed to allocate advertising impressions in a cost-effective manner will not display ads that are intended to be gender-neutral in a gender-neutral manner, but instead will favor cheaper - male - eyeballs.

Table 7: In general, women are more expensive to advertise to on social media: Competitive spillovers from other advertisers' decisions may explain our finding

	(1) Avg Suggested Bid	(2) Avg Suggested Bid	(3) Avg Suggested Bid
Female	0.0534*	0.0525*	-0.0464
	(0.0238)	(0.0247)	(0.0378)
Female × Age18-24			0.0648+
			(0.0376)
Female × Age25-34			0.174+
			(0.0935)
Female × Age35-44			0.150***
			(0.0429)
Female × Age45-54			0.0751
			(0.0544)
Female × Age55+			0.129**
			(0.0445)
Age18-24	-0.0100	-0.0100	-0.0421
	(0.0270)	(0.0282)	(0.0405)
Age25-34	0.0762	0.0763	-0.0105
	(0.0497)	(0.0519)	(0.0406)
Age35-44	0.0740*	0.0740*	-0.000557
	(0.0348)	(0.0364)	(0.0444)
Age45-54	0.0597	0.0589	0.0216
	(0.0389)	(0.0405)	(0.0557)
Age55+	0.0211	0.0198	-0.0446
	(0.0333)	(0.0347)	(0.0435)
Country Controls	No	Yes	Yes
Observations	2096	2096	2096
Log-Likelihood	-2096.5	-1219.8	-1215.0
R-Squared	0.00443	0.569	0.571

Ordinary least squares estimates. Dependent variable is average suggested bid. Omitted demographic groups are those aged between 13-17 and those of the male gender. Robust standard errors. + $p < 0.1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

8 Why are women such a prized demographic?

The next question is why women are such a prized demographic in non-employment sectors that this crowding out occurs. We first turn to broader findings that suggest that women, and especially younger women, are a highly prized demographic for advertisers. The business press reports that it is precisely the demographic of 25-34-year-old women which should be most prized by online advertisers, both because they are likely to engage with advertising and because they traditionally control household expenses.[15] More broadly, in the US, out of \$5.9 trillion in consumer spending, women control \$4.3 trillion (Silverstein and Sayre, 2009).

To further investigate the question we use completely separate data from a large retailer that sold a broad range of fashionable physical consumer products. Examples include a skateboard deck, a toothbrush holder, a picture frame or a coat rack. Though many of the products offered were gender-neutral, we have no data about the share of men and women in the retailer's customer base or the inherent appeal of these items to either gender.

The retailer used social media advertising to try and generate demand for its household goods which highlighted the discounts offered that day. It set up its advertising campaigns so that each campaign targeted a specific demographic by age and gender: Either men or women of a particular age group. We focus our analysis on instances where there was at least one campaign that was identical in terms of product, behavioral targeting and wording across men and women.

The data is on the campaign level and includes information on the number of impressions per campaign as well as the number of clicks and whether, upon arrival on the website, consumers added products to their shopping carts.[16]

[15] `http://www.businessinsider.com/young-women-are-most-valuable-mobile-ad-demographic-2012-2`

[16] Due to complications of tracking consumers across the security features demanded by the separate payment system within the retailer's webpage, we use whether a consumer added a product to their shopping cart as a proxy for conversion rather than the actual purchase. Unlike our earlier data, the data the retailer provides us is focused on the US.

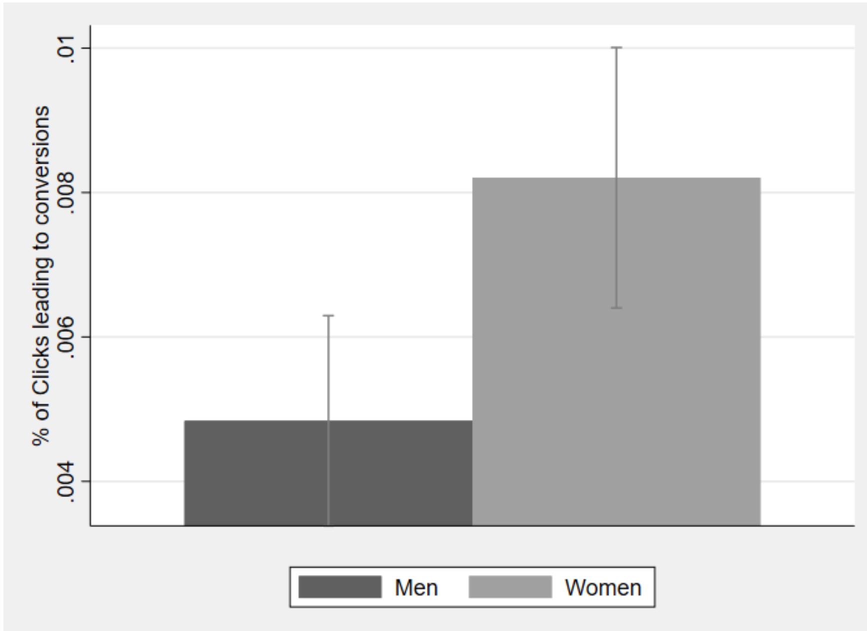

Figure 3:

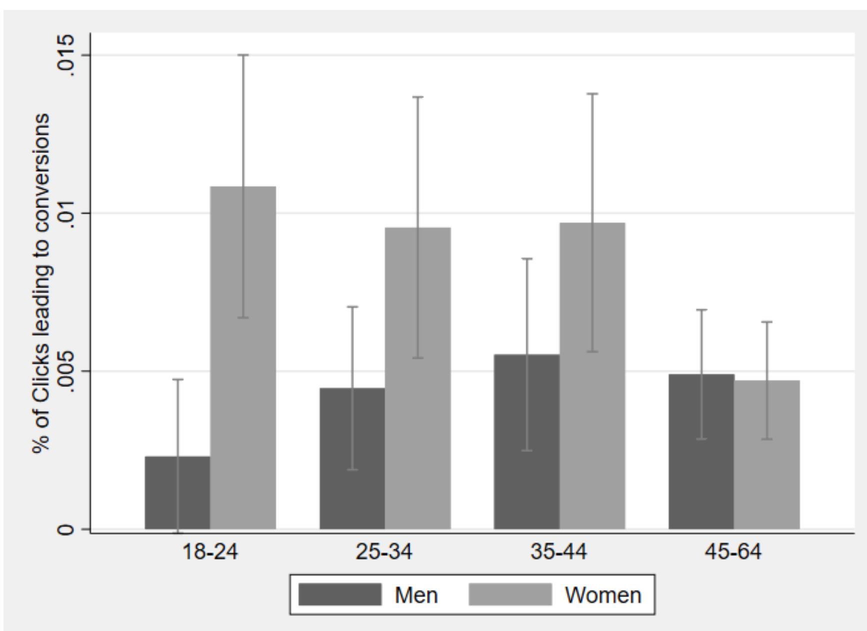

Figure 4:

Figure 3 emphasizes that for this retailer women are relatively more likely than men to convert after clicking. This is important as it emphasizes that the value, or ROI, of attracting women to click on an ad may be higher than that of attracting men - given that a click does not necessarily lead to a sale. Figure 4 displays this distribution across different age groups. The general pattern is relatively stable across different age groups, though the divergence is largest for younger men and women, the demographic groups for which we observe a divergence in impressions in our field test. Even if on average women were more interested in purchasing fashionable consumer items then men, conditioning on clicks means that we are focusing on the subset of men and women who show some interest in purchasing such products.

Though the data comes from only one advertiser for a variety of consumer items, it suggests that in this context conditional on clicking an ad, women are more likely to convert, a pattern that could explain why advertisers treat younger women as such a highly prized demographic and potentially are willing to pay more to show ads to them.

9 Do our insights generalize to other platforms?

Our field test was conducted on Facebook, which is a large social media site but not the only platform for online advertising. Therefore, we explored whether the pattern that we observed replicated on other platforms.[17]

We attempted to replicate the results of our field test on the Google display ad network, Twitter and Instagram.[18] We focused all these tests on the US market. The reason is that unlike Facebook, not all other platforms have the same reach internationally as they have in the US. Our intention was to see whether the result that we observed on Facebook replicated

[17]We thank our NBER discussant Ben Edelman for this very helpful suggestion.

[18]We were not able to conduct the test on LinkedIn, as LinkedIn does not break down ad performance by gender (or age), though it does allow targeting to these demographics.

in other settings.

9.1 Google AdWords

We ran a similar ad to that in Figure 1 on the Google Display Network, Google's network for distributing display ads across different websites.[19] This ad platform forced us to choose targeting criteria for showing the display ad, so we used keywords such as 'science jobs' and 'engineering careers' which were suggested by Google based on the website's content. Again, we did not restrict the bidding by gender and targeted all age groups above 18 years. We used a manual bid strategy where we bid 50 cents per click. We spent $181 for the campaign.

Table 8: 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

Table 8 displays summary statistics for the campaign. The data show a pattern that is reasonably similar to results of the field test on Facebook displayed earlier in Table 2. We find that 36% of impressions are displayed to women and 51% of impressions are displayed to men. A further category where the gender was unknown accounted for 13% of ad impressions. Consistent again with our earlier results, if they saw the ad women were far more likely than males to click on it. Women were slightly more expensive to advertise to, despite having far higher clickthrough rates, which in theory should exert downwards pricing pressure.

9.2 Instagram

We then replicated the campaign on Instagram. Instagram is owned by Facebook but is maintained consciously as a different social network. It does share a similar advertising

[19] The ad we use here had a slightly different STEM image to comply with Google's image resolution requirements.

platform, however. We again targeted all adults over 18 years of age and did not discriminate by gender. We set the budget to $100 and again allowed the algorithm to optimize the amount we bid per click. Table 9 reports the results.

On Instagram, only 15% of impressions were shown to women. However, Instagram is the one platform that we tested where men were more likely to click on the ad than women. It is possible that the disparity in click rate by gender may have exacerbated the algorithm's allocation of the ad. The fact that women were far more expensive to show ads to then men is consistent with such an interpretation.

Table 9: 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

9.3 Twitter

Last, we attempted to replicate the results on Twitter. On Twitter an advertiser has the option of posting a promoted tweet. That is, the advertiser instructs Twitter to show an advertising message in the form of a tweet to users (for details on advertising on Twitter see, Lambrecht et al. (2017)). We instructed Twitter to post a promoted tweet that said 'Find out more about STEM careers [url].' The tweet was purely textual and lacked an image. We bid a maximum price of $1.00 per engagement.[20] We spent $100 total on the campaign.

Table 10 reports the result of the field test on Twitter by gender. We were not able to obtain cost per click or click rate estimates by gender from the Twitter interface, as Twitter simply reports total spend for each gender group. However, the result obtained for the number of impressions displayed to women and to men echoes that of Table 2 in that, again, women were less likely to see the ad.

[20] On Twitter, bids are per 'engagement' which subsumes clicks, retweets and favorites of a promoted tweet.

Table 10: Results of test on Twitter

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

10 Implications

We use data from a field test of an ad on social media for STEM jobs that was explicitly intended to be gender-neutral in its delivery. Women were far less likely to be shown the ad than men - but not because they were less likely to click on it. If women ever saw the ad, they were more likely than men to click. The likelihood of showing ads to men rather than women also does not reflect World Bank measures of the culture of sexism within the country or the country's overall wealth.

Instead, we present suggestive evidence that the gender-imbalance reflects the fact that women are a prized demographic and as a consequence are more expensive to show ads to. This means that an ad algorithm which simply optimizes ad delivery to be cost-effective, can deliver ads which are intended to be gender-neutral in what appears to be a discriminatory way. Our finding suggests a nuanced view of the potential for apparent discriminatory outcomes even from 'neutral' algorithms.

Our findings also suggest multiple public policy challenges. In theory, one possible solution to the problem we identified would be for managers to simply run and manage campaigns and campaign budgets separately by gender. By actively managing separate campaigns, managers could ensure a balance in the distribution of ad impressions by gender. To validate this proposed solution, we attempted to run gender-specific campaigns for the same STEM website on Facebook as in the earlier field test. However, when the same creative used earlier was targeted by gender, the ad was not 'approved'. Figure 5 provides a screenshot of the ad that was not approved. Facebook did not approve these ads as they do not allow advertisers to exclude users of either gender when running an employment-related ad. The platform's website that explains why such ads are not approved emphasizes the need for advertisers to comply with Federal law regarding employment discrimination.[21]

The finding is important because it suggests a tension between algorithms, the use of targeting tools and the potential for discrimination that policy makers need to consider. Though it may seem a reasonable policy to prevent firms from using targeting techniques that can target or exclude certain demographic groups in areas such as employment to prevent discrimination, this kind of restriction also prevents firms from using targeting to try and correct any imbalances that the use of an algorithm may lead to. As algorithms become increasingly important in the distribution of digital content, it seems important for policy makers to clarify and ensure that regulation allows firms to use digital data and techniques to try and rectify imbalances that may be caused by algorithms.

The other more general policy implication that our research highlights is that some policy approaches which are currently being proposed or implemented to regulate algorithms online to prevent discrimination may not be fully effective. For example, advocates of an algorithmic transparency approach argue that by making the code of algorithms available for public scrutiny, policy makers may be able to prevent and identify instances of bias. However, in our setting such a policy would not have been effective because all that public scrutiny of the algorithm would have revealed is an algorithm that was trying to achieve the apparently reasonable aim of cost-minimization on behalf of advertisers.

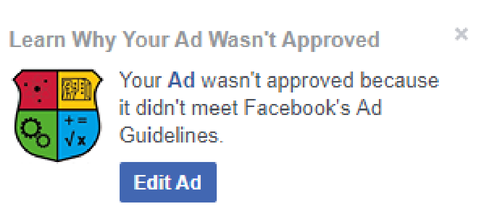

Figure 5: Facebook did not approve an employment-related ad targeted at a single gender

[21]See https://www.facebook.com/policies/ads/prohibited_content/discriminatory_practices for a description of the policy.

For managers, our paper emphasizes the difficulties that using new digital techniques such as targeting and automated algorithmic ad distribution can pose, especially in a context like employment, where traditionally discrimination has been a concern. There is no reason to assume that algorithms will lead to a balanced distribution of ads towards a protected group. However, it is not currently possible to use targeting techniques to try and correct for an algorithmically generated imbalance in order to better reach such a protected group. One potential solution to this is issue is for platforms themselves to offer advertisers the facility to equalize automatically how impressions are distributed across demographic groups.

There are of course limitations to our study. First, our field test consists of a single ad for STEM careers shown across multiple countries. Though it seems likely that our result would replicate across different ad designs and messages, we do not have data to test the possibility. Second, because we do not observe the workings of the actual ad algorithm, our result regarding the role of bidding decisions of other advertisers is suggestive rather than conclusive. Third, though our results suggest that the interaction between different economic actors can play an important role in leading to apparent discriminatory outcomes, we are unable to shed light on the extent to which algorithms themselves may be biased. Fourth, our results are descriptive and focus on explaining an empirical regularity. We are unable to test policy measures which may prevent the kind of outcomes we observe. Notwithstanding these limitations, we believe that our paper makes a useful contribution in that it documents not only an occasion when apparent 'algorithmic bias' may occur but also that it may occur even if there is no deliberate intent.

References

Altonji, J. G., P. Arcidiacono, and A. Maurel (2015). The analysis of field choice in college and graduate school: Determinants and wage effects. Technical report, National Bureau of Economic Research.

Arrow, K. (1973). The theory of discrimination. *Discrimination in labor markets 3*(10), 3–33.

Athey, S. and J. S. Gans (2010). The impact of targeting technology on advertising markets and media competition. *American Economic Review, Papers and Proceedings 100*(2).

Athey, S. and D. Nekipelov (2010). A structural model of sponsored search advertising auctions. In *Sixth ad auctions workshop*, Volume 15.

Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy 101*(3), 385–409.

Becker, G. S. (2010). *The Economics of Discrimination*. University of Chicago Press.

Bertrand, M. and E. Duflo (2017). Field experiments on discrimination. *Handbook of Economic Field Experiments 1*, 309–393.

Brown, C. and M. Corcoran (1997). Sex-based differences in school content and the male-female wage gap. *Journal of Labor Economics 15*(3), 431–465.

Cheryan, S., J. O. Siy, M. Vichayapai, B. J. Drury, and S. Kim (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science 2*(6), 656–664.

Cowgill, B. (2017). Automating judgement and decision-making: Theory and evidence from résumé screening.

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

Diekman, A. B., E. R. Brown, A. M. Johnston, and E. K. Clark (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science 21*(8), 1051–1057.

Edelman, B., M. Luca, and D. Svirsky (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics 9*(2), 1–22.

Edelman, B. G. and M. Luca (2014). Digital discrimination: The case of Airbnb.com.

Flath, D. and E. W. Leonard (1979). A comparison of two logit models in the analysis of qualitative marketing data. *Journal of Marketing Research 16*(4), pp. 533–538.

Ge, Y., C. R. Knittel, D. MacKenzie, and S. Zoepf (2016). Racial and gender discrimination in transportation network companies. Technical report, National Bureau of Economic Research.

Ghosh, A., R. P. McAfee, K. Papineni, and S. Vassilvitskii (2009). Bidding for representative allocations for display advertising. In *WINE*, Volume 5929, pp. 208–219. Springer.

Goldfarb, A. and C. Tucker (2011). Search engine advertising: Channel substitution when pricing ads to context. *Management Science 57*(3), 458–470.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review 104*(4), 1091–1119.

Gunderson, M. (1989). Male-female wage differentials and policy responses. *Journal of Economic Literature 27*(1), 46–72.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2017, February). Human decisions and machine predictions. Working Paper 23180, National Bureau of Economic Research.

Lambrecht, A., C. Tucker, and C. Wiertz (2017). Should you target early trend propagators? Evidence from Twitter. *Marketing Science, forthcoming*.

Mullainathan, S. and J. Spiess (2017, May). Machine learning: An applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 693–709.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group (NY).

Pope, D. G. and J. R. Sydnor (2011). What's in a picture? Evidence of discrimination from prosper.com. *Journal of Human Resources 46*(1), 53–92.

Scott Morton, F., F. Zettelmeyer, and J. Silva-Risso (2003). Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities? *Quantitative Marketing and Economics 1*(1), 65–92.

Shapiro, J. R. and A. M. Williams (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles 66*(3-4), 175–183.

Shaw, M. (2012, Feb). The state of social media. Presentation to Online Marketing Institute, London.

Silverstein, M. J. and K. Sayre (2009). The female economy. *Harvard Business Review 87*(9), 46–53.

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

Swire, P. (2014). Lessons from fair lending law for fair marketing and big data. *White Paper*.

Tucker, C. (2014a). Social Advertising: Social Advertising: How Advertising that Explicitly Promotes Social Influence Can Backfire. Mimeo, MIT.

Tucker, C. (2014b, October). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research 51*(5), 546–562.

Vermeren, I. (2015). Men vs. women: Who is more active on social media? *Brand Watch*.

Williams, W. M. and S. J. Ceci (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences 112*(17), 5360–5365.

Table A1: Women being shown fewer ad impressions than men is not driven by underlying gender disparity in labor market conditions in that country

	(1) Impressions	(2) Impressions	(3) Impressions	(4) Impressions	(5) Impressions
Female	-329.3**	-436.4***	-606.5***	-472.7***	-499.5***
	(123.6)	(102.2)	(125.6)	(103.9)	(131.8)
Female × High % Female Labor Part=1	-454.4*				
	(190.9)				
Female × High % Female Primary=1		-132.1			
		(233.5)			
Female × High % Female secondary=1			404.7*		
			(184.8)		
Female × High Female Equality Index (CPIA)=1				-48.32	
				(290.8)	
Female × High GDP=1					67.86
					(165.1)
Age18-24	2753.4***	2753.6***	2753.8***	2753.6***	2753.6***
	(248.1)	(248.1)	(248.1)	(248.1)	(248.1)
Age25-34	2132.2***	2132.3***	2132.6***	2132.4***	2132.4***
	(204.5)	(204.5)	(204.5)	(204.5)	(204.5)
Age35-44	920.3***	920.4***	920.7***	920.5***	920.5***
	(117.4)	(117.4)	(117.4)	(117.4)	(117.4)
Age45-54	492.2***	492.3***	492.5***	492.4***	492.4***
	(84.61)	(84.62)	(84.64)	(84.62)	(84.61)
Age55-64	108.8*	108.9*	109.2*	109.0*	109.0*
	(51.36)	(51.36)	(51.39)	(51.37)	(51.37)
Country Controls	Yes	Yes	Yes	Yes	Yes
Observations	2291	2291	2291	2291	2291
Log-Likelihood	-20250.1	-20254.4	-20251.2	-20254.8	-20254.7
R-Squared	0.488	0.486	0.487	0.485	0.486

Ordinary least squares estimates. Dependent variable is whether someone sees an ad impression. Omitted demographic groups are those aged 65+ and men. Robust standard errors. + $p < 0.1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

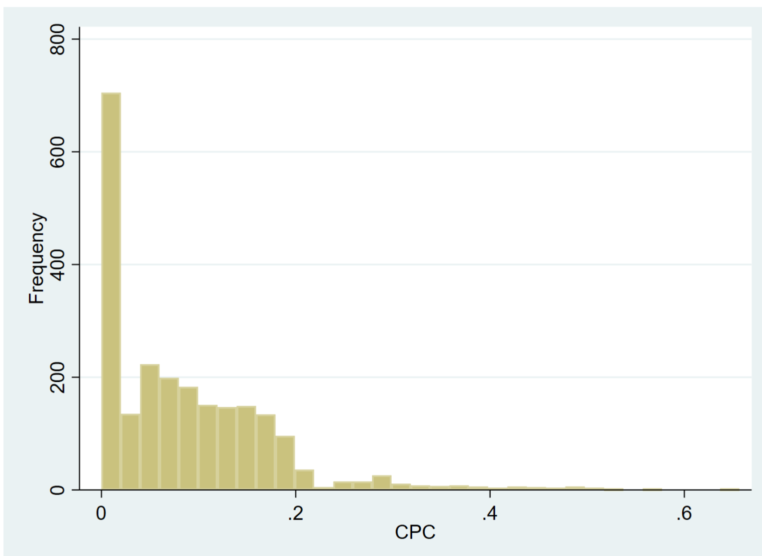

Figure A1: Histogram of average cost per country