

Does Mandating Women on Corporate Boards Backfire?*

Bo Bian[†]

Jingjing Li[‡]

Kai Li[§]

February 6, 2025

Abstract

This paper studies how external pressure to increase board gender diversity affects a firm's demand for female labor. Applying computational linguistic methods to job ads, we show that firms with a higher share of female directors write less female-friendly job ads. This is especially true in firms facing greater shareholder pressure, leaning toward Republican ideologies, or lacking female leadership. Using SB 826 in California as an exogenous shock, we find that board gender quotas can causally reduce firms' demand for female labor. Experimental evidence confirms this reduction and identifies psychological reactance and perceived gender norm violation as the primary mechanisms.

Keywords: Female Directors, Board Gender Quotas, Backlash, Job Postings, Female Labor Demand, Employee Reviews, SB 826, Computational Linguistic Methods.

JEL classification: *G30, G34, J16, J21, J23*

*We thank helpful comments from Audra Boone (discussant), Xiao Cen (discussant), Alex Edmans, Marina Gertsberg, Jing Gong, Todd Gormley, Shenje Hsieh (discussant), Wei Jiang, Ambrus Kecskes, Lenny Kostovetsky, Lena Larkin, Hamed Mahmudi, Feng Mai, David Matsa, Anya Mkrtchyan (discussant), Anh Nguyen, Alexandra Niessen-Ruenzi, Lin Peng, Felix von Meyerinck, Aazam Virani, Qiang Wu, Liu Yang, Xingtian Zhang, Jing Zhao, and seminar or conference participants at Baruch College, Central University of Finance and Economics, CKGSB, City University of Hong Kong International Finance Conference, KAIST, Hong Kong Polytechnic University, Hong Kong University of Science and Technology, Korea University, Northern Finance Association Meetings, Peking University, SFS Cavalcade North America, Stevens Institute of Technology, Texas A&M Young Scholars Symposium, University of Calgary, University of Hong Kong, and York University. We thank Ruiqi Ma and Yile Chen for excellent research assistance. Approval for the experiment in this paper was obtained from the Institutional Review Board (IRB) at University of Virginia under UVA IRB-SBS #6912. This study is registered with the American Economic Association RCT Registry (#15331). We acknowledge financial support from the Social Sciences and Humanities Research Council of Canada (Grant Number: 430-2021-00562). Bo Bian acknowledges financial support from the Phillips, Hager & North Centre at UBC. Kai Li acknowledges financial support from the Canada Research Chair in Corporate Governance.

[†]University of British Columbia, Sauder School of Business, bo.bian@sauder.ubc.ca.

[‡]University of Virginia, McIntire School of Commerce, jl9rf@virginia.edu.

[§]University of British Columbia, Sauder School of Business, kai.li@sauder.ubc.ca.

1 Introduction

In recent years, diversity, equity, and inclusion (DEI) initiatives have faced growing resistance. High-profile cases, such as the overturn of Nasdaq’s board diversity disclosure rules and the scaleback of their DEI commitments by major firms like Meta and McDonald’s, highlight the pushback these initiatives receive from both the public and corporations.¹ What triggers such pushback? Could well-intended policies backfire, undermining the very workplace equity and inclusion they aim to achieve? To provide an answer, this paper explores some unintended consequences of external pressure to increase board gender diversity by examining its impact on firms’ demand for female labor and shedding light on the underlying economic mechanisms.

Our empirical analysis faces a number of challenges. First, measuring gender-specific labor demand is difficult because firms do not typically disclose the gender composition of their workforce, and observed outcomes often obfuscate labor supply and demand factors. We overcome this by applying computational linguistic methods to online job ads to construct a proxy for gender-specific labor demand based on the usage of gendered language in job ads. Second, disentangling firms’ labor demand responses to external pressure from their voluntary endogenous behavior is crucial for causal inference. We address this by leveraging variations in external pressure from big institutional investors (Gormley et al., 2023) and, more importantly, exploiting board gender quotas implemented in California (Greene et al., 2020; Gertsberg et al., 2021; Hwang et al., 2021; von Meyerinck et al., 2021) and Europe (Ahern and Dittmar, 2012). Third, it is difficult, if not impossible, to gauge perceptions and intentions of individuals involved in hiring, which shape firms’ labor demand responses. To probe and distinguish among competing economic mechanisms, we run a randomized controlled experiment. The experiment provides clean causal evidence on how board gender quotas impact the demand for female labor, minimizing the influence of confounding factors that may present in real-world data.

¹See <https://abcnews.go.com/US/mcdonalds-walmart-companies-rolling-back-dei-policies/story?id=117469397> for a list of companies that have cut their DEI efforts.

We use job posting data from LinkUp, an employment intelligence company. We start with the gendered word list from Gaucher et al. (2011), who demonstrate that when job ads contain more masculine than feminine wording, women find those jobs less appealing. Our proxy for a firm’s demand for female labor is the share of feminine words out of all gendered words in a job ad. We also develop alternative measures by fine-tuning pre-trained Bidirectional Encoder Representations from Transformers (BERT, Devlin et al. (2018)) models to quantify the extent to which a job ad is female-tilted.

Our sample period is from 2014 to 2020, coinciding with the onset of growing pressure from stakeholders to improve board gender diversity. We document a negative and significant association between a firm’s share of female directors and its usage of feminine words in job ads, suggesting that a higher share of women on boards correlates with a weaker demand for female labor. This finding remains robust even when we hold constant the nature of a job by narrowing the comparison to job ads with the same occupation and identical job titles. Using an updated gendered word list developed by a leading job ad optimization company has minimal impact on our results.

This robust negative association between a firm’s share of female directors and its demand for female labor may seem counterintuitive. Firms adopting policies to increase female representation on boards may signal their greater openness to gender diversity. Moreover, having more female leadership helps improve female employment. Either scenario would suggest a positive, rather than a negative, association. How do we interpret this negative association? We propose “backlash” as a possible explanation, which traces its roots to the prominent psychological literature on reactance theory (Brehm, 1966) and the influential economics literature on the interplay between laws and social norms (Benabou and Tirole, 2011; Bowles and Polania-Reyes, 2012; Acemoglu and Jackson, 2017; Wheaton, 2022). These earlier papers demonstrate that when legislative changes clash with individuals’ preferred norms, to preserve those norms and regain their sense of autonomy, affected individuals may behave in direct opposition to that dictated by law.

In our setting, when there is external pressure (i.e., outside a company) to add more female directors on boards, members of that organization may push back as they perceive the external pressure as a threat to or loss of freedom, even if they support gender diversity in principle. Alternatively, such external pressure may clash with prevailing gender stereotypes, prompting members of that organization to push back by taking actions that undermine gender diversity elsewhere. It is worth noting that these two mechanisms could reinforce each other – when external pressure to add female directors triggers feelings of freedom being threatened or lost and at the same time conflicts with gender-related preferences and norms, reactance increases and resistance intensifies. In the end, empowering women through policies or mandates may backfire, which could manifest in a lower demand for female labor within affected firms.

Our first piece of evidence in support of these economic mechanisms comes from cross-firm variations in the amount of external pressure they face for board gender diversity. Our sample period 2014–2020 straddles the year 2017 when the three largest asset managers – State Street, BlackRock, and Vanguard – started to put pressure on their portfolio firms for board gender diversity. If the pushback against empowering women is driven by reactance to a perceived loss of autonomy, we expect it to be stronger in firms with a higher ownership by these three asset managers. For these firms, we indeed find that the negative association between a firm’s share of female directors and its demand for female labor is accentuated. We also find that the negative association becomes stronger when board gender diversity is more likely to clash with an organization’s gender norms. Specifically, we document more pushback in conservative, Republican-leaning firms or firms lacking female leadership.

To sharpen identification and provide further evidence consistent with pushback against external pressure, we use the passage of Senate Bill (SB) 826 in California (CA) in September 2018 as a laboratory. The bill required that all publicly listed companies headquartered in California have at least one female director by the end of 2019. We employ a difference-in-differences design in which we compare CA-headquartered firms (the treatment) with

non-CA-headquartered firms (the control). At the job ad level, we find that the adoption of SB 826 resulted in a four-to-five percentage-point drop in the share of feminine words out of all gendered words. We then explore different treatment intensities based on firms' ex-ante compliance with the requirements of SB 826. Firms facing greater compliance challenges show stronger negative responses, consistent with the backlash interpretation. We also estimate the dynamic effect of SB 826 and document little pre-trends but a clear negative post-trend.

Next, we run a randomized controlled experiment to help establish a direct link between subjects' exposure to board gender quotas and their gender-related hiring preferences. Importantly, the experiment reveals how individuals interpret quotas and incorporate them into hiring preferences, offering direct evidence for our proposed mechanisms and ruling out alternative mechanisms. These include moral licensing, whereby compliance with board gender quotas leads individuals to feel absolved from further promoting gender equality, and job insecurity, whereby quotas are viewed as threats to existing roles or job opportunities.

The experiment first manipulates the presence of a board gender quota policy to survey participants and examine how exposure to this policy affects treated subjects' hiring preferences. We next ask sets of questions guided by prior literature relating to different economic mechanisms: psychological reactance, norm violation, moral licensing, and job insecurity. The findings suggest reactance and norm violation as the dominant mechanisms, since treated subjects view quotas as infringing on their freedom of choice and misaligned with social and organizational norms. As a result, they ultimately shift preferences from hiring female to male candidates compared to control subjects. We further document that conservative subjects and male subjects exhibit stronger resistance toward hiring females, suggesting that reactance and norm violation could potentially reinforce each other in driving their responses. The experiment thus strengthens our interpretation of a proactive pushback against perceived deviations from established norms in the context of board gender quotas.

The evidence so far suggests a negative causal effect of adding more women on boards on

firms’ demand for female labor. Under the backlash interpretation, we expect such pushback to manifest not only in job ads but also in general treatment of female employees within a firm. To investigate this further, we analyze employee reviews to explore any attitude changes toward women. We show that after the passage of SB 826, compared to firms headquartered outside California, those headquartered in California received more negative employee reviews and poorer ratings of workplace accommodations that are important for female employees. This piece of evidence confirms that after the bill’s passage, the work environment became more hostile to female employees, consistent with backlash.

What happens to overall female employment after the bill’s passage? Aggregating the Current Population Survey data to state-worker class-month level over the period 2014–2020, we show that after the passage of SB 826, the share of female employment in the private sector in California experienced about a one percentage-point drop, all else being equal. In supplemental analysis, using staggered adoptions of board gender quotas in Europe over the period 2008–2020, we uncover a similar pattern, suggesting that the pushback extends to female employment.

We caution that this paper focuses on labor market outcomes for women, in particular firms’ demand for female labor, and we do not address broader efficiency or value implications. Many papers have examined the impact of board gender quotas on stock price and firm performance. These studies show negative announcement returns (Greene et al., 2020; Gertsberg et al., 2021; Hwang et al., 2021; von Meyerinck et al., 2021) and little or negative changes in firm performance (Ahern and Dittmar, 2012; Gopal, 2022). The pushback we document appears to stem more from personal preferences than economic considerations, potentially making affected firms less productive. This may partly explain the generally negative stock return and performance results associated with quotas.²

We conclude that board gender quotas can backfire, resulting in worse labor market outcomes for women in the corporate sector. Our analysis further suggests that the pushback

²Other papers suggest the shortage of suitable candidates, i.e., supply-side constraints, as an explanation. See Greene et al. (2020); Hwang et al. (2021); He and Whited (2023).

against those quotas is proactive in nature, driven by individuals’ perceptions that quotas jeopardise freedom of choice and violate existing gender-related preferences and norms. This stands in contrast to a more passive form of pushback, such as moral licensing. Delineating the underlying economic mechanisms is important because of their different policy implications. Proactive pushback implies resistance rooted in perceived loss of individual autonomy and deviations from norms, which may require targeted interventions, such as reframing policies to address concerns about the top-down approach associated with quotas and aligning them with social and organizational values.³

Related Literature. Our paper makes a number of contributions to the existing literature. First, our paper contributes to the growing economics and finance literature studying board gender quotas. Prior work produces mixed evidence on stock market responses to board gender quotas and their effects on operating performance and corporate innovation (Nygaard, 2011; Ahern and Dittmar, 2012; Matsa and Miller, 2013; Greene et al., 2020; Gertsberg et al., 2021; Hwang et al., 2021; von Meyerinck et al., 2021; Eckbo et al., 2022).⁴ There is, however, a scarcity of research exploring how quotas affect women’s labor market outcomes, especially for those not in leadership roles. We are among the first to employ granular multi-faceted data, spanning from hiring to workplace practices that affect the makeup and well-being of a firm’s labor force, to study whether and how mandating more women on corporate boards can generate some unintended consequences.

To this end, our paper is closely related to the seminal work by Bertrand et al. (2019) who examine the effect of board quotas on female labor market outcomes in Norway – one of the most progressive countries on gender equality – and conclude that there is no evidence suggesting that such quotas help improve female employment outcomes beyond women who

³For example, creating opportunities for open dialogue and involving the public in the policy making process can ameliorate feelings of external pressure and foster a sense of ownership. In addition, emphasizing the broader benefits of diversity for team performance and innovation, rather than focusing solely on rigid compliance, could help shift perceptions.

⁴Earlier pioneering work focuses on endogenously chosen gender diverse boards and their effects on firm outcomes, see Adams and Ferreira (2009) and Levi et al. (2014).

made it into boardrooms.⁵ We complement as well as extend Bertrand et al. (2019) in a number of ways. First, by applying computational linguistic methods to job ads in the US and hence separating labor demand from supply, we show that board gender quotas not only fail to deliver broad benefits, but also backfire by reducing firms’ demand for female labor. Second, we present evidence, including through an experiment, that suggests psychological reactance and gender norm violation as key mechanisms driving the pushback. Our findings help reconcile the many negative point estimates in Bertrand et al. (2019), which they describe as “counter-intuitive.” Third, our evidence from employee reviews highlights that the pushback extends to workplace poor treatment of female employees.⁶

Second, our paper adds to the fast-growing body of research that uses computational linguistic methods to gain insights into labor market dynamics and corporate outcomes. Deming and Kahn (2018) and Hershbein and Kahn (2018) are the pioneer users of job posting data to study skill requirements across firms and labor markets over time. Using job posting data, Babina et al. (2021) develop firm-level measures of AI investments to study firm growth and product innovation. Applying machine learning to earnings call transcripts, Li et al. (2021) develop measures of corporate culture and relate the presence of a strong culture to shareholder value. Leveraging tens of millions of job ads, our paper is one of the first to quantify firms’ preferences for one gender over another in their labor demands (see Bertrand (2020) for initial evidence). Our measures vary across different job positions within a firm and over time, thereby capturing firms’ demand for female labor in a granular and timely way. We expect that our method can be applied to job ads of private firms to study their labor market demand and/or extended to other countries and in different languages.

Finally, our paper contributes to the economics literature on backlash against regulatory and legislative efforts that push for social changes. A number of papers provide evidence

⁵Similarly, Maida and Weber (2022) evaluate the Italian board gender quota introduced in 2011 and find no evidence of it having effects on the representation of women in top corporate positions or as top earners.

⁶It is worth noting that the body of evidence in our paper is not inconsistent with Matsa and Miller (2011) and Tate and Yang (2015), who establish that women in power exert positive externalities on other women either in management roles or within the entire organization. We show that those positive externalities could be preceded and/or dominated by pushback from mandating “women in power.”

consistent with backlash against major social policy laws in the US, such as the 1975 revision to the Voting Rights Act (Ang, 2019) and the state Equal Rights Amendments in the 1970s (Wheaton, 2022). Existing work (Acemoglu and Angrist, 2001; Cullen and Pakzad-Hurson, 2021; Bailey et al., 2022) also cautions against laws targeting firms with social agendas, including gender equality, since firms’ responses might undo the intended effects or even backfire. In the context of empowering women in business, our findings suggest that legislative tools may have unintended consequences—they can hurt women’s advancement in business, because firms push back by reducing their demand for female labor.⁷ Our paper thus paints a nuanced picture of the economic consequences of board gender quotas and provides relevant input into future lawmaking and policymaking with a social agenda.

2 Data, Measurement, and Empirical Strategy

In this section, we first describe data sources and sample formation. Second, we construct a proxy for gendered labor demand by measuring gendered wording in job ads using two different computational linguistic methods. Finally, we explain our empirical strategies and provide summary statistics.

2.1 Data

Our job posting data come from LinkUp, an employment intelligence company. LinkUp has scraped job postings directly from public companies’ websites since 2007; its coverage has improved significantly since 2014. LinkUp provides reliable data for large companies that typically post jobs on their own websites. The data include job title, full-text description of each job, information about the employer, the date of the posting, the location in which a job is posted, and basic occupation information, such as the six-digit ONET Standard Occupation Classification (SOC) code. There are two advantages to obtaining job postings directly from employer websites as opposed to third-party job boards. First, employers

⁷Using a French law on academic hiring committees introduced in 2015, Deschamps (2023) documents that gender quotas in hiring committees can backfire and lower women’s probability of being hired due to reactions of men heading those committees. In contrast, Giannetti and Wang (2023) show that changes in societal views in favor of women are associated with an increase in board gender diversity, although they may not affect all firms equally due to different corporate cultures.

update their own websites more regularly than they update job boards. Second, employers post a job opening only once on their company’s website, as opposed to posting the same job multiple times across various job boards. As a result, LinkUp data suffer less from stale postings or duplicate postings compared to job posting data obtained from job boards.

For each job, we match LinkUp data to Compustat to obtain firm financial information; to BoardEx to obtain board characteristics, including the share of female directors in the job posting month; and to ExecuComp to obtain information on (up to) five top executives. Our US job-level sample comprises 30 million job ads from over 3,000 unique public companies over the period 2014–2020. We discuss the coverage of LinkUp in Online Appendix D.1, and show that using LinkUp data allows us to capture the largest firms in the economy without any industry or geographic bias.⁸ Table A1 in the Online Appendix provides the list of variables, their definitions, and data sources.

2.2 Measuring gendered wording in job ads

Words matter. Prior work documents that gendered language has implications for female labour market and other career-related outcomes. In a series of lab experiments, Gaucher et al. (2011) show that the mere presence of wording differences in job ads may be sufficient to change individuals’ appraisals of relevant jobs.⁹ Employing a gendered-grammar language, Spanish, in a large-scale field experiment, Del Carpio and Fujiwara (2023) show that job ads adopting gender-neutral language receive a higher share of female applicants. To delineate the mechanism, they run an experiment and find that gender-neutral language in job ads increases job appeal and helps form beliefs about a workplace’s culture of inclusion. These effects are economically large, comparable in magnitude to specifying that a job is remote and

⁸To facilitate regression analysis, most of our regression analyses employ a random one-fifteenth (1/15) of the full sample that comprises close to 2 million job ads. We confirm that the results remain robust across 1,000 random samples in Figure D11.

⁹For example, in their Study 4, masculinely-worded jobs are ranked 10%-20% less favorably than femininely-worded jobs by female subjects; the opposite pattern holds for male subjects. In other words, the share of female applicants in the pool of applicants can be reduced by 10%-20% due to language use in job ads. In their other experiments, when job ads include more masculine than feminine wording, participants perceive more men within those occupations (Study 3), and importantly, women find those jobs less appealing (Studies 4 and 5).

greater in magnitude than explicit diversity statements. Beyond job ads, Cohen et al. (2023) document that using gender-neutral language is associated with a significant improvement in the performance of women in Israel’s standardized college admission tests.

Bag-of-words (BOW) approach. Building on prior literature, we start with gendered word dictionaries that are pre-compiled word lists used to describe social traits and behaviors that differentiate females and males. These word lists are typically extracted from questionnaires given to college students to measure their self-concept and valuation of feminine and masculine characteristics (Bem, 1974; Rosenkrantz et al., 1968). We employ the gendered word dictionary created by Gaucher et al. (2011), based on published lists of communal (e.g., committed, supportive) and agentic words (e.g., individualistic, competitive) as well as feminine (e.g., compassionate, understanding) and masculine trait words (e.g., ambitious, assertive). Their list contains words and lemmas. Each lemma is followed by a wildcard (asterisk) that accepts all letters, hyphens, or numbers to incorporate inflectional or derivationally-related forms of a word. For example, *aggress** will match *aggressive* and *aggressiveness*. For our corpus of job ads, we take standard pre-processing steps, including lower-case transforming all words, removing trailing whitespace, line break, non-English words/posts, URLs, and special characters, etc.¹⁰ We then use regular expressions to reflect the aforementioned wildcard at the end of lemmas.

For each job ad, we count the number of feminine (masculine) words, and compute the share of feminine words over the total number of feminine and masculine words, as our proxy for female labor demand. Figure 1a (Figure 1b) shows the share of word frequency counts for the most frequent feminine (masculine) words in all feminine (masculine) word frequency counts over the entire sample period (2014–2020). We further document relatively stable shares of the most frequent gendered words over time in Figures 1c and 1d.

¹⁰Table D1 in the Online Appendix provides the complete dictionary of feminine and masculine words in Gaucher et al. (2011).

Large language models (LLMs). While using the gendered word dictionary from Gaucher et al. (2011) is transparent and easily replicable, this approach is subject to some limitations. First, the dictionary could be incomplete and outdated. We address this by augmenting the original list with additional words provided by a leading job ad optimization company.¹¹ Second, the bag-of-words approach ignores the order or context of words and does not capture relationships between words, which can be crucial for understanding nuanced meaning and complex text structure. To tackle this, we develop alternative measures of gendered wording in job ads based on LLMs. BERT (Devlin et al., 2018) is a pre-trained LLM that has demonstrated impressive performance on a wide range of natural language processing (NLP) tasks. To fine-tune a BERT model, we first form a representative sample of 4,500 job ads equally distributed in male-dominated, female-dominated, and gender-neutral occupations based on the share of female employees in a given occupation. We then have eight annotators to label them into feminine versus masculine job ads. Finally, we use the 80% of the labeled data as training data and the remaining 20% as testing data to fine-tune the model. Our chosen model is JobBERT, pre-trained on 3.2 million sentences from job postings (Zhang et al., 2022).¹² The fine-tuned model assigns a probability between 0 and 1 indicating the likelihood of each job ad being female-tilted. Our model has an area under the receiver operating characteristic curve (AUC) score over 0.9 in predicting the “feminine” probability. As expected, the LLM-based measure is positively correlated with the measure based on the gendered word dictionary. For details of our LLM-based method and statistics, see Online Appendix B.

Validation. We first establish that gendered language usage in job ads is a reasonable proxy for gendered labor demand. Figure D5 in the Online Appendix presents a sample overview of gendered language usage in different occupations based on a job’s two-digit ONET-SOC code. We show that across the six major occupational categories, the sales

¹¹See <https://www.ongig.com/gender-bias-in-job-descriptions>.

¹²For robustness, we also use DistilBERT, a distilled version of BERT that retains almost all the language understanding capabilities of BERT while being faster and smaller (Sanh et al., 2019).

and office sector and the service sector, traditionally associated with high shares of female employment, have the highest shares of feminine words in job ads, whereas the military sector, traditionally associated with high shares of male participation, has the lowest share, suggesting that gendered language usage could be a meaningful proxy for gendered labor demand.

We next examine whether and how gendered labor demand varies with occupation-level characteristics. Figure D6 splits the sample using four different occupation-level characteristics: job zone, mean wage, entry education requirement, and STEM occupations or not. We show that jobs requiring more preparation (Figure D6a), offering above-the-median pay (Figure D6b), requiring above-the-median levels of education (Figure D6c), or being in STEM sectors (Figure D6d) tend to employ more masculine words in job ads, and hence more male employees, than jobs requiring less preparation, offering below-the-median pay, requiring below-the-median levels of education, or not being in STEM sectors.

Other outcomes beyond job ads. We use job ads as our primary measure of gendered labor demand since they offer several unique advantages. First, job ads are ubiquitous. This feature is particularly valuable for studying backlash, as it allows us to capture hiring patterns in a wide range of occupations, firms, and industries. Second, job ads provide a widely accessible and consistent signal of hiring intent, whereas there is no economy-wide granular data on the gender gap in pay, promotions, and dismissal. Third, job ads reflect pure demand-side preferences, devoid of supply-side confounders that show up in measures like the gender gap in employment.

One potential concern is whether gendered language in job ads meaningfully discourages applications or influences actual hiring decisions. One may also argue that since our key measure is the share of feminine words over the total number of feminine and masculine words, its variation can be driven by changes in masculine words. To address these concerns, we draw on a diverse set of datasets and methodologies to supplement our job ad analysis with

additional measures that capture general treatment of women in workplace. Specifically, we examine employee reviews, gender-related labor violations, and female employment trends to provide a broader perspective beyond job ads. Additionally, in a randomized controlled experiment, we directly ask subjects about their hiring preferences, which are more closely tied to hiring outcomes than job ads.

2.3 Empirical strategy

Correlational analysis. For motivating evidence, we examine the correlation between board gender diversity and gendered language usage in job ads. We run the following regression employing a job-level sample:

$$Y_{p,t} = \gamma_i + \delta_{l,t} + \alpha_{k,t} + \vartheta_q + \beta \times \%Fem\ dir_{i,t} + \theta' X_{i,t} + \varepsilon_{p,t} \quad (1)$$

where p denotes job ad, i denotes firm, l denotes two-digit SIC industry, k denotes region (state/zip), q denotes occupation, and t denotes year-month when a job is posted. The dependent variable, $Y_{p,t}$, is the share of feminine words out of all gendered words in a job ad – an ad-level measure of gendered labor demand. Ad- and firm-level control variables, denoted by $X_{i,t}$, include log (# words in an ad), firm female leadership (i.e., women in C-suite), firm size, cash holdings, tangibility, Tobin’s Q, leverage, ROA, and capex. In Equation 1, the coefficient of interest, β , captures the relationship between female board representation and gendered labor demand.

Identification using SB 826 in California. We exploit the passage of SB 826 in California on September 30, 2018 for identification and to help pin down the economic mechanisms. The bill required companies listed on major US stock exchanges and headquartered in the state (over 600 firms) to have at least one female director on their boards by the end of 2019.¹³ We run the following difference-in-differences (DID) regression:

¹³It further required that by the end of 2021, all firms must have at least one female director if their boards have four directors or fewer, two female directors if their boards have five directors, and three female directors if their boards have six directors or more.

$$Y_{p,t} = \gamma_i + \delta_{l,t} + \alpha_{k,t} + \vartheta_q + \beta \times Flag_CA_i \times Post_2018m9_t + \theta' X_{i,t} + \varepsilon_{p,t} \quad (2)$$

where the notations are the same as those in Equation 1. The indicator variable, $Flag_CA_i$, takes the value of one for public firms headquartered in California, and zero otherwise. The indicator variable, $Post_2018m9_t$, takes the value of one after September 30, 2018, and zero otherwise. The coefficient of interest, β , captures the (reduced-form) effect of SB 826. Under the backlash interpretation, firms in California push back gender-balancing quotas in their boardrooms by reducing their demand for female labor elsewhere, we expect a negative β .

In both Equation 1 and Equation 2, the unit of observation is at the job-ad level. One benefit of such a specification is that it allows us to include high-dimensional fixed effects to control for occupation-level characteristics and local labor supply shocks. Specifically, we include firm and industry $\times$ time fixed effects (γ_i and $\delta_{l,t}$) to absorb firm-invariant characteristics and time-varying industry-specific shocks, respectively. Since labor markets are segmented by region, we add region $\times$ time fixed effects ($\alpha_{k,t}$) to account for time-varying regional labor supply shocks. We define region at either the state or zipcode level.¹⁴ Importantly, we also include occupation fixed effects (at the six-digit ONET-SOC code level) to allow for a comparison of job ads within a specific occupation across firms with different female director ratios.¹⁵ As an alternative to occupation fixed effects, we include more refined job title fixed effects to ensure that effectively we compare the same job to extract any possible change in the language used (instead of any change in the type of jobs advertised).

Experiment. Following the analyses with real-word data using the two strategies above, we run a randomized controlled experiment to help delineate how board gender quotas shape the demand for female labor. Subjects are randomly assigned to one of the two groups: a board gender quota-exposed group (treatment) or a group with no such exposure (control). Subjects are asked to imagine themselves as employees of a fictional company and are intro-

¹⁴Note that information for a zipcode is not always available, resulting in a smaller sample size when job zip $\times$ year-month fixed effects are included in the regression.

¹⁵There are around 1,000 unique occupations at the six-digit ONET-SOC code level. See <https://www.onetonline.org/find/all> for a list of occupations.

duced to the policy specific to their assigned group. The experiment involves 200 US-based participants with diverse demographic backgrounds, recruited via CloudResearch’s Connect platform. Eligibility criteria include age (18-65) and work experience (at least one year).¹⁶ Since we ask subjects’ preferences when they choose between male and female candidates for a non-leadership position, we address potential concerns that gendered wording in job ads may not accurately reflect the demand for female labor. For example, one may argue that such wording could be influenced by factors such as candidate qualifications, organizational fit, or other considerations unrelated to labor demand. The experiment allows us to capture hiring preferences explicitly tied to gender, as well as addresses potential confounding factors that are difficult to control for using real-world data.

Importantly, by eliciting the rationales behind subjects’ reactions to board gender quotas, the experiment sheds light on the underlying economic mechanisms. After the hiring preference question, we ask subjects questions designed to uncover the behavioral and psychological factors that shape their reactions. This helps us differentiate competing mechanisms. We describe the experiment in detail in Section 3.4 and Online Appendix C.

2.4 Summary statistics

Panel A of Table 1 presents the summary statistics for observations at the job level. The mean/median share of feminine words in a job ad is 58%/57%. The mean/median number of feminine (masculine) words in a job ad is 4.51/3.71 words (4/2 words). Panel B presents the summary statistics for job ads at the firm-year level, aggregated using all 30 million job ads posted by public firms in the fully-merged sample. A median firm in our merged sample posts 214 jobs per year, with a mean share of feminine words around 55%. Out of the 9 directors in a median firm, 17% are females.¹⁷ Panel B also presents the summary statistics for control variables used in the regression analysis of Equation 1. These firm characteristics

¹⁶Our sample may seem small, which biases against finding significant results. In contrast, in large, non-clustered experimental setting, even minor effects could appear significant.

¹⁷In our regression specification of Equation 1, the variable of interest is the share of female directors at the firm-month level, to accurately capture director turnovers that could take place at any month in a given year.

are largely consistent with those of large Compustat firms, given that LinkUp’s coverage leans toward large companies that typically post job ads on their own websites.

3 Gendered Labor Demand

We first examine the correlation between female board representation and gendered wording in job ads. We then explore possible economic mechanisms by exploiting cross-firm variations in the amount of external pressure faced and investigating changes around the passage of SB 826 in California.

3.1 Motivating evidence

Table 2 presents the regression results following Equation 1. In column (1), we include time-varying accounting variables and a rich set of fixed effects to control for firm-level time invariant characteristics, industry shocks, and state-level shocks that can affect labor supply in a given state. We find a negative and significant coefficient on $\%Fem\ dir$, pointing to a negative association between the share of female directors at a firm and the share of feminine words in job ads of the same firm. We continue by adding more granular fixed effects in columns (2) and (3). We compare within each occupation (at the six-digit ONET-SOC code level) for jobs posted by firms with different levels of female board representation in column (2). The estimated coefficient is similar in terms of magnitude and significance to that in column (1). To better control for local labor supply shocks, we add job zip $\times$ year-month fixed effects in column (3) and again find a negative and significant coefficient on $\%Fem\ dir$. Columns (4) and (6) further include female executive indicators and the length of each job ad as controls. The estimated coefficients change little.

In terms of economic significance, we show that a one-standard-deviation increase in the share of female directors (0.12) is associated with a 1.2 percentage-point drop in the share of feminine words in job ads (based on the most stringent specification in column (6)). This drop is economically significant given the annual increase of 0.7 percentage points in the share of feminine word usage in job ads before SB 826.¹⁸

¹⁸See Figure D7a for the temporal trend in the average share of feminine words in job ads from 2014 to 2017, a period less influenced by strong external pressure to increase female board directors.

Using LLM-based measures of gendered wording in job ads, we find similar results. In Panel A of Table B1, we show the results from correlational analysis using the bag-of-words (BOW) approach and the JobBERT method. Columns (1) and (2) start with the continuous measure. The coefficients are highly similar in terms of both economic and statistical significance. Given that the LLM-based measure has a rather skewed distribution, we also employ indicator variables based on quartiles and deciles of the continuous measure. We find stronger negative effects that are significant at the 1% level using those indicators for gendered wording in job ads, as shown in columns (4) and (6).¹⁹

Interpreting the negative association. As discussed earlier, if a higher share of female directors signal firms’ commitment to gender equality, and/or if more female leadership supports more demand for female labor, we would expect a positive association between a firm’s share of female directors and its gendered labor demand in favor of women. Instead, we observe a negative association. To explain this seemingly counterintuitive finding, we draw on two lines of research, which provide the microfoundation for backlash when imposed (legislative) changes conflict with prevailing norms.

First, the reactance theory in psychology (Brehm, 1966; Brehm and Brehm, 2013) posits that individuals will push back when there is a threat to or loss of freedom. In our setting, when there is external pressure to add more women on corporate boards, it may be perceived as a loss of autonomy and create a sense of constraint, triggering psychological reactance that leads to resistance. In response, members of that organization may attempt to restore their freedom of choice by acting in direct opposition to such a constraint, resulting in the negative association observed in our job ad analysis.

Second, a long-standing economics literature examines how laws and institutions interact with social norms and trigger pushback. Benabou and Tirole (2011) model the interplay between laws and norms and show that laws can crowd out and undermine social norms

¹⁹Results using DistilBERT are also similar, see Table B2.

in certain scenarios, especially when the expressive role of law is invoked.²⁰ Acemoglu and Jackson (2017) show that when laws clash with prevailing social norms – for example, when laws attempt to restrict people’s behavior relative to the distribution of anticipated payoff-relevant behavior (i.e., social norms) in society, more people will elect to break the law. In a similar vein, Wheaton (2022) highlights a social version of crowd-out (Bowles and Polania-Reyes, 2012) – individuals will push back against law as they seek to preserve their preferred norms. He warns that any aggressive push for social change through legislation may face significant costs in terms of countervailing cultural backlash. Empirically, prior work has documented backlash from many major US social policy laws.²¹ Mirroring those earlier findings in our context, board gender quotas may clash with gender stereotypes prevailing within an organization, prompting members of that organization to take actions that undermine gender equality.

Under either mechanism, empowering women through board gender quotas may backfire. It is worth noting that psychological reactance and perceived norm violation can work independently or norm violation may amplify reactance when individuals view quotas as against their long-held gender norms, leading to intensified feelings of a threat to or loss of autonomy.

However, it is empirically challenging to establish that firms are forced to have more female directors and link any pushback to the proposed mechanisms. The observed variation in the share of female directors could come from choices made by firms without any legal constraints or external pressure. We first tackle this challenge by focusing on cross-firm heterogeneities, varying by the amount of external pressure they face on board gender diversity. We then study law changes that mandate women on boards. Finally, we run a

²⁰Pioneering work in legal scholarship discusses the expressive role of law and its effects on cultural norms and attitudes (Sunstein, 1996; Cooter, 1998; Posner, 2009). Simply put, in addition to its functional role, law also plays a key role in expressing and shaping the values of a society.

²¹Using historical data and surveys, Ang (2019) examines the long-run effects of the federal oversight of election laws through the US Congress’ 1975 revision to the Voting Rights Act (VRA) and provides evidence of backlash among the white majority, who have reduced their support for the Democratic party. Using survey data, Wheaton (2022) shows that after the enactment of the state Equal Rights Amendments (ERA) of the 1970s, men in states that passed an ERA reacted by becoming less likely to believe in women’s equality.

randomized controlled experiment that manipulates the presence of board gender quotas to elicit subjects’ gender-related hiring preferences and rationales behind those preferences.

3.2 Cross-firm heterogeneities

External pressure from institutional investors. In 2017, “The Big Three” (Big3) asset managers (i.e., State Street, BlackRock, and Vanguard) launched campaigns to increase gender diversity on boards of their portfolio firms. Both the BoardEx population of firms (Figure D1a) and S&P 1500 firms (Figure D1b) show an accelerated rise in the share of female directors after 2017, suggesting a concerted pressure from major shareholders on board gender diversity. If external pressure triggers reactance due to a perceived loss of autonomy, we would expect the negative association between a firm’s share of female directors and its female labor demand to be accentuated in those firms facing the most pressure from their shareholders. If, instead, the variations in board gender diversity are entirely voluntary and signal firms’ commitment to gender equality, the association should be positive and/or invariant to external pressure.

To test the above conjecture, we introduce an indicator variable, *High Big3 ownership*, to flag firms whose Big3 ownership is in the top quartile. Column (1) of Table 3 presents the regression results following Equation 1 with an additional interaction term $\%Fem\ dir \times High\ Big3\ ownership$. The coefficient on this interaction term is negative and significant at the 1% level. The results indicate that for firms with the top quartile Big3 ownership, the effect of board gender diversity more than doubled. Interestingly, the effect for the remainder is small and insignificant, as captured by the coefficient on the standalone term $\%Fem\ dir$. Thus the observed negative association in Table 2 is mainly driven by firms facing great pressure from these largest asset managers.

Political leaning. Earlier, we argue that one mechanism underlying backlash is the board gender diversity policy’s clash with gender stereotypes prevailing within an organization. Given that supporters of the Republican party hold conservative views in general, we expect

Republican-leaning firms to exhibit more backlash compared to their Democratic-leaning peers. To capture a firm’s political leaning, we introduce an indicator variable, $> 50\%$ *Republican contribution*, flagging firms that contribute more to Republican candidates than to Democrat candidates based on data from the Federal Election Commission. Column (2) of Table 3 presents the results. Consistent with the norm violation mechanism, we find the negative association is concentrated in Republican-leaning firms.

Female leadership. The gender-specific preferences of female executives likely do not clash with the board gender diversity policy, which predicts less backlash. Moreover, the presence of female executives helps ease the bluntness of such policy and raises awareness of female employees’ skills and abilities, which again predicts less backlash (Dobbin and Kalev, 2016; Acemoglu and Jackson, 2017). Finally, female executives may also help curtail gender discrimination in the workplace by their roles as leaders.

By hand collecting relevant data, we identify genders of Chief Executive Officers (CEOs), Chief Financial Officers (CFOs), and Chief Operating Officers (COOs) based on their respective job titles, and construct an indicator variable, *Female Top3 exec*, flagging firms with at least one female top executive. We also introduce a second indicator variable, *Female COO*. The COO ensures that a company’s operations run smoothly. Their focus is on areas such as recruitment, training, payroll, legal, and administrative services. The COO is usually second in command to the CEO. Our focus on the COO is new in the literature. Columns (3) and (4) of Table 3 present the results. We find evidence that the presence of female top executives helps mitigate backlash, especially when the COO is a female, as shown by the positive coefficient on the interaction term.

Decentralization in drafting job ads. We expect that both psychological reactance and perceived norm violation can arise at different levels of a corporate hierarchy. Consequently, it should not matter who drafts job ads. Consistent with this conjecture, we show in Table D2 that pushback is invariant to firms’ different degrees of decentralization in drafting job ads,

suggesting that pushback occurs in a wide range of firms that differ in their hiring practices.²² The methodology for constructing different decentralization measures, along with a detailed interpretation of the results, are provided in Online Appendix D.2.

3.3 Identification through California’s SB 826

So far, we find that female representation on boards negatively correlates with female-friendly word usage in job ads in 2014–2020. The evidence from various heterogeneity tests underscores the role of external pressure firms face for board gender diversity during this time period, supporting the backlash interpretation. While informative, the correlational nature of the above analysis means that omitted variables or selection could bias the estimated coefficients. It is worth noting, however, the magnitude of the correlation goes up when we employ more stringent specifications in Table 2, suggesting that possible endogeneity issues likely result in an underestimation of the true correlation. The reason is that any unobserved factor tends to drive the share of female directors and the share of feminine words in job ads in the same direction, biasing the estimated coefficient upward, i.e., towards zero.²³

We resort to a law change in California during our sample period to better identify the causal impact of female board representation on gendered language usage in job ads. Studying a law change also aligns our analysis more closely with the proposed economic mechanisms – psychological reactance to perceived threat to or loss of autonomy and the misalignment between laws and norms. SB 826, enacted on September 30, 2018, applied to US and foreign firms listed on major US stock exchanges and headquartered in California. We first determine a firm’s headquarters’ state from Compustat and note that 17% of US firms have their headquarters in California.

In Table D4 in the Online Appendix, we first cross-validate whether the adoption of SB 826 resulted in an increase in the number/share of female directors for CA firms compared

²²We also group different jobs into high- and low-tier ones using four different metrics. In Online Appendix Table D3, we find a stronger pushback for lower-tier jobs that are far more frequently featured in job ads than higher-tier jobs, suggesting broad and pervasive resistance to hiring women.

²³For example, assume that we cannot measure a female-friendly culture at a firm. Such a culture is likely positively correlated with both a firm’s share of female directors and its share of feminine words in job ads.

to non-CA firms. We find that indeed, the adoption of SB 826 resulted in a 0.12-0.15 (1.7% - 1.8%) increase in the number (share) of female directors in CA firms compared to that in non-CA firms.

DID regression results. Figure 2 compares the time trend in the share of feminine words for CA-headquartered firms around the bill’s passing in September 2018 with that for non-CA-headquartered firms over the same period. We observe a distinct drop in the share of feminine words used by CA-headquartered firms in their job ads compared to that by non-CA-headquartered firms. Moreover, the drop in the share of feminine words used by CA firms appears to have started in early 2018. Based on the bill’s history from the state government website,²⁴ the first sign that it would become law appeared in April 2018, when it was read at the state assembly for the first time and passed with strong support.

We next employ a DID design in Equation 2 in which we compare firms headquartered *inside* and *outside* California, *before* and *after* the September 2018 law change. Panel A of Table 4 presents the results. We find that SB 826 reduced the share of feminine words by 4.2-5.4 percentage points. We note that, using a law change helps address potential endogeneity issues and underestimations associated with OLS regressions, leading to a significantly larger effect compared to that in Table 2.

In terms of economic significance, the drop in the usage of feminine words by four to five percentage points represents about 10% of the mean share of feminine words in job ads. We assume this estimate to be a crude proxy for the share of job ads tilting from femininely-worded to masculinely-worded. According to Gaucher et al. (2011), masculinely-worded job ads are ranked 10%-20% less favorably than femininely-worded job ads by female applicants. Combining both statistics suggests that the female applicant pool shrinks by 1%-2%. All else equal, this implies that 1%-2% fewer females would be hired, i.e., the share of female employment out of total employment would drop by 1%-2%. At the time of the bill’s

²⁴https://leginfo.legislature.ca.gov/faces/billVotesClient.xhtml?bill_id=201720180SB826.

passing, female accounted for 46.9% of the labor force, and a drop of 1%-2% corresponds to a drop in the share of females in the labor force by 0.5 to 1 percentage points.²⁵ Over the period 1950–2020, the share of females in the labor force rose from 28.6% to 46.8%, an annual increase of 0.26 percentage points. Taken together, this simple back-of-the-envelope exercise suggests that SB 826 set back the progress in female labor market participation by roughly two to four years.

We use LLM-based gendered wording measures to show robustness of our main findings. In Panel B of Table B1, we present the DID results using the bag-of-words (BOW) and JobBERT measures. In columns (1) and (2), we start with the continuous measure and find a significantly negative effect of SB 826 using the JobBERT method. This effect becomes even stronger when we replace the continuous measure with quartile or decile indicators, as shown in columns (4) and (6).

Treatment intensity. To address the concern that our DID results could be driven by concurrent events such as the #MeToo Movement in 2017, we explore different treatment intensities based on whether a CA firm was in compliance with SB 826 or not prior to the bill’s passing, or whether a CA firm experienced an increase in its number of female directors after the bill’s passing. Table 5 presents the results. Columns (1) and (2) compare firms just below the threshold set by SB 826 (one female director short) and firms at the threshold before the treatment date. We show a significantly negative effect for CA firms whose number of female directors fell short of the threshold. In contrast, those meeting the threshold experienced little pushback. Columns (3) and (4) compare firms that experienced an increase in their number of female directors after the treatment date versus those without an increase. The treatment effect only showed up in the former set of firms. The findings above further strengthen our identification strategy using SB 826.

²⁵See <https://www.dol.gov/agencies/wb/data/lfp/civilianlfbbysex> for information from the U.S. Bureau of Labor Statistics.

Dynamic effect. A key identifying assumption is the presence of parallel trends, i.e., in the absence of SB 826, the usage of gendered language in job ads would have followed similar trends for CA-headquartered and non-CA-headquartered firms. Figure 3 presents the temporal pattern based on regression results. To reduce estimation error, we group every three months within a corresponding quarter. We assign quarter -1 as the benchmark quarter and combine quarters prior to $t = -6$ (or after $t = 6$) into a single period. The dynamic DID coefficients confirm that the effect of SB 826 only materialized after there were strong expectations of the bill’s passing. The coefficients for all quarters prior to the second quarter of 2018 are mostly close to zero and insignificant. However, there is a clear negative post-quota trend, suggesting significant pushback after the bill’s passing.²⁶

3.4 Evidence from an experiment

We interpret our baseline findings as an act of proactive pushback, where external pressure triggers psychological reactance due to perceived imposition and loss of autonomy, and/or board gender quotas clashes with deep-rooted gender norms within an organization, leading individuals to resist such quotas by favoring male candidates in hiring decisions. However, one may argue that our findings are instead a consequence of moral licensing (Blanken et al., 2015; List and Momeni, 2021). When individuals initially behave in a moral way (e.g., supporting greater female representation on boards), they are later more likely to engage in behaviors that are immoral or unethical (e.g., not supporting broad gender equality in their workplace). Another possible explanation is that we simply document rational reactions from employees who want to protect their own jobs because with the board gender diversity policy/mandate, (incoming) female employees are “supported” (or “have it easy”).

Although the heterogeneity analysis along shareholder pressure, political view, and female leadership in Section 3.2 provides some suggestive evidence supporting our proposed

²⁶In 2019, three California taxpayers filed a lawsuit (*Crest v. Padilla I*), seeking to nullify SB 826. The plaintiffs’ main argument was that the bill violated the Equal Protection Clause of the California Constitution. The fact that the lawsuit was initiated shortly after the bill’s passing is consistent with backlash. A Los Angeles County Superior Court struck down SB 826 in May 2022. The California Secretary of State appealed and a final ruling is still pending. Although these recent legal developments are outside of our sample period, future research on their ramifications will have important policy implications.

mechanisms, it cannot rule out the aforementioned alternative mechanisms. Differentiating between these mechanisms is particularly challenging, as we do not typically observe individuals' motives behind their actions. To overcome this challenge and better delineate the economic mechanisms underlying the observed labor demand responses to board gender quotas documented earlier in the paper, we run an experiment which we manipulate the presence of a board gender quota policy.

The policy statement provided to the treatment group is as follows: *“The company you work for follows a **government-enforced gender quota policy** that **mandates** a specific percentage of women must be included among its board members. This policy **applies strictly to leadership-level hiring decisions** and limits the company’s ability to select candidates solely based on qualifications or experience for these roles. The policy is **non-negotiable**, and **compliance is required to avoid penalties**. However, this policy does not apply to **non-leadership positions**, where hiring decisions are made independently of the quota. The goal of the policy is to enforce gender diversity and ensure balanced representation in decision-making processes.”* Subjects in the control group are shown a statement without the gender quota component. Table C1 details both statements and the pre-trial instructions.

After manipulation checks, we ask subjects (both the treatment and control groups) their gender-specific hiring preferences with the following question: *“Imagine you are in charge of hiring for a non-leadership position, and your input will be critical in the decision-making process. Keeping the policy you just read in mind, to what extent do you agree with the following statements? (1) The recruitment policy shifts my attention toward male candidates for non-leadership roles; (2) I find myself considering male candidates more often for non-leadership positions in this context; (3) The recruitment policy makes me more inclined to favor male candidates for non-leadership positions.”* Subjects respond using a seven-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Figure C2b displays a screenshot of this question.

Finally, we ask questions adapted from existing literature to examine the following economic mechanisms: reactance, norm violation, moral licensing, and job insecurity. Delineating the underlying mechanisms is important because each mechanism entails a different policy prescription. Similar to the hiring preferences questions, we use a multi-item, seven-point Likert scale, with the specific questions listed in Table C2. For additional details on our experimental design, including the questions used for manipulation and attention checks, refer to Online Appendix C. Table C3 shows that subjects in the treatment group are comparable to those in the control group in age, race, work experience, or political view, validating our randomization design.

Table 6 presents the experimental evidence on subjects' hiring preferences and sheds light on the underlying mechanisms by comparing the treated and control groups. In column (1) of Panel A, we first document a significant manipulation effect, indicating that the treatment has successfully primed subjects as intended. In column (2), we examine the treatment's influence on hiring preferences by analyzing the extent to which subjects' preferences for male versus female candidates differ between the treatment and control groups. We find a significant, positive treatment effect, suggesting that subjects exposed to board gender quotas are more likely to favor male candidates for a non-leadership position than their counterparts not exposed to the policy. The effect is economically large, representing approximately 17% ($0.512/2.992$) of the average value on the Likert scale measuring a preference for male candidates. This result reinforces our earlier findings from job ads, providing consistent evidence of a reduced demand for female labor after the policy. Together, these findings strengthen our main argument in the paper that board gender quotas may unintentionally impact hiring decisions, disadvantaging women outside the boardroom.

In columns (3) to (6) of Panel A, we delve into the economic mechanisms. Columns (3) and (4) explore whether the treatment triggers psychological reactance and perceptions of gender norm violation. We find positive and large coefficients, significant at the 1% level. These findings indicate that treated subjects are significantly more likely to view the policy as

restricting their decision-making autonomy and conflicting with established workplace norms than control subjects. Treated subjects exhibit elevated levels of psychological reactance.²⁷ Columns (5) and (6) further explore moral licensing and job insecurity, respectively. Interestingly, for these two mechanisms, we find smaller and insignificant variations across the treatment and control groups, suggesting that they are not primary drivers of the observed responses to board gender quotas.

One potential concern about our experimental design is that some subjects may lack recruiting experience, raising questions about the real-world relevance of their survey responses. To address this, we exclude subjects with less than five years of work experience, as well as those who are self-employed, and re-run the analysis. Table C4 presents the results. Our main findings remain.

In Panel B of Table 6, we take one step further by exploring if the strength of subjects' reactions in their hiring preferences, as well as the underlying mechanisms, depend on their gender-related views and preferences. The idea is that the extent to which board gender quotas deviate from those preferences determines a subject's treatment intensity, allowing us to examine the heterogeneous treatment effects.

We first analyze the role of political view, comparing reactions from subjects with conservative or liberal views. To do this, we include an interaction term between treatment and political view (1-7 from liberal to conservative) in the regression. The treatment effect for hiring preferences is significantly stronger among conservative subjects (column (1)). Their reactions to the policy are also significantly more pronounced (column (3)) since these subjects are more likely to view the quota policy as misaligned with gender norms in their workplace (column (5)). We also examine whether the treatment effect varies by gender. The findings reveal notable differences, with male subjects showing stronger resistance to the policy in their hiring preferences than female subjects. This resistance is closely linked to

²⁷Table C5 presents the three subscores of reactance: experience of reactance, negative attitudes, and aggressive behavioral intentions. The treatment effects are estimated separately for each subscore, and we observe consistently significant and positive effects across all three.

the underlying mechanisms – we find greater psychological reactance among male subjects, in line with elevated perceptions of norm violation, compared to their female counterparts (columns (4) and (6)). These findings suggest that reactance and norm violation could potentially reinforce each other in driving subjects’ hiring preferences.

Taken together, the experimental evidence complements the job ad results by showing that subjects exposed to the board gender quota policy exhibit significantly weaker preferences for female candidates compared to their counterparts not exposed to such a policy. Moreover, the experiment allows us to pin down reactance and gender norm violation as the dominant economic mechanisms underlying our main findings.

3.5 Additional results and robustness checks

3.5.1 Gendered job ads vs. gendered jobs

One may argue that we might be capturing gendered jobs instead of gendered job ads because the language in job ads often correlates with job characteristics. For example, the word “analy*” shows up in all job ads concerning “analysts” and it is a masculine word. Our goal is to measure changes in the language used rather than changes in the type of jobs advertised. To achieve this, ideally one would compare the exactly same job advertised by different firms at a point in time or by the same firm over time. In earlier specifications, we include occupation fixed effects (at the six-digit ONET-SOC code level). However, one concern is that such a classification is too crude, and jobs within the same occupation might still differ in their gender shares, skill requirements, compensation, career prospects, etc. To alleviate this concern, we take several approaches.

First, we limit the comparison to job ads with the same job title, which is substantially more granular than the six-digit ONET-SOC code. We standardize job titles by keeping English words only, lower-case transforming all words, and removing stop words. In column (1) of Table 7, we include full job title fixed effects. One may argue that these job title fixed effects are too granular and that the same job can have slightly different titles. As a result, the sample of job ads that fall within the common support of control variables and fixed

effect units would be very thin. So instead of considering the full job title, we only keep the first three words in a job title to form fixed effect units in column (2).²⁸ In column (3), we further restrict the comparison to job ads with the same first three words in a job title, the same ONET-SOC code and, importantly, posted by the same firm – these jobs likely have characteristics that are nearly identical to each other in nature.²⁹ We continue to find that a firm’s share of female directors is negatively associated with its share of feminine words in job ads, similar to the results in Table 2. We also find that the effect of SB 826 is highly comparable to the results in Table 4.

Second, since the main concern centers around us capturing gendered jobs (instead of gendered labor demand), we limit our analysis to gender-neutral jobs. Specifically, we drop potentially gendered jobs based on the share of female employees in a given occupation.³⁰ For instance, if an occupation has a female share of 5% or 95%, jobs posted in this occupation category are likely gendered (male- or female-dominated). In column (4), we keep jobs in occupations whose female share is between 25% and 75%. In column (5), we impose a tighter restriction that we only keep occupations whose female share falls between 40% and 60%. This allows us to capture changes in the language used in job ads for gender-neutral jobs rather than gendered jobs (i.e., jobs with either a very high or a very low share of female employees). The estimated coefficients remain similar to the baseline estimate.

Third, certain words in our gendered word list may be more problematic than others in the sense that they may reflect the nature of the job rather than the gendered language choice. For example, “analy*” shows up in all job ads for the job “analyst,” while “analytical” is presumably a male trait and a language choice that discourages female applicants. However, a company cannot advertise “analyst” jobs without using the word “analyst.” To ensure

²⁸The most popular job titles include “restaurant general manager,” “customer service associate,” “retail sales associate,” and “assistant general manager.” Using the first three words in a job title reduces the number of unique fixed effect units by half.

²⁹Sometimes the same job title may refer to different jobs. For example, an “assistant general manager” job may be advertised by many firms and belong to different ONET-SOC codes.

³⁰We calculate the female share using the 2020 Labor Force Statistics from the Current Population Survey. See Table 39 from https://www.bls.gov/cps/cps_aa2020.htm.

that our results are not driven by a few selected feminine or masculine words, we recalculate our key measure of interest – the share of feminine words out of all gendered words in a job ad – by excluding each of the ten most popular feminine or masculine words in job ads. As shown in Panel A of Figures D8 and D9, the estimated coefficients using these new measures are similar to the baseline estimate. Furthermore, we examine the frequency of every word in our gendered word list in each job ad. If indeed words like “analy*” reflect the nature of the job, they should show up frequently in the job description. We thus re-compute our key measure by dropping the five most frequent job-specific gendered words from each job ad. Column (6) in Table 7 presents the results. We note that the estimated coefficient remains similar to the baseline estimate, ruling out the possibility that our results are driven by high-frequency job-specific gendered words.

These three sets of tests above lend support that the estimates capture changes in the language used in job ads instead of changes in the nature of the job advertised. Moreover, the evidence on hiring preference from the experiment does not depend on gendered language as a measure, mitigating potential concerns on measurement issues.

3.5.2 Robustness checks

We conduct extensive robustness checks for the baseline findings in Tables 2 and 4, with the results summarized in Figure 4.

Alternative sampling criteria. In our main analysis, we keep all job posting observations with non-missing data over the period 2014–2020. However, we need to address concerns that the limited coverage of LinkUp in 2014 and 2015 might bias the estimation, and that the COVID pandemic might affect labor demands across states and firms. We therefore use subsamples starting in 2016, or including only the pre-COVID period (before March 2020). During our sample period, a number of states in addition to California (Colorado, Maryland, Illinois, Washington, and New York) have also introduced board gender diversity interventions. Although these interventions differ markedly in their scope and strength compared

to SB 826 in California, we drop these states in robustness checks as a precautionary measure.³¹ We also remove firms with a very small or large board size (below the 1st percentile or above the 99th percentile). Since the length of job ads varies substantially, we exclude extremely long or short ads to eliminate the effect of such outliers. Figure 4 compares the estimated coefficients using these alternative samples with the baseline estimate. The estimates are similar in both the correlational analysis (Panel A) and the DID analysis using SB 826 (Panel B).

There are two additional issues concerning the applicability of SB 826. In recent years, some California firms moved their headquarters out of the state. To address the measurement error associated with a CA firm’s exposure to SB 826, we compare firms’ headquarters locations before and after SB 826 and drop those relocated over our estimation period. The results in the DID analysis remain highly similar (Figure 4 Panel B). SB 826 only applies to CA firms listed on a major US stock exchange. Our sample consists of large public firms in the US and over 99.5% of the job postings are by firms listed on NYSE, AMEX, or NASDAQ. Nevertheless, we repeat the DID analysis by using only firms listed on a major US stock exchange, and the results barely change (Figure 4 Panel B).

Alternative regression specifications. We also carry out robustness checks using alternative regression specifications. Our results are robust to excluding all control variables and controlling for occupation-specific shocks (at either the two-digit or six-digit ONET-SOC code level) or even shocks at the occupation $\times$ job zip level. However, mandating female directors could have the unintended consequence of causing firms to change their business operations and job composition, with the outcome of limiting female-dominant occupations. To rule out this possibility, we add firm $\times$ occupation (at the six-digit ONET-SOC code level) fixed effects. Figure 4 contrasts the estimated coefficients from these alternative specifications with the baseline estimate. The estimates are quite similar in both panels.

³¹The interventions in Maryland, Illinois, and New York only imposed disclosure requirements. The state of Washington took a less aggressive “comply or explain” approach. The intervention in Colorado was a non-binding resolution, encouraging more gender diverse boards.

Augmenting the feminine and masculine word list. One concern about our bag-of-words approach is that the gendered word list from Gaucher et al. (2011) may have become outdated. We thus augment the original word list with a gendered word list provided by Ongig, an industry leader committed to eliminating biased job ads. We recalculate our key measure using the augmented list and find similar results, as shown in the last row of each panel in Figure 4.

Random sampling. In our US job-level analysis, we use a randomly drawn subsample that is one fifteenth of the full sample. Although randomly drawing nearly two million observations from a sample of 30 million observations likely loses no generality, we also use other randomly drawn subsamples to confirm that our results are not driven by chance. Figure D11 presents the distribution of the estimated coefficients for 1,000 randomly drawn subsamples, each corresponding to one fifteenth of the full sample. In both the correlational analysis (Panel A) and the DID analysis using SB 826 (Panel B), we find that the estimated coefficients center around the baseline estimate and are far away from zero.

4 Other Female Labor Market Outcomes

So far, we have provided evidence on firms’ reduced demand for female labor in response to external pressure to improve board gender diversity. One may argue that through perceived cultural changes in firms subject to board gender quotas, affected firms might experience an increase in the supply of female labor, rendering no need for them to use female-friendly language in job ads to attract female applicants. To assess whether this argument has any merit, we examine whether the pushback against external pressure to increase board gender diversity also manifests in actual (poor) workplace treatment of female employees and (low) female employment in the corporate sector. Moreover, while we study the board quota in California; outside the US, ten European countries have adopted similar quotas since 2003, do we expect to observe a similar pattern in Europe? In this section, we provide additional evidence using employee reviews and ratings of female-friendly amenities and benefits from

Glassdoor, as well as aggregate female employment in the US and Europe.

4.1 Employee reviews of female-friendly amenities

Sockin (2022) identifies 50 amenities using textual analysis of Glassdoor employee reviews. For our purposes, we focus on eight amenities (5-Paid time off, 10-Work-life balance, 11-Hours, 12-Work schedule, 16-Teleworking, 19-Respect/abuse, 21-Support, and 31-Diversity/inclusion) that are important for female employees in the workplace. These “female-friendly” amenities not only reflect firm-level policies but also are shaped by supervisors and coworkers who interact with female employees. If quotas trigger backlash, we would expect the work environment becomes more hostile for female employees and they write more negative reviews about (lack of) workplace accommodations. Table D5 in the Online Appendix provides the word list of the eight amenities (anchor words and top 20 relevant words) that we use to obtain a word count of different amenities in the pros and cons sections of Glassdoor reviews. Table D6 Panel A in the Online Appendix provides the summary statistics for Glassdoor amenity measures.

We empirically test the above conjecture and the results are presented in Table 8. Panel A examines the relationship between a firm’s share of female directors and its employee reviews of female-friendly amenities. We find largely negative and significant associations between a firm’s share of female directors and its employee reviews of female-friendly amenities when the outcome variables are based on reviews in the pros section (columns (1) and (2)) and the difference between pros and cons sections (columns (5) to (8)). However, the relationship turns positive when we use employee reviews of amenities in the cons section (columns (3) and (4)). The evidence thus consistently suggests less favorable reviews of female-friendly amenities when firms have more female directors. Panel B of Table 8 further explores the effect of SB 826. We find robust evidence that the bill resulted in a substantial drop (increase) in positive (negative) reviews of female-friendly amenities in California firms. Although not all estimated coefficients are significant across the eight specifications, the directions of the estimates paint a consistent picture. Glassdoor also collects employee ratings of various

aspects of a firm’s operations on a scale of 1-5. For our purposes, we focus on both a firm’s overall rating and its specific ratings for compensation and benefits, culture and values, and work-life balance – benefits that likely matter (more) for female employees in the workplace. Table D7 provides strong evidence that the bill resulted in a significant drop in employee ratings of these benefits in California firms.³²

The above evidence indicates a shift toward a more hostile work environment for female employees after the passage of SB 826 – there were more negative reviews of workplace accommodations that are important for females. This could be interpreted as another avenue through which firms push back board gender quotas, suggesting that backlash spans from labor demand preferences to workplace practices.

4.2 Aggregate female employment in the US

In this section, we provide additional evidence on female employment in the US after the passage of SB 826. We use IPUMS CPS, which harmonizes microdata from the monthly U.S. labor force survey, or the Current Population Survey (CPS), over the period 2014–2020. The data set contains demographic and rich employment information, including labor force status, of approximately 65,000 households every month. The survey separates workers into different classes: self-employed, not incorporated; self-employed, incorporated; private, for profit; private, nonprofit; federal/state/local government employee; and unpaid family worker. We keep employed persons (age 15-64) and aggregate individual responses to compute the share of female employment for each state, worker class, and month.

We use a triple difference specification that compares the gap in the change in female employment around SB 826 between the more exposed worker class (private for profit) and less exposed classes in CA with the gap outside CA. Our variable of interest is the triple interaction term, $Flag_CA \times Post_2018m9 \times Private_For_Profit$. Table 9 presents the results. Columns (1) and (2) include all worker classes, and columns (3) and (4) include only the

³²The anonymous nature of reviews on Glassdoor allows employees to rate and/or discuss their (former) employers without fear of retaliation or repercussions. It is thus unlikely that our results are driven by female directors encouraging employees to write about workplace poor gender-related practices and misbehaviors.

following four worker classes: self-employed, not incorporated; self-employed, incorporated; private, for profit; and private, nonprofit. Across all specifications, we find that the share of female employment in the private sector in California experiences a significantly larger drop compared to that outside the private sector outside California. In terms of economic significance, using column (2) as an example, we show that the share of female employment in the private sector in California experienced a 1.1 percentage-point drop as a result of SB 826, all else being equal.³³ Note that we include state-year-month fixed effects in columns (2) and (4), so any other state-level shocks in California are unlikely to drive the results unless they impact the private for-profit worker class differently from other worker classes exactly around 2018.

The above evidence is consistent with our main findings of reduced female labor demand after the passage of SB 826, pointing to the possibility that backlash extends to the ultimate labor market outcome – employment. While gendered language in job ads is one tactic we emphasize in shaping female labor market outcomes due to data and measurement considerations, we provide suggestive evidence that other tactics or measures could also be at play, such as hostile work environments for female employees or discriminatory practices in screening, promotion, and termination.³⁴

4.3 External validity: Aggregate female employment in Europe

In this section, we exploit the staggered adoptions of board gender quotas across European countries at different points in time to study the effect of board gender quotas on female labor market outcome in very different institutional frameworks.

We obtain information on board gender quotas, including passage dates, enforcement methods, and scope, from the legislature offices. Based on the strength of the enforcement, the wording of the relevant law, and the institutional context in which the law is embedded,

³³This drop is larger than the magnitude (a 0.5-to-1 percentage-point drop) inferred in Section 3.3, which is based on the estimated drop in the share of feminine words in job ads. In practice, pushback can happen at different stages of the hiring process, such as candidate screening and interview evaluations, contributing to the larger drop in female employment observed in Table 9.

³⁴Additional analyses in Online Appendix D.3 focus on gender-related labor violations.

Mensi-Klarbach and Seierstad (2020) classify these board gender quotas into soft (e.g., Spain and Iceland) versus non-soft (e.g., Norway and France). Table D9 in the Online Appendix lists quotas in Europe introduced over the period 2003–2017. We obtain data on the share of female employment (men and women, age 15–64) from the Labor Force Survey (LFS) of Eurostat in each country-industry-year from 2008 to 2020. Table A2 in the Online Appendix provides the list of variables, their definitions, and data sources.

Panel A of Table 10 presents the results based on a staggered DID specification. We consistently find a negative and significant effect of a country’s introduction of a board gender quota on its share of female employment in various specifications from columns (1) to (6).³⁵ Figure D10 plots the dynamic effects of such quotas. We observe a clear post-quota downward trend and in the pre-quota years the estimated coefficients fluctuate around zero and are always insignificant.

Do all quotas have the same effect on female labor market participation? We divide all ten quotas into two regimes depending on whether a quota is enforced by the state as opposed to a voluntary target. Panel B of Table 10 presents the regression results. We show that the above effect is stronger when we focus on hard-/medium-hard quotas with strong and clear enforcement mechanisms. In contrast, the effect is close to zero when we focus on soft quotas. This contrast is consistent with backlash – strict quotas tend to provoke pushback compared to voluntary codes and this can happen across countries that differ in their cultural and institutional background.

5 Conclusion

In this paper, by applying computational linguistic methods to job ads in the US over the period 2014–2020, we uncover an intriguing negative correlation between a firm’s share of female directors and its demand for female labor. To understand this negative correlation,

³⁵Our findings remain when we cluster standard errors at the industry level or assign equal weight to each country-industry-year observation, as shown in columns (1) and (2) of Table D10. We also follow methodologies in Borusyak et al. (2021) and Callaway and Sant’Anna (2021) to re-estimate the average treatment effect in light of the recent criticism of two-way fixed effects regressions in staggered settings. We find similar results, see columns (3) and (4) of Table D10.

we first examine cross-sectional variations in the amount of external pressure firms face for board gender diversity. We find that the negative association is accentuated in firms with a higher ownership by the three largest asset managers – State Street, BlackRock, and Vanguard. We then rely on California’s Senate Bill 826 passed in 2018 that imposed board gender quotas on locally headquartered firms. We show that the bill resulted in a significant drop in California firms’ demand for female labor. Finally, we run a randomized controlled experiment to help establish a direct link between subjects’ exposure to board gender quotas and their gender-specific hiring preferences as well as delineate the economic mechanisms. We show that psychological reactance driven by the top-down approach of quotas, as well as the perceived misalignment between quotas and gender norms, are the likely culprit. Supplemental evidence on other labor market outcomes reveals less favorable reviews of workplace accommodations that are important for female employees after the passage of SB 826. Moreover, the negative outcome extends to female employment, which dropped significantly in California after SB 826 and in ten European countries after their adoption of board gender quotas. We conclude that board gender quotas/policies may backfire, resulting in worse labor market outcomes for women.

The findings of our paper highlight some complex welfare implications of gender equality policies that might have long-run distributive effects when (targeted) firms fail to internalize the social cost of such policies. More broadly, our findings suggest that when governments introduce laws that deviate from social norms, the potential for pushback from firms’ stakeholders, such as their employees, should be considered when evaluating laws’ overall effectiveness.

References

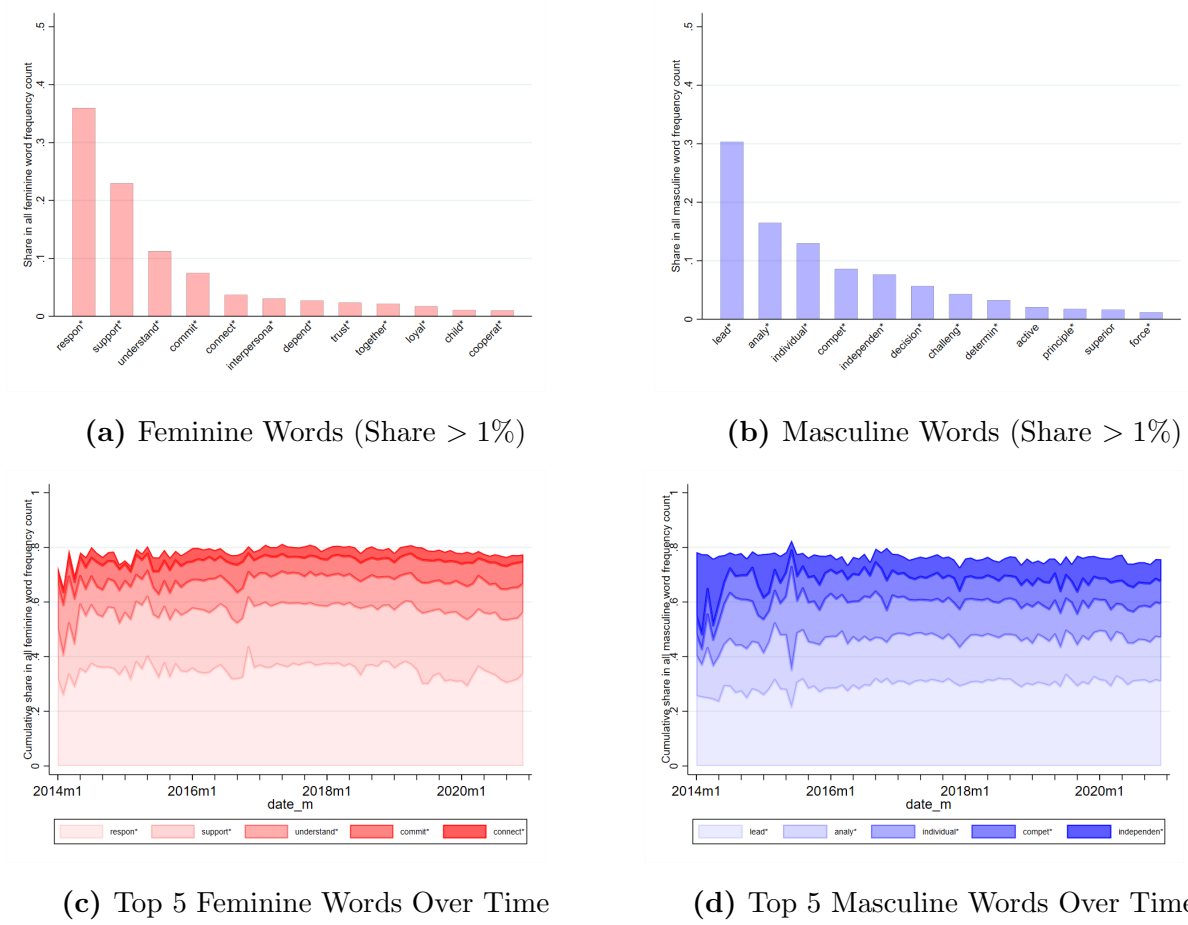
- Acemoglu, D. and J. D. Angrist (2001). Consequences of Employment Protection? the Case of the Americans with Disabilities Act. *Journal of Political Economy* 109(5), 915–957.
- Acemoglu, D. and M. O. Jackson (2017). Social Norms and the Enforcement of Laws. *Journal of the European Economic Association* 15(2), 245–295.
- Adams, R. B. and D. Ferreira (2009). Women in the Boardroom and Their Impact on Governance and Performance. *Journal of Financial Economics* 94(2), 291–309.

- Ahern, K. R. and A. K. Dittmar (2012). The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation. *Quarterly Journal of Economics* 127(1), 137–197.
- Ang, D. (2019). Do 40-year-old Facts Still matter? Long-run Effects of Federal Oversight under the Voting Rights Act. *American Economic Journal: Applied Economics* 11(3), 1–53.
- Babina, T., A. Fedyk, A. X. He, and J. Hodson (2021). Artificial Intelligence, Firm Growth, and Product Innovation. SSRN Scholarly Paper ID 3651052, Social Science Research Network, Rochester, NY.
- Bailey, M. J., T. Helgerman, and B. A. Stuart (2022). How the 1963 Equal Pay Act and 1964 Civil Rights Act Shaped the Gender Gap in Pay. University of California Los Angeles Working Paper.
- Bem, S. L. (1974). The Measurement of Psychological Androgyny. *Journal of Consulting and Clinical Psychology* 42(2), 155–162.
- Benabou, R. and J. Tirole (2011). Laws and Norms. Technical report, National Bureau of Economic Research.
- Bertrand, M. (2020). Gender in the Twenty-first Century. In *AEA Papers and proceedings*, Volume 110, pp. 1–24.
- Bertrand, M., S. E. Black, S. Jensen, and A. Lleras-Muney (2019). Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway. *Review of Economic Studies* 86(1), 191–239.
- Blanken, I., N. Van De Ven, and M. Zeelenberg (2015). A Meta-Analytic Review of Moral Licensing. *Personality and Social Psychology Bulletin* 41(4), 540–558.
- Borusyak, K., X. Jaravel, and J. Spiess (2021). Revisiting Event Study Designs: Robust and Efficient Estimation. *arXiv preprint arXiv:2108.12419*.
- Bowles, S. and S. Polania-Reyes (2012). Economic Incentives and Social Preferences: Substitutes or Complements? *Journal of Economic Literature* 50(2), 368–425.
- Brehm, J. W. (1966). A Theory of Psychological Reactance.
- Brehm, S. S. and J. W. Brehm (2013). *Psychological Reactance: A Theory of Freedom and Control*. Academic Press.
- Callaway, B. and P. H. Sant’Anna (2021). Difference-in-differences with Multiple Time Periods. *Journal of Econometrics* 225(2), 200–230.
- Cohen, A., T. Karelitz, T. Kricheli-Katz, S. Pumpian, and T. Regev (2023). Gender-neutral Language and Gender Disparities. Technical report, National Bureau of Economic Research.
- Cooter, R. (1998). Expressive Law and Economics. *The Journal of Legal Studies* 27(S2), 585–607.
- Cullen, Z. B. and B. Pakzad-Hurson (2021). Equilibrium Effects of Pay Transparency. Technical report, National Bureau of Economic Research.

- Del Carpio, L. and T. Fujiwara (2023). Do Gender-Neutral Job Ads Promote Diversity? Experimental Evidence from Latin America’s Tech Sector. Technical report, National Bureau of Economic Research.
- Deming, D. and L. B. Kahn (2018). Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals. *Journal of Labor Economics* 36(S1), S337–S369.
- Deschamps, P. (2023). Gender Quotas in Hiring Committees: A Boon or a Bane for Women? *Management Science*.
- Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2018). Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.
- Dobbin, F. and A. Kalev (2016). Why Diversity Programs Fail. *Harvard Business Review* 94(7), 14.
- Eckbo, B. E., K. Nygaard, and K. S. Thorburn (2022). Valuation Effects of Norway’s Board Gender-quota Law Revisited. *Management Science* 68(6), 4112–4134.
- 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(1), 109–128.
- Gertsberg, M., J. Mollerstrom, and M. Pagel (2021). Gender Quotas and Support for Women in Board Elections. Technical report, National Bureau of Economic Research.
- Giannetti, M. and T. Y. Wang (2023). Public Attention to Gender Equality and Board Gender Diversity. *Journal of Financial and Quantitative Analysis* 58(2), 485–511.
- Gopal, B. (2022). How Do Firms Respond to Gender Quotas? Evidence From California’s Senate Bill 826.
- Gormley, T. A., V. K. Gupta, D. A. Matsa, S. C. Mortal, and L. Yang (2023). The Big Three and Board Gender Diversity: The Effectiveness of Shareholder Voice. *Journal of Financial Economics* 149(2), 323–348.
- Greene, D., V. J. Intintoli, and K. M. Kahle (2020). Do board gender quotas affect firm value? Evidence from California Senate Bill No. 826. *Journal of Corporate Finance* 60, 101526.
- He, L. and T. M. Whited (2023). Underrepresentation of Women CEOs. *Available at SSRN 4615373*.
- Hershbein, B. and L. B. Kahn (2018). Do Recessions Accelerate Routine-biased Technological Change? Evidence from Vacancy Postings. *American Economic Review* 108(7), 1737–72.
- Hwang, S., A. Shivdasani, and E. Simintzi (2021). Mandating Women on Boards: Evidence from the United States. SSRN Scholarly Paper ID 3265783, Social Science Research Network, Rochester, NY.
- Jin, Y., A. Bhatia, and D. Wanvarie (2021). Seed Word Selection for Weakly-supervised Text Classification with Unsupervised Error Estimation. *arXiv preprint arXiv:2104.09765*.
- Levi, M., K. Li, and F. Zhang (2014). Director Gender and Mergers and Acquisitions.

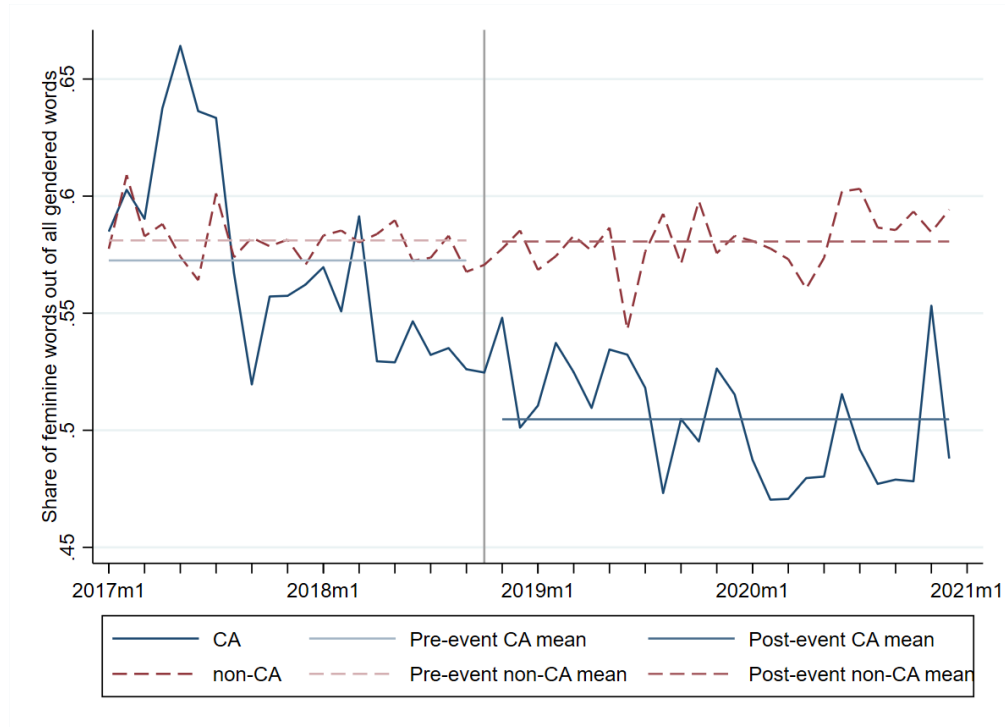
- Journal of Corporate Finance* 28, 185–200.
- Li, K., F. Mai, R. Shen, and X. Yan (2021). Measuring Corporate Culture using Machine Learning. *The Review of Financial Studies* 34(7), 3265–3315.
- List, J. A. and F. Momeni (2021). When Corporate Social Responsibility Backfires: Evidence from a Natural Field Experiment. *Management Science* 67(1), 8–21.
- Maida, A. and A. Weber (2022). Female Leadership and Gender Gap within Firms: Evidence from an Italian Board Reform. *ILR Review* 75(2), 488–515.
- Matsa, D. A. and A. R. Miller (2011). Chipping Away at the Glass Ceiling: Gender Spillovers in Corporate Leadership. *American Economic Review* 101(3), 635–639.
- Matsa, D. A. and A. R. Miller (2013). A Female Style in Corporate Leadership? Evidence from Quotas. *American Economic Journal: Applied Economics* 5(3), 136–169.
- Mensi-Klarbach, H. and C. Seierstad (2020). Gender Quotas on Corporate Boards: Similarities and Differences in Quota Scenarios. *European Management Review* 17(3), 615–631.
- Merritt, A. C., D. A. Effron, and B. Monin (2010). Moral Self-licensing: When Being Good Frees Us to be Bad. *Social and personality psychology compass* 4(5), 344–357.
- Nygaard, K. (2011). Forced Board Changes: Evidence from Norway. SSRN Scholarly Paper ID 1793227, Social Science Research Network, Rochester, NY.
- Posner, E. A. (2009). *Law and Social Norms*. Harvard University Press.
- Rosenkrantz, P., S. Vogel, H. Bee, I. Broverman, and D. M. Broverman (1968). Sex-role Stereotypes and Self-concepts in College Students. *Journal of Consulting and Clinical Psychology* 32(3), 287–295.
- Sanh, V., L. Debut, J. Chaumond, and T. Wolf (2019). Distilbert, a Distilled Version of bert: Smaller, Faster, Cheaper and Lighter. *arXiv preprint arXiv:1910.01108*.
- Shoss, M. K. (2017). Job Insecurity: An Integrative Review and Agenda for Future Research. *Journal of management* 43(6), 1911–1939.
- Sittenthaler, S., E. Traut-Mattausch, C. Steindl, and E. Jonas (2015). Salzburger State Reactance Scale (ssr Scale). *Zeitschrift für Psychologie*.
- Sockin, J. (2022). Show Me the Amenity: Are Higher-Paying Firms Better All Around?
- Sunstein, C. R. (1996). Social Norms and Social Roles. *Colum. L. Rev.* 96, 903.
- Tate, G. and L. Yang (2015). Female leadership and gender equity: Evidence from plant closure. *Journal of Financial Economics* 117(1), 77–97.
- von Meyerinck, F., A. Niessen-Ruenzi, M. Schmid, and S. Davidoff Solomon (2021). As California Goes, so Goes the Nation? Board Gender Quotas and Shareholders’ Distaste of Government Interventions. SSRN Scholarly Paper ID 3303798, Rochester, NY.
- Wheaton, B. (2022). Laws, Beliefs, and Backlash. The University of California Los Angeles Working Paper.
- Zhang, M., K. N. Jensen, S. D. Sonniks, and B. Plank (2022). Skillspace: Hard and Soft Skill Extraction from English Job Postings. *arXiv preprint arXiv:2204.12811*.

Figure 1: Frequency Counts of Top Feminine & Masculine Words



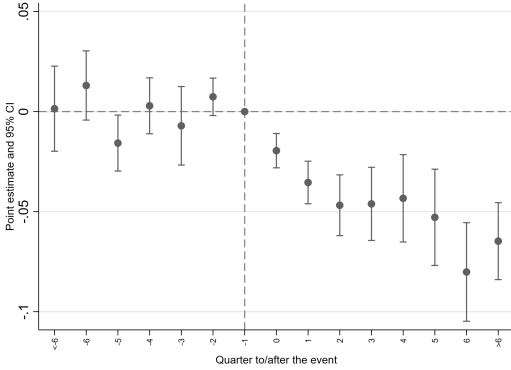
This figure shows frequency counts of the most frequent feminine and masculine words in job ads from the Gaucher et al. (2011) list. Figure 1a (Figure 1b) shows the share of the most frequent feminine (masculine) words out of all feminine (masculine) words over the sample period (2014-2020). Figure 1c (Figure 1d) shows the cumulative share of the top 5 feminine (masculine) words out of all feminine (masculine) words over time from 2014 to 2020. The asterisk denotes the acceptance of all letters, hyphens, or numbers following its appearance.

**Figure 2: California Statute and Gendered Language Usage in Job Ads
– Dynamics (Univariate)**

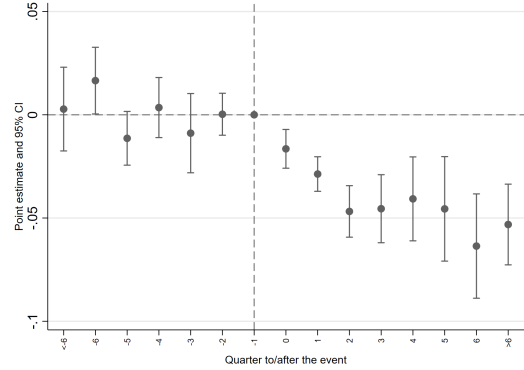


This figure shows the monthly movement in the average share of feminine words across all jobs posted by CA headquartered and non-CA headquartered firms around the California board gender diversity statute (SB 826). The share of feminine words out of all gendered words in a job ad is shown on the y-axis. The months from January 2017 to December 2020 are shown on the x-axis. The vertical line indicates the event date (the passage of the bill in September 2018).

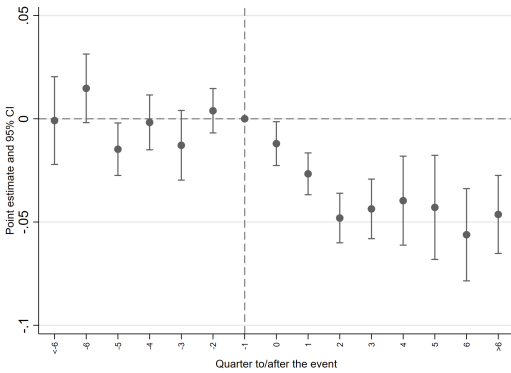
**Figure 3: California Statute and Gendered Language Usage in Job Ads
– Dynamics (Regression)**



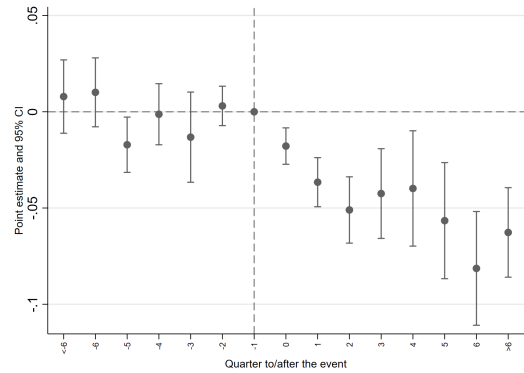
(a) Column (1) of Table 4



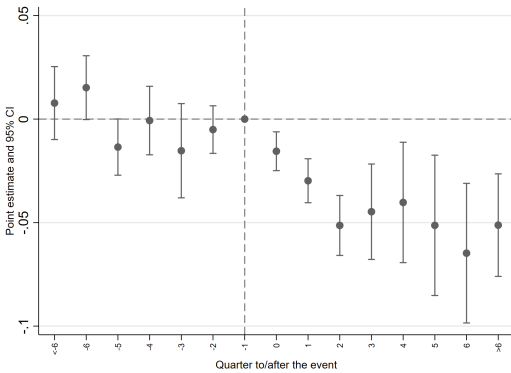
(b) Column (2) of Table 4



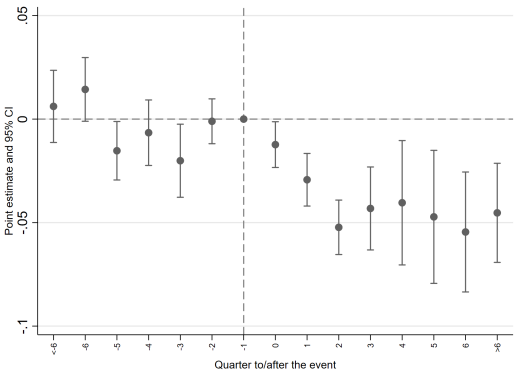
(c) Column (3) of Table 4



(d) Column (4) of Table 4



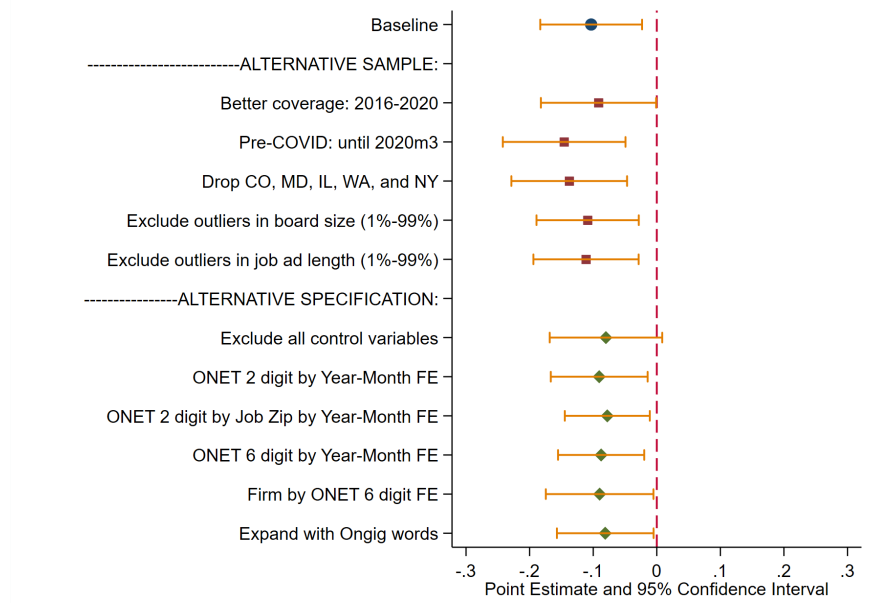
(e) Column (5) of Table 4



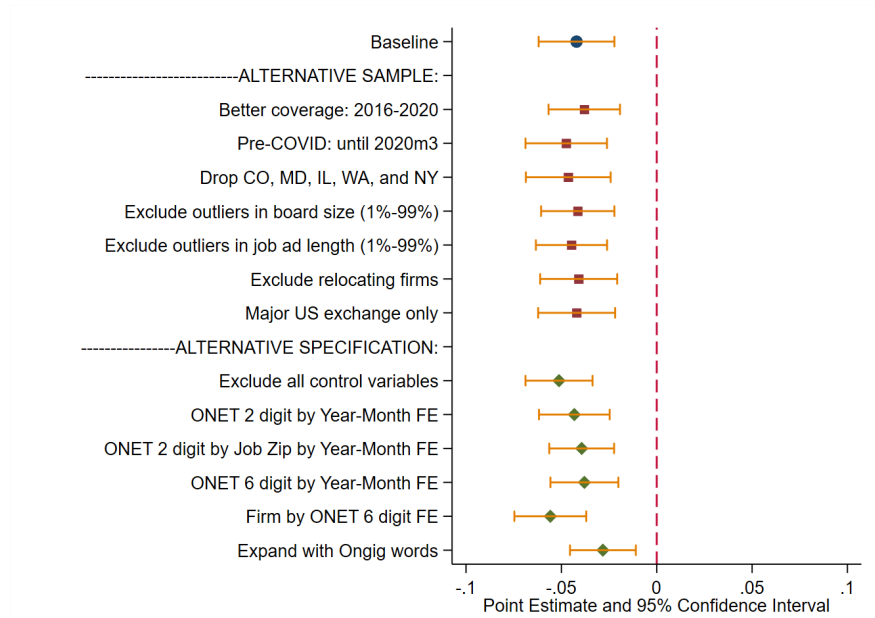
(f) Column (6) of Table 4

This figure shows the dynamic effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads. Each figure corresponds to a specification used in the six columns of Table 4, which include different sets of controls and fixed effects. Quarter -1 is the first quarter of 2018 – by the second quarter of 2018, there were high expectations of the Bill's passing – and is the omitted category. Every three months in a corresponding quarter are grouped together to reduce estimation error.

Figure 4: Robustness - Alternative Samples and Specifications



(a) Correlational Analysis



(b) CA DID

This figure shows the robustness of the job-level regression results to using alternative samples and specifications. Both the relationship between the share of female directors and gendered language usage in job ads (Figure 4a) and the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads (Figure 4b) are examined. The regression specifications are the same as those in column 6 of Table 2 and Table 4. The top row of each panel shows the baseline estimate in column 6 of Table 2 and Table 4.

Table 1: Summary Statistics**Panel A: US Job Level**

Variable	N	Mean	Std. Dev.	Median	p1	p99
%Fem words	1,857,301	0.58	0.27	0.57	0.00	1.00
#Fem words	1,857,156	4.51	3.84	4.00	0.00	17.00
#Mas words	1,857,156	3.71	4.04	2.00	0.00	18.00
#Words	1,857,156	425.87	245.96	389.00	79.00	1139.00
Job zone	1,296,275	2.74	1.01	2.00	1.00	5.00
Mean wage	1,359,316	28.48	16.30	22.74	12.36	74.49
Entry education	1,360,709	2.01	2.02	1.00	0.00	7.00
STEM	1,360,709	0.07	0.25	0.00	0.00	1.00

Panel B: US Firm-Year Level

Variable	N	Mean	Std. Dev.	Median	p1	p99
#Job	14,601	2,223	14,050	214	1	36,956
%Fem words [mean]	14,601	0.547	0.135	0.541	0.205	0.91
#Dir	14,601	9.07	2.35	9.00	5.00	15.00
%Fem dir	14,601	0.18	0.12	0.17	0.00	0.50
Female CEO	8,888	0.07	0.25	0.00	0.00	1.00
Female CFO	8,888	0.13	0.34	0.00	0.00	1.00
Female COO	8,888	0.02	0.14	0.00	0.00	1.00
Firm Size (log assets)	8,888	8.40	1.74	8.32	4.72	12.80
Cash holdings	8,888	0.14	0.15	0.08	0.00	0.68
Tangibility	8,888	0.23	0.23	0.14	0.00	0.88
Tobin's Q	8,888	2.16	1.76	1.59	0.77	9.61
Leverage	8,888	0.28	0.23	0.26	0.00	0.95
ROA	8,888	0.11	0.12	0.11	-0.18	0.37
Capex	8,888	0.04	0.04	0.02	0.00	0.20

The table presents summary statistics for the main variables used in our analysis. Panel A summarizes measures at the US job level based on a random 1/15 of the full sample of job ads. Panel B summarizes variables at the firm-year level.

Table 2: Female Directors and Gendered Language Usage in Job Ads

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)					
%Fem dir	-0.071*	-0.085**	-0.091**	-0.076*	-0.094**	-0.103**
	[0.040]	[0.039]	[0.034]	[0.044]	[0.045]	[0.040]
Log (#Words)				-0.037***	-0.026***	-0.025**
				[0.013]	[0.009]	[0.010]
Female CEO				-0.061***	-0.047**	-0.041**
				[0.021]	[0.023]	[0.018]
Female CFO				-0.023	-0.023	-0.024**
				[0.014]	[0.014]	[0.012]
Female COO				0.014	0.019	0.008
				[0.025]	[0.023]	[0.020]
Firm size	0.016	0.013	0.01	0.013	0.01	0.007
	[0.012]	[0.011]	[0.009]	[0.013]	[0.012]	[0.010]
Cash holdings	-0.101***	-0.090***	-0.075**	-0.155***	-0.137***	-0.121***
	[0.032]	[0.032]	[0.029]	[0.038]	[0.038]	[0.034]
Tangibility	-0.129	-0.152*	-0.142*	-0.202**	-0.222**	-0.209**
	[0.078]	[0.087]	[0.079]	[0.079]	[0.091]	[0.084]
Tobin's Q	0.007**	0.008***	0.006**	0.008**	0.010**	0.007**
	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]	[0.003]
Leverage	0.008	0.015	0.01	-0.013	-0.001	-0.008
	[0.020]	[0.020]	[0.018]	[0.022]	[0.024]	[0.021]
ROA	-0.148**	-0.149***	-0.112***	-0.224***	-0.222***	-0.183***
	[0.058]	[0.053]	[0.042]	[0.071]	[0.067]	[0.059]
Capex	0.219	0.237	0.311**	0.445*	0.463*	0.496**
	[0.176]	[0.174]	[0.148]	[0.233]	[0.231]	[0.213]
Firm FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job State $\times$ Year-Month FE	YES	YES	-	YES	YES	-
Job Zip $\times$ Year-Month FE	NO	NO	YES	NO	NO	YES
Occupation (ONET 6 digit) FE	NO	YES	YES	NO	YES	YES
Obs	1,802,618	1,802,618	1,802,618	1,629,816	1,629,816	1,629,816
Adj R2	0.355	0.414	0.429	0.361	0.419	0.434

This table shows the relationship between the share of female directors and gendered language usage in job ads. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. The outcome variable is the share of feminine words out of all gendered words in a job ad. Robust standard errors double clustered at the firm and state (state of a job) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

**Table 3: Female Directors and Gendered Language Usage in Job Ads
– Heterogeneity: “The Big Three”, Partisanship, and Female Leadership**

	(1)	(2)	(3)	(4)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)			
%Fem dir	-0.067 [0.045]	-0.037 [0.064]	-0.127*** [0.046]	-0.119*** [0.043]
%Fem dir $\times$ High Big3 ownership	-0.118*** [0.044]			
%Fem dir $\times$ >50% Republican contribution		-0.203*** [0.052]		
%Fem dir $\times$ Female Top3 exec			0.113** [0.055]	
%Fem dir $\times$ Female COO				0.438*** [0.122]
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES
Occupation (ONET 6-digit) FE	YES	YES	YES	YES
Obs	1,143,596	767,802	1,143,596	1,143,596
Adj R2	0.463	0.435	0.463	0.463

This table shows heterogeneities in the relationship between the share of female directors and gendered language usage in job ads across firms. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. The outcome variable is the share of feminine words out of all gendered words in a job ad. In column (1), female director share (% Fem dir) is interacted with a dummy indicating firms that have high Big Three ownership (top quartile). The Big Three institutional investors are Vanguard, BlackRock, and State Street. In column (2), the interaction term is a dummy indicating whether more than 50% of a firm’s political contributions go to Republican candidates. In column (3), the interaction term is a dummy indicating whether a firm has at least one female top executive (CEO, CFO, or COO). In column (4), the interaction term is a dummy indicating whether a firm has a female COO. All regressions include an interaction between female director share (% Fem dir) and job zone as an control variable. Robust standard errors double clustered at the firm and state (state of a job) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 4: Effect of California Statute on Gendered Language Usage in Job Ads

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)					
Flag_CA $\times$ Post_2018m9	-0.053*** [0.008]	-0.046*** [0.008]	-0.041*** [0.008]	-0.054*** [0.010]	-0.047*** [0.011]	-0.042*** [0.010]
Controls	NO	NO	NO	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job State $\times$ Year-Month FE	YES	YES	-	YES	YES	-
Job Zip $\times$ Year-Month FE	NO	NO	YES	NO	NO	YES
Occupation (ONET 6 digit) FE	NO	YES	YES	NO	YES	YES
Obs	1,816,477	1,816,477	1,816,477	1,642,550	1,642,550	1,642,550
Adj R2	0.354	0.414	0.429	0.36	0.419	0.434

This table shows the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. Flag_CA is a dummy that equals 1 for public firms headquartered in California, and 0 otherwise. Post_2018m9 is a dummy that equals 1 after September 30, 2018, and 0 otherwise. The outcome variable is the share of feminine words out of all gendered words in a job ad. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

**Table 5: Effect of California Statute on Gendered Language Usage in Job Ads
– Treatment Intensity**

	(1)	(2)	(3)	(4)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)			
Subsamples	Met the threshold in SB 826 before 2018m9?		Change in female directors around 2018m9?	
	Just below the threshold	At the threshold	Increase	No Increase
Flag_CA $\times$ Post_2018m9	-0.041*** [0.008]	0.000 [0.015]	-0.049*** [0.010]	-0.006 [0.015]
p-value (diff)	0.002		0.008	
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES
Occupation (ONET 6 digit) FE	YES	YES	YES	YES
Obs	377,637	630,272	512,965	1,082,906
Adj R2	0.545	0.357	0.463	0.448

This table shows heterogeneities in the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads across firms that are differentially treated. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. Columns (1) and (2) compare firms just below the threshold set by SB 826 (one female director short) and firms at the threshold before the treatment date. Columns (3) and (4) compare firms that experience an increase in their number of female directors after the treatment date versus firms without an increase. Flag_CA is a dummy that equals 1 for public firms headquartered in California, and 0 otherwise. Post_2018m9 is a dummy that equals 1 after September 30, 2018, and 0 otherwise. The outcome variable is the share of feminine words out of all gendered words in a job ad. p-val (diff) reports the p-value for testing the difference between the coefficients yielded by different subsamples. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 6: Experimental Evidence on Hiring Preference and Mechanisms**Panel A: Comparing Treated and Control Subjects**

	(1)	(2)	(3)	(4)	(5)	(6)
			Mechanisms			
Dep. Var.	Manipulation	Hiring Preference	Reactance	Norm Violation	Moral Licensing	Insecurity
Treatment	1.779*** [0.179]	0.512** [0.250]	1.311*** [0.205]	0.817*** [0.249]	0.240 [0.203]	0.344 [0.223]
Controls	YES	YES	YES	YES	YES	YES
Dep. Var. Mean	5.047	2.992	2.891	2.690	4.112	2.205
Obs	199	199	199	199	199	199
Adj R2	0.335	0.104	0.320	0.182	0.047	0.206

Panel B: Heterogeneity - Political View and Gender

	(1)	(2)	(3)	(4)	(5)	(6)
			Mechanisms			
Dep. Var.	Hiring Preference		Reactance		Norm Violation	
Treatment $\times$ Political View	0.298** [0.139]		0.233** [0.110]		0.240* [0.133]	
Treatment $\times$ Male		0.732 [0.498]		1.266*** [0.410]		1.111** [0.497]
Treatment	-0.531 [0.535]	0.150 [0.324]	0.496 [0.411]	0.683** [0.289]	-0.022 [0.450]	0.267 [0.350]
Controls	YES	YES	YES	YES	YES	YES
Dep. Var. Mean	2.992	2.992	2.891	2.891	2.690	2.690
Obs	199	199	199	199	199	199
Adj R2	0.126	0.111	0.335	0.357	0.194	0.203

Panel A reports the treatment effect estimated from the experiment. The outcome variable is seven-point Likert scale for the manipulation questions in column (1), for questions on hiring preference (favoring men) in column (2), and for questions related to different mechanisms in columns (3) to (6). Panel B reports heterogeneous treatment effects based on political view (1-7 from liberal to conservative) and gender. The following control variables are included: age, gender, ethnicity, employment status, years of work experience, political view, type of device used for the survey, and the hour of the survey. The unit of observation is a subject. The sample includes 199 subjects who passed the attention test. Treatment is a dummy that equals 1 for the treatment group and 0 for the control group. Robust standard errors adjusted for heteroskedasticity are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 7: Female Directors and Gendered Language Usage in Job Ads
Gendered Job Ads vs. Gendered Jobs

Panel A: Correlation with Female Directors

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)					
	Same job title			Drop gendered jobs		Drop job-specific gendered words
%Fem dir	-0.089* [0.053]	-0.085* [0.046]	-0.083* [0.049]	-0.084** [0.039]	-0.088** [0.041]	-0.117** [0.045]
Job Title FE (full)	YES	-	NO	NO	NO	NO
Job Title FE (first 3 words)	NO	YES	-	NO	NO	NO
Job Title $\times$ Occupation $\times$ Firm FE	NO	NO	YES	NO	NO	NO
Firm FE	YES	YES	-	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Occupation (ONET 6 digit) FE	NO	NO	-	YES	YES	YES
Obs	1,641,670	1,641,670	1,641,670	1,364,478	1,019,866	1,639,456
Adj R2	0.714	0.665	0.778	0.446	0.431	0.399

Panel B: Effect of California Statute

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)					
	Same job title			Drop gendered jobs		Drop job-specific gendered words
Flag_CA $\times$ Post_2018m9	-0.049*** [0.011]	-0.047*** [0.010]	-0.075*** [0.011]	-0.034*** [0.008]	-0.042*** [0.007]	-0.029** [0.012]
Job Title FE (full)	YES	-	NO	NO	NO	NO
Job Title FE (first 3 words)	NO	YES	-	NO	NO	NO
Job Title $\times$ Occupation $\times$ Firm FE	NO	NO	YES	NO	NO	NO
Firm FE	YES	YES	-	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Occupation (ONET 6 digit) FE	NO	NO	-	YES	YES	YES
Obs	1,641,670	1,641,670	1,641,670	1,354,218	1,009,552	1,639,691
Adj R2	0.714	0.665	0.778	0.447	0.433	0.399

This table shows the relationship between the share of female directors and gendered language usage in job ads (Panel A) and the effect of the California board gender diversity statute (SB 826) (Panel B) on gendered language usage in job ads by comparing job ads that are similar in nature. Columns (1) to (3) include fixed effects related to job titles. Column (4) excludes jobs in occupations where the female worker share is below 25% or above 75%. Column (5) excludes jobs in occupations where the female worker share is below 40% or above 60%. Column (6) excludes the top 5 gendered words (in frequency) within each job ad when calculating the outcome variable. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. In Panel B, Flag_CA is a dummy that equals 1 for public firms headquartered in California, and 0 otherwise. Post_2018m9 is a dummy that equals 1 after September 30, 2018, and 0 otherwise. The outcome variable is the share of feminine words out of all gendered words in a job ad in both panels. Robust standard errors clustered at the firm level are in brackets in Panel A. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets in Panel B. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 8: Female Directors and Employee Reviews of Female-Friendly Amenities

Panel A: Correlation with Female Directors								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Coverage of 8 amenities in Glassdoor job reviews (word frequency)							
	Pros	Pros	Cons	Cons	Pros - Cons	Pros - Cons	Pros - Cons (weighted)	Pros - Cons (weighted)
%Fem dir	-0.242** [0.115]	-0.238* [0.132]	0.241* [0.136]	0.283 [0.183]	-0.241*** [0.088]	-0.298*** [0.114]	-0.268** [0.110]	-0.346** [0.152]
List of words for amenities	Anchor	Anchor&Top20	Anchor	Anchor&Top20	Anchor	Anchor&Top20	Anchor	Anchor&Top20
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Obs	54,814	54,814	54,814	54,814	54,814	54,814	54,814	54,814
Adj R2	0.054	0.057	0.084	0.089	0.074	0.075	0.081	0.084

Panel B: Effect of California Statute								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Coverage of 8 amenities in Glassdoor job reviews (word frequency)							
	Pros	Pros	Cons	Cons	Pros - Cons	Pros - Cons	Pros - Cons (weighted)	Pros - Cons (weighted)
Flag_CA $\times$ Post_2018m9	-0.035 [0.023]	-0.019 [0.022]	0.048* [0.028]	0.067 [0.048]	-0.043*** [0.012]	-0.041* [0.021]	-0.044** [0.020]	-0.048 [0.033]
List of words for amenities	Anchor	Anchor&Top20	Anchor	Anchor&Top20	Anchor	Anchor&Top20	Anchor	Anchor&Top20
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Obs	52,856	52,856	52,856	52,856	52,856	52,856	52,856	52,856
Adj R2	0.056	0.059	0.087	0.092	0.077	0.077	0.083	0.085

This table shows potential backlash against female employees in the workplace in terms of the provision of female-friendly amenities. The unit of observation is an employee review. The sample includes all Glassdoor reviews in English from full-time former employees whose review text in the pros and cons sections has more than 50 characters. Panel A shows the relationship between the share of female directors and employee reviews of female-friendly amenities. Panel B shows the effect of the California board gender diversity statute (SB 826) on employee reviews of female-friendly amenities. The outcome variables are the frequency counts of words related to eight female-friendly amenities in the review text in the pros section (columns (1) and (2)) or the cons section (columns (3) and (4)). Columns (5) and (6) use the difference in word counts between the pros and cons sections. Columns (7) and (8) take into account the length of the review text and use the weighted difference in word counts between the pros and cons sections. Columns (1), (3), (5), and (7) count anchor words while columns (2), (4), (6), and (8) count both anchor and top 20 words in the respective amenities group in Appendix Section I of Sockin (2022). Out of 50 amenities, we identify eight amenities that are important for female employees in the workplace (5-Paid time off, 10-Work-life balance, 11-Hours, 12-Work schedule, 16-Teleworking, 19-Respect/abuse, 21-Support, and 31-Diversity/inclusion). Table D5 in the Online Appendix provides the word list of the eight amenities. Robust standard errors clustered at the firm level are in brackets in Panel A. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets in Panel B. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 9: Effect of California Statute on Female Labor Market Participation

	(1)	(2)	(3)	(4)
Dep. Var.	Share of female employment out of total employment (%Fem Emp $\times$ 100)			
Flag_CA $\times$ Post_2018m9 $\times$ Private_For_Profit	-1.153*** [0.259]	-1.119*** [0.286]	-1.001*** [0.330]	-1.001*** [0.332]
Sample	All worker classes		Self-employed & private	
State FE	YES	-	YES	-
Year-Month FE	YES	-	YES	-
Worker class FE	YES	-	YES	-
State-Year-Month FE	NO	YES	NO	YES
State-Worker class FE	NO	YES	NO	YES
Worker class-Year-Month FE	NO	YES	NO	YES
Obs	29,096	29,096	15,708	15,708
Adj R2	0.753	0.812	0.801	0.832

This table shows the effect of the California board gender diversity statute (SB 826) on the share of female employment. The unit of observation is a state-worker class-year-month. Columns (1) and (2) include all worker classes in Current Population Survey (CPS): self-employed, not incorporated; self-employed, incorporated; private, for profit; private, nonprofit; federal/state/local government employee; unpaid family worker. Columns (3) and (4) exclude public sector worker classes and include only the following four worker classes: self-employed, not incorporated; self-employed, incorporated; private, for profit; and private, nonprofit. Flag_CA is a dummy that equals 1 for public firms headquartered in California, and 0 otherwise. Post_2018m9 is a dummy that equals 1 after September 30, 2018, and 0 otherwise. Private_For_Profit is a dummy that equals 1 for the worker class “private, for profit”, and 0 otherwise. The outcome variable across all columns is the share of female employment out of total employment, which is scaled by 100 for easy interpretation of the regression coefficients. All regressions are weighted by total employment in a state-worker class-year-month. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 10: External Validity – Effect of Board Gender Quotas in Europe on Female Labor Market Participation

Panel A: All Quotas

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Share of female employment out of total employment (%Fem Emp $\times$ 100)					
Quota	-0.289*** [0.105]	-0.252** [0.106]	-0.228** [0.092]	-0.358*** [0.130]	-0.294** [0.127]	-0.299*** [0.098]
%Fem population				0.268 [0.268]	0.276 [0.238]	0.156 [0.139]
Log (population)				7.311** [2.991]	6.052** [2.834]	4.244* [2.160]
Population growth				0.066 [0.050]	0.025 [0.081]	0.088* [0.044]
Log (GDP per capita)				0.959 [0.807]	1.543* [0.872]	0.991* [0.548]
GDP per capita growth				-0.002 [0.012]	0.004 [0.019]	-0.007 [0.011]
Industry FE	YES	-	-	YES	-	-
Country FE	YES	YES	-	YES	YES	-
Region $\times$ Year FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year FE	NO	YES	YES	NO	YES	YES
Country $\times$ Industry FE	NO	NO	YES	NO	NO	YES
Obs	29,734	29,734	29,688	29,734	29,734	29,688
Adj R2	0.918	0.917	0.991	0.918	0.917	0.991

Panel B: Hard vs. Soft Quotas

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Share of female employment out of total employment (%Fem Emp $\times$ 100)					
Hard quota	-0.311** [0.117]	-0.265** [0.110]	-0.238** [0.095]	-0.385*** [0.139]	-0.311** [0.130]	-0.313*** [0.101]
Soft quota	0.194 [0.119]	0.009 [0.277]	-0.013 [0.307]	0.093 [0.163]	-0.012 [0.286]	-0.06 [0.300]
%Fem population				0.22 [0.264]	0.247 [0.239]	0.132 [0.143]
Log (population)				7.074** [2.949]	5.929** [2.847]	4.132* [2.171]
Population growth				0.067 [0.048]	0.025 [0.081]	0.089* [0.044]
Log (GDP per capita)				1.046 [0.834]	1.605* [0.887]	1.038* [0.555]
GDP per capita growth				-0.003 [0.012]	0.004 [0.019]	-0.007 [0.011]
Industry FE	YES	-	-	YES	-	-
Country FE	YES	YES	-	YES	YES	-
Region $\times$ Year FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year FE	NO	YES	YES	NO	YES	YES
Country $\times$ Industry FE	NO	NO	YES	NO	NO	YES
Obs	29,734	29,734	29,688	29,734	29,734	29,688
Adj R2	0.918	0.917	0.991	0.918	0.917	0.991

This table shows the effect of board gender quotas on the share of female employment. The unit of observation is a country-industry-year. The sample includes 89 two-digit NACE industries in 32 European countries from 2008 to 2020. Panel A considers all board gender quotas adopted by the ten European countries. Panel B compares gender quotas with stronger enforcement and more precise policy wording (“hard” quotas) relative to ones with weaker enforcement and less precise policy wording (“soft” quotas). The outcome variable across all columns is the share of female employment, which is scaled by 100 for easy interpretation of the regression coefficients. All regressions are weighted by total employment in a country-industry-year. Robust standard errors clustered at the country level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

For Online Publication:

Internet Appendix to “Does Mandating Women on
Corporate Boards Backfire?”

This appendix has four sections. Section A contains additional information on variable definition and construction. Section B provides details of the LLM-based method and statistics. Section C discusses the details of the experimental design. Section D covers additional analysis, including tables and figures.

Table of Contents

Appendix A	Definition of Variables	A-4
Appendix B	LLM-based Gendered Wording Measures	A-5
B.1	Data collection and annotation	A-5
B.2	Data preprocessing and model selection	A-6
Appendix C	Experimental Design	A-10
C.1	Participants, manipulation, and survey questions	A-10
C.2	Validation of randomization and manipulation	A-11
Appendix D	Additional Analysis, Figures, and Tables	A-18
D.1	Coverage of LinkUp	A-18
D.2	Does “who drafts job ads” matter?	A-19
D.3	Gender-related labor violations	A-19
List of Appendix Figures		
C1	Landing Page of the Experiment	A-12
C2	Screenshots of the Experiment	A-13
D1	Female Board Representation from 2000 to 2020	A-21
D2	LinkUp Coverage from 2014 to 2020	A-22
D3	LinkUp vs. BoardEx/ExecuComp	A-23
D4	Sample Representativeness: LinkUp $\cap$ BoardEx/ExecuComp – Industry and Geography	A-24
D5	Gendered Language Usage in Different Occupations – High-Level Aggregation	A-25
D6	Gendered Language Usage in Different Occupations – Occupation-Level Char- acteristics	A-26
D7	Progress in Female Labor Market Participation Over Time	A-27
D8	Robustness - Excluding Top Feminine Words	A-28
D9	Robustness - Excluding Top Masculine Words	A-29
D10	Effect of Board Gender Quotas in Europe on Female Labor Market Partici- pation – Dynamics	A-30
D11	Robustness - Random Sampling	A-31
List of Appendix Tables		
A1	List of Variables, Definitions, and Data Sources - US Analysis	A-4
A2	List of Variables, Definitions, and Data Sources - European Analysis	A-5
B1	Gendered Language Usage in Job Ads: LLM-based Approach JobBERT	A-8
B2	Gendered Language Usage in Job Ads: LLM-based Approach DistilBERT	A-9
C1	Stimuli Used in the Experiment	A-14

C2	Measures Used in the Experiment	A-15
C3	Covariate Balance between Treatment and Control Subjects	A-16
C4	Experimental Evidence on Hiring Preference and Mechanisms – Restricting Subjects to Experienced Employees	A-16
C5	Decomposing Reactance	A-17
D1	Dictionary of Masculine and Feminine Words in Gaucher et al. (2011) . . .	A-32
D2	Female Directors and Gendered Language Usage in Job Ads – Heterogeneity: Decentralization in Drafting Job Ads	A-33
D3	Female Directors and Gendered Language Usage in Job Ads – Heterogeneity: Occupation Characteristics	A-34
D4	Effect of California Statute on Female Directors	A-35
D5	Descriptions of Female-Friendly Amenities	A-36
D6	Additional Summary Statistics	A-37
D7	Employee Ratings of Female-Friendly Benefits	A-38
D8	Female Directors and Gender-Related Labor Violations	A-39
D9	Board Gender Quotas in Europe	A-40
D10	Effect of Board Gender Quotas in Europe on Female Labor Market Partici- pation – Robustness to Alternative Specifications and Estimations	A-40

A Definition of Variables

Table A1: List of Variables, Definitions, and Data Sources - US Analysis

Variable	Definition	Source
%Fem words	Share of feminine words out of all gendered words in a job ad	LinkUp
#Fem words	Number of feminine words in a job ad	LinkUp
#Mas words	Number of masculine words in a job ad	LinkUp
#Words	Number of words in a job ad	LinkUp
Job zone	Job zones: 1 (less preparation required) to 5 (most preparation required)	ONET 26.2 Database
Mean wage	Average hourly wage of an occupation	BLS OEWS
Entry education	Typical entry-level education requirement: 0 (no formal educational credential) to 7 (Doctoral or professional degree)	BLS OEWS
STEM	Equals 1 if an occupation is in STEM fields, and 0 otherwise	BLS OEWS
#Job	Number of job ads in a firm-year	LinkUp
#Dir	Number of directors	BoardEx
%Fem dir	Share of female directors out of all directors	BoardEx
Female CEO	Equals 1 if a firm has a female CEO, and 0 otherwise	ExecuComp
Female CFO	Equals 1 if a firm has a female CFO, and 0 otherwise	ExecuComp
Female COO	Equals 1 if a firm has a female COO, and 0 otherwise	ExecuComp
Firm size	Size of a firm: log of total assets (ta)	Compustat
Cash holdings	Cash holdings: che/at	Compustat
Tangibility	Tangibility: ppent/at	Compustat
Tobin's Q	Tobin's Q: (at+prcc-f*csho-ceq-txditc)/at	Compustat
Leverage	Book leverage: (dltt+dlc)/at	Compustat
ROA	Profitability: oibdp/at	Compustat
Capex	Capital Investment: capx/at	Compustat
High Big3 ownership	Equals 1 if firms have top quartile Big3 (Vanguard, BlackRock, and State Street) ownership	13F
High concentration	Equals 1 for firms in industries with a high level of concentration (top quartile HHI defined at the two-digit SIC code level), and 0 otherwise	Compustat
>50% Republican contribution	Equals 1 if more than 50% of a firm's political contributions go to Republican candidates	Federal Election Commission
Female Top3 exec	Equals 1 if a firm has at least one female top executive (CEO, CFO, or COO), and 0 otherwise	ExecuComp
Std Dev (%Fem Words)	Standard deviation of gendered language usage (%Fem words) within each two-digit ONET-SOC occupation in a firm, then average across all occupations in a firm	Compustat & LinkUp
Std Dev (#Words)	Standard deviation of gendered language usage (#Words) within each two-digit ONET-SOC occupation in a firm, then average across all occupations in a firm	Compustat & LinkUp
Decentralized (%Fem Words)	Equals 1 if Std Dev (%Fem Words) is above median, and 0 otherwise	Compustat & LinkUp
Decentralized (#Words)	Equals 1 if Std Dev (#Words) is above median, and 0 otherwise	Compustat & LinkUp
Pros	Frequency counts of words related to eight female-friendly amenities in the review text in the pros section	Glassdoor
Cons	Frequency counts of words related to eight female-friendly amenities in the review text in the cons section	Glassdoor
Pros - Cons	Difference in frequency counts of words related to eight female-friendly amenities in the pros and cons section	Glassdoor
Pros - Cons (weighted)	Difference in frequency counts of words related to eight female-friendly amenities in the pros and cons section, weighted by length of the review	Glassdoor
Ratings: Overall	Overall employee ratings (1-5)	Glassdoor
Ratings: Compensation and benefits	Employee ratings for compensation and benefits (1-5)	Glassdoor
Ratings: Culture and values	Employee ratings for culture and values (1-5)	Glassdoor
Ratings: Work-life balance	Employee ratings for work-life balance (1-5)	Glassdoor
#Violation	Number of gender-related labor violations	Violation Tracker
log (1 + \$ Penalties)	Log of 1 plus dollar amount of penalty incurred by gender-related labor violations	Violation Tracker
%Fem emp	Share of female employment out of total employment at the state-worker class-year-month level	IPUMS CPS

The table lists variables used in the US job-, review-, and firm-level analysis.

Table A2: List of Variables, Definitions, and Data Sources - European Analysis

Variable	Definition	Source
%Fem emp	Share of female employment out of total employment at the country-industry (NACE)-year level	Eurostat LSF
Quota	Equals 1 if a country has implemented a board gender quota policy, and 0 otherwise	Legislative files
Hard quota	Equals 1 if a country has implemented a hard or medium-hard board gender quota policy, and 0 otherwise	Legislative files
Soft quota	Equals 1 if a country has implemented a soft board gender quota policy, and 0 otherwise	Legislative files
Total emp (1,000)	Total employment at the country-industry (NACE)-year level	Eurostat LSF
%Fem population	Share of female population	WorldBank
Log (population)	Log of total population	WorldBank
Population growth	Year-on-year population growth	WorldBank
Log (GDP per capita)	Log of GDP per capita	WorldBank
GDP per capita growth	Year-on-year GDP per capita	WorldBank

The table lists variables used in the European country-industry-level analysis.

B LLM-based Gendered Wording Measures

Constructing the gendered wording measure involves classifying a job ad into high versus low female-tilted categories, making it a supervised classification task. Given our focus on text classification, we leverage two pre-trained BERT models to classify job ads based on their gendered language characteristics. JobBERT (Zhang et al., 2022) is a BERT model pre-trained on 3.2 million sentences from job postings. DistilBERT (Sanh et al., 2019), is a distilled version of BERT, known for its balance of performance and computational efficiency, retaining 97% of BERT’s language understanding capabilities while being 60% faster and 40% smaller (Sanh et al., 2019). This appendix outlines our methodology for constructing and validating the gendered wording measure, ensuring transparency and replicability for future research.

B.1 Data collection and annotation

To fine-tune a pre-trained BERT model for a specific classification task, we need a labeled data set of job ads with ground truth labels. We start with a representative sample, comprising 4,500 job ads (extracted from our LinkUp sample) evenly distributed in male-dominated, female-dominated, and gender-neutral occupations based on the share of female employees in a given occupation. Each category contains 1,500 randomly sampled job ads to ensure diversity across industries and roles. Additionally, for each category, we require an equal

number of ads (500) to receive a high, medium, or low share of gendered words as identified by Gaucher et al. (2011).

Annotations are performed by eight trained research assistants, balanced by gender (four females and four males) and minority representation (four Whites and four minority individuals). A custom data labeling platform, hosted at <https://nlpuva.lighttag.io/>, is developed to manage the annotation workflow (This platform was discontinued in August 2024; all annotated data and metadata were preserved).

Job ads are evaluated using a single question adapted from Gaucher et al. (2011): “How many women work in the position being advertised?” Responses are recorded on a Likert scale ranging from 0 (0% women) to 20 (100% women) in 5% increments. Each job ad is labeled by three annotators, and disagreements are resolved by majority vote, followed by a final review conducted by a fourth annotator. Weekly meetings are held to monitor annotators’ productivity, ensure annotation quality, address issues, and refine guidelines as needed.

To simplify the process, the continuous Likert scale is converted into binary labels. Job ads with scores above 50% are categorized as female-tilted (1), while those with scores of 50% or below are categorized as male-titled (0). After removing incomplete entries, the final data set comprises 4,436 labeled job ads, with 2,984 categorized as more feminine-worded and 1,452 as more masculine-worded.

B.2 Data preprocessing and model selection

Before fine-tuning a BERT model, job ads are preprocessed for consistency. Tokenization and cleaning are performed using the Hugging Face tokenizer, ensuring compatibility with downstream tasks. The annotated data set is split into training (60%), validation (20%), and testing (20%) subsets.

The training process involves fine-tuning JobBERT and DistilBERT to adapt them to the job ad context. Key hyperparameters include a batch size of 32 and 8 epochs. Optimization is performed using the AdamW optimizer with a linear learning rate schedule and 10% of

training steps as warm-up steps, with an initial learning rate of $5e-5$ and a weight decay of 0.01. Model performance is evaluated using standard classification metrics, including accuracy and AUC.

JobBERT achieves an accuracy of 88.84% and an AUC score of 94.00% while DistilBERT achieves an accuracy of 95.60% and an AUC score of 97.71%.

The fine-tuned JobBERT and DistilBERT models are then applied to a separate data set to classify job ads. Each model outputs a probability between 0 and 1, indicating the likelihood of a job ad being female-tilted. A probability closer to 1 suggests a more feminine ad, while a probability closer to 0 indicates a more masculine ad.

**Table B1: Gendered Language Usage in Job Ads: LLM-based Approach
JobBERT**

Panel A: Correlation with Female Directors

	(1)	(2)	(3)	(4)	(5)	(6)
	BoW	JobBERT	BoW	JobBERT	BoW	JobBERT
Dep. Var.	Continuous		4 Group		10 Group	
%Fem dir	-0.103** [0.040]	-0.112*** [0.030]	-0.337** [0.146]	-0.729*** [0.147]	-0.847** [0.360]	-1.656*** [0.365]
Controls	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Occupation (ONET 6-digit) FE	YES	YES	YES	YES	YES	YES
Obs	1,642,316	1,573,705	1,642,316	1,573,705	1,642,316	1,573,705
Adj R2	0.434	0.57	0.409	0.594	0.416	0.638

Panel B: Effect of California Statute

	(1)	(2)	(3)	(4)	(5)	(6)
	BoW	JobBERT	BoW	JobBERT	BoW	JobBERT
Dep. Var.	Continuous		4 Group		10 Group	
Flag_CA $\times$ Post_2018m9	-0.042*** [0.010]	-0.018*** [0.007]	-0.174*** [0.043]	-0.095*** [0.035]	-0.463*** [0.101]	-0.309*** [0.090]
Controls	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Occupation (ONET 6-digit) FE	YES	YES	YES	YES	YES	YES
Obs	1,642,550	1,573,942	1,642,550	1,573,942	1,642,550	1,573,942
Adj R2	0.434	0.57	0.409	0.594	0.416	0.638

This table uses alternative measures based on LLM, and more specifically, a fine-tuned JobBERT method. Columns (1), (3), and (5) use the bag-of-word approach used in the main specification while Columns (2), (4), and (6) use the JobBERT method. The outcome variable is a continuous measure between 0 and 1 in columns (1) and (2). Columns (3) and (4) ((5) and (6)) categorize the continuous measure into 4 (10) groups. Panel A shows the relationship between the share of female directors and gendered language usage in job ads. Panel B shows the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. Flag_CA is a dummy that equals 1 for public firms headquartered in California, and 0 otherwise. Post_2018m9 is a dummy that equals 1 after September 30, 2018, and 0 otherwise. Robust standard errors clustered at the firm level are in brackets in Panel A. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets in Panel B. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

**Table B2: Gendered Language Usage in Job Ads: LLM-based Approach
DistilBERT**

Panel A: Correlation with Female Directors

	(1)	(2)	(3)	(4)	(5)	(6)
	BoW	BERT	BoW	BERT	BoW	BERT
Dep. Var.	Continuous		4 Group		10 Group	
%Fem dir	-0.103** [0.040]	-0.049 [0.056]	-0.337** [0.146]	-0.415*** [0.148]	-0.847** [0.360]	-1.168*** [0.390]
Controls	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Occupation (ONET 6-digit) FE	YES	YES	YES	YES	YES	YES
Obs	1,642,316	1,606,553	1,642,316	1,606,553	1,642,316	1,606,553
Adj R2	0.434	0.441	0.409	0.448	0.416	0.489

Panel B: Effect of California Statute

	(1)	(2)	(3)	(4)	(5)	(6)
	BoW	BERT	BoW	BERT	BoW	BERT
Dep. Var.	Continuous		4 Group		10 Group	
Flag_CA $\times$ Post_2018m9	-0.042*** [0.010]	-0.067*** [0.008]	-0.174*** [0.043]	-0.148*** [0.029]	-0.463*** [0.101]	-0.422*** [0.081]
Controls	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES	YES	YES
Occupation (ONET 6-digit) FE	YES	YES	YES	YES	YES	YES
Obs	1,642,550	1,606,790	1,642,550	1,606,790	1,642,550	1,606,790
Adj R2	0.434	0.482	0.409	0.448	0.416	0.489

This table uses alternative measures based on LLM, and more specifically, a fine-tuned DistilBERT method. Columns (1), (3), and (5) use the bag-of-word approach used in the main specification while Columns (2), (4), and (6) use the DistilBERT method. The outcome variable is a continuous measure between 0 and 1 in columns (1) and (2). Columns (3) and (4) ((5) and (6)) categorize the continuous measure into 4 (10) groups. Panel A shows the relationship between the share of female directors and gendered language usage in job ads. Panel B shows the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. Flag_CA is a dummy that equals 1 for public firms headquartered in California, and 0 otherwise. Post_2018m9 is a dummy that equals 1 after September 30, 2018, and 0 otherwise. Robust standard errors clustered at the firm level are in brackets in Panel A. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets in Panel B. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

C Experimental Design

C.1 Participants, manipulation, and survey questions

Our experiment employs a between-subjects design with two conditions: board gender quota policy (treatment) and no board gender quota policy (control). US-based participants are recruited via CloudResearch’s Connect platform. We require subjects to be between 18 and 65 years old and have at least one year of work experience. Our subject pool comprises 200 participants including 96 females, 101 males, and 3 non-binary or prefer not to disclose. The median age is 36 ($SD = 9.99$), and age ranges from 19 to 65 years old.

Subjects participate the experiment via an online survey. The experiment starts with a landing page shown in Figure C1.³⁶ Subjects are asked to imagine themselves as employees of a fictitious company. Each subject is then randomly assigned to one of the two policy conditions (see Table C1 for details). One condition (treatment) mentions and explains the board gender quota policy while the other (control) does not. To ensure adequate attention paid to each condition, a delay is introduced so that subjects spend at least five seconds reading the policy. To check subjects’ understanding of the policy, we include a manipulation check (using a three-item, seven-point Likert scale, see Table C2 and Figure C2a).

After completing the manipulation check, subjects are asked a set of questions related to their hiring preferences (see Table C2 and Figure C2b). We then ask questions related to the different economic mechanisms, listed in Table C2, to shed light on the rationale underlying subjects’ hiring preferences. We first explore whether the board gender-quota policy, which may be perceived as a threat to or loss of freedom, triggers psychological reactance among subjects. Reactance, in this context, refers to a defensive response when individuals feel their freedom is threatened or eliminated. The second mechanism is norm violation, i.e., subjects view the quota policy (treatment) as a violation of gender norms in their workplace.

In addition to reactance and norm violation, we consider two other mechanisms that could shape subjects’ hiring preferences. Under moral licensing, compliance with the quota

³⁶The full survey can be found at https://commercevirginia.col.qualtrics.com/jfe/form/SV_6DstycBnLo3hygC.

policy in one context might lead subjects to feel justified in acting contrary to gender equality goals in other contexts. We employ an adapted version of the scale in Merritt et al. (2010) to assess this mechanism (Figure C2c). Under job insecurity, the quota policy causes subjects to worry about their own jobs or career prospects, particularly if they view the policy as favoring women at the expense of merit-based hiring or reducing opportunities for men. These concerns of job insecurity may lead subjects to go against hiring female candidates in non-leadership roles as a way to safeguard their own positions or organizational stability. To evaluate this mechanism, we adopt the scale from Shoss (2017).

To eliminate order effects, we randomize the mechanisms in their presentation. We include a check requiring subjects to select “Strongly Disagree” to ensure their attention (Figure C2d). We end the experiment by asking subjects demographic questions, including age, gender, ethnicity, language, employment status, years of work experience, and political view (from liberal to conservative).

C.2 Validation of randomization and manipulation

Table C3 confirms our randomization design among subjects. Participants in the treatment group are comparable to those in the control group across key characteristics, with no significant differences in age, race, work experience, or political view. The result in column (1) of Table 6 Panel A reveals a significant manipulation effect. Subjects under the treatment condition view the policy as significantly more binding and impactful on board gender composition and diversity compared to those under the control condition.

Figure C1: Landing Page of the Experiment

Welcome to the Policy Study!

In this survey, you will review a recruitment policy from your company and assist in the recruitment process for a specific position. You will also be asked to answer some questions for classification purposes.

Before beginning this study, please read the following general instructions:

- Complete this study in a quiet space continuously and in one sitting.
- You must NOT be working on any other tasks or have another browser open.
- Turn your cellphone off and refrain from having conversations with others.
- Read all questions carefully and answer honestly and to the best of your ability.
- Keep the browser window maximized.
- Do NOT press the back or refresh buttons on the browser.
- Please go through all pages.

NOTE: There are attention check questions in this study. These are included to ensure that you are devoting your full attention to the study.

In the question below, please indicate that you are prepared to meet all study requirements. If you are not, please close your browser and complete the study another time.

Agreement:

I acknowledge that I have read all of the requirements and agree to participate in this research study.

(If you do not agree, please close the browser and exit the study)

☐ Agree

This figure presents the landing page of the experiment when a subject starts the experiment.

Figure C2: Screenshots of the Experiment

Thinking about your company's current policy for recruitment, to what extent do you agree with the following statements:

	Strongly disagree 1	2	3	4	5	6	Strongly agree 7
My company is required by law to follow this policy.	☐	☐	☐	☐	☐	☐	☐
My company's policy will impact the composition of its top leadership team in the near future.	☐	☐	☐	☐	☐	☐	☐
My company's policy will lead to a shift in gender diversity in the company.	☐	☐	☐	☐	☐	☐	☐



(a) Manipulation Check

Imagine you are in charge of hiring for a **non-leadership position (this position is not affected by the gender quota policy)**, and your input will be critical in the decision-making process. Keeping the policy you just read in mind, to what extent do you agree with the following statements?

	Strongly disagree 1	2	3	4	5	6	Strongly agree 7
The recruitment policy shifts my attention toward male candidates for non-leadership roles	☐	☐	☐	☐	☐	☐	☐
I find myself considering male candidates more often for non-leadership positions in this context.	☐	☐	☐	☐	☐	☐	☐
The recruitment policy makes me more inclined to favor male candidates for non-leadership positions.	☐	☐	☐	☐	☐	☐	☐



(b) Dependent Variable: Hiring Decisions

Thinking about your company's current policy for recruitment, to what extent do you agree with the following statements:

	Strongly disagree 1	2	3	4	5	6	Strongly agree 7
My company has earned enough moral credit through its actions for women.	☐	☐	☐	☐	☐	☐	☐
My company has done enough to support women.	☐	☐	☐	☐	☐	☐	☐
My company is morally justified in its actions regarding women's representation.	☐	☐	☐	☐	☐	☐	☐



(c) Mechanism - Moral Licensing

Please select 'strongly disagree' to show you are paying attention to this question.

Strongly disagree ☐	Disagree ☐	Slightly disagree ☐	Neutral ☐	Slightly agree ☐	Agree ☐	Strongly agree ☐
--	-----------------------------------	--	----------------------------------	---	--------------------------------	---



(d) Attention Check

This figure presents screenshots from various sections of the experiment. Figure C2a lists the manipulation check questions. Figure C2b lists the questions related to subjects' hiring preferences. Figure C2c lists the questions related to one of the mechanisms - moral licensing. Figure C2d lists our attention check question.

Table C1: Stimuli Used in the Experiment

Instruction
On the next page, you will read the recruitment policy from the company that you are working for. Then, you will be asked to share your opinions on it. While you are reading, please keep imagining yourself actually working for that company. Please note that there will be a short delay before the forward “»” button appears. This is so you have sufficient time to read the scenario.
Treatment (with Board Gender Quota Policy)
The company you work for follows a government-enforced gender quota policy that mandates a specific percentage of women must be included among its board members. This policy applies strictly to leadership-level hiring decisions and limits the company’s ability to select candidates solely based on qualifications or experience for these roles. The policy is non-negotiable , and compliance is required to avoid penalties . However, this policy does not apply to non-leadership positions , where hiring decisions are made independently of the quota. The goal of the policy is to enforce gender diversity and ensure balanced representation in decision-making processes.
Control (without Board Gender Quota Policy)
The company you work for follows a recruitment policy that prioritizes selecting candidates based on their qualifications, professional experience, and how well they align with the company’s goals . This policy applies equally to all positions, including both leadership and non-leadership positions , and emphasizes the importance of individual merit in hiring decisions. The policy allows full flexibility for the company to select the best candidates for each role. The goal of the policy is to promote fairness by ensuring that hiring decisions are based on objective criteria and organizational needs .

This table presents the instruction for participating in our experiment, followed by the two policy conditions (treatment vs. control) that are randomly assigned to subjects.

Table C2: Measures Used in the Experiment

Measures	Questions and Scales
Manipulation Check	Thinking about your company's current policy for recruitment, to what extent do you agree with the following statements: My company is required by law to follow this policy. My company's policy will impact the composition of its top leadership team in the near future. My company's policy will lead to a shift in gender diversity in the company. 1 = Strongly Disagree, 7 = Strongly Agree; Cronbach's $\alpha = 0.71$
Dependent Variable: Hiring Preference	Imagine you are in charge of hiring for a non-leadership position, and your input will be critical in the decision-making process. Keeping the policy you just read in mind, to what extent do you agree with the following statements: The recruitment policy shifts my attention toward male candidates for non-leadership roles. I find myself considering male candidates more often for non-leadership positions in this context. The recruitment policy makes me more inclined to favor male candidates for non-leadership positions. 1 = Strongly Disagree, 7 = Strongly Agree; Cronbach's $\alpha = 0.95$
Mechanism: Reactance	
- Experience of Reactance	To what extent do you perceive the recruitment policy as a restriction of your freedom to make hiring decisions: Are you frustrated about the recruitment policy? How much does the recruitment policy annoy you? To what extent are you offended/disturbed by the recruitment policy?
- Negative Attitudes	Thinking about your company's current policy for recruitment, answer the following questions: Do you think that this recruitment policy could also have prejudices against men? Do you think the recruitment policy reflects discriminatory practices in hiring? How likely do you think the recruitment policy unfairly advantages women over others?
- Aggressive Behavioral Intentions	Thinking about your company's current policy for recruitment, answer the following questions: Would you consider publicly criticizing the recruitment policy? How strong is your wish to file a formal complaint about the recruitment policy? How much would you advise colleagues or peers push back the recruitment policy? 1 = Not at all, 7 = Very much; Adapted from Sittenthaler et al. (2015); Cronbach's $\alpha = 0.97$
Mechanism: Norm Violation	Thinking about your company's current policy for recruitment, to what extent do you agree with the following statements: This policy goes against my beliefs about the roles men and women should have in the workplace. This policy challenges my understanding of how gender representation in the workplace has traditionally developed over time. This policy conflicts with my personal assumptions about how workplace roles should be distributed between men and women. 1 = Strongly Disagree, 7 = Strongly Agree; Cronbach's $\alpha = 0.93$
Mechanism: Moral Licensing	Thinking about your company's current policy for recruitment, to what extent do you agree with the following statements: My company has earned enough moral credit through its actions for women. My company has done enough to support women. My company is morally justified in its actions regarding women's representation. 1 = Strongly Disagree, 7 = Strongly Agree; Adapted from: Merritt et al. (2010); Cronbach's $\alpha = 0.85$
Mechanism: Job Insecurity	Thinking about your company's current policy for recruitment, to what extent do you agree with the following statements: Chances are, I will soon lose my job. I feel insecure about my future job. I think I might lose my job in the near future. 1 = Strongly disagree, 7 = Strongly agree; Adapted From: Shoss (2017); Cronbach's $\alpha = 0.95$

This table lists the multiple-item, seven-point Likert scale used in the manipulation check, hiring preference (dependent variable), and mechanism tests. The mechanisms are randomized in their presentation to eliminate order effects.

Table C3: Covariate Balance between Treatment and Control Subjects

	Control		Treatment		Diff	T-stat
	Mean	SD	Mean	SD		
Age	36.394	10.296	37.581	9.767	-1.187	-0.832
Male (0/1)	0.511	0.503	0.505	0.502	0.006	0.082
Ethnicity: White (0/1)	0.734	0.444	0.714	0.454	0.020	0.310
Fulltime (0/1)	0.691	0.464	0.686	0.466	0.006	0.087
Work Experience (>5 years)	0.787	0.411	0.819	0.387	-0.032	-0.560
Political View (1-7)	3.277	1.822	3.619	2.007	-0.342	-1.262

This table shows the balance of covariates between subjects in the treatment group and those in the control group.

**Table C4: Experimental Evidence on Hiring Preference and Mechanisms
– Restricting Subjects to Experienced Employees**

	(1)	(2)	(3)	(4)	(5)	(6)
			Mechanisms			
Dep. Var.	Manipulation	Hiring Decisions	Norm Violation	Reactance	Moral Licensing	Insecurity
Treatment	1.857*** [0.204]	0.638* [0.331]	1.468*** [0.249]	1.173*** [0.314]	-0.002 [0.268]	0.338 [0.274]
Controls	YES	YES	YES	YES	YES	YES
Dep. Var. Mean	4.99	2.982	2.827	2.638	4.177	1.984
Obs	128	128	128	128	128	128
Adj R2	0.373	0.008	0.327	0.167	0.056	0.081

This table reports the treatment effect estimated from the experiment by removing subjects with fewer than five years of work experience or self-employed. The outcome variable is seven-point Likert scale for the manipulation questions in column (1), for questions on hiring preference (favoring men) in column (2), and for questions related to different mechanisms in columns (3) to (6). The following control variables are included: age, gender, ethnicity, employment status, years of work experience, political view, type of device used for the survey, and the hour of the survey. The unit of observation is a subject. The sample includes 128 subjects who passed the attention test. Treatment is a dummy that equals 1 for the treatment group and 0 for the control group. Robust standard errors adjusted for heteroskedasticity are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table C5: Decomposing Reactance

Panel A: Comparing Treated and Control Subjects

	(1)	(2)	(3)
	Reactance		
Dep. Var.	Experience of Reactance	Negative Attitudes	Aggressive Behavior Intentions
Treatment	1.735*** [0.251]	1.712*** [0.245]	1.056*** [0.243]
Controls	YES	YES	YES
Dep. Var. Mean	3.077	3.33	2.523
Obs	199	199	199
Adj R2	0.291	0.312	0.305

Panel B: Heterogeneity - Political View and Gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Reactance					
Dep. Var.	Experience of Reactance		Negative Attitudes		Aggressive Behavior Intentions	
Treatment × Political View	0.300** [0.139]		0.133 [0.131]		0.320** [0.136]	
Treatment × Male		1.429*** [0.510]		0.752 [0.503]		0.789 [0.482]
Treatment	0.686 [0.529]	1.027*** [0.358]	1.247** [0.505]	1.339*** [0.315]	-0.064 [0.477]	0.665** [0.320]
Controls	YES	YES	YES	YES	YES	YES
Dep. Var. Mean	3.077	3.077	3.33	3.33	2.523	2.523
Obs	199	199	199	199	199	199
Adj R2	0.308	0.322	0.311	0.317	0.327	0.312

This table examines the three subscales of reactance - experience of reactance, negative attitudes, and aggressive behavior intentions. Panel A reports the treatment effect estimated from the experiment. Panel B reports heterogeneous treatment effects based on political views (1-7 from liberal to conservative) and gender. The following control variables are included: age, gender, ethnicity, employment status, years of work experience, political view, type of device used for the survey, and the hour of the survey. The unit of observation is a subject. The sample includes 199 subjects who passed the attention check. Treatment is a dummy that equals 1 for the treatment group and 0 for the control group. Robust standard errors adjusted for heteroskedasticity are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

D Additional Analysis, Figures, and Tables

D.1 Coverage of LinkUp

Figure D2a shows that LinkUp’s coverage of job ads grows over time. The number of job ads increases from fewer than 0.2 million per month in 2014 to around 0.6 million per month toward the later half of our sample period. Figure D2b shows LinkUp’s monthly coverage of public firms. Their coverage grows from fewer than 1,000 firms at the beginning of 2014 to close to 2,500 firms at the end of 2020.

Figure D3a shows that the merged sample of LinkUp and BoardEx firms (blue bar) spans between 1,647 to 2,447 Compustat firms over the period 2014-2020, capturing between 37% to 64% of BoardEx firms. While this coverage ratio does not seem particularly high, LinkUp tends to cover large companies that are also big employers. The blue bar in Figure D3b shows the assets-weighted share of public firms in the merged sample of LinkUp and BoardEx. This ratio varies between 71% to 85%.

Next, when we focus on large companies by further merging the data with ExecuComp (ExecuComp covers public firms that are the constituents of S&P 1500), the resulting sample has between 1,180 to 1,453 S&P 1500 firms over the period 2014–2020, as shown by the green bar in Figure D3a. This represents a majority of S&P 1500 firms, both in terms of the number of firms and the assets-weighted share. The green bar in Figure D3b shows that between 74% to 88% of the aggregate assets value of S&P 1500 firms in ExecuComp are covered by LinkUp.

Our sample using LinkUp data captures the most significant firms in the economy. We next examine if there is any industry or geographic tilt. We compare the distribution of firms in our merged sample across industries and states with the distribution of firms in BoardEx or ExecuComp. Figure D4 suggests that the sample merged with LinkUp has very similar industry (Figures D4a and D4b) or geographic (Figures D4c and D4d) coverage to that of BoardEx and ExecuComp.

D.2 Does “who drafts job ads” matter?

While it is difficult to pinpoint the individuals responsible for drafting each job ad within an organization, creating measures to capture how decentralized this process is and examining whether these measures change the negative association are informative. Specifically, for a given occupation, we can assess the degree of similarity across job ads posted by a firm. If those job ads are very similar (different), likely the firm has a centralized (decentralized) process of drafting job ads.

We create four measures at the firm level to capture the degree of decentralization in a firm’s drafting of job ads. We start by computing the standard deviation of the share of feminine words and the length of a job ad at the occupation level. We then average across different occupations to generate two firm-level continuous measures. A larger value indicates more dispersion, and hence a higher degree of decentralization, in a firm’s writing of job ads. For each continuous measure, we further create an indicator variable for an above-median level of decentralization.

Table D2 presents the results using the above measures. We show that the pushback effect is invariant to firms’ different degrees of decentralization in drafting job ads, suggesting that pushback occurs in a wide range of firms that differ in their hiring practices. In decentralized firms, a large number of employees who provide inputs for job ads can all push back as gender stereotypes or ideological preferences prevail within the organization. In centralized firms, a few conservative decision-makers may have a dominant voice and drive the pushback. This set of results provides suggestive evidence that the pushback against empowering women is prevalent across different levels of a corporate hierarchy.

D.3 Gender-related labor violations

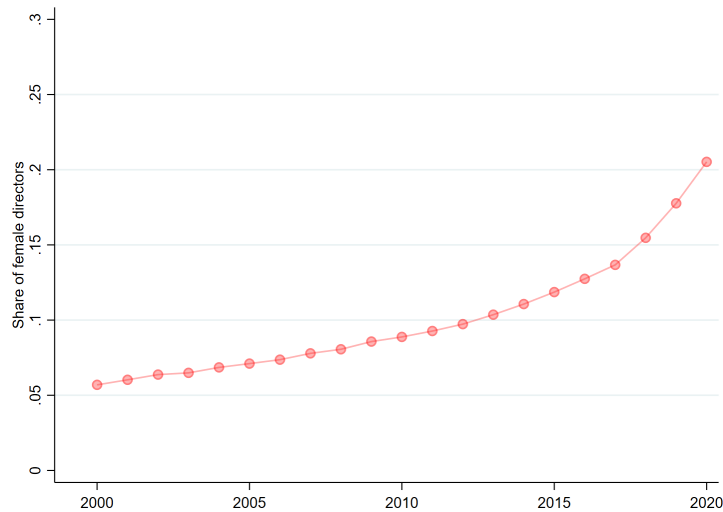
We explore whether backlash also manifests in the unfair treatment of female employees in the workplace. We start with all employment-related violation cases from Violation Tracker and employ two approaches to identifying gender-related labor violations. First, a violation is considered gender-related if it mentions any of the following seed words in the “description”

or “secondary offense” fields: female, gender, hostile, hours, leave, sex, sexual, support, and vacation. Second, we expand this initial word list by adding words from the above two fields with high associations in terms of pmi-freq scores (Jin et al., 2021) with the seed words: applicants, women, job, positions, discrimination, environment, medical, sick, and harassment. We then aggregate case-level information to firm-year level and calculate the number of gender-related violations and the dollar amount of penalties paid for those violations.³⁷ Table D6 Panel C in the Online Appendix provides the summary statistics. An average firm has 0.011 gender-related violations in a year based on the expanded word list.

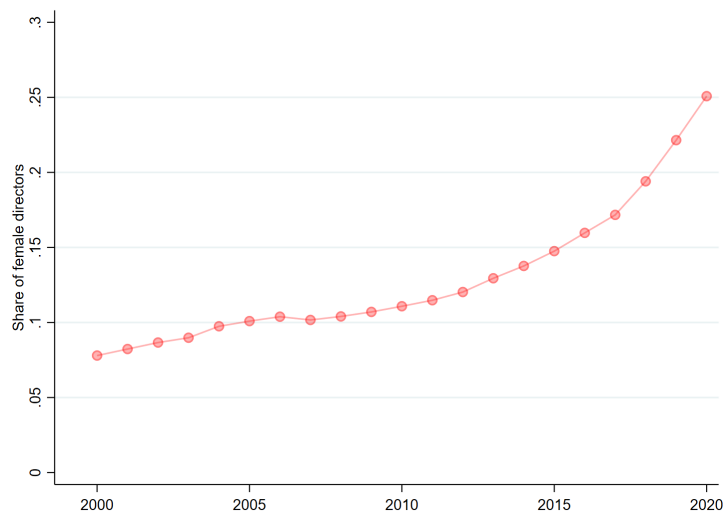
Table D8 presents the results. Panel A documents a positive association between a firm’s share of female directors and its number of (the amount of penalties paid for) gender-related labor violations. While we lack statistical power in columns (1) and (2), the results become significant at the 5% level when the expanded word list helps identify a larger set of relevant violations. Table D8 Panel B further exploits SB 826 for identification. We find consistent evidence that mandating women on corporate boards led to a significant increase in both the number of gender-related violations and the amount of penalties paid for those violations. The evidence suggests that SB 826-triggered drop in firms’ demand for female labor was accompanied by more female employees experiencing unjust workplace treatments, such as gender discrimination and sexual harassment.

³⁷These cases are mostly brought by the Equal Employment Opportunity Commission (EEOC) or the Office of Federal Contract Compliance Programs (OFCCP).

Figure D1: Female Board Representation from 2000 to 2020



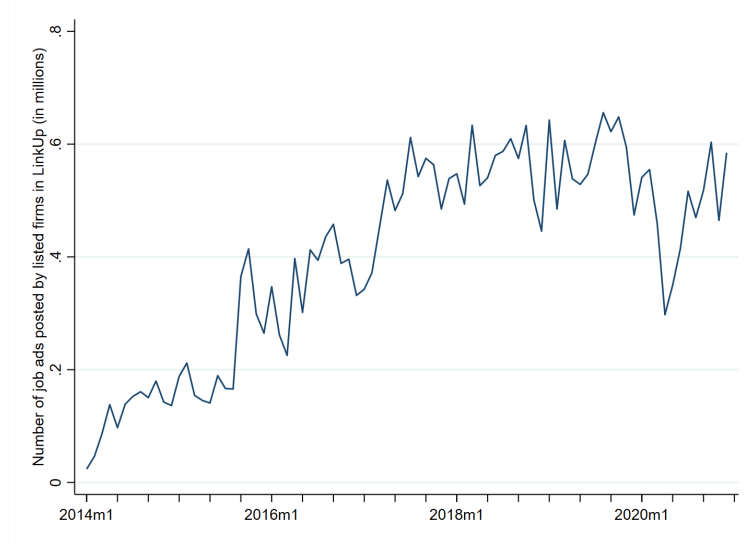
(a) BoardEx Firms



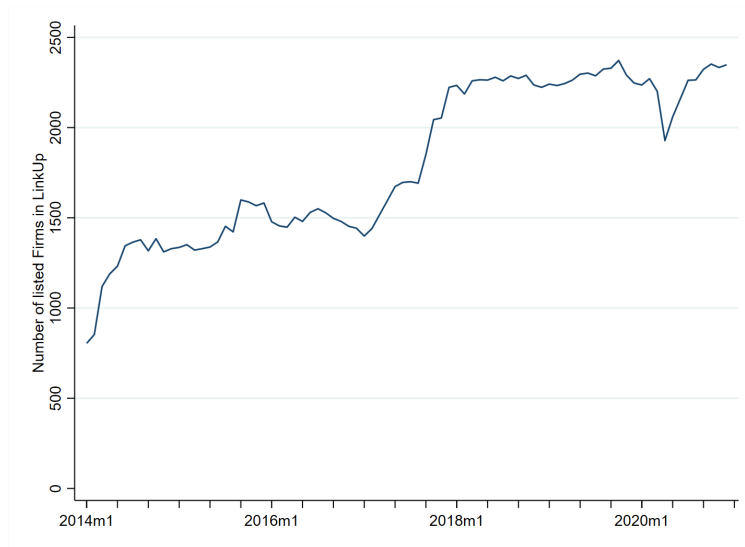
(b) S&P 1500 Firms

This figure shows the time trend in female board representation from 2000 to 2020. Figure D1a plots the average share of female directors in public firms with available data in Compustat and BoardEx. Figure D1b plots the average share of female directors in S&P 1500 firms.

Figure D2: LinkUp Coverage from 2014 to 2020



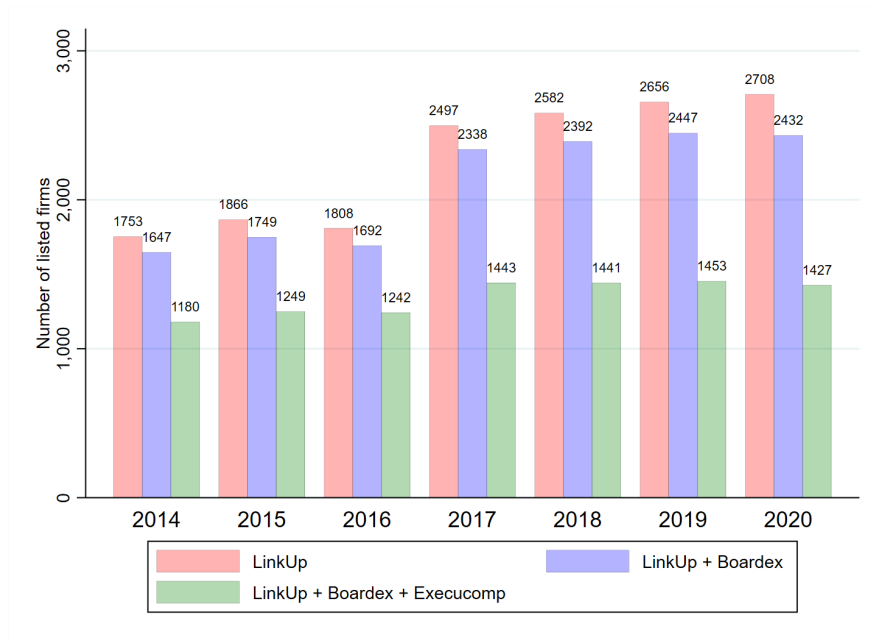
(a) Number of Job Ads in LinkUp



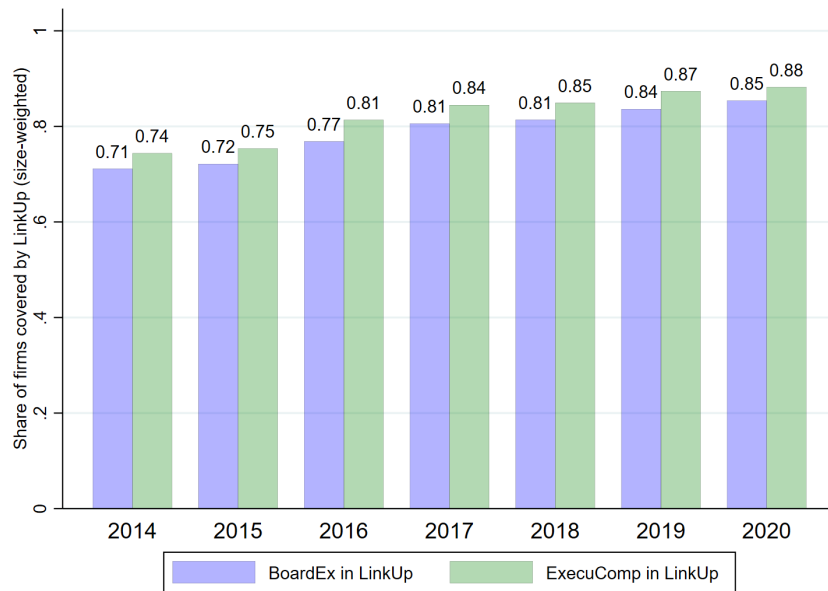
(b) Number of Listed Firms in LinkUp

This figure shows the coverage of LinkUp from 2014 to 2020. Figure D2a plots the number of job ads covered by LinkUp. Figure D2b plots the number of public firms covered by LinkUp.

Figure D3: LinkUp vs. BoardEx/ExecuComp



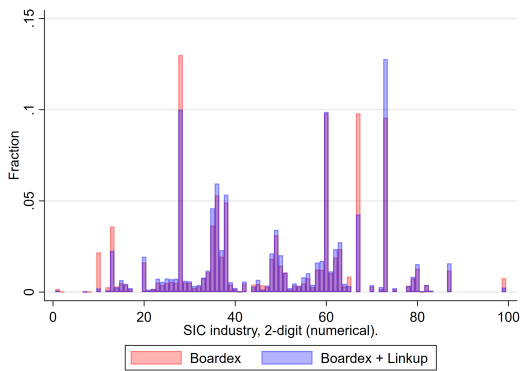
(a) Number of Public Firms in LinkUp and $\text{LinkUp} \cap \text{BoardEx/ExecuComp}$



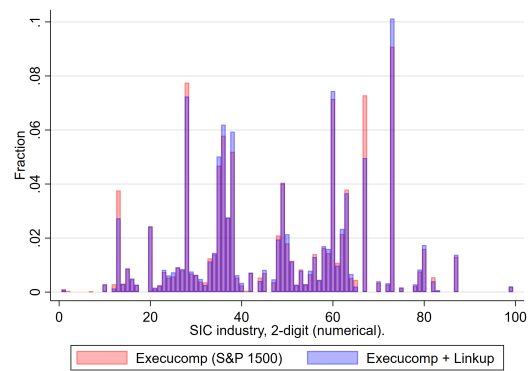
(b) Share of BoardEx/ExecuComp Firms Covered by LinkUp

This figure shows the coverage of LinkUp relative to BoardEx and ExecuComp. Figure D3a shows the number of public firms in LinkUp relative to the number of firms covered by both LinkUp and BoardEx, or the number of firms covered by LinkUp, BoardEx, and ExecuComp. Figure D3b shows the assets-weighted share of BoardEx (ExecuComp) firms covered by LinkUp.

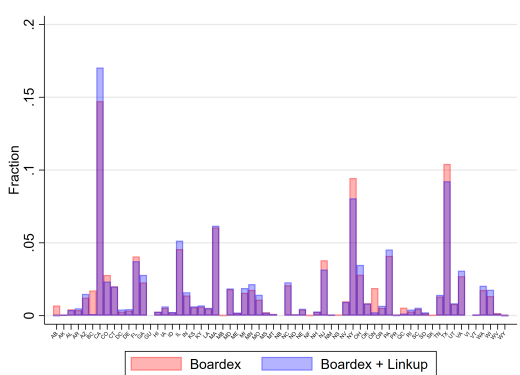
Figure D4: Sample Representativeness: $\text{LinkUp} \cap \text{BoardEx}/\text{ExecuComp}$ – Industry and Geography



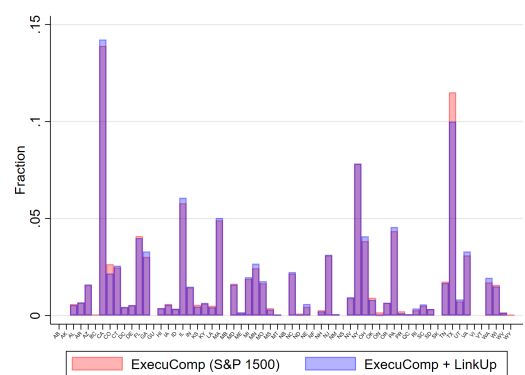
(a) Industry, BoardEx vs. BoardEx + LinkUp



(b) Industry, ExecuComp vs. ExecuComp + BoardEx + LinkUp



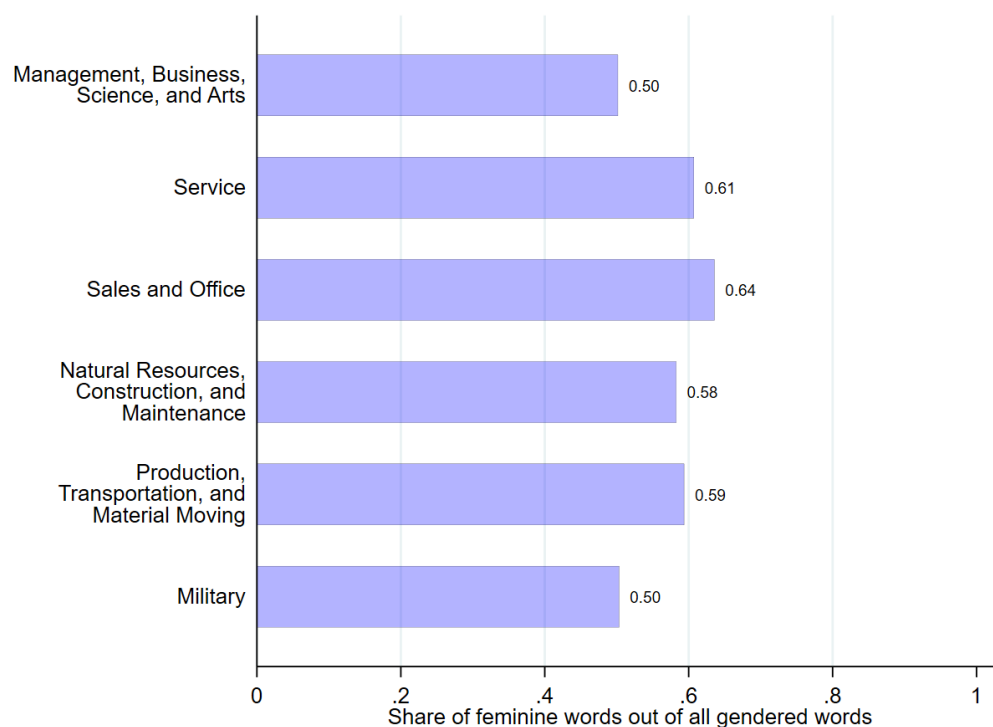
(c) HQ State, BoardEx vs. BoardEx + LinkUp



(d) HQ State, ExecuComp vs. ExecuComp + BoardEx + LinkUp

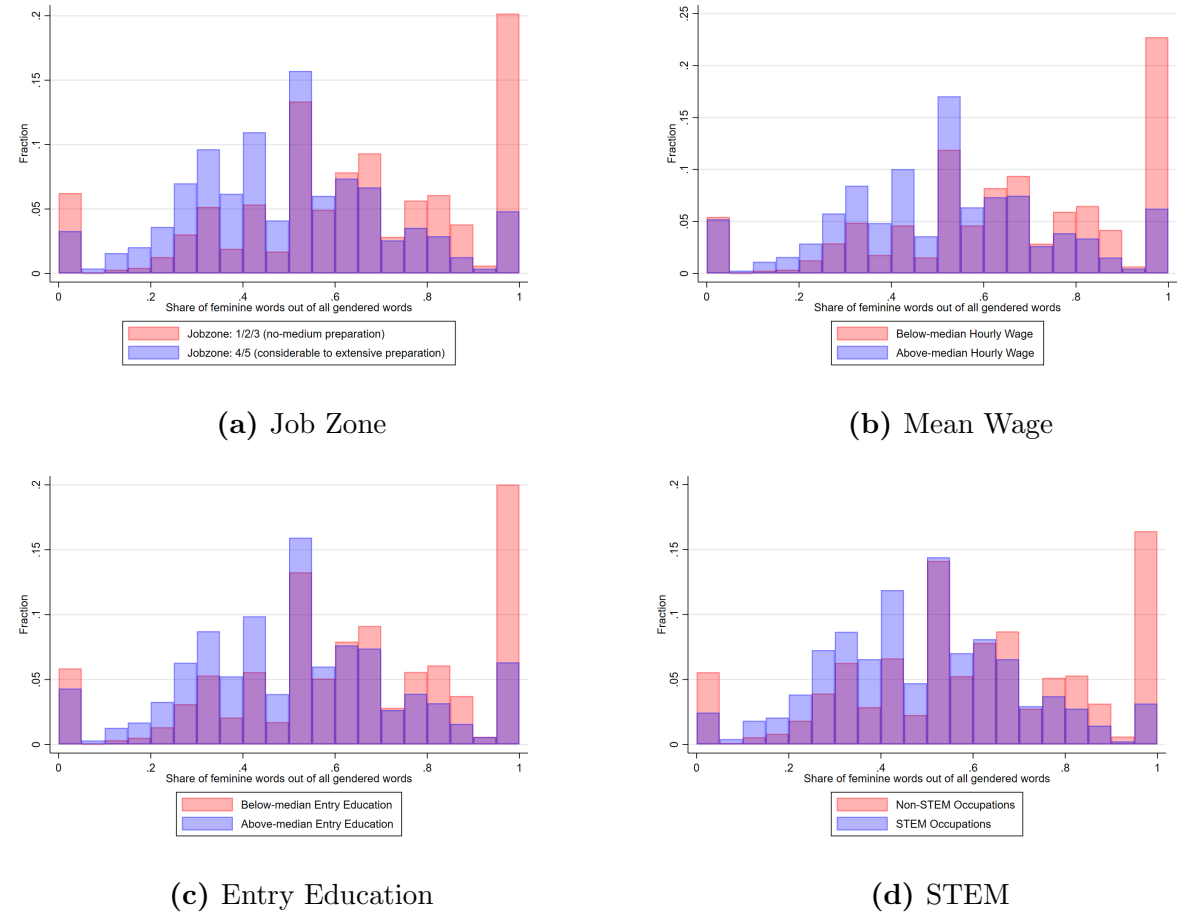
This figure shows the industry (two-digit SIC) and geographical coverage of the merged data set of LinkUp and BoardEx (and ExecuComp). Figure D4a (Figure D4c) compares the fraction of public firms in each two-digit SIC industry (in each headquarters state) in the BoardEx data set with the corresponding fraction in the merged data set of LinkUp and BoardEx. Figure D4b (Figure D4d) compares the fraction of public firms in each two-digit SIC industry (in each headquarters state) in the ExecuComp data set with the corresponding fraction in the merged data set of LinkUp, BoardEx, and ExecuComp.

**Figure D5: Gendered Language Usage in Different Occupations
– High-Level Aggregation**



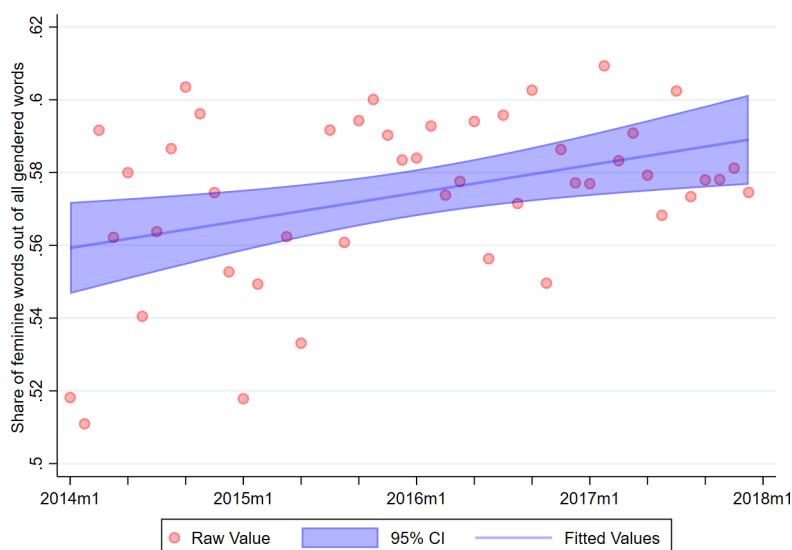
This figure shows gendered language usage in six occupations, based on the two-digit ONET-SOC code (11-29, 31-39, 41-43, 45-49, 51-53, 55). The definitions are from Table 6 of the Standard Occupational Classification (SOC) User Guide. See <https://www.aila.org/File/Related/17122005a.pdf>

Figure D6: Gendered Language Usage in Different Occupations
– Occupation-Level Characteristics

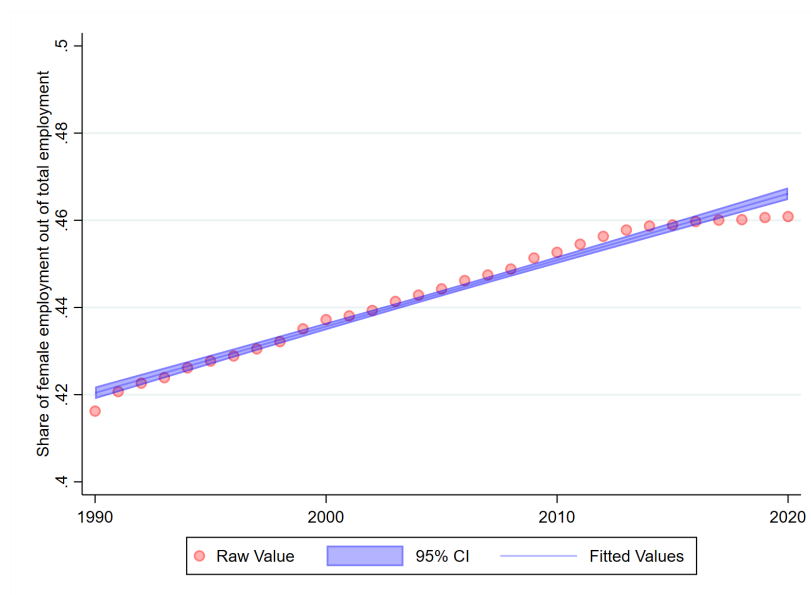


This figure shows how gendered language usage in job ads varies with occupation-level characteristics. Figure D6a splits the sample by the value of job zone, which is a group indicator from 1 to 5 for occupations that are similar in their education, experience, and training requirements. The higher the value of a job zone, the more complex skills a job requires. Information on job zone is from the ONET Database. Figure D6b splits the sample by the median of the average hourly wage for each occupation, using the 2020 national Occupational Employment and Wage Statistics (OEWS) data sets from BLS. Figure D6c splits the sample by the median of the entry education requirement for each occupation, using OEWS' typical entry-level educational requirement data sets. Figure D6d splits the sample into STEM and non-STEM occupations, using OEWS' STEM data sets.

Figure D7: Progress in Female Labor Market Participation Over Time



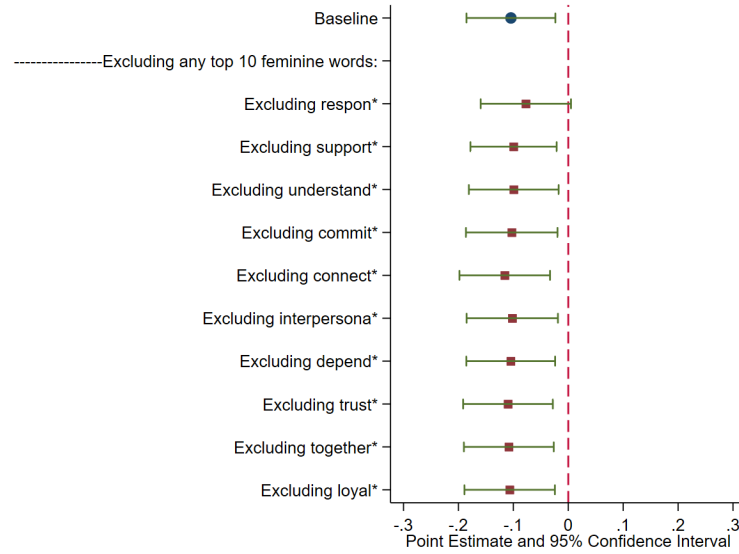
(a) US Gendered Language Usage in Job Ads: %Fem words



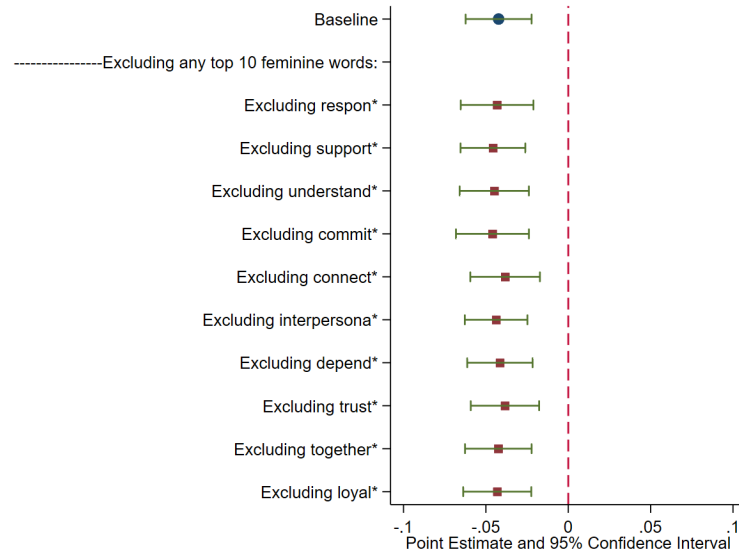
(b) EU Share of Female Employment

This figure shows progress in European female labor market participation and gendered language usage in US job ads over time. Figure D7a plots the monthly average share of feminine words out of all gendered words in job ads covered by LinkUp from 2014 to 2017, before the California Statute (SB 826). Figure D7b plots the female share of the labor force in EU from 1990 to 2020 using data from the World Bank (ILOSTAT database).

Figure D8: Robustness - Excluding Top Feminine Words



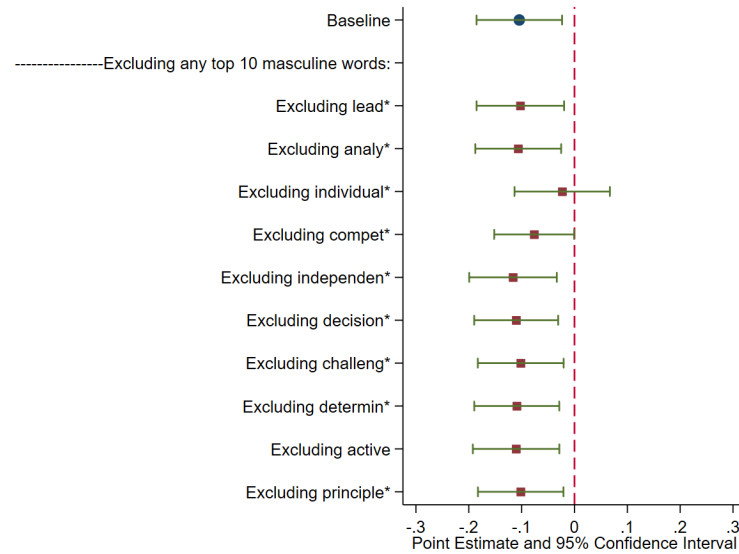
(a) Correlational Analysis



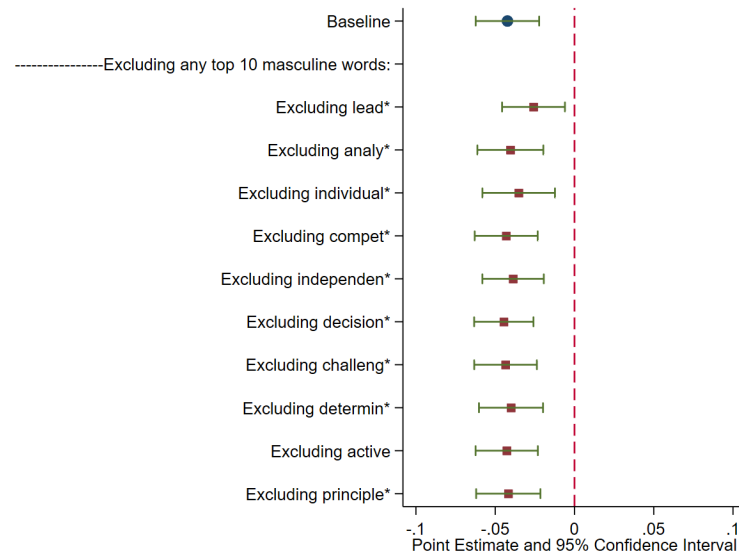
(b) CA DID

This figure shows the robustness of the job-level regression results to excluding top feminine words when computing gendered language usage in each job ad. The asterisk denotes the acceptance of all letters, hyphens, or numbers following its appearance. Both the relationship between the share of female directors and gendered language usage in job ads (Figure D8a) and the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads (Figure D8b) are examined. The regression specifications are the same as those in column 6 of Table 2 and Table 4. The top row of each panel shows the baseline estimate in column 6 of Table 2 and Table 4.

Figure D9: Robustness - Excluding Top Masculine Words



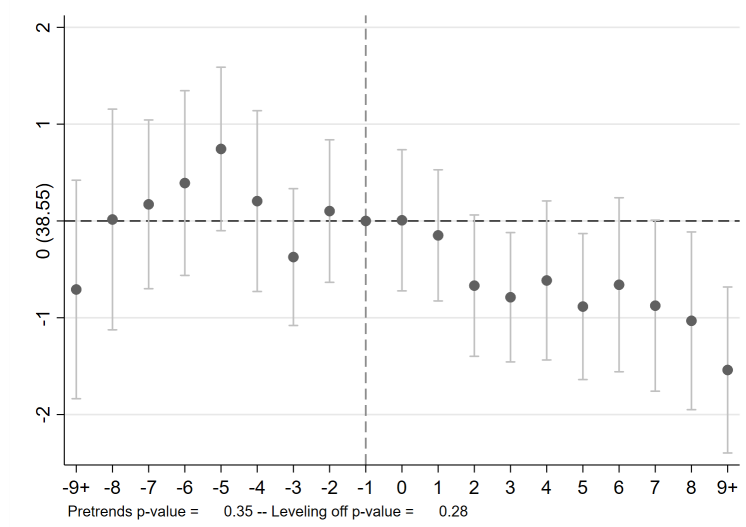
(a) Correlational Analysis



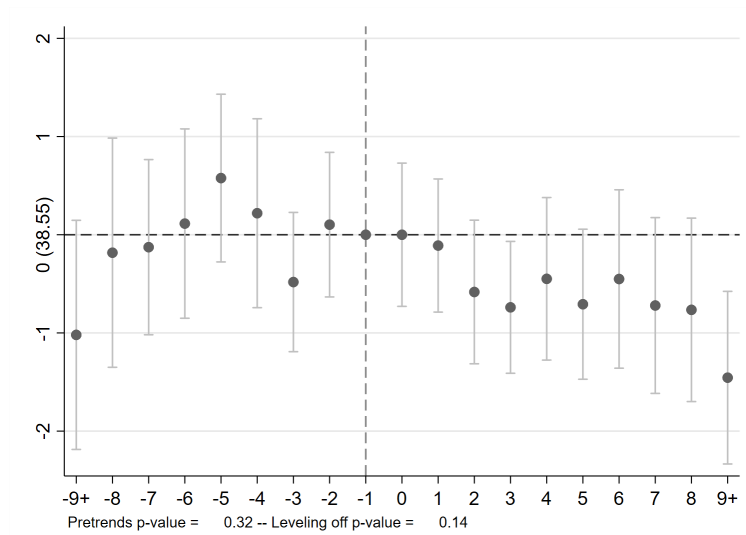
(b) CA DID

This figure shows the robustness of the job-level regression results to excluding top masculine words when computing gendered language usage in each job ad. The asterisk denotes the acceptance of all letters, hyphens, or numbers following its appearance. Both the relationship between the share of female directors and gendered language usage in job ads (Figure D9a) and the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads (Figure D9b) are examined. The regression specifications are the same as those in column 6 of Table 2 and Table 4. The top row of each panel shows the baseline estimate in column 6 of Table 2 and Table 4.

Figure D10: Effect of Board Gender Quotas in Europe on Female Labor Market Participation – Dynamics



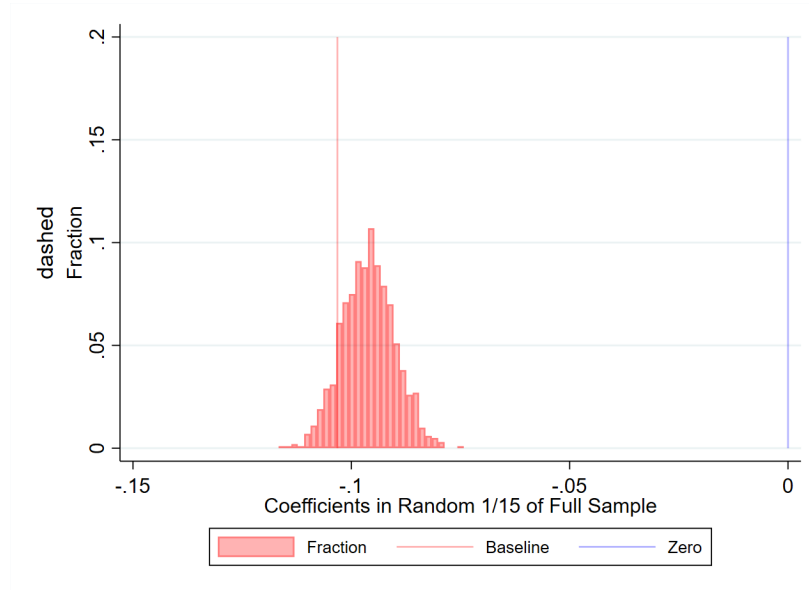
(a) Without Controls



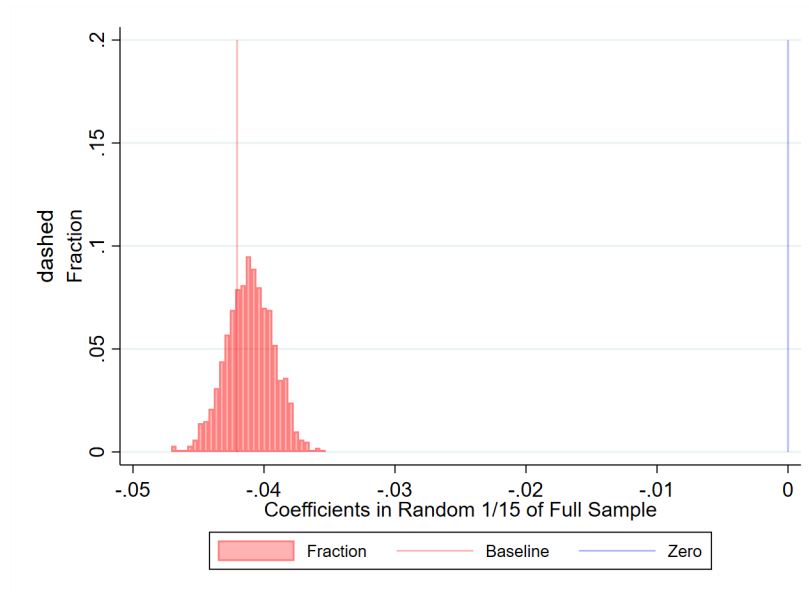
(b) With Controls

This figure plots the dynamic effect of board gender quotas on the share of female employment. Year minus one is the year before the announcement of board gender quotas and is the omitted category. The outcome variable is scaled by 100 for easy interpretation of the regression coefficients. Figure D10a (Figure D10b) presents the estimated coefficients without (with) control variables in the regression. P-values for testing for the absence of pre-event effects and p-values for testing the null that dynamics have leveled off are included. Pointwise 95 percent confidence intervals are included.

Figure D11: Robustness - Random Sampling



(a) Correlational Analysis



(b) CA DID

This figure shows the robustness of the job-level regression results to alternative draws from the full sample. A random 1/15 of the full sample is drawn each time, for 1000 times. Both the relationship between the share of female directors and gendered language usage in job ads (Figure D11a) and the effect of the California board gender diversity statute (SB 826) on gendered language usage in job ads (Figure D11b) are examined. The regression specifications are the same as those in column 6 of Table 2 and Table 4. The red line in each panel indicates the baseline estimate in column 6 of Table 2 and Table 4. The blue line in each panel indicates zero treatment effect.

Table D1: Dictionary of Masculine and Feminine Words in Gaucher et al. (2011)

Masculine	Feminine
active	affectinate
adventurous	child*
aggress*	cheer*
ambitio*	commit*
analy*	communal
assert*	compassion*
athlet*	connect*
autonom*	considerate
boast*	cooperat*
challeng*	depend*
compet*	emotiona*
confident	empath*
courag*	feminine
decide	flatterable
decisive	gentle
decision*	honest
determin*	interpersonal
dominant	interdependen*
domina*	interpersona*
force*	kind
greedy	kinship
headstrong	loyal*
hierarch*	modesty
hostil*	nag
impulsive	nurtur*
independen*	pleasant*
individual*	polite
intellect*	quiet*
lead*	respon*
logic	sensitiv*
masculine	submissive
objective	support*
opinion	sympath*
outspoken	tender*
persist	together*
principle*	trust*
reckless	understand*
stubborn	warm*
superior	whin*
self-confiden*	yield*
self-sufficien*	
self-relian*	

The table lists all the masculine and feminine words identified in Gaucher et al. (2011). The asterisk denotes the acceptance of all letters, hyphens, or numbers following its appearance.

Table D2: Female Directors and Gendered Language Usage in Job Ads
– Heterogeneity: Decentralization in Drafting Job Ads

	(1)	(2)	(3)	(4)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)			
%Fem dir	-0.103** [0.040]	-0.103** [0.040]	-0.136*** [0.046]	-0.111** [0.052]
%Fem dir $\times$ Std Dev (%Fem words)	0.009 [0.041]			
%Fem dir $\times$ Std Dev (#Words)		-0.008 [0.043]		
%Fem dir $\times$ Decentralized (%Fem words)			0.066 [0.074]	
%Fem dir $\times$ Decentralized (#Words)				0.015 [0.078]
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES
Occupation (ONET 6-digit) FE	YES	YES	YES	YES
Obs	1,642,315	1,642,315	1,642,315	1,642,315
Adj R2	0.434	0.434	0.434	0.434

This table shows heterogeneities in the relationship between the share of female directors and gendered language usage in job ads across firms with different levels of decentralization in drafting job ads. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. The outcome variable is the share of feminine words out of all gendered words in a job ad. Decentralization is measured by a lack of consistency in the language used in job ads. For each two-digit ONET-SOC occupation in a firm, we first calculate the standard deviation of gendered language usage (%Fem words) and length of job ads (#Words). We then take the average across all occupations to obtain a firm-level score that captures the degree of decentralization in drafting job ads. A higher score means that a firm writes different ads for functionally similar jobs, suggesting a high level of decentralization. In columns (1) and (2), female director share (% Fem dir) is interacted with the standard deviation of gendered language usage (%Fem words) and length of job ads (#Words), respectively. In columns (3) and (4), the interaction term is a dummy indicating firms with a high level of decentralization (above median). Robust standard errors double clustered at the firm and state (state of a job) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

**Table D3: Female Directors and Gendered Language Usage in Job Ads
– Heterogeneity: Occupation Characteristics**

	(1)	(2)	(3)	(4)
Dep. Var.	Share of feminine words out of all gendered words (%Fem words)			
%Fem Dir	-0.231*** [0.083]	-0.146** [0.061]	-0.137** [0.053]	-0.088** [0.042]
%Fem Dir $\times$ Job Zone	0.046** [0.021]			
%Fem Dir $\times$ Mean Wage		0.002* [0.001]		
%Fem Dir $\times$ Entry Education			0.021** [0.010]	
%Fem Dir $\times$ STEM				0.043* [0.025]
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES
Job Zip $\times$ Year-Month FE	YES	YES	YES	YES
Occupation (ONET 6-digit) FE	YES	YES	YES	YES
Obs	1,152,891	1,207,466	1,208,683	1,208,683
Adj R2	0.461	0.455	0.455	0.455

This table shows heterogeneities in the relationship between the share of female directors and gendered language usage in job ads across different occupations. The unit of observation is a job ad. The sample includes a random 1/15 of the full sample of job ads. The outcome variable is the share of feminine words out of all gendered words in a job ad. In column (1), female director share (% Fem dir) is interacted with the job zone, which is a group indicator from 1 to 5 for occupations that are similar in their education, experience, and training requirements. The higher the value of job zone, the more complex skills a job requires. Information on job zone is from the ONET Database. In column (2), the interaction term is the average hourly wage for each occupation, using the 2020 national Occupational Employment and Wage Statistics (OEWS) data sets from BLS. In column (3), the interaction term is the job entry requirement for each occupation, using OEWS' typical entry-level educational requirement data sets. In column (4), the interaction term is a dummy indicating STEM occupations, using OEWS' STEM data sets. All interaction variables are defined at the six-digit ONET-SOC code level. Robust standard errors double clustered at the firm and state (state of a job) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table D4: Effect of California Statute on Female Directors

	(1)	(2)	(3)	(4)
Dep. Var.	# Fem dir		%Fem dir	
Flag_CA $\times$ Post_2018m9	0.147*** [0.029]	0.122*** [0.038]	0.018*** [0.003]	0.017*** [0.004]
Controls	NO	YES	NO	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES
Firm State $\times$ Year-Month FE	YES	YES	YES	YES
Obs	136,194	86,330	136,194	86,330
Adj R2	0.841	0.834	0.815	0.807

This table shows the effect of the California board gender diversity statute (SB 826) on the number or share of female directors. The unit of observation is a firm-year. Flag_CA is a dummy that equals 1 for public firms headquartered in California, and 0 otherwise. Post_2018m9 is a dummy that equals 1 after September 30, 2018, and 0 otherwise. In columns (1) and (2), the outcome variable is the number of female directors on corporate boards. In columns (3) and (4), the outcome variable is the share of female directors on corporate boards. Robust standard errors clustered at the state (headquarter state of a firm) level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table D5: Descriptions of Female-Friendly Amenities

#	Category	Amenity	Anchor words	Top 20 words
5	Fringe benefits	Paid time off	vacation, pto, sick days, leave, pay time off	leave, pto, vacation, pay time off, sick days, bereavement, pto sick, benefit unlimited, good amount, must use, paternal, lot desire, generous amount, benefit pay, accumulation, hard take, maternal, benefit generous, vto, pay time off sick
10	Working conditions	Work-life balance	work life balance, work life	work life balance, work life, balance ability, maintain healthy, promote healthy, balance none, balance limited, hard balance, imbalance, difficult maintain, culture good, hard maintain, good balance, culture benefit, balance, long hours little, balace, benefit culture, flexibility good, ballance
11	Working conditions	Hours	hours, full time, part time	hours, full time, part time, work full time, part timer, part time employee, college student, get hours, seasonal, get full time, cut hours, position available, full time position, hours cut, require long, normal business, part time job, work part time, hours hours, benefit flexible
12	Working conditions	Work schedule	hours, shift, schedule, flex time	hours, schedule, shift, flex time, scheduling, availability, hour shift, early morning, pick extra, swap, pay flexible, monday friday, late night, inflexible, super flexible, offer flexible, night shift, extremely flexible, schedule change, week advance
16	Working conditions	Teleworking	telecommute, telework, work home, home office, remote	work home, remote, home office, telework, telecommute, one day week, flexible work, schedule ability, set hours, flexible work schedule, option available, flex schedule, days per week, hours ability, make schedule, equipment provide, flexibility ability, provide equipment, remote position, benefit ability
19	Working conditions	Respect/abuse	respect, dignity, abuse, harass, hostile	respect, abuse, hostile, harass, reason, lie, joke, upper management, write, literally, dignity, quit, unless, promise, woman, speak, blame, absolutely, claim, ignore
21	Working conditions	Support	help, support, supportive, encourage	help, support, encourage, supportive, always available, always ready, always happy, nice willing, student need, worker willing, manager willing, available need, wiling, pay school, further education, class size, support teacher, assist need, lots training, administrative support
31	Working conditions	Diversity/inclusion	diversity, ethnic, multicultural, inclusive, lgbtq, inclusion, equality, diverse	diversity, diverse, inclusive, inclusion, equality, lgbtq, multicultural, ethnic, gender, diversity equity, lack diversity, ethnicity, inclusivity, focus diversity, commitment diversity, race gender, patient population, inclusiveness, nationality, student body

This table provides the anchor and top 20 word list of the eight female-friendly amenities from Sockin (2022).

Table D6: Additional Summary Statistics**Panel A: Employee Reviews of Female-Friendly Amenities**

Variable (word frequency)	List of words for amenities	N	Mean	Std. Dev.	Median	p1	p99
Pros	Anchor	54,814	0.49	0.83	0.00	0.00	4.00
Pros	Anchor&Top20	54,814	0.61	0.93	0.00	0.00	4.00
Cons	Anchor	54,814	0.61	1.08	0.00	0.00	5.00
Cons	Anchor&Top20	54,814	0.90	1.37	0.00	0.00	7.00
Pros - Cons	Anchor	54,814	-0.06	0.68	0.00	-2.50	1.50
Pros - Cons	Anchor&Top20	54,814	-0.17	0.88	0.00	-3.75	1.78
Pros - Cons (weighted)	Anchor	54,814	-0.14	0.82	0.00	-3.00	2.00
Pros - Cons (weighted)	Anchor&Top20	54,814	-0.30	1.14	0.00	-4.97	2.08

Panel B: Employee Ratings of Female-Friendly Benefits

Variable	N	Mean	Std. Dev.	Median	p1	p99
Ratings: Overall	180,247	3.15	1.30	3.00	1.00	5.00
Ratings: Compensation and benefits	172,010	3.06	1.49	3.00	0.00	5.00
Ratings: Culture and values	171,542	2.81	1.61	3.00	0.00	5.00
Ratings: Work-life balance	171,960	2.76	1.53	3.00	0.00	5.00

Panel C: European Country-Industry-Year Level

Variable	N	Mean	Std. Dev.	Median	p1	p99
%Fem emp	29,734	0.39	0.23	0.36	0.03	0.91
%Fem dir	29,734	0.19	0.10	0.17	0.03	0.45
Qutoa	29,734	0.24	0.42	0.00	0.00	1.00
Hard quota	29,734	0.15	0.36	0.00	0.00	1.00
Soft quota	29,734	0.08	0.28	0.00	0.00	1.00
Total emp (1,000)	29,734	105.00	305.00	22.40	0.80	1409.00
%Fem population	29,734	0.51	0.01	0.51	0.50	0.54
Log (population)	29,734	16.00	1.44	16.00	12.70	18.20
Population growth	29,734	0.32	0.79	0.30	-1.67	2.40
Log (GDP per capita)	29,734	10.20	0.73	10.20	8.56	11.60
GDP per capita growth	29,734	0.98	3.70	1.38	-8.60	8.37

The table presents additional summary statistics. Panels A and B present summary statistics for variables constructed from employee reviews and ratings of female-friendly amenities and benefits from Glassdoor. Panel C summarizes variables at the European country-industry-year level.

Table D7: Employee Ratings of Female-Friendly Benefits**Panel A: Correlation with Female Directors**

	(1)	(2)	(3)	(4)
Dep. Var.	Ratings (1-5)			
	Overall	Compensation and benefits	Culture and values	Work-life balance
%Fem dir	-0.163 [0.104]	-0.043 [0.112]	0.037 [0.129]	0.1 [0.106]
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES
Obs	180,247	172,010	171,542	171,960
Adj R2	0.129	0.09	0.072	0.067

Panel B: Effect of California Statute

	(1)	(2)	(3)	(4)
Dep. Var.	Ratings (1-5)			
	Overall	Compensation and benefits	Culture and values	Work-life balance
Flag_CA $\times$ Post_2018m9	-0.052*** [0.018]	-0.111*** [0.030]	-0.174*** [0.025]	-0.079*** [0.025]
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year-Month FE	YES	YES	YES	YES
Obs	173,019	164,936	164,479	164,890
Adj R2	0.128	0.094	0.074	0.07

This table shows potential backlash against female employees in the workplace in terms of the provision of female-friendly benefits. The unit of observation is an employee review. Panel A shows the relationship between the share of female directors and employee ratings of female-friendly benefits (on a scale of 1-5). Panel B shows the effect of the California board gender diversity statute (SB 826) on employee ratings of female-friendly benefits. The outcome variables are the overall rating, and ratings for compensation and benefits, culture and values, and work-life balance. Robust standard errors clustered at the firm level are in brackets in Panel A. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets in Panel B. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table D8: Female Directors and Gender-Related Labor Violations**Panel A: Correlation with Female Directors**

	(1)	(2)	(3)	(4)
Dep. Var.	Seed words only		Expanded word list	
	#Violations	log (1 + \$ Penalties)	#Violations	log (1 + \$ Penalties)
%Fem Dir	0.01 [0.011]	0.131 [0.145]	0.041** [0.018]	0.470** [0.213]
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year FE	YES	YES	YES	YES
Obs	11,182	11,182	11,182	11,182
Adj R2	0.124	0.105	0.133	0.116

Panel B: Effect of California Statute

	(1)	(2)	(3)	(4)
Dep. Var.	Seed words only		Expanded word list	
	#Violations	log (1 + \$ Penalties)	#Violations	log (1 + \$ Penalties)
Flag_CA $\times$ Post_2018m9	0.005** [0.002]	0.056*** [0.021]	0.005* [0.003]	0.128*** [0.031]
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Industry $\times$ Year FE	YES	YES	YES	YES
Obs	10,967	10,967	10,967	10,967
Adj R2	0.126	0.106	0.134	0.117

This table shows potential backlash against female employees in the workplace in terms of gender-related labor violations and penalties. The unit of observation is a firm-year. We use case-level data from Violation Tracker and aggregate them to firm-year level. Panel A shows the relationship between the share of female directors and gender-related labor violations and penalties. Panel B shows the effect of the California board gender diversity statute (SB 826) on gender-related labor violations and penalties. In columns (1) and (2), the outcome variable is the number of or the penalties incurred by labor violations that can be labeled as gender-related. These violations mention any of the following seed words in the “description” or “secondary offense” fields: female, gender, hostile, hours, leave, sex, sexual, support, and vacation. In columns (3) and (4), the outcome variable is the number of or the penalties incurred by labor violations that can be labeled as gender-related using an expanded word list. The following words are added as they have high associations with the above seed words in terms of pmi-freq scores (Jin et al., 2021): applicants, women, job, positions, discrimination, environment, medical, sick, and harassment. Robust standard errors clustered at the firm level are in brackets in Panel A. Robust standard errors clustered at the state (headquarters state of a firm) level are in brackets in Panel B. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table D9: Board Gender Quotas in Europe

Country	Passage Date	Quota	Quota Law	Hard/Soft Quota?
NO	12/19/2003	40%	Norwegian Company Act	Hard
ES	3/12/2007	40%	Spanish Equality Act	Soft
IS	3/4/2010	40%	Icelandic Company Act	Soft
FR	1/13/2011	40%	French Cope Zimmerman Company Act	Medium Hard
NL	6/6/2011	30%	Dutch Civil Code	Soft
IT	6/28/2011	33%	Italian Golfo Mosca Company Act	Hard
BE	6/30/2011	33%	Belgian Company Act	Hard
DE	3/15/2015	30%	German Company Act	Hard
AT	7/26/2017	30%	Austrian Company Act	Medium Hard
PT	8/1/2017	33.30%	Portuguese Law	Hard

The table lists board gender quotas in Europe introduced over the period 2003 - 2017, including each quota's passage dates and the relevant quota law. Whether a quota is hard, medium hard, or soft is classified by Mensi-Klarbach and Seierstad (2020).

Table D10: Effect of Board Gender Quotas in Europe on Female Labor Market Participation – Robustness to Alternative Specifications and Estimations

	(1)	(2)	(3)	(4)
Dep. Var.	Share of female employment out of total employment (%Fem Emp)			
Quota	-0.299** [0.119]	-0.399** [0.185]	-0.469** [0.225]	-0.670** [0.311]
Alternative Estimator	Cluster by Industry	Equal-weighted	Borusyak et al. (2021)	CSDID
Controls	YES	YES	YES	YES
Region $\times$ Year FE	YES	YES	YES	NO
Industry $\times$ Year FE	YES	YES	YES	NO
Country $\times$ Industry FE	YES	YES	YES	YES
Year FE	-	-	-	YES
Obs	29,688	29,688	27,545	26,909
Adj R2	0.991	0.943	0.858	0.858

This table shows the effect of board gender quotas on the share of female employment using different specifications or alternative DID estimators. The unit of observation is a country-industry-year. The sample includes 89 two-digit NACE industries in 32 European countries from 2008 to 2020. Column (1) uses alternative clustering at the industry level instead of the country level. Column (2) puts no weight to a country-industry-year in the regression analysis. Column (3) uses the imputation approach in Borusyak et al. (2021). Column (4) uses the estimator in Callaway and Sant'Anna (2021). All outcome variables are scaled by 100 for easy interpretation of the regression coefficients. Robust standard errors clustered at the country level are in brackets. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.