NBER WORKING PAPER SERIES

DO GENDER-NEUTRAL JOB ADS PROMOTE DIVERSITY? EXPERIMENTAL EVIDENCE FROM LATIN AMERICA'S TECH SECTOR

Lucia Del Carpio
Thomas Fujiwara

Working Paper 31314
http://www.nber.org/papers/w31314

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2023, Revised January 2025

Research funding from INSEAD's Diversity and Inclusion Fund is gratefully acknowledged. We are particularly grateful to Get on Board and Laboratoria for being excellent partners and to Gaspar Arias, Pedro Cabra-Acela, Mylene Feuillade, Gurcan Gulersoy, Eugenio Piga, and Edgar Sanchez-Cueva for excellent research assistance. We thank Jose Maria Barrero, Ruben Durante, Dylan Glover, Emilio Gutierrez, Adrian Rubli, and seminar participants at GRADE, INSEAD, ITAM, and Universidad del Pacifico for their helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Do Gender-Neutral Job Ads Promote Diversity? Experimental Evidence from Latin America's Tech Sector
Lucia Del Carpio and Thomas Fujiwara
NBER Working Paper No. 31314
June 2023, Revised January 2025
JEL No. J16, J7, M14, Z13

ABSTRACT

Gendered-grammar languages are spoken by 39% of the global population. We conduct two experiments studying the effects of gender-neutral language in job advertisements and its treatment spillovers. In a Spanish-speaking tech job platform, ads randomly assigned to gender-neutral language attract more female applicants, but only when a small proportion of other ads considered by applicants is treated. In a second experiment, gender-neutral language in ads affects beliefs about job characteristics. Our results suggest that female applicants interpret gender-neutral language as a signal about job amenities and that scalability is limited: if most ads were gender-neutral, the effects would be negligible.

Lucia Del Carpio
ITAM
Mexico
lucia.delcarpio@itam.mx

Thomas Fujiwara
Department of Economics
Princeton University
131 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and NBER
fujiwara@princeton.edu

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/10076
and
https://www.socialscienceregistry.org/trials/5509

1 Introduction

Language can shape cognition and decisions. Gendered distinctions, in particular, have been hypothesized to make gendered divisions of labor seem more natural (Whorf, 1956). In English, the generic use of "he" (as opposed to "she/he") leads subjects to imagine male referents (Moulton et al., 1978, Cole et al., 1983, Gastil, 1990). Women recall information better when instructions refer to women (Crawford and English, 1984). Gender-neutral language in Israeli college entrance exams improves female performance (Cohen et al., 2023). Jakiela and Ozier (2018) provides global evidence showing that speakers of gendered-grammar languages have lower female labor force participation and educational attainment.

In recent years, advocacy and controversy have grown around using inclusive language to enhance diversity, yet evidence about its effectiveness remains scarce. We conducted two randomized experiments to study the effects of gender-neutral language in job ads within Latin America's tech sector. Women make up only 7% of the tech workforce in the region (Del Carpio and Guadalupe, 2021), which has seen both informal adoption of gender-neutral language and government interventions supporting or opposing its use (see Appendix A).

Our first experiment includes all ads posted on an online job board, enabling us to examine treatment spillovers: whether the effects of gender-neutral language become more muted as more ads in applicants' consideration sets are also gender-neutral. Our second experiment investigates the mechanisms at play by studying how gender-neutral language in ads affects female tech sector workers' beliefs about the position. A growing literature explores how interventions on the content and language of recruitment materials affect the composition of the applicant pool. To our knowledge, our study provides the first evaluation of gender-neutral language and the first examination of treatment spillovers for any type of content.[1]

In Spanish, like many gendered-grammar languages spoken by 39% of the world population (Jakiela and Ozier, 2018), nouns have a male or female gender. The traditional default is to use the masculine form as a "generic" when referring to an unspecified sex. For example, there are words for "male engineer" (*ingeniero*) and "female engineer" (*ingeniera*), but no word referring to an engineer without conveying gender, so job ads only mention *ingeniero*.[2]

[1]For experiments, see Abraham et al. (2024), Coffman et al. (2024), Del Carpio and Guadalupe (2021), Delfino (2024), Flory et al. (2015, 2021), Gaucher et al. (2011), Ibañez and Riener (2018), Leibbrandt and List (2015, 2018), Mas and Pallais (2017), Samek (2019). Papers based on observational data are Card et al. (2024), Helleseter et al. (2020), Kuhn et al. (2020), Kuhn and Shen (2023). None study treatment spillovers, general equilibrium effects, or gender-neutral language. Most experimental papers vary the content of ads *from a single firm*, making it difficult to study spillovers. An exception is Kuhn and Shen (2023), a non-experimental paper on a Chinese board's ban on explicit gender requests affecting all ads simultaneously. It estimates effects both on ads directly and non-directly affected (without requests before the ban). We study a different type of spillovers, leveraging random variation in the share of ads for similar positions that were concurrently treated. Appendix B further discusses these papers.

[2]Plurals are also gendered (*ingenieros* and *ingenieras*). Appendix A discusses gendered grammar in Spanish and further issues (and controversies) related to gender-neutral language.

Our first experiment was conducted in partnership with Get on Board, a widely used website for tech sector job ads in Latin America. From April to November 2020, *all* ads submitted to the platform (over 2,000) were randomly assigned to a gender-neutral treatment or control ("business as usual") condition. Treated ads were edited to include only gender-neutral language following a protocol based on government guidelines. For example, "*ingeniero*" was revised to "*ingeniera/o*." Potential applicants were unaware of the experiment; they observed that some ads used gender-neutral language while others did not.

A key contribution in our study regards treatment spillovers, which are crucial for identifying mechanisms and understanding scalability: whether effects would persist if most (or all) ads used gender-neutral language (List, 2022). In particular, applicants may interpret gender-neutral language as a signal about firm characteristics and job amenities. For example, they might expect that firms that use gender-neutral language are also more likely to offer flexible work hours or employ more women. This updating mechanism suggests that a policy mandating that all ads use gender-neutral language would have no effect, as applicants have nothing new to infer from gender-neutral ads.

On the other hand, the mechanisms behind the effects of gender-neutral language may be more psychological ("behavioral") in nature, such as imagining female referents or better recall (Moulton et al., 1978, Cole et al., 1983, Gastil, 1990, Crawford and English, 1984). The implications for spillovers are less clear, and it is possible that a policy mandating gender-neutral language in all ads could still have significant effects.

To study spillovers, we define, for each ad, a set of *neighbor ads* that applicants likely saw listed together when using the platform. In practice, given the specifics of the platform's user experience, these are ads with similar job titles posted within a week of each other. Given random and independent treatment assignment, our measure of treatment spillovers (share of neighbor ads treated) is also random. This allows us to estimate how treatment effects differ by share of neighbor ads leveraging random variation both in the treatment itself and in the relevant margin of treatment heterogeneity. Using causal forests (Athey et al., 2019), we confirm the share of neighbor ads treated as the key source of treatment heterogeneity.

Results indicate that treatment increases the share of women who applied to the position, but only for ads randomly assigned to small shares of neighbor ads treated, and thus likely to be perceived by applicants as a relatively rare gender-neutral ad. The effect is mostly driven by more women applying. We also find suggestive evidence that the effects are uniform throughout the "candidate quality" distribution (both more and less qualified women apply to the position) and that treatment increases the share of women that advance to later stages of the recruitment process, and perhaps are hired.[3]

Consistent with the presence of treatment spillovers, effects for ads with larger shares of neighbor ads treated are negative but statistically indistinguishable from zero. Our results thus suggest that the effects of gender-neutral language have limited scalability and that policies that increase its prevalence to an extent that it becomes "common" are unlikely to have substantial effects. Indeed, we find a zero treatment for the entire sample (where 52% of ads are gender-neutral).[4]

To investigate underlying mechanisms, we conducted a second experiment in partnership with Laboratoria, an NGO that trains Latin American women for tech sector jobs. In an online survey sent to its alumni, each respondent saw two ads that were randomly assigned to use gender-neutral or generic masculine language. Subjects were asked about their propensity to apply for and their beliefs about the position. To study spillovers, the randomization protocol made gender-neutral language more noticeable in the second ad than the first: respondents either saw a gender-neutral ad followed by a generic-masculine ad, or vice versa.[5]

The (female) respondents reported being more likely to apply and believe they are suitable for the job and likely to be hired for jobs with a gender-neutral ad. Moreover, they also stated that the firm with a gender-neutral ad was more likely to have an inclusive culture, promote work-life balance, and employ a larger share of women. Additionally, in a cross-randomized design, we varied whether ads stated the position was remote and whether it included a statement about the company's commitment to diversity. The effects of gender-neutral language were substantially larger than those of diversity statements and comparable or larger than the effect of making the position remote.

Corroborating the importance of spillovers, the effects of gender-neutral language are substantially larger when respondents previously saw an ad with a different status, relative to when they evaluated a first ad without being provided a clear comparison ad. We do not find a similar pattern for the diversity statement or remote position treatments.

Results from both experiments are thus consistent with female applicants interpreting the use of gender-neutral language as a signal and using it to update their beliefs about the firm's characteristics and job amenities. Corroborating this, we do not find effects of gender-

[3]Specifically, treatment increases the share of female applicants by 3.9 percentage points if the ad's share of neighbors ads treated falls in the first quartile of its distribution. The average share of female applicants in the control group is 14.6%. Ads falling in the first quartile have, on average, 7.7 neighbor ads, with 20% of them assigned to treatment. Sections 2.1 and 3 discuss caveats about our measure of candidate quality and how information on applicants advancing to later stages is observed for a selected sample of firms.

[4]The effect for the entire sample is 0.0002 p.p. (SE=0.0068). The 52% share of gender-neutral ads is not a "business as usual" scenario since it includes treated ads. Effects discussed here are intent-to-treat and Section 3 discusses treatment-on-treated effects.

[5]Ads were fictional but subjects were told the goal of the experiment was to calibrate future job advertisements to incentivize truthful answers (Kessler et al., 2019). Respondents were not made aware the survey involved evaluating gender-neutral language.

neutral language on a question about how respondents evaluated themselves (whether they met the job's requirements, which were clearly stated in the ad). The spillover results are also consistent with this updating mechanism: when applicants perceive gender-neutral ads as more common, they update to a lesser extent. Thus the "updating mechanism" can explain the entirety of our results. The same is not as clear for "psychological mechanisms" such as imagining female referents or better recall, which do not have clear predictions for spillovers or immediately explain the Laboratoria experiment's results.

This paper speaks to two strands of literature. The first is how content in job advertisements affects the diversity of the applicant pool, particularly in fields where women and minorities are under-represented (see Footnote 1 and Appendix B). As mentioned earlier, we innovate on two fronts: by providing the first evaluation of gender-neutral language in recruitment materials and the first exploration of treatment spillovers for any type of content. The latter is key in understanding scalability (List, 2022). Indeed, our results suggest that a smaller-scale experiment treating a fraction of ads would suggest gender-neutral language can promote diversity, but our overall results indicate that this is unlikely to scale. Second, literature dating back to the Whorf (1956) hypothesis studies how language affects cognition and behavior (see the first paragraph of this introduction and Appendix A). We contribute to it by testing how an intervention that only affects language affects job applications and shedding light on the mechanisms at play.

The paper is organized as follows. Section 2 describes the two experimental designs. Section 3 provides the results for the Get On Board experiment, while Section 4 does so for the Laboratoria experiment. Section 5 concludes.

2 Experimental Designs

2.1 First experiment: Get on Board

The platform. Get on Board (getonbrd.com) is one of the largest online job boards focused on tech sector professionals in Latin America. By 2024, almost one million professionals had submitted over 2.8 million applications to more than 12,000 registered companies via the website. More than 90% of posted jobs are full-time.

To post job ads, companies pay a submission fee or subscribe to a service allowing multiple postings. All ads are first submitted for moderation where Get On Board staff ensures they comply with quality standards. Ads are presented in a standardized format with the job title prominently displayed. Ads have a header describing the company, followed by a description of job roles, candidate features that are "required" or "desirable", and job benefits (Figure 1).

Companies with a subscription have access to a personalized evaluation board where they can rank candidates who apply for their jobs, such as which ones to discard, pass the first round of screening, or select for the job. Not all companies use this tool (Section 3 discusses this further). We describe the user experience for job applicants later in this section, as it plays a key role in our analysis of treatment spillovers.

Scope and randomization. Between April 17 and November 27, 2020, *all* 2,535 job advertisements submitted to the platform were assigned to either a control or treatment status. Treatment assignment had a 50% probability and was *independently* drawn for each ad. An ad under *control* status is treated as the platform usually treats its ads. An ad under *treatment* underwent the same process plus the additional protocol described below.[6]

Firms that submitted ads assigned to treatment received the message below:

> *This job has been randomly selected for gender-neutral moderation. We are evaluating requiring gender-neutral language to all jobs. For a brief period, we are selecting jobs at random, and our moderation team is making sure they comply with gender-neutral language guidelines. This requires no action on your part.*
>
> ***Ok, keep this job in the study (default)***
> ***Remove this job from the study***

Only two ads (out of 1242 assigned to treatment) chose to opt out of the experiment.

Treatment protocol. Ads assigned to treatment were edited by Get On Board staff to comply with a gender-neutral language protocol before being posted. This editing process was integrated into the business-as-usual moderation stage of a job posting, which ensures that all ads, including those in the control group, meet basic standards. This allows our treatment to occur naturally within the usual advertiser experience.

The gender-neutral language protocol was based on recommendations provided by South American governments (Appendix A) and consisted of two ranked guidelines. The first (preferred) involved the use of strategies that avoid using the "generic masculine" form: e.g., replacing them with (gender-neutral) relative pronouns, imperative verbs, and nouns with no gender assigned.[7] Second, when it was not possible to avoid "generic masculines," the ad gave visibility to both genders by doubling the word in the feminine first and the masculine second (e.g., "ingeniero" should be changed to "ingeniera/o.").

[6] The experiment was registered with the AEA's RCT registry in March 2020 (AEARCTR-0005509).

[7] For example, when instructing candidates meeting requirements to send a CV, "*Los candidatos* que cumplan con los requisitos deberán enviar su CV" should be changed to "*Envíe* su CV si cumple con los requisitos" (replacement of a masculine noun with an imperative form). When telling dynamic and innovative candidates to apply, "si eres *dinámico e innovador...*" should be changed to "si eres *una persona dinámica e innovadora*" since "persona" (person) is a noun that applies to both genders.

Figure 1 provides an example of the same ad under control and treatment status. Table A.15 shows key examples of the protocol and Appendix F contains the exact guidelines used by Get On Board staff.

Data. We collected data on the entire text of the ads, which provides information on the position (e.g., job title, seniority level, whether it is remote or in-person), as well as data on the applicants themselves. Applicants' gender (male or female) was coded based on their first names.[8] We also observe a measure of candidate quality: the "badness scores." To apply for a job, professionals must register with the platform. Get on Board evaluates professionals based on their history recorded in the evaluation boards, creating an index internally referred to as their "badness score." The score evolves as they go through different recruitment processes: each time an applicant is rejected or moves on to the next stage, the score goes up or down, respectively. A lower badness score signals a "better" applicant from the revealed preference of companies' hiring processes.[9]

Applicants' user experience. The most common way potential applicants find job ads is by using a prominently displayed search bar that prompts users to "search for jobs." Searching for a particular job title (e.g., "desarrollador full stack") will then provide a list of ads with similar (but not necessarily the same) job title. Figure 2 provides an example. The search algorithm simultaneously handles job titles in both English and Spanish. For example, a search for "web developer" will also return jobs titled "desarrollador web." Users can also browse through a predetermined set of 12 fields that Get On Board uses to classify ads, although this is not as common as searching. Table A.4 lists the 12 fields and provide their representation in our sample.

Importantly for our analysis, both when searching or browsing, ads are essentially listed in chronological order (more recently posted positions are listed first).

Spillovers and share of neighbor ads treated. We study treatment spillovers between ads that applicants see listed together when using the platform. The key variable operationalizing this is the *share of neighbor ads treated*. For each ad i, we define a set of *neighbor ads* which are all ads in the sample that were i) posted within a 7-day window (same day or 3 days before or after) of ad i and ii) belong to the same *job title group* as ad i.

We classify ads into 16 job title groups. Each group represents a set of job titles with similar on-the-job roles and tasks. Moreover, they reflect the ads applicants would see listed

[8]First names in Spanish-speaking countries are more straightforward to assign a gender than in English-speaking countries. Only 1.62% of applicants had a name that could not be easily assigned to a gender.

[9]Applicants cannot observe their own scores, which is used internally by Get On Board and subscribing firms. In 2021 (after our experiment) the platform stopped its use of the scores. We tracked applications until all ads in our sample until they were "closed" and stopped accepting further applications.

together after their search results. For example, job titles such as "UX/UI Designer," "Diseñador UI," "Diseñador/a UX," and "Diseñador UX/UI" are grouped into the *designer* group, capturing that searches for, e.g., "Diseñador UI," would also provide ads for the other listed positions. Ads are classified solely by the text of their title before assignment to treatment, thus group composition is similar for control and treated ads. Table A.3 lists the 16 groups and Appendix C describes the procedure for assigning ads to job title groups.[10]

Thus, the share of *neighbor ads assigned to treatment* measures the intensity of treatment spillovers. Intuitively, it combines both timing and job title information to provide a proxy for the share of treated ads among those that Get On Board users see listed together with ad i.[11]

Since treatment is assigned to each ad *independently*, ad i's share of neighbor ads assigned to treatment is a random variable (following a binomial distribution) that is independent of ad i's characteristics and ad i's own treatment assignment. This is a key advantageous feature of our experimental design. We estimate treatment effect heterogeneity identified from random variation both in the treatment itself and in the intensity of treatment spillovers, which is the relevant dimension of treatment effect heterogeneity.[12]

Summary statistics and balance. Since the share of neighbor ads treated is a key variable in our analysis, we exclude from the sample 334 ads for which its value is missing.[13] Thus our main sample includes 2,201 ads from 792 unique companies. The share of treated units was 48.7%.[14] These ads received a total of 104,680 applications, of which 47.3% were for the ads in the treated condition. The average ad received 9.2 applications from women and 38.2 applications from men. The distribution of number of applications is right-skewed, with a few ads receiving several hundred or even over a thousand applications.

Figure A.1 shows the number of ads posted by week of the experiment, indicating balance by treatment status and also that the overall number of ads posted in the platform increased over time. Table A.1 presents the average characteristics of the control and treatment ads.

[10] As another example, job titles such as "Back-end Developer Java Node," "desarrollador Back-end Python," and "Back-end Developer" are grouped into *back-end developer* job title group.

[11] We use a 7-day window as our baseline "specification" since it approximates the size of ads listed on the page (based on our experience testing different searches on the website) and it averages out day of the week of effects (every 7-day window includes one Monday, one Saturday, and so on). Section 3 discusses the robustness to different windows.

[12] Formally, ad i with n_i neighbor ads has a share of neighbors treated following a binomial distribution $B(n_i, 0.5)$. Ad i's set of neighbor ads (and thus n_i) is determined before i's treatment assignment and cannot be affected by it.

[13] Of the 334 ads removed, 103 were removed because they could not be assigned to a specific job title group (see Appendix C) and 103 ads did not have at least one neighbor ad (ads without no ad from the same job title group posted within a 7-day window or ads that could not be assigned to a job title group).

[14] This number differs from the expected 50% but is consistent with our random assignment. The probability of an equal or larger deviation from a 50%-50% split in a binomial distribution with 2,201 draws and 0.5 probability in each draw is 21%.

Control and treatment ads are balanced in terms of seniority of the position, whether they presented a wage range (and its value), whether the position is remote, and the number of neighbor ads. An omnibus test of joint orthogonality following Kerwin et al. (2024) does not reject the null of balance across all available covariates (p-value = 0.34, see Appendix D). Roughly 40% of the positions are remote, given the experimental period coincided with the first months of the Covid-19 pandemic. Information on the country of the firm posting the ad is not available for remote positions. Amongst non-remote positions, 86.9% of ads are for positions based in Chile and 9.3% for positions in Peru. Argentina, Brazil, Colombia, Costa Rica, and the United States are also represented.[15]

As discussed above, ad i's share of neighbors treated is a random variable orthogonal to ad i's characteristics and own treatment status. Corroborating this, Table A.1 shows that the average share of neighbors treated is similar in the control and treated ads. Moreover, Table A.2 shows that the share of neighbor treated is uncorrelated with ad characteristics.

Construal and subject perceptions. (Potential) job applicants were not aware an experiment was taking place or that some ads were chosen by Get On Board to implement gender-neutral language, making this a "natural field experiment" (Harrison and List, 2004). From the applicants' point of view, some ads on the platform were gender-neutral and some were not, and the most plausible interpretation is that it was the choice of the companies themselves to write gender-neutral ads. It would thus be natural for them to make inferences about the company from its use of language.

Variation in the share of treated neighbor ads can influence applicants' perceptions of gender-neutral language usage among potential employers. For example, if ad i posted by employer e uses gender-neutral language while few of its neighbor ads do, potential applicants might perceive that employer e made an uncommon choice and use it to infer that e differs from other firms. Conversely, if most of ad i's neighbor ads also use gender-neutral language, there's less reason to see e as distinctive.

2.2 Second Experiment: Laboratoria

While the Get On Board experiment allows us to estimate the effects of gender-neutral language in a "natural field experiment" (Harrison and List, 2004), the Laboratoria experiment was designed to delve into the mechanisms at play by examining how perceptions of job amenities and company characteristics are influenced by the use of gender-neutral language in ads.

A nonprofit organization founded in Peru in 2015, Laboratoria has expanded to Chile,

[15] See Appendix C for the definition or remoteness status and further information.

Mexico, Colombia, Ecuador, and Brazil. It offers six-month coding boot camps in Web Development and UX Design to build *female* trainees' technical and life skills. Over 85% of graduates secure jobs in the tech sector upon graduation. As of 2022, Laboratoria had an alumni network of over 2,500 women.

The experiment took place in September and October 2022. The survey and all communications with participants were in Spanish, except for alumni of the Brazilian boot camp, which was in Portuguese (also a gendered-grammar language). Appendix G provides all the experimental materials.[16]

Scope and invitations. Laboratoria runs an e-mail newsletter recommending a selection of jobs available on various online platforms to its alumni. Within this newsletter, Laboratoria sent an invitation inviting them to collaborate on "*a study that seeks to find out how job advertisements published on various job platforms in the technology sector are perceived*" to "*promote better quality in the selection of recommended ads, allowing more people to find the job they are looking for.*" Participation included an entry into a draw to win an Amazon Kindle. Neither the invitation nor the survey explicitly mentioned gender-neutral language in any manner to avoid priming the subjects and minimize potential demand effects. Since Laboratoria's alumni are exclusively female, our sample consists only of women.

Experimental Design. Each respondent was shown two fictitious job ads in their field of graduation. To avoid deception, respondents were informed they were fictitious. However, they had a motivation to respond truthfully since their answers would impact the future job recommendations they receive from Laboratoria. Kessler et al. (2019) employs a similar strategy to incentivize employers rating resumes. To make them realistic, ads were written to closely mimic those on Get on Board (see Figure 1 for an example).

The survey was structured so that each respondent viewed both a non-gender-neutral and a gender-neutral ad, with the order of presentation randomly assigned with equal probability. The content of non-gender-neutral and gender-neutral versions of the ads was identical, except that the latter adhered to the protocol used by Get On Board. Specifically, the ads were crafted so that the title (e.g., "desarrollador" versus "desarrollador/a") and two sentences in the main body were presented in a masculine form for the non-gender-neutral version and in a gender-neutral form for the gender-neutral version.

Additionally, we cross-randomized two other ad variations: whether the advertised position was remote and whether it included a statement about the company's commitment to workplace diversity (a "diversity statement"). Ads under the diversity statement condi-

[16]The experiment was pre-registered with the AEA's RCT registry under number 10076.

tion had an additional sentence at the end of the first paragraph.[17] Ads under the remote condition stated "remote" saliently under the job title (as opposed to "in-person") and also re-stated that the job was remote (as opposed to as in-person) at the bottom under a "remote work policy" section. See Appendix G for a full description of ad variations and their text.

The diversity statement and remote status variations were independently assigned with a 50% probability each time a respondent viewed an ad.[18] This factorial ($2\times2\times2$) design achieves two goals. First, it ensures the sample better reflects the context as many Get On Board ads have diversity statements and involve remote positions. Second, it allows us to compare the effects of gender-neutral language to those of diversity statements, an intervention studied by previous papers (Ibañez and Riener, 2018, Leibbrandt and List, 2018, Flory et al., 2021), and of a valuable workplace amenity.

The experiment was not intended to estimate treatment interactions and may lack the statistical power to do so. Indeed, our AEA pre-registration states the goal of the experiment was to compare the effect of gender-neutral language to that of diversity statements and remote status, and not to estimate interactions. Appendix E provides further discussion.

Survey and outcomes. After introductory questions on graduation year, country of residence, boot camp field, and whether they had a job in the tech sector or were searching for one, respondents were shown an ad, asked the eleven questions below, shown another ad, and asked the same questions again, and the survey ended.

The first nine questions were statements with sliders for a Likert scale of 0-10 on whether they fully disagreed (0) to entirely agreed (10):

- I find this job attractive (*"Job appeal"*)
- I think this company would be a good employer (*"Good employer"*)
- I have the required qualifications for this job (*"Meet requirements"*)
- I would apply for this job if I have the required qualifications (*"Probability of applying"*)
- I think this company is looking for someone like me (*"Suitability"*)
- If I applied, I would have a high probability of being chosen (*"Probability of being chosen"*)
- I think this company offers a good salary (*"Good salary"*)

[17]Either "*At 'name of company' we are committed to diversity and do not accept any type of discrimination*" or "*'Company name' is a forthcoming company and we do not accept any type of discrimination.*"

[18]Specifically, all ads, regardless of gender-neutral status, had a 0.25 probability of being assigned to each of the four combinations of remote-by-diversity-statement status.

- I think this company offers a good work/life balance (*"Work-Life Balance"*)
- I think this company has an inclusive/diverse culture (*"Inclusive culture"*)

The final two questions asked what respondents thought was the proportion of women in the entire company and in the advertised position, with six categorical answers.[19]

As mentioned above, participants had a motivation to respond truthfully since their answers would impact future job recommendations. Kessler et al. (2019) employs a similar strategy to incentivize employers to rate resumes without deception.

Testing for spillovers. The randomization was designed to make salient the gender-neutral status of the second ad shown to respondents, compared to the first. Respondents either saw a gender-neutral ad followed by a non-gender-neutral ad, or vice versa. This sequencing makes the change in language more noticeable in the second ad. For example, finding larger effects of gender-neutral language for the second ads would support our hypothesis on treatment spillovers and is consistent with the results from the Get On Board experiment, which found larger effects for ads with a lower share of neighbor ads treated.

Summary statistics and balance. We received 546 responses (1,092 ad impressions) from approximately 2,500 invitations. Over 80% of the respondents work in the tech sector (and essentially all that do not were looking for a job in the tech sector). The median respondent took seven minutes to complete the survey, with 95% spending more than three minutes. In Section 4, we highlight results that serve as "attention checks." Table A.10 presents the summary statistics and covariate balance.[20]

3 Get on Board Experiment Results

We begin by reporting the effect of treatment assignment on the use of gender-neutral language (first-stage results). Next, we discuss intent-to-treat estimates. We then present treatment-on-treated effects, using our treatment assignment to instrument the use of gender-neutral language. We follow with additional results for applicants advancing to later hiring stages and conclude with findings on treatment effect heterogeneity, robustness, and placebo tests.

[19] Very low (0-10%), low (11-20%), relatively low (21-30%), median (31-40%), relatively high (41-50%), a majority (over 51%).

[20] Approximately 25% of respondents were alumni from the UX design boot camp and the remainder from web development. Alumni from the Chilean, Peruvian, and Mexican boot camps account for 25% of responses each. Brazilian alumni, who received the Portuguese version of the survey, account for 8% of the sample.

Effect on use of gender-neutral language. We use two classifications of whether an ad uses gender-neutral language. In both cases, ads are classified into three categories (English, Spanish gender-neutral, and Spanish non-gender-neutral). The first uses only job titles (as only these are listed in the platform when browsing and appear saliently in larger font at the top of ads). Note that the "Spanish gender-neutral" category includes both active gender neutrality (e.g. "desarrollador/a"), and passive gender neutrality (e.g. "analista").[21]

The second classification is based on the full text of the ad. We code an ad as "gender-neutral" if it complies entirely with the protocol: every noun, pronoun, article, and adjective is gender-neutral. If an ad has an English title and gender-neutral Spanish text, it is coded as "Spanish gender-neutral." Both classifications were done by the researchers separately from the implementation of the treatment by Get On Board.

Table 1 provides the number of ads by gender-neutral language categories and treatment status. There are four noteworthy points. First, about half of all ads use a job title in English (e.g., "designer" instead of "diseñador"), but over 85% of ads have their text in Spanish. Second, some firms choose to use gender-neutral language on their own and thus 25% of the control ads have their full text in Spanish gender-neutral. Third, some treated ads are Spanish non-gender-neutral, as the Get On Board staff did not perfectly implement the treatment. This is rare for job titles but more common for the full text, in particular sections that were not as salient such as the company description. Fourth, more ads are classified as Spanish gender-neutral by their full text than by their title only, since an ad with an English job title and Spanish gender-neutral text is classified as "English" by their title and "Spanish gender-neutral" in the full text.

Since English has non-gendered grammar, the overall first-stage estimate (treatment effects on gender-neutral language) can be inferred from subtracting control from treatment percentages in the "Spanish not GN" column in Table 1. For job titles, this figure is 33.4 p.p. and the magnitudes for the full-text classification are similar (31.4 p.p.). We return to the estimation of first-stages when we discuss treatment-on-treated (2SLS) results.

Machine Learning Identifies Share of Neighbor Ads Treated as Key Predictor of Treatment Effect Heterogeneity. We examine treatment effect heterogeneity by the share of neighbor ads treated (SNT_i, see Section 2.1). This choice is driven by our focus on spillovers, scalability, and underlying mechanisms. However, we also confirm the importance of SNT_i for treatment effect heterogeneity with causal forests (Athey et al., 2019). When applied to our data, it finds that SNT_i has the highest "variable importance" among available covariates in predicting heterogeneity in the treatment effect on the share of female applicants (Figure 3). "Variable importance" indicates how frequently a variable is used in tree splits.

[21] In Spanish, some nouns in male and female form are spelled the same. For example, "analista" refers to both a male of female analyst (see Appendix A).

A common caveat in interpreting it as a driver of treatment effect heterogeneity is that if two covariates are correlated, the trees may split on one but not the other, even if both are relevant in the data-generating process. However, this is not a concern for SNT_i since it is random and uncorrelated with other covariates included in the causal forest. Appendix D provides additional discussion and results.

Estimation: intent-to-treat effects and spillovers. Our main estimating equation is:

$$
\begin{aligned}
y_i = \alpha + \beta_0 T_i + \beta_M T_i \cdot MidQuartiles_i^{SNT} + \beta_T T_i \cdot TopQuartile_i^{SNT} + \\
+ \gamma_M MidQuartiles_i^{SNT} + \gamma_T TopQuartile_i^{SNT} + X_i'\theta + \epsilon_i \quad (1)
\end{aligned}
$$

where i indexes ads, y_i is an outcome variable (e.g., the share of female applicants), T_i is a dummy indicating the ad was assigned to treatment, and X_i is a vector of controls. $MidQuartiles_i^{SNT}$ is a dummy equal one if i's share of neighbor ads treated (SNT_i) is in the two middle quartiles of its distribution, while $TopQuartile_i^{SNT}$ is an indicator if SNT_i is in the top quartile. The parameter β_0 thus provides the intent-to-treat (ITT) effect on ads in the bottom quartile of SNT_i. The effect on ads with intermediate shares of neighbor ads treated (in the middle quartiles) is $\beta_0 + \beta_M$. The effect on ads in the top quartile of SNT_i is $\beta_0 + \beta_T$. The average treatment effect for all ads is $\beta_0 + 0.5\beta_M + 0.25\beta_H$. The parameters γ_M and γ_T capture a "direct" effect of the share of neighbor ads treated. Note the distinction between spillovers as drivers of treatment effect heterogeneity (the effect of treating i differs by SNT_i, the βs) and the "direct" treatment spillovers (SNT_i directly affects y_i, the γs).[22]

The median value of SNT_i is 0.5 and its first and third quartile are 0.34 and 0.63, respectively. Thus $MidQuartiles^{SNT} = \mathbb{1}(0.34 < SNT_i \leq 0.63)$ and $TopQuartile^{SNT} = \mathbb{1}(SNT_i > 0.63)$. Panel (a) of Figure 4 shows that the average SNT_i in the bottom and top quartiles is close to 20% and 80%, while it is (as expected) close to 50% for the medium quartiles. It also shows that had we used longer time windows to define neighbor ads instead of three days before and after (a 7-day window), the differences in the average share of neighbors treated across groups defined by quartiles would become smaller. Panel (a) of Figure 4 thus highlights that the variation in our SNT_i comes from the "small sample size" of neighbor ads in the 7-day window.[23]

The variables SNT_i, $MidQuartiles_i^{SNT}$, and $TopQuartile_i^{SNT}$ are randomly determined

[22]Externalities in other contexts, such as contagious diseases, occur primarily as "direct" spillovers. For example, Miguel and Kremer (2004) focuses on the "direct" spillovers from deworming.

[23]As the time window to define neighbor ads increase, the number of neighbors n_i for each ad i becomes larger. Since the share of neighbors treated has a binomial distribution ($SNT_i \sim B(n_i, 0.5)$), it converges to 0.5 as n_i grows. Using the baseline 7-day window, n_i varies between 1 and 24 but is between 4 and 10 for 53% of the sample.

and uncorrelated with ad characteristics. They are also orthogonal to T_i since treatment was *independently* assigned to each ad. See Section 2.1 for further discussion and Tables A.1 and A.2 for corroborating evidence. We thus estimate treatment effect heterogeneity identified from random variation in the treatment itself and the relevant dimension of heterogeneity. Intuitively, one does not have to worry if the heterogeneity in treatment effects is driven by SNT_i or some other (potentially unobservable) correlated variable, because SNT_i is random and expected to be uncorrelated with any other variable.

Throughout the paper, we report results using two sets of controls (X_i). The "baseline" includes month dummies interacted with a dummy indicating whether the ad is for a remote position, given the experiment took place as the first months of the covid-19 pandemic evolved. We also use the post-double-selection (PDS) LASSO from Belloni et al. (2014) to select controls from a set of month dummies, a dummy if the ad posted a salary range, dummies for required seniority, day-of-the-week dummies (Sunday, Monday, etc.). All these variables are further allowed to interact with a dummy for remote positions. We also include the *number* of neighbor ads.[24]

The specification in equation (1) fits the treatment spillover setting studied by Borusyak and Hull (2023). However, given the orthogonality of SNT_i to T_i and X_i, equation (1) naturally implements their recommended "recentered treatment" procedure.[25]

We report (heteroskedasticity-robust) standard errors and also two-sided randomization-inference p-values. Our randomization inference follows Borusyak and Hull (2023). For each draw of the entire assignment vector, we recalculate not only T_i but also SNT_i and the variables defined by them (i.e., all variables in equation (1) except y_i and X_i), and re-estimate equation (1). We use 1,000 draws and take as p-values the share of placebo coefficients that exceed the realized one in absolute value. This procedure takes into account the dependencies created by the spillovers we study (e.g., the treatment assigned status of ad j can affect ad i if j and i are neighbors).

Intent-to-treat results. The top panel of Table 2 reports the results from the estimation of equation (1) for our four main outcomes of interest: share of female applicants, the number of female applicants, the number of male applicants, and the average badness score (our measure of applicants' quality). The bottom panel provides the linear combination of parameters for the *implied treatment effects* for ads in different quartiles of share of neighbor ads treated

[24]Summary statistics for these variables are provided in Table A.1.

[25]The Borusyak and Hull (2023) procedure is implemented by substituting the variables in equation (1) with the differences between their realized and expected values. However, in our setting the relevant expected values are constants given that treatment was independently assigned to each ad. For example, $\mathbb{E}(T_i) = 0.5$, $\mathbb{E}(TopQuartile_i^{SNT}) = 0.25$, and $\mathbb{E}(T_i \cdot TopQuartile_i^{SNT}) = 0.5 \cdot 0.25 = 0.125$ for all i. Subtracting the expected values of variables does not affect its estimated coefficient in an OLS regression (given the inclusion of a constant and applying the Frisch-Waugh-Lovell theorem).

(SNT_i). As described above, it includes the two-sided randomization-inference p-values that account for dependencies created by spillovers. Odd columns report results using baseline controls, while even columns show results using PDS-LASSO controls.

Columns (1) and (2) show large and significant effects of treatment on the share of female applicants for ads in the bottom quartile (with a share of neighbor ads treated lower than 34%). The implied treatment effect in column (2) is 3.9 p.p. or a 27% increase relative to the control mean of 14.6%. In contrast, implied effects for ads in the middle and top quartiles ($\beta_0 + \beta_M$ and $\beta_0 + \beta_t$) are negative, smaller in magnitude, and not statistically significant. The top panel indicates the differences between the effect for the bottom quartile and other quartiles (β_M and β_T) are themselves statistically significant.

The average effect for the entire sample ($\beta_0+0.5\beta_M+0.25\beta_H$) based on column (2) is 0.0002 p.p. (SE=0.0068). Thus a 95% confidence interval does not include effects with a magnitude of 1.3 p.p. or larger for the whole sample. Figure A.2 presents the distribution of the share of female applicants for ads in different quartiles of the SNT_i distribution, indicating that the effect for those in the bottom quartile is present throughout the distribution (a "right-shift" in the cumulative distribution of treated versus control ads). See Appendix D for further discussion. The direct spillovers (γs) are smaller in magnitude and we cannot reject they are equal to zero.

Columns (3) to (6) of Table 2 report results for the inverse hyperbolic sine of the number of female and male applicants. Although noisily estimated (the number of female and male applicants are outcomes with larger variance than the share of female applicants), the point estimates indicate a percent increase in the number of female applications that is 2.5 times larger than the reduction in male applications. We thus interpret our results as being primarily driven by more women applying to treated ads in the bottom quartile.[26]

Columns (7) and (8) report treatment effects on the average quality of applicants (as measured by badness scores) that are close to zero, regardless of the share of neighbor ads treated. The default badness score set for a new user is 1500. To facilitate exposition, we re-scale badness scores by dividing them by one hundred, so it has a mean of 15.06 and a standard deviation of 1.92 across all applicants in our sample. Thus even the significant effect for ads in the top quartile has a small magnitude (less than 0.09 of a standard deviation). Figure A.3 and A.4 show the distribution of applicants' badness scores by gender. Male and

[26]We use inverse hyperbolic sines since 27% of ads in our sample have zero female applicants. Thus our estimates are weighed averages of extensive and intensive margin effects. For ads in the bottom quartile, the intensive margin (effect on a dummy if at least one woman applied) is 4.5 p.p. (SE=3.6). The extensive margin is 0.10 (SE=0.13), estimated using log(number of female applicants) as the outcome while dropping ads with zero female applicants. Both are estimated using the right-hand side from column (4). Only 6 (out of 2,201) ads have zero male applicants and thus effects for male applicants are essentially the same when using logs. Using inverse hyperbolic sines and/or logs is appropriate since the distribution of the number of applicants is right-skewed (Section 2.1).

female quality distributions in control and treated ads are remarkably similar, indicating no effects at different points of the distribution (e.g., treatment does not increase applications for particularly high- or low-quality applicants of either gender). These patterns hold for each quartile of SNT_i.

Given a positive effect on the share of female applicants for ads in the bottom quartile of SNT_i, these results suggest that treatment increases the share of women applying without affecting the quality distribution of applicants, indicating that the larger share of female applicants comes from across the quality spectrum. This implies effects on the share of female applicants at any given quality threshold. Intuitively, firms that only consider applicants with badness scores above a certain cutoff would see a larger share of female applicants above that cutoff as a result of the treatment, for any cutoff. See Appendix D for further discussion.

Treatment-on-treated effects. To interpret effects' magnitudes, we estimate treatment-on-treated (ToT) effects of gender-neutral language in the following 2SLS framework:

$$y_i = \alpha^{2SLS} + \beta_0^{2SLS} GN_i + \beta_M^{2SLS} GN_i \cdot MidQuartiles_i + \beta_T^{2SLS} GN_i \cdot TopQuartile_i + \\ + \gamma_M^{2SLS} MidQuartiles_i + \gamma_T^{2SLS} TopQuartile_i^{SNT} + X_i' \theta^{2SLS} + \epsilon_i \quad (2)$$

where y_i is the share of female applicants and GN_i is a dummy equal to one if the full text of ad i is gender neutral.[27] The three endogenous variables are GN_i and its interactions with $MidQuartiles_i^{SNT}$ and $TopQuartile_i^{SNT}$ and the three excluded instruments are the treatment dummy (T_i) and its interaction with $MidQuartiles_i^{SNT}$ and $TopQuartile_i^{SNT}$.

Table 3 presents the treatment-on-treated (2SLS) effect of gender-neutral language for ads with share of neighbor ads treated (SNT_i) falling in the bottom quartile, middle quartiles, and top quartile of the SNT_i distribution: β_0^{2SLS}, $\beta_0^{2SLS} + \beta_M^{2SLS}$, and $\beta_0^{2SLS} + \beta_T^{2SLS}$, respectively. The effect for the bottom quartile is 10.4 p.p. or 11.5 p.p., depending on the controls used, and significant at the 5% level. This a sizable effect, representing more than a 70% increase over the control mean. However, we highlight that more modest effects (such as a 3 p.p. increase) are also included in the 95% confidence interval. The effects for the middle and top quartiles are, as expected, negative and statistically insignificant.[28] The average effect for the entire sample ($\beta_0^{2SLS} + 0.5\beta_M^{2SLS} + 0.25\beta_H^{2SLS}$) based on column (2) is -0.0015 p.p.

[27]Effects based on the gender-neutrality of ad *titles* are similar given that the first-stage on titles and full-text classifications of are similar (Table 1). Later in this section, we discuss evidence suggesting that the use of gender-neutral language in the text ad itself, and not only the titles, drives the results.

[28]While the ITT effect for middle quartiles is not significant (Table 2), the treatment-on-treated effect is significant at the 10% level. While this may appear puzzling, there is not necessarily a relationship between the significance of reduced form and 2SLS estimates. See Appendix A of Lochner and Moretti (2004) for a formal argument. Estimating 2SLS effects for the number of applicants is less informative given the outcome has a larger variance and is less precisely estimated.

(SE=0.0022).[29]

Effects on selected and hired candidates. As discussed in Section 2.1, companies may use an evaluation board provided in the Get On Board platform to assist with their selection process. It allows companies to sort candidates into different categories: "discarded," "selected," and "hired." However, not all companies use the evaluation board and we observe which candidates advance in the selection process for only a subset of ads.

Table 4 thus replicates our main ITT results (columns (1) and (2) of Table 2) but also results on the share of female candidates that firms sort as "not discarded," "selected," or "hired". With caveats about selection into using the board and smaller sample sizes, results are consistent with a higher share of women moving forward on the selection process for ads treated and with a share of neighbor ads treated in the bottom quartile. In particular, the effect on the share of female candidates "not discarded" is 4.6 p.p.for the bottom quartile (with smaller and insignificant effect for the middle and top quartiles). We also observe a large effect on the share of female applicants actually hired, although this is imprecisely estimated and based on less than a quarter of all ads in the sample.[30]

Overall, we find suggestive evidence that effects on female shares of applicants translate into higher female shares up to the final stages of the selection process. This is consistent with female under-representation in the tech sector stemming from women *not applying* to certain positions, which bolsters the policy relevance of using ad language that increases female representation in the applicant pool.

Additional results. Table A.6 investigates treatment effect heterogeneity among two dimensions. First, it replicates our main ITT estimates (Table 2) separately for ads with titles in English and in Spanish. As shown in Table 1, approximately half of the ads has a job title in English (e.g., "programmer" instead of "programador" or "programadora/o"). Among these, 80% have their main text in Spanish. Moreover, English titles are, by default, gender neutral: "programmer" refers to both male and female programmers and our protocol indicated an ad with a title in English should thus not be edited.

Thus, by exploring treatment heterogeneity by title language, we can test if the effects are driven by only changing the title or the text of the entire ad. The results in columns

[29]The estimation of the coefficients in equation (2) as well as the three related first-stage regressions are provided in Table A.5 and discussed in Appendix D.

[30]For each category, we define the share of female applicants in the category and only include in the sample ads where we can observe the firm using the evaluation board for the category (labeling at least one candidate). For example, columns (5)-(6) use as the outcome the share of female candidates among those labeled "selected" and only have 774 observations since only this number of ads had at least one candidate labeled as "selected." The sample sizes does decreases along the selection process: starting from a total sample of 2,201 ads to 1,714 "not discarded," 774 "selected," and 508 "hired."

(1)-(4) of Table Table A.6 suggest similar effects for ads with texts in English or Spanish, indicating that the use of gender-neutral in the main text of the ad plays a role.

Second, columns (5)-(8) explore whether effects differ by ads that were remote or not, and also do not find a clear pattern of heterogeneity. Table A.7 provides evidence that being assigned to treatment does not affect subsequent behavior on the platform: treatment does not increase the number of future ads posted or the chance firms choose, on their own, to use gender-neutral language on subsequent ads. See Appendix F for further discussion.

Robustness checks and placebo tests. Figure 4 examines how our main ITT results (equation 1) are influenced by different time windows used to define neighbor ads. Our baseline specification considers ad i's neighbors to be all other ads in the same job title group posted three days before or after ad i. Panel (b) shows that for ads in the bottom quartile of the SNT_i distribution, using a five- or seven-day window yields similar results. However, as the time window increases, the effects converge to zero. For the middle and top quartiles in panels (c) and (d), the effect is not statistically significant regardless of the window used.

The pattern for the bottom quartile (panel b) supports our interpretation of the results. Applicants see ads posted around the same time together, so spillovers from ads posted 3-7 days before or after are more relevant than those posted 15-30 days before or after. Additionally, as the time window increases, the difference in SNT_i between quartiles diminishes as the number of neighbors for each ad increases. This indicates that differences in SNT_i across quartiles indeed drive the results.[31]

Tables A.8 and A.9 present placebo tests. Table A.8 re-estimates the main ITT results from equation (1) and Table 2, but defines SNT_i based on "future" neighbors. In columns (1)-(2), SNT_i is defined "30 days ahead"; instead of being based on ads in the same job title group posted 3 days before or after ad i, it is based on ads in the same job title group posted 27 to 33 days after. Columns (3)-(4) perform a similar "60 days ahead" exercise. The results indicate there is no treatment heterogeneity by "future SNT_i."

Table A.9 replicates the main ITT results, but instead of exploring heterogeneity in SNT_i, it examines heterogeneity based on the female representation in the job title group. We focus on this dimension of heterogeneity because it is the second most important factor identified in our causal forest analysis (Figure 3) and because the gender composition of an occupation can be predictive of gender bias (Galos and Coppock, 2023). We do not find treatment effect heterogeneity across this dimension. Appendix D provides further details on the implementation and interpretation of the two placebo tests.

[31] As previously discussed in this section, SNT_i has a binomial distribution ($SNT_i \sim B(n_i, 0.5)$) which converges to 0.5 as the number of neighbors n_i grows, as depicted in Panel (a).

Tables A.8 and A.9, along with Figure 4, support the conclusion that the (randomly assigned) share of treated ads in the job title group that were posted around similar dates drives heterogeneity in treatment effects, rather than fixed characteristics of job title groups.

Interpretation of results. The results highlight the key role of treatment effect spillovers. Gender-neutral language in ads significantly increases the share of female candidates applying when likely listed next to a few other gender-neutral ads. However, when ads are among a larger share of gender-neutral ads, the effects are zero or negative, and likely zero if most ads were treated.

These results are consistent with applicants using gender-neutral language as a signal to infer job characteristics. However, as gender-neutral language becomes more common from the point of view of the applicant, this signal may lose its informativeness. The Laboratoria experiment, discussed in the next section, directly tests whether gender-neutral language in ads influences applicants' beliefs about the firm and the position.

4 Laboratoria Experiment Results

We start by discussing straightforward mean comparisons that pool both ads shown first or second to respondents. We then explore the heterogeneity by ad order (and its implications for estimating treatment spillovers), and conclude the section discussing potential experimenter demand effects.

"Raw" averages. Figure 5 provides simple averages for all eleven outcomes described in Subsection 2.2. It does so separately for the three treatments. Since the experiment has a $2 \times 2 \times 2$ factorial design with equal probability, other treatment conditions are balanced when making two-way comparisons.[32]

Positive impacts of using gender-neutral language are visible for all outcomes, with one exception. Gender-neutral language makes subjects report they are 10% more likely to apply for a job (a 0.54-point increase over a control mean of 5.2 on a 0-10 Likert scale). Similarly, it makes respondents report they are 16% more "suitable" for the job (agree the company is "looking for someone like me") and 7% more likely to be hired. Moreover, gender-neutral language increases beliefs about the company's inclusive culture and promotion of work-life balance by 25% and 10%, respectively. It also makes respondents believe the company is

[32]For example, when comparing gender-neutral to non-gender-neutral ads, in both groups the share of ads that are remote and have a diversity statement is 25%, that is non-remote and has a diversity statement is 25%, and so on.

more likely to employ a larger share of women. All these effects are statistically significant at the 5% level, and most at the 1% level.[33]

The effect on respondents stating they meet requirements is small and close to zero. This is consistent with gender-neutral language leading respondents to update their beliefs about the company, but not on whether they meet requirements clearly specified in the ad.

The impacts of diversity statements are closer to zero, though large for beliefs about the firms' culture of inclusiveness, indicating the statements were not ignored by respondents. This suggests that gender-neutral language sends stronger signals about the company than explicit statements. For five outcomes (job appeal, suitability, good salary, and percent of women in the position and company), we can reject the hypothesis that the effect of gender-neutral language and diversity statements are the same at the 5% significance level.[34]

The impact of remoteness is significant and larger than the use of gender-neutral language for some outcomes. It increases the appeal of the job and views about the company's culture and work-life balance, but not whether the respondents meet requirements, are likely to be hired, or believe more women work in it. The effects of gender-neutral language are larger for suitability for the job, inclusive culture, and the percent of women in the company and position, while remote status has a larger effect on views about work-life balance (for these five outcomes, we can reject the hypothesis that the effect of gender-neutral language and remote status are the same at the 5% level).

Appendix E presents the cumulative distribution functions (CDF) for each of the eleven outcomes by the three different treatment statuses, essentially replicating for CDFs what Figure 5 does for averages (Figures A.5, A.6, and A.7). In cases we find effects on averages, they are driven by broad changes throughout the distribution of outcomes (e.g., a broader "right shift" in the CDF). Appendix E also provides the table counterpart of Figure 5 (Table A.11) and also replicates it splitting the sample by whether the respondents are alumni of the web development or the UX design boot camps (Tables A.12 and A.13). Results are similar in magnitude, suggesting little heterogeneity by field. Table A.14 replicates Table A.11 adding respondent fixed effects. As expected given the experimental design, these within-estimates are quite similar to other estimates. Appendix E also discusses the interpretation of the results in light of recent research on factorial designs (Muralidharan et al., 2023).[35]

[33]Throughout this section, we use heteroskedasticity-robust standard errors for inference. We obtain similar p-values when using randomization inference based on 1,000 draws, but we omit them from the figures and tables here and in Appendix E to economize on space.

[34]The same applies to the probability of applying at the 10% level.

[35]In unreported regressions, we find that the results are also robust to excluding the Brazilian boot camp alumni (who answered a version of the survey in Portuguese) and excluding respondents that answered the survey "too quickly" (e.g., less than three or five minutes).

Estimating equation and spillovers. Our main estimating equation is:

$$\begin{aligned} y_{ia} = \alpha_1 + \alpha_2 2ndAd_{ia} + \beta_1 GNeutral_{ia} + \beta_2 GNeutral_{ia} \times 2ndAd_{ia} + \\ + \gamma_1 Diversity_{ia} + \gamma_2 Diversity_{ia} \times 2ndAd_{ia} + \\ + \delta_1 Remote_{ia} + \delta_2 Remote_{ia} \times 2ndAd_{ia} + \epsilon_{ia} \end{aligned} \tag{3}$$

where i indexes respondents and a indexes the ads they see. Each respondent sees two ads and thus with 546 respondents we have up to 1092 observations to be used. y_{ia} is an outcome variable (e.g., whether respondent i answered she would apply to job ad a). $GNeutral_{ia}$, $Diversity_{ia}$, and $Remote_{ia}$ are dummies indicating whether the ad a shown to i was randomly assigned to be gender-neutral, have a diversity statement, or advertise a remote position, respectively. $2ndAd_{ia}$ is a dummy indicating whether the ad is the second one seen by respondent i. Thus, β_1 provides the effect of using gender-neutral language in the first ad, and $\beta_1 + \beta_2$ provides the effect for the second ad. The γs and δs parameters provide analogous effects of diversity statements and remote status. α_2 provides the effect of being the second ad for an ad assigned to non-gender-neutral, non-remote, and without a diversity statement.

As discussed in Subsection 2.2, the randomization was designed to highlight the gender-neutral status of the second ad compared to the first. Since respondents either saw a gender-neutral ad followed by a non-gender-neutral ad, or vice versa, the change in gender-neutral language is more noticeable in the second ad. Given this, we interpret a positive β_2 as evidence of spillovers: the effect of gender-neutral language is stronger when the respondent just saw a non-gender-neutral ad before, compared to when they first see a gender-neutral ad and evaluate it without being provided a clear comparison ad.[36]

The results from equation (3) differ from the previously discussed results from Figure 5 on two dimensions. First, Figure 5 provides two-way comparisons of means, while equation (3) estimates the effects of our three treatments jointly. This decision makes a negligible difference, as expected from a factorial design that ensures the three treatments are uncorrelated with each other.[37] Second, and more importantly, it allows us to estimate the effects of first and second ads separately.

Table 5 provides the results. Overall, it shows that the effects of gender-neutral language

[36] Given the independent draws for diversity statements and remote status (Subsection 2.2), γ_2 and δ_2 do not have a similar interpretation. For example, half of the respondents exposed to a remote second ad also saw a remote first ad.

[37] Moreover, the factorial design makes it so that "contamination bias" from multiple treatments is not an issue for our estimates (Goldsmith-Pinkham et al., 2022). Such bias arises from cases where treatments are correlated with each other (e.g., not independently drawn, such as when the design is not factorial and units receive either one treatment or another) and including covariates (such as strata fixed effects) are required in estimation. Neither of these situations applies to our design.

are substantially larger for second ads when compared to first ads: β_2 is positive and significant for nine (out of eleven) outcomes.[38] This pattern suggests the presence of spillovers of gender-neutral language, similar to the Get On Board results. The effects of gender-neutral language are stronger when respondents previously saw an ad with a different status, relative to when they evaluate the first ad without being provided a clear comparison ad.

No similar pattern is present for the diversity statement and remote treatments. In the cases it does have an effect, it is similar for both the first and second ad the respondent sees (i.e., γ_2 and δ_2 are relatively small we cannot reject they are zero). These provide a "placebo test," indicating it is not the case that all effects are simply stronger for second ads for reasons unrelated to spillovers.

Experimenter demand effects. Five factors suggest experimenter demand effects cannot explain our results. First, as described in Section 2, subjects had no reason to believe the experiment involved evaluating gender-neutral language (or that ad texts varied randomly). They saw different ads without knowing what were the possible variations and treatments. Second, the small and insignificant effect of gender-neutral language for meeting requirements for the job provides evidence against demand effects or any other mechanism leading respondents to give higher ratings for all outcomes. Third, we find small or zero effects of diversity statements. Presumably, any demand effects mechanism that operates for gender-neutral language would also operate for related treatments. Fourth, it is unclear why experimenter demand effects would create stronger effects of gender-neutral language on the second ad (while not doing the same for the remote and diversity statement treatments). Fifth, respondents had an incentive to respond with their true evaluations since their answers would impact the future job recommendations they received from Laboratoria.

Interpretation of results. Overall, our results are consistent with respondents interpreting the use of gender-neutral language as a signal that the firm is a more appealing employer. Indeed, the only outcome that is not affected is a question that does not involve beliefs about employer characteristics (whether respondents believe they meet the requirements for the job, which are clearly explained in the ad). Given we observe effects for almost all outcomes we study, the results do not shed light on which firm characteristics and job amenities respondents update the most about. The substantially larger effects for the second ads corroborate the importance of spillovers, as the effect of gender-neutral language is stronger after respon-

[38] These effects are significant at the 1% level, with one exception: the probability of being chosen, significant at the 10% level. Of the two outcomes where β_2 is not statistically distinct from zero, one is "meet requirements" which, as previously discussed, is not affected by gender-neutral language. Only one outcome ("suitability") presents a pattern consistent with the effect being the same on the first and second ads. As a graphical counterpart, Figures A.8 and A.9 replicate Figure 5 for first ads and second ads only.

dents saw a non-gender-neutral ad (compared to the first ads, which respondents evaluate without a clear comparison ad).

5 Conclusion

This paper provides, to our knowledge, the first evaluation of gender-neutral language in job ads and the first exploration of treatment spillovers in interventions that alter the language or content of recruitment materials. Our results suggest that gender-neutral ads attract more female applicants when a small proportion of other ads concurrently considered by applicants are also gender-neutral. However, this effect would likely substantially diminish and even become zero if all or most ads were gender-neutral.

Studying spillovers is crucial for scalability. Our results suggest that a smaller-scale experiment treating only a fraction of ads would indicate that gender-neutral language can promote diversity. However, it wouldn't reveal whether these effects would persist if a higher share of ads were treated.

In a second experiment, gender-neutral language in ads influenced beliefs about job characteristics, particularly when the comparison to non-gender-neutral ads was salient. Overall, the results from both experiments are consistent with female applicants interpreting gender-neutral language as a signal, using it to update their beliefs about the firm's characteristics and job amenities.

While the overall policy conclusion on gender-neutral language may seem negative due to limited scalability, some results suggest it can positively impact diversity in certain circumstances. We find suggestive evidence that when it affects the diversity of the applicant pool, it also influences the diversity of candidates advancing in the selection process and potentially getting hired. This underscores the importance of studying interventions, especially those that are light-touch and virtually costless as the one we study, that enhance applicant pool diversity. We hope future research will further investigate this issue, including other aspects of inclusive language and contexts beyond Spanish-speaking countries.

References

Abraham, L., J. Hallermeier, and A. Stein (2024): "Words matter: Experimental evidence from job applications," *Journal of Economic Behavior & Organization*, 225, 348–391.

Athey, S., J. Tibshirani, and S. Wager (2019): "Generalized Random Forests," *The Annals of Statistics*, 47, 1148–1178.

Banfi, S. and B. Villena-Roldan (2019): "Do high-wage jobs attract more applicants? Directed search evidence from the online labor market," *Journal of Labor Economics*, 37, 715–746.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81, 608–650.

Belot, M., P. Kircher, and P. Muller (2019): "Providing advice to jobseekers at low cost: An experimental study on online advice," *The review of economic studies*, 86, 1411–1447.

Borusyak, K. and P. Hull (2023): "Nonrandom exposure to exogenous shocks," *Econometrica*, 91, 2155–2185.

Card, D., F. Colella, and R. Lalive (2024): "Gender preferences in job vacancies and workplace gender diversity," *Review of Economic Studies*, rdae085.

Chen, K. M. (2013): "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets," *American Economic Review*, 103, 690–731.

Coffman, K. B., M. R. Collis, and L. Kulkarni (2024): "Whether to apply," *Management Science*, 70, 4649–4669.

Cohen, A., T. Kricheli-Katz, T. Regev, T. Karelitz, and S. Pumpian (2023): "Gender-Neutral Language and Gender Disparities," *Available at SSRN*.

Cole, M. C., F. A. Hill, and L. J. Dayley (1983): "Do masculine pronouns used generically lead to thoughts of men?" *Sex Roles*, 9, 737–750.

Crawford, M. and L. English (1984): "Generic versus specific inclusion of women in language: Effects on recall," *Journal of Psycholinguistic Research*, 13, 373–381.

Danziger, S. and R. Ward (2010): "Language Changes Implicit Associations Between Ethnic Groups and Evaluation in Bilinguals," *Psychological Science*, 21, 799–800.

Del Carpio, L. and M. Guadalupe (2021): "More Women in Tech? Evidence from a Field Experiment addressing Social Identity," *Management Science*, 68.

Delfino, A. (2024): "Breaking gender barriers: Experimental evidence on men in pink-collar jobs," *American Economic Review*, 114, 1816–1853.

Fitsimons, G. M. and A. C. Kay (2004): “Language and Interpersonal Cognition: Causal Effects of Variations in Pronoun Usage on Perceptions of Closeness,” *Personality and Social Psychology Bulletin*, 30, 547–557.

Flory, J. A., A. Leibbrandt, and J. A. List (2015): “Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions,” *The Review of Economic Studies*, 82, 122–155.

Flory, J. A., A. Leibbrandt, C. Rott, and O. Stoddard (2021): “Increasing Workplace Diversity: Evidence from a Recruiting Experiment at a Fortune 500 Company,” *Journal of Human Resources*, 56, 73–92.

Galos, D. R. and A. Coppock (2023): “Gender composition predicts gender bias: A meta-reanalysis of hiring discrimination audit experiments,” *Science Advances*, 9, eade7979.

Gastil, J. (1990): “Generic Pronouns and Sexist Language: The Oxymoronic Character of Masculine Generics,” *Sex Roles*, 23, 629–643.

Gaucher, D., J. Friesen, and A. C. Kay (2011): “Evidence that gendered wording in job advertisements exists and sustains gender inequality,” *Journal of Personality and Social Psychology*, 101, 109–128.

Gee, L. K. (2018): “The More You Know: Information Effects on Job Application Rates in a Large Field Experiment,” *Management Science*, 65, 2077–2094.

Goldsmith-Pinkham, P., P. Hull, and M. Kolesár (2022): “Contamination bias in linear regressions,” Tech. rep., National Bureau of Economic Research.

Harrison, G. W. and J. A. List (2004): “Field experiments,” *Journal of Economic literature*, 42, 1009–1055.

Helleseter, M. D., P. Kuhn, and K. Shen (2020): “The age twist in employers’ gender requests: Evidence from four job boards,” *Journal of Human Resources*, 55, 428–469.

Hellinger, M. and H. Bussmann (2015): “Gender across languages: The linguistic representation of women and men,” *Gender across languages*, 1–26.

Ibañez, M. and G. Riener (2018): “Sorting through affirmative action: Three field experiments in Colombia,” *Journal of Labor Economics*, 36, 437–478.

Jakiela, P. and O. Ozier (2018): “Gendered Language,” *World Bank Policy Research Working Paper No. 8464*.

Kerwin, J., N. Rostom, and O. Sterck (2024): "Striking the Right Balance: Why Standard Balance Tests Over-Reject the Null, and How to Fix It," Tech. rep., Institute of Labor Economics (IZA).

Kessler, J. B., C. Low, and C. D. Sullivan (2019): "Incentivized resume rating: Eliciting employer preferences without deception," *American Economic Review*, 109, 3713–3744.

Kuhn, P. and K. Shen (2023): "What happens when employers can no longer discriminate in job ads?" *American Economic Review*, 113, 1013–1048.

Kuhn, P. J., K. Shen, and S. Zhang (2020): "Gender-targeted job ads in the recruitment process: Facts from a Chinese job board," *Journal of Development Economics*, 147, 1025–1031.

Leibbrandt, A. and J. A. List (2015): "Do women avoid salary negotiations? Evidence from a large-scale natural field experiment," *Management Science*, 61, 2016–2024.

——— (2018): "Do equal employment opportunity statements backfire? Evidence from a natural field experiment on job-entry decisions," Tech. rep., National Bureau of Economic Research.

List, J. A. (2022): *The voltage effect: How to make good ideas great and great ideas scale*, Crown Currency.

Lochner, L. and E. Moretti (2004): "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports," *American economic review*, 94, 155–189.

Mas, A. and A. Pallais (2017): "Valuing alternative work arrangements," *American Economic Review*, 107, 3722–3759.

Miguel, E. and M. Kremer (2004): "Worms: identifying impacts on education and health in the presence of treatment externalities," *Econometrica*, 72, 159–217.

Moulton, J., G. M. Robinson, and C. Elias (1978): "Reducing Women's Lack of Fit with Leadership? Effects of the Wording of Job Advertisements," *American Psychologist*, 33, 1032–1036.

Muralidharan, K., M. Romero, and K. Wüthrich (2023): "Factorial designs, model selection, and (incorrect) inference in randomized experiments," *Review of Economics and Statistics*, 1–44.

Ogunnaike, O., Y. Dunham, and M. R. Banaji (2010): "The language of implicit preferences," *Journal of Experimental Social Psychology*, 46, 999–1003.

Osborne, M. J. and A. Rubinstein (1994): *A course in game theory*, MIT press.

——— (2020): *Models in Microeconomic Theory ('She' Edition)*, Cambridge, UK: Open Book Publishers.

Pérez, E. and M. Tavitz (2019): "Language Heightens the Political Salience of Ethnic Divisions," *Journal of Experimental Political Science*, 6, 131–140.

Samek, A. (2019): "A University-Wide Field Experiment on Gender Differences in Job Entry Decisions," *Management Science*, 65, 3072–3281.

Tibshirani, J., S. Athey, R. Friedberg, V. Hadad, D. Hirshberg, L. Miner, E. Sverdrup, S. Wager, M. Wright, and M. J. Tibshirani (2024): *grf: Generalized Random Forests*, R package version 2.3.2.

Whorf, B. L. (1956): *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*, MIT press.

Figure 1: Example of Same Ad Under Control and Treatment Status

CloudSystems
April 15, 2022

Ingeniero DevOps

Remote | Full time | SysAdmin / DevOps / QA

Somos CloudSystems, empresa líder en la provisión de soluciones de nueva generación basadas en la nube, con aplicaciones de contabilidad, nómina y factura electrónica para pequeñas y medianas empresas en Latinoamérica. Estamos buscando al profesional responsable de automatizar la infraestructura y herramientas de la compañía para acelerar el desarrollo de productos, su calidad y el lanzamiento de los mismos. Tenemos un entorno innovador y una cultura horizontal, y buscamos ingenieros DevOps dinámicos, con capacidad de trabajar en equipo y críticos con su trabajo.

Funciones

Buscamos ingenieros especialistas en el rol de Devops y automatización de procesos de desarrollo e infraestructura. Deberás:

· Instalar y promover la cultura DevOps bajo metodologías agile en conjunto con el equipo de desarrolladores.

· Proveer y monitorear infraestructura 100% Cloud para soportar el desarrollo de software.

· Dominar ampliamente los mejores estándares de automatización de pipelines CI / CD.

Requisitos

· Ingeniero de Sistemas, Programación o carreras afines.

· Experiencia relevante y comprobable de al menos 3 años.

· Herramientas para creación de pipelines CI/CD: Jenkins, GitLab

· Experiencia con sistemas operativos: Unix / Linux

· Conocimiento en plataformas Cloud: Oracle, AWS, Azure

· Manejo de contenedores: Docker o Kubernetes.

· Experiencia trabajando con desarrolladores en metodologías agile (Scrum, Kanban).

· Internet velocidad mínima de bajada: 500 Mbps y de subida: 10 Mbps y espacio aislado de ruido para trabajar remotamente.

Deseables

· Experiencia con SQL Server, PostgreSQL y NoSQL

· Manejo de control de versiones de código: GIT

Beneficios

· Sueldo competitivo

· Bono de conectividad para trabajo 100% Remoto. Cambia de proveedor o trabaja desde el mejor cowork en tu ciudad.

· Horario flexible

· Día de cumpleaños libre

· Bono/Aguinaldo Fiestas Patrias y Navidad

Flexible hours
Flexible schedule and freedom for attending family needs or personal errands.

Paid sick days
Sick leave is compensated (limits might apply).

Vacation on birthday
Your birthday counts as an extra day of vacation.

Remote work policy

Fully remote
Candidates can reside anywhere in the world.

Agile · Amazon Web Services · Azure · CI/CD · Cloud Computing · Continuous Integration · DevOps · Docker · Jenkins · Kanban · Kubernetes · Linux · Oracle · Scrum · Virtualization

CloudSystems
April 15, 2022

Ingeniera/o DevOps

Remote | Full time | SysAdmin / DevOps / QA

Somos CloudSystems, empresa líder en la provisión de soluciones de nueva generación basadas en la nube, con aplicaciones de contabilidad, nómina y factura electrónica para pequeñas y medianas empresas en Latinoamérica. Estamos buscando a la o el profesional responsable de automatizar la infraestructura y herramientas de la compañía para acelerar el desarrollo de productos, su calidad y e lanzamiento de los mismos. Tenemos un entorno innovador y una cultura horizontal y buscamos ingenieras/os DevOps dinámicas/os, con capacidad de trabajar en equipo y críticas/os con su trabajo.

Funciones

Buscamos ingenieras e ingenieros especialistas en el rol de Devops y automatizació de procesos de desarrollo e infraestructura. Deberás:

· Instalar y promover la cultura DevOps bajo metodologías agile en conjunto con el equipo de desarrolladores.

· Proveer y monitorear infraestructura 100% Cloud para soportar el desarrollo de software.

· Dominar ampliamente los mejores estándares de automatización de pipelines CI / CD.

Requisitos

· Estudios universitarios en Ingeniería de Sistemas, Programación o carreras afines.

· Experiencia relevante y comprobable de al menos 3 años.

· Herramientas para creación de pipelines CI/CD: Jenkins, GitLab

· Experiencia con sistemas operativos: Unix / Linux

· Conocimiento en plataformas Cloud: Oracle, AWS, Azure

· Manejo de contenedores: Docker o Kubernetes.

· Experiencia trabajando con desarrolladoras y desarrolladores en metodologías agile (Scrum, Kanban).

· Internet velocidad mínima de bajada: 500 Mbps y de subida: 10 Mbps y espacio aislado de ruido para trabajar remotamente.

Deseables

· Experiencia con SQL Server, PostgreSQL y NoSQL

· Manejo de control de versiones de código: GIT

Beneficios

· Sueldo competitivo

· Bono de conectividad para trabajo 100% Remoto. Cambia de proveedor o trabaja desde el mejor cowork en tu ciudad.

· Horario flexible

· Día de cumpleaños libre

· Bono/Aguinaldo Fiestas Patrias y Navidad

Flexible hours
Flexible schedule and freedom for attending family needs or personal errands.

Paid sick days
Sick leave is compensated (limits might apply).

Vacation on birthday
Your birthday counts as an extra day of vacation.

Remote work policy

Fully remote
Candidates can reside anywhere in the world.

Agile · Amazon Web Services · Azure · CI/CD · Cloud Computing · Continuous Integration · DevOps · Docker · Jenkins · Kanban · Kubernetes · Linux · Oracle · Scrum · Virtualization

Figure 2: Example of Neighbor Ads

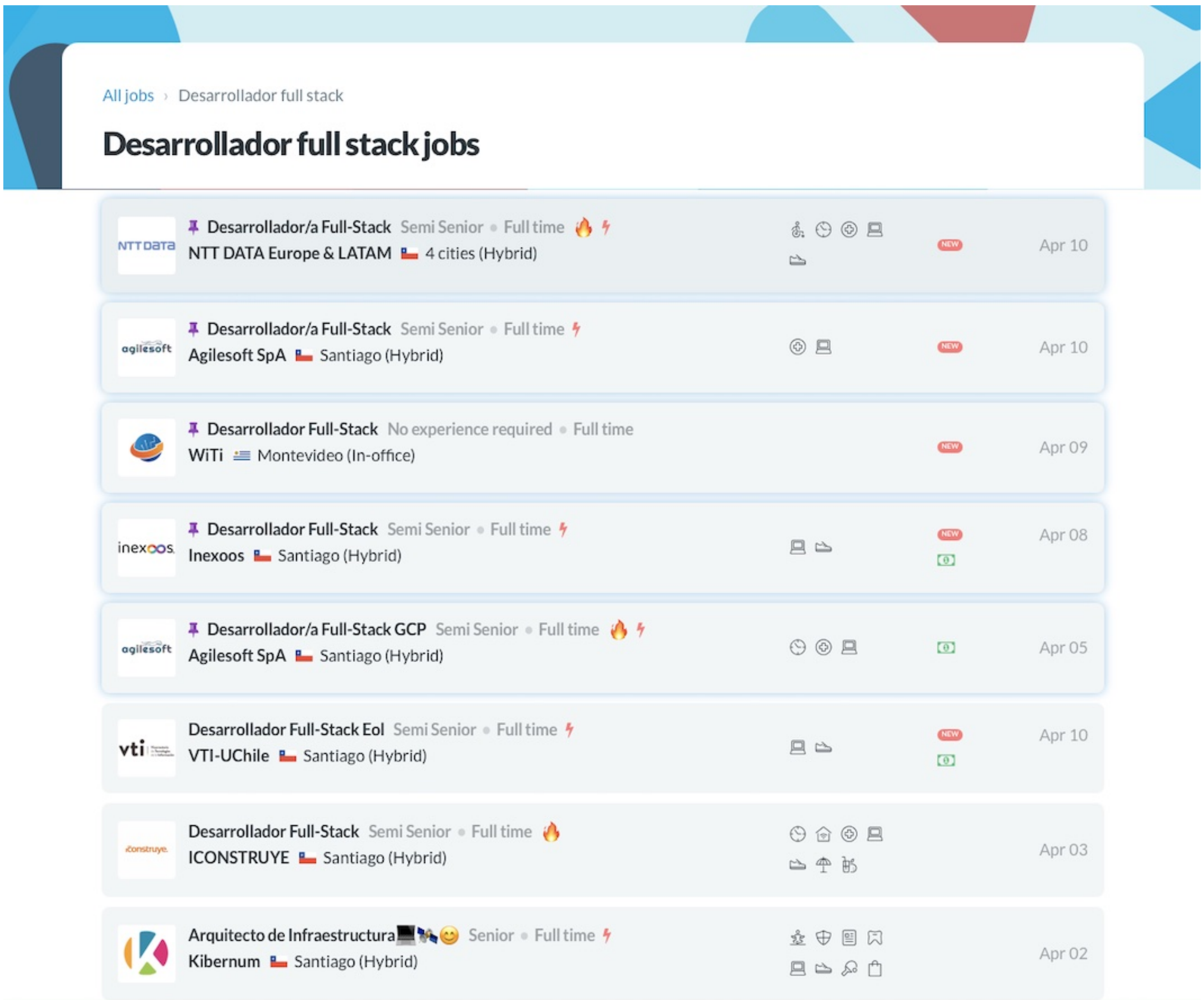
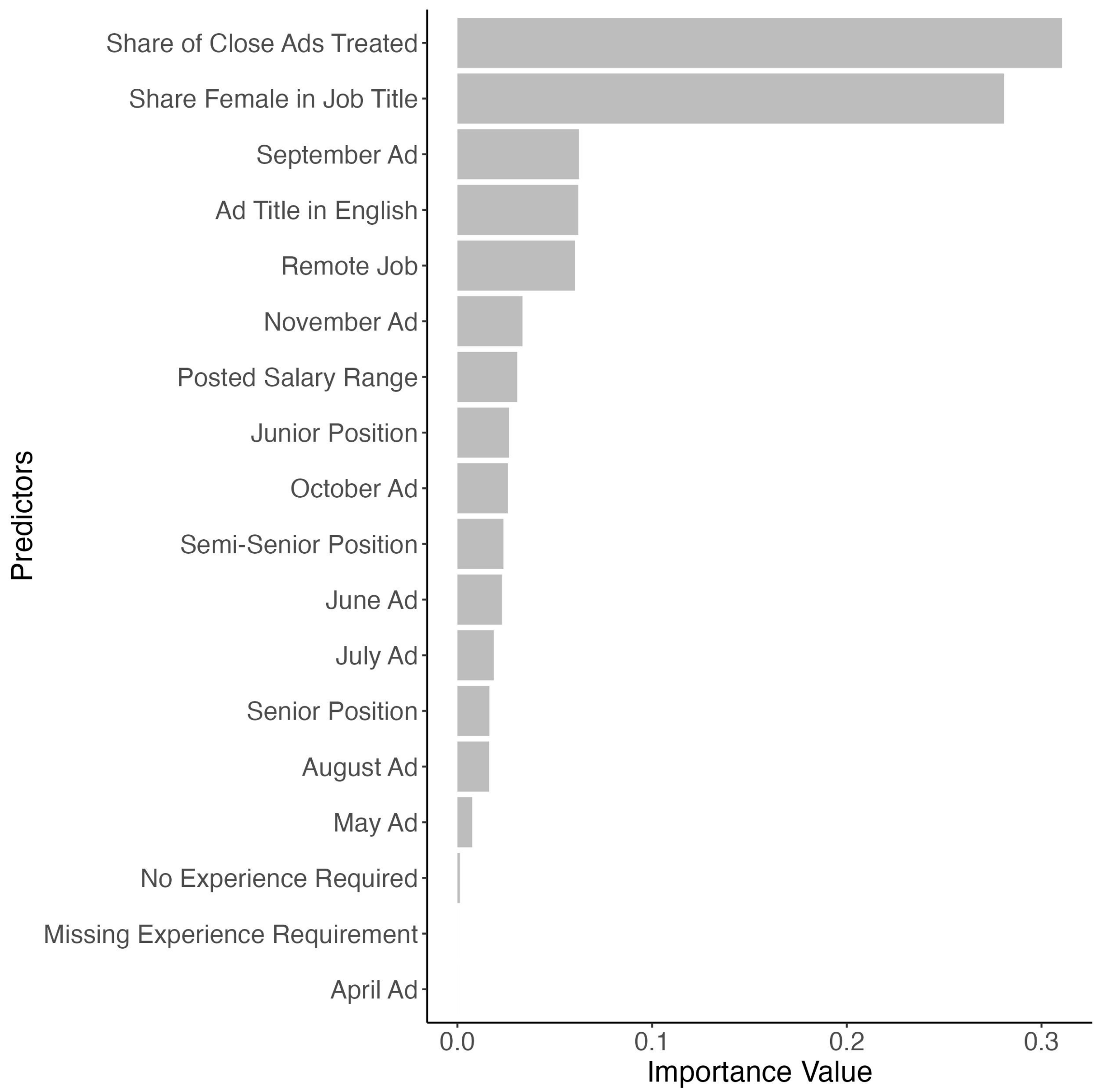
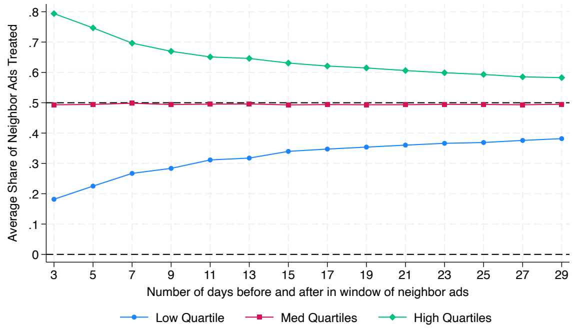


Figure 3: Covariates' Importance

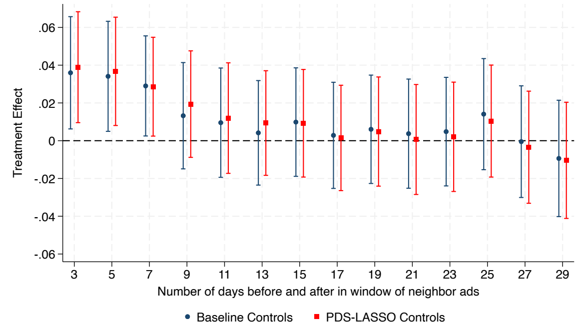


Notes: The unit of observation is an ad (2,201 observations). The figure provides the "variable importance" of each covariate used to fit a causal forest (Athey et al., 2019). We use the GRF package in R (Tibshirani et al., 2024) and its "variable_importance" function, which provides a measure of how often the variable was used in tree splits. The outcome is the share of applicants to ad i that are female and we estimate heterogeneous effects of assigned treatment (an intent-to-treat analysis). The set of covariates that can potentially predict effect heterogeneity include an indicator if the ad title is in English, a set of month dummies, the share of female applicants in the job title group (constructed only using the control group), and all variables listed in Table A.1 (except the minimum and maximum of salary range, which is missing for ads that did not post a range). See Appendix D for further information.

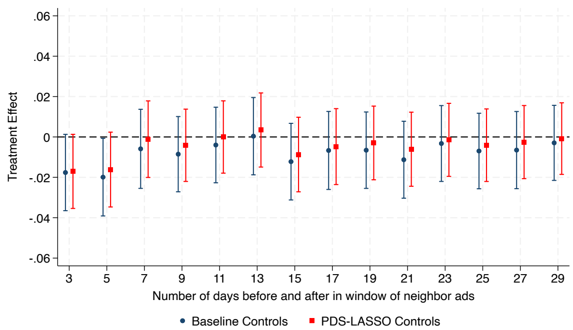
Figure 4: Treatment Effects for Different Time Windows Used in Defining Neighbor Ads - Get On Board



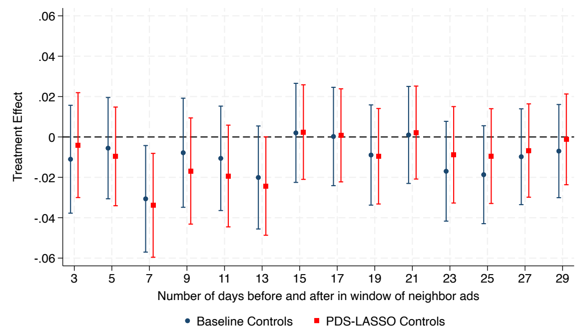
(a) Avg. Share of Neighbor Ads Treated



(b) Bottom Quartile of % Neighbor Ads Treated



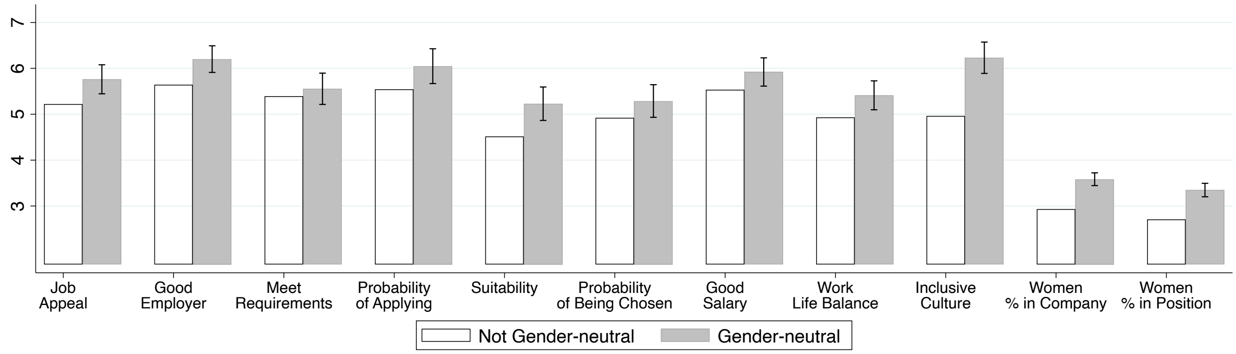
(c) Mid Quartiles of % Neighbor Ads Treated



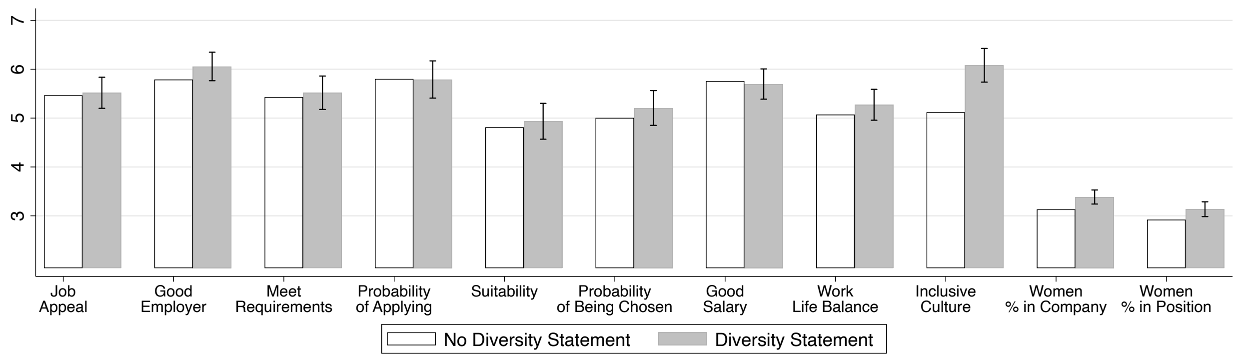
(d) Top Quartile of % Neighbor Ads Treated

Notes: The unit of observation is an ad. Panel (a) shows the average share of neighbor ads treated (SNT_i) in the bottom quartile, middle quartiles, and top quartile of the SNT_i distribution for different time windows. Moving rightward along the x-axis, the estimates are provided using longer time windows to define neighbor ads. Our baseline is 3 days before and after (thus a 7-day window around the ad), the leftmost point in the panel. Panels (b), (c), and (d) respectively show the intent-to-treat effect of treatment for ads with shares of neighbor ads treated (SNT_i) in the bottom quartile, middle quartiles, and top quartile of the SNT_i distribution. In particular, they respectively show β_0, $\beta_0 + \beta_M$, and $\beta_0 + \beta_T$. Thus the leftmost markers (the 3 days before or after window) match the estimates in columns 1-2 of the bottom panel of Table 2. Circles are estimates using baseline controls (month dummies interacted with remote status), while squares use controls selected by PDS-LASSO. The whiskers present the 95% confidence intervals.

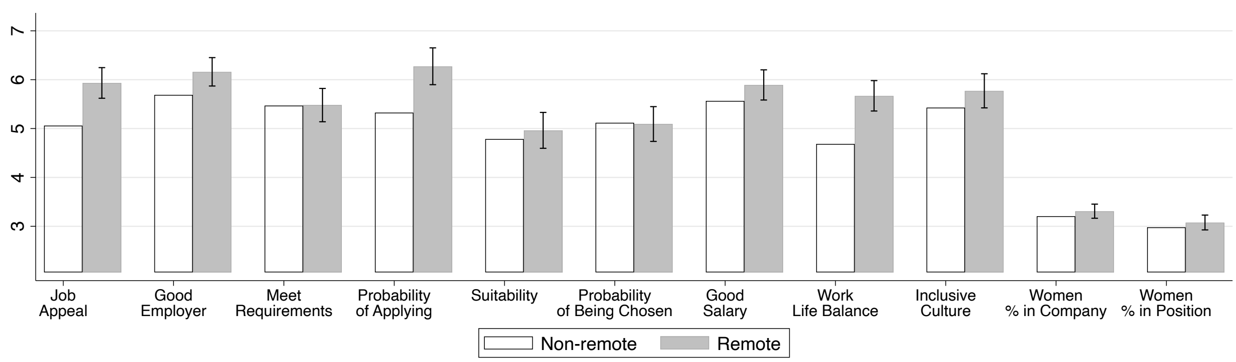
Figure 5: Outcome Averages by Different Treatment Statuses - Laboratoria



(a) Gender Neutral Language Treatment



(b) Diversity Statement Treatment



(c) Remote Job Treatment

Notes: The unit of observation is a response to an ad (each of the 546 respondents sees two ads). Figures provide the "raw" averages for the eleven outcomes collected in the survey (see text for definitions), by different treatment statuses. Whiskers present the 95% CI of the difference between averages (the treatment effect), based on robust standard errors. All observations are included (e.g., Panel (a) includes all observations regardless of remote or diversity statement status).

Table 1: Gender-Neutrality by Treatment Status - Get On Board

Classification Based on Ads' Titles

	English	Spanish GN	Spanish not GN	Total
Control	589	130	411	1,130
	(52.12%)	(11.50%)	(36.37%)	
Treatment	517	522	32	1,071
	(48.27%)	(48.74%)	(2.99%)	

Classification Based on Ads' Full Text

	English	Spanish GN	Spanish not GN	Total
Control	135	283	712	1,130
	(11.95%)	(25.04%)	(63.01%)	
Treatment	143	590	338	1,071
	(13.35%)	(55.09%)	(31.56%)	

Notes: Unit of observation is an ad. The use of gender-neutral language is classified in two manners. The top panel classifies job ads by considering only the text in the title. The lower panel classifies ads using the title and entire text of the ad. See the main text for further details. The table lists the number of ads in each category of gender-neutrality (English, Spanish gender-neutral, Spanish not gender-neutral) and assigned treatment status (treatment and control). Numbers in parentheses provide the ratio between the number of ads and the "total" in the same row (e.g., the share of control ads that have titles in English, titles in Spanish gender-neutral language, and so on).

Table 2: Intent-to-Treat Effects - Get on Board

	(1) Fem. Share Applicants	(2) Fem. Share Applicants	(3) asinh(Fem. Applicants)	(4) asinh(Fem. Applicants)	(5) asinh(Male Applicants)	(6) asinh(Male Applicants)	(7) Avg. Badness Score	(8) Avg. Badness Score
Treatment (β_0)	0.036**	0.039***	0.184	0.183	-0.072	-0.079	0.034	0.025
	(0.015)	(0.015)	(0.134)	(0.133)	(0.101)	(0.101)	(0.049)	(0.048)
Treat × Mid. Quartiles of	-0.054***	-0.056***	-0.238	-0.225	0.140	0.148	-0.018	0.001
% Neighbors Treated (β_M)	(0.018)	(0.018)	(0.160)	(0.157)	(0.121)	(0.121)	(0.062)	(0.061)
Treat × Top Quartile of	-0.047**	-0.043**	-0.236	-0.199	0.057	0.064	0.136*	0.134*
% Neighbors Treated (β_T)	(0.020)	(0.020)	(0.184)	(0.181)	(0.141)	(0.141)	(0.072)	(0.071)
Mid. Quartiles of	-0.010	0.024*	-0.191*	0.103	-0.111	-0.043	0.072*	0.040
% Neighbors Treated (γ_M)	(0.012)	(0.012)	(0.114)	(0.115)	(0.084)	(0.086)	(0.043)	(0.044)
Top Quartile of	-0.004	0.001	-0.047	0.018	-0.040	-0.028	-0.039	-0.040
% Neighbors Treated (γ_T)	(0.014)	(0.014)	(0.131)	(0.128)	(0.100)	(0.100)	(0.050)	(0.049)
Implied Treatment Effects								
Bottom Quartile of %	0.036	0.039	0.184	0.183	-0.072	-0.079	0.034	0.025
Neighbors Treated (β_0)	(0.015)	(0.015)	(0.134)	(0.133)	(0.101)	(0.101)	(0.049)	(0.048)
	[0.044]**	[0.022]**	[0.243]	[0.240]	[0.519]	[0.462]	[0.530]	[0.638]
Mid. Quartiles of %	-0.018	-0.017	-0.055	-0.042	0.068	0.069	0.016	0.026
Neighbors Treated (β_0+β_M)	(0.010)	(0.009)	(0.087)	(0.083)	(0.067)	(0.067)	(0.038)	(0.038)
	[0.140]	[0.141]	[0.595]	[0.673]	[0.341]	[0.327]	[0.685]	[0.503]
Top Quartile of %	-0.011	-0.004	-0.052	-0.016	-0.015	-0.014	0.170	0.159
Neighbors Treated ($\beta_0 + \beta_T$)	(0.014)	(0.013)	(0.127)	(0.123)	(0.099)	(0.097)	(0.053)	(0.053)
	[0.522]	[0.811]	[0.712]	[0.906]	[0.889]	[0.876]	[0.002]***	[0.000]***
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.146	0.146	-	-	-	-	15.121	15.121
N	2,201	2,201	2,201	2,201	2,201	2,201	2,201	2,201

Notes: Unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status), while even-numbered columns include controls selected by PDS-LASSO. Outcomes are the share of applicants that are female (columns 1-2), the inverse hyperbolic sine of the number of female and male applicants (columns 3-4 and 5-6, respectively), and the applicants' average "badness score" (a measure of applicant quality, columns 7-8). The top panel provides the estimated coefficients from equation (1). The independent variables are a dummy for treatment assignment, two dummies indicating if the ad's share of neighbor ads treated (SNT_i) falls in the middle quartiles or the top quartile of the SNT_i distribution, and their interactions with treatment. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with SNT_i in the bottom quartile, medium quartiles, and top quartile. The last two rows provide the average of the outcome variable for control ads and the number of observations. Standard errors are in parentheses and randomization inference p-values are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Treatment-on-Treated (2SLS) Effects of Gender-Neutrality - Get on Board

	(1) Fem. Share Applicants	(2) Fem. Share Applicants
Bottom Quartile of % Neighbors Treated	0.104**	0.115***
	(0.044)	(0.044)
Mid Quartiles of % Neighbors Treated	-0.055*	-0.053*
	(0.031)	(0.030)
Top Quartile of % Neighbors Treated	-0.039	-0.015
	(0.050)	(0.050)
Baseline Controls?	YES	
PDS-LASSO Controls?		YES
Control Mean	0.146	0.146
N	2,201	2,201

Notes: Unit of observation is an ad. Column (1) includes baseline controls (month dummies interacted with remote status), while column (2) includes controls selected by PDS-LASSO. The outcome (dependent variable) in both columns is the share of applicants that are female. The table presents the linear combinations that provide the treatment-on-treated effects of an ad being gender-neutral (based on the full-text classification) for ads with a share of neighbor ads treated (SNT_i) falling in the bottom quartile, middle quartile, and top quartile of the SNT_i distribution. In particular, the table presents β_0^{2SLS}, $\beta_0^{2SLS} + \beta_M^{2SLS}$, and $\beta_0^{2SLS} + \beta_T^{2SLS}$ from equation (2) estimated via 2SLS where the three excluded instruments are the treatment assignment and its interaction with two dummies indicating if SNT_i falls in the middle quartiles or the top quartile of its distribution. Table A.5 presents the estimates of β_0^{2SLS}, β_0^{2SLS}, and β_0^{2SLS} and the related first-stage regressions. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effect by Discarded, Selected, Hired Status
- Get on Board

	All Applicants		Not Discarded		Selected		Hired	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment (β_0)	0.036**	0.039***	0.044**	0.046**	0.050	0.053	0.130*	0.101
	(0.015)	(0.015)	(0.019)	(0.018)	(0.041)	(0.042)	(0.071)	(0.070)
Treat × Mid. Quartiles of	-0.054***	-0.056***	-0.057**	-0.059***	-0.101**	-0.107**	-0.210**	-0.187**
% Neighbors Treated (β_M)	(0.018)	(0.018)	(0.022)	(0.022)	(0.050)	(0.050)	(0.087)	(0.084)
Treat × Top Quartile of	-0.047**	-0.043**	-0.035	-0.033	-0.050	-0.048	-0.079	-0.022
% Neighbors Treated (β_T)	(0.020)	(0.020)	(0.026)	(0.026)	(0.057)	(0.056)	(0.094)	(0.094)
Mid. Quartiles of	-0.010	0.024*	-0.015	0.026*	-0.022	0.016	0.059	0.036
% Neighbors Treated (γ_M)	(0.012)	(0.012)	(0.016)	(0.016)	(0.034)	(0.035)	(0.059)	(0.058)
Top Quartile of	-0.004	0.001	-0.009	0.001	-0.036	-0.033	-0.042	-0.082
% Neighbors Treated (γ_T)	(0.014)	(0.014)	(0.017)	(0.017)	(0.038)	(0.038)	(0.063)	(0.060)
Implied Treatment Effects								
Bottom Quartile of %	0.036	0.039	0.044	0.046	0.050	0.053	0.130	0.101
Neighbors Treated (β_0)	(0.015)	(0.015)	(0.019)	(0.018)	(0.041)	(0.042)	(0.071)	(0.070)
	[0.044]**	[0.022]**	[0.044]**	[0.029]**	[0.257]	[0.246]	[0.064]*	[0.151]
Mid. Quartiles of %	-0.018	-0.017	-0.014	-0.013	-0.051	-0.054	-0.080	-0.086
Neighbors Treated (β_0+β_M)	(0.010)	(0.009)	(0.012)	(0.012)	(0.027)	(0.027)	(0.048)	(0.047)
	[0.140]	[0.141]	[0.341]	[0.362]	[0.090]*	[0.072]*	[0.096]*	[0.078]*
Top Quartile of %	-0.011	-0.004	0.009	0.013	0.001	0.005	0.051	0.079
Neighbors Treated ($\beta_0+\beta_T$)	(0.014)	(0.013)	(0.018)	(0.018)	(0.039)	(0.037)	(0.062)	(0.063)
	[0.522]	[0.811]	[0.682]	[0.528]	[0.991]	[0.906]	[0.484]	[0.256]
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.146	0.146	0.157	0.157	0.175	0.175	0.202	0.202
N	2,201	2,201	1,714	1,714	774	774	508	508

Notes: Unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status), while even-numbered columns include controls selected by PDS-LASSO. The outcome (dependent variable) in all columns is the share of applicants that are female, calculated using all applicants (columns 1-2) or only those marked by the firm as "not discarded," "selected," and "hired" on Get On Board's personalized evaluation board (columns 3-4, 5-6, and 7-8, respectively). The number of observations changes across columns since not all companies use the evaluation boards for all their ads. The top panel provides the estimated coefficients from equation (1). The independent variables are a dummy for treatment assignment, two dummies indicating if the ad's share of neighbor ads treated (SNT_i) falls in the middle quartiles or the top quartile of the SNT_i distribution, and their interactions with treatment. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with SNT_i in the bottom quartile, medium quartiles, and top quartile. The last two rows provide the average of the outcome variable for control ads and the number of observations. Standard errors are in parentheses and randomization inference p-values are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Treatment Effects by Ad Order - Laboratoria

	(1) Job Appeal	(2) Good Employer	(3) Meet Require-ments	(4) Probability of Applying	(5) Suitability	(6) Probability of Being Chosen	(7) Good Salary	(8) Work Life Balance	(9) Inclusive Culture	(10) Women % Company	(11) Women % Position
Gender-neutral	0.146	-0.055	0.030	-0.269	0.636**	0.035	-0.039	-0.031	0.493**	0.298***	0.351***
	(0.215)	(0.202)	(0.246)	(0.283)	(0.266)	(0.261)	(0.223)	(0.224)	(0.239)	(0.095)	(0.102)
GN × 2nd Ad	0.798***	1.243***	0.266	1.553***	0.167	0.667*	0.862***	1.037***	1.579***	0.716***	0.585***
	(0.309)	(0.286)	(0.348)	(0.376)	(0.371)	(0.363)	(0.305)	(0.307)	(0.332)	(0.138)	(0.146)
Remote	0.797***	0.406**	-0.057	0.819***	0.238	-0.131	0.198	0.879***	0.282	0.098	0.083
	(0.216)	(0.202)	(0.246)	(0.283)	(0.267)	(0.262)	(0.223)	(0.223)	(0.239)	(0.095)	(0.102)
Remote × 2nd Ad	0.109	0.100	0.138	0.226	-0.137	0.211	0.216	0.177	0.109	0.004	0.021
	(0.310)	(0.286)	(0.349)	(0.376)	(0.371)	(0.364)	(0.305)	(0.307)	(0.332)	(0.138)	(0.146)
Diversity Statement	-0.160	0.010	-0.070	-0.075	-0.003	0.095	-0.216	-0.023	0.912***	0.158*	0.115
	(0.216)	(0.203)	(0.246)	(0.283)	(0.267)	(0.262)	(0.223)	(0.224)	(0.240)	(0.096)	(0.103)
Diversity × 2nd Ad	0.424	0.518*	0.336	0.149	0.246	0.222	0.281	0.460	0.099	0.196	0.192
	(0.310)	(0.286)	(0.349)	(0.376)	(0.371)	(0.364)	(0.305)	(0.307)	(0.332)	(0.138)	(0.147)
2nd Ad	0.572*	0.072	-0.720**	0.065	0.552	-0.532	0.532*	0.303	0.355	-0.069	0.044
	(0.320)	(0.297)	(0.352)	(0.391)	(0.372)	(0.367)	(0.309)	(0.302)	(0.325)	(0.138)	(0.144)
Control Mean - 1st Ad	4.786	5.232	5.871	5.246	4.250	5.186	5.314	4.290	4.174	2.786	2.571
Observations	1,090	1,090	1,089	1,089	1,086	1,088	1,089	1,088	1,085	1,089	1,085

Notes: Unit of observation is a response to an ad (each respondent sees two ads). Each column presents an estimate from equation (3) for a different outcome (see text for definitions). Gender-neutral, Remote, and Diversity Statements are dummies indicating the ad was assigned to the respective status. 2nd Ad is a dummy indicating whether the ad was the second shown. The control mean is the outcome mean for the first ads shown under the non-gender-neutral language, non-remote treatment, and no diversity treatment status. The number of observations varies across columns due to missing data on outcomes (a few instances when respondents did not answer a survey question). Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix

Appendix A discusses gendered grammar in Spanish (and Portuguese), as well as issues related to the adoption of gender-neutral language and its effects.
Appendix B discusses papers examining the effects language and content in job ads.
Appendix C provides additional information on data construction and variable definitions for the Get On Board experiment.
Appendix D presents additional results, tables, and figures for the Get On Board experiment.
Appendix E presents additional results, tables, and figures for the Laboratoria experiment.
Appendix F provides the experimental materials related to the Get On Board experiment.
Appendix G provides the experimental materials related to the Laboratoria experiment.

A Gendered Grammar and Gendered Languages

Gendered grammar. Languages differ on their treatment of gender. At one extreme, some languages do not make gender distinctions (e.g., Finnish), while at the other are languages that assign gender to all nouns, including inanimate objects (e.g., Spanish, Portuguese, French, Italian). English is situated in the "middle of the spectrum," since most nouns do not have a gender and it has non-gendered third-person pronouns ("it" and "they") and articles ("the" and "an").

We refer to languages such as Spanish and Portuguese as having *gendered grammar* (Hellinger and Bußmann, 2015). English, given the distinctions described above, does not fit this definition. Jakiela and Ozier (2018) documents the presence or absence of gendered grammar in more than 4,000 languages that account for more than 99% of the world's population and find that 39% of the world's population speaks a gendered grammar language.

Gendered grammar in Spanish. This section describes the traditional grammar in Spanish, but all the issues here apply equally to Portuguese (the language used by roughly 8% of respondents in the Laboratoria experiment). In Spanish, *every* noun is gendered. For example, "*ingeniero*" and "*ingeniera*" mean "male engineer" and "female engineer," respectively. There is no traditional and widely accepted way to refer to an engineer without implying a gender. The same applies to job candidates ("*candidato*" versus "*candidata*") or the person hired ("*contratado*" versus "*contratada*").

Moreover, all articles are gendered in order to match the gender of the noun. Indefinite articles in Spanish are the male and female "*un*" and "*una*" (and the plurals "*unos*" and "*unas*"). Similarly, definite articles are the female "*la*'' (plural "*las*") and the male "*el*" (plural "*los*") and "*lo*." This implies one refers to "*el ingeniero*" or "*una ingeniera*." A group of engineers of both genders would be referred as "*los ingenieros*," which is the exact same as one would refer to an all-male group of engineers. "*Las ingenieras*" would imply an all-female group of engineers.

The examples above indicate the "generic masculine" that is traditional and widely common in Spanish. In situations where no gender must be specified (such as a job ad searching for an engineer), the standard is to state that a company is looking to hire one "*ingeniero*" or multiple "*ingenieros.*"

Moreover, inanimate objects have gender too. For example, a car ("*un coche*") is male and a house ("*una casa*") is female. Third-person pronouns are also gendered ("*él*" and "*ellos*', "*ella*" and "*ellas*"). There are no third-person non-gendered pronouns like "it" or "they" in English.

Some nouns have their male and female form spelled the same way. For example, "*analista*" refers to a male or female analyst, and "*economista*" refers to a male or female economist. However, given gendered pronouns, these nouns are also gendered. For example, "the company is hiring an economist" can either be translated to "*la empresa esta contratando* ***un*** *economista*" (implying a male economist) or "*la empresa esta contratando* ***una*** *economista*" (implying a female economist). A similar issue applies with plurals ("*unas economistas*" versus "*unos economistas*").

Gender-neutral language in Latin America. In recent years, a growing movement has advocated for the use of gender-neutral language throughout the continent. However, there is no consensus on the method to make Spanish gender-neutral. For example, some advocate that instead of using the male "*amigos*" or female "*amigas*" to refer to "friends," one should use "x" or "e" to create non-gendered nouns: "*amigxs*" or "*amigues.*" American readers may be familiar with the term "*latinx*" to avoid the generic masculine "*latino*" and thus be gender-neutral. This is a substantial departure from "traditional" Spanish grammar (e.g., what most Latin Americans learn at school).

Both our experiments follow what is arguably a less radical approach, which is also the one advocated by some Latin American governments. In particular, our gender-neutral language protocol is based on a set of guidelines published by the Ministry of Women and Vulnerable Populations in Peru in 2017.[39] Note that our partner organizations (Get On Board and Laboratoria) are based in Peru.

The adoption of gender-neutral language has attracted substantial controversy and government intervention in Latin America. For example, in July 2022 the city government in Buenos Aires (Argentina) banned primary and secondary school teachers from using any gender-neutral words during class and in communications with parents, claiming it violated Spanish grammar rules and adversely affected students' reading comprehension. There was

[39] https://www.mimp.gob.pe/files/direcciones/dgteg/Guia-de-Lenguaje-Inclusivo_v2.pdf

no official policy regarding gender-neutral language in Buenos Aires, and some teachers had informally adopted it.

Similarly, since 2021 bill prohibiting the use of gender-neutral language in schools has been proposed in 80% of Brazilian state legislatures. Three different states (Amazonas, Paraná, and Rondônia) have enacted such bills into law. Individual municipalities in Brazil also enacted laws imposing fines and withdrawal of government support to schools that used gender-neutral language. Three Brazilian supreme court decisions (in 2021, 2023, and 2024) stated that such prohibitions and fines are unconstitutional on the grounds that only the federal government can legislate on such matters.

Literature on gendered languages. A large body of research, across multiple disciplines, studies how language shapes human decisions and cognition. For example, speakers of languages that demarcate the future from the present have been shown to save less than those whose language makes no such distinction (Chen, 2013), and bilinguals display different subconscious attitudes when tested in different languages (Ogunnaike et al., 2010, Danziger and Ward, 2010). Speaking minority tongues primes ethnic divisions (Pérez and Tavitz, 2019). The use of plural pronouns impacts perceptions of a relationship (Fitsimons and Kay, 2004).

The closest literature to the issue in this paper refers to how people interpret *masculine generics*. Moulton et al. (1978) found evidence that when the terms "he, him, and man" were expressed in a supposedly gender-neutral way, people more often thought of male referents than they did when explicitly neutral alternative forms such as feminine-masculine word pairs were used. Crawford and English (1984) provide evidence that women recall information better when instructions specifically include reference to women. Gastil (1990) found that the feminine-masculine word pairs were perceived as generic, leading subjects to recall roughly the same amount of female, male, and mixed images, whereas the masculine form appeared to bias the reader toward imagining male referents. Cohen et al. (2023) studies the introduction of gender-neutral language in college entrance in Israel, and finds that it raised female performance on quantitative questions, but had no effect on female performance on verbal questions or male performance on either type of questions.

Jakiela and Ozier (2018) provides an overview of definitions and a survey the literature on gendered language.

A digression on gendered language in the economics profession. The difficulties of dealing with gendered language and generic masculines are neither new nor foreign to academic economists, who tend to refer to agents in abstract models by the pronouns "she/her/hers." An illustrative example comes from two textbooks, written over 25 years apart.

In the 1994 textbook *A Course on Game Theory* (Osborne and Rubinstein, 1994), the authors provide a "*note on personal pronouns*" where Rubinstein advocates for the use of "he" as a "*neutral*" pronoun, stating the use of "she" would "*divert the readers' attention.*" His co-author Osborne takes issue with this position and argues that "*a wealth of evidence*" indicates that "*'he' is not generally perceived to encompass both females and males,'*' and his preference is to refer to agents as "she." The note ends with "*To conclude, we both feel strongly on this issue; we both regard the compromise that we have reached as highly unsatisfactory. When referring to specific individuals, we sometimes use 'he' and sometimes 'she'.*" However, both authors agree that "*language is extremely important in shaping our thinking.*"

In the 2020 textbook *Models in Microeconomic Theory* (Osborne and Rubinstein, 2020), the same authors state that, although "*during our thirty years of collaboration we have often debated the use of gendered pronouns in academic material,*" their opinions on the topic "*remain unchanged.*" However, they find a different solution: "*this book has two editions, one that uses feminine pronouns and one that uses masculine pronouns. We leave it to you to make your choice.*" As of December 2024, the female-pronoun version of the book's second edition had a slightly larger number of downloads than the male-pronoun version, according to the book's website.

B Literature on Job Ad Content and Language

As discussed in the introduction, a growing body of literature examines how interventions on the content and language of ads and recruitment materials affect the composition of the applicant pool. To our knowledge, our study is the first to evaluate gender-neutral language and to examine treatment spillovers for any type of content. This appendix provides further information on this literature. See also Kuhn and Shen (2023) for a related discussion.

Experimental papers. Several papers study randomized experiments where the intervention involves changing the content or language of job ads. Some interventions provide clear and factual descriptions about relatively objective job characteristics, such as indicating flexible work hours (Mas and Pallais, 2017), negotiable salaries (Leibbrandt and List, 2015), competitive compensation regimes (Flory et al., 2015, Samek, 2019), or information on the share of workers receiving high evaluations (Delfino, 2024). Another set of papers varies the posted wage rate, focusing on job search models' predictions rather than applicant pool diversity (e.g., Belot et al., 2019, Banfi and Villena-Roldan, 2019).

Another set of papers involves interventions that can be interpreted as changes in language, which do not directly provide information about job characteristics but can still signal them to potential applicants. These include explicit diversity statements (Ibañez and Riener, 2018, Leibbrandt and List, 2018, Flory et al., 2021), removing optional qualifications and superfluous language (Abraham et al., 2024), reducing ambiguity around required qualifications (Coffman et al., 2024), or the gender of workers depicted in photographs (Delfino, 2024).[40]

Gee (2018) examines the effects of providing information on the number of competing applicants for a job. Del Carpio and Guadalupe (2021) investigate a multifaceted intervention aimed at recruiting Latin American women to tech sector boot camps. This intervention included emphasizing female role models, providing information on returns, and offering access to female networks.

The papers mentioned above involve researchers partnering with a single firm or creating a job position and posting ads themselves (e.g., hiring research assistants within a university or on online platforms). A common design is randomizing at the potential applicant level (i.e., which information about a job is provided to them individually). Thus, their experimental designs do not allow for the study of the type of spillovers that is the focus of this paper. Our study leverages multiple ads from different firms being treated, creating (random) variation in the share of treated ads that different applicants consider.[41]

Non-experimental papers. Card et al. (2024), Helleseter et al. (2020), Kuhn et al. (2020), Kuhn and Shen (2023) use observational data to study explicit gender requests (ads explicitly stating they prefer applications from men or women). These papers do not examine spillovers or general equilibrium effects.

The exception is Kuhn and Shen (2023) study of a ban on employers' explicit gender requests in a Chinese job board that affected all its ads simultaneously. It estimates its effects both on ads directly affected and on non-directly affected ads that did not have requests before the ban. We study a different type of spillovers, leveraging random variation in the share of ads for similar positions that were concurrently treated.

C Additional Information on Variable Definitions

Procedure to create *job title groups*. The definition and intuition behind the job title groups, a key variable defining the neighbor ads, is discussed in Section 3. This appendix section describes the procedure used to create the groups. Based on our reading of a random sample of titles, we created an initial set of seven job title groups labeled *admin*, *developer*,

[40]Gaucher et al. (2011) studies how university students respond to hypothetical job ads, varying whether words associated with male (e.g., "dominant") or female stereotypes (e.g., "support") are used in the ads.

[41]The exceptions are Gee (2018), which randomizes at the user level, with treated users seeing the number of applicants for all job postings on the platform, and Belot et al. (2019), which posts fictitious ads on an online job board, but never exceeding 2% of all posted ads.

programmer, *designer*, *engineer*, *analyst*, and *other*. We then assigned every ad in our data following the procedure below:

1. Assign ad i to *admin* if at least one of the following holds: i) the ad title's first word includes "adm" or "jefe"; ii) the second or the last word includes "manag".
2. Assign ad i to *developer* if the first, second, or last word of its title included "desar" or "deve".
3. Assign ad i to *programmer* if its title's first word included "progra".
4. Assign ad i to *designer* if at least one of the following holds: i) the first word included "dise"; ii) the first, second, or the last word included "desi".
5. Assign ad i to *engineer* if at least one of the following holds: i) its title's first word started with "ing"; ii) the first, the second, or the last word in its title started with "eng".
6. Assign ad i to *analyst* if the first, second, or last word in its title started with "ana".
7. Assign ad i to *other* if it was not assigned to any of the six categories above or if it was assigned to more than one.

In step two, we prompted the ChatGPT large language model by providing the full list of job ad titles in our data and prompting the query "*I will provide you with a list of job titles. Your task is to simplify the job titles making them as general as possible, similar to other relevant titles as possible whilst merging them where possible. In the simplified job titles, there is no need to differentiate the different software or tools involved for the jobs; as long as the roles are similar, they should have the same job title.*"

While we did not simply use ChatGPT's suggestion unchanged, its suggestions informed the creation of additional groups and substituting two initial ones, as described below.

ChatGPT's suggestion involved six categories with the word "developer" in its group title: *web developer*, *front-end developer*, *back-end developer*, *mobile developer*, *full-stack developer*, and *other developer*. We assigned ad i to such groups as suggested by ChatGPT if ad i had originally been assigned to the *developer* and/or *other* group in step one. This implied that the original *developer* group was substituted by six distinct groups.

We then assigned ad i to step one's *engineer* group if the ad had been assigned to *other* in step one and ChatGPT's suggestion for its job title group included the word "engineer." We also assigned to a new group *architect* the ads that remained in the *other* group and had "architect" in its title. We assigned to a new group *data science* the remaining ads in the *other* group that included "data science", "data scientist", "científico de datos", or "científica/o de datos", in their titles. We also assigned the remaining ads in the *other* group to a new group *scrum* if they included "scrum" in their titles.

We manually broke down the ads originally in the *admin* group into two separate groups (*sysadmin* and *bizadmin*). This implied that the original *admin* group was entirely substituted by the two new groups. The rationale is to separate administrators of business operations from (software) system administrators. Lastly, amongst the ads remaining in the *other* category, we manually assigned some to *marketing/customers*. By "manually," we mean we asked a research assistant to read the relevant job titles and make a decision regarding the assignment. We independently performed the task and reached the same assignment.

The procedure above resulted in the creation of 16 job title groups, not including step one's *other* group. Out of the 2,535 ads in our original sample, 231 remained in the *other* group at the end of the procedure. These 231 ads are not used in our main analysis given that defining meaningful job title groups and thus neighbor ads are an essential part of the analysis (see Section 3).

The 16 job title groups, their representation in the sample, and the share of applicants to its positions that are women are provided in Table A.3.

Remoteness. Our experiment was conducted while mobility restrictions due to the covid-19 pandemic were still in place, and a large portion of the ads listed a remote position (at least temporarily). Get On Board asked firms to state how their ad fitted into three mutually exclusive categories: *temporarily remote* jobs, expected to become in-person after restrictions were lifted; *locally remote* jobs that were fully remote but required a person living in a specific country; and *fully remote* jobs that had no restrictions on the location of the employee. We classify as "remote" all the positions listed as locally remote or fully remote. Jointly, they constitute 40% of our sample.[42]

D Additional Results, Tables, and Figures - Get On Board

Covariate Balance. As discussed in Section 2.1, Table A.1 provides summary statistics and balance checks. As a test of the overall balance in our sample, we report an omnibus test suggested by Kerwin et al. (2024). Specifically, we estimate a regression where the dependent variable is the treatment dummy indicator and the independent variables are all the variables listed in Table A.1, a set of nine country dummies, a set of 16 job group title dummies, and a set of 12 field dummies.[43] We report the randomization inference (permutation) p-values

[42]Before the Covid-19 pandemic, only 6% of ads on the platform were remote.

[43]See Tables A.3 and A.4 for the list of job group titles and fields. The set of country dummies includes a dummy equal one if the ad did not specify a country of work (which is common amongst remote positions).

based on randomly reassigning the treatment (i.e., the p-value is the share of draws where the computed F-statistic is larger than the actual F-statistic computed with the actually realized treatment assignment). We use our entire sample (2,201 observations) and 1,000 repetitions. The p-value from the test is 0.338, indicating we cannot reject the null of joint covariate balance.[44]

Similar omnibus tests using only the set of country dummies, only the set of job group title dummies, or only the set of field dummies, indicate that treatment and control groups are balanced in terms of country, job group titles, and fields. The respective p-values are 0.241, 0.286, and 0.281.

Causal forests and treatment effect heterogeneity. As discussed in Section 3, machine learning confirms the importance of share of neighbor ads treated (SNT_i) for treatment effect heterogeneity. Figure 3 provides the results, with SNT_i being the variable with the largest "variable importance." Specifically, using our entire sample (2201 observations), we fit a causal forest (Athey et al., 2019) using the share of applicants to ad i that are female as the outcome and T_i as the treatment (i.e., an intent-to-treat analysis). We use the GRF package in R (Tibshirani et al., 2024) and its the "variable_importance" function, which provides a measure of how often the variable was used in tree splits.

The set of covariates that can potentially predict effect heterogeneity include an indicator if the ad title is in English, a set of month dummies, the share of female applicants in the job title group, and all variables listed in Table A.1 (except the minimum and maximum of salary range, which is missing for ads that did not post a range). The share of female applicants in the job title group is constructed only using ads assigned to control. For each job title group, we calculate the average share of applicants to control female ads. We then assign that value to all ads in that job title group. This variable thus measures the gender balance in a job title group in a baseline scenario in a manner not directly affected by our treatment. We include this variable as it allows us to test if the effects are heterogeneous based on whether the type of position is more gender-balanced, which is motivated by female representation in an occupation being predictive of gender bias in a meta-analysis of audit studies (Galos and Coppock, 2023). Table A.3 provides the value of the female share of applicants by job title group. It is particularly low for developers, but higher for bizadmin, designers, and marketing/customers positions.

The set of covariates that can potentially predict treatment effect heterogeneity differs

[44]Simulations in Kerwin et al. (2024) indicate that using the F-statistic from such regressions and the use of randomization inference (permutation p-values, instead of sampling-based) yields tests of correct size. Of the variables from Table A.1, the minimum and maximum of the salary range are not included, since it is missing for ads that did not post a salary range. Each new draw of our simulation also involved recomputing the share of neighbor ads treated used as an explanatory variable since this variable is a function of the treatment status of neighbor ads.

slightly from the set of covariates we use as potential controls in our PDS-LASSO specification when estimating equation (1), discussed in Section 3. Using that as the covariate set, we again find that the share of neighbors treated has the highest variable importance (34.3%). The number of neighbor ads has an importance of 20.7%, and every other variable has an importance below 4.1%.

Effects on the Distribution of the Share of Female Applicants. As discussed in Section 3, Figure A.2 provides the cumulative distribution function (CDF) of the share of female applicants in ads assigned to control and treatment status. The unit of observation in the distributions is an ad. The figures do not involve the use of any controls. It does so for the entire sample and separately for ads in the bottom quartile, middle quartiles, and top quartile of the share of neighbor ads treated (SNT_i) distribution. It thus replicates for CDFs what columns (1)-(2) of Table 2 do for averages. The treatment CDF is most clearly "shifted to the right" of the control CDF in panel (b): the case of ads in the bottom quartile. This indicates that the effects of treatment appear relatively constant throughout the distribution.

Effects on the Distribution of Applicants' Quality. As discussed in Section 3, Figure A.3 provides the CDF of badness scores in control and treatment groups. It does so separately for male and female applicants. Note that, differently from Figure A.2, the unit of observation is a job applicant (and not an ad). It thus shows the distributions of applicant quality (as measured by the badness scores) that applied to the entire pool of treated and control ads. Hence, the figures allow us to test if treatment ads attract or repel applicants from lower or upper parts of the quality distribution (i.e., effects beyond the average badness scores). The CDFs have a remarkable overlap, indicating that the distribution of badness score is not affected by treatment in the overall sample, for either gender. An "excess mass" is visible at the badness score of 15 (which is the default score assigned to Get On Board users when they first create an account).

Figure A.4 repeats the exercise but separately for ads in the bottom quartile, middle quartiles, and top quartile of the share of neighbor ads treated (SNT_i) distribution. Again, the CDFs have a remarkable overlap in all cases.

For ads in the bottom quartile of SNT_i, there is an effect on the share of female applicants (columns 1-2 of Table 2). Panels (a)-(b) of Figure A.4 show that the distribution of male and female applicant quality in control and treatment ads is similar for these ads. These two results combined suggest that treatment increases the share of women applying without affecting the quality distribution of applicants, indicating that the larger share of female applicants comes from across the quality spectrum. This implies effects on the share of female applicants at any given quality threshold. For example, firms that only consider applicants with badness scores above a certain cutoff would see a larger share of female applicants *above the cutoff* as a result of the treatment, regardless of the cutoff.

Treatment-on-treated (2SLS) effects. As discussed in Section 3, Columns (1) and (2) of Table A.5 present the results from 2SLS estimation of equation (2). In particular, it provides the estimates of β_0^{2SLS}, β_0^{2SLS}, and β_0^{2SLS} that inform the linear combinations reported on Table 3 and discussed in the main text. Columns (3)-(8) present the first-stage estimates. Since we have three endogenous variables and three excluded instruments on equation (2), there are three first-stages reported.

We highlight three points about the first stages. First, they show a roughly 30% first-stage effect, consistent with the bottom panel of Table 1. Second, for each first stage, the "relevant coefficient" is roughly 30% but the other two are close to zero and insignificant. For example, when the instrument is gender-neutral ad interacted with a dummy for *middle* quartiles of SNT_i, the "relevant coefficient" is the of treatment interacted with a dummy for *middle* quartiles, which is approximately 30%. The coefficients on non-interacted treatment and its interaction with the top quartile dummy are essentially zero. Since this is expected given random assignment, but can also be interpreted as a check on the randomization protocol. Also consistent with randomization, we cannot reject the null that the "relevant coefficients" are the same across all columns. Third, the first stage is strong, with the "relevant" coefficients having t-statistics ranging from 6.7 to 11.0.

Additional results: heterogeneity by title language and remote position. As discussed in Section 3, Table A.6 presents treatment effect heterogeneity by whether the ad's title is in English or in Spanish and whether the position is remote. In particular, the results on title language suggest that gender-neutral language in the text of the ad (beyond its title) matters. Given titles in English are already gender-neutral, our treatment does not affect how they are written (see Section 3).

Additional results: subsequent ads. As discussed in Section 3, Table A.7 examines whether receiving treatment affects the *subsequent* ads that a firm posts on the platform. In particular, columns (1) and (2) report estimates from a firm-level regression:

$$y_f = \delta_0 + \delta_1 FirstAdTreated_f + \epsilon_f \tag{4}$$

where f indexes firms in the sample and $FirstAdTreated_f$ is a dummy equal one if the first ad the firm posted on the platform during the experimental period was randomly selected for treatment. We examine two outcomes (y_f): a dummy if the firm posted a second ad, and the total number of ads the firm posted in the sample period. δ_1 thus tests if being selected for treatment makes the firm use the platform less or more intensely. Our regression includes 711 firms that posted at least one ad in the sample period. We exclude from the sample 293 ads that could not be assigned to a given company, given missing data on the official name of the company as they registered on Get On Board. We estimate a δ_1 close to zero, indicating treatment does not affect the number of ads a firm posts on the platform.

Columns (3)-(6) present results from the following ad-level regression:

$$GN_i = \theta_0 + \theta_1 FirstAdTreated_i + \epsilon_i \tag{5}$$

where i indexes ads and $FirstAdTreated_i$ is a dummy equal one if the first ad that the firm that posted ad i was randomly selected for treatment. The sample only includes ads that are the second or higher order posted by a firm in the sample period, which restricts us to 527, since we also exclude 293 ads that could not be assigned to a firm. In columns (3) and (4) we further restrict to only the second ad (163 observations). We examine two outcomes (GN_i): whether ad i's title was gender-neutral, or whether its entire text was gender-neutral (see Table 1 and related discussion in Section 3). Standard errors are clustered at the firm level. We estimate a θ_1 that is close to zero and insignificant. This indicates that, after having their first ad treated, firms are not more likely to post more ads using gender-neutral language.

Robustness checks and placebo tests. As discussed in Section 3, Figure 4 examines how our main ITT results (equation 1 and Table 2) are influenced by different time windows used to define neighbor ads. Table A.8 presents a placebo test by re-estimating the same main ITT results but defining SNT_i based on "future" neighbors. Sample sizes are smaller in Table A.8 than Table 2 since ads in the last 30 and 60 days of the sample must be dropped from columns 1-2 and 3-4, respectively.

Table A.9 presents another placebo test. It replicates the same main ITT results from equation (1) but instead of exploring heterogeneity in SNT_i, it examines heterogeneity based on the female representation in the job title group. We focus on this dimension of heterogeneity because it is the second most important factor identified in our causal forest analysis (Figure 3) and because the gender composition of an occupation can be predictive of gender bias (Galos and Coppock, 2023). The share of female applicants in the job title group is constructed only using ads assigned to control. For each job title group, we calculate the average share of applicants to control female ads. We then assign that value to all ads in that job title group. This variable thus measures the gender balance in a job title group in a baseline scenario in a manner not directly affected by our treatment. Table A.3 provides the value of the female share of applicants by job title group.

E Additional Results, Tables, and Figure - Laboratoria

Balance and summary statistics. Table A.10 provides the sample averages by each treatment arm (three treatment combinations), indicating randomization successfully achieved covariate balance. See the table notes for an omnibus test of covariate balance.

Effects on outcome distributions. Figures A.5, A.6, and A.7 present the cumulative distribution function (CDF) for each of the eleven outcomes. It does so separately by each treatment. Since the experiment has a $2 \times 2 \times 2$ factorial design with equal probability, other treatment conditions are balanced when making two-way comparisons. In other words, Figures A.5, A.6, and A.7 do for outcomes' CDFs what Figure 5 does for outcomes' averages. In cases where we find positive effects, we can see they are driven by broad changes throughout the distribution of outcomes (e.g., a broader "right shift" in the CDF), implying effects along the entire distribution of outcomes.

Results in table format. Table A.11 presents the results from the following regression:

$$y_{ia} = \alpha + \beta GNeutral_{ia} + \gamma Diversity_{ia} + \delta Remote_{ia} + \epsilon_{ia} \tag{6}$$

where i indexes respondents and a indexes the ads they see. Each respondent sees two ads and thus with 546 respondents we have up to 1092 observations to be used. y_{ia} is an outcome variable (e.g., whether respondent i answered she would apply to job ad a). The three right-hand side variables are dummies indicating whether the ad shown was randomly assigned to be gender-neutral, have a diversity statement, and have remote status. We use heteroskedasticity-robust standard errors but obtain similar p-values for all estimates when using randomization inference based on 1,000 draws (which we omit from this and other related tables to economize on space).

Since the results discussed in the main text from Figure 5 are based on estimating treatment effects separately by two-way comparisons of means, equation (6) probes robustness to estimating them jointly. Results indicate this decision makes a negligible difference, as expected from a factorial design that ensures the three treatments are uncorrelated with each other. As mentioned in Section 4, this design also makes it so that "contamination bias" from multiple treatments is not an issue for our estimates (Goldsmith-Pinkham et al., 2022). Such bias arises from cases where treatments are correlated with each other (e.g., not independently drawn, such as when the design is not factorial and units receive either one treatment or another) and including covariates (such as strata fixed effects) are required in estimation. Neither of these situations applies to our design.

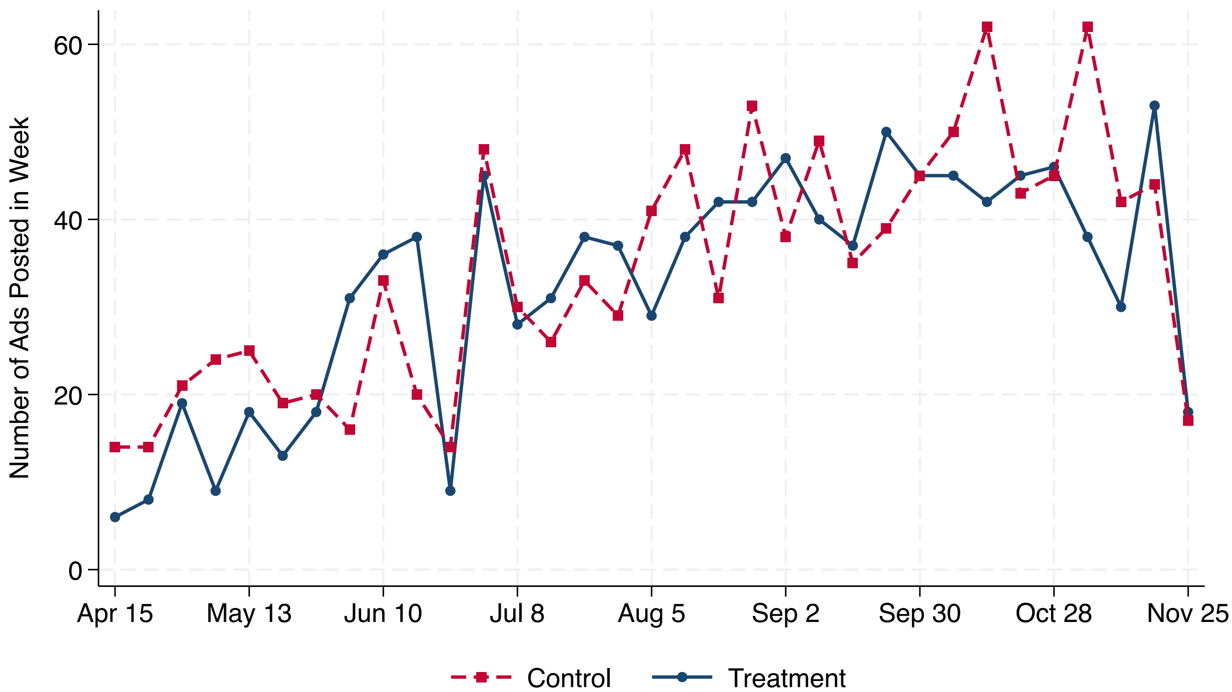
In the terminology of Muralidharan et al. (2023), equation (6) and equation (3) that is reported on Table 5 estimate a "short model," as opposed to a fully interacted "long model." As discussed in Section 2.2, the "short model" is the appropriate choice in this context. The experiment's factorial design was designed to i) allow us to compare the effects of gender-neutral language to explicit diversity statements and a valuable job amenity (working remotely), and ii) to ensure the sample reflected Get On Board ads (of which many have diversity statements and involve remote positions). Thus we are not as interested in effect interactions (for which we have less statistical power). Indeed, our pre-registration states that the experiment was designed to *compare* the effects of gender-neutral language to the other two treatments, and does not mention the interaction of effects.

Muralidharan et al. (2023) discusses related issues on the estimation from experiments with factorial designs. Note, however, that their discussion is centered on cases where researchers are testing new policies that are "new" or not common in their context, and thus estimating interacted effects from "long models" is perhaps more suitable. In our context, all treatments represent relatively common practices in our context, and the factorial design aims to make the sample more representative of the context.

Robustness checks and heterogeneity. Tables A.12 and A.13 replicate Table A.11 splitting the sample by whether the respondents are alumni of the web development or the UX design boot camps, respectively. Results are similar in magnitude, suggesting little heterogeneity by field. Table A.14 replicates Table A.11 adding respondent fixed effects. As expected given the experimental design, these within-estimates are quite similar to other estimates. Note that we cannot estimate the effects of gender-neutral language by ad order (i.e., equation (3) reported in Table 5) while using respondent fixed effects. Given that all respondents see both a gender-neutral and a non-gender-neutral ad, respondent fixed effects are collinear with the interaction between $GNeutral_{ia}$ and $2ndAd_{ia}$.

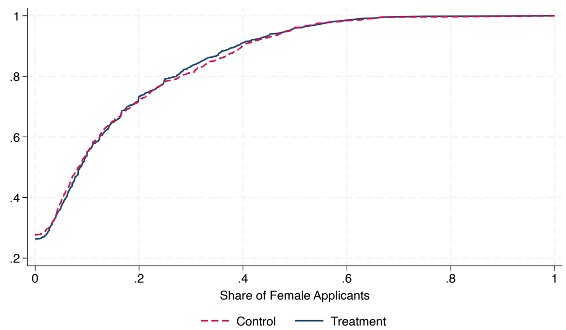
In unreported regressions, we find that the results are also robust to excluding the Brazilian boot camp alumni (who answered a version of the survey in Portuguese) and excluding respondents who answered the survey "too quickly" (e.g., less than three or five minutes).

Figure A.1: Weekly Number of Ads Posted Over Time - Get On Board

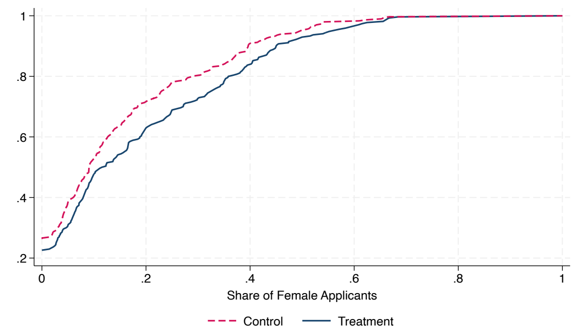


Notes: Figure provides the weekly number of ads posted during the experimental period (April 17 to November 27, 2020), by treatment assignment. Labels on the x-axis refer to the day a week starts (e.g., Apr 15 is the week of April 15-21).

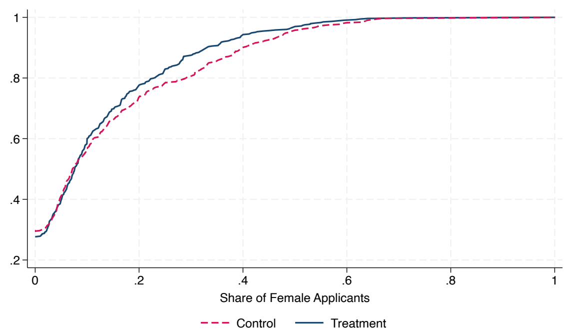
Figure A.2: Share of Female Applicants Distribution - Get On Board



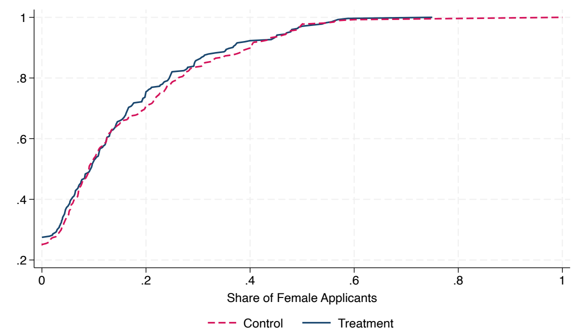
(a) All Ads



(b) Bottom Quartile of % Neighbors Treated



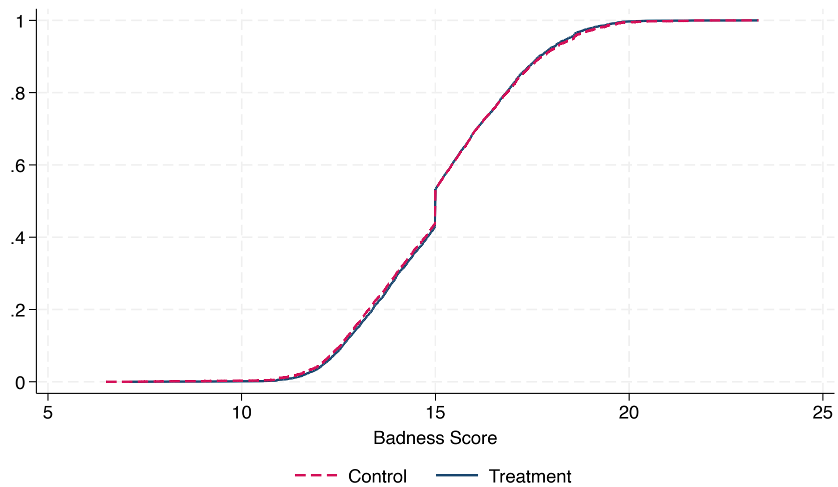
(c) Mid Quartiles of % Neighbors Treated



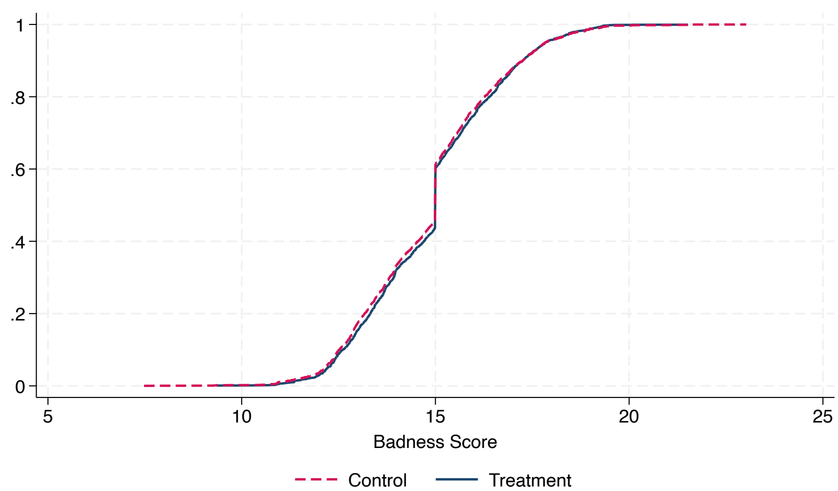
(d) Top Quartile of % Neighbors Treated

Notes: Unit of observation is an ad. Figures provide the cumulative distribution function (CDF) of the share of female applicants to control and treated ads, for all ads (Panel a) and separately by whether the ad's share of neighbor ads treated (SNT_i) falls in the bottom quartile, middle quartiles, or the top quartile of the SNT_i distribution (Panels (c)-(d), respectively).

Figure A.3: Badness Score Distribution - Get On Board



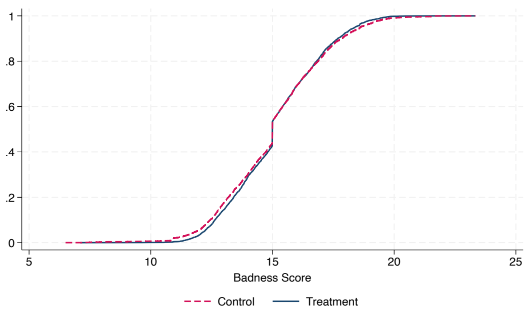
(a) Male Applicants, Full Sample



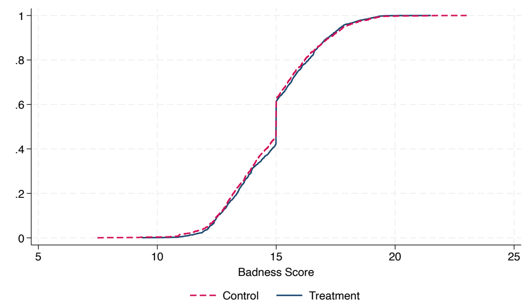
(b) Female Applicants, Full Sample

Notes: The unit of observation is an applicant. Figures provide the cumulative distribution function (CDF) of the "badness scores" of applicants to control and treated ads, separately by applicant gender (see text for details).

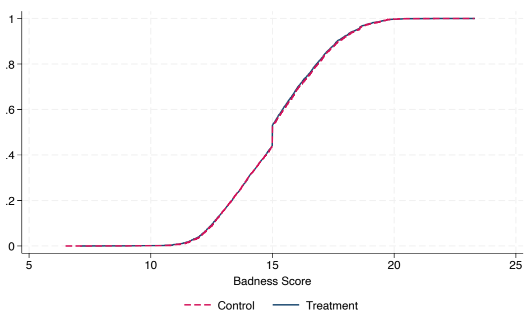
Figure A.4: Badness Score Distribution by Share of Neighbors Ads Treated - Get On Board



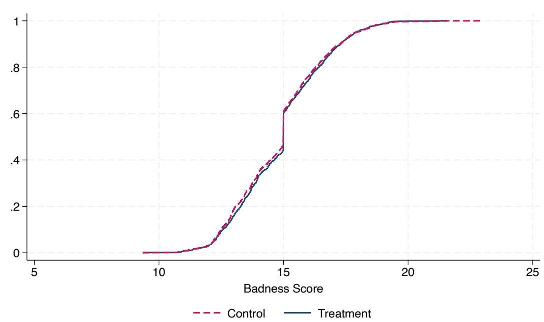
(a) Male Applicants,
Bottom Quartile of % Neighbors Treated



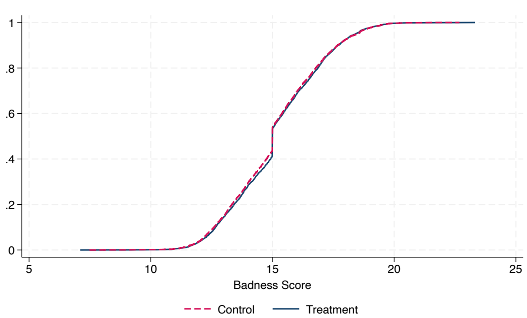
(b) Female Applicants,
Bottom Quartile of % Neighbors Treated



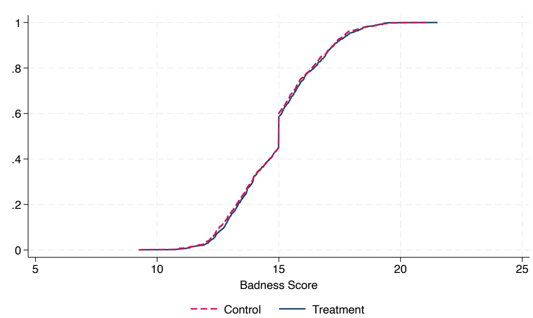
(c) Male Applicants,
Mid Quartiles of % Neighbors Treated



(d) Female Applicants,
Mid Quartiles of % Neighbors Treated



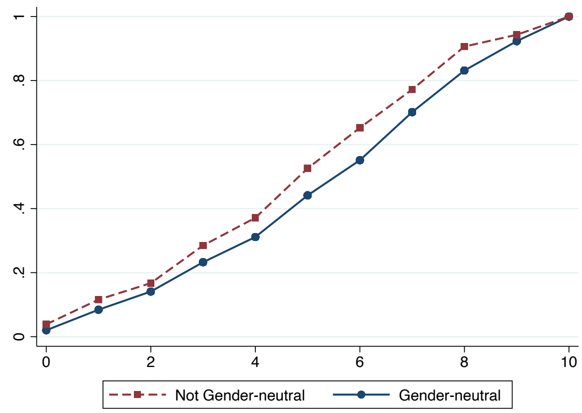
(e) Male Applicants,
Top Quartile of % Neighbors Treatedd



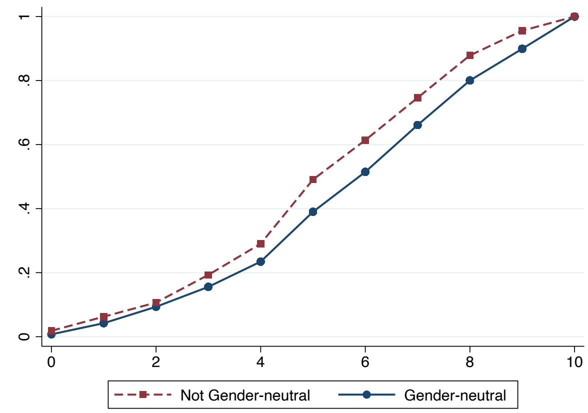
(f) Female Applicants,
Top Quartile of % Neighbors Treated

Notes: The unit of observation is an applicant. Figures provide the cumulative distribution function (CDF) of the "badness scores" of applicants to control and treated ads, separately by applicant gender and whether the ad's share of neighbor ads treated (SNT_i) falls in the bottom quartile, middle quartiles, or the top quartile of the SNT_i distribution (see text for details).

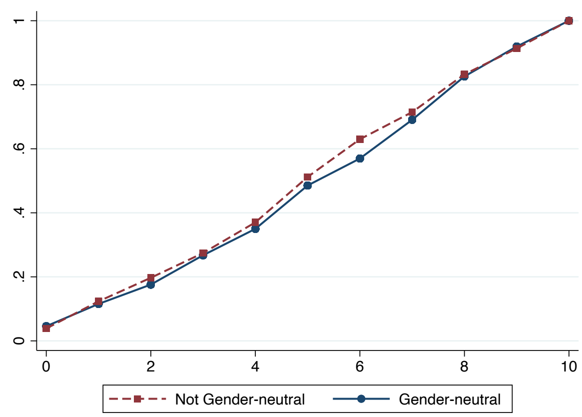
Figure A.5: Outcomes Distribution in Laboratoria Experiment, by Gender-Neutral Treatment Status



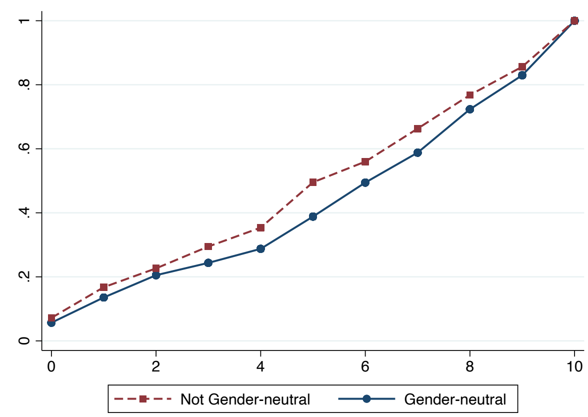
(a) Job Appeal



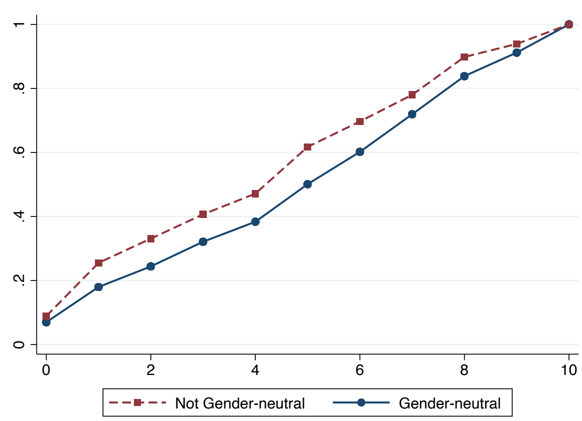
(b) Good Employer



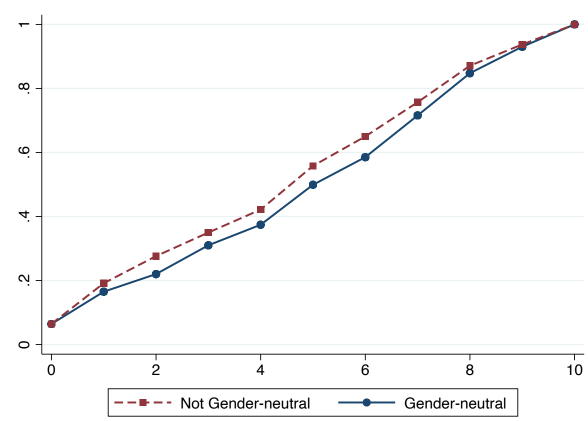
(c) Meet Requirements



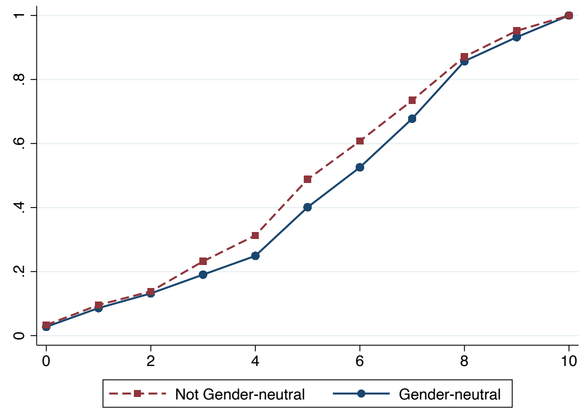
(d) Probability of Applying



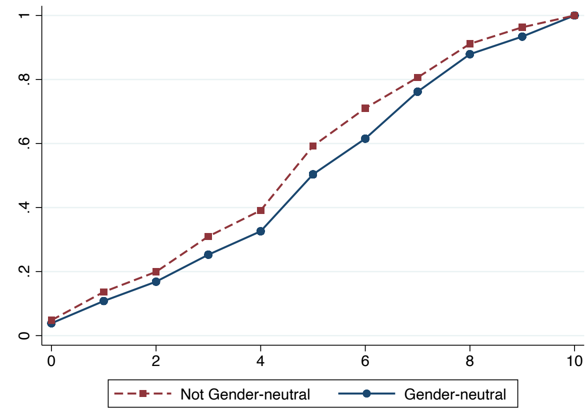
(e) Suitability



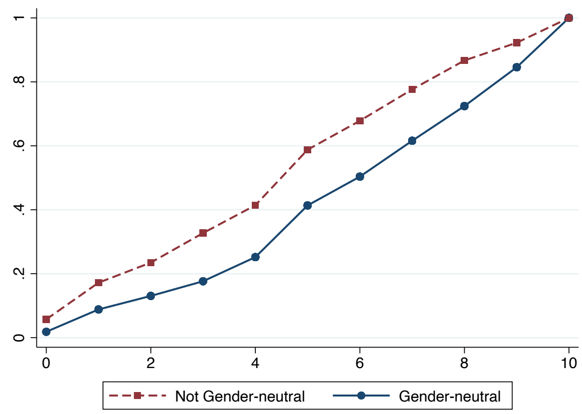
(f) Probability of Being Chosen



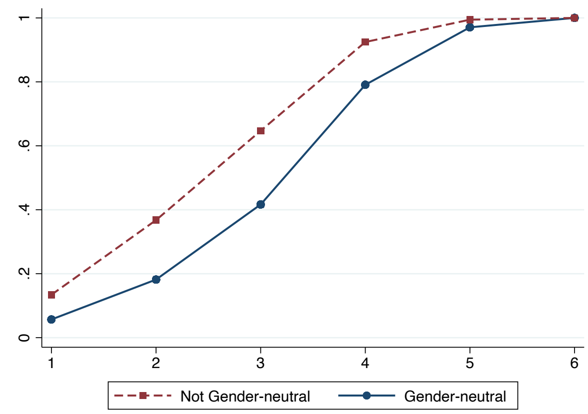
(g) Good salary



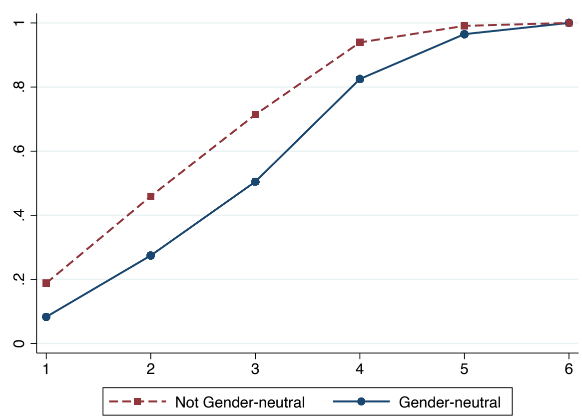
(h) Work-life Balance



(i) Inclusive Culture



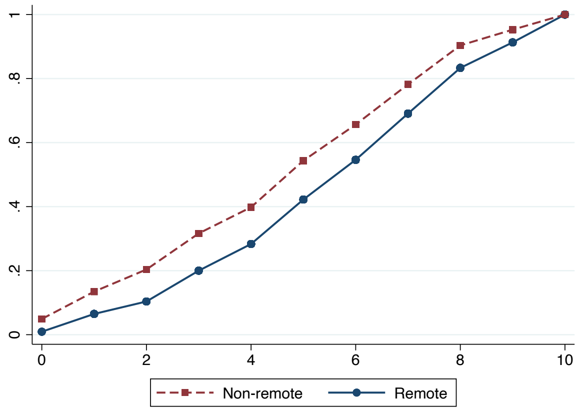
(j) Women Percentage Company



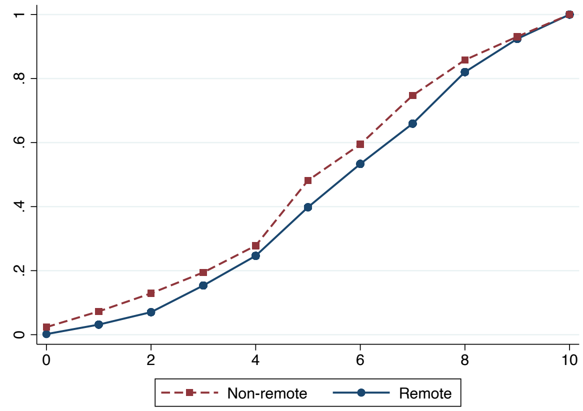
(k) Women Percentage Position

Notes: The unit of observation is a response to an ad (each respondent sees two ads). Figures provide the cumulative distribution function (CDF) for the eleven outcomes collected in the survey (see text for definitions). All observations are included (regardless of remote or diversity statement status).

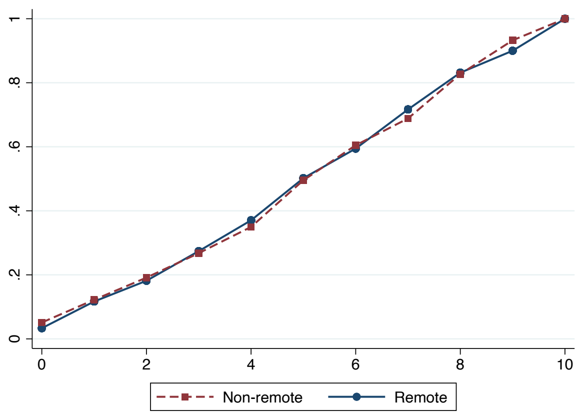
Figure A.6: Outcomes Distribution in Laboratoria Experiment,
by Remote Treatment Status



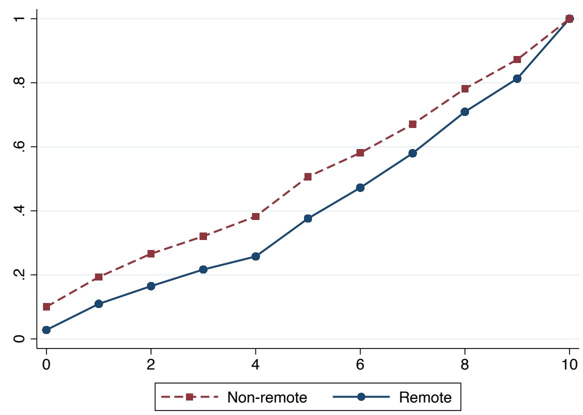
(a) Job Appeal



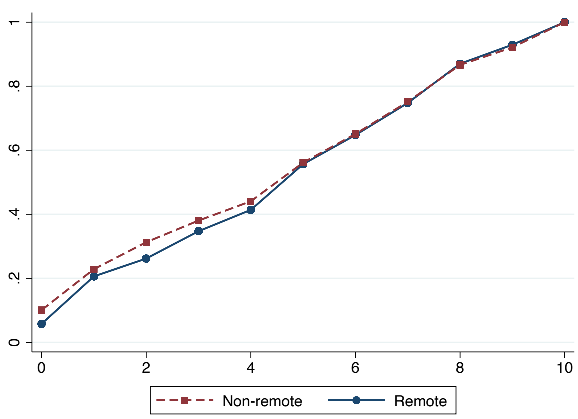
(b) Good Employer



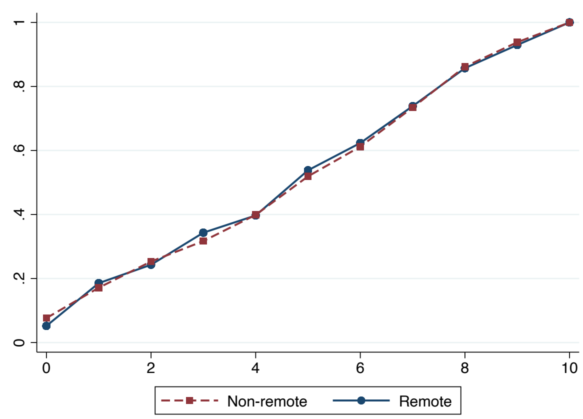
(c) Meet Requirements



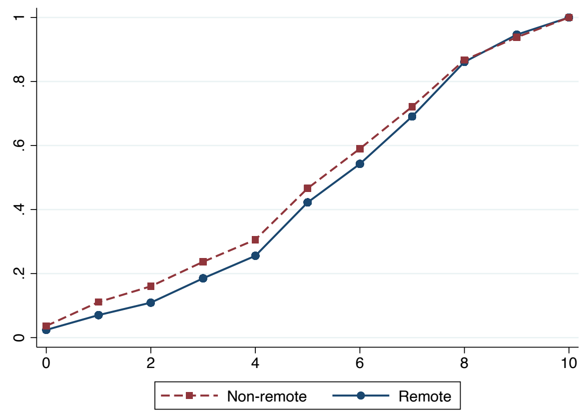
(d) Probability of Applying



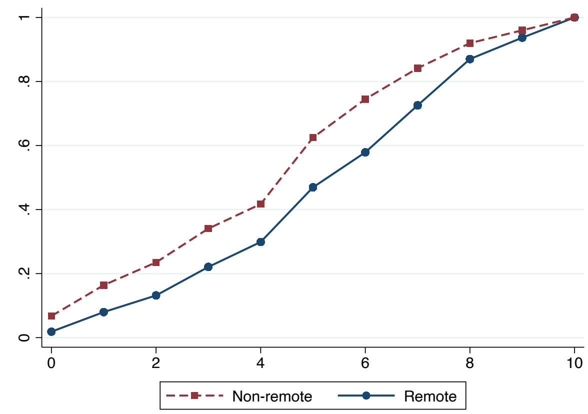
(e) Suitability



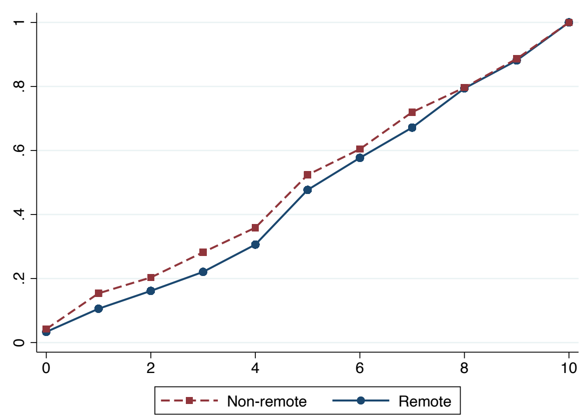
(f) Probability of Being Chosen



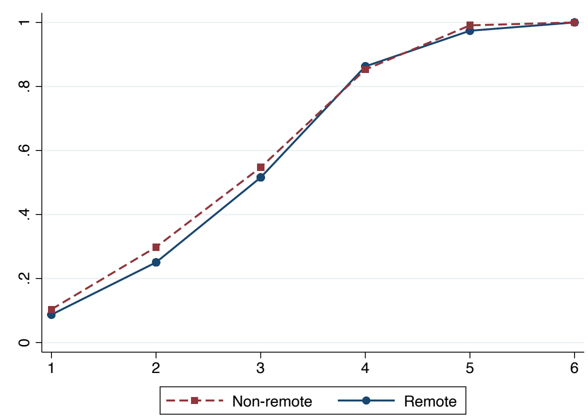
(g) Good salary



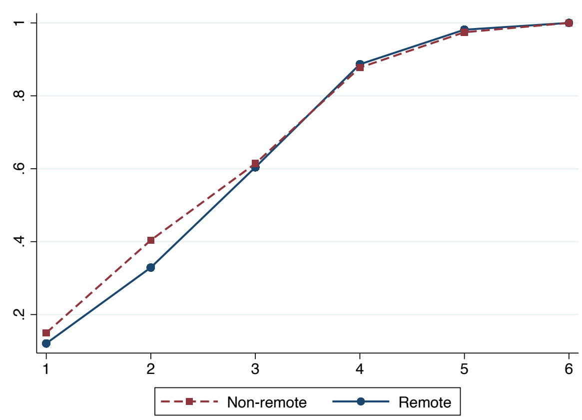
(h) Work-life Balance



(i) Inclusive Culture



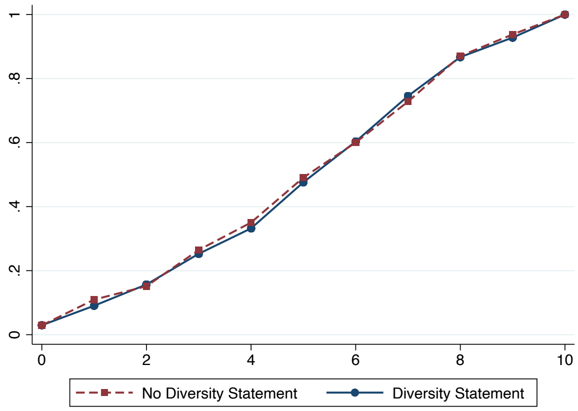
(j) Women Percentage Company



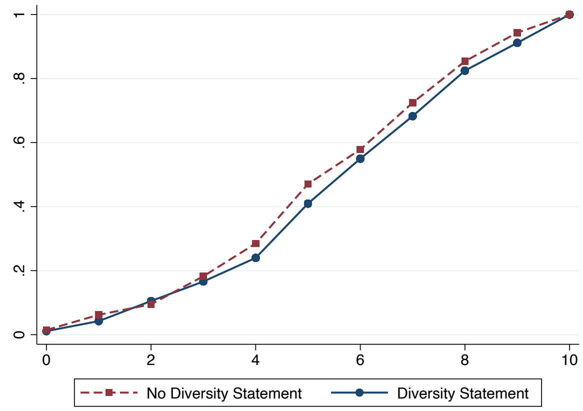
(k) Women Percentage Position

Notes: The unit of observation is a response to an ad (each respondent sees two ads). Figures provide the cumulative distribution function (CDF) for the eleven outcomes collected in the survey (see text for definitions). All observations are included (regardless of gender-neutral or diversity statement status).

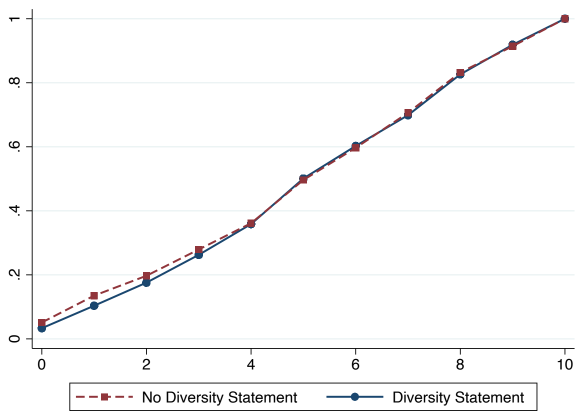
Figure A.7: Outcomes Distribution in Laboratoria Experiment,
by Diversity Statement Treatment Status



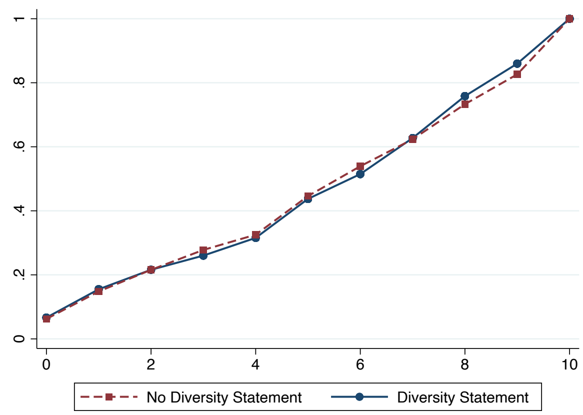
(a) Job Appeal



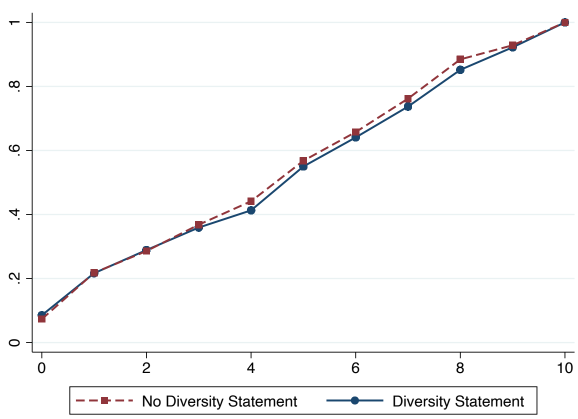
(b) Good Employer



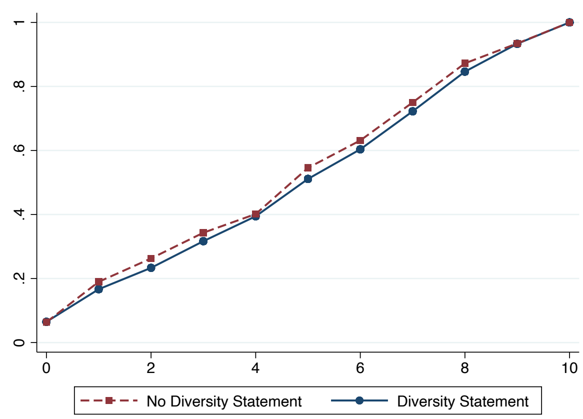
(c) Meet Requirements



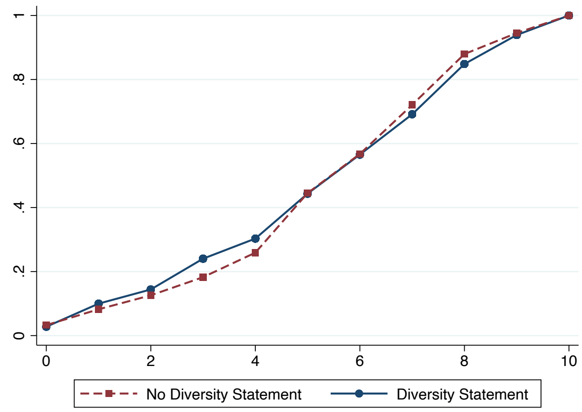
(d) Probability of Applying



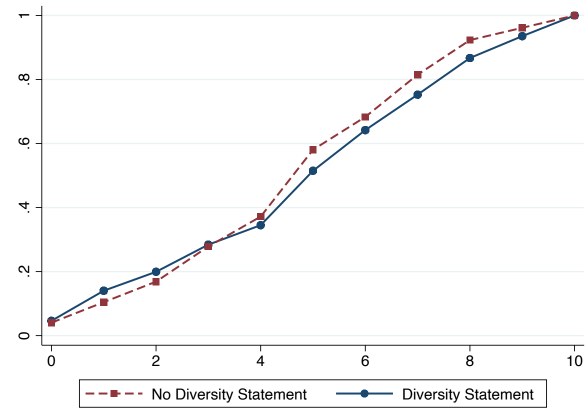
(e) Suitability



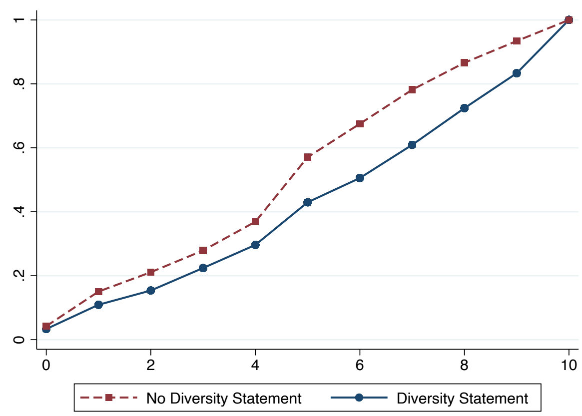
(f) Probability of Being Chosen



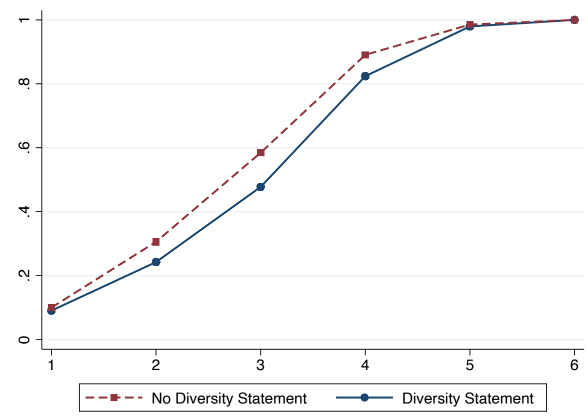
(g) Good salary



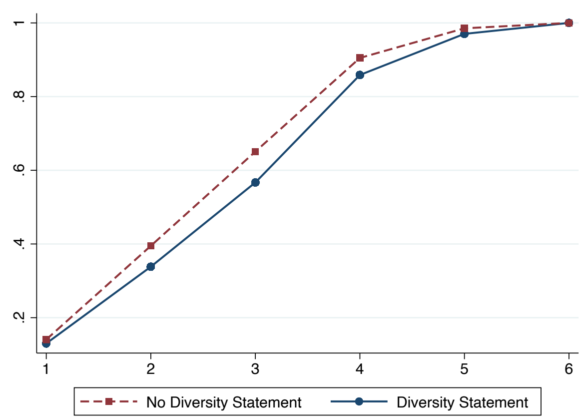
(h) Work-life Balance



(i) Inclusive Culture



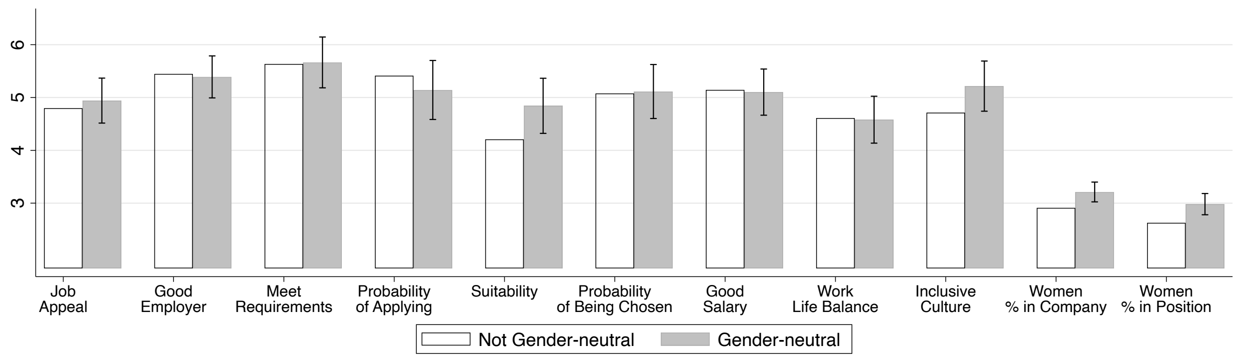
(j) Women Percentage Company



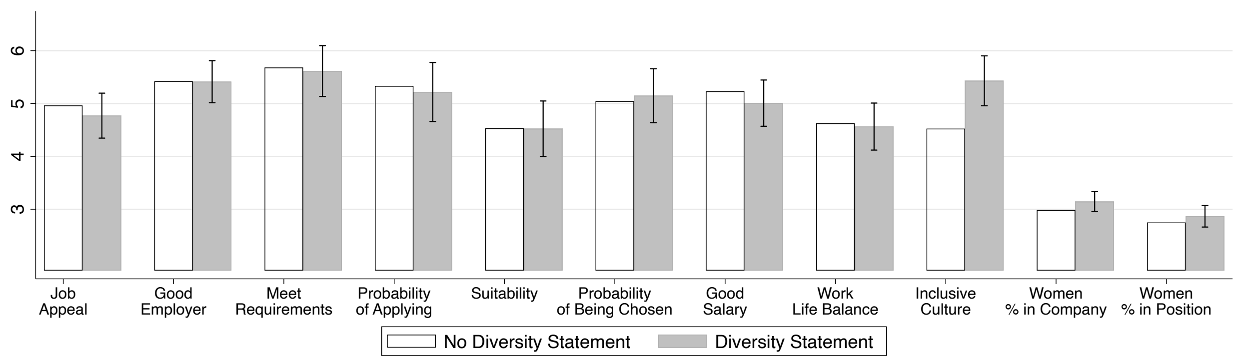
(k) Women Percentage Position

Notes: The unit of observation is a response to an ad (each respondent sees two ads). Figures provide the cumulative distribution function (CDF) for the eleven outcomes collected in the survey (see text for definitions). All observations are included (regardless of remote or gender-neutral statement status).

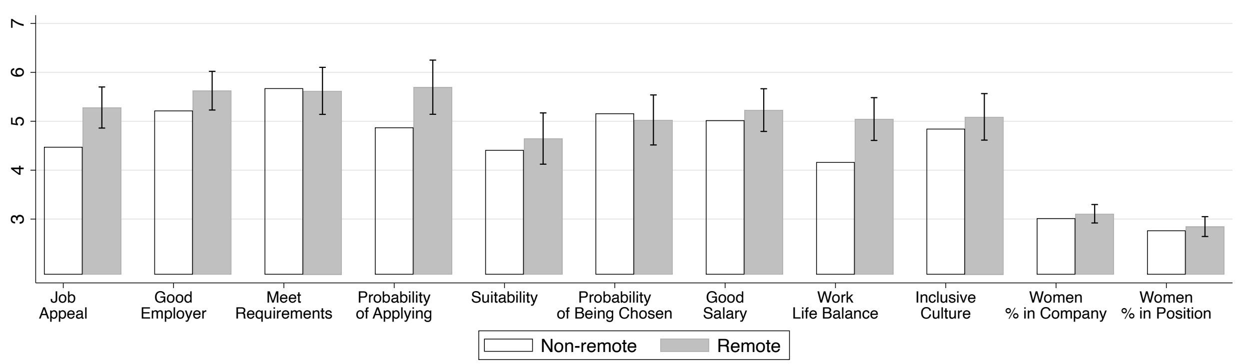
Figure A.8: Outcome Averages by Different Treatment Statuses - Laboratoria (First Ads Only)



(a) Gender Neutral Language Treatment



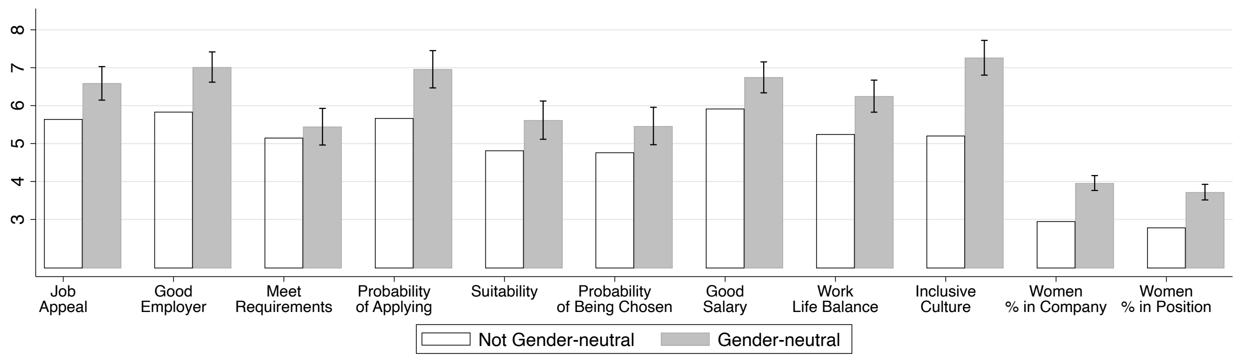
(b) Diversity Statement Treatment



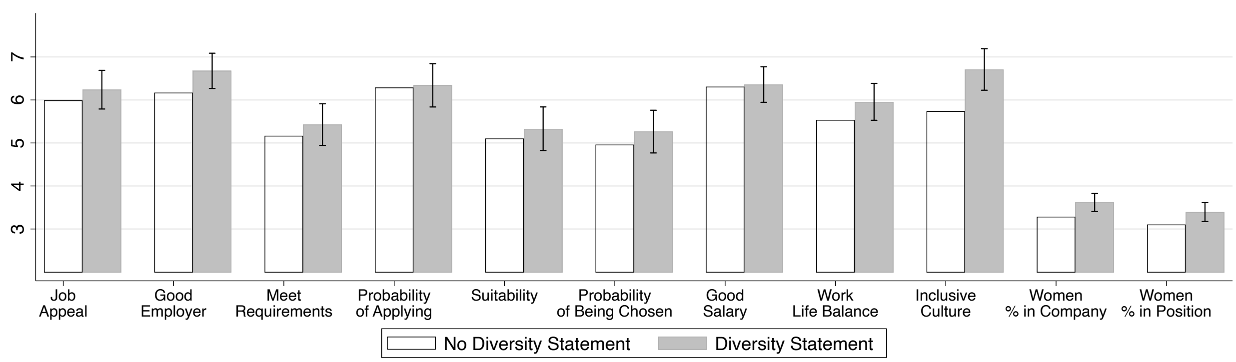
(c) Remote Job Treatment

Notes: The unit of observation is a response to the first ad a respondent sees (each of the 546 respondents sees two ads). Figures provide the "raw" averages for the eleven outcomes collected in the survey (see text for definitions) by different treatment statuses. Whiskers present the 95% CI of the difference between averages (the treatment effect) based on robust standard errors. All observations are included (e.g., Panel (a) includes all observations regardless of remote or diversity statement status).

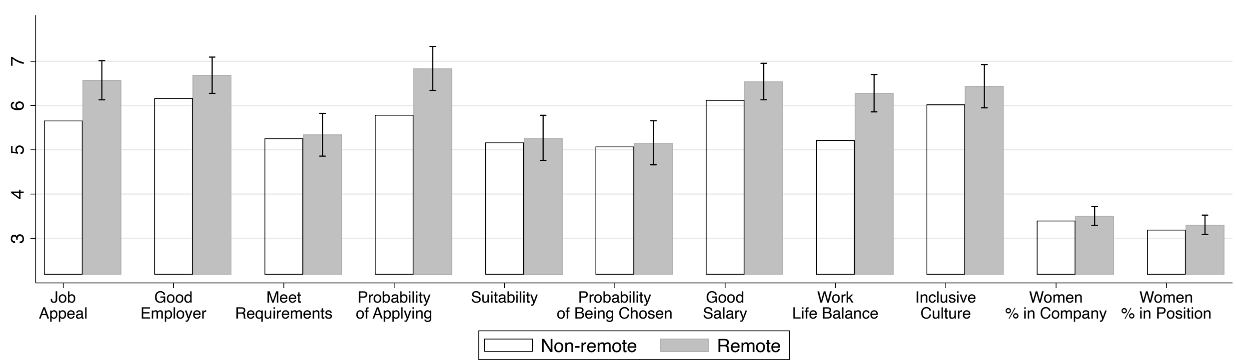
Figure A.9: Outcome Averages by Different Treatment Statuses - Laboratoria (Only Second Ads)



(a) Gender Neutral Language Treatment



(b) Diversity Statement Treatment



(c) Remote Job Treatment

Notes: The unit of observation is a response to the second ad a respondent sees (each of the 546 respondents sees two ads). Figures provide the "raw" averages for the eleven outcomes collected in the survey (see text for definitions) by different treatment statuses. Whiskers present the 95% CI of the difference between averages (the treatment effect) based on robust standard errors. All observations are included (e.g., Panel (a) includes all observations regardless of remote or diversity statement status).

Table A.1: Summary Statistics and Covariate Balance - Get On Board

Variable	Mean (C)	Mean (T)	Difference (T-C)	SE	p-value	Obs
Remote	0.394	0.411	0.017	0.021	0.416	2,201
Junior Position	0.199	0.166	-0.033	0.016	0.046	2,201
Semi-Senior Position	0.582	0.567	-0.016	0.021	0.461	2,201
Missing Experience Requirement	0.010	0.007	-0.003	0.004	0.403	2,201
No Experience Required	0.033	0.041	0.008	0.008	0.301	2,201
Posted Salary Range	0.441	0.426	-0.015	0.021	0.480	2,201
Salary Range (Min, USD 1,000s)	1.799	1.871	0.072	0.055	0.190	954
Salary Range (Max, USD 1,000s)	2.393	2.487	0.094	0.076	0.218	954
Share of Neighbor Ads Treated	0.482	0.492	0.010	0.010	0.353	2,201
Number of Neighbor Ads	7.598	7.886	0.288	0.216	0.183	2,201

Notes: The unit of observation is an ad. The table provides means for ads assigned to treatment and control status, as well as their difference (and its standard error and p-value). Remote, Junior Position, Semi-Senior Position, Missing Experience Required, and Posted Salary Range are dummy indicators. The minimum and maximum of the posted monthly salary range are measured in thousands of US dollars. See text for further variable definitions.

Table A.2: Share of Neighbor Ads Treated is Uncorrelated with Ad Characteristics

Variable	Coeff	SE	p-value
Remote	0.002	0.043	0.956
Junior Position	0.025	0.034	0.468
Semi-Senior Position	-0.017	0.043	0.702
Senior Position	0.004	0.034	0.915
Missing Experience Requirement	-0.006	0.008	0.441
No Experience Required	-0.006	0.017	0.736
Posted Salary Range	-0.018	0.043	0.677
Salary Range (Min, USD 1,000)	0.138	0.116	0.232
Salary Range (Max, USD 1,000)	0.241	0.172	0.161

Notes: The unit of observation is an ad. Each row provides the coefficient, standard error, and p-value of a separate regression where the dependent variable is listed in the first column and the independent variable is the share of neighbor ads treated (SNT_i). All regressions have 2,201 observations, except those for the minimum and maximum of the salary range (954 observations). Remote, Junior Position, Semi-Senior Position, Missing Experience Required, and Posted Salary Range are dummy indicators. The minimum and maximum of the posted monthly salary range are measured in thousands of US dollars.

Table A.3: Share of Female Applicants by Job Title Group - Get On Board

Job Title Group	Fem. Share Applicants (Control)	Share of Sample
Full-Stack Developer	0.043	0.152
Mobile Developer	0.044	0.052
Architect	0.045	0.005
Back Developer	0.058	0.074
Web Developer	0.062	0.021
Other Developer	0.089	0.115
Programmer	0.094	0.018
Data Scientist	0.095	0.006
Engineer	0.102	0.172
Front Developer	0.114	0.082
Sysadmin	0.188	0.060
Analyst	0.245	0.067
Scrum	0.272	0.006
Bizadmin	0.298	0.058
Designer	0.391	0.086
Marketing/Customers	0.400	0.026

Notes: For each job title group, we provide the average share of female applicants using data from the control group only, as well as the share of ads in each field (in the entire sample). See main text and Appendix C for definitions and construction of job title groups.

Table A.4: Share of Female Applicants by Field - Get On Board

Field	Fem. Share Applicants (Control)	Share of Sample
Mobile	0.035	0.054
Programming	0.069	0.570
Data Analytics	0.149	0.047
Sysadmin	0.177	0.090
Operations	0.224	0.047
Innovation/Agile	0.270	0.020
Sales	0.308	0.015
Customer Support	0.321	0.024
Advertising/Media	0.376	0.006
Design	0.389	0.087
Digital Marketing	0.410	0.038
Human Resources	0.495	0.003

Notes: For each job field, we provide the average share of female applicants using data from the control group only, as well as the share of ads in each field (in the entire sample).

Table A.5: 2SLS and First-Stage Estimates for Treatment-on-Treated Effects - Get on Board

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fem. Share Applicants	Fem. Share Applicants	GN Ad	GN Ad	GN Ad × Mid Quartiles of % Neighbors Treated	GN Ad × Mid Quartiles of % Neighbors Treated	GN Ad × Top Quartile of % Neighbors Treated	GN Ad × Top Quartile of % Neighbors Treated
GN Ad	0.104**	0.115***						
	(0.044)	(0.044)						
GN Ad × Mid Quartiles of % Neighbors Treated	-0.159***	-0.168***						
	(0.054)	(0.053)						
GN Ad × Top Quartile of % Neighbors Treated	-0.143**	-0.131**						
	(0.067)	(0.066)						
Treatment			0.350***	0.344***	0.003	-0.001	0.001	-0.003
			(0.039)	(0.039)	(0.004)	(0.003)	(0.003)	(0.002)
Treat × Mid Quartiles of % Neighbors Treated			-0.032	-0.026	0.316***	0.319***	-0.001	0.002
			(0.048)	(0.048)	(0.029)	(0.029)	(0.004)	(0.002)
Treat × Top Quartile of % Neighbors Treated			-0.081	-0.082	-0.005	-0.003	0.274***	0.272***
			(0.057)	(0.057)	(0.005)	(0.004)	(0.041)	(0.041)
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
N	2,201	2,201	2,201	2,201	2,201	2,201	2,201	2,201

Notes: The unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status), while even-numbered columns include controls selected by PDS-LASSO. Columns 1-2 present the results from the 2SLS estimation of equation (2). The three excluded instruments are the treatment dummy and its interaction with two dummies indicating if the share of neighbor ads treated (SNT_i) falls in the middle quartiles or the top quartile of its distribution. The linear combinations presented in Table 3 are based on these estimated coefficients. Columns 3-8 provide the related first-stage estimates for the three endogenous variables: a dummy equal one if the ad is gender-neutral (based on the full-text classification) and its interaction with two dummies indicating if the share of neighbor ads treated (SNT_i) falls in the middle quartiles or the top quartile of its distribution. Standard errors in parentheses. All regressions include dummies indicating if SNT_i falls in the middle quartiles and the top quartile of its distribution (omitted to economize on space). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Intent-to-Treat Effects by Title's Language and Remote Status - Get on Board

	Job Title in English		Job Title in Spanish		Remote Job		Non-remote Job	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants
Treatment (β_0)	0.043**	0.051**	0.030	0.030	0.044*	0.044*	0.031	0.035*
	(0.022)	(0.021)	(0.022)	(0.021)	(0.024)	(0.024)	(0.019)	(0.019)
Treat × Mid. Quartiles of	-0.053**	-0.061**	-0.056**	-0.058**	-0.082***	-0.080***	-0.035	-0.039*
% Neighbors Treated (β_M)	(0.025)	(0.024)	(0.026)	(0.026)	(0.029)	(0.028)	(0.023)	(0.023)
Treat × Top Quartile of	-0.046	-0.052*	-0.046*	-0.037	-0.047	-0.037	-0.050**	-0.050**
% Neighbors Treated (β_T)	(0.031)	(0.029)	(0.028)	(0.027)	(0.034)	(0.033)	(0.025)	(0.025)
Mid. Quartiles of	-0.022	0.024	0.003	0.028	-0.003	0.028	-0.014	0.020
% Neighbors Treated (γ_M)	(0.016)	(0.016)	(0.019)	(0.019)	(0.019)	(0.019)	(0.016)	(0.016)
Top Quartile of	0.007	0.018	-0.012	-0.011	0.006	0.007	-0.010	-0.002
% Neighbors Treated (γ_T)	(0.019)	(0.019)	(0.020)	(0.020)	(0.023)	(0.023)	(0.018)	(0.017)
Implied Treatment Effects								
Bottom Quartile of %	0.043	0.051	0.030	0.030	0.044	0.044	0.031	0.035
Neighbors Treated (β_0)	(0.022)	(0.021)	(0.022)	(0.021)	(0.024)	(0.024)	(0.019)	(0.019)
	[0.072]*	[0.031]**	[0.167]	[0.161]	[0.098]*	[0.091]*	[0.138]	[0.080]*
Mid. Quartiles of %	-0.010	-0.010	-0.026	-0.028	-0.038	-0.036	-0.004	-0.004
Neighbors Treated ($\beta_0+\beta_M$)	(0.013)	(0.013)	(0.014)	(0.014)	(0.015)	(0.014)	(0.013)	(0.012)
	[0.525]	[0.508]	[0.109]	[0.073]*	[0.020]**	[0.027]**	[0.777]	[0.771]
Top Quartile of %	-0.003	-0.002	-0.016	-0.007	-0.002	0.007	-0.020	-0.015
Neighbors Treated ($\beta_0+\beta_T$)	(0.021)	(0.021)	(0.018)	(0.017)	(0.023)	(0.022)	(0.017)	(0.016)
	[0.920]	[0.952]	[0.459]	[0.732]	[0.917]	[0.775]	[0.341]	[0.461]
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.146	0.146	0.145	0.145	0.147	0.147	0.145	0.145
N	1,106	1,106	1,095	1,095	885	885	1,316	1,316

Notes: The unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status), while even-numbered columns include controls selected by PDS-LASSO. The outcome (dependent variable) in all columns is the share of applicants that are female. Columns 1-2 only use ads with titles in English, while columns 3-4 only use ads with titles in Spanish. Columns 5-6 only use ads for remote positions, while columns 7-8 only use ads for non-remote (in-person) positions. The independent variables are a dummy for treatment assignment, two dummies indicating if the ad's share of neighbor ads treated (SNT_i) falls in the middle quartiles or the top quartile of the SNT_i distribution, and their interactions with treatment. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with SNT_i in the bottom quartile, medium quartiles, and top quartile. The last two rows provide the average of the outcome variable for control ads and the number of observations. Standard errors are in parentheses and randomization inference p-values are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Treatment Effects on Subsequent Ads - Get On Board

	(1) Posted 2nd Ad	(2) # of Ads	(3) GN Ad Title	(4) GN Ad Title	(5) GN Ad Text	(6) GN Ad Text
First Ad Treated	-0.032	0.144	0.028	-0.024	0.003	-0.057
	(0.037)	(0.296)	(0.077)	(0.077)	(0.080)	(0.071)
Sample: Firms	YES	YES				
Sample: 2nd ads			YES		YES	
Sample: 2nd or later ads				YES		YES
Control Mean	0.435	2.418	0.635	0.661	0.446	0.426
N	711	711	163	527	163	527

Notes: The independent variable in all regressions is a dummy equal to one if the first ad the firm posted in the sample period was assigned to treatment. The unit of observation in columns 1-2 is a firm. The dependent variables are, respectively, a dummy equal one if the firm posted a second ad and the total number of ads the firm posted in the sample period. The unit of observation in columns 3-6 is an ad. Columns 4 and 6 restrict the sample to ads that were the second or higher-order ads that a firm posted in the sample period. Columns 3 and 5 further restrict the sample only to second ads. The dependent variable in columns 3-4 is a dummy if the ad had a gender-neutral title, and in 5-6, it is a dummy equal one if the ad has a gender-neutral full text. See Appendix D for further details. Standard errors clustered at the firm level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Placebo Tests: Effect by Share of Future Neighbor Ads Treated - Get on Board

	Close Ads Window 30 Days Ahead		Close Ads Window 60 Days Ahead	
	(1)	(2)	(3)	(4)
	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants
Treatment (β_0)	0.007	0.010	0.010	0.009
	(0.015)	(0.015)	(0.016)	(0.016)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.011	-0.014	-0.010	-0.008
	(0.018)	(0.018)	(0.020)	(0.020)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.010	-0.014	-0.001	0.000
	(0.023)	(0.023)	(0.023)	(0.022)
Mid. Quartiles of % Neighbors Treated (γ_M)	-0.045***	-0.024*	-0.018	0.002
	(0.012)	(0.012)	(0.014)	(0.013)
Top Quartile of % Neighbors Treated (γ_T)	-0.005	-0.006	-0.014	-0.016
	(0.016)	(0.016)	(0.015)	(0.015)
Implied Treatment Effects				
Bottom Quartile of % Neighbors Treated (β_0)	0.007	0.010	0.010	0.009
	(0.015)	(0.015)	(0.016)	(0.016)
Mid. Quartiles of % Neighbors Treated (β_0+β_M)	-0.005	-0.004	-0.000	0.001
	(0.009)	(0.009)	(0.011)	(0.011)
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	-0.003	-0.004	0.009	0.009
	(0.017)	(0.017)	(0.016)	(0.015)
Baseline Controls?	YES		YES	
PDS-LASSO Controls?		YES		YES
Control Mean	0.142	0.142	0.141	0.141
N	1,913	1,913	1,499	1,499

Notes: The unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status), while even-numbered columns include controls selected by PDS-LASSO. The outcome (dependent variable) in all columns is the share of applicants that are female. Columns 1-2 use as ad i's neighbors the ads posted between 27 and 33 days after ad i's date. Columns 3-4 use as ad i's neighbors the ads posted between 57 and 63 days after ad i's date. The independent variables are a dummy for treatment assignment, two dummies indicating if the ad's share of neighbor ads treated (SNT_i) falls in the middle quartiles or the top quartile of the SNT_i distribution, and their interactions with treatment. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with SNT_i in the bottom quartile, medium quartiles, and top quartile. The number of observations differs across columns (and from Table 2) since ads at the last 30 and 60 days of our sample must be dropped from columns 1-2 and 3-4, respectively. The last two rows provide the average of the outcome variable for control ads and the number of observations. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Placebo Tests: Effects by Share Female Applicants in Job Title Group - Get on Board

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fem. Share Applicants	Fem. Share Applicants	asinh(Fem. Applicants)	asinh(Fem. Applicants)	asinh(Male Applicants)	asinh(Male Applicants)	Avg. Badness Score	Avg. Badness Score
Treatment (β_0)	-0.003	-0.003	0.064	0.058	0.093	0.083	0.072**	0.076**
	(0.005)	(0.005)	(0.057)	(0.057)	(0.058)	(0.058)	(0.034)	(0.034)
Treat × Mid. Quartiles of % Fem. in Title Group (β_M)	0.008	0.006	-0.131	-0.117	-0.278**	-0.259**	-0.037	-0.041
	(0.017)	(0.017)	(0.153)	(0.152)	(0.125)	(0.125)	(0.056)	(0.056)
Treat × Top Quartile of % Fem. in Title Group (β_T)	0.022	0.019	-0.136	-0.140	-0.252*	-0.216	-0.084	-0.089
	(0.017)	(0.017)	(0.144)	(0.147)	(0.134)	(0.137)	(0.058)	(0.056)
Mid. Quartiles of % Fem. in Title Group (γ_M)	0.168***	0.170***	1.729***	1.660***	0.685***	0.617***	-0.067*	-0.065
	(0.012)	(0.012)	(0.103)	(0.107)	(0.083)	(0.085)	(0.038)	(0.041)
Top Quartile of % Fem. in Title Group (γ_T)	0.315***	0.317***	2.481***	2.411***	0.572***	0.527***	-0.211***	-0.212***
	(0.012)	(0.012)	(0.096)	(0.103)	(0.085)	(0.090)	(0.038)	(0.037)
Implied Treatment Effects								
Bottom Quartile of % Fem. in Title Group (β_0)	-0.003	-0.003	0.064	0.058	0.093	0.083	0.072	0.076
	(0.005)	(0.005)	(0.057)	(0.057)	(0.058)	(0.058)	(0.034)	(0.034)
Mid. Quartiles of % Fem. in Title Group ($\beta_0 + \beta_M$)	0.004	0.003	-0.066	-0.060	-0.185	-0.177	0.035	0.035
	(0.017)	(0.016)	(0.142)	(0.141)	(0.111)	(0.110)	(0.044)	(0.044)
Top Quartile of % Fem. in Title Group ($\beta_0 + \beta_T$)	0.018	0.017	-0.072	-0.083	-0.158	-0.133	-0.012	-0.013
	(0.016)	(0.016)	(0.132)	(0.136)	(0.121)	(0.124)	(0.047)	(0.045)
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.146	0.146	-	-	-	-	15.121	15.121
N	2,201	2,201	2,201	2,201	2,201	2,201	2,201	2,201

Notes: Unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status), while even-numbered columns include controls selected by PDS-LASSO. Outcomes are the share of applicants that are female (columns 1-2), the inverse hyperbolic sine of the number of female and male applicants (columns 3-4 and 5-6, respectively), and the applicants' average "badness score" (a measure of applicant quality, columns 7-8). The top panel provides the estimated coefficients from a regression where the independent variables are a dummy for treatment assignment, two dummies indicating if the ad's share of female applicants in the job title group falls in the middle quartiles or the top quartile of its distribution, and their interactions with treatment. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with a share of female applicants in the job title group in the bottom quartile, medium quartiles, and top quartile. The share of female applicants in the job title group is constructed only using the control group observations (see Appendix D for details). The last two rows provide the average of the outcome variable for control ads and the number of observations. Standard errors are in parentheses and randomization inference p-values are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Summary Statistics by Treatment Status - Laboratoria

	(1) GN_R_D	(2) GN_R_ND	(3) GN_NR_D	(4) GN_NR_ND	(5) NGN_R_D	(6) NGN_R_ND	(7) NGN_NR_D	(8) NGN_NR_ND
Years of Experience	5.855	5.964	6.050	6.169	5.985	5.934	6.207	5.912
	(1.836)	(1.875)	(1.750)	(1.783)	(1.857)	(1.868)	(1.690)	(1.829)
Tech Sector	0.794	0.864	0.820	0.757	0.773	0.796	0.800	0.869
	(0.406)	(0.344)	(0.385)	(0.430)	(0.421)	(0.405)	(0.401)	(0.339)
Looking for Tech Sector	0.466	0.400	0.424	0.478	0.432	0.482	0.471	0.380
	(0.501)	(0.492)	(0.496)	(0.501)	(0.497)	(0.502)	(0.501)	(0.487)
Share of entire sample (in %) from country of boot camp and treatment arm:								
Chile	2.56	3.39	3.94	3.39	3.11	3.75	3.02	3.39
Colombia	1.37	1.10	0.92	1.19	0.64	1.47	1.19	1.28
Equador	0.18	0.00	0.00	0.09	0.09	0.00	0.09	0.09
Mexico	3.48	3.30	3.21	3.48	4.03	3.48	3.39	2.56
Peru	4.12	3.66	2.93	3.57	3.75	2.66	3.66	4.21
Brazil	0.92	1.01	0.92	1.10	1.19	1.10	0.64	1.01
Country not specified	0.09	0.00	0.09	0.00	0.00	0.09	0.09	0.00
Observations	131	140	139	136	132	137	140	137

Notes: The unit of observation is a response to an ad (each of the 546 respondents sees two ads). Each column presents the averages for one of the eight different treatment arms from a $2 \times 2 \times 2$ design. GN, R, and D indicate the gender-neutral, remote, and diversity statement statuses, respectively. NGN, NR, ND, indicate the non-gender-neutral, non-remote, and no diversity statement statuses, respectively. For example, column (6) provides the averages for NGN-R-ND (non-gender-neutral, remote, no diversity statement). Standard deviations in parentheses.

Variable definitions: Years of Experience is years since graduating from the Laboratoria boot camp. Tech Sector and Looking for Tech Sector are dummy indicators for whether the respondent currently has a job and is searching for a job in the tech sector, respectively. The survey allowed those with a current job in the sector to report they are searching for another job (Appendix G). The bottom panel provides the share (in percentage points) of respondents in each treatment arm by country of boot camp graduation cell (i.e., all numbers in the panel add up to 100).

Balance tests: For each variable in the table rows (including country indicators), we cannot reject the hypothesis that averages are the same across columns at usual significance levels. We do so by regressing the variable in question against all eight treatment arm dummies and performing a joint F-test. p-values range from 0.31 to 0.94, except for working in the tech sector (p=0.14).

Table A.11: Treatment Effects (Full Sample) - Laboratoria

	(1) Job Appeal	(2) Good Employer	(3) Meet Require-ments	(4) Probability of Applying	(5) Suitability	(6) Probability of Being Chosen	(7) Good Salary	(8) Work Life Balance	(9) Inclusive Culture	(10) Women % Company	(11) Women % Position
Gender-neutral	0.538***	0.559***	0.161	0.504***	0.715***	0.367**	0.387**	0.480***	1.274***	0.653***	0.639***
	(0.159)	(0.147)	(0.174)	(0.192)	(0.186)	(0.181)	(0.157)	(0.158)	(0.171)	(0.071)	(0.075)
Remote	0.874***	0.477***	0.011	0.948***	0.181	-0.022	0.325**	0.989***	0.359**	0.107	0.101
	(0.159)	(0.147)	(0.174)	(0.191)	(0.186)	(0.182)	(0.157)	(0.158)	(0.171)	(0.071)	(0.075)
Diversity Statement	0.072	0.280*	0.090	0.010	0.131	0.204	-0.054	0.223	0.976***	0.257***	0.215***
	(0.159)	(0.147)	(0.174)	(0.192)	(0.186)	(0.182)	(0.157)	(0.158)	(0.171)	(0.071)	(0.075)
Control mean	4.800	5.148	5.304	5.157	4.346	4.822	5.370	4.284	4.269	2.676	2.515
Observations	1,090	1,090	1,089	1,089	1,086	1,088	1,089	1,088	1,085	1,089	1,085

Notes: Unit of observation is a response to an ad (each respondent sees two ads). Each column presents an estimate from equation (6) for a different outcome (see text for definitions). Gender-neutral, Remote, and Diversity Statements are dummies indicating the ad was assigned to the respective status. The control mean is the outcome mean for ads under the non-gender-neutral language, non-remote treatment, and no diversity treatment status. The number of observations varies across columns due to missing data on outcomes (a few instances when respondents did not answer a survey question). Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$..

Table A.12: Treatment Effects (Alumni of Web Development Boot Camp Only) - Laboratoria

	(1) Job Appeal	(2) Good Employer	(3) Meet Require-ments	(4) Probability of Applying	(5) Suitability	(6) Probability of Being Chosen	(7) Good Salary	(8) Work Life Balance	(9) Inclusive Culture	(10) Women % Company	(11) Women % Position
Gender-neutral	0.580***	0.662***	0.179	0.506**	0.682***	0.455**	0.340*	0.541***	1.270***	0.670***	0.680***
	(0.184)	(0.171)	(0.188)	(0.222)	(0.207)	(0.203)	(0.184)	(0.185)	(0.198)	(0.083)	(0.083)
Remote	0.934***	0.455***	0.106	1.024***	0.258	-0.010	0.297	1.152***	0.366*	0.179**	0.191**
	(0.184)	(0.171)	(0.188)	(0.222)	(0.207)	(0.204)	(0.184)	(0.184)	(0.198)	(0.083)	(0.083)
Diversity Statement	0.140	0.303*	-0.096	0.010	0.001	0.003	-0.040	0.193	0.920***	0.237***	0.189**
	(0.184)	(0.171)	(0.188)	(0.222)	(0.207)	(0.204)	(0.185)	(0.185)	(0.199)	(0.083)	(0.083)
Control mean	4.696	5.176	4.461	5.129	3.870	4.196	5.431	4.208	4.198	2.515	2.194
Observations	820	820	819	819	816	818	819	818	815	819	815

Notes: Unit of observation is a response to an ad (each respondent sees two ads). Each column presents an estimate from equation (6) for a different outcome (see text for definitions). Sample includes only responses from alumni of the web development boot camp. Gender-neutral, Remote, and Diversity Statements are dummies indicating the ad was assigned to the respective status. The control mean is the outcome mean for ads under the non-gender-neutral language, non-remote treatment, and no diversity treatment status. The number of observations varies across columns due to missing data on outcomes (a few instances when respondents did not answer a survey question). Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$..

Table A.13: Treatment Effects (Alumni of UX Design Boot Camp Only) - Laboratoria

	(1) Job Appeal	(2) Good Employer	(3) Meet Require-ments	(4) Probability of Applying	(5) Suitability	(6) Probability of Being Chosen	(7) Good Salary	(8) Work Life Balance	(9) Inclusive Culture	(10) Women % Company	(11) Women % Position
Gender-neutral	0.414	0.247	0.113	0.494	0.819**	0.098	0.530*	0.290	1.283***	0.597***	0.515***
	(0.320)	(0.290)	(0.261)	(0.383)	(0.362)	(0.308)	(0.300)	(0.303)	(0.343)	(0.125)	(0.129)
Remote	0.693**	0.541*	-0.343	0.715*	-0.093	-0.123	0.412	0.481	0.327	-0.124	-0.191
	(0.320)	(0.290)	(0.260)	(0.383)	(0.362)	(0.308)	(0.300)	(0.303)	(0.342)	(0.125)	(0.128)
Diversity Statement	-0.147	0.215	0.448*	0.013	0.401	0.646**	-0.102	0.321	1.124***	0.297**	0.235*
	(0.320)	(0.290)	(0.263)	(0.382)	(0.361)	(0.309)	(0.300)	(0.302)	(0.342)	(0.125)	(0.129)
Control mean	5.121	5.061	7.909	5.242	5.788	6.758	5.182	4.515	4.485	3.182	3.515
Observations	270	270	270	270	270	270	270	270	270	270	270

Notes: Unit of observation is a response to an ad (each respondent sees two ads). Each column presents an estimate from equation (6) for a different outcome (see text for definitions). Sample includes only responses from alumni of the UX design boot camp. Gender-neutral, Remote, and Diversity Statements are dummies indicating the ad was assigned to the respective status. The control mean is the outcome mean for ads under the non-gender-neutral language, non-remote treatment, and no diversity treatment status. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$..

Table A.14: Treatment Effects (Full Sample, with Respondent FEs) - Laboratoria

	(1) Job Appeal	(2) Good Employer	(3) Meet Require-ments	(4) Probability of Applying	(5) Suitability	(6) Probability of Being Chosen	(7) Good Salary	(8) Work Life Balance	(9) Inclusive Culture	(10) Women % Company	(11) Women % Position
Gender-neutral	0.545***	0.559***	0.184**	0.502***	0.712***	0.374***	0.400***	0.498***	1.300***	0.655***	0.640***
	(0.115)	(0.108)	(0.091)	(0.140)	(0.123)	(0.107)	(0.114)	(0.122)	(0.144)	(0.056)	(0.055)
Remote	0.916***	0.357**	0.208	0.769***	0.405**	0.396**	0.284*	1.088***	0.257	0.187**	0.219***
	(0.172)	(0.158)	(0.126)	(0.197)	(0.165)	(0.154)	(0.157)	(0.175)	(0.202)	(0.077)	(0.077)
Diversity Statement	0.300	0.502***	0.012	0.276	0.248	0.071	0.194	0.273	1.224***	0.206**	0.165*
	(0.188)	(0.173)	(0.130)	(0.217)	(0.178)	(0.157)	(0.181)	(0.189)	(0.227)	(0.089)	(0.085)
Control mean	4.800	5.148	5.304	5.157	4.346	4.822	5.370	4.284	4.269	2.676	2.515
Observations	1,090	1,090	1,089	1,089	1,086	1,088	1,089	1,088	1,085	1,089	1,085

Notes: Unit of observation is a response to an ad (each respondent sees two ads). Each column presents an estimate from equation (6) for a different outcome (see text for definitions), with the addition of respondent fixed effects. Gender-neutral, Remote, and Diversity Statements are dummies indicating the ad was assigned to the respective status. The control mean is the outcome mean for ads under the non-gender-neutral language, non-remote treatment, and no diversity treatment status. The number of observations varies across columns due to missing data on outcomes (a few instances when respondents did not answer a survey question). Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$..

F Experimental materials - Get On Board Experiment

This appendix provides the experimental materials related to the Get On Board experiment. We provide both the original instructions in Spanish and an English translation (specific nouns used as examples cannot be translated, given that English does not have gendered grammar). Table A.15 provides key examples of how the gendered-language protocol works. Figure A.10 contains the exact instructions provided to Get On Board staff to implement the protocol (a one-page document in Spanish). Section F.1 translates this protocol to English, but maintains some key words in Spanish (since English has primarily non-gendered nouns making the exact translation impossible).

Table A.15: Treatment Protocol Examples - Get On Board

Non-inclusive	**Inclusive**
Rule 1:	
Los candidatos que pasen el primer filtro seran entrevistados	*Quienes* pasen el primer filtro seran entrevistados
Los candidatos que cumplan con los requisitos deberan enviar su CV	*Envíe* su CV si cumple con los requisitos
El area de I+D esta buscando *un Ingeniero Civil* para ocupar el cargo de *gerente*	El area de I+D esta buscando *Profesionales en Ingenieria* para ocupar la *gerencia*
Si eres *dinamico e innovador* para resolver problemas	Si eres *una persona dinamica e innovadora* para resolver problemas
Rule 2: for articles, nouns, quantifiers and adjectives	
En Novartis estamos buscando *programadores*	En Novartis estamos buscando *programadoras y programadores*
Rule 2: For isolated adjectives	
Requisitos: Titulado	*Requisitos: Titulada/o*

Notes: Examples in Spanish for each of our treatment protocol rules. Words in italics replaced in each case.

Figure A.10: Gender-Neutral Language (Treatment) Protocol Used by Get On Board

Checklist para Lenguaje Incluyente

1. La prioridad es neutralizar el género haciendo uso de estrategias de redacción tales como:

   - □ El uso de los pronombres relativos "quien" o "quienes".

   No Inclusivo: Los candidatos que pasen el primer filtro serán entrevistados.
   Inclusivo: Quienes pasen el primer filtro serán entrevistadas/os.

   - □ Modificar los verbos o usar la forma imperativa.

   No Inclusivo: Quien será el líder del área comercial.
   Inclusivo: Quien liderará el área comercial.

   No Inclusivo: Los candidatos que cumplan con los requisitos deberán enviar su hoja de vida al correo.
   Inclusivo: Envíe su hoja de vida si cumple con los requisitos.

   - □ El uso de sustantivos con doble marca de género (profesional, especialista, personal, Jefatura, Junta Directiva, gerencia, etc.).

   No Inclusivo: El área de I+D está buscando un Ingeniero Civil para ocupar el cargo de gerente.
   Inclusivo: El área de I+D está buscando Profesionales en Ingeniería para ocupar la gerencia.

   - □ El uso particular del sustantivo "persona."

   No Inclusivo: Si eres dinámico e innovador para resolver problemas.
   Inclusivo: Si eres una persona dinámica e innovadora para resolver problemas.

2. Posteriormente, se pretende visibilizar ambos géneros de la siguiente manera:

   - □ Para el uso de pronombres, artículos, cuantificadores, sustantivos y adjetivos que acompañen a estos últimos, se propone el uso del "desdoblamiento" en la redacción.

   No Inclusivo: El área de I+D está buscando un Ingeniero Civil para ocupar el cargo de gerente.
   Inclusivo: El área de I+D está buscando una Ingeniera o un Ingeniero Civil para ocupar el cargo de gerenta o gerente.

   - □ Para el uso de adjetivos aislados (sin un sustantivo acompañando) se propone el uso de barras oblicuas (/):

   No Inclusivo: Requisitos: Titulado
   Inclusivo: Requisitos: Titulada/o

3. Finalmente, para cambiar algunas prácticas que siempre colocan a los hombres en primer lugar de las enumeraciones, se propone ubicar a las mujeres al inicio de la redacción:

   - □ Alternancia de los géneros en las enumeraciones

   No Inclusivo: En Novartis estamos buscando programadores y programadoras.
   Inclusivo: En Novartis estamos buscando programadoras y programadores.

F.1 English Translation of Neutral Language (Treatment) Protocol Used by Get on Board

1. The priority is to neutralize the gender making use of writing strategies such as:

   ☐ The use of the relative pronouns *quien* or *quienes*.

   Non-Inclusive: *Los candidatos* who pass the first filter will be interviewed.

   Inclusive: *Quienes* who pass the first filter will be interviewed.

   ☐ Modify the verbs or use the imperative form.

   Non-Inclusive: Who will be *el líder* of the commercial area.

   Inclusive: Who will lead *el área comercial*.

   Non-Inclusive: *Los candidatos* who meet the requirements must send their resume by mail.

   Inclusive: Submit your resume if you meet the requirements.

   ☐ The use of nouns with a double gender mark (professional, specialist, personal, headquarters, board of directors, management, etc.).

   Non-Inclusive: The R&D area is looking for *un Ingeniero Civil* to fill the position of *gerente*.

   Inclusive: The R&D area is seeking *Profesionales en Ingeniería* to fill the management position.

   ☐ The use of the noun *persona*.

   Non-Inclusive: If you are *dinámico e innovador* to solve problems.

   Inclusive: If you are a *persona dinámica e innovadora* to solve problems.

2. Subsequently, it is intended to make both genders visible in the following way:

   ☐ For the use of pronouns, articles, quantifiers, nouns and adjectives that accompany the latter, the use of "unfolding" in the writing is proposed.

   Non-Inclusive: The R&D area is looking for *un Ingeniero* to fill the position of *gerente*.

   Inclusive: The R&D area is looking for *una Ingeniera o un Ingeniero* to fill the position of *gerenta o gerente*.

   ☐ For the use of isolated adjectives (without an accompanying noun) the use of oblique bars (/) is proposed:

   Non-Inclusive: Requirements: *Titulado*

Inclusive: Requirements: *Titulada/o*

3. Finally, to change some practices that always place men in the first place in the lists, it is proposed to place women at the beginning of the writing:

    - ☐ Alternation of genders in enumerations.

        Non-Inclusive: At Novartis we are looking for *programadores y programadoras.*

        Inclusive: At Novartis we are looking for *programadoras y programadores.*

G Experimental Materials - Laboratoria

This section provides the materials (invitation e-mail, survey instruments, ads shown to subjects) from the Laboratoria experiment. All materials are originally in Spanish, except those sent to alumni of the boot camps Laboratoria performed in Brazil, which were all in Portuguese. Only 43 of the 546 responses we obtained were from Brazilian alumni (partly reflecting that about 14% of Laboratoria's alumni are from the Brazilian boot camps).

G.1 Invitation e-mail - Laboratoria

English translation. The following is the translation of the e-mail sent to Laboratoria alumni inviting them to the survey. It also included a link to the survey website.

> Hello [*subject name*] Hope all is well with you. We're sending this email to invite you!
>
> Laboratoria had the opportunity to collaborate with researchers from INSEAD (France) and Princeton (USA) universities in a study that seeks to find out how job advertisements published on various job platforms in the technology sector are perceived. This survey is intended to help promote better quality of recommended ads, allowing more people to find the job they are looking for!
>
> Given that you are a key part of this industry, we would love it if you could help us with this research project by answering a short survey in which we show you job advertisements in your field and you give us your opinion about them.
>
> All guests who respond to the survey will enter a Kindle draw. We will draw two Kindles and if more than 700 alumni answer the survey, we will draw an additional Kindle for every 100 responses above 700 (for example, if 900 respondents answer, we will draw a total of 4 Kindles). In addition, all guests will have access to the results of the research project.
>
> Your participation in this survey is voluntary and your responses will be recorded in a secure system that can only be accessed by the research team. None of your personal data will appear in publications based on this research. If you have questions about this research, you can contact the principal investigators: lucia.delcarpio@insead.edu or fujiwara@princeton.edu, or contact the ethics review board directly: irb@princeton.edu
>
> Thank you very much for your attention! If you are interested in participating, click the button below to accept your participation and begin the survey.

Original version in Spanish. The original invitation in Spanish is below. A similar version in Portuguese was sent to the alumni of the Brazilian boot camp (but only mentioned that one single Kindle would be awarded, given the smaller number of Brazilian alumni).

> Hola [*subject name*] Esperamos que estés muy bien. Te enviamos este mail ya que ¡queremos extenderte una invitación!
>
> Como Laboratoria, tenemos la oportunidad de colaborar junto con investigadores de las universidades INSEAD (Francia) y Princeton (EEUU), en un estudio que busca conocer cómo se perciben los anuncios de ofertas laborales que se publican en diversas plataformas de trabajo en el sector tecnológico. Esta investigación tiene como objetivo ayudar a promover una mejor calidad en la selección de anuncios que se recomiendan, ¡permitiendo que más personas accedan al trabajo que buscan!
>
> Dado que eres parte fundamental de esta industria, nos encantaría que nos pudieras apoyar en esta investigación respondiendo una breve encuesta en la cual te compartiremos dos anuncios de trabajo en tu área laboral, para que nos des tu opinión sobre ellos.
>
> Entre todas aquellas egresadas que contesten la encuesta, estaremos sorteando dos Kindles y si más de 700 egresadas contestan la encuesta, sortearemos un Kindle adicional por cada 100 respuestas por encima de 700 (por ejemplo, si 900 contestan, sortearemos un total de 4 Kindles). Además de que todas podrán tener acceso a los resultados de la investigación.
>
> Tu participación respondiendo esta encuesta es voluntaria y tus respuestas se recogen con una aplicación segura a la que sólo podrá acceder el equipo de investigación. Ninguno de tus datos personales aparecerá en los informes posteriores de este estudio. Si tienes preguntas sobre la investigación, puedes ponerte en contacto con los investigadores principales: lucia.delcarpio@insead.edu o fujiwara@princeton.edu, o contactar directamente a la Junta de Revisión Institucional: irb@princeton.edu
>
> ¡Desde ya muchas gracias por tu atención! Si estás interesada en participar, marca el siguiente botón para aceptar tu participación y comenzar con la encuesta.

G.2 Survey Instrument - Laboratoria

English translation. The following is a translation of the survey used in the Laboratoria experiment. Originals were in Spanish or Portuguese. Text in *italics* provide further context and were not shown to participants.

> Hello! At Laboratoria, together with researchers from INSEAD (France) and Princeton (USA) universities, we are carrying out a study to find out how the

advertisements of job offers that are listed on various job platforms in the Tech sector are perceived. This will help us to promote a better quality of ads and better select those that we recommend. Now we are going to show you two ads in your field so that you can give us your opinion about them. Important: These ads do not represent current job openings. They are built based on a representative sample of ads listed in the past. We remind you that participation in this survey is voluntary. Your answers are collected with a secure application and will only be accessible by the research team. None of your personal data will appear in subsequent reports of this study. If you have questions about the research, you can contact the principal investigators: lucia.delcarpio@insead.edu or fujiwara@princeton.edu, or contact the Institutional Review Board directly: irb@princeton.edu

If you decide to participate in the survey, please check the button below to see the announcements.

Which Laboratoria boot camp you graduated from?

- Web Developer
- UX Designer

[*The answer to this question directed the respondent to see an ad in their field.*]

Graduation year?

[*Options were between 2015 and 2022.*]

Country of boot camp?

- Chile
- Colombia
- Peru
- Mexico
- Ecuador

[*Question only asked in the Spanish-version of survey. Alumni of the Brazilian boot camp received a separate invitation e-mail for a survey in Portuguese.*]

Currently:

Do you work in the tech sector?

- Yes
- No

Are you searching fora job in the tech sector?

- Yes
- No

Please read this advertisement and click the arrow when you are done:

[*Subjects were shown the first randomly selected ad. The questions below appeared after clicking the arrow. Questions 1-9 had sliders for a scale 0-10 on whether they fully disagreed (0) to entirely agreed (10) and questions 10-11 were multiple choice.*]

- I find this job attractive
- I think this company would be a good employer
- I have the required qualifications for this job
- I would apply for this job if I have the required qualifications
- I think this company is looking for someone like me
- If I applied, I would have a high probability of being chosen
- I think this company offers a good salary
- I think this company offers a good work/life balance
- I think this company has an inclusive/diverse culture

And about the composition of human talent in this company, would you think that:

- The proportion of women in the entire company is:
- The proportion of women in the type of position advertised is:

- Very low (0 to 10%)
- Low (11 to 20%)
- Relatively low (21 to 30%)
- Medium (31 to 40%)
- Relatively high (41 to 50%)

- Majority (more than 50%)

[*After answering the questions, another ad was provided and another round of similar questions asked. The survey ended after that, asking respondents to provide an e-mail solely for the purposes of the Kindle draw.*]

Original survey instrument in Spanish. The following is the original survey instrument in Spanish. The text in *italics* provides further context and were not shown to participants. A similar version in Portuguese was used for the alumni of the Brazilian boot camps.

¡Hola! En Laboratoria, junto con investigadores de las universidades INSEAD (Francia) y Princeton (EEUU), estamos haciendo un estudio para conocer cómo se perciben los anuncios de ofertas de trabajo que se listan en diversas plataformas de trabajo en el sector Tech. Esto nos ayudará a promover una mejor calidad de anuncios y seleccionar mejor aquellos que te recomendamos. Ahora te vamos a mostrar dos anuncios en tu campo para que nos des tu opinión sobre ellos. Importante: estos anuncios no representan ofertas laborales actuales. Están construidos en base a una muestra representativa de anuncios listados en el pasado. Te recordamos que la participación en esta encuesta es voluntaria. Tus respuestas se recogen con una aplicación segura y sólo serán accesibles por el equipo de investigación. Ninguno de tus datos personales aparecerá en los informes posteriores de este estudio. Si tienes preguntas sobre la investigación, puedes ponerte en contacto con los investigadores principales: lucia.delcarpio@insead.edu o fujiwara@princeton.edu, o contactar directamente a la Junta de Revisión Institucional: irb@princeton.edu

Si decides participar en la encuesta, por favor marca el botón siguiente para ver los anuncios.

boot camp que seguiste en Laboratoria:

- Web Developer
- UX Designer

Año de graduación

[*Options were between 2015 and 2022*]

País del boot camp:

- Chile
- Colombia
- Perú
- México
- Ecuador

[*Question only asked in the Spanish-version of survey. Alumni of the Brazilian boot camp received a separate invitation e-mail for a survey in Portuguese*]

Actualmente:
Trabajas en el sector Tech?

- Sí
- No

Estás buscando empleo en el sector Tech?

- Sí
- No

Lee por favor este anuncio y marca la flecha cuando hayas terminado:

[*Subjects were shown the first randomly selected ad. The questions below appeared after clicking the arrow. Questions 1-9 had sliders for a scale 0-10 on whether they fully disagreed (0) to entirely agreed (10) and questions 10-11 were multiple choice.*]

- Este empleo me parece atractivo
- Creo que esta compañía sería un buen empleador
- Tengo las calificaciones requeridas para este trabajo
- Postularía a este trabajo de tener las calificaciones requeridas
- Creo que esta empresa está buscando a alguien como yo
- De postular, creo que tendría altas probabilidades de ser elegida/o
- Creo que esta compañía ofrecería un buen salario
- Creo que esta compañía ofrecería un buen equilibrio trabajo/vida personal
- Creo que esta compañía tiene una cultura inclusiva/diversa

Y sobre la composición del talento humano en esta empresa, pensarías que:

- La proporción de mujeres en toda la empresa es:
- La proporción de mujeres en el tipo de puesto anunciado es:
- Muy baja (0 a 10%)
- Baja (11 a 20%)
- Relativamente baja (21 a 30%)
- Mediana (31 a 40%)
- Relativamente alta (41 a 50%)
- Mayoritaria (más de 50%)

[*After answering the questions, another ad was provided and another round of similar questions asked. The survey ended after that, asking respondents to provide an e-mail solely for the purposes of the Kindle draw.*]

G.3 Ads used in Laboratoria experiment

We prepared two separate sets of field-specific ads (UX Design and Web Development), the two boot camp fields that Laboratoria provides. In each set, two ads were prepared (since each respondent saw two separate ads, and we used different company names, descriptions, etc). Since each ad has eight variations (a 2×2×2 factorial design), we created 32 ads in Spanish and 32 (very similar) ads in Portuguese.

Since we believe presenting 64 different ads in this appendix is not productive, Figure A.11 provides an ad for a position in the web development field with non-gender-neutral language, no diversity statement, and for a non-remote position, and compares to the same ad version with gender-neutral language, a diversity statement, and for a remote position. The other six combinations of these three binary treatment conditions of the ad can be inferred from them. Figures A.12, A.13, and A.14 provide the text for the other position in the web development field and the two ads for a job in the UX design field. It shows the version under gender-neutral, with a diversity statement, and non-remote condition. (which is the most general, and other treatment conditions can be inferred from them). We present the Spanish version. Translation to Portuguese is straightforward given the similarity of the two languages.

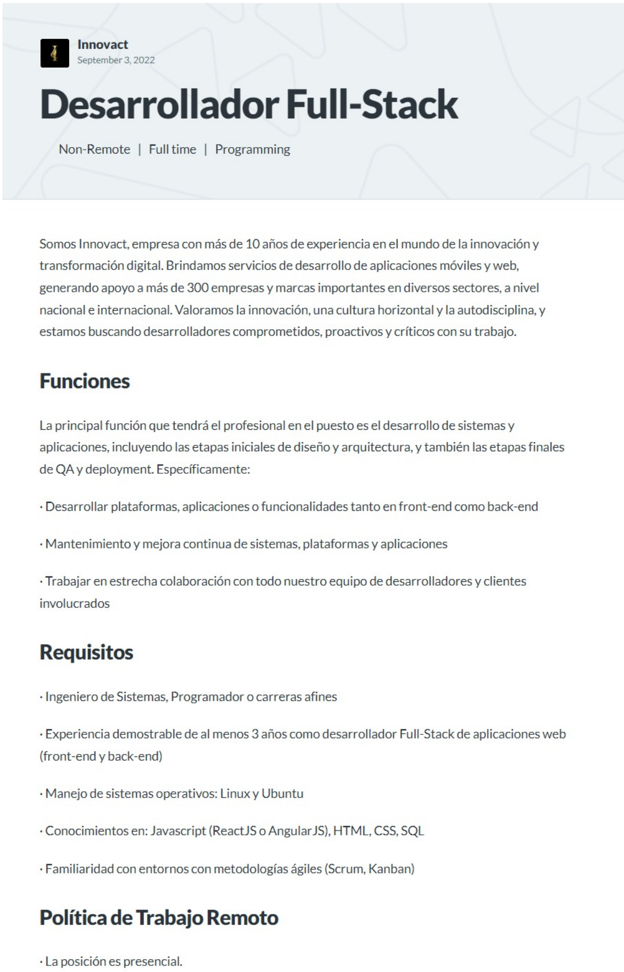
Differences between "treatments" and "controls." The differences created under each treatment status are:

1. If gender-neutral, the title is "*desarrollador/a Full Stack*" or "*diseñador/a UX UI*", while if non-gender-neutral ads would only show the masculine form "*desrollador*" and "*diseñador*." Another two gender-neutral (or masculine form) sentences also appear as the first bullet point under "*funciones*" (tasks) and under "*requisitos*" (requisites).

2. Under the diversity statement condition, one additional sentence is added to the end of the first paragraph ("*At 'name of company' we are committed to diversity and do not accept any type of discrimination*" or "*'Company name' is a forthcoming company and we do not accept any type of discrimination.*");

3. Under remote status, the word "remote" appears under the title and an explicit statement ("this position is remote" or "*Esta posición es remota*") appears at the bottom under "remote work policy" ("*Política de Trabajo Remoto*"). Under non-remote status, the word "non-remote" appears under the title and the remote work policy states "the position is in-person" ("*La posición es presencial*").

Figure A.11: Example of Ads in Laboratoria Experiment



Innovact
September 3, 2022

Desarrollador Full-Stack

Non-Remote | Full time | Programming

Somos Innovact, empresa con más de 10 años de experiencia en el mundo de la innovación y transformación digital. Brindamos servicios de desarrollo de aplicaciones móviles y web, generando apoyo a más de 300 empresas y marcas importantes en diversos sectores, a nivel nacional e internacional. Valoramos la innovación, una cultura horizontal y la autodisciplina, y estamos buscando desarrolladores comprometidos, proactivos y críticos con su trabajo.

Funciones

La principal función que tendrá el profesional en el puesto es el desarrollo de sistemas y aplicaciones, incluyendo las etapas iniciales de diseño y arquitectura, y también las etapas finales de QA y deployment. Específicamente:

· Desarrollar plataformas, aplicaciones o funcionalidades tanto en front-end como back-end

· Mantenimiento y mejora continua de sistemas, plataformas y aplicaciones

· Trabajar en estrecha colaboración con todo nuestro equipo de desarrolladores y clientes involucrados

Requisitos

· Ingeniero de Sistemas, Programador o carreras afines

· Experiencia demostrable de al menos 3 años como desarrollador Full-Stack de aplicaciones web (front-end y back-end)

· Manejo de sistemas operativos: Linux y Ubuntu

· Conocimientos en: Javascript (ReactJS o AngularJS), HTML, CSS, SQL

· Familiaridad con entornos con metodologías ágiles (Scrum, Kanban)

Política de Trabajo Remoto

· La posición es presencial.

(a) Non-gender-neutral, no diversity statement, non-remote

Innovact
September 3, 2022

Desarrollador/a Full-Stack

Remote | Full time | Programming

Somos Innovact, empresa con más de 10 años de experiencia en el mundo de la innovación y transformación digital. Brindamos servicios de desarrollo de aplicaciones móviles y web, generando apoyo a más de 300 empresas y marcas importantes en diversos sectores, a nivel nacional e internacional. Valoramos la innovación, una cultura horizontal y la autodisciplina, y estamos buscando desarrolladoras/es comprometidas/os, proactivas/os y críticas/os con su trabajo. Innovact es una empresa abierta y no aceptamos ningún tipo de discriminación.

Funciones

La principal función que tendrá la o el profesional en el puesto es el desarrollo de sistemas y aplicaciones, incluyendo las etapas iniciales de diseño y arquitectura, y también las etapas finales de QA y deployment. Específicamente:

· Desarrollar plataformas, aplicaciones o funcionalidades tanto en front-end como back-end

· Mantenimiento y mejora continua de sistemas, plataformas y aplicaciones

· Trabajar en estrecha colaboración con todo nuestro equipo de desarrolladoras/es y clientas/es involucradas/os

Requisitos

· Formación en Ingeniería de Sistemas, Programación o carreras afines

· Experiencia demostrable de al menos 3 años como desarrollador/a Full-Stack de aplicaciones web (front-end y back-end)

· Manejo de sistemas operativos: Linux y Ubuntu

· Conocimientos en: Javascript (ReactJS o AngularJS), HTML, CSS, SQL

· Familiaridad con entornos con metodologías ágiles (Scrum, Kanban)

Política de Trabajo Remoto

· Esta posición es remota.

(b) Gender-neutral, diversity statement, remote

Both ads are for a position in the web development field. The ad on the left is non-gender-neutral, while the ad on the right is gender-neutral (see title, first sentence under "*funciones*," and first bullet point under "*requisitos*."). The ad on the left is also for a non-remote position, while the ad on the right is for a remote position (see immediately below the title and the bottom "remote work policy."). The ad on the left does not have a diversity statement, while the one on the right does (see the last sentence in the first paragraph).

Figure A.12: Example of Ad in Laboratoria Experiment (Web Development)

NERV
September 3, 2022

Desarrollador/a Full-Stack

Non-remote | Full time | Programming

Somos NERV, empresa líder a nivel nacional e internacional en el desarrollo de tecnología para el sector eléctrico. Brindamos asesoría en la entrega de soluciones a organizaciones para que puedan gestionar su energía de forma activa e inteligente. Actualmente trabajamos con empresas de distintos tamaños y en rubros tales como: industrial, inmobiliario, logística, transporte, vinícola, salud y sector público. Buscamos desarrolladoras/es motivadas/os, críticas/os y comprometidas/os a brindar las mejores soluciones a nuestros/as clientes/as. En NERV estamos comprometidos con la diversidad y no aceptamos ningún tipo de discriminación.

Funciones

Buscamos una desarrolladora o un desarrollador full-stack para incorporarse al equipo (2 front-end, 3 back-end y 1 full-stack) y tomar la responsabilidad de desarrollar nuestras soluciones tecnológicas para el sector eléctrico. Específicamente:

· Liderar el equipo en el desarrollo de plataformas, aplicaciones o funcionalidades tanto en front-end como back-end
· Identificar, diseñar e implementar las mejores soluciones de software para los distintos problemas u oportunidades del negocio
· Servir de mentor/a a las desarrolladoras y los desarrolladores más junior
· Mantenimiento y mejora continua de sistemas, plataformas y aplicaciones

Requisitos

· Formación en Ingeniería de Sistemas, Programación o carreras afines
· Experiencia demostrable de al menos 5 años como desarrollador/a Full-Stack de aplicaciones web (front-end y back-end)
· Manejo de sistemas operativos: Linux y Ubuntu
· Manejo de control de versiones de código: GIT
· Conocimientos en: Javascript (ReactJS), HTML, CSS, SQL
· Experiencia en entornos con metodologías ágiles (Scrum, Kanban)

Política de Trabajo Remoto

· Esta posición es presencial.

This ad is under the gender-neutral, non-remote, with a diversity statement conditions.

Figure A.13: Example of Ad in Laboratoria Experiment (UX Design)

WheCode
September 3, 2022

Diseñador/a UX UI

Non-Remote | Full time | Design / UX

Somos WheCode, un equipo apasionado por lo que hacemos: productos digitales con enfoque centrado en las/los usuarias/os. Brindamos servicios de desarrollo de aplicaciones móviles y web, generando apoyo a más de 300 empresas y marcas importantes en diversos sectores, a nivel nacional e internacional. Valoramos la innovación, una cultura horizontal y la autodisciplina, y estamos buscando diseñadoras y diseñadores proactivas/os, con sensibilidad estética y críticas/os con su trabajo. WheCode es una empresa abierta y no aceptamos ningún tipo de discriminación.

Funciones

Buscamos diseñadoras y diseñadores UX/UI con conocimientos en investigación de usuarias/os, arquitectura de información y diseño de interfaces e interacción. Deberás:

· Investigar el negocio, mercado y perfil de las/los usuarias/os, para definir una estrategia de experiencia

· Diseñar la experiencia de uso del producto para que sea intuitiva y se presente con fluidez

· Diseñar soluciones para resolver problemas específicos de nuestras/os clientas/es a través de prototipos para testear con sus usuarias/os

· Realizar diagnósticos web: benchmark, análisis heurísticos

· Definir la arquitectura de información y flujos de interacción del usuario con el producto

Requisitos

· Formación en Diseño Gráfico, Industrial, Visual o afines.

· Experiencia relevante de al menos 3 años

· Portafolio web (Behance, Adobe, etc.) de trabajos anteriores

· Herramientas de diseño visual: Adobe Suite (Illustrator, Photoshop), Figma

· Experiencia en Diseño Centrado en Usuario, benchmark y usabilidad

· Herramientas de prototipado: Sketch, Invision, Axure

· Dominio del inglés oral y escrito

Política de Trabajo Remoto

· Esta posición es presencial.

This ad is under the gender-neutral, non-remote, with a diversity statement conditions.

Figure A.14: Example of Ad in Laboratoria Experiment (UX Design)

Tekadan
September 3, 2022

Diseñador/a UX UI

Non-Remote | Full time | Design / UX

Somos Tekadan, empresa líder en servicios de desarrollo de software, ecommerce, integración tecnológica y transformación digital. Acompañamos a más de 200 firmas en diversos sectores en todo el proceso de transformación digital, desde etapas iniciales hasta la implementación y optimización de las soluciones web. Tenemos un entorno innovador y una cultura horizontal, y estamos buscando ampliar nuestro equipo con diseñadoras y diseñadores creativas/os y con capacidad de trabajar en equipo, que compartan nuestra visión. En Tekadan estamos comprometidos con la diversidad y no aceptamos ningún tipo de discriminación.

Funciones

Buscamos diseñadoras y diseñadores UX/UI junior con sensibilidad estética y orientación a usuarias/os, capaces de resolver interfaces de modo atractivo y funcional. Deberás:

· Participar en la etapa de Research de cada proyecto asignado

· Realizar benchmarking para levantar hipótesis y pruebas de usabilidad

· Generar wireframes y prototipados con sus respectivos test de usuarias y usuarios

· Diseñar la identidad visual de productos y servicios digitales

Requisitos

· Formación en Diseño Gráfico, Industrial, Visual o afines

· Experiencia relevante y comprobable de al menos 1 año

· Herramientas de diseño visual: Adobe Suite (Illustrator, Photoshop), Figma

· Experiencia en Diseño Centrado en Usuario, benchmark y usabilidad

· Herramientas de prototipado: Sketch, Invision, Axure

· Dominio del inglés oral y escrito

Política de Trabajo Remoto

· Esta posición es presencial.

This ad is under the gender-neutral, non-remote, with a diversity statement conditions.