STEM and teens: An algorithm bias on a social media

Grazia Cecere[*,†] Clara Jean[*,‡] Fabrice Le Guel[*] Matthieu Manant[*]

July 10, 2018

Abstract

We study whether online platforms might reproduce offline stereotypes of girls in the STEM disciplines, and if this bias can be reduced. For this purpose, we estimate ad distribution on a popular social media platform via a field experiment by setting up a randomized online ad campaign on behalf of a French computer science school. The ad campaign targeted students in high schools in France. The treatment aims to estimate whether a message aimed at prompting girls is displayed to girls more than to boys. The article contributes to work that aims to shed light on the possible biases generated by algorithms. Our results show that on average, girls are less likely to see the ad than boys.The treatment ad which was aimed to be shown to more girls had a crowding-out effect, since overall, it was displayed less to both boys and girls.

Keywords: Gender-gap, discrimination, algorithm bias, STEM education.

JEL Codes: J16, I24.

[*]RITM, University of Paris-Sud, 54, bd Desgranges - 92330 Sceaux, France

[†]Télécom École de Management, Institut Mines-Télécom, 9, rue Charles Fourier - 91000 Évry, France

[‡]EPITECH, 24, rue Pasteur - 94270 Le Kremlin-Bicêtre, France

We are grateful to Ananya Sen, Maya Bacache, Anna D'Annunzio, Babur De los Santos, Nicolas Jacquemet, Amadeo Piolatto, and Emmanuel Vespa for their helpful comments and suggestions. We give special thanks to the association Pascaline, and to the school Efrei Paris-Sud for their support. An earlier version of this paper was presented at the 2017 Boston MIT Conference on Digital Experimentation (CODE) Poster Session, 2018 Research seminar Università di Torino, 2018 TSE Biennal Postal Conference, 2018 10^{th} ICT Paris Conference. Financial supports from MSH Paris-Saclay under grant 16-OT-6 and DAPCODS/IOTics ANR 2016 project (ANR-16-CE25-0015). All errors are the sole responsibility of the authors.

1 Introduction

While the presence of women in higher education has increased over the years in the OECD countries, the number of women enrolled in Science, Technology, Engineering, and Mathematics (STEM) education programs continues to be significantly lower than the number of men (OECD, 2017). An explanation for this societal trend might be that girls' education choices follow stereotypes (Bordalo *et al.*, 2016), or social norms (Akerlof and Kranton, 2010). Although, in a context of digitized information which allows easier access to information for all, we would expect a better matching between girls and education programs, and particularly, an increasing participation of girls in STEM education. However, the use by data-driven algorithms of individuals' personal information (such as gender, age, and education, and users' interactions) can lead to unanticipated correlations, and can reproduce offline discrimination and stereotypes (Tucker, 2017) with offset the potential benefits of digitized information. Here we attempt to study the behavior of such algorithms.

We rely on the previous results by Lambrecht and Tucker (2018) who ran an ad campaign on STEM jobs, and show apparent discriminatory outcomes to the detriment of women. We set up a field experiment using a popular social media to study whether an ad related to STEM studies prompting girls is equally distributed to girls and boys. We ran an online ad campaign that targeted high school students in relation to STEM education, on a popular social media platform, on behalf of a post-secondary French computer science school. On the one hand the design of the experiment allows us to estimate the counterfactual distribution of the ad by analyzing the effect of a treatment ad, and on the other hand to match the ad performance data with administrative data at the high school level, which enables us to study characteristics of the high schools potentially learned by the algorithm. The ad campaign ran for a two-week period, and targeted French high school students aged between 16 and 19 years with accounts on the social network platform, making no distinction between girls and boys. The computer science school's goal was to attract high school students. The ad provided information on job market access and average first wage after graduating from this school. We wanted first to study the ad display of the algorithm at the high school level according to gender. Second, we examined how "girl content" affected the ad display it produced.

We conducted a field experiment that included 101 simultaneous ad campaigns, one for each high school targeted. In order to test our hypothesis on the effect of "girl" content on the distribution of the ad by the platform algorithm, high schools were randomly assigned to the treatment. The ad images displayed were the same for both groups but the heading descriptions differed slightly: the treatment group received the ad with headings that included girl content, whilst the headings on the ad displayed to the control group made no reference to gender. Our results suggest that on average, girls received 24 fewer impressions than boys but girls aged between 18 and 19 were more likely to click on the ad if they come across it. The difference in the number of impressions received between girls and boys cannot be attributed to cost of the ad since in our field experiment girls and boys were prized equally. The treatment has a crowding-out effect with the ad overall, being less shown to all the students. Our results are robust to the control group and several specifications.

Our experimental design allows us to investigate different issues. First, we measure whether girls in high schools received more impressions than boys. Second, we estimate how including a message that prompted girls, modifies the number of impressions made by the algorithm. Third, we exploit the variation over the number of girls enrolled in the main tracks (Science, Economics and Social Science, Literature) to assess the difference in the number of impressions received by girls enrolled in these tracks. Fourth, we measure the interest of teens in the ad by gender based on whether or not they clicked on it. Finally, we highlight implications for both policy and managers by measuring whether an ad campaign with a general interest is biased by data-driven algorithms.

We contribute to three literature strands. First, we complement the literature on algorithmic biases by estimating casual relationships between ad distribution and the characteristics of targeted individuals. On the one hand, algorithms can improve ad effectiveness and reduce the cost of prediction, making decision algorithms more valuable (Agrawal *et al.*, 2016). On the other hand, they can reproduce offline stereotypes based on race (Angwin *et al.*, 2016) and gender (Lambrecht and Tucker, 2018).[1]

Several studies investigate biases produced by algorithms on online platforms. For example, Sweeney (2013) highlights that algorithms with machine learning capability might express unintentional biases linked to individual sociodemographic data. Based on Google searches, she observed that compared to white-identifying names, black-identifying names received more displays of an ad for criminal records services. Following this study, others have confirmed the existence of such biases on ad algorithm (see e.g. Datta *et al.*, 2014). O'Neil (2016) argues that algorithms may generate biases because they are trained with biased data. In particular, algorithms can reproduce apparent discrimination or stereotypes observed and learned from individual data. Lambrecht and Tucker (2018) explore and try to explain these biases. They ran a country level field test on a social media platform with a gender-neutral ad for STEM jobs. Their results suggest apparent discriminatory outcomes to the detriment of women. They underline two possible causes for this bias. First, women are a prized demographic, suggesting more expensive "eyeballs". Second, there are spillovers associated to targeting by other advertisers, which might incite the algorithm to reproduce gender biases against women. We rely on this work to conduct our experiment, which provides slightly different findings.

From yet another point of view, a growing strand of this literature is aimed at showing the effectiveness of algorithms, and how supervised algorithms can potentially help reduce biases and discrimination. Kleinberg *et al.* (2018) studied the use of a calibrated and trained algorithm in the context of legal court decisions. They found that machine learning helped to

[1] The literature describes these processes as 'feature engineering'. See Datta *et al.* (2014) and OECD (2017) for more information.

reduce criminality, and generated fairer decisions in the case of Afro-Americans and Hispanics compared to human decisions. The challenge is how to set up an experiment which identifies counterfactual evidence of correcting a given bias (Cowgill and Tucker, 2017).

Second, a large number of approaches have been developed to study the gender gap in STEM education (e.g. Fryer and Levitt, 2010; Croson and Gneezy, 2009; Guiso *et al.*, 2008). We are interested in how social media might indirectly influence the participation of girls in STEM education. A large literature in economics attempts to explain the gender gap issue in STEM education (Fryer and Levitt, 2010) and how it can be reduced. The literature examines also how the gender gap in math shapes individual education and career choices, and the impact on labor market integration (Chetty *et al.*, 2011; Leibbrandt and List, 2015). While teens are spending increasing amounts of time on social media,[2] research on its mediating role to promote information related to STEM education is limited. Our paper contributes to this body of work by shedding light on the implications of using intelligent algorithms to achieve this, and specifically, on the potential effect and behavior of data-driven algorithms in relation to an online gender gap.

Finally, we investigate the relative importance of personal data for improving the matching between individuals and services. The literature on the economics of privacy highlights its importance for Internet companies in terms of exploiting personal data to target consumers. Individuals disclose their personal information in order to obtain immediate access to services, or to get in touch with peers (Acquisti *et al.*, 2016). Our understanding of the potential future spillovers from data disclose at given time is limited. These data can be used by algorithms to target particular groups of consumers, and could generate unexpected negative spillovers. Since data is likely to persist in digital identities, data disclosed at one point in time can be unexpectedly used in the future (Tucker, 2017). Goldfarb and Tucker (2011) show that limiting firms' data collection capabilities can induce negative effects resulting in

[2]Statistics are available for the US (Pew Research Center, 2016) and for Europe (Livingstone *et al.*, 2011; Eurostat, 2015).

less effective advertising campaigns. Our paper contributes by highlighting that the personal information disclosed can influence how an algorithm categorizes the individual and controls the type of information he or she has access to.

Our paper differ from Lambrecht and Tucker (2018)'s in several ways. First, we prompt social media's algorithm by using a randomization process to check whether a minor manipulation of the ad content addressed to girls (rather than boys) affects the behavior of the ad algorithm. Second, we focus on teenagers – specifically high school pupils aged between 16 and 19. Third, we run the ad campaigns at the high school level in order to match experimental data with administrative data. Fourth, in order to reach the largest number of individuals, the cost paid for the experiment was based on the cost per thousand impressions (CPM) rather than the cost per click. **Finally, if Lambrecht and Tucker (2018) highlight a market bias for STEM ads online explained by more expensive"eyeballs" according to women, we emphasize a real algorithm bias according to such ads for girls on social media.**

Policymakers and education institutions strive to spark girls interest in STEM education and jobs, and dissemination to teenagers of information on STEM is an important issue. More generally, policymakers need a better understanding of the extent to which data-driven algorithms reproduce offline social bias.

The remainder of the paper is organized as follows. Section 2 describes the data set. Section 3 presents empirical evidences on the results of the field experiment. Section 4 presents the econometric estimations. Section 5 concludes the paper.

2 Research design

We want to understand how an ad message prompting girls affects the ad's distribution among teens. We designed a field experiment with simultaneous ad campaigns distinguishing between a neutral ad addressed to the control group and an ad with "girl content" addressed to the treatment group. We hypothesize that the algorithm takes account of the "girl content" and displays the ad more to girls than boys. In this section, we describe the French education system, our experimental design, and the data used.

2.1 Context: French high schools

The aim of the computer science school was to encourage the enrollment of new students after high school, and especially girls who are under-represented in STEM disciplines. High school education in the French system lasts for three years.[3]

There are three main types of high schools: general, vocational, and polyvalent high schools. Within the general high school (and polyvalent high school), there are three main tracks: Literature (L), Economics and Social sciences (ES), and Science (S).[4] Vocational high school provides essentially non-academic syllabus, teens are specialized in manual or clerical jobs. To access post-secondary education, 12th-grade students are required to sign up on a government platform to state their education preferences. In 2017, students were required to declare their option on the official government platform by March 21^{st}. We ran our field experiment from March 11^{th} to March 26^{th} 2017.

2.2 Experimental design

The design of our experiment consisted of targeting French high school students aged between 16 and 19. Our field experiment includes a random sample of high schools with a social network page. We targeted both girls and boys without distinguishing between genders. We ran 101 simultaneous ad campaigns – one for each high school targeted – over a two-week

[3] In France, the education system comprises three stages. Most children attend high school (15 to 17 years) for three years after spending five years in primary school (age 6 to 10), and four years in middle school (11 to 14 years). Education is compulsory up to age 16.

[4] Other available tracks include Management science and technologies (STMG), Biotechnologies and physics in the laboratory (STL), and Music and dance techniques track (TMD).

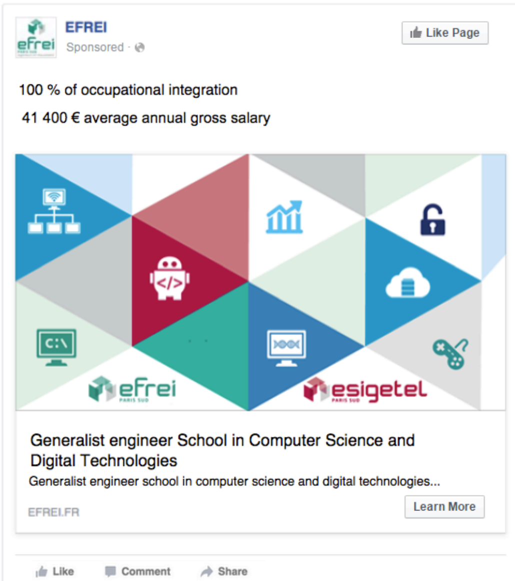
period, on a social network platform. Our main interest was in the behavior of the social media ad algorithm.

We conducted a randomized campaign offering two nearly identical ads. For this purpose, we created two groups of high schools, treatment and control, with similar observable characteristics. They were shown the same picture but with slightly different texts. The text presented to the treatment group was: "100% of occupational integration. €41 400 average annual gross salary for women" (Fig. 2).[5] The salary proposed is the real wage offered to graduates from this institution. The ad addressed to the control group displayed the average salary without reference to gender; the ad had the slightly different heading: "100% of occupational integration. €41 400 average annual gross salary" (Fig. 1). We had two objectives. First, we wanted to study whether the social media algorithm distributed the ad equally to girls and to boys. Second, we wanted to test the effect of two slightly different messages by random separation of the high schools into two groups.[6]

We set a daily budget of €2 for each high school, the bid auction was fixed at the CPM. The aim was to reach the maximum number of students, and we were able to see whether mass distribution of an ad would target girls and boys equally.

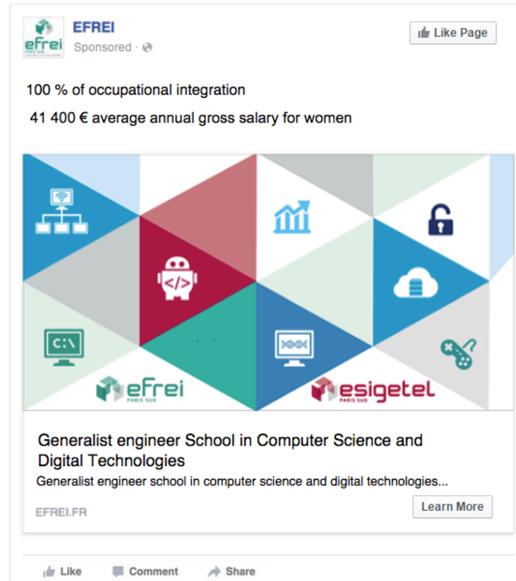
[5]The original French ads are available on request.

[6]In 2016, the difference in average yearly annual gross salary between men and women after the graduation from an engineer school was about 1800 euros. See https://start.lesechos.fr/rejoindre-une-entreprise/actu-recrutement/pour-les-femmes-ingenieures-les-inegalites-debutent-des-le-diplome-8738.php



Note: Translation from French of the neutral ad we presented to high school students in the control group

Figure 1: Control ad



Note: Translation from French of the ad with "girl content" that was presented to high school students in the treatment group

Figure 2: Treatment ad

2.3 High school data and randomization procedure

The design of our field experiment allows us to match ad performance (see Section 3 for data) with publicly available administrative data. We collected high school administrative data related to the academic year 2015-2016 from Etalab, a French national project that provides open data in France.[7] Table 1 summarizes the descriptive statistics.

For each high school, we have information on the overall proportion of girls (*Proportion of girls overall*), and on the proportion of girls enrolled in each main track. While the proportion of girls enrolled in Literature track is 63.4%, the proportion of girls enrolled in Science track is 43.6% compared to boys. *Graduation rate* is the average proportion for all tracks of students who obtain the degree. In our sample, the graduation rate of students who passed their final exam is 91.6%, which is in line with national average.[8] The set of variables

[7]Data are available upon request from Etalab. For more information on Etalab data, see `https://www.data.gouv.fr/fr/search/?q=lycees`, last retrieved January 12, 2018.

[8]In 2016, the graduation rate for the general education was 91.5%, see http://www.education.gouv.fr/cid126752/resultats-definitifs-de-la-session-2016-du-baccalaureat-stabilite-de-la-reussite-dans-les-voies-generale-et-technologique-progression-dans-la-voie-professionnelle.html

Educational stages distribution measures the proportion of students enrolled in each grade. Enrollment per high school in *10th-grade* (resp. 11th-grade and 12th-grade) is the average proportion of *10th-grade* (resp. 11th-grade and 12th-grade) students enrolled in each high school. The set of variables associated with *Overall tracks distribution* presents the overall proportion of students in each main track. *Public school* is a dummy variable indicating whether the high school is public. There are 79.2% of public schools. *Paris school* indicates if the school is located in Paris. *General high school* indicates if a school offers general education courses, *Vocatioqnal high school* indicates if a school offers only technical courses. *Polyvalent high school* takes value 1 if a high school offers both general and technical courses. *Average enrollment* measures the average number of students enrolled in a high school. The average enrollment in each high school is 869 students.

Before launching our ad campaigns, we randomly attributed high schools with a social media account to either the control, or the treatment group. Table 2 presents the pre-treatment statistics. We estimate the average baseline characteristics for high schools in the treatment and control groups. To test for balanced groups, we compute the equality of the means for each characteristic for each covariate. The last column in the table presents the p-values. There is no mean difference between the two groups. All observable characteristics proved to be balanced between the control and the treatment groups at the conventional levels.

Table 1: Descriptive statistics for high schools

Variable	Mean	Std. Dev.	Min.	Max.	N
Proportion of girls overall	0.514	(0.137)	0	0.729	101
Science track	0.436	(0.152)	0	0.750	101
Economics and Social Science track	0.530	(0.203)	0	0.828	101
Literature track	0.634	(0.319)	0	0.950	101
Other tracks	0.074	(0.079)	0	0.271	101
Graduation rate	0.916	(0.090)	0.420	1	101
Science track	0.848	(0.254)	0	1	101
Economics and Social Science track	0.819	(0.301)	0	1	101
Literature track	0.756	(0.376)	0	1	101
Other tracks	0.140	(0.123)	0	0.445	101
Educational stages distribution					
12th-grade	0.321	(0.022)	0.261	0.395	101
11th-grade	0.335	(0.028)	0.243	0.468	101
10th-grade	0.344	(0.039)	0.222	0.448	101
Overall tracks distribution					
Science track	0.411	(0.182)	0	0.838	101
Economics and Social Science track	0.239	(0.114)	0	0.480	101
Literature track	0.108	(0.081)	0	0.508	101
Other tracks	0.098	(0.105)	0	0.475	101
Proportion of household income					
High household income	0.467	(0.241)	0	0.960	101
Middle high household income	0.111	(0.052)	0	0.224	101
Middle low household income	0.202	(0.086)	0	0.395	101
Low household income	0.177	(0.143)	0	0.625	100
Public school†	0.792	(0.408)	0	1	101
Paris school‡	0.327	(0.471)	0	1	101
General high school	0.743	(0.439)	0	1	101
Vocational high school	0.030	(0.171)	0	1	101
Polyvalent high school$^\sharp$	0.228	(0.421)	0	1	101
Average enrollment§	868.802	(372.010)	189	2391	101

Notes: This table reports overall mean estimates for high schools in our sample. Graduation rate is the average proportion for all tracks of students who obtain the degree. Enrollment per high school in 10th-grade (resp. 11th-grade and 12th-grade) is the average proportion of 10th-grade (resp. 11th-grade and 12th-grade) students enrolled in each high school. Standard deviations are in parentheses.

† Public school takes value 1 if the high school is public, 0 otherwise

‡ Paris school takes value 1 if the high school is located in Paris, 0 otherwise

$^\sharp$ High schools which provide both vocational and general education, takes the value 1 if the high school is doing both, 0 otherwise

§ Average number of students enrolled in a high school

3 Results

In this section, we describe the raw data related to our experiment. We depict ad campaigns performance, and the effect of the treatment on ad displays.

Table 2: Pre-treatment summary statistics for high schools in our sample

Variable	**Control** Mean	Std. dev.	N	**Treatment** Mean	Std. dev.	N	**p-value**
Proportion of girls	0.507	(0.112)	52	0.522	(0.159)	49	0.573
Science track	0.447	(0.111)	52	0.424	(0.187)	49	0.465
Economics and Social Science track	0.561	(0.162)	52	0.497	(0.237)	49	0.117
Literature track	0.606	(0.328)	52	0.663	(0.310)	49	0.376
Other track	0.072	(0.080)	52	0.076	(0.079)	49	0.778
Graduation rate	0.916	(0.075)	52	0.915	(0.104)	49	0.954
Science track	0.892	(0.162)	52	0.803	(0.319)	49	0.078
Economics and Social Science track	0.870	(0.227)	52	0.766	(0.359)	49	0.083
Literature track	0.743	(0.392)	52	0.771	(0.362)	49	0.713
Other	0.148	(0.122)	52	0.132	(0.125)	49	0.527
Educational stages distribution							
12th-grade	0.321	(0.020)	52	0.322	(0.025)	49	0.742
11th-grade	0.334	(0.033)	52	0.336	(0.022)	49	0.632
10th-grade	0.346	(0.039)	52	0.342	(0.040)	49	0.598
Tracks distribution							
Science track	0.439	(0.161)	52	0.383	(0.200)	49	0.123
Economics and Social Science track	0.247	(0.103)	52	0.230	(0.124)	49	0.459
Literature track	0.100	(0.071)	52	0.116	(0.091)	49	0.329
Other	0.106	(0.111)	52	0.090	(0.100)	49	0.453
Proportion of household income							
High household income	0.462	(0.240)	52	0.472	(0.244)	49	0.831
Middle high household income	0.117	(0.054)	52	0.103	(0.049)	49	0.187
Middle low household income	0.206	(0.088)	52	0.198	(0.086)	49	0.623
Low household income	0.184	(0.133)	51	0.171	(0.153)	49	0.653
Public school	0.827	(0.382)	52	0.755	(0.434)	49	0.379
Paris school	0.327	(0.474)	52	0.327	(0.474)	49	0.997
General high school	0.692	(0.065)	52	0.796	(0.058)	49	0.238
Vocational high school	0.019	(0.019)	52	0.041	(0.029)	49	0.528
Polyvalent high school	0.288	(0.063)	52	0.163	(0.053)	49	0.136
Average enrollment	897.827	(403.081)	52	838.000	(337.35)	49	0.422

Notes: This table reports mean estimates for high schools in our sample for both the treatment and control groups. Standard deviations are in parentheses.

3.1 Ad campaigns performance

The social media platform provided us with detailed information on the ad campaigns. In our data set an observation is at the high school level, and is by gender and by age category 16-17 and 18-19. Hence, our total sample includes 5,333 observations. From these data we derive five main measures for ad campaign performance: *Impressions*, *Ad clicks*, *CPM*, *Reach*, and *Frequency*. We focus mainly on the number of impressions which is the number of times an ad is displayed to an age group, and to a gender group in a given high school.

Similarly, *Ad clicks* is a dummy variable that indicates whether there was at least one click on an ad by an individual. In the marketing literature, clicks are used to proxy for individuals' interests. *CPM* indicates the amount we spent for a thousand impressions per high school and per day.[9] *Reach* refers to the number of distinct individuals who saw the ad. Since a single individual could have seen the ad several times, *Frequency* measures the average number of times an individual saw the ad.

During the ad campaigns, the algorithm displayed a total of 1,226,929 impressions, and reached 31,302 different individuals, that is an average of 227 impressions for each high school.[10] In relation to individual reactions to the ads, 24.7% of high schools recorded at least one click on the ads. This low click rate is in line with that recorded for similar ad campaigns on social media described in the literature.[11]

3.2 Is ad display different for girls?

We investigate whether there is a difference overall between girls and boys in the display of ads. Table 3 presents the statistics of overall ad performances by gender. Girls were likely to receive fewer impressions (212.7) than boys (241.2), and benefited from fewer ad repetitions (4.026) than boys (4.185). The mean differences are statistically significant $p < 0.001$ for both t-tests.[12] It might be assumed that the difference in ad costs between boys and girls explains this difference.[13] However, the statistical evidence suggests the difference between

[9]While we set a daily budget of €2 per day for each high school, we actually paid €1.564 per day on average as the ad system of the social network platform is based on a second price auction.

[10]The number of individuals reached is obtained from the feedback given by the social media ad manager.

[11]Chatterjee *et al.* (2003) and Lambrecht and Tucker (2018) also found a low individual click probability.

[12]According to Kolmogorov-Smirnov test, there is a significant difference in distributions of the number of impressions between girls and boys ($p = 0.000$). Similarly, the Wilcoxon rank-sum test rejects the null hypothesis of equal distributions and central tendency ($p = 0.000$).

[13]Lambrecht and Tucker (2018) underline that advertisers pay higher prices for females, which can justify the difference in the number of impressions between men and women in their study. In our field experiment, we targeted teens. Data from the digital advertisement manager Adespresso, owned by Hootsuite, show that although boys under 17 are more likely priced more than girls, costs per clicks (CPC) for women over 24 are higher than for men. See `https://adespresso.com/blog/instagram-ads-cost`, last retrieved March 11, 2018.

girls' CPM (€0.395) and boys' CPM (€0.389) is not significant ($p = 0.618$). Thus, the difference in the number of impressions received by girls and boys may be attributable not to a difference in cost but rather to a negative stereotyping of girls in relation to STEM education.

Raisonnement sur le reach

Table 3: Ad display according to gender (boys vs girls)

	Overall		**Boys**		**Girls**		**p-value**
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	
Impressions	226.899	(238.093)	241.247	(245.479)	212.739	(229.740)	0.000
Ad clicks	0.247	(0.431)	0.254	(0.435)	0.239	(0.427)	0.208
Reach	47.191	(37.335)	47.262	(37.573)	47.121	(37.106)	0.890
Frequency	4.418	(4.251)	4.815	(5.312)	4.026	(2.783)	0.000
CUM overall	0.003	(0.002)	0.003	(0.002)	0.003	(0.002)	0.000
CUM 16-17	0.003	(0.002)	0.003	(0.002)	0.004	(0.002)	0.002
CUM 18-19	0.003	(0.002)	0.003	(0.002)	0.003	(0.002)	0.000
Sample size	5,333		2,684		2,649		

Notes: This table reports the means of the regressors and highlights the differences between boys and girls. The last column presents the p-values. Standard deviations in parentheses.

To complement, the last two rows in Table 3 present the differences in cost (CPM) by age category. Across all 101 campaigns, the average *CPM 16-17* was €0.228 for boys aged between 16 and 17, and €0.225 for the same age category of girls. Similarly, the average *CPM 18-19* (students aged between 18 and 19) was €0.558 for boys and €0.538 for girls. The mean differences are not statistically different at conventional $p < 0.05$ t-tests.[14]

3.3 Is ad display different for the treatment group?

The design of the field experiment allows us to test the effect of a text with "girl content" on ad display. Table 4 presents the mean differences in the ad performance variables between the control and treatment groups. The last column reports the p-values of the mean differences between these two groups. All mean differences are statistically significant with overall lower

[14]This result confirms Lambrecht and Tucker (2018).

means for high schools in the treatment group, which suggests that prompting the algorithm towards girls resulted in reducing the ad display for both girls and boys.[15]

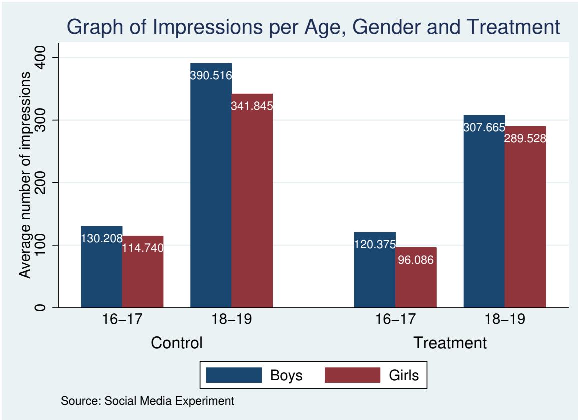
Table 4: Ad display according to group (control vs treatment)

	Control		**Treatment**		**p-value**
	Mean	Std. dev.	Mean	Std. dev.	
Impressions	245.702	(253.268)	205.301	(217.423)	0.000
Ad clicks	0.261	(0.439)	0.230	(0.421)	0.009
Reach	50.302	(37.729)	43.619	(36.560)	0.000
Frequency	4.251	(2.207)	4.610	(5.759)	0.002
CUM	0.003	(0.002)	0.003	(0.002)	0.000
CUM 16-17	0.003	(0.002)	0.004	(0.002)	0.003
CUM 18-19	0.003	(0.002)	0.003	(0.002)	0.020
Sample size	2,851		2,482		

Notes: This table highlights the differences between the control and the treatment group according to ad performance. The last column presents the p-values. Standard deviations in parentheses.

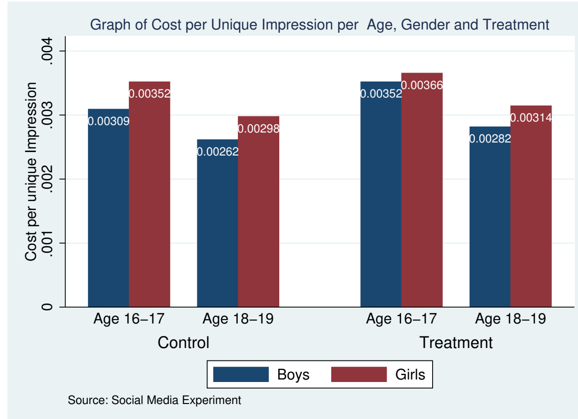
A graphical depiction of the distribution of the number of impressions between girls and boys, and treatment group is informative. Figure 3 presents the average number of impressions by gender and age group across the control and treatment groups. It shows that the number of impressions displayed is always lower for girls than for boys regardless of age or treatment group. We observe a tendency for this gap to decrease for the treatment group although this group seems also to experience a crowding-out effect with fewer impressions for all age groups and genders. Figure 4 depicts the CPM by gender and age group across the treatment and control groups. On average, overall, boys are slightly more costly than girls, with girls aged between 18 and 19 costing €0.057 more than boys of the same age only in the treatment group.

[15] According to Kolmogorov-Smirnov test, there is a significant difference in distributions of the number of impressions between control and treatment group ($p = 0.000$). Similarly, the Wilcoxon rank-sum test rejects the null hypothesis of equal distributions and central tendency ($p = 0.000$).



Notes: The figure presents differences in the number of impressions by age group and gender. Control group statistics are shown on the left side of the horizontal axis, and those for the treatment group are on right side of the horizontal axis. Standard deviations are: (102.770), (96.654), (305.877), (284.226) for the control group and (121.197), (81.556), (251.271), (256.709) for the treatment group.

Figure 3: Bar graph of impressions



Notes: The figure depicts the differences in Cost per Unique Impressions by age group and gender. The left side of the horizontal axis shows the average CPM value for the control group, and the right side shows the average CPM value for the treatment group. Standard deviations are: (0.141), (0.130), (0.259), (0.278) for the control group and (0.172), (0.131), (0.284), (0.325) for the treatment group.

Figure 4: Bar graph of CUM

4 Econometric approach

We explore the effect of treatment on the ad performance, and then a range of potential explanations for the difference in the number of impressions between girls and boys. First, we estimate the number of impressions focusing on the ad performance data. Second, we augment our ad data with administrative data at the high school level. Then, we test whether individuals' ad clicks did influence the algorithm in displaying ads more to girls than to boys.

4.1 Main result: Did we prompt the algorithm?

We are interested in the determinants of ad display, that is the effect on the number of impressions of being a girl and of the treatment. To this end, we estimate an OLS regression on the pooled dataset of high schools, and model the number of impressions for a demographic group (gender and age group) i and a high school j at time t:

$$\begin{aligned} Impressions_{ijt} =& \beta_0 + \beta_1 Girls\ SN_i + \beta_2 Age_i 18\text{-}19 + \beta_3 Treatment_j + \beta_4 (Girls\ SN_i \times Age_i 18\text{-}19) \\ &+ \beta_5 (Girls\ SN_i \times Treatment_j) + \alpha_j + \lambda_t + \epsilon_{ijt}, \end{aligned} \tag{1}$$

where the variable $Girls\,SN$ indicates whether or not the demographic group includes girls, Age18-19 is a set of dummy variables indicating the age of high school teens within the group, and $Treatment$ is a dummy variable that measures whether the high school belongs to the treatment group. We include also two interaction terms. $Girls\ SN \times Age$ 18-19 estimates the joint effect of being in the group of girls and being aged between 18 and 19. $Girls\,SN \times Treatment$ estimates the joint effect of being a girl and receiving the treatment ad. We include a vector of high school fixed effects α_j to capture heterogeneity among high schools, and add a vector of time fixed effects λ_t to account for variations due to a different time period (14 days). ϵ_{ijt} is the error term.

Table 5 presents the main estimations of the number of impressions. All regressions include day fixed effects. Columns (1), (2), and (4) include high school fixed effects, and columns (3), (5), (6), and (7) include the dummy variable $Treatment$ without high school fixed effects in order to avoid collinearity.

Column (1) shows that girls are likely to receive 26 fewer impressions than boys. More important for our purposes, column (3) shows that the treatment group of teens is likely to have seen 41 fewer impressions than individuals in the control group. This suggests the presence of a crowding-out effect due to the treatment – overall, fewer impressions were shown to high schools in the treatment group. Column (6) reports the full equation presented in equation (1). It includes the interaction term $Girls\,SN \times Treatment$ which is not statically significant, i.e. the treatment ad was not shown more to girls. This result suggests that unless the "girl content" message directly aims at prompting the algorithm in favor of girls, the ad algorithm did not consider the message as oriented towards girls, but more as less appropriate to all subjects. Column (7) adds the interaction term $Girls\,SN \times Treatment \times Age$ 18-19, and shows that high schools in the treatment group with a large majority of girls between 18-19 received more impressions.

Column (2) suggests that teens aged between 18 and 19 years received more impressions

Table 5: Main result - Effect of prompting the algorithm

	Impressions						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girls SN	-25.591***	-23.984***	-26.704***	-13.319***	-32.175***	-24.262***	-15.060***
	(4.768)	(3.583)	(5.688)	(4.472)	(8.204)	(6.544)	(5.303)
Age 18-19		226.253***	219.280***	236.802***	219.290***	227.165***	260.839***
		(3.617)	(5.628)	(5.176)	(5.629)	(8.215)	(11.901)
Treatment			-41.259***		-47.177***	-47.228***	-10.113
			(5.653)		(8.270)	(8.276)	(6.316)
Girls SN × Age 18-19				-20.938***		-15.643	-33.611**
				(7.125)		(11.247)	(16.221)
Girls SN × Treatment					11.756	11.830	-8.280
					(11.308)	(11.315)	(7.971)
Age 18-19 × Treatment							-72.547***
							(16.228)
Girls SN × Age 18-19 × Treatment							38.873*
							(22.301)
High school fixed effects	Yes	Yes	No	Yes	No	No	No
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	438.223***	325.692***	219.588***	320.360***	222.345***	218.379***	201.221***
	(15.532)	(14.782)	(13.883)	(14.761)	(14.273)	(13.963)	(13.459)
Observations	5,333	5,333	5,333	5,333	5,333	5,333	5,333
R-squared	0.480	0.705	0.242	0.705	0.242	0.243	0.246

Notes: OLS estimations. The dependent variable is number of *Impressions*. Columns (1), (2), (4) include high school fixed effects. Robust standard errors are reported in parentheses. Columns (3), (5), (6), (7) do not include high school fixed effects since they are collinear with the dummy variable of treatment. All the regressions include day fixed effects. Significance at 1%; 5% and 10% levels indicated respectively by ***,** and *.

(227) than those aged between 16 and 17. Column (4) shows that although overall teens between 18 and 19 are likely to see more impressions than teens between 16 and 17, the interaction term *Girls SN* × *Age* 18-19 is negative and statistically significant, suggesting that girls between 18 and 19 saw 21 fewer impressions than boys.

4.2 Did the algorithm match with administrative data?

In this part, we investigate whether the algorithm learned about social media users' gender. To this end, we augment our social media data with high-school administrative data to estimate an effect of the distribution of girls and boys in high schools. The key rationale behind this approach is that we assume comparable distributions for social media data and for administrative data. We use the same model as in the previous part, and add a vector of high school characteristics. These data include the proportion of girls enrolled in each track:

Science track, Literature track, Economics and Social Science track, and Other tracks .We add also in the regressions the set of high school characteristics (such as Public School, Paris School, General high school, Vocational high school and Polyvalent high school).

Table 6: Effects of prompting the algorithm on science oriented high schools

	Impressions		**Ad clicks**	
	(1)	(2)	(3)	(4)
Girls SN	-16.912***	-17.070***	-0.040**	-0.039**
	(4.078)	(4.096)	(0.018)	(0.018)
Age 18-19	229.046***	228.873***	0.131***	0.131***
	(7.769)	(7.756)	(0.016)	(0.016)
Treatment	-48.796***	-6.189	-0.029**	-0.151***
	(5.621)	(17.547)	(0.012)	(0.049)
Girls in Science track	-68.547***	2.840	-0.228***	-0.427***
	(21.943)	(38.700)	(0.055)	(0.095)
Girls in Economics and Social Science track	-5.499	8.728	0.074	0.032
	(24.652)	(26.749)	(0.046)	(0.049)
Girls in Literature track	18.445	9.366	0.021	0.047
	(15.820)	(17.233)	(0.027)	(0.029)
Girls in Other tracks	-489.071***	-506.446***	-0.244***	-0.198**
	(35.552)	(36.980)	(0.084)	(0.085)
Girls SN × Age 18-19	-17.423*	-17.268	0.045*	0.045*
	(10.587)	(10.581)	(0.023)	(0.023)
Girls in Science track × Treatment		-93.839**		0.266***
		(39.440)		(0.103)
Constant	294.377***	261.200***		
	(17.556)	(21.797)		
High school characteristics	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
R-squared	0.332	0.332		
Observations	5,333	5,333	5,333	5,333

*Notes:*Columns 1 and 2 present the estimations for the number of impressions from an OLS regression. Columns 3 and 4 present the marginal effects of the estimations for the probability to click on the ad. All the regressions include day fixed effects and high school characteristics. Robust standard errors reported in parentheses. Significance at 1%; 5% and 10% indicated respectively by ***, ** and *.

Table 6 presents the results of the regressions. Columns (1) and (2) provide the estimates for the number of impressions, where Column (2) adds the interaction term *Girls in Science track* × *Treatment*, and Columns (3) and (4) present the probability to click on the ad. Columns (1) shows that the number of impressions received in a high school decreases with the proportion of girls in Science track in the high school. The interaction term *Girls in Science track* × *Treatment* in Column (2) is negative and significant, meaning that there is some non-linearity in the effect of the variables on the number of impressions performed by the algorithm, suggesting that high schools with a greater proportion of girls in Science track received fewer impressions if they belong to the treatment group. While high schools in the treatment group with a greater proportion of girls in Science track received significantly fewer impressions, column (4) shows that they clicked significantly more on the ad, suggesting a greater interest in it.

4.3 Did the algorithm learn from individuals' ad clicks?

A possible interpretation of our main result is that the ad algorithm is more likely to distribute the ad to teens that it is assumed will be interested in the ad content. To measure individuals' interest in the ad, we estimate the probability of clicking on it using a probit estimation. We model *Ad clicks* using the same set of variables as the model of *Impressions*. *Ad clicks* is a binary dependent variable measuring whether at least one teen in the focal high school clicked on the ad. Table 7 presents the marginal effects of the probability to click. Column (1) includes only the indicator of the demographic group of girls, Column (2) adds the dummy variable *Age* 18-19 showing that this group of students is more likely to click on the ad. Columns 3 includes the dummy variable $Treatment$, and Column (4) adds the interaction term *Girls SN* $\times$ *Age* 18-19. All regressions include both high schools and day fixed effects.

While girls in high school received statistically fewer impressions than boys, girls aged between 18 and 19 are significantly more likely to click on the ad with a probability of 4.1%. Column (5) includes the interaction term *Girls SN* $\times$ $Treatment$. Column (6) includes the interaction term *Age* 18-19 $\times$ $Treatment$. The main result still holds. Column (7) includes all the regressors and shows that the treatment does not have a significant impact on the probability to click for girls regardless of the age category.

Table 7: Probability to click

	Ad clicks						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girls SN	-0.012	-0.012	-0.015	-0.036**	-0.015	-0.096***	-0.038
	(0.012)	(0.011)	(0.012)	(0.017)	(0.016)	(0.019)	(0.024)
Age 18-19		0.164***	0.152***	0.144***	0.152***		0.147***
		(0.011)	(0.011)	(0.016)	(0.011)		(0.022)
Treatment			-0.031***		-0.030*	-0.063***	-0.008
			(0.012)		(0.016)	(0.019)	(0.025)
Girls SN × Age 18-19				0.041*		0.149***	0.040
				(0.023)		(0.019)	(0.032)
Girls SN × Treatment					-0.000	-0.005	-0.008
					(0.023)	(0.023)	(0.036)
Age 18-19 × Treatment						0.063***	-0.039
						(0.020)	(0.033)
Girls SN × Age 18-19 × Treatment							0.013
							(0.047)
High school fixed effects	Yes	Yes	No	Yes	No	No	No
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,113	5,113	5,333	5,113	5,333	5,333	5,333

Notes : Dependent variable is the binary variable Ad clicks. The table reports the marginal effects of the probit estimations. All the regressions include day fixed effects. Columns (1), (2), (4) include high school fixed effects. Columns (3), (5), (6) and (7) do not include high school fixed effects since they are collinear with the dummy variable of treatment. Robust standard errors are reported in parentheses.
5,113 observations are explained by the fact that high school have not receive any click and have been omitted.
Significance at 1%; 5% and 10% indicated respectively by ***,** and *.

4.4 Robustness check: Estimations on the subsample of girls

In the previous estimations, we compared girls to boys. To check the robustness of our model, we estimated the model for the subsample of girls. This allows us to compare girls enrolled in different tracks while controlling also for high school characteristics and time period.

Table 8 displays the results of the regressions of Table 6 but for the subsample of girls[16]. Columns (1) and (2) present the estimates for the number of impressions. Columns (3) and (4) estimate the probability to click. In columns (2) and (4) we add the interaction term between proportion of girls enrolled in science track and the treatment variable $Treatment$. Overall, girls aged between 18 and 19 years are more likely to see impressions and to click on the ads, compared to girls aged between 16 and 17 years.

[16] Girls on social network who received impressions

Table 8: Effects of prompting the algorithm on science oriented high schools: subsample of girls

	Impressions		**Ad clicks**	
	(1)	(2)	(3)	(4)
Age 18-19	211.853***	211.818***	0.171***	0.171***
	(7.185)	(7.172)	(0.015)	(0.015)
Treatment	-47.900***	36.333	-0.037**	0.005
	(7.793)	(22.389)	(0.017)	(0.059)
Girls in Science track	-57.433**	76.575	-0.128*	-0.062
	(27.623)	(47.715)	(0.075)	(0.113)
Girls in Economics and Social Science track	46.514	70.668**	0.073	0.085
	(28.775)	(30.360)	(0.062)	(0.064)
Girls in Literature track	51.491***	34.325*	0.048	0.039
	(19.207)	(20.432)	(0.038)	(0.040)
Girls in other tracks	-400.385***	-437.026***	-0.193*	-0.212*
	(47.772)	(48.832)	(0.116)	(0.120)
Girls in Science track × Treatment		-185.317***		-0.092
		(50.907)		(0.124)
Constant	105.982***	44.936*		
	(20.429)	(26.145)		
High school characteristics	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
R-squared	0.337	0.340		
Observations	2,684	2,684	2,684	2,684

Notes:Columns 1 and 2 present the estimations for the number of impressions from an OLS regression. Columns 3 and 4 present the marginal effects of the estimations for the probability to click on the ad. All the regressions include day fixed effects and high school characteristics. Robust standard errors reported in parentheses. This set of regressions is estimated on a smaller sample, the subsample of girls. Significance at 1%; 5% and 10% indicated respectively by ***, ** and *.

Column (1) shows that on average, girls enrolled in the science track received significantly fewer impressions and they are less likely to click. Column (2) adds the interaction effect *Girls in Science track × Treatment* which is negative suggesting that girls in the science track who benefited from the treatment ad, received fewer impressions. Column (2) shows that girls enrolled in high school with a large proportion of girls in *Economics and Social Science track* and *Literature track* are likely to see impressions.

4.5 Robustness check: Estimation of the price of our target

According to previous results, we find that girls benefit from a lower number of impressions compared to boys while girls between 18-19 click more on the ad if they came across it. In order to better identify the mechanism behind this display, we analyze the cost of our target according to age and gender. We include different measures of the price. First the Cost per Unique Impression (CUM)[17] then, the daily cost per thousand impressions (CPM daily)[18], third the Cost per thousand impressions estimated by the social media[19], and the Cost per Click (CPC)[20]. While we optimize our display by impressions in order to reach the maximum number of people, we see in Table **??** with the *CUM* and *CPM estimate* that the 'youngest' teens[21] were more expensive to advertise than the 'oldest' ones. Among teens between 18-19, Girls are the one more expensive to advertise, here we retrieve the same result as Lambrecht and Tucker (2018). However, if we look at the CPC estimate we see that Girls are also the more expensive. Indeed, if we consider a profit maximization (that seems to be in the case of impressions), the CPC would have been the lowest for girls between 18-19 because they click more. A question remain, does the algorithm takes in account the click during the campaign while we ask it to maximize by impressions not by click?

This is one of our concern. Even if girls are more reactive on the platform by clicking more on the ad which may justify the lower number of impressions for them, the facts remains they were our target of interest and the social media algorithm should have displayed more to them even if they click more. This mechanism, however, allows us to explain our equal reach among girls and boys.

[17] equal to the number of impressions divided by the amount effectively paid during the camapign

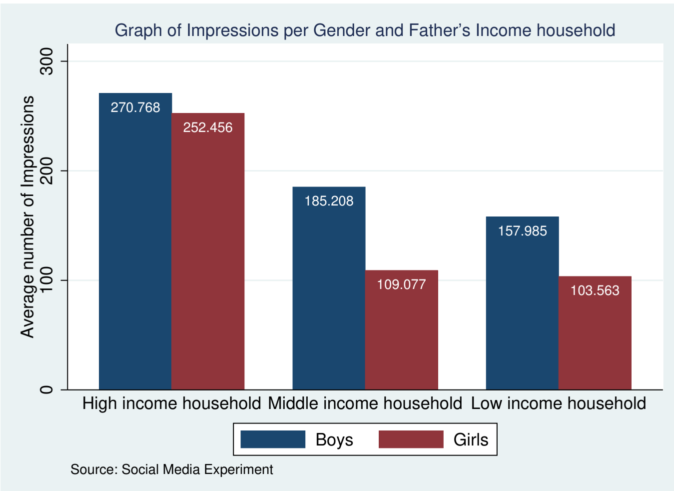
[18] amount paid during the campaign

[19] the amount we would have paid if the social media has effectively displayed 1,000 impressions

[20] the amount we would have paid if we maximize per click

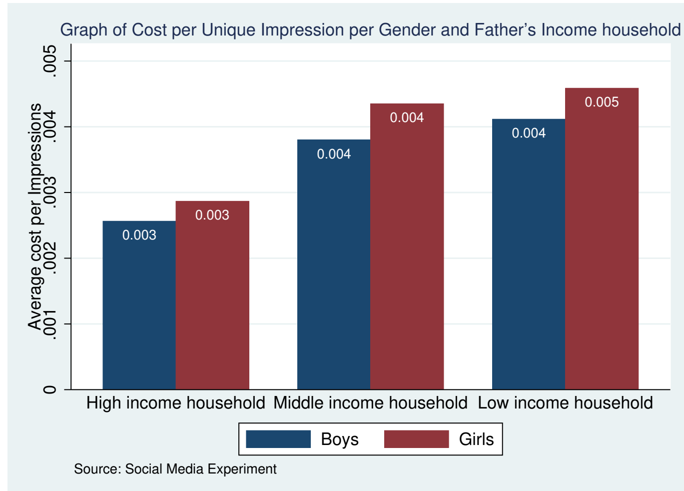
[21] Teens between 16-17

4.6 Low household income: more expensive to advertise but benefit from the lowest number of impressions



Note: High schools with a high proportion of low household income level receive less impressions than to the others level of household income. According to high household income level girls were less displayed impressions compared to boys. The highest gap of impressions between boys and girls is noticed in Middle income household

Figure 5: Average number of impressions per gender and household income level



Note: High schools with a high proportion of low income household are the highest category to advertise while they benefit from the lowest number of impressions among household income categories.

Figure 6: Average Cost per Unique Impressions according to gender and household income

Table 9: Impression, Household income and Science oriented HS

	Impressions			
Girls SN	-25.109***	-26.490***	-25.234***	-25.190***
	(5.397)	(5.622)	(5.459)	(5.396)
Age 18-19	220.937***	218.855***	221.163***	221.042***
	(5.358)	(5.555)	(5.427)	(5.363)
Treatment	-49.320***	-37.389***	-53.120***	-50.427***
	(5.421)	(5.606)	(5.514)	(5.468)
Girls in Science	35.643*	172.679***	249.883***	288.787***
	(19.550)	(15.549)	(16.657)	(16.322)
High income household	76.079***			(basis)
	(12.050)			
Girls in Science x High income household	152.843***			
	(27.205)			
Middle income household		57.004**		78.355***
		(24.074)		(24.387)
Girls in Science x Middle income household		-322.197***		-428.561***
		(48.505)		(49.140)
Low income household			122.276***	131.014***
			(34.079)	(34.062)
Girls in Science x Low income household			-532.956***	-570.334***
			(66.273)	(66.273)
Constant	97.086***	146.119***	136.281***	125.917***
	(16.318)	(15.229)	(14.905)	(14.622)
Time fixed effects	Yes	Yes	Yes	Yes
R-squared	0.319	0.260	0.303	0.319
Observations	5,333	5,333	5,333	5,333

Notes: All the regressions include day fixed effects. Robust standard errors reported in parentheses. Significance at 1%; 5% and 10% are respectively indicated by ***,** and *.

	Impressions		
Girls SN	-46.450***	-23.455***	-20.768***
	(8.585)	(5.921)	(6.337)
Age 18-19	221.017***	219.132***	219.551***
	(5.388)	(5.597)	(5.550)
Treatment	-48.989***	-40.546***	-49.689***
	(5.432)	(5.628)	(5.662)
High household income	130.656***		
	(8.482)		
Girls SN × High household income	28.268***		
	(10.856)		
Middle household income		-55.662***	
		(14.596)	
Girls SN × Middle household income		-52.676***	
		(18.774)	
Low household income			-112.807***
			(9.886)
Girls SN × Low household income			-25.321**
			(12.409)
Constant	122.941***	222.957***	239.216***
	(14.361)	(13.890)	(13.922)
Time fixed effects	Yes	Yes	Yes
R-squared	0.312	0.250	0.276
Observations	5,333	5,333	5,277

Notes: All the regressions include day fixed effects. Robust standard errors reported in parentheses. Significance at 1%; 5% and 10% are respectively indicated by ***,** and *.

5 Conclusion

Our study interest was in the role of data-driven social media algorithms. To our knowledge, the link between potential algorithm biases resulting from machine learning algorithms, and gender gap persistence in STEM education has not been investigated.

First, our main contribution is on the economics of privacy and algorithm bias. Teenagers who subscribe to a social media reveal their gender unaware that this information potentially could be used to discriminate them in the future. Second, in enrollment in STEM higher education programs girls continue to be under-represented compared to boys. Our paper contributes by adding the role of social networks as media used to inform teens about education and future jobs.

We conducted this online experiment to test whether the ad algorithm of a social network reproduces stereotypes. We ran 101 simultaneous ad campaigns at the high school level. The experiment was conducted over 2 weeks and aimed at boys and girls aged between 16 and 19 years. High schools were split into two groups, control and treatment, to test the ad display formulated by the algorithm. The aim of our field experiment was twofold. First, to measure whether the algorithms distribute STEM ads equally to girls and boys according to the percentage of girls enrolled in each high school. Second, to estimate whether a girl-oriented message is distributed to girls more than boys. Our study extends a previous experiment conducted by Lambrecht and Tucker (2018), and contributes additional findings on this topic.

Our paper is related closely to Lambrecht and Tucker (2018). However, while they observe a cost difference between boys and girls, which might explain the differences in ad delivery, we found no such evidence. We also conducted several robustness checks, which showed that the algorithm made fairer decisions towards girls enrolled in a science track. These findings allow us to (i) verify the existence of bias, but (ii) conflict with previous findings that cost explains ad display disparities between boys and girls. Finally, we highlight that algorithms might be an effective and fairer mean of targeting girls attending science-oriented high schools who will likely be more interested in the ad set up.

This study makes several contributions. First, it adds to work on the existence of biases and the reproduction of stereotypes related to teenagers interest in STEM fields. Second, we provide evidence that ad algorithm cannot help reduce the gender gap in STEM as they might distribute biased information to teens.

This research raises several questions about the role of machine learning algorithms for ad effectiveness and ad personalization. Although, intelligent algorithms might generate apparent discrimination and lower levels of ad display to girls, they are able to decide whether to show the ad to those girls likely to be more interested in the ad content.

6 Implications

7 Appendix

PAY LING COMMENT:

- Find what the algorithm is maximizing: reach
- Girls in science schools click on the ad because they already know about this school or because they compare the different choice of school
- modifier l'intro(couterfactuel)
- le raisonnement sur le reach
- insertion des incomes
- discussion ”implications” sur le contrefactuel

Table 10: Effects of prompting the algorithm on science oriented high schools: Missing excluded

	Impressions		**Ad clicks**	
	(1)	(2)	(3)	(4)
Girls SN	-31.092***	-31.038***	-0.072***	-0.072***
	(5.899)	(5.895)	(0.025)	(0.025)
Age 18-19	204.834***	204.957***	0.114***	0.115***
	(9.672)	(9.676)	(0.022)	(0.022)
Treatment	7.866	-50.589	0.044***	-0.065
	(7.120)	(41.047)	(0.017)	(0.095)
Girls in Science	-379.190***	-442.145***	-0.424***	-0.540***
	(48.458)	(65.804)	(0.121)	(0.162)
Girls in Economics and Social Science	-30.849	-31.584	-0.125	-0.131
	(46.960)	(46.878)	(0.118)	(0.118)
Girls in Literature	352.671***	365.904***	0.131	0.152
	(49.619)	(51.193)	(0.122)	(0.124)
Girls in other tracks	-191.938***	-191.977***	-0.068*	-0.069*
	(20.299)	(20.233)	(0.040)	(0.040)
Girls SN x Age 18-19	1.105	1.051	0.086***	0.086***
	(13.681)	(13.679)	(0.032)	(0.032)
Girls in Science x Treatment		120.046		0.226
		(81.123)		(0.192)
Constant	239.663***	259.752***		
	(45.087)	(46.512)		
Time fixed effects	Yes	Yes	Yes	Yes
High school characteristics	Yes	Yes	Yes	Yes
R-squared	0.346	0.347		
Observations	2,776	2,776	2,776	2,776

Notes:Columns 1 and 2 present the estimations for the number of impressions from an OLS regression. Columns 3 and 4 present the marginal effects of the estimations for the probability to click on the ad. All the regressions include day fixed effects. Robust standard errors reported in parentheses. This set of regressions is estimated on a smaller sample than the previous regressions due to missing variables in the vector of high school characteristics. Significance at 1%; 5% and 10% indicated respectively by ***, ** and *.

Table 11: Effects of prompting the algorithm on science oriented high schools: Sub-sample of girls, Missing excluded

	Impressions		**Ad clicks**	
	(1)	(2)	(3)	(4)
Age 18-19	205.758***	205.929***	0.192***	0.192***
	(9.636)	(9.617)	(0.021)	(0.021)
Treatment	-22.081**	-163.467***	0.008	-0.090
	(9.886)	(55.149)	(0.023)	(0.132)
Girls in Science	-328.895***	-481.431***	-0.364**	-0.466**
	(66.920)	(91.199)	(0.170)	(0.220)
Girls in Economics and Social Science	22.792	20.969	-0.068	-0.073
	(70.827)	(70.522)	(0.162)	(0.162)
Girls in Literature	372.566***	404.789***	0.207	0.226
	(71.854)	(74.039)	(0.169)	(0.172)
Girls in other tracks	-136.406***	-136.479***	0.014	0.013
	(27.406)	(27.153)	(0.055)	(0.056)
Girls in Science x Treatment		290.358***		0.203
		(111.680)		(0.267)
Constant	116.183*	164.873**		
	(64.265)	(66.792)		
Time fixed effects	Yes	Yes	Yes	Yes
R-squared	0.332	0.334		
Observations	1,403	1,403	1,403	1,403

Notes: Columns 1 and 2 present the estimates for number of impressions using an OLS. Columns 3 and 4 present the marginal effect for the probability to click on the ad. All the regressions include day fixed effects. Robust standard errors reported in parentheses. Significance at 1%; 5% and 10% are respectively indicated by ***,** and *.

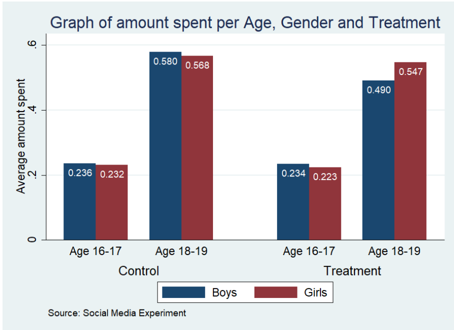


Figure 7: Amount daily spend on average by Treatment, gender and Age

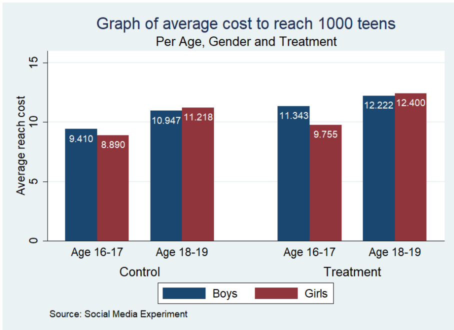


Figure 8: Average cost to reach 1,000 teens by Treatment, gender and Age

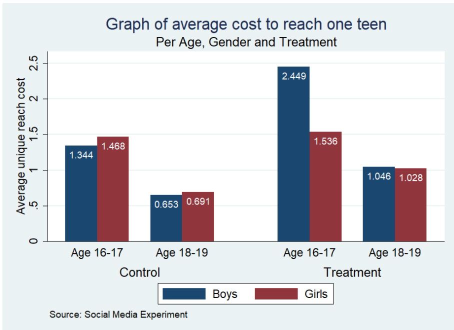


Figure 9: Average cost to reach 1 teen by Treatment, gender and Age

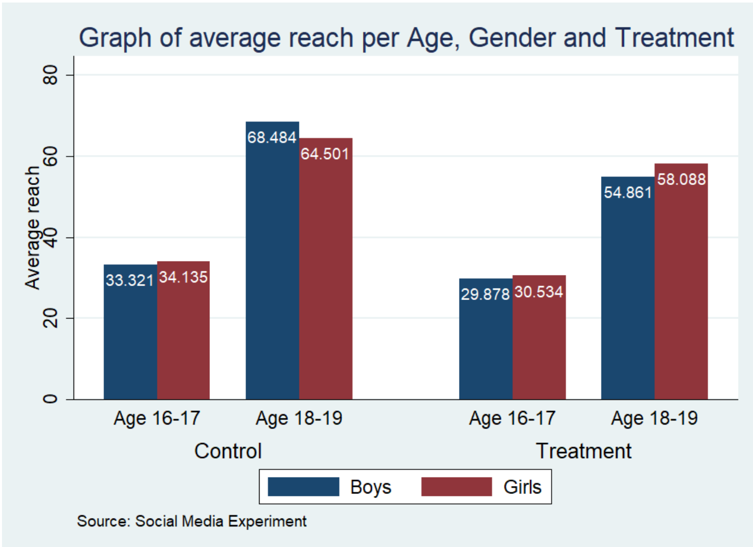


Figure 10: Average reach

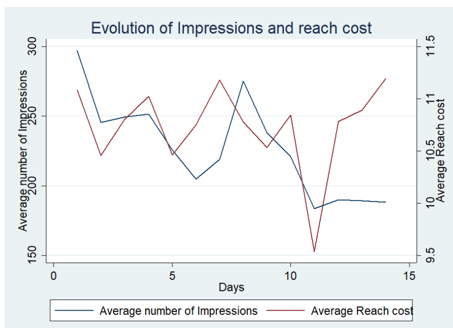


Figure 11: Average number of Impression and reach cost among days

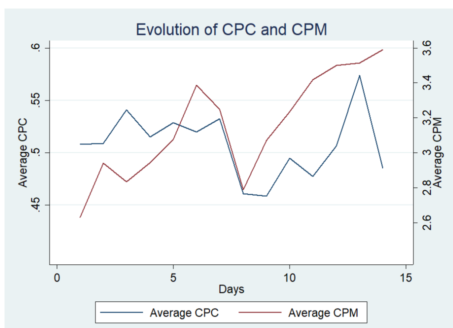


Figure 12: Average CPC and CPM among days

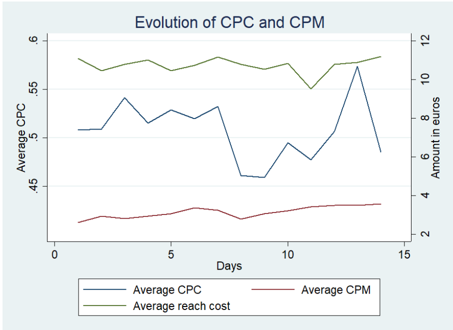


Figure 13: Average CPC, CPM and CPR among days

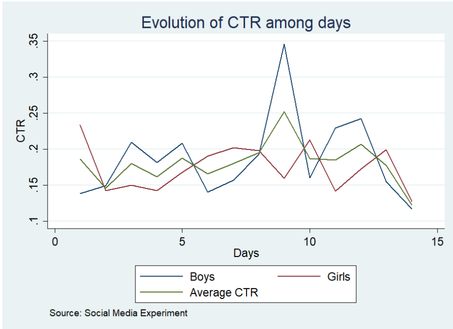


Figure 14: Average CTR among days

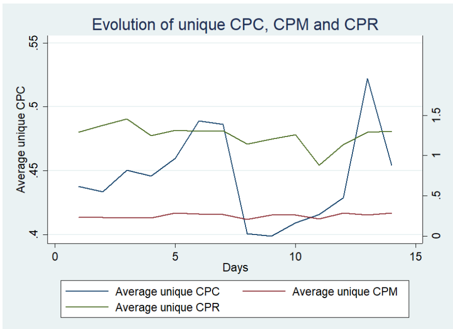


Figure 15: Average unique cost among days

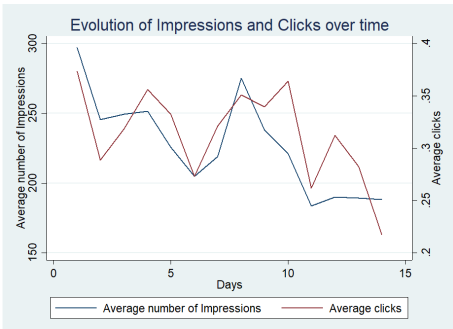


Figure 16: Evolution of clicks and Impressions over time

Table 12: Main result - Effect of prompting the algorithm

	Reach						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girls SN	0.369	0.591	0.068	1.809***	-0.915	-0.909	0.822
	(0.633)	(0.463)	(0.934)	(0.578)	(1.265)	(1.263)	(1.253)
Age 18-19		31.157***	29.749***	32.362***	30.457***	33.452***	35.171***
		(0.466)	(0.927)	(0.635)	(1.313)	(1.562)	(1.779)
Treatment			-6.810***		-8.644***	-5.342***	-3.448**
			(0.937)		(1.331)	(1.355)	(1.378)
Girls SN x Age 18-19				-2.392***	-1.398	-1.384	-4.805*
				(0.919)	(1.853)	(1.852)	(2.531)
Girls SN x Treatment					3.640*	3.598*	-0.140
					(1.873)	(1.872)	(1.862)
Age 18-19 x Treatment						-6.454***	-10.156***
						(1.856)	(2.628)
Girls SN x Age 18 19 x Treatment							7.353**
							(3.711)
Constant	56.026***	40.530***	37.043***	39.921***	37.541***	36.015***	35.138***
	(1.615)	(2.017)	(1.832)	(2.035)	(1.882)	(1.872)	(1.871)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
High school fixed effects	Yes	Yes	No	Yes	No	No	No
R-squared	0.627	0.800	0.168	0.800	0.169	0.171	0.171
Observations	5,333	5,333	5,333	5,333	5,333	5,333	5,333

Notes: OLS estimations. The dependent variable is number of *Reach*. Columns (1), (2), (4) include high school characteristics. Robust standard errors are reported in parentheses. Columns (3), (5), (6), (7) do not include high school fixed effects since they are collinear with the dummy variable of treatment. All the regressions include day fixed effects. Significance at 1%; 5% and 10% levels indicated respectively by ***,** and *.

References

Acquisti, A., Taylor, C. and Wagman, L. (2016). The Economics of Privacy. *Journal of Economic Literature*. 54(2), 442–92.

Agrawal, A., Gans, J. and Goldfarb, A. (2016). The simple economics of machine intelligence. *Harvard Business Review*.

Akerlof, G. A. and Kranton, R. E. (2010). *Identity economics: How our identities shape our work, wages, and well-being*. Princeton University Press.

Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016). Machine Bias: There's software used across the country to predict the future criminials and its biased against blacks. *ProPublica*.

Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*. 131(4), 1753–1794.

Chatterjee, P., Hoffman, D. L. and Novak, T. P. (2003). Modeling the Clickstream: Implications for Web-Based Advertising Efforts. *Marketing Science*. 22(4), 520–541.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. and Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics*. 126(4), 1593–1660.

Cowgill, B. and Tucker, C. (2017). Algorithmic Bias: A Counterfactual Perspective.

Croson, R. and Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*. 47(2), 448–74.

Datta, A., Tschantz, M. C. and Datta, A. (2014). Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination. *ArXiv e-prints*.

Eurostat (2015). *Being young in Europe today*. Eurostat Statistical books.

Fryer, R. G. and Levitt, S. D. (2010). An Empirical Analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics*. 2(2), 210–40.

Goldfarb, A. and Tucker, C. E. (2011). Privacy Regulation and Online Advertising. *Man-*

agement Science. 57(1), 57–71.

Guiso, L., Monte, F., Sapienza, P. and Zingales, L. (2008). Culture, Gender, and Math. *Science*. 320(5880), 1164–1165.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J. and Mullainathan, S. (2018). Human Decisions and Machine Predictions. *The Quarterly Journal of Economics*. 133(1), 237–293.

Lambrecht, A. and Tucker, C. E. (2018). Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Management Science, forthcoming*.

Leibbrandt, A. and List, J. A. (2015). Do Women Avoid Salary Negotiations? Evidence from a Large-Scale Natural Field Experiment. *Management Science*. 61(9), 2016–2024.

Livingstone, S., Ólafsson, K. and Staksrud, E. (2011). Social networking, age and privacy.

OECD (2017). Going Digital: The Future of Work for Women. *Policy brief on the future of work*.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY, USA: Crown Publishing Group.

Pew Research Center, P. (2016). Social Media Update. (pp. 1–18). November.

Sweeney, L. (2013). Discrimination in Online Ad Delivery. *Queue*. 11(3), 10:10–10:29.

Tucker, C. (2017). Privacy, Algorithms and Artificial Intelligence. In *Economics of Artificial Intelligence*. University of Chicago Press.