NBER WORKING PAPER SERIES

SPILLOVERS AND SCALABILITY IN JOB AD EXPERIMENTS: EVIDENCE FROM GENDER-NEUTRAL LANGUAGE

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 February 2026

Research funding from INSEAD’s Diversity and Inclusion Fund is gratefully acknowledged. A previous version of this paper was circulated under the title “Do Gender-Neutral Job Ads Promote Diversity?Experimental Evidence from Latin America’s Tech Sector.” 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, Seema Jayachandran, Adrian Rubli, and seminar participants at Boston U, GRADE, INSEAD, ITAM, LACEA Meeting, Stanford, Universidad del Pacifico, U Maryland, U Pitt, and the World Bank 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.

Spillovers and Scalability in Job Ad Experiments: Evidence from Gender-Neutral Language
Lucia Del Carpio and Thomas Fujiwara
NBER Working Paper No. 31314
June 2023, Revised February 2026
JEL No. J16, J7, M14, Z13

ABSTRACT

Gendered-grammar languages are spoken by 39% of the global population, and recent years have seen increasing advocacy for adopting gender-neutral language to promote diversity. We present evidence from two experiments on the effects of gender-neutral language in job advertisements and its treatment spillovers. In a field experiment encompassing all job postings on a Spanish-language tech platform, ads randomly assigned gender-neutral language attracted more female applicants—but only when few other ads applicants viewed were also treated. A second experiment shows that gender-neutral language shapes female tech workers' beliefs about job characteristics, particularly when the contrast with gendered language is salient. These findings are consistent with applicants interpreting gender-neutral language as a signal about job attributes, with effects that diminish as treatment becomes widespread. Short-run scalability is thus limited: small-scale interventions may produce meaningful impacts, but large-scale adoption may have negligible effects. However, longer-term effects may exist that our designs cannot capture.

Lucia Del Carpio
Instituto Tecnológico Autónomo de
México (ITAM)
lucia.delcarpio@itam.mx

Thomas Fujiwara
Princeton University
Department of Economics,
and School of Public and
International Affairs
and NBER
fujiwara@princeton.edu

Randomized controlled trials registry entries are available at
https://www.socialscienceregistry.org/trials/10076 and
https://www.socialscienceregistry.org/trials/5509

1 Introduction

Language can shape cognition and decision-making. Gendered grammar, in particular, may reinforce traditional gender roles (Whorf, 1956). In English, the generic "he" prompts male imagery (Moulton et al., 1978, Cole et al., 1983, Gastil, 1990), while women recall information better when instructions refer to women (Crawford and English, 1984). Gender-neutral language improves female performance on college entrance exams (Cohen et al., 2023), and bilinguals express more support for gender equality when surveyed in a gender-neutral language (Pérez and Tavitz, 2019b). Globally, speakers of gendered languages show lower female labor force participation and educational attainment (Jakiela and Ozier, 2022).

In recent years, advocacy and debate over inclusive language have intensified, yet evidence on its effectiveness remains scarce. We conducted two randomized experiments to study the effects of gender-neutral language in job ads in Latin America's tech sector, where women make up only 7% of the tech workforce (Del Carpio and Guadalupe, 2021). The region has seen both informal adoption of gender-neutral language and government interventions supporting or opposing its use (see Appendix A).

The 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. The second experiment investigates mechanisms by studying how gender-neutral language in ads affects female tech workers' beliefs about job characteristics. A growing literature explores how the content and language of recruitment materials affect applicant pool composition. To our knowledge, ours is the first study to evaluate gender-neutral language and the first to experimentally examine treatment spillovers for any type of content.[1]

In Spanish, as in many gendered-grammar languages spoken by 39% of the global population (Jakiela and Ozier, 2022), nouns carry masculine or feminine gender. The traditional default is to use the masculine form as a "generic" when gender is unspecified. For example, there are distinct words for male engineer ("ingeniero") and female engineer ("ingeniera"), but no gender-neutral term, so job ads traditionally use only "ingeniero."[2]

Our first experiment was conducted in partnership with Get on Board, a widely used job platform for the tech sector in Latin America. From April to November 2020, *all* ads submitted to the platform (over 2,000) were randomly assigned to treatment (gender-neutral language) or control (business as usual). Treated ads were edited to use gender-neutral

[1] For job ad content 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) and Samek (2019). None study treatment spillovers, general equilibrium effects, or gender-neutral language. See Appendix B for further discussion.

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

language following a protocol based on government guidelines. For example, "ingeniero" was revised to "ingeniera/o." Potential applicants were unaware of the experiment; they simply observed that some ads used gender-neutral language while others did not.

A key contribution of our study concerns 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). Applicants may interpret gender-neutral language as a signal about firm characteristics and job amenities, particularly in a setting like ours where gender-neutral ads are not entirely uncommon. For example, they might infer that firms using gender-neutral language are more likely to offer flexible work hours or employ more women. This updating mechanism suggests that a policy mandating such language in all ads would have no effect, as applicants would have nothing new to infer from such ads.

Alternatively, the mechanisms behind gender-neutral language may be more psychological ("behavioral") in nature, such as evoking female referents or improving recall. The implications for spillovers are less clear, and a policy mandating gender-neutral language in all ads could still have substantial effects.

To study spillovers, we define for each ad a set of *neighbor ads* that applicants likely viewed alongside it. Given the platform's user experience, ad i's neighbors are ads with similar job titles posted three days before or after ad i. Because treatment was assigned randomly and independently, the share of neighbor ads treated is also random. This allows us to estimate how treatment effects vary with this share, leveraging random variation both in the treatment itself and the dimension of heterogeneity. Using causal forests (Athey et al., 2019), we confirm that the share of neighbor ads treated is the key predictor of treatment effect heterogeneity.

Results indicate that treatment increases the share of women who apply, but only for ads with a low share of treated neighbors—those likely perceived by applicants as relatively rare gender-neutral ads. The effect is driven primarily by more women applying. We also find suggestive evidence that effects are uniform across the quality distribution (both more and less qualified women apply), that treatment increases the share of women advancing to later stages of the hiring process and potentially being hired, and that effects reflect net new female applications rather than diversion from control ads.[3]

Consistent with spillovers, effects for ads with larger shares of treated neighbors are smaller and statistically indistinguishable from zero. Our results thus suggest that the effects of gender-neutral language have limited scalability: policies that increase its prevalence to

[3]Treatment increases the share of female applicants by 3.7 p.p. for ads in the bottom quartile of the share of neighbors treated distribution (average of 20% of neighbors treated), compared to a control mean of 14.6%. Sections 2.1 and 3 discuss caveats regarding our measure of candidate quality and the fact that information on later-stage outcomes is observed only for a selected sample of firms.

an extent it becomes "common" are unlikely to have substantial effects, at least in the short run. Indeed, we find a zero treatment effect for the entire sample, where half of the ads are treated.[4]

To investigate mechanisms, we conducted a second experiment in partnership with Laboratoria, a leading NGO that trains Latin American women for tech sector jobs. Laboratoria alumnae resemble typical Get on Board users, though they are exclusively female and apply for junior to semi-senior positions. In an online survey, each respondent viewed two ads randomly assigned to use gender-neutral or generic masculine language and was asked about her propensity to apply and her beliefs about the position. Ads were fictional, but respondents were told the survey would help calibrate future job recommendations, incentivizing truthful answers (Kessler et al., 2019). Respondents were not informed that the survey involved evaluating gender-neutral language.[5]

Respondents reported being more likely to apply to gender-neutral ads and believed they were more suitable for the position and more likely to be hired. They also perceived firms using gender-neutral language as more likely to have an inclusive culture, promote work-life balance, and employ more women. In a cross-randomized design, we varied whether ads offered remote work and whether they included a diversity statement. The effects of gender-neutral language were substantially larger than those of diversity statements and comparable to or larger than those of remote work.

Our randomization protocol makes gender-neutral language more salient in the second ad: respondents see either a gender-neutral ad followed by a non-neutral ad, or vice versa. Confirming the importance of contrast, effects are substantially larger when respondents previously saw an ad with different language than when they evaluated a first ad without a comparison.

Results from both experiments are consistent with female applicants interpreting gender-neutral language as a signal about firm characteristics and job amenities. Corroborating this, we find no effects on a question about self-assessment (whether respondents met the job's requirements, which are clearly stated in the ads). The spillover results in our first experiment also support this mechanism: when gender-neutral ads are more common, applicants update less. Thus, an updating mechanism can explain the entirety of our results. The same is not as clear for psychological mechanisms such as evoking female referents or improving recall, which do not yield clear predictions for spillovers and do not immediately explain the second experiment's results.

[4]The effect for the entire sample is -0.001 (SE = 0.0069). Effects discussed here are intent-to-treat; Section 3 discusses treatment-on-treated effects.

[5]Surveying Get on Board users directly was not feasible. Section 4 discusses experimenter demand effects and social desirability bias.

Related literature. This paper contributes to three strands of literature. The first, primarily experimental, examines how job ad content influences applicant pool diversity (see Footnote 1). Closest to our work are studies on how subtle language choices, rather than explicit statements, affect female representation—for example, Abraham et al. (2024) on optional qualifications and superfluous language and Coffman et al. (2024) on ambiguity in required qualifications. We contribute in two ways: providing the first evaluation of gender-neutral language in recruitment materials and the first experimental analysis of treatment spillovers for any ad content. The latter is essential for understanding scalability (List, 2022).

This literature is limited in its capacity to study spillovers as it involves a single employer randomizing at the applicant level. Our findings suggest that while such interventions can yield effects when few employers adopt them, impacts may diminish, or even disappear, when the treatment is scaled across many employers. See Appendix B for further discussion and a brief description of papers.

A second strand uses difference-in-differences strategies and observational data to study job ads containing *explicit statements* about employer preferences over applicant gender. Closest to our work is Kuhn and Shen (2023), which examines a ban on such *explicit gender requests* on a Chinese job board. Their study explores a different type of spillover: the ban's impact on ads that did not include gender requests. Card et al. (2024) analyzes a 2005 Austrian reform that effectively ended the practice. While it does not examine spillovers, it studies hiring outcomes directly.

Both studies find that removing explicit requests for male applicants increases the share of women applying (China) or hired (Austria). At first glance, this contrasts with our null effect in the overall sample. However, the updating mechanism reconciles these findings. In China and Austria, only 12.5% and 20% of pre-reform ads explicitly requested men, making such language a distinctive signal. In our setting, gender-neutral language is more common even without treatment, and during the experiment, half of the ads were treated, making it less distinctive and thus less informative. Consistent with this, we find effects only when gender-neutral language stands out, that is, when few neighboring ads are treated.[6]

Of course, other important differences remain: we study a more subtle change in ad content (language rather than *explicit* statements of employer preferences), on a different continent, and in the tech sector (where acquiring skills in the short run is difficult). We further compare our findings to those of Kuhn and Shen (2023) and Card et al. (2024) in Section 3 and Appendix B.[7]

Third, a literature dating to Whorf (1956) studies how language affects cognition and behavior (see the first paragraph of this introduction and Appendix A). We contribute by

[6]See Section 3.2 and Table 1 for details on gender-neutral language prevalence.

[7]See also Kuhn and Shen (2013), Kuhn et al. (2020), Helleseter et al. (2020), and Arceo-Gomez et al. (2022) on ads with explicit gender (and age) requests.

providing a first study of gender-neutral language in recruitment materials and shedding light on the underlying mechanisms.

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 a job board focused on the tech sector in Latin America. By early 2026, over a million professionals had submitted more than three million applications to more than 14,000 registered companies via the website.

To post job ads, companies pay a submission fee or subscribe to a service allowing multiple postings. All ads are submitted for moderation, where Get on Board staff ensure they comply with quality standards. Ads are presented in a standardized format: ad title (the position advertised), followed by a company description and sections on job description and candidate requirements, desirable candidate attributes, and 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 an offer. Not all companies use this tool (see Section 3.2). We describe the user experience for job applicants later in this section, as it plays a key role in our analysis of treatment spillovers. While companies can post ads elsewhere, such as on LinkedIn or their own hiring pages, those advertising on Get on Board rarely use other job boards, likely reflecting Get on Board's dominance in the markets where it operates.

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.[8]

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*

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

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 generic masculine forms by replacing them with (gender-neutral) relative pronouns, imperative verbs, and nouns with no gender assigned.[9] 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.").

Figure 1 displays the same ad under control and treatment status, with differing words highlighted. Table A.22 provides key examples of the protocol, and Appendix F contains the exact guidelines used by Get on Board staff.

Data. We collected data on the text of the ads, the positions (e.g., job title, seniority level, whether it is remote or in-person), and applicants, whose gender (male or female) is coded based on first names.[10] We also observe a measure of applicant quality. 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. Lower scores imply "better" applicants from employers' revealed preferences.[11]

[9]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.

[10]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.

[11]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.

Applicants' user experience. Most applicants find job ads by entering job titles into a prominently displayed search bar. The platform uses a semantic search algorithm: querying for a particular title (e.g., "desarrollador full stack") returns ads with similar, though not necessarily identical, titles. Figure 2 provides an example. The algorithm handles both English and Spanish, so a search for "web developer" also returns jobs titled "desarrollador web." The platform does not use user-specific information (e.g., past searches or location), meaning any user entering the same query at the same time sees the same set of ads. Users can also browse ads through a predetermined set of 12 fields. Appendix C describes these fields, while Section 3.3 discusses the robustness and interpretation of our results to different assumptions about how users explore the platform.

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 on the same day or 3 days before or after 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 roles and tasks. Moreover, they reflect the ads that applicants would see listed 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. Table A.1 lists the 16 groups and Appendix C describes the classification procedure.[12]

For example, suppose ad i and belonging to the "designer" job title group was posted on May 25. Suppose further that 7 other ads belonging to the "designer" job title group were posted in the May 22-28 period, and 3 out of these 7 ads were randomly assigned to treatment. Then ad i's SNT_i is $3/7 = 43\%$.

Thus, the *share of neighbor ads treated* captures the intensity of treatment spillovers. It combines timing and ad titles to proxy the share of treated ads likely seen alongside ad i by Get on Board users. While we do not observe applicants' search queries or the platform's algorithm directly, our interactions with the platform indicate that our proxy closely approximates actual search results. Search results list ads based on their titles in

[12] 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.

roughly chronological order and are not customized based on the applicant's information.[13]

Since treatment is assigned to each ad *independently*, ad i's share of neighbor ads assigned to treatment is a random variable (deriving from a binomial distribution) that is *uncorrelated with 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.[14]

Note that i) ads are assigned to job title groups solely by the text of their title, and ii) applicants' search behavior or the platform algorithm do not enter the computation of SNT_i. This variable is entirely determined by ads' titles and submission dates.

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.[15] Thus our main sample includes 2,201 ads from 792 unique companies. The share of treated units was 48.7%.[16] These ads received a total of 104,680 applications. The average ad received 9.2 applications from women and 38.2 applications from men. The distribution of the number of applications is right-skewed, with a few ads receiving several hundred 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.2 presents the average characteristics of the control and treatment ads. 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. Table A.3 shows that randomization also generated balance in job group title composition. An omnibus test of joint orthogonality following Kerwin et al. (2024) does not reject the null of balance across all available covariates (p-value = 0.338, see Appendix D). Roughly 40% of the positions are remote, given that 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

[13] We use a window of 3 days before and after as our baseline since it approximates the size of ads listed on the page (based on our experience testing different searches) and it averages out day of the week effects (every window includes one Monday, one Saturday, and so on). Section 3.3 discusses robustness to different windows and an alternative measure of neighbor ads defined by the 12 fields used for browsing ads.

[14] Formally, ad i with n_i neighbor ads has a number 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.

[15] Of the 334 ads removed, 231 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 any other ads in the same job title group posted 3 days before or after or ads that could not be assigned to a job title group).

[16] 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%.

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.

As discussed above, ad i's share of neighbors treated is a random variable orthogonal to ad i's characteristics, its own treatment status, and its job title group. Corroborating this, Table A.2 shows that the average share of neighbors treated is similar in the control and treated ads. Moreover, Table A.4 shows that the share of neighbors treated is uncorrelated with ad characteristics, while Table A.5 shows it is also uncorrelated with job title groups.

Construal and subject perceptions. Get on Board users were not informed that an experiment was taking place or that some ads were selected by the platform to adopt gender-neutral language, making this a "natural field experiment" (Harrison and List, 2004). From applicants' perspective, some ads were gender-neutral and some were not, with the most plausible interpretation being that companies themselves chose their language. Use of gender-neutral language on the platform before the experiment (and in control ads) was not entirely uncommon, making it natural for applicants to draw inferences about employers from their language choices (see Section 3.2 for details on gendered language use in the control group).

Variation in the share of treated neighbor ads can shape how applicants perceive gender-neutral language. Suppose ad i uses gender-neutral language while few of its neighbors do. Applicants may perceive the employer posting i as having made an uncommon choice and infer it differs from other firms. Conversely, if most of ad i's neighbors also use gender-neutral language, there is less reason to view the employer as distinctive.[17]

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 study mechanisms by examining how perceptions of job attributes and company characteristics are influenced by the use of gender-neutral language in ads.[18]

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

[17] This logic assumes many applicants lack strong priors and update beliefs about the prevalence of gender-neutral language based on their latest search. Alternatively, narrow bracketing (Read et al., 2000) could lead applicants to make decisions in isolation. Our analysis can be interpreted as a test of weak priors and/or narrow bracketing: SNT_i should not affect applications to ad i in their absence. Lastly, an applicant who sees *all* ads as gender-neutral might infer this reflects Get on Board policy, but this is unlikely—only 2.7% of treated ads have $SNT_i = 1$ (Section 3.2).

[18] We selected Laboratoria because its alumnae resembles the typical users of the platform—though they are exclusively women and apply for junior to semi-senior positions—after Get on Board confirmed that surveying its users was not feasible.

of graduates secure jobs in the tech sector upon graduation. In 2022, Laboratoria had an alumnae 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 alumnae of the Brazilian boot camp, which was in Portuguese (also a gendered-grammar language). Appendix G provides all the experimental materials.[19]

Scope and invitations. Laboratoria distributes an email newsletter to its alumnae featuring curated job listings from various online platforms, with the majority sourced from Get on Board. 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 alumnae are exclusively female, our sample consists only of women.

Experimental Design. Each respondent was shown two fictitious job ads in their bootcamp field. To avoid deception, respondents were informed that the ads were fictitious. Moreover, respondents had an incentive to answer truthfully since their responses would influence the job recommendations they receive from Laboratoria, similar in spirit to the incentivized resume-rating design of Kessler et al. (2019). To ensure realism, ads 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-

[19]The experiment was pre-registered with the AEA's RCT registry in September 2022 (AEARCTR-0010076).

tion had an additional sentence at the end of the first paragraph.[20] 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.[21] This factorial (2×2×2) 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"*)

[20] 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.*"

[21] 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.[22]

As discussed above, participants had an incentive to respond truthfully since their answers would influence future job recommendations. Moreover, stated or hypothetical choices have been shown to correlate with actual job application behavior (Maestas et al., 2023, Wiswall and Zafar, 2018).

Salience of comparisons. 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.[23]

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.17 presents the summary statistics and covariate balance.[24]

3 Get on Board Experiment Results

Section 3.1 presents our empirical strategy. Section 3.2 reports the effect of treatment assignment on the use of gender-neutral language (first-stage results) and the main results on job applications. We conclude with findings on treatment effect heterogeneity, placebo tests, and robustness checks (Section 3.3).

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

[23] This comparison is not intended to replicate the spillover environment of Get on Board; rather, it provides a complementary test of the signaling mechanism.

[24] Approximately 25% of respondents are alumnae from the UX design boot camp and the remainder from web development. Alumnae from the Chilean, Peruvian, and Mexican boot camps account for 25% of responses each. Brazilian alumnae, who received the Portuguese version of the survey, account for 8%.

3.1 Empirical Strategy: Effects and Spillovers

We examine treatment effect heterogeneity by the share of neighbor ads treated (SNT_i; see Section 2.1). This choice reflects our focus on spillovers, scalability, and underlying mechanisms. We later use causal forests (Athey et al., 2019) to confirm that SNT_i is a key source of treatment effect heterogeneity. 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 indicates assignment to treatment, and X_i is a vector of controls. $MidQuartiles_i^{SNT}$ is a dummy equal to one if i's share of treated neighbor ads (SNT_i) falls in the middle two quartiles of its distribution, while $TopQuartile_i^{SNT}$ indicates whether SNT_i is in the top quartile. The parameter β_0 captures the intent-to-treat (ITT) effect for ads in the bottom quartile of SNT_i. The effect for ads in the middle quartiles is $\beta_0 + \beta_M$, and for ads in the top quartile it is $\beta_0 + \beta_T$. The average treatment effect across all ads is $\beta_0 + 0.5\beta_M + 0.25\beta_T$. The parameters γ_M and γ_T capture the effect of SNT_i on *control* ads. Note the distinction between spillovers as a source of treatment effect heterogeneity (captured by the βs) and "direct" spillovers on control ads (captured by the γs).[25]

The median value of SNT_i is 0.5, with first and third quartiles of 0.34 and 0.63, respectively. Thus, $MidQuartiles_i^{SNT} = \mathbb{1}(0.34 < SNT_i \leq 0.63)$ and $TopQuartile_i^{SNT} = \mathbb{1}(SNT_i > 0.63)$. Panel (a) of Figure 4 shows that the average SNT_i in the bottom and top quartiles is close to 20% and 80%, respectively, while it is close to 50% for the middle quartiles. The figure also shows that using longer time windows to define neighbor ads would compress these differences, as the variation in SNT_i stems from the relatively small number of neighbors within the three-day window.[26]

SNT_i is randomly determined and uncorrelated with ad characteristics. It is also orthogonal to T_i since treatment was *independently* assigned to each ad. See Section 2.1 for further discussion and Tables A.2–A.5 for corroborating evidence. We thus estimate treatment effect heterogeneity identified from random variation in both treatment assignment and the dimension of heterogeneity itself. Intuitively, one need not worry whether the heterogeneous effects are driven by SNT_i or some correlated unobservable, because SNT_i is random and therefore uncorrelated with any other variable in expectation.

[25]Externalities in other contexts, such as treating contagious diseases, may occur primarily as direct spillovers (e.g., Miguel and Kremer, 2004).

[26]As the time window increases, the number of neighbors n_i grows. Since the number of treated neighbors follows a binomial distribution $B(n_i, 0.5)$, SNT_i converges to 0.5 as n_i increases. Using our baseline window of three days before and after, n_i ranges from 1 to 24, and 53% of ads have between 4 and 10 neighbors.

That said, $MidQuartiles_i^{SNT}$ and $TopQuartile_i^{SNT}$ are correlated with ad i's *number* of neighbors, which can raise identification concerns for the direct spillover effects (γ_M and γ_T), as discussed in Borusyak and Hull (2023). To address this, we implement that paper's "recentered treatment" procedure: all estimates of equation (1) include in the vector X_i two variables, Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$]. This correction ensures that effects are identified from random variation in treatment assignment and eliminates the "non-random shock exposure" stemming from $MidQuartiles_i^{SNT}$ and $TopQuartile_i^{SNT}$ being correlated with the number of neighbors (Borusyak and Hull, 2023). It also implies that controlling for the number of neighbors is unnecessary and unlikely to affect results, which we verify when discussing Tables 2 and A.14. See Appendix D for further discussion.[27]

We augment the vector of controls X_i in two ways. The baseline specification includes month dummies interacted with a dummy for remote positions, given that the experiment took place during the early months of the Covid-19 pandemic. Alternatively, we use the post-double-selection (PDS) LASSO procedure from Belloni et al. (2014) to select controls from the following set: month dummies, a dummy for whether the ad posted a salary range, dummies for required seniority, and day-of-the-week dummies. All these variables interact with a dummy for remote positions. We also include the *number* of neighbor ads in the candidate set.

Inference. We report heteroskedasticity-robust standard errors and also provide p-values obtained via randomization inference, following a procedure from Borusyak and Hull (2023). This approach accounts for dependencies induced by spillovers, thereby avoiding the need to rely on clustered standard errors to address such dependencies. For each draw of the full assignment vector, we recalculate not only T_i but also SNT_i and all variables derived from 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 report two-sided p-values: the proportion of placebo coefficients that exceed the observed coefficient in absolute value.

Share of Neighbor Ads Treated as the Key Predictor of Treatment Effect Heterogeneity. We confirm the importance of SNT_i for treatment effect heterogeneity using causal forests (Athey et al., 2019). Applied to our data, causal forests identify SNT_i as having 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. A common concern in interpreting it as

[27] As discussed in Section 2.1, ad i's number of treated neighbors follows a binomial distribution with n_i draws (number of neighbors) and probability of success 0.5 (treatment). We can thus directly calculate Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$]. For example, Prob[$TopQuartile_i^{SNT} = 1$] = Prob[$SNT_i > 0.63$] is 14.5% when $n_i = 8$ (6 or more treated neighbors) and 22.7% when $n_i = 7$ (5 or more treated neighbors).

a driver of treatment effect heterogeneity is that correlated covariates may substitute for one another: trees may split on one but not the other, even if both matter in the data-generating process. This is not a concern for SNT_i, however, since it is random and uncorrelated with other covariates. Appendix D provides additional discussion and results.

3.2 Main Results

"First stage" results: effects on use of gender-neutral language. We begin by documenting the prevalence of gendered language and how treatment affects it. We counted the number of gendered words in each section of treated and control ads, independently of the platform's implementation of the treatment. In executing the protocol, the platform staff rarely edited four gendered words (and their plurals): "usuario" (user, often in the phrase "experiencia del usuario," a loan translation of user experience), "nosotros" and "nuestro" (we/our, referring to the company), and "cliente" (client). We exclude these from our counts (i.e., treat them as non-gendered) to better focus on how treatment affects ad content.[28]

Column (1) of Table 1 presents control means, characterizing the prevalence of gendered language absent treatment. Over a third of control ad titles have gendered words, and 18.7% contain zero gendered words throughout. The average number of gendered words in the entire text is 2.85 (this average would rise to 4.4 if the four excluded words were included).

Columns (2)-(4) present results from estimating equation (1). Panel A indicates that treatment substantially reduces the number of gendered words: the estimated β_0 implies a reduction of 1.18 gendered words in the entire ad text. Although this coefficient is identified from ads in the bottom quartile of SNT_i, the interaction terms are small and statistically insignificant, indicating similar effects across all levels of treatment intensity.

The effect is concentrated in the most salient parts of the ad: in treated ads, average gendered word counts in the title and job description and requirements section are close to zero, with more than 80% containing no gendered words in either section. Figures A.2 and A.3 present the corresponding histograms, comparing the distribution of word counts across treated and control ads.

Panel A also indicates that implementation was incomplete: some gendered words remain in treated ads, particularly in sections describing the company and benefits (e.g., "gerente" when referring to the firm's manager). Editing was most thorough in sections referring directly to the position and candidate. Consistent with this pattern, Table A.6 shows that the most frequently edited words are occupational and role-related terms such as those for

[28]While the gendered status of a word is usually clear-cut, native Spanish speakers may disagree on borderline cases (e.g., whether "clienta" should be used for a female client or whether "cliente" is already gender-neutral). We use the Royal Spanish Academy's dictionary to determine whether a word is gendered.

developers, engineers, designers, and programmers.[29]

Panel B of Table 1 confirms these patterns with binary indicators. The share of ads with zero gendered words in the title and job description and requirements section increases from 43.5% in the control group to more than 85% under treatment, and the share with no gendered words throughout doubles with treatment. The table also reports patterns in language choice. While 52% of control ads have an English title, only 11% are written entirely in English, reflecting the common practice of combining English titles with Spanish text. There are no significant differences between treatment and control groups in English usage, consistent with the protocol specifying that English titles should not be edited.

Finally, throughout the table, the interaction effects (β_M and β_T) are statistically insignificant and substantially smaller than β_0. This is expected: there is no reason for treatment to affect ad text differently based on how many neighbors are treated.[30]

Main results: ITT effects on the share of female applicants. Columns (1) and (2) of the top panel of Table 2 report the results from the estimation of equation (1) for our main outcome: the share of female applicants. 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 (SNT_i). It includes the randomization-inference p-values that account for dependencies created by spillovers. Odd columns report results using baseline controls; even columns show results using PDS-LASSO controls.

The effect for the entire sample ($\beta_0 + 0.5\beta_M + 0.25\beta_H$) based on column (1) is -0.001 (SE=0.0069). Thus, a 95% confidence interval does not include effects with a magnitude of 1.2 p.p. or larger for the whole sample. However, this average effect masks substantial heterogeneity by share of neighbors treated (SNT_i).

Columns (1) and (2) show positive 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 (1) is 3.7 p.p. increase relative to the control mean of 14.6%. We further discuss effect magnitudes when presenting treatment-on-treated effects below. 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 that the differences between the effect for the bottom quartile and other quartiles (β_M and β_T) are themselves statistically significant.

Taking point estimates at face value, leaving standard errors and p-values aside, Table 2

[29] For each gendered word w appearing in the ads' texts, we compute N_w^{control} and N_w^{treat}: the total number of times w appears across all ads assigned to control and treatment, respectively. Table A.6 reports $N_w^{\text{control}} - N_w^{\text{treat}}$ for the twenty words with the highest values.

[30] To economize on space, Table 1 omits the direct spillovers (γ_M and γ_T). Like the interactions, they are close to zero and insignificant. It also omits randomization inference p-values, as the conclusions from it and the reported standard errors are similar.

suggests that treatment has a positive effect for ads in the bottom quartile of SNT_i but a smaller negative effect for the medium and top quartiles. This averages to a zero effect. A more nuanced interpretation given statistical uncertainty is that there is a detectable positive effect for the bottom quartile, which dissipates as more neighbor ads are treated, becoming small and statistically undetectable. For the entire sample, the point estimate is zero, and large effects can be ruled out.

Figure A.4 presents the cumulative distribution function (CDF) of the share of female applicants for treated and control ads across quartiles of the SNT_i distribution. For ads in the bottom quartile, the effect is driven by a broad rightward shift of the treated CDF relative to the control CDF (see Appendix D for further discussion). The direct spillovers (γs) are smaller in magnitude, and we cannot reject that they equal zero; we return to their interpretation when discussing spillovers and mechanisms.

ITT Effects on the Number and "Quality" of Applicants. 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 applicants has a larger variance than the female share), the point estimates indicate a percent increase in female applications that is at least two 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.[31]

Columns (7) and (8) report effects on the average quality of applicants (as measured by badness scores) that are close to zero, regardless of the SNT_i quartile. 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.5 and A.6 show the distribution of applicants' badness scores by gender. Male and 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, the results suggest that treatment increases the share of women applying without affecting the quality distribution of applicants, indicating that the larger share of female

[31]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 extensive margin (effect on a dummy if at least one woman applied) is 0.042 (SE=0.036). The intensive margin is 0.093 (SE=0.131), 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 of column (4). Only 6 ads have zero male applicants. Thus effects are essentially the same when using logs. We use inverse hyperbolic sines and/or logs since the distribution of the number of applicants is right-skewed (Section 2.1).

applicants comes from across the quality distribution. 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. See Appendix D for further discussion.

Interpreting results: spillovers and scalability. The results in Table 2 are consistent with substantial spillovers and limited scalability. A smaller-scale experiment treating only a small fraction of ads seen by applicants would likely find significant effects on the share of female applicants. However, treating a larger share of ads yields effects close to zero. In the entire sample, where 50% of ads are treated, we find a point estimate of zero. As discussed in the introduction and Appendix B, previous experiments on job ad content are precisely such smaller-scale experiments: they typically involve a single employer and randomize at the applicant level, limiting their ability to detect spillovers.

For ads with SNT_i in the bottom quartile, where we find an effect, a natural question arises: does treatment generate net new applications by women, or does it divert applications from control ads to treated ones? Our experimental design is not ideally suited to answer this question, given randomization at the ad level, but the evidence favors net new applications over diversion.[32]

If diversion were the main mechanism, one would expect that increasing SNT_i for a control ad would *lower* its share and number of female applicants. However, we find no evidence of this: the direct spillovers (γ_M and γ_T) are small, statistically insignificant, and, if anything, *positive*. The most straightforward interpretation is that treatment generates net new applications by women when SNT_i is low.[33]

Treatment-on-treated effects. Given the intent-to-treat nature of the previous results, it is natural to consider treatment-on-treated effects. However, as the discussion of Table 1 suggests, there is no single obvious measure of treatment intensity. Our goal here is not to identify the "correct" treatment-on-treated specification, but to illustrate some possibilities. Without additional assumptions, we cannot determine whether gender-neutral language in different sections of the ad (e.g., job description versus benefits) matters differentially, or whether effects are nonlinear (e.g., moving from one gendered word to zero may matter more

[32] An experiment with individual-level randomization—where respondents see either few or many treated ads and total applications are tracked—would better address this question, but would be logistically difficult to implement as a natural field experiment (Harrison and List, 2004) and would not speak to the policy-relevant question of how treatment affects an ad's applicant pool from the firm's perspective.

[33] We find that increasing SNT_i reduces female applications for *treated* ads (i.e., negative β_M and β_T). Standard models of updating (e.g., Bayesian) do not necessarily imply that such an effect on treated ads must be accompanied by a negative effect on control ads; the predicted sign depends on applicants' priors, the cost of applying, and other parameters we cannot observe.

than moving from two to one). That said, Section 3.3 presents evidence suggesting that gender-neutral language in the ad text, not only in titles, drives the results.

One approach is to compute elasticities from removing gendered words. Focusing on the bottom quartile of SNT_i, one can use the β_0 from column (1) of Table 2 and the β_0 for the number of gendered words in the entire text from Table 1, along with the relevant control means, to compute an elasticity of $\frac{0.037}{-1.18} \cdot \frac{2.852}{0.146} = -0.61$. This suggests that a 10% increase in the number of gendered words lowers the share of female applicants by 6%.

This calculation assumes that the effect of a gendered word is the same regardless of where it appears (e.g., in the title, when describing the position, or in the benefits section). Different assumptions yield different magnitudes. For example, an elasticity based only on the title and job description and requirements section—implicitly assuming only words in these sections matter—would have half the magnitude: -0.31.

Tables A.7 and A.8 present treatment-on-treated results from 2SLS estimation where the endogenous (instrumented) treatment is a dummy equal to one if the ad title and job description and requirements section is gender-neutral (i.e., contains zero gendered words). These are arguably the most salient sections for applicants and also the ones Get on Board staff focused on most heavily when implementing the treatment (Table 1). Treatment-on-treated effects for the bottom quartile of SNT_i are substantial—close to 9 p.p., or nearly 60% of the control mean of 14.6%. However, more modest effects, such as a 1.5 p.p. increase (about 10% of the control mean), also fall within the confidence interval. See Appendix D for further discussion and details of the 2SLS implementation.

Effect sizes in related literature. Other studies report similarly large effects of job ad content. Kuhn and Shen (2023) finds that removing explicit male preferences increases the share of female applicants by 89% (a 4.95 p.p. increase from a 5.6% baseline). Card et al. (2024) does not observe applications but finds that eliminating stated male preferences raises the probability of hiring a woman by 167% (a 5 p.p. increase from a 3% baseline). Experimental studies in this literature are not directly comparable, as they estimate individual-level effects from a sample of interested applicants (see Appendix B). For instance, Coffman et al. (2024) finds that removing ambiguity surrounding required qualifications with a prescriptive statement increases the likelihood of qualified women applying by 28 p.p., from a control mean of 6%.

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 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.

Columns (1) and (2) of Table 3 first replicate our main ITT results, restricting the sample to ads where the posting firm used the evaluation board. Results are similar to those in Table 2, with standard errors and p-values increasing modestly given that the sample size is smaller (1,714 instead of 2,201 ads). Columns (3) to (8) then provide ITT effects 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 4.3 p.p. for the bottom quartile (with a smaller and insignificant effects 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.[34]

Overall, we find suggestive evidence that effects on the female share of applicants translate into more female representation in 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 their representation in the applicant pool.

3.3 Placebo Tests, Robustness Checks, and Additional Results

Placebo tests. Figure 4 examines how the main ITT results (equation (1) and Table 2) are influenced by different time windows used to define neighbor ads. Our baseline specification considers as ad i's neighbors 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 window of five or seven days before and after 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

[34] 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." Columns (1) and (2) restrict the sample to be the firms that used the "not discarded" category. The sample sizes thus decrease along the selection process.

across quartiles indeed drive the results.[35]

Table A.9 provides another placebo test. It re-estimates the main ITT results 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." These results and 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 effect heterogeneity, rather than other characteristics of job title groups.

Using ad field to define neighbors. As discussed in Section 2.1, most applicants find ads by searching job titles in a search bar, but the platform also allows users to browse ads through a predetermined set of 12 fields (listed in Appendix D). Table A.10 replicates our main ITT results (Table 2), but instead of using job title groups to define neighbors, it uses fields (i.e., ad i's neighbors are ads in the same *field* posted 3 days before or after, and we calculate an analogous SNT_i using this set of neighbors: SNT_i^{field}).

The results are similar to the main ITT results: point estimates have the same sign and comparable magnitudes. In particular, effects on the female share of applicants are similar but somewhat smaller (e.g., the effect for the bottom quartile of SNT_i is 3.7 p.p. in Table 2 versus 2.6 p.p. in Table A.10). There are two non-mutually exclusive interpretations for this result. First, the share of users browsing fields is smaller than the share searching job titles (Section 2.1), making SNT_i a stronger predictor of effect heterogeneity than SNT_i^{field}. Second, SNT_i may be the only variable that truly drives effect heterogeneity, but it is correlated with SNT_i^{field}. While job title groups and fields do not map perfectly onto each other, they are strongly associated, and the correlation between SNT_i and SNT_i^{field} is large.[36]

Are effects driven by the ad title or its main text? Table A.11 replicates our main ITT estimates (Table 2) separately for ads with titles in English and Spanish. As Table 1 shows, roughly half of the ads have a job title in English (e.g., "programmer" instead of "programador" or "programadora/o"). Among these, 80% have their main text in Spanish. English titles are, by default, gender neutral. 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 Table A.11 suggest similar effects for ads with texts in English

[35] As previously discussed in this section, SNT_i converges to 0.5 as the number of neighbors n_i grows, as depicted in panel (a).

[36] e.g., of the 115 ads in the "mobile developer" job title group, 103 are in the "mobile" field and 12 in "programming." The Cramer's V statistic of association between the two group categories is 0.58 (p-value < 0.001). A regression of SNT_i against SNT_i^{field} yields a coefficient of 0.497 ($SE = 0.031$, $R^2 = 0.15$).

or Spanish, indicating that gender-neutral language in the *main text* of the ad plays a role.[37]

Effect heterogeneity by share of female applicants in job title. Table A.12 replicates the main ITT results, but instead of exploring heterogeneity in SNT_i, it examines heterogeneity based on the share of female applicants in the job title group. This dimension of heterogeneity is the second-most-important factor identified in the causal forest (Figure 3). Moreover, Galos and Coppock (2023) shows that the gender composition of an occupation predicts gender bias. However, Table A.12 does not show that effects vary by quartiles of female share of applicants, indicating it does not drive effect heterogeneity on its own.[38]

A potential reason that the causal forest finds this variable "important" is that it predicts heterogeneity when interacted with SNT_i. Table A.13 provides suggestive evidence in this regard. It replicates the main ITT results but separately for ads in job title groups with a share of female applicants below and above its median. Although effects are noisily estimated, given the smaller sample size, the results suggest a stronger effect for ads in the bottom quartile of SNT_i in job title groups with higher female representation. This is consistent with gender-neutral language providing a stronger signal in more gender-inclusive occupations.

Additional results. As noted in the discussion equation (1) above, adding the *number* of neighbor ads as a control is unlikely to affect results. Table A.14 verifies this by replicating Table 2 with this variable included. As expected, the results are virtually the same. See Appendix D for further discussion. Columns (1)-(2) of Table A.15 replicate our main ITT estimates, weighing observations by the number of applications.[39] Columns (3)-(4) drops from the sample ads with main text in English. The estimates are similar to the unweighted baseline estimates. Columns (5)-(8) explore whether effects differ whether the ad is for a remote position. We do not find a clear pattern of heterogeneity. Table A.16 provides evidence that 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 D for further discussion.

Taking stock: interpretation of results. The results in this section highlight the role of treatment effect spillovers. Gender-neutral language in ads substantially increases the share of female applicants when likely listed next to a few other gender-neutral ads. However, when

[37] Table A.15 presents results dropping the 266 ads that are entirely in English - and for which our treatment protocol would involve fewer changes to the ad. As expected, the results are similar to the main sample.

[38] The share of female applicants in job title variable used here is constructed solely using the control group, so it is not affected by treatment. See Appendix C.

[39] The rationale for this robustness test is that the distribution of applications is right-skewed (Section 2.1). We caveat this exercise by noting that the number of applications is itself a potential outcome of treatment.

the ad is among a larger number of gender-neutral ads, the effects become more muted. In the overall sample where half the ads are treated, the point estimate is zero.

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 conclude the section discussing potential experimenter demand effects and social desirability bias.

"Raw" averages. Figure 5 provides simple averages for all eleven outcomes described in Section 2.2. It does so separately for the three treatments. Since the experiment has a $2\times2\times2$ factorial design, other treatment conditions are balanced in these two-way comparisons.[40]

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 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.[41]

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

[40]e.g., when comparing gender-neutral to non-gender-neutral ads, 25% of ads in both groups are remote and have a diversity statement, 25% are non-remote with a diversity statement, and so on.

[41]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.

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.[42]

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 the eleven outcomes by the three different treatment statuses, replicating for CDFs what Figure 5 does for averages (Figures A.7, A.8, and A.9). In cases we find effects on averages, they are driven by broad changes throughout the distribution of outcomes (e.g., a broad rightward shift in the CDF). Panel B of Table 4 provides the table counterpart of Figure 5. Appendix E provides further discussion and also replicates it splitting the sample by whether the respondents are alumnae of the web development or the UX design boot camps (Tables A.18 and A.19). Results are similar in magnitude, suggesting little heterogeneity by field. Table A.20 adds respondent fixed effects. As expected, given the experimental design, these within-estimates are 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).[43]

Estimating equation. Our main estimating equation is:

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

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

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

[43] In unreported regressions, we find that the results are also robust to excluding the Brazilian boot camp alumnae (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).

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 provide analogous effects of diversity statements and remote status. α_2 provides the effect of being the second ad assigned to non-gender-neutral, non-remote, and no-diversity-statement status.

As discussed in Section 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 consistent with 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. The design does, however, does not allow us to separate the mechanisms behind such spillovers (updating on the prevalence of gender-neutral ads versus a "salience effect").

Equation (2) differs from Figure 5 on two dimensions. First, Figure 5 provides two-way comparisons of means, while equation (2) estimates the effects of the three treatments jointly. This is inconsequential, as expected from a factorial design that ensures the three treatments are uncorrelated with each other. Second, and more importantly, it allows us to estimate the effects of first and second ads separately.[44]

Panel A of Table 4 provides the results. Overall, it shows that the effects of gender-neutral language are substantially larger for second ads when compared to first ads: β_2 is positive and significant for nine (out of eleven) outcomes.[45] As discussed above, this pattern is consistent with 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 we find effects, they are similar for both the first and second ad (i.e., γ_2 and δ_2 are relatively small and we cannot reject that they are zero). These provide a "placebo test," in the sense that it is not the case that all effects are simply stronger for second ads for reasons unrelated to spillovers. However, we caveat that, given independent draws for these

[44]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.

[45]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.10 and A.11 replicate Figure 5 for first ads and second ads only.

two treatments (Section 2.2), γ_2 and δ_2 do not have a similar interpretation as β_2 (e.g., half of the respondents exposed to a remote second ad also saw a remote first ad). For comparison, Panel B of Table 4 provides estimates of equation (2) without interactions with ad order.

Experimenter demand effects and social desirability bias. Five factors suggest that experimenter demand effects or social desirability bias cannot explain our results. First, as described in Section 2.2, 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 experimenter demand or social desirability mechanism that operates for gender-neutral language would also operate for such a closely related treatment. Fourth, it is unclear why experimenter demand or social desirability 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 some incentive to respond truthfully since their answers would impact the future job recommendations they received from Laboratoria.

Interpretation. Overall, our results are consistent with respondents using gender-neutral language as a signal about job characteristics. Indeed, the only outcome not affected by it is a question that does not involve beliefs about employer characteristics (whether respondents believe they meet the requirements for the job, which are clearly stated 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 attributes 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 respondents see a non-gender-neutral ad (compared to the first ads, which respondents evaluate without a clear comparison).

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 would not reveal whether these effects would persist if a higher share of ads were treated.

In a second experiment, gender-neutral language in ads influences beliefs about job attributes, particularly when the comparison to non-gender-neutral ads is salient. Overall, the results from both experiments are consistent with female applicants using gender-neutral language as a signal and using it to infer about job amenities and firm characteristics.

While the overall policy implications may seem negative due to limited scalability, some results suggest gender-neutral language can positively affect diversity in certain circumstances. We find suggestive evidence that when treatment affects applicant pool diversity, it does so by generating new applications on net and also influences who advances in the selection process and potentially gets hired. Moreover, inclusive language may have longer-term effects our designs cannot capture—for example, signaling to women that entering the field and accumulating human capital is worthwhile. This underscores the value of studying light-touch, virtually costless job ad interventions like the one we examine. We hope future research investigates 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.

Arceo-Gomez, E. O., R. M. Campos-Vazquez, R. Y. Badillo, and S. Lopez-Araiza (2022): "Gender stereotypes in job advertisements: What do they imply for the gender salary gap?" *Journal of Labor Research*, 43, 65–102.

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 (2022): "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 (2013): "Gender discrimination in job ads: Evidence from China," *The Quarterly Journal of Economics*, 128, 287–336.

——— (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.

Maestas, N., K. J. Mullen, D. Powell, T. Von Wachter, and J. B. Wenger (2023): "The value of working conditions in the United States and implications for the structure of wages," *American Economic Review*, 113, 2007–2047.

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.

Papadopoulos, B. (2022): "A brief history of gender-inclusive Spanish," *Deportate, esuli, profughe*, 48, 31–48.

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

——— (2019b): "Language Influences Public Attitudes toward gender equality," *The Journal of Politics*, 81, 81–93.

Pérez, E. O. and M. Tavits (2017): "Language shapes people's time perspective and support for future-oriented policies," *American Journal of Political Science*, 61, 715–727.

Read, D., G. Loewenstein, M. Rabin, G. Keren, and D. Laibson (2000): "Choice bracketing," *Elicitation of preferences*, 171–202.

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.

Wiswall, M. and B. Zafar (2018): "Preference for the workplace, investment in human capital, and gender," *The Quarterly Journal of Economics*, 133, 457–507.

Figure 1: Example Ad in Control and Treatment Versions, with Differences Highlighted

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 el 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. En CloudSystems estamos comprometidos con la diversidad y no aceptamos ningún tipo de discriminación.

Funciones

Buscamos ingenieras e 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 desarrolladoras/es.

· 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

· Espacio intelectualmente desafiante, diverso, inclusivo, y comprometido con su entorno.

· 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

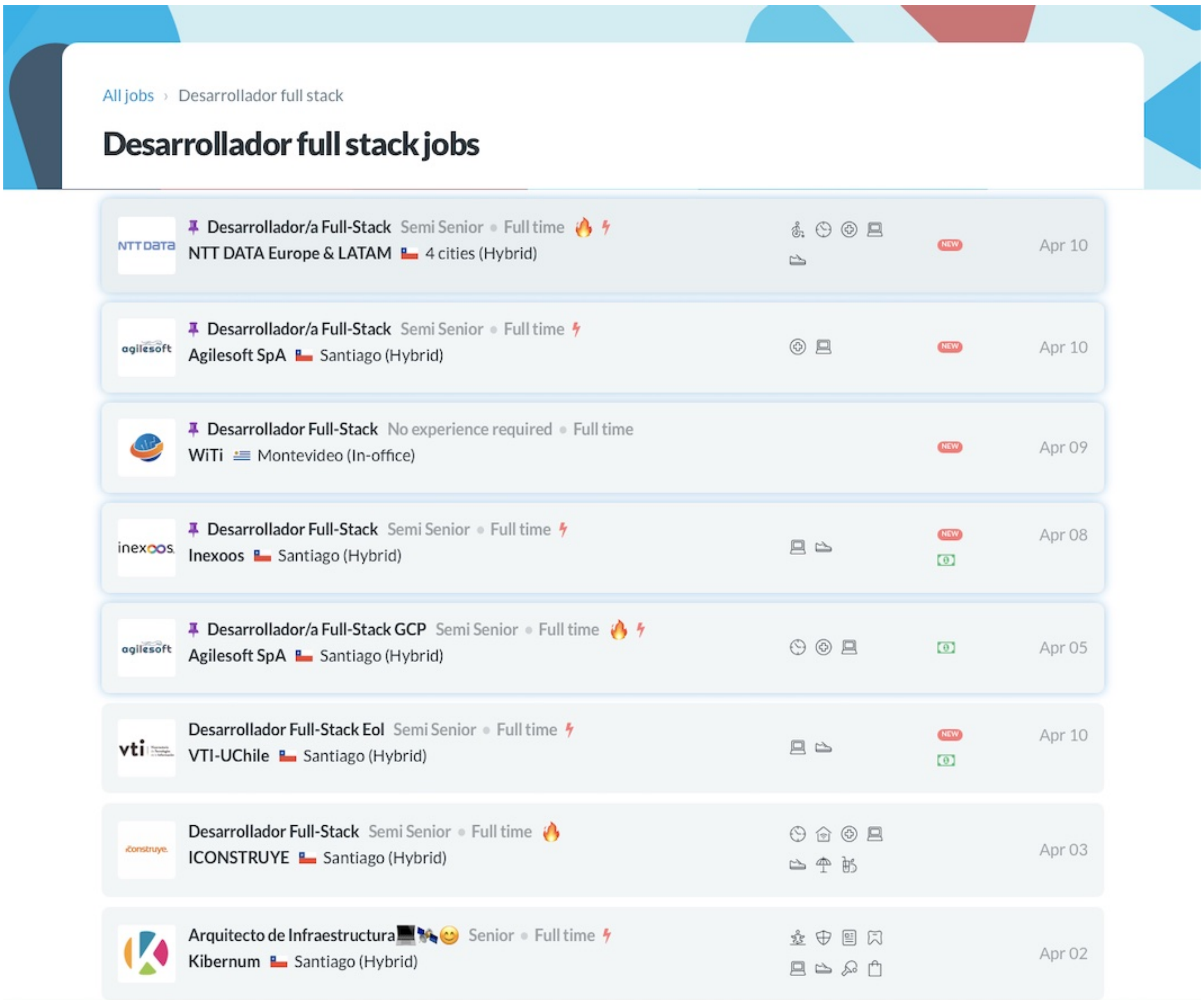
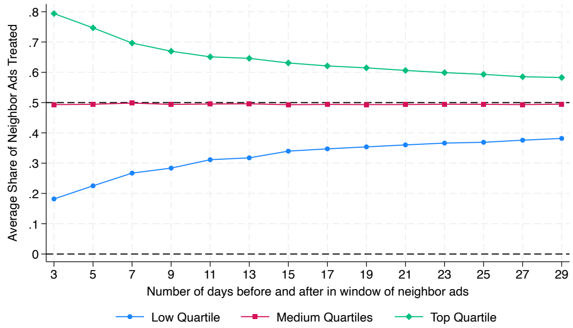


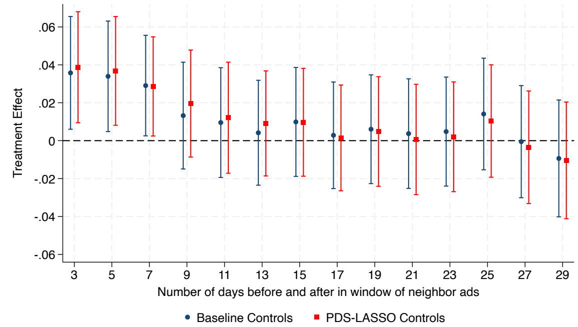
Figure 3: Causal Forest - Covariates' Importance

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.2 (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.

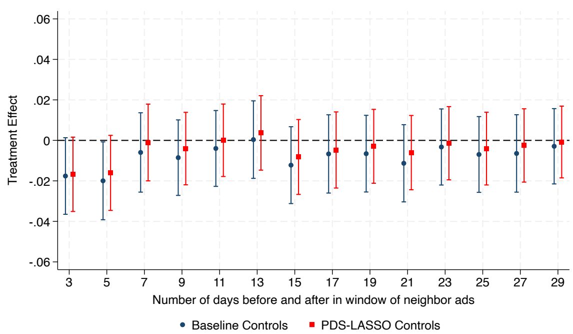
Figure 4: Treatment Effects for Different Time Windows Used in Defining Neighbor Ads - Get on Board



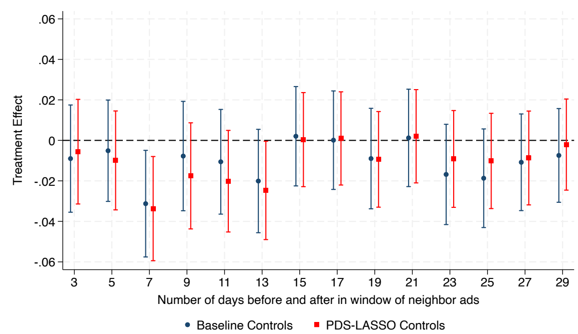
(a) Avg. Share of Neighbor Ads Treated



(b) Bottom Quartile of % Neighbor Ads Treated



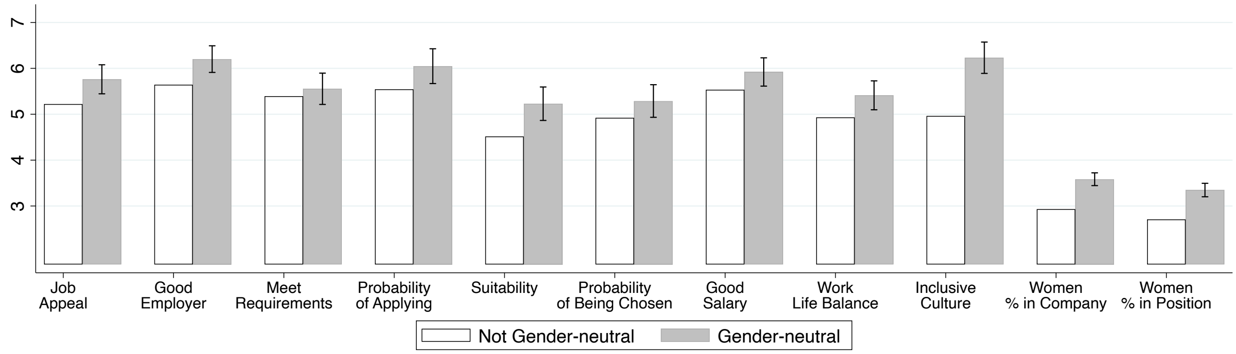
(c) Mid Quartiles of % Neighbor Ads Treated



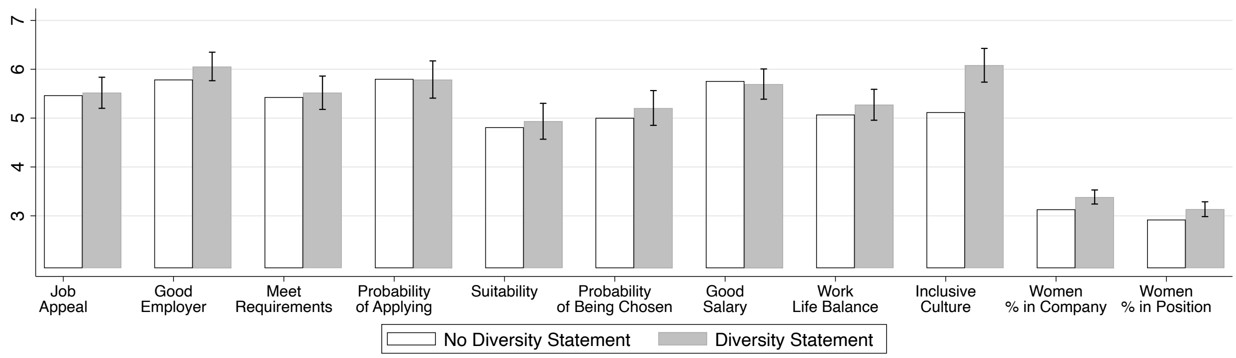
(d) Top Quartile of % Neighbor Ads Treated

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, 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.

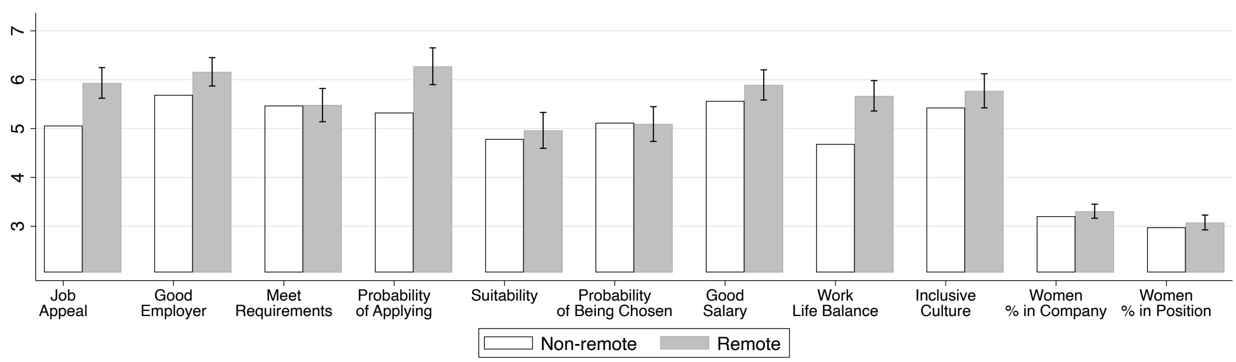
Figure 5: Outcome Averages by Different Treatment Statuses - Laboratoria



(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: Treatment Effects on Ad Text and Language - Get on Board

	(1) Control Mean	(2) Treatment (β_0)	(3) Treat × Mid. Quartiles of % Neighbors Treated (β_M)	(4) Treat × High Quartile of % Neighbors Treated (β_T)
Panel A: Outcomes are the number of gendered words by ad section				
Ad Title	0.380	-0.365***	0.048	0.024
		(0.033)	(0.040)	(0.046)
Company Description Section	1.421	-0.211*	-0.087	-0.069
		(0.125)	(0.158)	(0.183)
Job Description and Requirements Section	0.685	-0.531***	0.075	0.032
		(0.083)	(0.101)	(0.118)
Desirables Section	0.107	-0.073**	0.008	0.062
		(0.034)	(0.046)	(0.045)
Benefits Section	0.259	-0.000	-0.041	0.026
		(0.072)	(0.088)	(0.096)
Entire Ad Text	2.852	-1.180***	0.002	0.075
		(0.182)	(0.230)	(0.260)
Panel B: Outcomes are dummy indicators for				
Title and Job Desc. and Requir. Section is GN	0.435	0.421***	-0.026	-0.033
		(0.036)	(0.045)	(0.052)
Entire Text is Gender-neutral	0.187	0.169***	-0.006	-0.071
		(0.036)	(0.045)	(0.051)
Title is in English	0.521	-0.056	0.025	0.004
		(0.041)	(0.051)	(0.059)
Title is in Spanish (Gendered)	0.365	-0.352***	0.051	0.022
		(0.031)	(0.038)	(0.044)
Title is in Spanish (Gender-neutral)	0.113	0.408***	-0.075*	-0.026
		(0.034)	(0.043)	(0.050)
Entire Ad Text is in English	0.112	0.022	-0.009	-0.011
		(0.026)	(0.032)	(0.038)

Notes: Each row reports a separate regression for the outcome listed in the leftmost column, based on equation (1). The unit of observation is an ad. 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; only the treatment dummy and interactions are reported. All regressions include 2,201 observations, baseline covariates (month dummies interacted with remote status), as well as Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. The control mean is the average of the outcome variable for control ads. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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)	0.037**	0.039***	0.194	0.175	-0.075	-0.090	0.032	0.022
	(0.015)	(0.015)	(0.135)	(0.133)	(0.101)	(0.101)	(0.049)	(0.048)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.053***	-0.055***	-0.237	-0.216	0.147	0.150	-0.018	-0.002
	(0.018)	(0.018)	(0.160)	(0.157)	(0.121)	(0.121)	(0.062)	(0.061)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.046**	-0.044**	-0.226	-0.188	0.059	0.076	0.134*	0.134*
	(0.020)	(0.020)	(0.183)	(0.182)	(0.141)	(0.141)	(0.072)	(0.071)
Mid. Quartiles of % Neighbors Treated (γ_M)	0.021	0.024*	0.080	0.093	-0.075	-0.073	0.019	0.010
	(0.013)	(0.013)	(0.123)	(0.120)	(0.090)	(0.089)	(0.046)	(0.046)
Top Quartile of % Neighbors Treated (γ_T)	0.002	0.003	0.007	0.005	-0.032	-0.058	-0.050	-0.054
	(0.014)	(0.014)	(0.131)	(0.129)	(0.101)	(0.099)	(0.050)	(0.049)
Implied Treatment Effects								
Bottom Quartile of % Neighbors Treated (β_0)	0.037	0.039	0.194	0.175	-0.075	-0.090	0.032	0.022
	(0.015)	(0.015)	(0.135)	(0.133)	(0.101)	(0.101)	(0.049)	(0.048)
	[0.035]**	[0.022]**	[0.210]	[0.261]	[0.504]	[0.413]	[0.558]	[0.701]
Mid. Quartiles of % Neighbors Treated (β_0+β_M)	-0.016	-0.016	-0.043	-0.041	0.072	0.061	0.014	0.020
	(0.009)	(0.009)	(0.086)	(0.083)	(0.067)	(0.067)	(0.038)	(0.038)
	[0.163]	[0.162]	[0.665]	[0.679]	[0.314]	[0.392]	[0.709]	[0.622]
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	-0.009	-0.005	-0.033	-0.013	-0.017	-0.014	0.165	0.156
	(0.013)	(0.013)	(0.124)	(0.124)	(0.099)	(0.098)	(0.052)	(0.052)
	[0.608]	[0.757]	[0.829]	[0.932]	[0.873]	[0.892]	[0.002]***	[0.003]***
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.146	0.146	1.764	1.764	3.718	3.718	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); 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 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). Reported 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with SNT_i in the bottom quartile, the middle quartiles, and the 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. p-values from a randomization inference procedure that accounts for dependencies induced by spillovers are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: 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)	0.035**	0.036**	0.043**	0.044**	0.044	0.050	0.133*	0.109
	(0.016)	(0.016)	(0.019)	(0.018)	(0.041)	(0.042)	(0.072)	(0.071)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.045**	-0.047**	-0.054**	-0.056**	-0.095*	-0.105**	-0.215**	-0.195**
	(0.019)	(0.019)	(0.022)	(0.022)	(0.050)	(0.050)	(0.088)	(0.085)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.028	-0.029	-0.032	-0.033	-0.038	-0.045	-0.081	-0.029
	(0.022)	(0.022)	(0.026)	(0.026)	(0.057)	(0.056)	(0.095)	(0.094)
Mid. Quartiles of % Neighbors Treated (γ_M)	0.008	0.013	0.016	0.022	0.001	0.008	0.082	0.060
	(0.014)	(0.014)	(0.017)	(0.016)	(0.037)	(0.037)	(0.066)	(0.062)
Top Quartile of % Neighbors Treated (γ_T)	-0.011	-0.007	-0.004	0.000	-0.040	-0.034	-0.034	-0.070
	(0.015)	(0.014)	(0.017)	(0.017)	(0.039)	(0.038)	(0.065)	(0.062)
Implied Treatment Effects								
Bottom Quartile of % Neighbors Treated (β_0)	0.035	0.036	0.043	0.044	0.044	0.050	0.133	0.109
	(0.016)	(0.016)	(0.019)	(0.018)	(0.041)	(0.042)	(0.072)	(0.071)
	[0.064]*	[0.050]*	[0.044]**	[0.034]**	[0.321]	[0.275]	[0.054]*	[0.123]
Mid. Quartiles of % Neighbors Treated ($\beta_0+\beta_M$)	-0.010	-0.011	-0.012	-0.012	-0.051	-0.055	-0.082	-0.086
	(0.010)	(0.010)	(0.012)	(0.012)	(0.027)	(0.027)	(0.048)	(0.047)
	[0.408]	[0.377]	[0.424]	[0.403]	[0.086]*	[0.068]*	[0.092]*	[0.073]*
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	0.007	0.008	0.011	0.011	0.007	0.005	0.052	0.079
	(0.015)	(0.014)	(0.018)	(0.018)	(0.039)	(0.037)	(0.062)	(0.062)
	[0.700]	[0.663]	[0.604]	[0.565]	[0.890]	[0.910]	[0.479]	[0.259]
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.151	0.151	0.157	0.157	0.175	0.175	0.202	0.202
N	1,714	1,714	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); 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. Columns 1-2 replicate the first two columns of Table 2 restricting the sample to ads where the firm used the evaluation board. The top panel provides the estimated coefficients from equation (1). Reported 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. 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. p-values from a randomization inference procedure that accounts for dependencies induced by spillovers are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Treatment Effects by Ad Order - Laboratoria

	Job Appeal (1)	Employer (2)	Meet Requirements (3)	Probability Of Applying (4)	Suitability (5)	Probability Of Being Chosen (6)	Salary (7)	Work Life Balance (8)	Culture (9)	Women % Company (10)	Women % Position (11)
Panel A: Treatment Dummies Interacted with Ad Order											
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)
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)
Gender-neutral × 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	-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)
Panel B: Treatment Dummies Without Interactions with Ad Order											
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	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 - 1st Ad	4.815	5.061	4.692	5.062	4.446	4.431	5.431	4.277	4.369	2.561	2.455
N	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 in Panel A presents estimates from equation (2) for a different outcome (see text for definitions). Panel B provides similar regressions without interacting treatments with ad order. 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 in their treatment of gender. Some languages do not make gender distinctions (e.g., Finnish), while others assign gender to all nouns, including inanimate objects (e.g., Spanish, Portuguese, French, Italian). English is 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 (2022) 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 finds that 39% of the world's population speaks a gendered grammar language. Additional examples are German, Russian, Arabic, Hindi, Somali, Hebrew, and Urdu.

Gendered grammar in Spanish. This section describes the traditional grammar in Spanish, but all statements 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 to match the gender of the noun. Indefinite articles in Spanish are the masculine and feminine "un" and "una" (and the plurals "unos" and "unas"). Similarly, definite articles are the feminine "la" (plural "las") and the masculine "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” implies an all-female group of engineers.

This *generic masculine* form is traditional and common in Spanish. In situations where no gender must be specified (such as a job ad for an engineer), the standard is to state that a company is looking to hire an “ingeniero” or multiple “ingenieros.”

Inanimate objects have gender: e.g., 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 articles, these nouns are also gendered. For example, “the company is hiring **an** economist” can either be translated to “la empresa está contratando **un** economista” (implying a male economist) or “la empresa está contratando **una** economista” (implying a female economist). A similar issue applies with plurals (“unas economistas” versus “unos economistas”).

Gender-neutral language in Latin America. Advocacy for gender-neutral language in Spanish dates to at least the 1970s and has intensified across Latin America in recent years, generating both institutional adoption and public debate (Papadopoulos, 2022). 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.

Both our experiments follow an arguably less radical approach, which is also the one advocated by some Latin American governments. Our approach is thus gender-neutral in the sense it includes both male and female genders, but (unlike some other forms of inclusive gender) does not address potentially non-binary genders. 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.[1]

The adoption of gender-neutral language has attracted substantial controversy and government intervention in Latin America. For example, in July 2022 the Buenos Aires government 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 no official policy regarding gender-neutral language in Buenos Aires, and some teachers had informally adopted it.

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

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. Several studies across disciplines study 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 attitudes and opinions when surveyed in different languages (Ogunnaike et al., 2010, Danziger and Ward, 2010, Pérez and Tavits, 2017, Pérez and Tavitz, 2019a). 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 generic masculine forms. Moulton et al. (1978) found evidence that when the terms "he," "him," and "man" were expressed as generic masculines, respondents 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) provides 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. Pérez and Tavitz (2019b) finds that bilinguals are less supportive of gender equality when interviewed in a language with gendered grammar. Jakiela and Ozier (2022) 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 generic masculines are not 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 *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 February 2026, the feminine-pronoun version of the book's second edition had a larger number of downloads, 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 the content and language of job ads affect the composition of the applicant pool. To our knowledge, this is the first study to evaluate gender-neutral language and the first to experimentally 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 on the content of job ads. The closest work to ours studies interventions that can be interpreted as changes in *language*, without *explicitly* changing information about job attributes or employer preferences but can still signal them to potential applicants, such as removing optional qualifications and superfluous language (Abraham et al., 2024) or reducing reducing ambiguity on required qualifications (Coffman et al., 2024).[2]

Another set of papers more explicitly suggests employer preferences via diversity statements (Ibañez and Riener, 2018, Leibbrandt and List, 2018, Flory et al., 2021) or varying the gender of workers depicted in photographs (Delfino, 2024). In Latin America, Del Carpio and Guadalupe (2021) investigates 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.

Another set of papers involves experimental interventions that provide factual descriptions about objective job characteristics, such as indicating flexible work hours (Mas and

[2] 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.

Pallais, 2017), negotiable salaries (Leibbrandt and List, 2015), competitive compensation regimes (Flory et al., 2015, Samek, 2019), information on the share of workers receiving high evaluations (Delfino, 2024), or providing information on the number of competing applicants for a job (Gee, 2018). Other papers vary posted wages to test job search models' predictions rather than applicant diversity (e.g., Belot et al., 2019, Banfi and Villena-Roldan, 2019).

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. For example, after an applicant expresses interest in the position, she receives an individualized e-mail. The treatment is embedded in the content of such an e-mail, allowing applicant-level randomization. Thus, their experimental designs do not allow for the study of the type of spillovers that is the focus of this paper. We exploit multiple ads from different firms being treated, creating (random) variation in the share of treated ads that applicants consider.[3]

Non-experimental papers on *explicit* gender requests. As discussed in the introduction, Kuhn and Shen (2023) and Card et al. (2024) use difference-in-differences designs to study reforms that eliminate *explicit gender requests* or *stated gender preferences* in job ads. Compared to our intervention, these studies examine a more direct and overt change in ad content, where employers explicitly indicate a preference for male or female applicants.

Kuhn and Shen (2023) examines treatment spillovers since its difference-in-differences strategy allows estimating the reform's effect on ads that did not include gender requests before the policy change. This is a meaningful and policy-relevant spillover, but it differs from ours in that it does not directly address scalability, i.e., whether the effects vary depending on the share of ads with gender requests in an applicant's choice set. In contrast, Card et al. (2024) does not study spillovers; its design uses ads with stated gender preferences as the treatment group and those without as the control group. However, both papers capture general equilibrium effects, as their difference-in-differences designs estimate impacts in settings where the share of ads with explicit gender requests changes across the entire market due to the reform.

We now expand on a comparison briefly mentioned in the introduction. While we find a zero point estimate for our overall sample, both studies find that removing explicit requests for male applicants from an ad increases the share of women applying (in China) and hired (in Austria). All three studies, including ours, estimate effects at the individual ad level, not the aggregate impact of the ban on the platform or market

The "updating mechanism" can reconcile the findings of Kuhn and Shen (2023), Card

[3]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.

et al. (2024), and this paper. It relies on applicants interpreting gender-neutral language or explicit gender requests as signals about firm characteristics or job amenities (e.g., they expect that firms using gender-neutral language are more likely to offer flexible hours or employ more women).

In the Chinese job board studied by Kuhn and Shen (2023), 12.5% of pre-reform ads requested male applicants; in the Austrian setting analyzed by Card et al. (2024), the figure was 20%. Post-reform, these shares dropped to nearly zero, as the reforms aimed to ban gender requests. In our setting, gender-neutral language is more common even without treatment: 19% of control ads have no gendered word in their entire text, and 44% have no gendered word in their title or job description and requirements section (Table 1). During the experiment, half of the ads were treated, making gender-neutral language less distinctive and thus less informative.

Thus, the updating mechanism may explain what initially appears to be a divergence in findings: had explicit male requests been more prevalent in the Chinese or Austrian settings, the marginal effect of removing any single request would likely have been smaller. Moreover, consistent with the mechanism, we find effects only when gender-neutral language is distinctive, i.e., when few neighbor ads are treated.

Several caveats apply. There may be no reason to expect similar results in the first place: our intervention involves a more subtle language change, whereas the other studies examine the removal of overt and explicit statements over gender preferences. Moreover, the contexts differ not only geographically (Latin America vs. China and Austria), but also in the nature of the jobs advertised: ours focus is on the tech sector, where acquiring the necessary skills may be more difficult in the short run.[4]

Lastly, see also Kuhn and Shen (2013) and Kuhn et al. (2020) on explicit gender requests in China and Helleseter et al. (2020) on requests on applicant *age* in the same context. Arceo-Gomez et al. (2022) uses gender-targeted advertisements in Mexico to predict whether non-targeted ads are directed toward men or women, based on the language they use, and how they differ from those effectively targeted toward men.

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 2.1. We describe

[4]Both Kuhn and Shen (2023) and Card et al. (2024) also examine the effects of explicit requests for *female* applicants. We do not focus on these in the discussion above (or in the introduction), as they are less relevant to our treatment and the language variation observed in our data, which includes only masculine-form and gender-neutral ads.

here 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*, *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). Table A.1 provides the 16 job title groups, their representation in the sample, and the share of female applicants.

Remoteness. Our experiment was conducted while mobility restrictions due to the Covid-19 pandemic were still in place and several ads listed a remote position. 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.[5]

Fields. As discussed in Section 2.1, users can browse through a predetermined set of 12 fields that Get on Board uses to classify ads, although this is not as common as searching. The fields (and the share of ads in the sample they represent) are Mobile (5.4%), Programming (57.0%), Data Analytics (4.7%), Sysadmin (9.0%), Operations (4.7%), Innovation/Agile (2.0%), Sales (1.5%), Customer Support (2.4%), Advertising/Media (0.6%), Design (8.7%), Digital Marketing (3.8%), and Human Resources (0.3%).

Share of female applicants in job title group. The share of female applicants in the job title group is a variable used only in the causal forest analysis (Figure 3) and Tables A.12 and A.13. It is constructed only using ads assigned to control. For each job title group, we

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

calculate the average share of female applicants to *control* ads. We then assign that value to all ads in that job title group. This variable thus measures female representation in a job title group in a baseline scenario in a manner not directly affected by our treatment. Table A.1 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.

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

Covariate balance. As discussed in Section 2.1, Tables A.2 and A.3 provide 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.2, a set of nine country dummies, a set of 16 job group title dummies, and a set of 12 field dummies.[6] We report the randomization inference (permutation) p-values 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 F-statistic computed with the actually realized treatment assignment). We use our entire sample (2,201 observations) and 1,000 repetitions. The p-value is 0.338, thus we cannot reject the null of joint covariate balance.[7]

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 also indicate covariate balance. The respective p-values are 0.241, 0.286, and 0.281.

Causal forests and treatment effect heterogeneity. As discussed in Section 3.1, 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 (2,201 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

[6] The set of country dummies includes a dummy equal to one if the ad did not specify a country of work (which is common for remote positions).

[7] 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.2, the minimum and maximum of the salary range are not included, since it is missing for ads that did not post a salary range.

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.2 (except the minimum and maximum of salary range, which is missing for ads that did not post a range). Appendix C discusses the construction of the share of female applicants in the job title group. 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.1 provides its values.

The set of covariates that can potentially predict effect heterogeneity differs slightly from the set of covariates we use as potential controls in the PDS-LASSO specification of equation (1), discussed in Section 3.1. 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%.

Implementing the Borusyak-Hull recentered treatment estimator. This appendix expands on the econometric issues raised in Section 3.1. Equation (1) implements the approach of Borusyak and Hull (2023), which shows that their "recentered instrument" estimator can be obtained by including the *expected* treatment variables as regression controls, and that a single linear control suffices for identification.[8] In our setting, the relevant variables are dummies, so the expected treatments are Prob$[MidQuartiles_i^{SNT} = 1]$ and Prob$[TopQuartile_i^{SNT} = 1]$. Note that the expected value of treatment (T_i) is always 0.5 and thus collinear with the constant. As discussed in Section 3.1, these probabilities can be calculated directly using binomial formulas.[9]

Importantly, implementing Borusyak and Hull (2023) does not require controlling for the expected value of the interactions between T_i and $MidQuartiles_i^{SNT}$ or $TopQuartile_i^{SNT}$; moreover, including them would complicate interpretation. The intuition is as follows. The variables $MidQuartiles_i^{SNT}$ and $TopQuartile_i^{SNT}$ contain both random variation (from treatment assignment) and non-random variation (from correlation with i's number of neighbors). Controlling for their expected values ensures the direct spillover effects (γ_M and γ_T) are estimated from random variation only (Borusyak and Hull, 2023). However, the interaction terms (β_M and β_T) are already identified from random variation alone. To see this, note that equation (1) can be estimated separately for ads in each quartile of SNT_i. For example, estimating (1) using only ads where $TopQuartile_i^{SNT} = 1$ identifies the effect for the top

[8] Borusyak and Hull (2023) use the terminology "recentered instrument" broadly, encompassing reduced-form or intent-to-treat regressions where the instruments and treatment variables coincide.

[9] Since treatment is assigned independently to each ad with the same probability, Prob$[MidQuartiles_i^{SNT} = 1]$ and Prob$[TopQuartile_i^{SNT} = 1]$ depend only on the number of neighbors, though in a nonlinear and non-monotonic fashion.

quartile ($\beta_0 + \beta_T$), and within this subsample, treatment and control groups are balanced, in expectation, on the number of neighbors.

Finally, the discussion above implies that controlling for the number of neighbors in equation (1) is unnecessary for implementing the Borusyak and Hull (2023) estimator and unlikely to affect results. We verify this by re-estimating our main results (Table 2) with the number of neighbors included as a control. For the PDS-LASSO specifications, we force its inclusion by adding it to the amelioration set. Table A.14 presents these results, showing that estimates are virtually the same to the baseline.

Effects on the distribution of the share of female applicants. As discussed in Section 3.2, Figure A.4 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 3.2, Figure A.5 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.4, 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 (the default score assigned to users when they register).

Figure A.6 repeats the exercise 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.6 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 briefly discussed in Section 3.2, we estimate treatment-on-treated effects in the following 2SLS framework:

$$y_i = \alpha^{2SLS} + \beta_0^{2SLS} GN_i + \beta_M^{2SLS} GN_i \cdot MidQuartiles_i + \beta_T^{2SLS} GN_i \cdot TopQuartile_i + \\ + \gamma_M^{2SLS} MidQuartiles_i + \gamma_T^{2SLS} TopQuartile_i^{SNT} + X_i'\theta^{2SLS} + \epsilon_i \quad (3)$$

where y_i is the share of female applicants and GN_i is a dummy equal to one if the title and job description and requirements section is gender-neutral (has zero gendered words). 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 A.7 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 8.8 p.p. or 9.3 p.p., depending on the controls used, and significant at the 5% level. The effects for the middle and top quartiles are, as expected, negative and statistically insignificant. 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 -0.0017 p.p. (SE=0.017).[10]

Columns (1) and (2) of Table A.8 present the estimates of β_O^{2SLS}, β_M^{2SLS}, and β_T^{2SLS} that inform the linear combinations reported on Table A.7. Columns (3)-(8) present the first-stage estimates. With three endogenous variables and three excluded instruments, there are three first stages.

We highlight three points about the first stages. First, they show a roughly 40 p.p. first-stage effect, consistent with Table 1. Second, for each first stage, the "relevant coefficient" is roughly 40 p.p. but the other two are close to zero and insignificant. For example, when the instrument is interacted with a dummy for *middle* quartiles of SNT_i, the "relevant coefficient" is the treatment interacted with a dummy for *middle* quartiles. The coefficients on the non-interacted treatment and its interaction with the top quartile dummy are essentially zero. This is expected given random assignment and can be interpreted as a "randomization check."

[10] 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.

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 larger than 10.

Additional results: subsequent ads. Table A.16 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 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 and job description and requirements is gender-neutral, or whether its title and job requirements and description section is gender-neutral (see Table 1 and related discussion in Section 3.2). Standard errors are clustered at the firm level. We estimate a θ_1 that is close to zero and insignificant. This suggests that, after having their first ad treated, firms are not more likely to post more ads using gender-neutral language.

E Additional Results, Tables, and Figures - Laboratoria

Balance and summary statistics. Table A.17 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.7, A.8, and A.9 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.7, A.8, and A.9 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 broad rightward shift in the CDF), implying effects along the entire distribution of outcomes.

Results in table format. Panel B of Table 4 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).[11]

In the terminology of Muralidharan et al. (2023), equation (6) and equation (2) that is reported on Table 4 estimate a "short model," as opposed to a fully interacted "long

[11] 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.

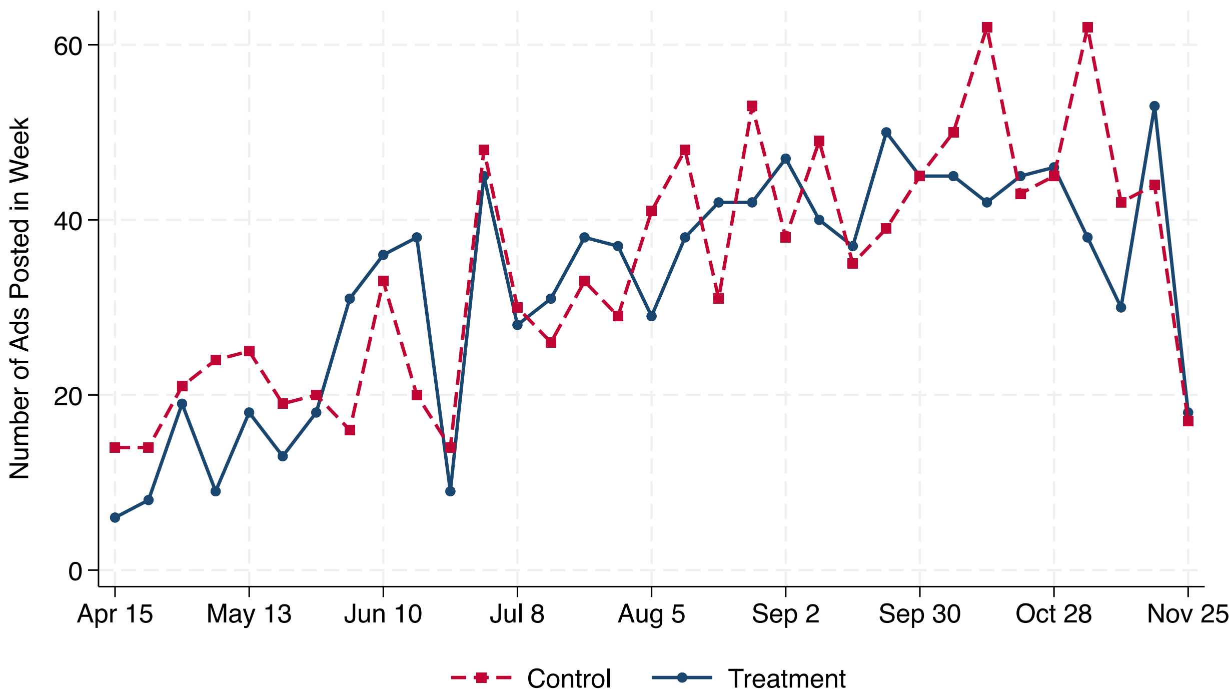
model." As discussed in Section 2.2, the "short model" is the appropriate choice in this context. The experiment's factorial design was designed to i) allow us to compare the effects of gender-neutral language to explicit diversity statements and a valuable job amenity (working remotely), and ii) to 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 AEA 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 estimation in experiments with factorial designs. Within their focus on cases where researchers are testing new policies that are "new" to their context, estimating interacted effects from "long models" is perhaps more suitable. However, in our context, all treatments represent relatively common practices, and the factorial design aims to make the sample more representative of "real" ads. With these caveats and considerations in mind, Table A.21 reports the fully-interacted "long model." Overall, the results suggest the effects of gender-neutral language have limited interaction with other effects (i.e., the main "non-interacted" effect of gender-neutral language is similar to the one reported in the short model, and its interaction with other treatments has smaller point estimates, which are not statistically significant).

Robustness checks and heterogeneity. Tables A.18 and A.19 replicate Panel B of Table 4 splitting the sample by whether the respondents are alumnae of the web development or the UX design boot camps, respectively. Results are similar in magnitude, suggesting little heterogeneity by field. Table A.20 replicates Panel B of Table 4 adding respondent fixed effects. As expected given the experimental design, these within-estimates are similar to those on Table 4. It is not possible to estimate effects by ad order (i.e., equation (2) reported in Table 4) 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 $2ndAd_{ia}$.

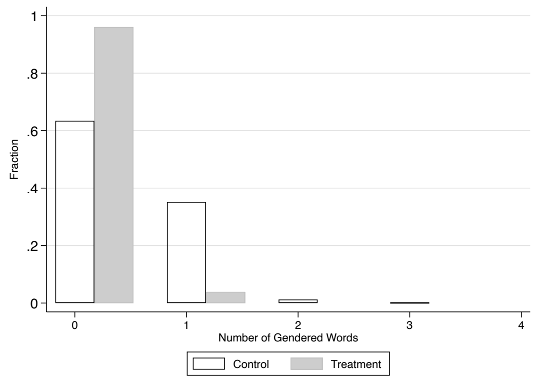
In unreported regressions, we find that the results are also robust to excluding the Brazilian boot camp alumnae (who answered a version of the survey in Portuguese) and excluding respondents who answered the survey "too quickly" (e.g., less than three or five minutes).

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

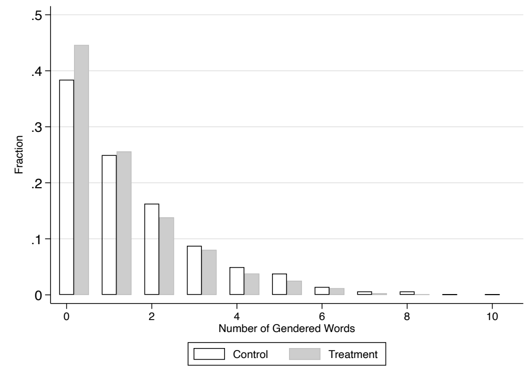


Notes: Figure provides the weekly number of ads posted during the experimental period (April 17 to November 27, 2020), by treatment assignment. Labels on the x-axis refer to the day a week starts (e.g., Apr 15 is the week of April 15-21). The drop at the final week (Nov 25) is due to the sample ending halfway during that week.

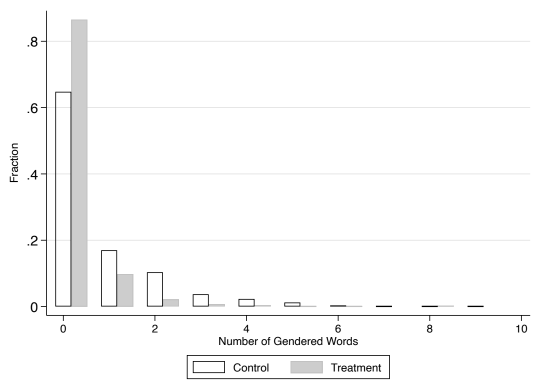
Figure A.2: Distribution of Gendered Word Counts in Ads - Get on Board



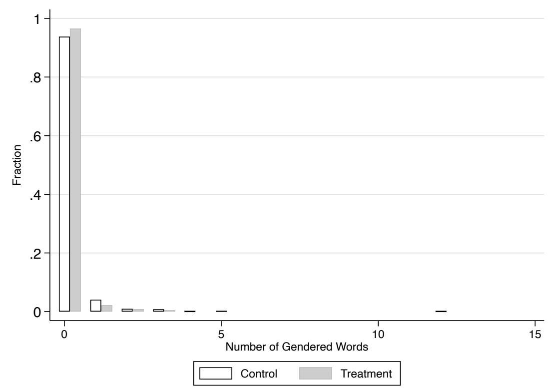
(a) Ad Title



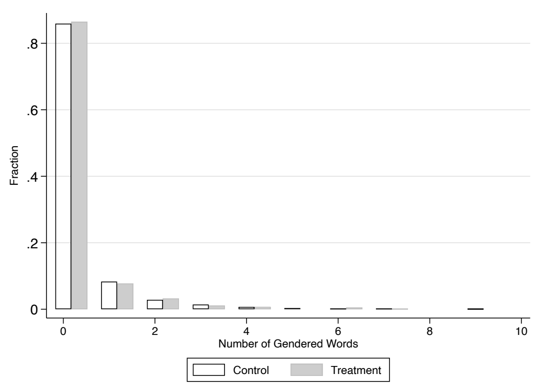
(b) Company Description Section



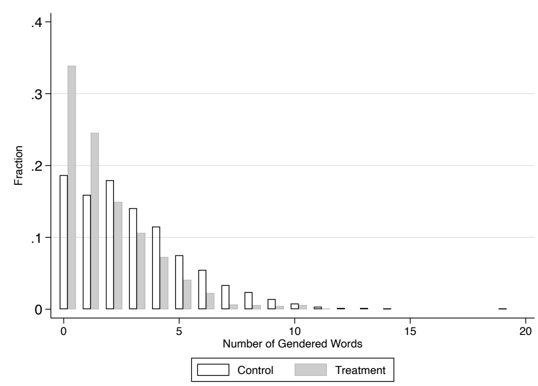
(c) Job Description and Requirements Section



(d) Desirables Section



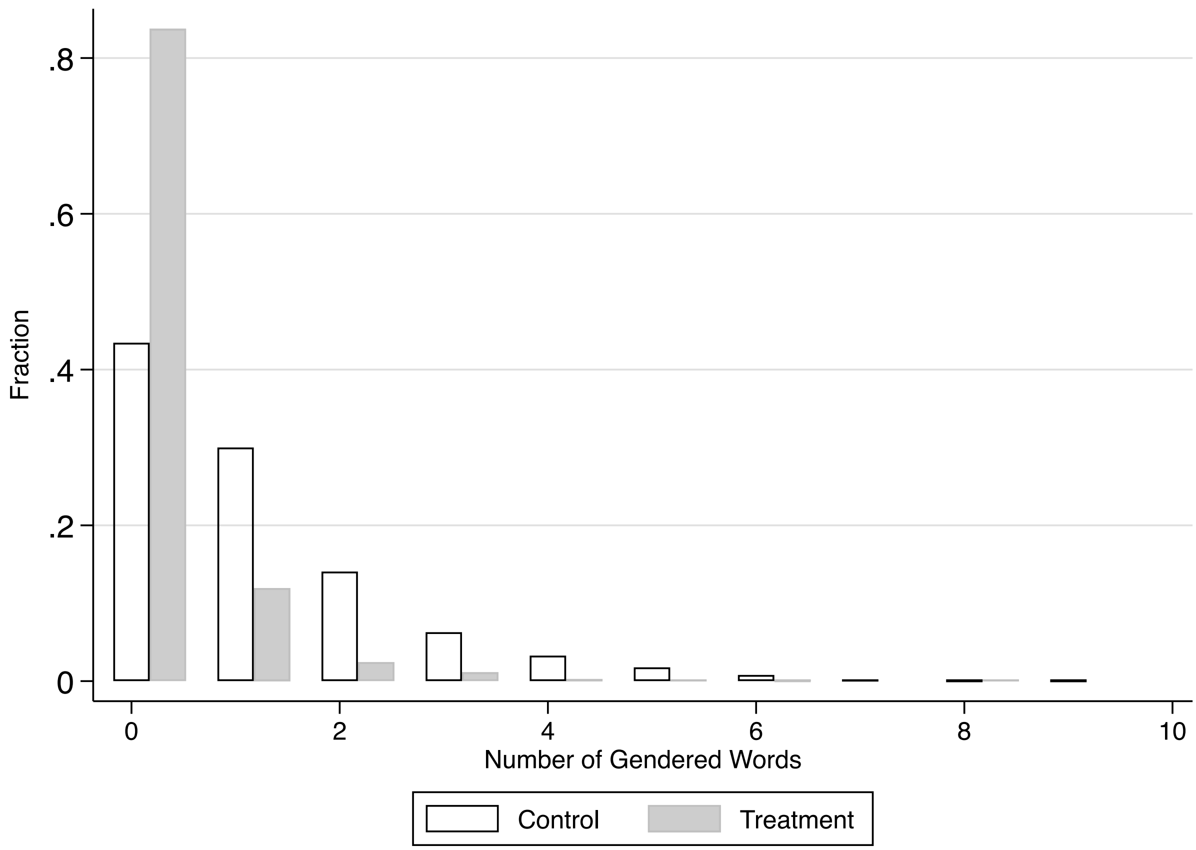
(e) Benefits Section



(f) Entire Ad Text

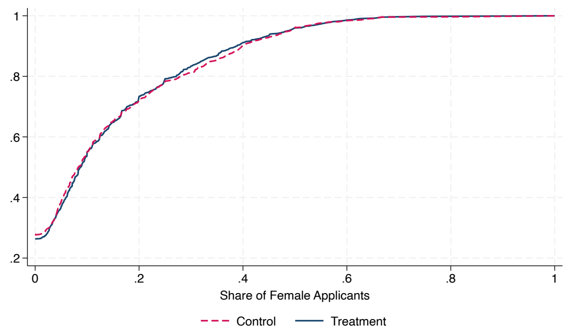
Notes: Unit of observation is an ad. Figures provide the histogram of the count of gendered words in the respective sections of the ad, separately for control and treated ads. Panel (f) does so for the entire text of the ad.

Figure A.3: Distribution of Gendered Word Counts in Ad Title and Job Description and Requirements Section - Get on Board

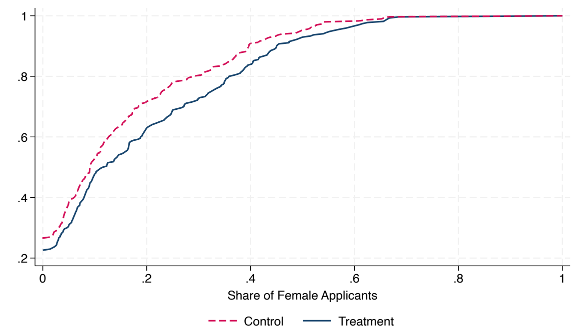


Notes: Unit of observation is an ad. The figure provide a histogram of the count of gendered words in the ad *title* and *job description and requirements section* of control and treated ads.

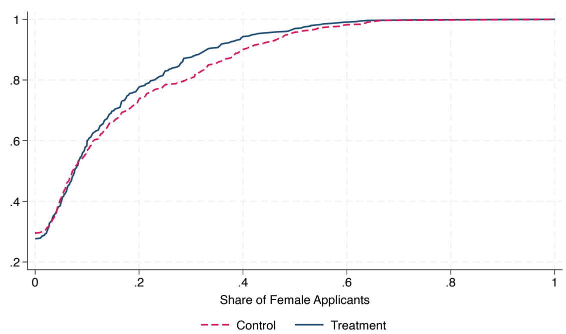
Figure A.4: Share of Female Applicants Distribution - Get on Board



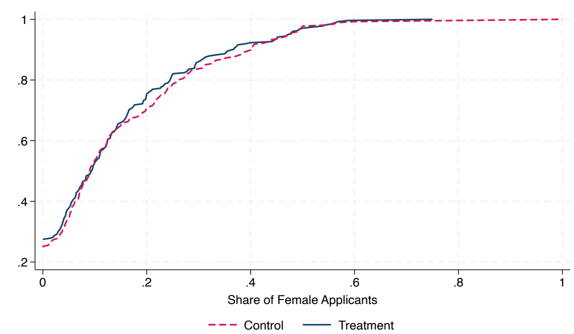
(a) All Ads



(b) Bottom Quartile of % Neighbors Treated



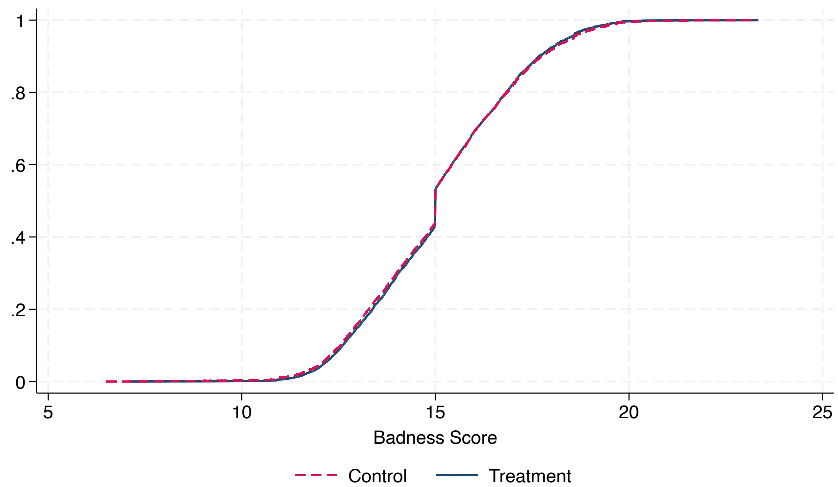
(c) Mid Quartiles of % Neighbors Treated



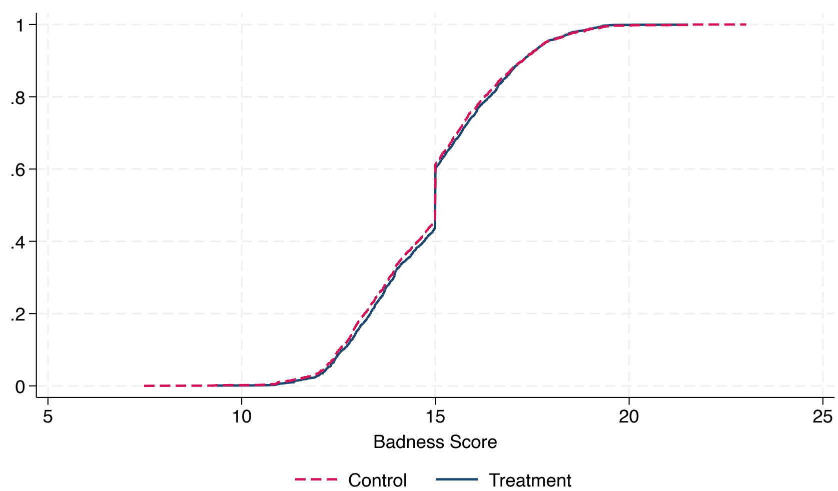
(d) Top Quartile of % Neighbors Treated

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.5: Badness Score Distribution - Get on Board



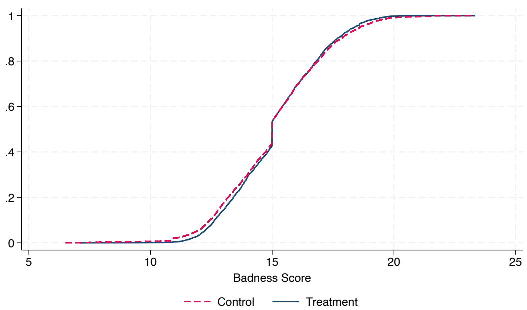
(a) Male Applicants, Full Sample



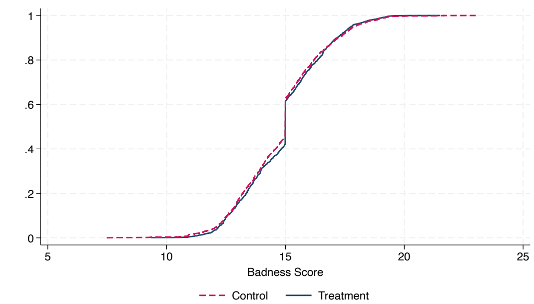
(b) Female Applicants, Full Sample

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).

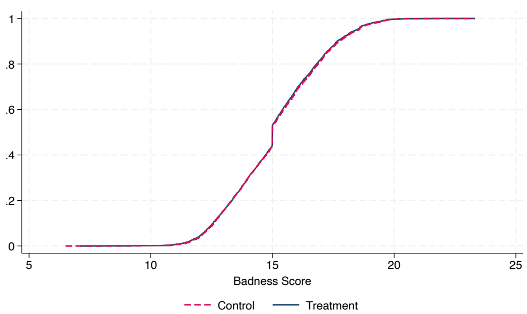
Figure A.6: Badness Score Distribution by Share of Neighbors Ads Treated - Get on Board



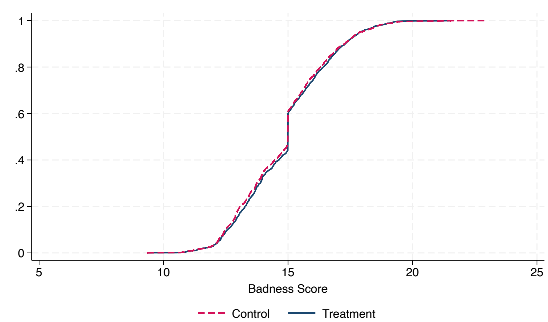
(a) Male Applicants,
Bottom Quartile of % Neighbors Treated



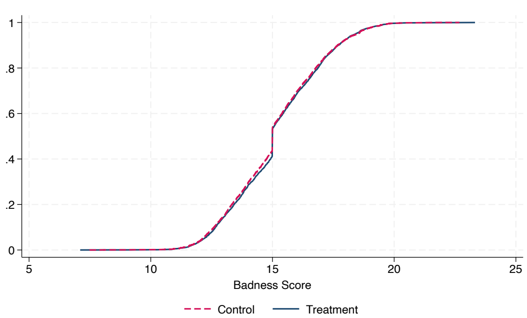
(b) Female Applicants,
Bottom Quartile of % Neighbors Treated



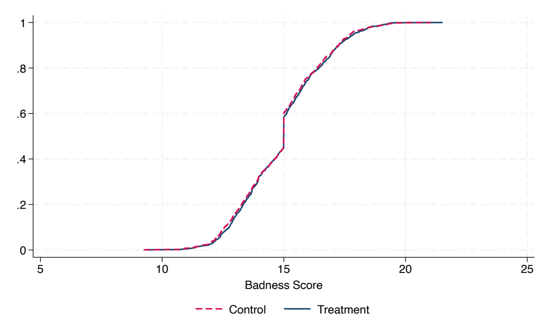
(c) Male Applicants,
Mid Quartiles of % Neighbors Treated



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



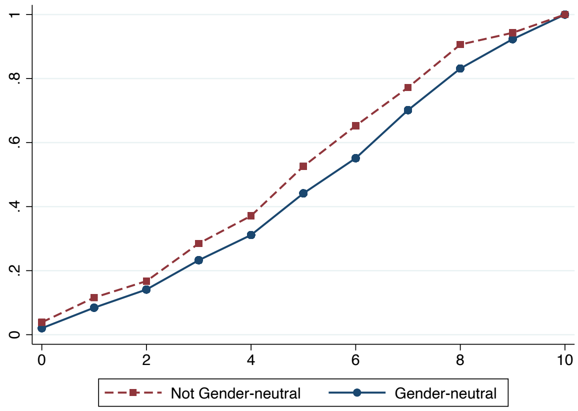
(e) Male Applicants,
Top Quartile of % Neighbors Treatedd



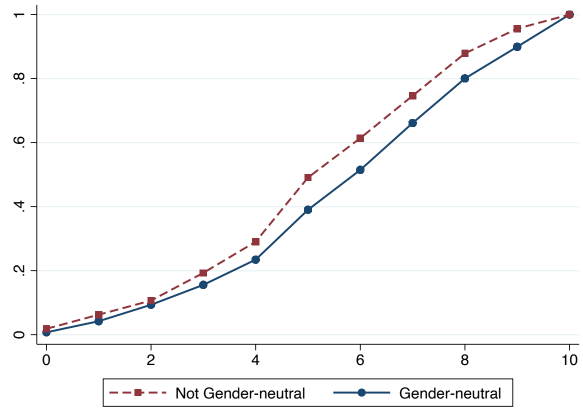
(f) Female Applicants,
Top Quartile of % Neighbors Treated

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).

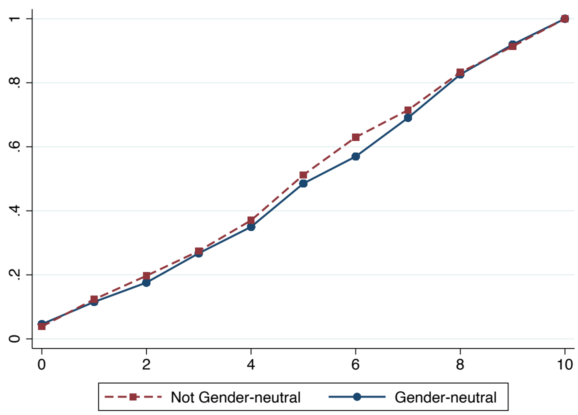
Figure A.7: Outcomes Distribution in Laboratoria Experiment,
by Gender-Neutral Treatment Status



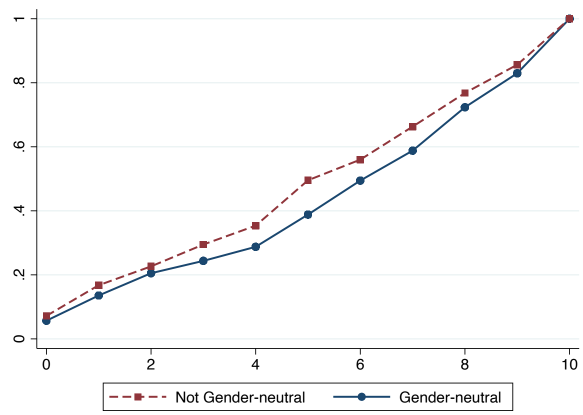
(a) Job Appeal



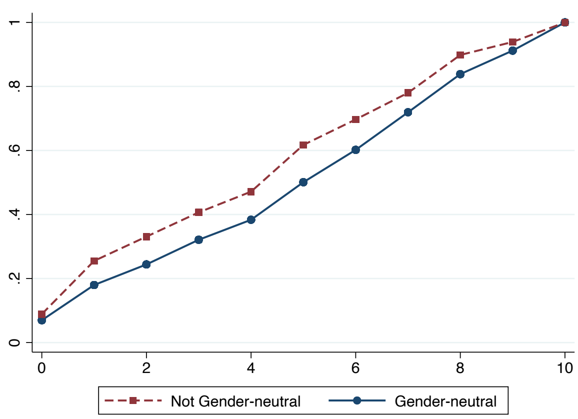
(b) Good Employer



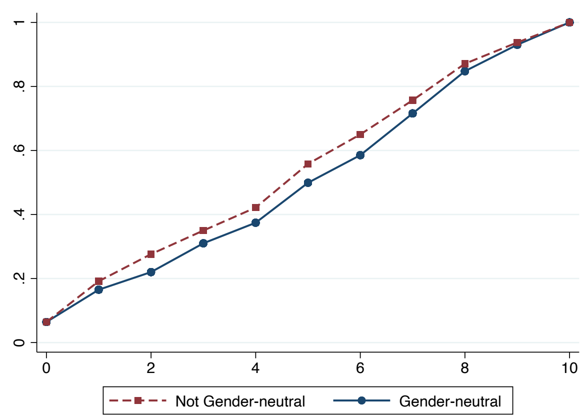
(c) Meet Requirements



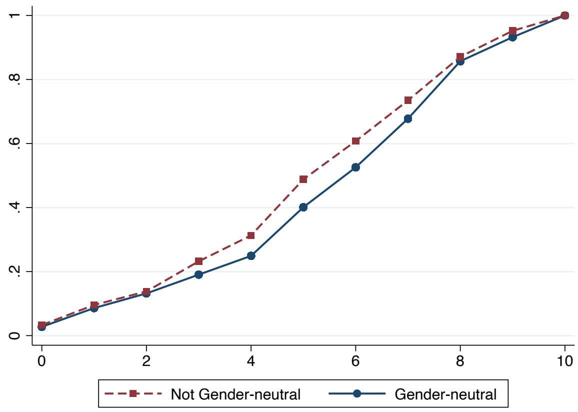
(d) Probability of Applying



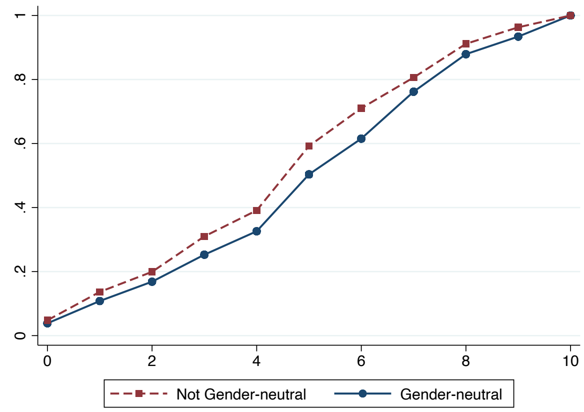
(e) Suitability



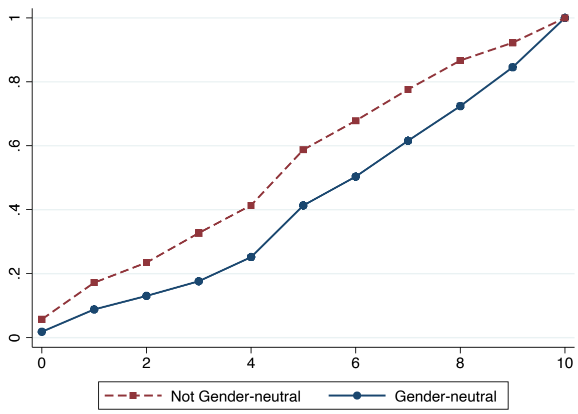
(f) Probability of Being Chosen



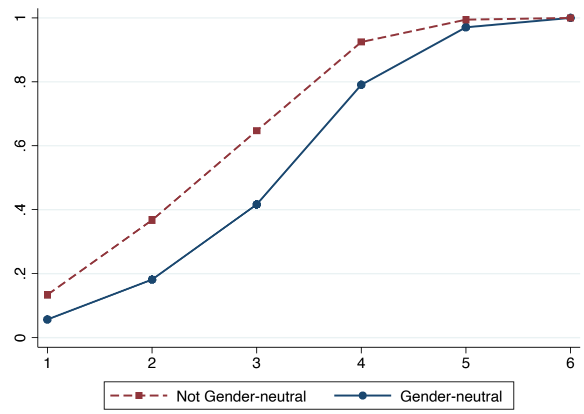
(g) Good salary



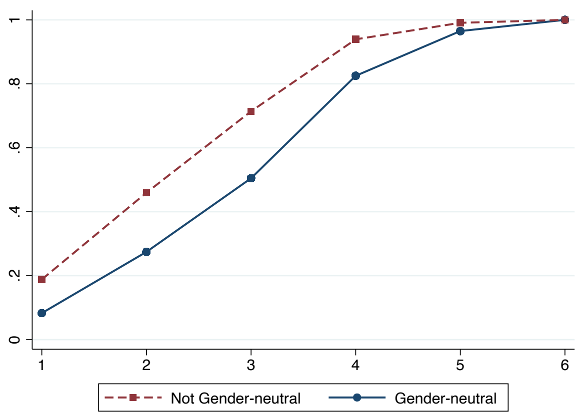
(h) Work-life Balance



(i) Inclusive Culture



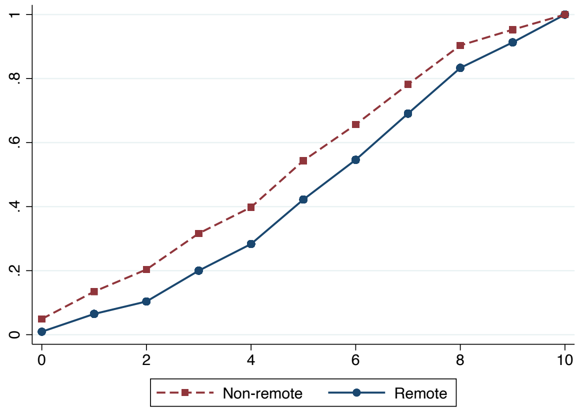
(j) Women Percentage Company



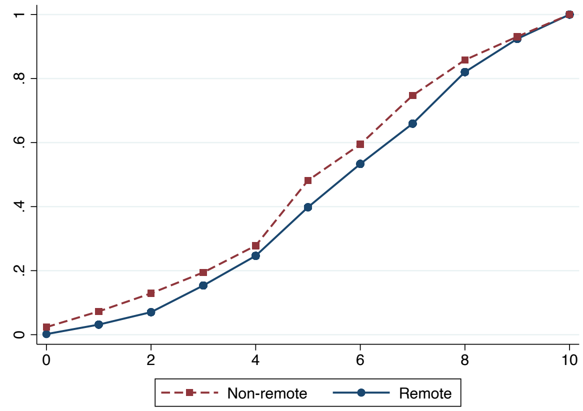
(k) Women Percentage Position

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).

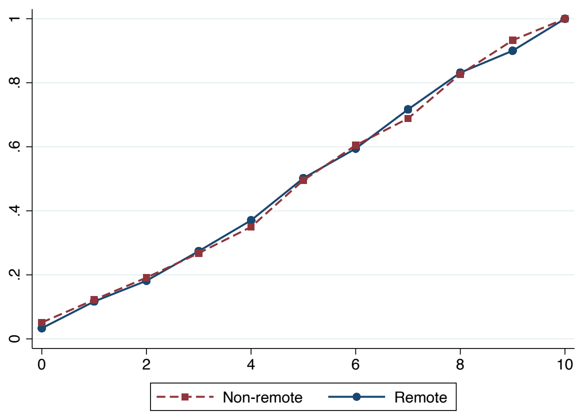
Figure A.8: Outcomes Distribution in Laboratoria Experiment, by Remote Treatment Status



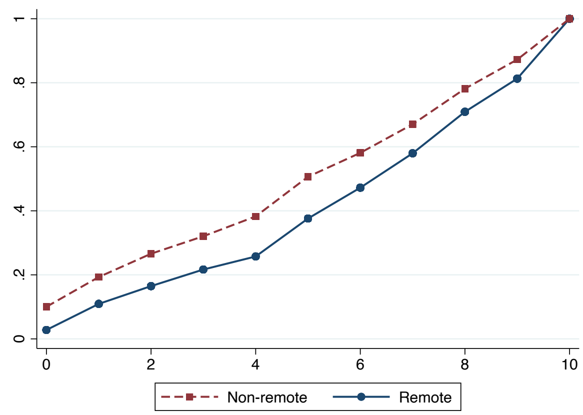
(a) Job Appeal



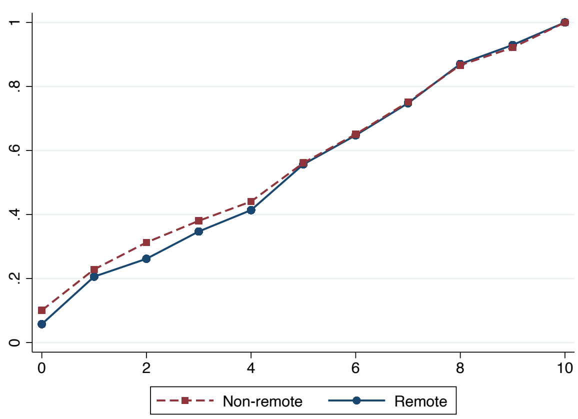
(b) Good Employer



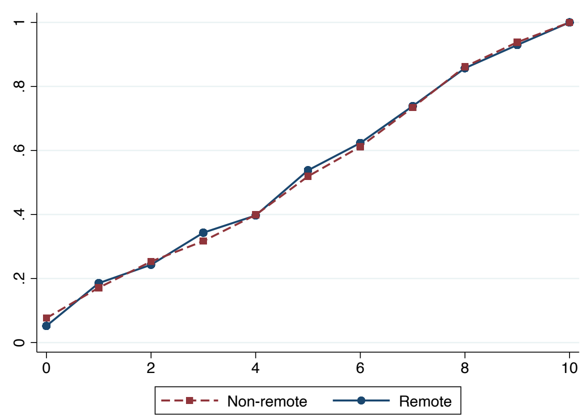
(c) Meet Requirements



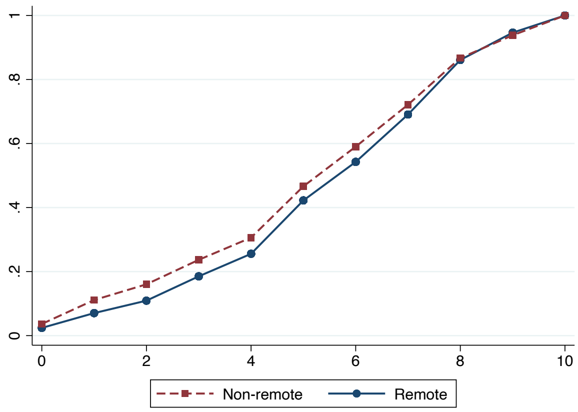
(d) Probability of Applying



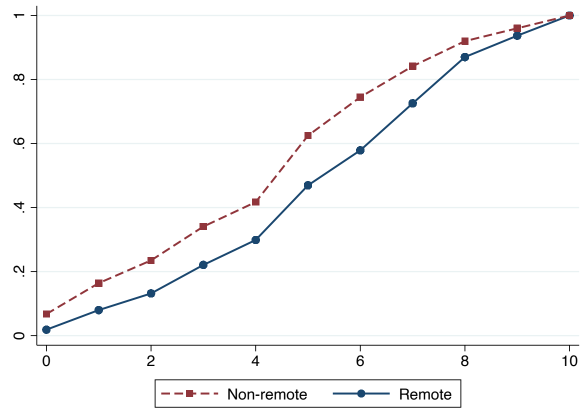
(e) Suitability



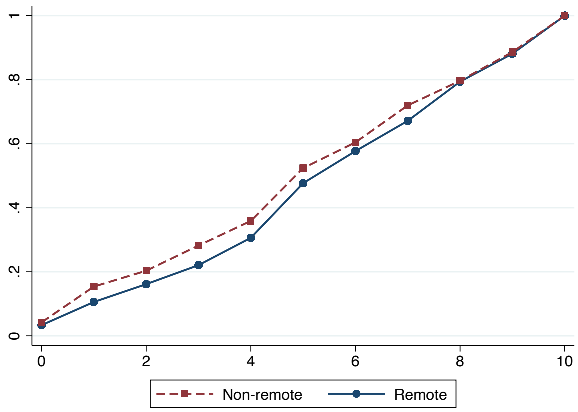
(f) Probability of Being Chosen



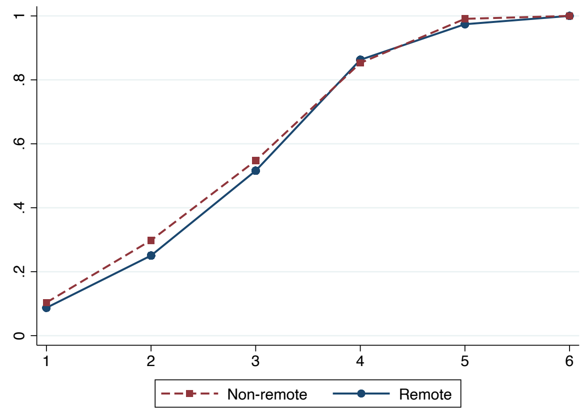
(g) Good salary



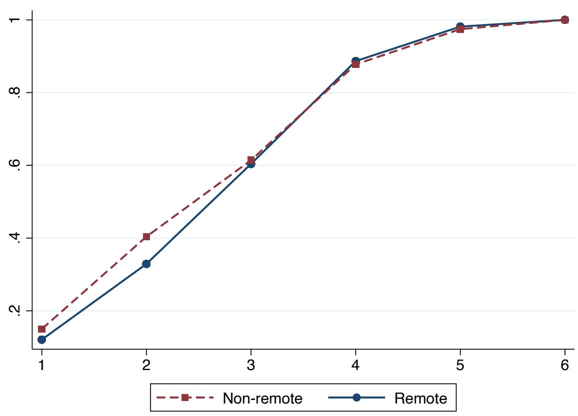
(h) Work-life Balance



(i) Inclusive Culture



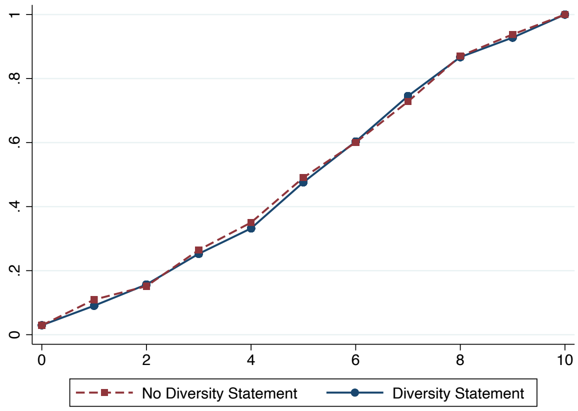
(j) Women Percentage Company



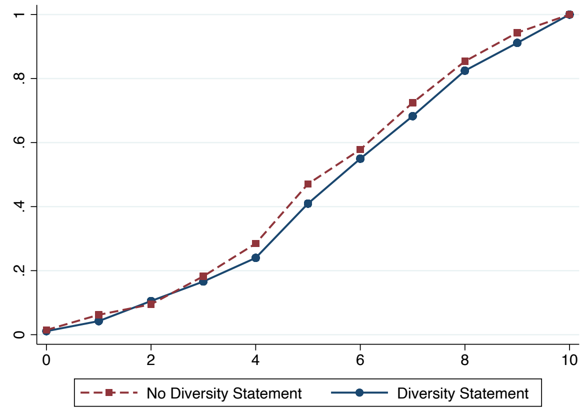
(k) Women Percentage Position

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).

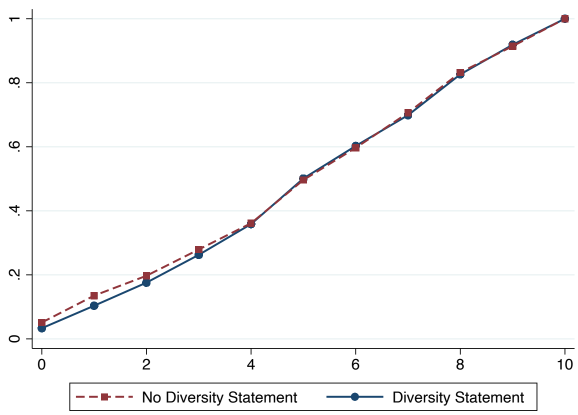
Figure A.9: Outcomes Distribution in Laboratoria Experiment,
by Diversity Statement Treatment Status



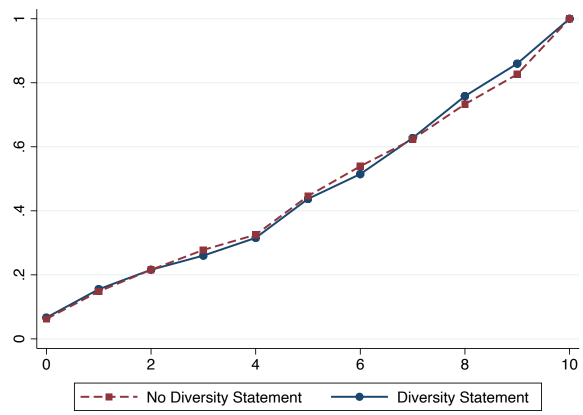
(a) Job Appeal



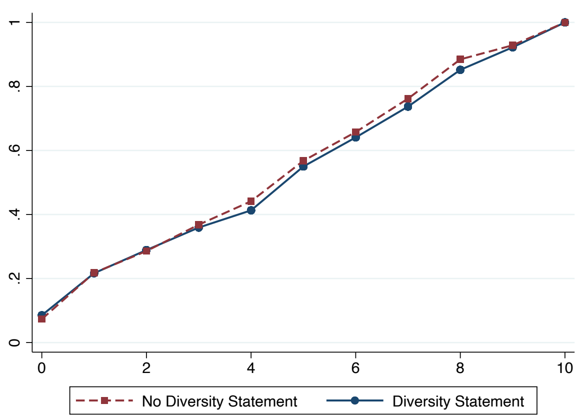
(b) Good Employer



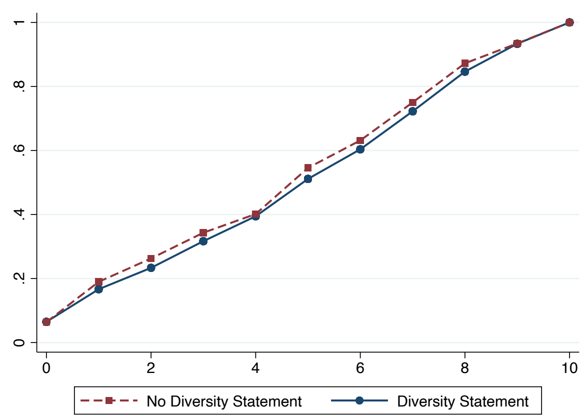
(c) Meet Requirements



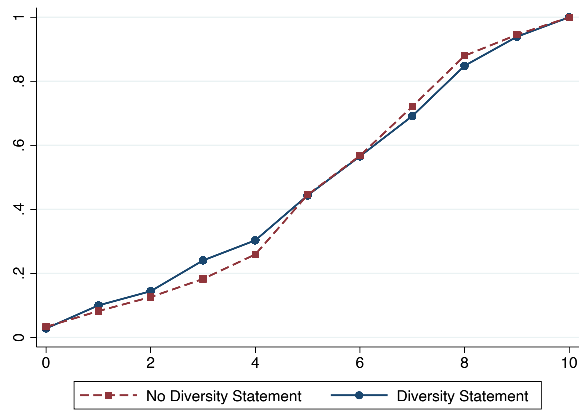
(d) Probability of Applying



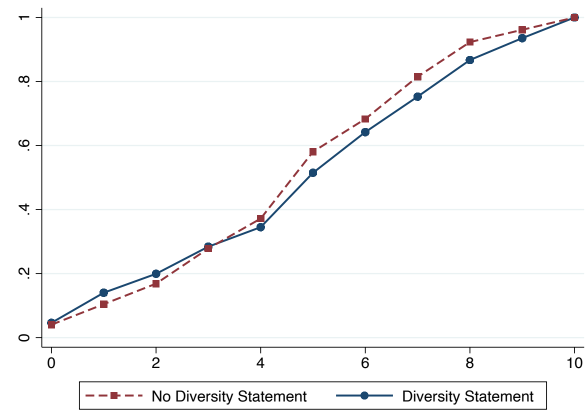
(e) Suitability



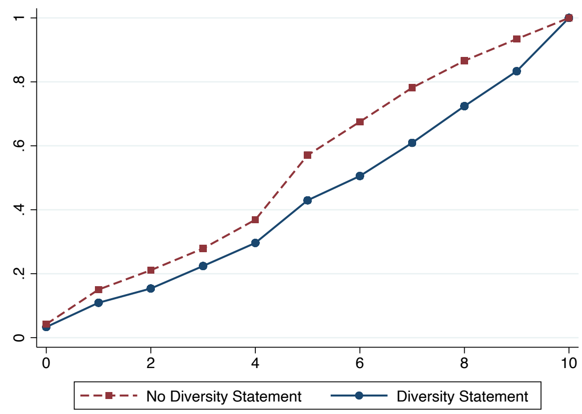
(f) Probability of Being Chosen



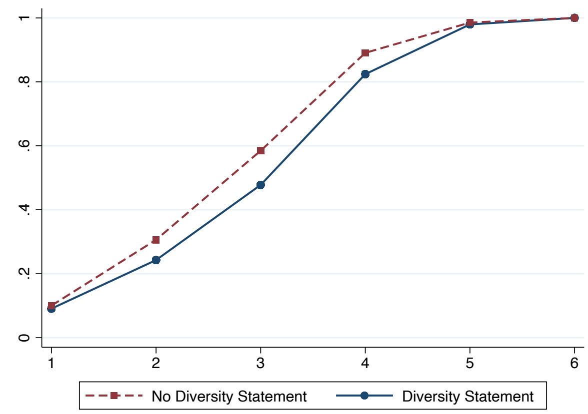
(g) Good salary



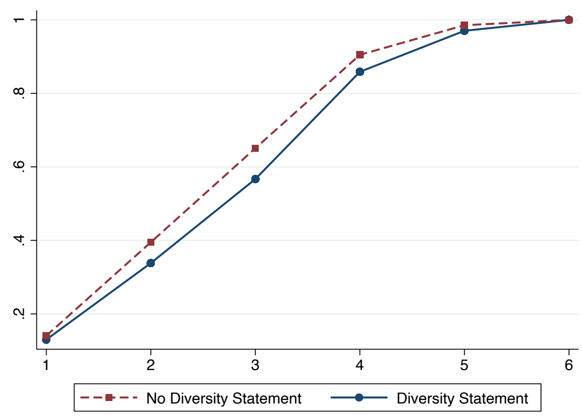
(h) Work-life Balance



(i) Inclusive Culture



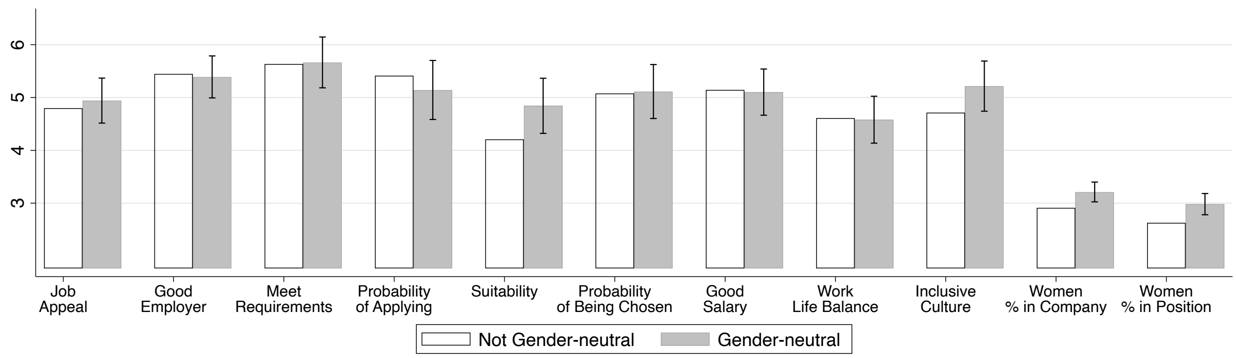
(j) Women Percentage Company



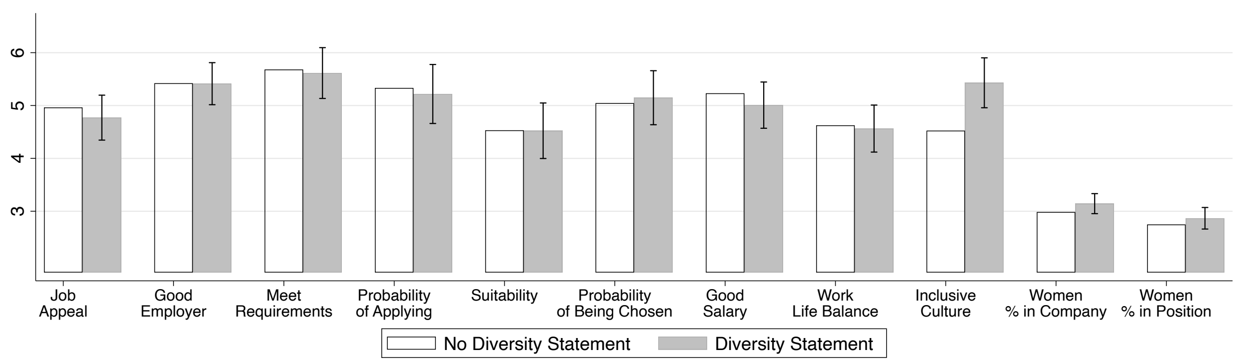
(k) Women Percentage Position

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).

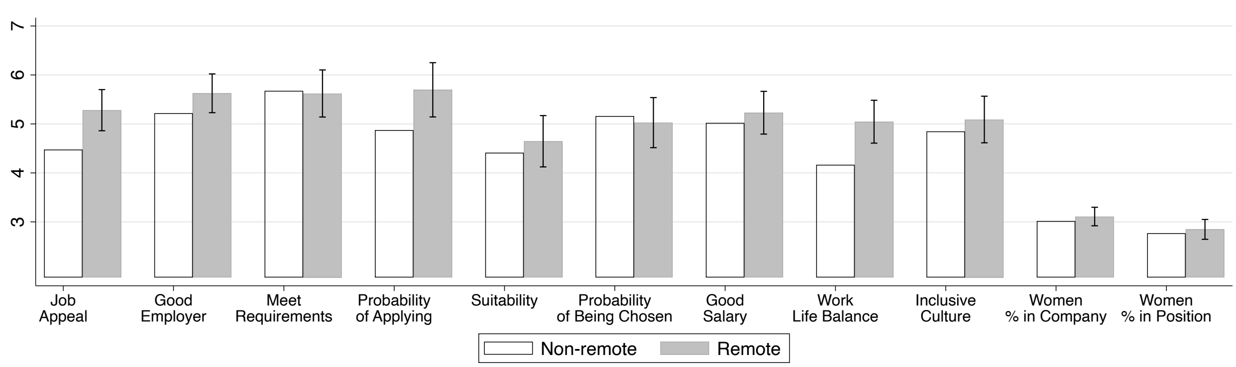
Figure A.10: Outcome Averages by Different Treatment Statuses - Laboratoria (First Ads Only)



(a) Gender Neutral Language Treatment



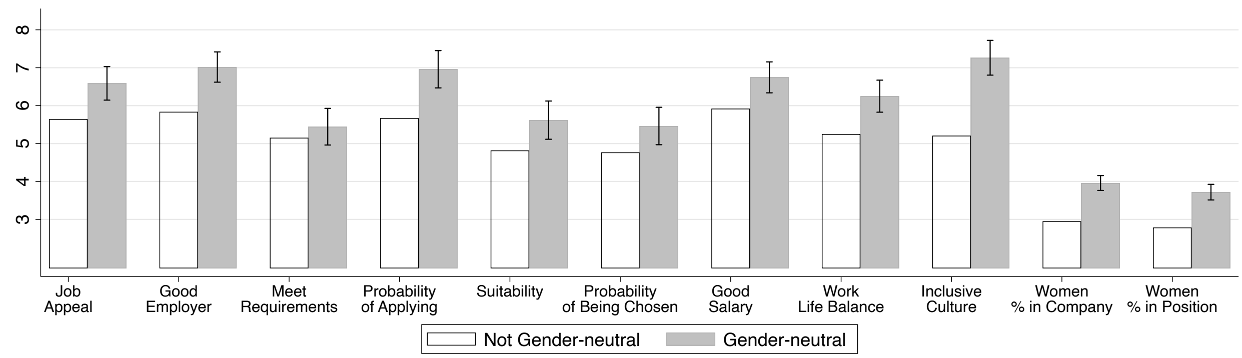
(b) Diversity Statement Treatment



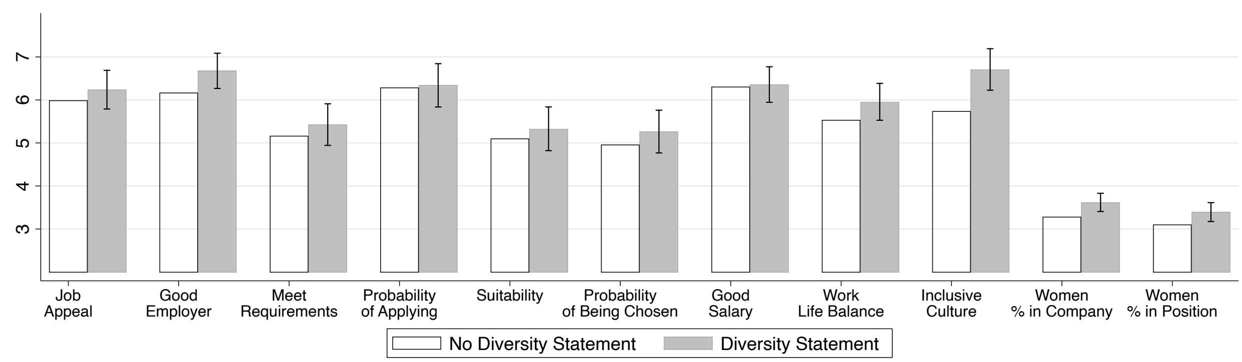
(c) Remote Job Treatment

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).

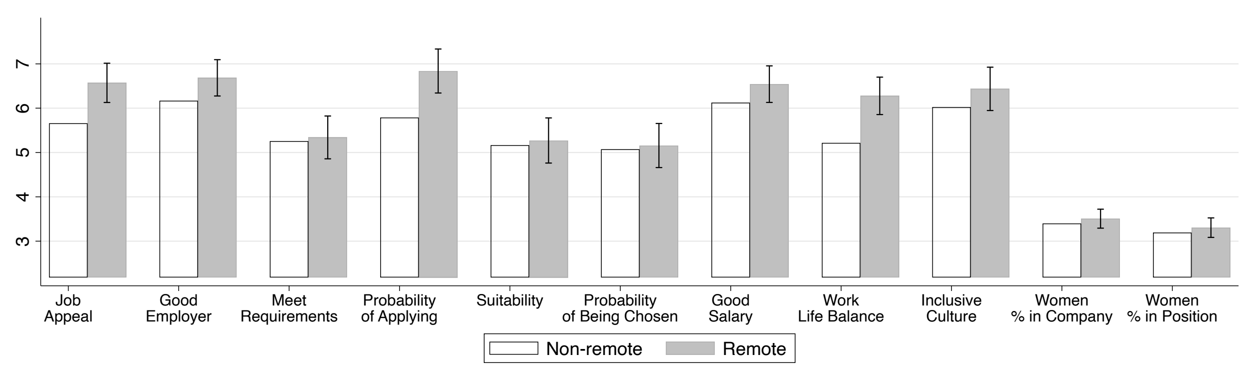
Figure A.11: Outcome Averages by Different Treatment Statuses - Laboratoria (Only Second Ads)



(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: 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.2: 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. The Kerwin et al. (2024) omnibus test of overall covariate balance across yields a p-value of 0.338 (see Appendix D). See text for further variable definitions.

Table A.3: Treatment Balance - Job Title Groups

Variable	Mean (C)	Mean (T)	Difference (T-C)	SE	p-value	Obs
Programmer	0.018	0.018	0	0.006	0.994	2,201
Other Developper	0.102	0.129	0.027	0.014	0.047	2,201
Designer	0.090	0.082	-0.008	0.012	0.499	2,201
Engineer	0.162	0.182	0.020	0.016	0.212	2,201
Analyst	0.063	0.072	0.009	0.011	0.397	2,201
Web Developer	0.020	0.021	0.001	0.006	0.854	2,201
Front Developer	0.087	0.077	-0.009	0.012	0.430	2,201
Back Developer	0.073	0.076	0.003	0.011	0.784	2,201
Mobile Developer	0.066	0.037	-0.029	0.009	0.002	2,201
Full-Stack Developer	0.155	0.148	-0.006	0.015	0.675	2,201
Sysadmin	0.057	0.063	0.006	0.010	0.558	2,201
Bizadmin	0.060	0.055	-0.005	0.010	0.609	2,201
Marketing/Customers	0.028	0.024	-0.004	0.007	0.553	2,201
Architect	0.006	0.005	-0.002	0.003	0.626	2,201
Data Scientist	0.006	0.006	-0.001	0.003	0.856	2,201
Scrum	0.007	0.005	-0.002	0.003	0.458	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). The variable on each row is a dummy equal to one if the ad's job title group is the one listed in the first column. The Kerwin et al. (2024) omnibus test of overall balance across all listed variables yields a p-value of 0.286 (see Appendix D). See main text and Appendix C for definitions and construction of job title groups.

Table A.4: 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.5: Share of Neighbor Ads Treated is Uncorrelated with Job Group Titles

Variable	Coeff	SE	p-value
Programmer	0.021	0.132	0.877
Other Developper	0.090	0.117	0.443
Designer	0.000	0.118	1.000
Engineer	0.056	0.117	0.632
Analyst	0.049	0.118	0.682
Web Developer	0.026	0.127	0.837
Front Developer	0.004	0.118	0.975
Back Developer	0.031	0.118	0.793
Mobile Developer	-0.128	0.120	0.287
Full-Stack Developer	0.017	0.117	0.882
Sysadmin	0.059	0.119	0.621
Bizadmin	-0.002	0.119	0.985
Marketing/Customers	-0.003	0.128	0.978
Architect	-0.045	0.150	0.764
Scrum	-0.077	0.170	0.651

Notes: The unit of observation is an ad. Each row provides the coefficient, standard error, and p-value of the coefficients of single regression with the share of neighbor ads treated (SNT_i) as the independent variable and a set of job group title dummy indicators as the explanatory variables. The "data scientist" job title group is the omitted dummy (due to collinearity). It has an average SNT_i of 0.462. The regression has 2,201 observations. The Kerwin et al. (2024) omnibus test of overall balance across all 16 job group titles yields a p-value of 0.215 (see Appendix D).

Table A.6: Twenty Most Frequently Edited Words due To Treatment

Word	Total Count in Control Minus Total Count in Treatment
Desarrollador	332
Un/El/Lo	152
Ingeniero	148
Diseñador	81
Ello/Al	56
Programador	56
Experto	40
Proactivo	31
Informático	27
Titulado	25
Comprometido	24
Candidato	17
Interesado	15
Familiarizado	14
Gráfico	13
Creativo	13
Extranjero	11
Enfocado	11
Administrador	11
Consultor	10

Notes: This table lists the 20 most frequently edited words due to the treatment. For each gendered word w appearing in the ads' texts, we compute N_w^{control} and N_w^{treat}: the total number of times w appears across all ads assigned to control (1,030 ads) and treatment (1,071 ads), respectively. The table reports $N_w^{\text{control}} - N_w^{\text{treat}}$ for the twenty words with the highest values, which serves as a proxy for how often each word was edited in response to treatment. We combine certain articles ("un/el/ello") and pronouns ("ello/al") and treat them as a single word.

Table A.7: Treatment-on-Treated (2SLS) Effects of Gender-Neutrality - Get on Board

	(1) Fem. Share Applicants	(2) Fem. Share Applicants
Bottom Quartile of % Neighbors Treated	0.088**	0.093**
	(0.037)	(0.036)
Mid Quartiles of % Neighbors Treated	-0.041*	-0.043*
	(0.024)	(0.024)
Top Quartile of % Neighbors Treated	-0.022	-0.014
	(0.034)	(0.034)
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 who are female. The table presents the linear combinations that provide the treatment-on-treated effects of an ad's title and job description section being gender-neutral (i.e., containing zero gendered words) 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 (3) 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.8 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 A.8: 2SLS and First-Stage Estimates for Treatment-on-Treated Effects - Get on Board

	(1) Fem. Share Applicants	(2) Fem. Share Applicants	(3) Gender Neutral (GN)	(4) Gender Neutral (GN)	(5) GN × Mid Quart. of % Neighbors Treated	(6) GN × Mid Quart. of % Neighbors Treated	(7) GN × Top Quart. of % Neighbors Treated	(8) GN × Top Quart. of % Neighbors Treated
Gender Neutral	0.088**	0.093**						
(GN)	(0.037)	(0.036)						
GN × Mid	-0.129***	-0.135***						
Quartiles of % Neighbors Treated	(0.044)	(0.044)						
GN × Top	-0.110**	-0.107**						
Quartile of % Neighbors Treated	(0.050)	(0.049)						
Treatment			0.421***	0.416***	0.001	-0.001	0.000	-0.001
			(0.036)	(0.036)	(0.004)	(0.002)	(0.002)	(0.002)
Treat × Mid			-0.026	-0.019	0.393***	0.398***	-0.001	0.001
Quartiles of % Neighbors Treated			(0.045)	(0.045)	(0.027)	(0.026)	(0.003)	(0.002)
Treat × Top			-0.033	-0.029	-0.006	-0.004	0.396***	0.396***
Quartile of % Neighbors Treated			(0.052)	(0.052)	(0.005)	(0.003)	(0.038)	(0.038)
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); even-numbered columns include controls selected by PDS-LASSO. Columns 1-2 present the results from the 2SLS estimation of equation (3). 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 A.7 are based on these estimated coefficients. Columns 3-8 provide the related first-stage estimates for the three endogenous variables: a dummy equal to one if the ad has no gendered word in its title and job description section, 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.9: 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) Fem. Share Applicants	(2) Fem. Share Applicants	(3) Fem. Share Applicants	(4) Fem. Share Applicants
Treatment (β_0)	0.008	0.007	0.009	0.008
	(0.015)	(0.015)	(0.016)	(0.016)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.013	-0.012	-0.010	-0.008
	(0.018)	(0.018)	(0.020)	(0.020)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.015	-0.010	0.000	0.001
	(0.023)	(0.023)	(0.023)	(0.022)
Mid. Quartiles of % Neighbors Treated (γ_M)	-0.034***	-0.029**	-0.010	0.003
	(0.013)	(0.013)	(0.014)	(0.013)
Top Quartile of % Neighbors Treated (γ_T)	-0.016	-0.008	-0.011	-0.015
	(0.016)	(0.016)	(0.015)	(0.015)
Implied Treatment Effects				
Bottom Quartile of % Neighbors Treated (β_0)	0.008	0.007	0.009	0.008
	(0.015)	(0.015)	(0.016)	(0.016)
Mid. Quartiles of % Neighbors Treated ($\beta_0+\beta_M$)	-0.005	-0.004	-0.001	0.001
	(0.009)	(0.009)	(0.011)	(0.011)
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	-0.006	-0.003	0.009	0.010
	(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); even-numbered columns include controls selected by PDS-LASSO. The outcome (dependent variable) in all columns is the share of applicants who 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. Reported 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. 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.10: Intent-to-Treat Effects Using *Fields* to Define Neighbor Ads - 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)	0.026*	0.026*	0.211	0.214*	0.056	0.077	0.025	0.027
	(0.015)	(0.014)	(0.131)	(0.127)	(0.103)	(0.101)	(0.050)	(0.049)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.038**	-0.035**	-0.274*	-0.283*	-0.085	-0.097	0.060	0.056
	(0.016)	(0.016)	(0.150)	(0.146)	(0.122)	(0.120)	(0.063)	(0.062)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.033	-0.031	-0.247	-0.257	-0.014	-0.060	-0.002	0.008
	(0.020)	(0.020)	(0.181)	(0.177)	(0.143)	(0.142)	(0.071)	(0.070)
Mid. Quartiles of % Neighbors Treated (γ_M)	0.020	0.015	0.089	0.067	0.024	0.033	-0.038	-0.023
	(0.012)	(0.012)	(0.113)	(0.108)	(0.093)	(0.091)	(0.046)	(0.045)
Top Quartile of % Neighbors Treated (γ_T)	0.022	0.017	0.147	0.121	0.019	0.032	0.033	0.041
	(0.015)	(0.014)	(0.129)	(0.126)	(0.104)	(0.102)	(0.050)	(0.049)
Implied Treatment Effects								
Bottom Quartile of % Neighbors Treated (β_0)	0.026	0.026	0.211	0.214	0.056	0.077	0.025	0.027
	(0.015)	(0.014)	(0.131)	(0.127)	(0.103)	(0.101)	(0.050)	(0.049)
	[0.163]	[0.166]	[0.183]	[0.175]	[0.595]	[0.476]	[0.601]	[0.580]
Mid. Quartiles of % Neighbors Treated ($\beta_0+\beta_M$)	-0.012	-0.010	-0.063	-0.068	-0.028	-0.021	0.086	0.082
	(0.008)	(0.007)	(0.074)	(0.071)	(0.067)	(0.066)	(0.039)	(0.039)
	[0.756]	[0.794]	[0.843]	[0.824]	[0.858]	[0.910]	[0.305]	[0.327]
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	-0.008	-0.005	-0.036	-0.043	0.043	0.016	0.023	0.035
	(0.014)	(0.014)	(0.125)	(0.123)	(0.100)	(0.100)	(0.050)	(0.049)
	[0.279]	[0.439]	[0.577]	[0.496]	[0.462]	[0.754]	[0.474]	[0.287]
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.144	0.144	1.745	1.745	3.713	3.713	15.123	15.123
N	2,172	2,172	2,172	2,172	2,172	2,172	2,172	2,172

Notes: This table replicates the main ITT results from Table 2, but using *fields* instead of *job title groups* to define neighbors. The top panel thus provides the estimated coefficients from equation (1) using SNT_i^{field} instead of SNT_i (see text for details). The unit of observation is an ad. The number of observations differs from Table 2 since ads with no neighbors defined at the field level must be dropped. Odd-numbered columns include baseline controls (month dummies interacted with remote status); 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). Reported independent variables are a dummy for treatment assignment, two dummies indicating if the ad's SNT_i^{field} falls in the middle quartiles or the top quartile of the SNT_i^{field} distribution, and their interactions with treatment. All regressions include Prob[$MidQuartiles_i^{SNT}=1$] and Prob[$TopQuartile_i^{SNT}=1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with SNT_i^{field} 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Intent-to-Treat Effects by Title's Language and Remote Status - Get on Board

	Job Title in English		Job Title in Spanish	
	(1) Fem. Share Applicants	(2) Fem. Share Applicants	(3) Fem. Share Applicants	(4) Fem. Share Applicants
Treatment (β_0)	0.037*	0.045**	0.036*	0.035*
	(0.022)	(0.021)	(0.021)	(0.021)
Treat × Mid. Quartiles of	-0.042	-0.051**	-0.063**	-0.065***
% Neighbors Treated (β_M)	(0.025)	(0.025)	(0.026)	(0.025)
Treat × Top Quartile of	-0.042	-0.042	-0.049*	-0.048*
% Neighbors Treated (β_T)	(0.030)	(0.030)	(0.027)	(0.027)
Mid. Quartiles of	0.014	0.020	0.025	0.030
% Neighbors Treated (γ_M)	(0.018)	(0.018)	(0.019)	(0.019)
Top Quartile of	0.009	0.010	-0.002	-0.002
% Neighbors Treated (γ_T)	(0.020)	(0.019)	(0.020)	(0.020)
Implied Treatment Effects				
Bottom Quartile of %	0.037	0.045	0.036	0.035
Neighbors Treated (β_0)	(0.022)	(0.021)	(0.021)	(0.021)
	[0.112]	[0.048]**	[0.107]	[0.116]
Mid. Quartiles of %	-0.004	-0.006	-0.027	-0.030
Neighbors Treated (β_0+β_M)	(0.013)	(0.013)	(0.014)	(0.014)
	[0.767]	[0.671]	[0.075]*	[0.049]**
Top Quartile of %	-0.004	0.003	-0.014	-0.013
Neighbors Treated ($\beta_0 + \beta_T$)	(0.021)	(0.021)	(0.017)	(0.017)
	[0.847]	[0.880]	[0.523]	[0.530]
Baseline Controls?	YES		YES	
PDS-LASSO Controls?		YES		YES
Control Mean	0.145	0.145	0.146	0.146
N	1,072	1,072	1,129	1,129

Notes: This table replicates the main ITT results from Table 2, but for separate subsamples. The unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status); even-numbered columns include controls selected by PDS-LASSO. The outcome (dependent variable) in all columns is the share of applicants who are female. Columns 1-2 only use ads with titles in English, while columns 3-4 only use ads with titles in Spanish. Reported 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. 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. p-values from a randomization inference procedure that accounts for dependencies induced by spillovers are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Effects by Share Female Applicants in Job Title Group - 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)	-0.003	-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: This table replicates the main ITT results from Table 2, but but exploring effect heterogeneity by share of female applicants in job title group (instead of SNT_i). The unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status); 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. 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 C). 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-values from a randomization inference procedure that accounts for dependencies induced by spillovers are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Intent-to-Treat Effects by Female Representation in Job Title - Get on Board

	Low % Fem. in Title Group		High % Fem. in Title Group	
	(1)	(2)	(3)	(4)
	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants	Fem. Share Applicants
Treatment (β_0)	0.000	0.000	0.029	0.033
	(0.012)	(0.011)	(0.021)	(0.020)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.001	-0.001	-0.051**	-0.055**
	(0.015)	(0.014)	(0.025)	(0.024)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.010	-0.008	-0.027	-0.024
	(0.017)	(0.017)	(0.028)	(0.027)
Mid. Quartiles of % Neighbors Treated (γ_M)	0.013	0.013	-0.005	-0.001
	(0.011)	(0.010)	(0.019)	(0.018)
Top Quartile of % Neighbors Treated (γ_T)	0.012	0.013	-0.036*	-0.034*
	(0.014)	(0.014)	(0.019)	(0.019)
Implied Treatment Effects				
Bottom Quartile of % Neighbors Treated (β_0)	0.000	0.000	0.029	0.033
	(0.012)	(0.011)	(0.021)	(0.020)
	[0.972]	[0.980]	[0.186]	[0.141]
Mid. Quartiles of % Neighbors Treated (β_0+β_M)	-0.001	-0.000	-0.021	-0.022
	(0.009)	(0.009)	(0.013)	(0.013)
	[0.908]	[0.955]	[0.196]	[0.167]
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	-0.009	-0.008	0.003	0.009
	(0.013)	(0.013)	(0.018)	(0.018)
	[0.473]	[0.537]	[0.901]	[0.688]
Baseline Controls?	YES		YES	
PDS-LASSO Controls?		YES		YES
Control Mean	0.060	0.060	0.212	0.212
N	950	950	1,251	1,251

Notes: This table replicates the main ITT results from Table 2, but for separate subsamples. The unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status); 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 include ads with a share of female applicants in the job title group below the sample median, while columns 3-4 only include ads with a share of female applicants in the job title group equal or above the sample median. Reported 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. 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 share of female applicants in the job title group is constructed only using the control group observations (see Appendix C). 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-values from a randomization inference procedure that accounts for dependencies induced by spillovers are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: Intent-to-Treat Effects Controlling for Number of Neighbor Ads - 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)	0.035**	0.039***	0.181	0.175	-0.079	-0.090	0.031	0.022
	(0.015)	(0.015)	(0.133)	(0.133)	(0.101)	(0.101)	(0.049)	(0.048)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.051***	-0.055***	-0.216	-0.216	0.153	0.150	-0.017	-0.002
	(0.018)	(0.018)	(0.158)	(0.157)	(0.121)	(0.121)	(0.062)	(0.061)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.043**	-0.044**	-0.200	-0.188	0.066	0.076	0.134*	0.134*
	(0.020)	(0.020)	(0.182)	(0.182)	(0.141)	(0.141)	(0.072)	(0.071)
Mid. Quartiles of % Neighbors Treated (γ_M)	0.024*	0.024*	0.103	0.093	-0.069	-0.073	0.020	0.010
	(0.013)	(0.013)	(0.121)	(0.120)	(0.090)	(0.089)	(0.046)	(0.046)
Top Quartile of % Neighbors Treated (γ_T)	0.003	0.003	0.009	0.005	-0.031	-0.058	-0.050	-0.054
	(0.014)	(0.014)	(0.130)	(0.129)	(0.100)	(0.099)	(0.050)	(0.049)
Implied Treatment Effects								
Bottom Quartile of % Neighbors Treated (β_0)	0.035	0.039	0.181	0.175	-0.079	-0.090	0.031	0.022
	(0.015)	(0.015)	(0.133)	(0.133)	(0.101)	(0.101)	(0.049)	(0.048)
	[0.040]**	[0.022]**	[0.237]	[0.261]	[0.478]	[0.413]	[0.560]	[0.701]
Mid. Quartiles of % Neighbors Treated ($\beta_0+\beta_M$)	-0.015	-0.016	-0.036	-0.041	0.074	0.061	0.014	0.020
	(0.009)	(0.009)	(0.085)	(0.083)	(0.067)	(0.067)	(0.038)	(0.038)
	[0.177]	[0.162]	[0.725]	[0.679]	[0.297]	[0.392]	[0.708]	[0.622]
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	-0.008	-0.005	-0.019	-0.013	-0.013	-0.014	0.165	0.156
	(0.013)	(0.013)	(0.124)	(0.124)	(0.099)	(0.098)	(0.052)	(0.052)
	[0.662]	[0.757]	[0.891]	[0.932]	[0.897]	[0.892]	[0.002]***	[0.003]***
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.146	0.146	1.764	1.764	3.718	3.718	15.121	15.121
N	2,201	2,201	2,201	2,201	2,201	2,201	2,201	2,201

Notes: The table replicates Table 2, but includes the number of neighbor ads as an additional control in all columns. Odd-numbered columns include baseline controls (month dummies interacted with remote status); 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). Reported 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. The bottom panel presents the linear combinations that provide the estimated treatment effects for ads with SNT_i in the bottom quartile, the middle quartiles, and the 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. p-values from a randomization inference procedure that accounts for dependencies induced by spillovers are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: ITT Effects: Weighted Regressions, Dropping Ads with Full-Text in English, and Heterogeneity by Remote Status - Get on Board

	Using Weights		Dropping Ads in English		Remote Job		Non-remote Job	
	(1) Fem. Share Applicants	(2) Fem. Share Applicants	(3) Fem. Share Applicants	(4) Fem. Share Applicants	(5) Fem. Share Applicants	(6) Fem. Share Applicants	(7) Fem. Share Applicants	(8) Fem. Share Applicants
Treatment (β_0)	0.035**	0.038**	0.038**	0.040**	0.039	0.043*	0.034*	0.035*
	(0.015)	(0.015)	(0.016)	(0.016)	(0.024)	(0.024)	(0.020)	(0.019)
Treat × Mid. Quartiles of % Neighbors Treated (β_M)	-0.053***	-0.057***	-0.050***	-0.054***	-0.077***	-0.080***	-0.036	-0.039*
	(0.018)	(0.018)	(0.019)	(0.019)	(0.028)	(0.028)	(0.023)	(0.023)
Treat × Top Quartile of % Neighbors Treated (β_T)	-0.036*	-0.035*	-0.050**	-0.048**	-0.037	-0.037	-0.054**	-0.051**
	(0.020)	(0.020)	(0.021)	(0.021)	(0.033)	(0.033)	(0.025)	(0.025)
Mid. Quartiles of % Neighbors Treated (γ_M)	0.018	0.023*	0.021	0.025*	0.030	0.030	0.015	0.019
	(0.014)	(0.014)	(0.014)	(0.014)	(0.020)	(0.020)	(0.018)	(0.017)
Top Quartile of % Neighbors Treated (γ_T)	-0.006	-0.005	0.010	0.011	0.006	0.006	-0.000	-0.000
	(0.014)	(0.014)	(0.015)	(0.015)	(0.023)	(0.023)	(0.018)	(0.018)
Implied Treatment Effects								
Bottom Quartile of % Neighbors Treated (β_0)	0.035	0.038	0.038	0.040	0.039	0.043	0.034	0.035
	(0.015)	(0.015)	(0.016)	(0.016)	(0.024)	(0.024)	(0.020)	(0.019)
	[0.056]*	[0.030]**	[0.032]**	[0.024]**	[0.143]	[0.108]	[0.092]*	[0.086]*
Mid. Quartiles of % Neighbors Treated (β_0+β_M)	-0.017	-0.018	-0.013	-0.014	-0.038	-0.036	-0.002	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.012)	(0.012)
	[0.155]	[0.124]	[0.295]	[0.255]	[0.018]**	[0.027]**	[0.896]	[0.792]
Top Quartile of % Neighbors Treated ($\beta_0+\beta_T$)	-0.001	0.003	-0.013	-0.008	0.003	0.007	-0.019	-0.016
	(0.013)	(0.013)	(0.014)	(0.014)	(0.022)	(0.022)	(0.016)	(0.016)
	[0.964]	[0.841]	[0.495]	[0.636]	[0.923]	[0.783]	[0.358]	[0.424]
Baseline Controls?	YES		YES		YES		YES	
PDS-LASSO Controls?		YES		YES		YES		YES
Control Mean	0.146	0.146	0.146	0.146	0.147	0.147	0.145	0.145
N	2,201	2,201	1,935	1,935	885	885	1,316	1,316

Notes: This table replicates the main ITT results from Table 2, but for separate subsamples. The unit of observation is an ad. Odd-numbered columns include baseline controls (month dummies interacted with remote status); even-numbered columns include controls selected by PDS-LASSO. Outcomes are the share of applicants that are female. Columns 1-2 weighted each observation (ad) by the number of applications it received. Columns 3-4 exclude from the sample ads that were entirely (both title and its text) in English. Columns 5-6 only use ads for remote positions, and columns 7-8 only include non-remote positions. The top panel provides the estimated coefficients from equation (1). Reported 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. All regressions include Prob[$MidQuartiles_i^{SNT} = 1$] and Prob[$TopQuartile_i^{SNT} = 1$] as controls, thus implementing the Borusyak and Hull (2023) recentered estimator. 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. p-values from a randomization inference procedure that accounts for dependencies induced by spillovers are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.16: 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	-0.027	-0.019	-0.049
	(0.037)	(0.296)	(0.077)	(0.077)	(0.080)	(0.072)
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	0.498
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 to 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 no gender-neutral words in its title, and in 5-6, it is a dummy equal to one if the ad has no gendered words in its title and job description and requirements section. See Appendix D for further information. Standard errors clustered at the firm level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17: 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.18: Treatment Effects (Alumnae 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 alumnae 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.19: Treatment Effects (Alumnae 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 alumnae 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.20: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$..

Table A.21: “Long” or “Fully Saturated” Model - 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	0.700**	0.837***	0.101	0.659	0.684*	0.317	0.563*	0.584**	1.450***	0.765***	0.776***
	(0.336)	(0.299)	(0.362)	(0.402)	(0.377)	(0.372)	(0.315)	(0.297)	(0.330)	(0.134)	(0.145)
Remote	0.660**	0.560*	0.025	0.639	0.107	-0.092	0.279	1.025***	0.048	0.178	0.113
	(0.309)	(0.288)	(0.357)	(0.392)	(0.362)	(0.359)	(0.300)	(0.292)	(0.313)	(0.134)	(0.136)
Diversity Statement	-0.114	0.445	0.221	-0.228	-0.001	0.178	-0.068	0.109	1.034***	0.360**	0.253*
	(0.332)	(0.308)	(0.345)	(0.401)	(0.379)	(0.366)	(0.326)	(0.319)	(0.351)	(0.144)	(0.153)
Gender-neutral × Remote	-0.074	-0.245	0.249	-0.041	0.264	0.274	-0.142	-0.093	0.377	-0.069	-0.147
	(0.455)	(0.412)	(0.501)	(0.550)	(0.517)	(0.511)	(0.429)	(0.421)	(0.456)	(0.193)	(0.203)
Gender-neutral × Diversity	-0.127	-0.408	0.014	-0.178	0.381	0.186	-0.204	0.210	-0.353	-0.132	-0.199
	(0.463)	(0.428)	(0.493)	(0.563)	(0.536)	(0.512)	(0.459)	(0.452)	(0.497)	(0.200)	(0.218)
Remote × Diversity	0.632	-0.016	-0.133	0.758	0.480	0.236	0.244	0.355	0.628	-0.047	0.050
	(0.446)	(0.408)	(0.494)	(0.545)	(0.526)	(0.515)	(0.445)	(0.439)	(0.484)	(0.199)	(0.206)
Gender-neutral × Remote × Diversity	-0.255	0.192	-0.293	-0.194	-1.199	-0.745	-0.021	-0.673	-0.771	-0.054	0.146
	(0.637)	(0.589)	(0.698)	(0.766)	(0.745)	(0.727)	(0.629)	(0.632)	(0.685)	(0.283)	(0.299)
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). Regressions in this table correspond to the “long” or “fully saturated” model. See Appendix E for further discussion, including that the experiment was originally designed to estimate effects with a “short” model without the interactions. 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.22 provides key examples of how the gendered-language protocol works. Figure A.12 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.22: 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.12: 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 alumnae of the boot camps Laboratoria performed in Brazil, which were all in Portuguese. Only 43 of the 546 responses we obtained were from Brazilian alumnae (partly reflecting that about 14% of Laboratoria's alumnae 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 alumnae 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 alumnae 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 alumnae of the Brazilian boot camp (but only mentioned that one single Kindle would be awarded, given the smaller number of Brazilian alumnae).

> 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. Alumnae 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 for a 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 alumnae 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. Alumnae 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.13 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.14, A.15, and A.16 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 "*desarrollador*" 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.13: 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.14: 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.15: 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.16: Example of Ad in Laboratoria Experiment (UX Design)

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