The Hidden Cost of Flat Hierarchies
on Applicant Pool Diversity: Evidence from Experiments*

Reuben Hurst
Ross School of Business, University of Michigan
Ann Arbor, MI 48109
hursre@umich.edu

Saerom (Ronnie) Lee
The Wharton School, University of Pennsylvania
Philadelphia, PA 19104
saeroms@wharton.upenn.edu

Justin Frake
Ross School of Business, University of Michigan
Ann Arbor, MI 48109
jfrake@umich.edu

June 6, 2022

Draft: Please do not circulate or cite without the permission of the authors.

*All authors contributed equally. We thank Matthew Bidwell, Seth Carnahan, Martin Ganco, Lindred Greer, Megan Lawrence, Chris Rider, Olav Sorenson, and Chiara Spina for providing valuable comments on the earlier versions of this manuscript. We also acknowledge the helpful suggestions from seminar participants of Yonsei University, the Wharton School, the Utah/BYU Winter Strategy Conference (2022), the People and Organizations Conference (2021), and the Consortium on Competitiveness and Cooperation Conference Entrepreneurship and Innovation Brownbag Series (2021). Saerom (Ronnie) Lee acknowledges generous financial support from the Wharton Dean’s Research Fund, the Wharton Behavioral Lab, the University of Michigan’s Ross School of Business, and the University of Michigan’s Rackham School of Graduate Studies. All errors in this paper are the sole responsibility of the authors.
The Hidden Cost of Flat Hierarchies on Applicant Pool Diversity: Evidence from Experiments

Abstract

Research on organizational design suggests that, by removing managerial layers, flatter hierarchies can help motivate employees and improve firm performance. Despite these potential benefits, we argue that these organizational structures have an unintended cost of decreasing gender diversity in the applicant pool. Using a field experiment, we show that a flatter hierarchy does not significantly increase the size of the applicant pool but decreases its female representation. Our follow-up survey experiment suggests that this decrease occurs because, compared to men, women perceive flatter hierarchies to offer fewer career opportunities, to be more difficult for them to fit into, and to disproportionately burden them with more work. These findings imply that flatter hierarchies could inadvertently undermine firm performance by curtailing employee diversity.

Keywords: organizational structure; human capital; gender; employee diversity; entrepreneurship
1 Introduction

The longstanding literature on organizational design demonstrates that formal hierarchy—the vertical division of tasks represented by the layers of management (Burton and Obel 2004:75–77, Simon 1997:7)—plays a central role in organizational coordination (Puranam 2018) and adaptation (Ethiraj and Levinthal 2004). However, multi-layered, “tall” organizational structures have received criticism for limiting employee autonomy (Lee 2022), breeding conflict (Greer et al. 2018), and reinforcing organizational rigidity (Csaszar 2013). This criticism has triggered calls for “flatter” structures with fewer hierarchical layers (Burton and Obel 2004:76–77). Recent studies argue that, by removing layers of managerial oversight, flatter hierarchies act as “egalitarian” structures that provide autonomy to employees (Lee and Edmondson 2017, Raveendran et al. 2022), ameliorate in-group biases (Reitzig and Sorenson 2013), and foster organizational flexibility and innovation (Burns and Stalker 1961, Keum and See 2017). This line of argument suggests that, because modern workers tend to prefer more egalitarian workplaces, flatter organizational structures may appeal to potential employees and thus generate larger applicant pools (Reitzig 2022, Sorenson 2022).

Despite these potential benefits, however, we propose that flatter hierarchies can have an important, unintended consequence in attracting human capital. Specifically, we argue that, because a flatter hierarchy can imply a lack of not only managerial supervision but also formal procedures, women may perceive this organizational structure to hinder their career progression within the organization and thus be less attracted than men to a flatter organization. This perception can stem from three potential sources regarding organizational culture, task allocation, and rewards distribution. First, by allowing for more discretion and gender homophilous interaction (Kleinbaum et al. 2013), a flatter hierarchy can breed an insular “bro culture” where women may find it difficult to “fit in” (Chang 2018, Kanter 1977). Second, by providing more latitude for gender-stereotypic expectations about who should perform which tasks (Doering and Thébaud 2017), it can allow men to secure the most rewarding tasks, while disproportionately allocating less desirable ones to women (Babcock et al. 2017, Reskin 1988:65, Wynn 2020:124–125). Finally, this organizational structure can lead employees to engage in more informal competition over rewards distribution, which women, compared to men, are more reluctant to enter into (Flory et al. 2015, Niederle and Vesterlund 2011) and less likely to succeed in (Gneezy et al. 2003, Stuhlmacher and Walters 1999). Hence, if women
perceive that they will face difficulty in fitting in the organizational culture, be burdened with undesirable tasks, and engage in more intense competition, they will perceive flatter organizations to be less amenable to their career progression and will thus be less inclined to apply to these firms. Because the likelihood of hiring women is typically proportional to their representation in the applicant pool (Fernandez and Abraham 2011), this under-representation in the recruitment stage can curtail gender diversity in the workplace (Fernandez and Campero 2017) and ultimately hinder firm performance (Dezső and Ross 2012, Hoogendoorn et al. 2013, Østergaard et al. 2011).

We assess the effect of an organization’s formal hierarchy on its applicant pool in the context of the U.S. labor market, where we document a growing tendency of firms to highlight their flat organizational structures in their recruiting efforts (as shown in Figure 1). Partnering with a U.S. healthcare company, we conducted a realistic, pre-registered field experiment, in which we randomized the description of the hiring firm’s hierarchical structure to around 8,200 job seekers. In contrast to the predictions suggested by prior work (Reitzig 2022, Sorenson 2022), we find that a flatter hierarchy has no statistically significant effect on the size of the applicant pool. However, our results show that it does significantly reduce the proportion of interested female candidates by 18% and lowers the share of female applicants by 25%.

To examine the underlying mechanisms regarding job seekers’ perceptions of flatter hierarchies, we conduct a complimentary survey experiment with approximately 8,500 subjects. Consistent with our arguments, we find that job seekers (regardless of gender) perceive flatter organizations to lack formal procedures but that, compared to men, women perceive flatter organizations to offer disproportionately fewer opportunities for career advancement. Our results provide suggestive evidence that this is due to women perceiving themselves as less likely to fit into the organizational culture and more likely to be allocated heavier workloads. However, our results do not suggest that competition is a significant factor. Furthermore, we find no evidence to suggest that our arguments are contingent on the founder/CEO’s gender. In addition, our results cast doubt on other potential mechanisms regarding perceptions of autonomy, employee voice, and firm quality. Lastly, in a free-response question, our survey participants offer qualitative evidence that women are disinclined to apply to flatter organizations because they perceive less cultural fit, more workload, and fewer career advancement opportunities in firms with flatter hierarchies.

Our work primarily contributes to the literature on organizational design by unveiling the
hidden cost of flat hierarchies in acquiring human capital. Although flatter organizational structures may help motivate existing employees and facilitate their idea generation (Keum and See 2017, Lee and Edmondson 2017, Raveendran et al. 2022, Reitzig 2022), we find that these structures can hinder the recruitment of a diverse set of new employees. As greater employee diversity enhances firm performance and innovation (Dezső and Ross 2012, Hoogendoorn et al. 2013, Østergaard et al. 2011), this finding implies that employee sorting at the recruitment stage may act as an important yet under-explored mechanism that ties organizational structure to firm performance.

Our study also speaks to the literature on entrepreneurship, workplace diversity, and human capital. First, by identifying an unintended consequence of flat hierarchies, we challenge the received wisdom in entrepreneurship research that startups should have a flat hierarchy (Burns and Stalker 1961, Reitzig 2022). Instead, we show that a flat hierarchy can lead to a failure in cultivating a gender-diverse team in a firm’s early years, which can perpetuate “diversity debt” (Wu 2017) that becomes more difficult to reverse as time passes (Alexy et al. 2021, Beckman and Burton 2008). Second, this study adds to research on workplace diversity by revisiting the role of formal hierarchy in gender inequality and segregation. Though our results confirm Acker’s (1990) insight that “organizational structure is not gender-neutral” (p.139), these results counter the proposed solution of creating “non-hierarchical, egalitarian organizations” (p.141). Instead, our findings imply that a flat hierarchy may, in fact, exacerbate workplace gender inequality and segregation. Lastly, by exploring how individuals perceive and self-select into organizations, we answer the recent calls in human capital research to examine the processes underlying team formation (Shah et al. 2019).

2 Theory and hypotheses

2.1 Formal hierarchy

The literature on organizational design defines formal hierarchy as the vertical division of tasks represented by the layers of management in an organization (Burton and Obel 2004:75–77, Puranam 2018:106–126, Simon 1997:7). By adding managerial levels and dividing decision-making responsibilities across these levels, taller hierarchical structures reduce the number of subordinates that each manager directly supervises (Graicunas 1937, Lee 2022). With a smaller span of control and clearer supervisory responsibilities, managers are better able to integrate the efforts of their
subordinates and resolve conflicts among them (Colombo and Grilli 2013, Puranam 2018:113). By fostering coordination, these multi-layered structures can enable organizations to efficiently adapt to their environment (Ethiraj and Levinthal 2004).

Despite these benefits, critics have argued that taller hierarchies can restrict employee autonomy (Lee and Edmondson 2017, Oldham and Hackman 1981), impede idea generation and sharing (Keum and See 2017, Reitzig and Sorenson 2013), generate conflict (Greer et al. 2018), and cause organizational rigidity (Csaszar 2013, Reitzig and Maciejovsky 2015). In the midst of these criticisms, flatter organizational structures with fewer managerial layers have garnered much interest among practitioners and academics, alike (Foss and Klein 2022, Lee 2022, Reitzig 2022).

By imposing fewer layers of management, flatter organizations seek to create distinct organizational cultures and unique approaches to task allocation and rewards distribution (Puranam 2018, Reitzig 2022). By reducing status differentiation and more evenly distributing power among employees, flatter hierarchies may help establish an “egalitarian” culture (Foss and Klein 2014:73), where employees are induced to recognize one another as equals (Siegel et al. 2013:1174). Furthermore, by decreasing managerial oversight and increasing employee autonomy (Lee and Edmondson 2017, Oldham and Hackman 1981), these structures can allow employees to self-organize and allocate their efforts toward tasks that they perceive to match their skills and to be more meaningful (Raveendran et al. 2022, Saxenian 1996:76). Lastly, flatter organizational structures may empower employees to evaluate each other’s task performance and distribute rewards (i.e., pay and promotion) to their peers (Lee and Edmondson 2017:47). As Valve Corporation, a video game company known for its flat hierarchy (Alexy et al. 2021, Lee 2022), illustrated in its “Handbook for New Employees”:

A flat structure removes every organizational barrier. . . . Since Valve is flat, people don’t join projects because they’re told to. Instead, you’ll decide what to work on after asking yourself the right questions. . . . Don’t believe that anyone holds authority over the decision you’re trying to make. . . . Compensation gets adjusted to fit an employee’s internal peer-driven valuation (Valve Corporation 2012).

Such egalitarian organizational culture, autonomous task allocation, and peer-driven rewards distribution in flat organizations can help motivate employees to exert extra effort, generate more creative ideas, voice their opinions, and enjoy a work-life balance (Barbulescu and Bidwell 2013, Keum and See 2017, Lind and Tyler 1988, Morgeson et al. 2005, Reitzig and Maciejovsky 2015). By motivating employees, flatter organizations may foster creativity and innovation and more flexibly
adapt to their rapidly changing environments (Burns and Stalker 1961, Reitzig 2022).

Because formal hierarchy plays such an important role in shaping workers’ experiences, many organizations highlight their hierarchical structure in their recruiting materials (as shown in Figure 1). For example, panel (a) of Figure 1 displays a recent job posting highlighting the firm’s “very flat hierarchy.” In panel (b), we demonstrate using a sample of 6.4 million companies that roughly 9% mention their hierarchical structures in their job postings. Though this proportion has remained relatively consistent over the last decade, the share of those specifically highlighting a flat organizational structure has roughly doubled. In what follows, we elaborate on the potential effect of this organizational structure on the size and diversity of its applicant pool. Specifically, we discuss how job seekers perceive an organization’s hierarchical structure to affect their experience (i.e., pay and promotion, organizational culture, task allocation, and rewards distribution) and how these perceptions influence their attraction to the organization.

[Figure 1 about here.]

2.2 The effect of flatter hierarchies on the size of the applicant pool

Labor markets are arenas for matching workers and employers. Workers seek employers that fit their skills and preferences, while employers seek workers that fit their tasks and workplace (Jovanovic 1979). Because job seekers have limited information about potential employers, they frequently observe information in job postings and infer monetary opportunities (e.g., pay and promotion; Baron et al. 1986) and non-monetary characteristics (e.g., organizational culture, task allocation, and rewards distribution; Abraham 2017, Barbulescu and Bidwell 2013, Judge and Cable 1997) within those organizations. Based on these perceptions, job seekers make their decision to apply (Rynes et al. 1991, Lievens and Highhouse 2003).

Although modern workers are unlikely to have homogeneous perceptions and preferences on organizational characteristics (Barbulescu and Bidwell 2013, Croson and Gneezy 2009, Reitzig 2022:63–78, Wiswall and Zafar 2018), recent studies suggest that they tend to strongly oppose “authoritarian” tall hierarchies with a strict chain of command and tight supervision (Lee and Edmondson 2017:37). Instead, employees tend to perceive work as more meaningful when they have greater latitude and control over their task allocation and rewards distribution (Lind and Tyler 1988, Turco 2016), and thus desire egalitarian culture and autonomous work environment (Bartling et al.
2014, Deci and Ryan 2000, Puranam 2018:125). Hence, if job seekers perceive flatter hierarchies to provide such a workplace, these organizational structures may attract more job applicants and thus increase the size of the applicant pool. In this regard, Reitzig (2022) posits that “[f]lat structures, by providing people with more interesting and satisfying jobs, may … help to recruit and retain employees” (Sorenson 2022:1). Thus, the above discussion leads to the following, pre-registered baseline hypothesis:

**Hypothesis 0 (H0):** Job seekers will be more attracted to jobs offered by organizations they perceive to have a flatter hierarchy.

### 2.3 The effect of flatter hierarchies on the diversity of the applicant pool

As women and men tend to differentially perceive and value organizational characteristics (Barbulescu and Bidwell 2013, Croson and Gneezy 2009, Wiswall and Zafar 2018), formal hierarchy can be consequential to the gender composition of its applicant pool. In line with this reasoning, theories of gendered organizations suggest that, compared to men, women will be disproportionately attracted to a flatter hierarchy, because they perceive this structure as an egalitarian alternative to the authoritarian tall hierarchy that hinders their career progression and reinforces male dominance. These theories argue that tall hierarchies act as a “scientific organization of inequality” (Ferguson 1984:7) that conceals “a gendered substructure” (Acker 1990:154) masked by “a veneer of fairness” (Eagly and Carli 2007:139). In particular, tall hierarchies can calcify gendered assumptions about the “ideal worker” to allocate specific tasks, and impose narrow criteria for rewards distribution that do not account for the societal expectation of domestic responsibilities that women disproportionately bear (DiPrete 1989, Glass and Estes 1997). Thus, flattening the organization’s hierarchy may shift organizational culture, task allocation, and rewards distribution in a way that ameliorates gender inequality and segregation (Acker 1990, Ferguson 1984, Kalev 2009, Ianello 1992).

Besides redressing the gender bias that may be inherent to these multi-layered hierarchies, flatter organizational structures may offer unique organizational characteristics that appeal disproportionately to female job seekers. Most notably, the egalitarian organizational culture, autonomous task allocation, and peer-driven rewards distribution in flat organizations may be more attractive to women (Barbulescu and Bidwell 2013, Wiswall and Zafar 2018), who are often forced to choose between the fulfillment of domestic responsibilities and the pursuit of high-promotability tasks.
(Glass and Estes 1997). In this vein, Ridgeway (2011:175) argues that flatter hierarchies are more appealing to “women . . . who seek equal outcomes with their male colleagues” because they offer more equitable access to resources and opportunities while allowing them to avoid “bad actors” with biased perceptions and agendas. Hence, if women share these optimistic perceptions, they should be relatively more inclined than men to apply to flatter organizations.

Contrary to this perspective, however, we argue that women will be less inclined than men to apply to flatter organizations, because they will have pessimistic perceptions regarding career opportunities within these firms. As one female practitioner highlighted: “[t]his attempt to be more inclusive [i.e., a flatter hierarchy] actually leaves more women left out . . . [and] actually holds them back” (Hamburger 2021; comment added). Next, we outline several potential reasons why women may perceive flatter organizations to be less attractive.

As flatter organizations are less formalized in terms of their vertical division of tasks (Burton and Obel 2004:75–77, Puranam 2018:106–126, Simon 1997:7), job seekers may infer these organizations lack formalized rules and procedures (or often referred to as “formalization”; Burton and Obel 2004:78–80). In organizations with less formal hierarchy and less formalized procedures, social order is called into question (Freeman 1972). Consequently, status disagreement and power contestation can arise, resulting in the emergence of a powerful informal structure (Greer and van Kleef 2010). Consistent with this logic, a former female employee of Valve Corporation elaborated:

There is actually a hidden layer of powerful management structure in the company. . . . There are popular kids that have acquired power in the company, then there are the trouble makers who actually want to make a difference (Maier 2013).

In such informal organizations, men are more likely than women to exercise informal power (Eagly and Karau 1991), partly because men tend to be more vocal and are stereotypically perceived to have greater leadership ability (Berger et al. 1972, Lockheed and Hall 1976), whereas women tend to face legitimacy problems when stepping up as leaders (Carli 1999). As one female worker described: “When a power struggle occurs, it puts women at a disadvantage. Not only do they lack the authority to lead the project or team, but they may also be silenced or sidelined compared to their male coworkers, who may have a more insistent personality type” (Hamburger 2021).

The informal structure in flatter organizations may hinder women’s career progression in several ways. First, because much social interaction is gender-homophilous (Kleinbaum et al. 2013), this informal structure with less managerial supervision and less formalized rules may allow male
employees to create an insular “bro culture” (or an “old boys’ club”) which women may find it
difficult to fit into (Chang 2018, Kanter 1977). The validity of this concern has been articulated by
several female practitioners (e.g., Diamond 2019, Finley 2014, Flower Horne 2014, Hamburger 2021,
Hunt 2017), one of whom noted that “a flat organization has no [formal] defense from becoming a
good ol’ boys club” (Hunt 2017).

Second, the informal structure in flatter organizations may enable male employees to allocate
high-promotability tasks among themselves, while saddling women with disproportionately more,
low-value work (Babcock et al. 2017, Wynn 2020:124–125). Given the persistent gender ratios of
jobs, individuals tend to have gender-stereotypic expectations about who should perform which
job (Doering and Thébaud 2017). When employees self-organize with less managerial oversight
and less formal procedures, these expectations can more easily surface, relegating women to “the
least desirable, most poorly rewarded work . . . such as cleaning [and] disposing of waste” (Reskin
1988:65). Supporting this line of reasoning, Vassallo et al. (2017) found in their survey that 47%
of the female workers in Silicon Valley were asked to perform lower-level tasks (e.g., note-taking
or ordering food) that their male employees were not asked to perform. One female practitioner
offered a more detailed account of her experience in a flat organization:

Women get stuck with the cleaning. . . . While they’re busy doing these very necessary
tasks that keep the organization running smoothly, the community’s more privileged
members have that time free to spend on interesting, high-value projects that raise their
personal profile within the organization. (Flower Horne 2014)

Lastly, the informal structure may cause employees to informally compete against their peers
for pay raises and promotions. Informal competitions tend to be more intense in flatter hierarchies
because there are fewer potential levels of promotion (Baron et al. 1986) but more potential
candidates competing for each promotion opportunity and for a given manager’s attention and
recognition (Graicunas 1937, Halac and Prat 2016). Compared to men, women are more reluctant
to enter into such intense competitions (Flory et al. 2015, Niederle and Vesterlund 2011, Barrymore
et al. 2022) and are less likely to succeed (Gneezy et al. 2003, Stuhlmacher and Walters 1999) due,
in part, to their reluctance to self-promote, positively frame their work, or ask for what they need,
Explaining this process, a female worker commented that a flatter organization “disadvantages
newcomers, women, and minorities who may be less likely to ask for help and advocate for themselves,
furthering a potential experience of inequity” (Diamond 2019).

In sum, we posit that, compared to men, female job seekers will be disinclined to apply to a flatter organization because they will perceive relatively fewer opportunities for career advancement within such an informal organization. We argue that this perception stems from expectations regarding organizational culture, task allocation, and rewards distribution. First, women may recognize a greater likelihood of not fitting into the organization’s insular culture. Second, they may foresee themselves being disproportionately burdened with low-value tasks. Lastly, they may expect more intense informal competition over scarce career advancement opportunities. Hence, we propose the following pre-registered hypothesis:

**Hypothesis 1 (H1):** Compared to men, women will be less attracted to jobs offered by organizations they perceive to have a flatter hierarchy.

A potential boundary condition for this hypothesis could be the founder/CEO gender (Barrymore et al. 2022, Dezső and Ross 2012). Underpinning this hypothesis is the idea, in a flat organization, formal hierarchy will be supplanted by informal structure, where men may secure informal power and reinforce male dominance (Castilla 2011). If job seekers perceive women to have formal authority at the highest levels of the organization, this tendency could be mitigated (Dezső et al. 2016), and our arguments for H1 could be attenuated or even reversed. We test this potential boundary condition in Study 2.

3 Methodological approach

There are two primary obstacles to using observational data to examine how job seekers perceive an organization’s formal hierarchy and thus select into or out of its applicant pool. First, it is difficult to identify the population of job seekers that applied (applicant pool) or were qualified to apply (labor pool) to a given job opening and to observe their perceptions of formal hierarchy. Second, any observed correlations would likely be confounded by selection effects, because an organization’s decision to adopt a flatter hierarchy and an applicant’s decision to apply to the organization are likely correlated with other organizational characteristics (e.g., firm age, size, and industry).

To overcome these empirical challenges, we implement two complementary labor market
experiments, both of which were approved by the Institutional Review Board and pre-registered.¹ First, partnering with a U.S. healthcare startup, we run a field experiment to identify our main effects of interest in a context with high external validity. Second, as it is difficult to directly observe the underlying mechanisms using a field experiment (Chatterji et al. 2016), we conduct a follow-up survey experiment to assess potential causal mechanisms that may be driving our observed effects. In the following two sections, we present each experiment’s design and results.

4 Study 1: Field experiment

4.1 Sampling

For our field experiment, we partnered with a U.S. healthcare company, which allowed us to recruit for two positions: a software engineer and a business development representative.² For these two positions, we identified a pool of job candidates on ZipRecruiter, a widely-used online job search platform. Using this platform’s search function, we found candidates who had updated their profile within the last 30 days and completed at least a bachelor’s degree. For the software engineer position, we limited our search to candidates who included “software engineer” in their past or current job titles and indicated “Healthcare” or “Engineering” as their areas of interest. For the business development position, we limited the search to those who included “business development” in their past or current job titles and indicated “Healthcare” or “Sales and Biz Dev” as their areas of interest. After applying these criteria, we gathered the first 4,200 individuals returned in each search (sorted by the date they last updated their profile) and thus collected a total of 8,400 candidates.³

¹The pre-analysis plan for Study 1 is available at: https://osf.io/8qdu4/?view_only=a78682025f57499b877c09a68a34b. For Study 2, see: https://osf.io/cyztj/?view_only=edd2a7da9bb54261a8c16973df6c1774.
²We chose these two positions because either position, alone, did not offer enough qualified candidates on ZipRecruiter to achieve our target sample size and because they are common across industries. In a post-hoc analysis, we found no significant difference in the main (H0) or interaction (H1) effects between the two jobs.
³We determined the sample size by conducting a pilot study with 100 subjects (i.e., 50 subjects per condition) and using the R package DeclareDesign to calculate the power necessary to detect an effect at the 95% confidence level. It is important to note that we did not select an equal number of women and men from ZipRecruiter’s search results for two reasons. First, ZipRecruiter does not specify job seekers’ gender, presumably due to a conscious decision to mitigate gender discrimination by employers. Second, oversampling on one gender to achieve gender balance would introduce other omitted variables. For example, if the pool of qualified candidates contains more females, oversampling male candidates would likely result in a male sample that differs from the female sample in terms of qualifications.
4.2 Experimental manipulation

As in previous reverse audit studies (e.g., Abraham and Burbano 2022, Flory et al. 2019), we embedded our experimental manipulation in emails sent once to each of the 8,400 potential applicants. Our emails consisted of four paragraphs (see Figure 2). Here, the second paragraph features the randomized manipulation, which is included for the treatment group (i.e., “Flatter” condition) but excluded for the control group (i.e., “No Information” condition). This manipulation states that the company has “a flat organizational structure” with “fewer levels of management than similarly sized startups in our industry.” This statement was intended to reduce the number of hierarchical levels that subjects perceived the firm to have, without introducing other confounding characteristics.

[Figure 2 about here.]

The control condition for the “Flatter” treatment could be operationalized in several ways. One approach would have been to describe a “taller” hierarchy with “more management levels than similarly sized startups in our industry.” But, this approach is problematic because firms rarely state such a description in their job postings (i.e., 0.0003% of companies in our analysis for Figure 1) and thus subjects may negatively respond to the abnormality of the statement. Another approach involves using an “inert” organizational description of a similar length that does not reference formal hierarchy. However, any description of the organization may not be inert, and instead may have an unintended direct effect on the participant’s attraction to the organization. For these reasons, we elected to use a control condition that makes no mention of the hierarchical structure. We address the concerns that our treatment effects are driven by differences in length or the mere mention of hierarchy in a follow-up survey experiment (see Section 5.7), where we demonstrate that the “Flatter” condition is, in fact, perceived by subjects to have fewer levels of management than the “No Information” condition. The follow-up experiment also includes a “Taller” condition to rule out the explanation that our results are driven by the mere mention of the organization’s hierarchy.

The remaining paragraphs were identical for both treatment and control groups. In the first paragraph, we invited candidates to apply for one of the job positions. The third paragraph held constant information on the firm’s founding year, its number of employees, and its desire to grow, because a flatter hierarchy may imply that the startup is exceptionally young (Stinchcombe 1965),
small (Burton and Obel 2004:168), or has no intention to grow (Lee and Kim 2022). The last paragraph included a hyperlink to the partner company’s application web page, which subjects were instructed to visit if they were interested in applying.

Because the randomized manipulation is embedded in our emails, this experiment works if the subjects receive and read these emails. To maximize delivery rates, we used an email delivery service (Mailgun.com), which more than 225 thousand businesses employ for their mass-email marketing. Using this service, we sent an email once to each subject, tracked whether the email was successfully delivered, and removed from the sample the 132 subjects that did not receive the emails. Among the remaining 8,268 email recipients, some subjects could have decided not to read our emails under any circumstance (i.e., “never-takers”; Angrist and Pischke 2008:158). However, because subjects decide whether to read the email before having any information about which condition they have been randomly assigned to, the non-compliance should be unrelated to our treatment assignments. Hence, this non-compliance will bias the results towards zero, making it more difficult for us to find statistically significant effects.  

4.3 Measurement

**Dependent variables:** We measured attraction in two ways. First, our pre-registered outcome \( \text{Click}_i \) captures whether subject \( i \) clicked the hyperlink, embedded at the bottom of the email, to the partner company’s application web page. \(^5\) We used this measure for two main reasons. First, because the treatment and the hyperlink are both embedded in the email, visiting the application website is the first behavioral outcome that we can observe in terms of the subject’s attraction. Second, this outcome is unlikely to be affected by other information about the company, which subjects may observe after visiting the company’s application web page but before deciding whether to apply. Despite these two benefits, this measure may not accurately reflect a subject’s interest in applying to the company. For example, even if they are not interested, subjects in the “Flatter” condition might click on the hyperlink because they are curious to learn about a flat organization.

---

\(^4\)This non-compliance could also bias our results if women and men vary in their behavior of reading emails. However, in a large-scale study of more than two million users exchanging 16 billion emails, Kooti et al. (2015) find no significant variation in email behavior by gender.

\(^5\)To identify who clicked the hyperlink, each subject’s hyperlink included a unique URL with an identifier. Thus, when measuring \( \text{Click} \), we were able to include the subjects who copied and pasted its URL into a browser, instead of clicking on the hyperlink in the email. All of the hyperlinks were directed to the same web page, regardless of the assigned treatment condition.
To address this concern, we also measure $Apply_i$, which is equal to one if subject $i$ submitted an application via our partner company’s web page.\textsuperscript{6}

**Independent variable:** Our treatment variable $Flatter_i$ indicates whether subject $i$ received an email stating that the organization’s hierarchy is flatter than its competitors’ (see Figure 2).

**Individual attributes:** As ZipRecruiter does not collect information on job seekers’ gender (presumably to reduce gender discrimination by employers), we used the website Gender API’s machine-learning algorithm to predict their gender ($FemaleSubject_i$) with their full name.\textsuperscript{7} In addition, we accounted for whether the subject was invited to the software engineer position ($Software_i$), as opposed to the business development position. To assess whether our randomization procedures achieved balance between the treatment and control groups, we also collected applicant characteristics using self-reported information on ZipRecruiter (e.g., short bios and résumés). These attributes include the subjects’ most recent job title (i.e., whether it is a managerial position), work experience (in months), and highest level of education (i.e., undergraduate, master’s, or doctorate).

### 4.4 Summary statistics and randomization balance

Among the 8,400 emails we sent to the job candidates, 132 were not delivered because these emails were blocked by the subjects’ email servers. In addition, Gender API’s algorithm was unable to predict the gender of 97 candidates. Thus, our final sample was 8,171 candidates. The first column of Table 1 shows the summary statistics across all subjects. Of these 8,171 subjects, 28.5% showed interest in the open positions by clicking on the hyperlink, and 9.8% applied to the company. Approximately 26.9% of subjects were women, and 31.8% held a managerial position in their most recent job. In terms of the highest level of education, 35.9% had a master’s degree, while 1.7% had completed a doctorate. In turn, the second and third columns provide the summary statistics by treatment condition (i.e., “No Information” vs. “Flatter”). The differences in observable characteristics of candidates between conditions are small and insignificant, suggesting that randomization was successful and that our estimates can be interpreted as causal.

\textsuperscript{6}We pre-registered this as an exploratory outcome because the partner company could not guarantee our access to all applications prior to conducting the experiment.

\textsuperscript{7}According to its website, Gender API has a database of more than six million names in 189 countries. In an independent study, Sebo (2021) finds that this algorithm is one of the two most accurate tools to predict gender based on names (with a misclassification rate of only 1.8%).
4.5 Estimation

To assess the effect of a flatter hierarchy on attraction (H0), we estimate the following equation:

\[ \text{Attraction}_i = \beta_0 + \beta_1 \text{Flatter}_i + \beta_2 \text{Software}_i + \varepsilon_i \]  

where \( \text{Attraction}_i \) stands for one of the two dependent variables (i.e., \( \text{Click}_i \) or \( \text{Apply}_i \)).\(^8\) The term \( \beta_1 \) represents the average treatment effect of a flatter hierarchy relative to the “No Information” condition. The terms \( \beta_0 \) and \( \varepsilon_i \) are the intercept and the random error term, respectively. In turn, the heterogeneous treatment effects with respect to gender (H1) are estimated using the equation:

\[ \text{Attraction}_i = \beta_0 + \beta_1 \text{Flatter}_i + \beta_2 \text{FemaleSubject}_i + \beta_3 \text{Flatter}_i \times \text{FemaleSubject}_i + \beta_4 \text{Software}_i + \varepsilon_i \]  

where \( \beta_3 \) is the difference in the average treatment effect of the “Flatter” condition for women relative to men. Because our two measures of attraction are binary, we estimate Equations 1 and 2 using a linear probability model (LPM) with robust standard errors.

4.6 Results

The results of Study 1 are reported in Table 2 (for bar graphs, see Figure 5 in Online Appendix A1). Here, Models 1 to 4 measure attraction in terms of whether subjects showed interest in the open positions by clicking on the email’s hyperlink to the company’s application web page (i.e., \( \text{Click} \)). In turn, Models 5 to 8 measure attraction in terms of whether subjects applied to the job (i.e., \( \text{Apply} \)).\(^9\)

We begin by examining the effect of a flatter hierarchy on subjects’ propensity to click in Models 1 to 4. Contrary to H0, Model 1 shows that the estimated treatment effect is small, negative and statistically insignificant at canonical levels (\( p = 0.724 \)). This null result could imply one of the following. First, it may suggest that our treatment failed to manipulate the participants’ perceptions of hierarchy. However, we believe this is unlikely because Study 2, which uses a similar treatment,\(^8\) in the regression for \( \text{Apply} \), we do not condition on or control for \( \text{Click} \) because it is an intermediate outcome of our treatment (i.e., \( \text{Flatter} \)) and is thus a “bad control” that can bias the estimates (Angrist and Pischke 2008:64–68).\(^9\) These results are consistent when using logistic regression (see Table 6 in Online Appendix A2).
shows that the “Flatter” condition is perceived by subjects to have fewer levels of management than the “No Information” condition (see Section 5.7). Second, given that our manipulation very likely succeeded, this finding may indicate that a flatter hierarchy does not impact in any meaningful way the subjects’ propensity to click on the hyperlink. Lastly, this insignificant result may mean that a flatter hierarchy does affect the subjects’ propensity to click but elicits offsetting responses. For example, as argued in H1, a flatter hierarchy may increase men’s propensity to click but decrease women’s propensity to click. We explore this explanation and find supportive evidence below.

In line with H1, Model 2 indicates that, compared to men, women were five percentage points less likely to click the hyperlink and show interest in the “Flatter” condition compared to the “No Information” condition ($p = 0.023$). To assess this divergence in the propensity to click, Models 3 and 4 estimate the treatment effect of a flatter hierarchy by splitting the observations by gender. Model 3 reveals that women were approximately four percentage points less likely to click in the “Flatter” condition compared to the “No Information” condition ($p = 0.032$). In contrast, Model 4 shows that men were one percentage point more likely to click in the “Flatter” condition, but the 95% confidence interval overlaps zero. Put differently, while 28% of those who clicked the hyperlink in the “No Information” condition were women, only 23% of those who clicked in the “Flatter” condition were women. Thus, perceptions of a flatter hierarchy reduced the proportion of interested female candidates by approximately 18%.

In Models 5 to 8, we examine the effect of a flatter hierarchy on subjects’ propensity to apply for the job. These results are remarkably consistent with Models 1 to 4. In line with Model 1, Model 5 exhibits no main effect of a flatter hierarchy on the propensity to apply. Thus, we find no evidence to support H0 that a flatter hierarchy attracts more applicants. However, like Models 2 to 4, Models 6 to 8 support H1 that a flatter hierarchy decreases the female representation in the applicant pool. Model 6 reveals that, compared to men, women were three percentage points less likely to apply in the “Flatter” condition than in the “No Information” condition ($p = 0.015$). Similarly, Model 7 shows that women were three percentage points less likely to apply in the “Flatter” condition than in the “No Information” condition ($p = 0.015$). In contrast, Model 8 indicates that men were about half a percentage point more likely to apply in the “Flatter” condition, but the 95% confidence interval overlaps with zero. Simply put, whereas 27% of the applicants were women in the “No Information” condition, only 20% were women in the “Flatter” condition. Hence, a flatter hierarchy
reduced the share of female applicants by 26%.

Taken together, Study 1 demonstrates that a flatter organizational structure with fewer levels of management does not substantively affect the size of the applicant pool, but does significantly decrease the representation of women in it.

5 Study 2: Survey experiment

We complemented our field study with a pre-registered survey experiment on MTurk. This follow-up experiment had four main objectives. First, it allows us to test our theory on a broader population of subjects (those without college degrees, those with lower-paying jobs, etc.). Second, this experiment enables us to address the concern that the results in Study 1 were driven by the mere mention of hierarchy, not by perceptions of how flat the organization’s hierarchy is. We do so by (1) including a “Taller” condition in addition to the “Flatter” and “No Information” conditions and (2) performing a manipulation check using a direct measure of the number of hierarchical levels that the subjects perceived in response to these conditions. Third, the survey experiment allows us to examine the mechanisms that we were unable to explore in a field study. Lastly, it enables us to manipulate the founder/CEO’s gender and study whether it is a potential boundary condition of our theory (as mentioned in Section 2.3).

One common concern with using MTurk is whether its workers are representative of U.S. workers. Recent studies (e.g., Difallah et al. 2018, Moss et al. 2020, Snowberg and Yariv 2021) have shown that MTurk workers are similar to the U.S. population in terms of occupational and racial composition, but they tend to be younger, more likely to be female, more educated, and more likely to be employed. We also observe these comparable demographic attributes in our 8,498 subjects (see Table 7 in Online Appendix A3). Although their median household income tends to be less than that of the U.S. population (e.g., $47,000 vs. $57,000; Difallah et al. 2018), most of them characterize their MTurk work as a paid leisure (56%) and a financial source for non-essential expenses (69%), whereas only 8% consider it a full-time job (Moss et al. 2020). Despite these differences, Snowberg and Yariv (2021) demonstrate that MTurk workers and a representative sample of the U.S. population show similar comparative statics and correlations between behaviors. Furthermore, Kees et al. (2017) find that, compared to other survey data sources (e.g., student and professional panel samples), MTurk offers similar or more reliable data. Albeit not without
limitations, we believe MTurk offers a broad and reliable sample of U.S. workers to test our theory.

5.1 Sampling

In this experiment, we posed as a human resources analytics company and recruited approximately 9,000 MTurk workers in the U.S.\textsuperscript{10} These subjects were asked to complete a task entitled “Give feedback on recruiting material,” which described a job posting for a part-time, remote copy-editor position. We chose a remote position for two reasons: (1) to increase the realism of the experiment, as the subjects are already working remotely and our intended sample size was too large to be geographically focused, and (2) to increase the generalizability of our findings to non-traditional workers. This remote work, however, may be a conservative setting because the perceived effects of hierarchy may be muted when subjects do not expect in-person interactions with coworkers or may not aspire to be promoted. MTurk workers may also have stronger preferences for informality and autonomy that flatter organizations may be perceived to offer.

5.2 Experimental manipulations

As in Study 1, we embedded experimental manipulations in the recruiting material. This job posting consisted of four sections (see Figure 3). The first section described the open position (i.e., “part-time, remote copyeditor”). The second section (“About Us”) provided information on the company. The third and fourth sections listed the job responsibilities and requirements, respectively. Here, the first, third, and fourth sections remained identical across all conditions.

![Figure 3 about here.]

The second section included our manipulations (for details, see Figure 4). As in our field study (see Figure 2), the “Flatter” condition stated that the firm has a “flat organizational structure” with “fewer levels of management than similarly sized startups in [the] industry.” For the “No Information” condition, this statement was omitted. As discussed in Section 4.2, we chose to provide no information as a control condition in Study 1 because any attempt to include an “inert” condition that described some organizational characteristic might itself have an unintended effect on attraction. To address the concern that the observed effects of the “Flatter” condition in Study 1 may be due

\textsuperscript{10}For the sample size, we ran a pilot study with 450 subjects (i.e., 50 subjects per condition) and used the R package \texttt{DeclareDesign} to compute the power necessary to detect an effect at the 95% confidence level.
to the differences in length or the mere mention of a flatter hierarchy, we also include a “Taller” condition. For this condition, the second section mentioned that the firm has a “tall organizational structure” with “more management levels.” To ensure that we successfully manipulate perceptions of hierarchy, the “Flatter” and “Taller” conditions included illustrative graphics.\footnote{Because hierarchical structures are commonly represented by two geometric shapes (i.e., a ladder or a pyramid) and these shapes can have different consequences on subjects’ perceptions (Yu et al. 2019), the geometric shape of both illustrative graphics is kept constant (i.e., a pyramid) but with a different number of layers.}

Because the effects of an organization’s hierarchy may depend on the gender of its leadership (as discussed in Section 2.3), we test this boundary condition by manipulating the founder/CEO’s gender. In the second section (i.e., “About Us” in Figure 3), this gender manipulation was implemented by including a gender-typical name (i.e., “Jessica Chandler” or “Michael Chandler”), along with a typically male or female face photo (for details, see Figure 4). These pictures were chosen from the Chicago Face Database, which provides high-resolution, standardized face photographs. We chose two photos that are similar in terms of perceived age, race, attractiveness, and smile intensity.

Overall, our survey experiment consists of nine interventions because the organization’s hierarchical structure (i.e., no information vs. flatter vs. taller) and the founder/CEO’s gender (i.e., no information vs. female vs. male) each have three conditions.

5.3 Procedure

For this survey experiment, subjects were asked to first review the recruiting material.\footnote{To prevent bots from contaminating this experiment, we inserted reCAPTCHA before this recruiting material.} After reviewing this material with its experimental manipulations, they advanced to a series of survey questions. The first set of questions checked whether the two treatments succeeded in manipulating the theoretical constructs of interest. For the organization’s hierarchical structure, subjects were asked to indicate how many management levels they believed the firm has between its founder/CEO and entry-level employees. In turn, for the founder/CEO’s gender, they were asked to specify their perception of the founder/CEO’s gender, the direct supervisor’s gender, and the proportion of female employees. Then, the second set asked about the extent to which the subjects find the firm attractive and their perceptions of various organizational characteristics (e.g., informality, fit, workload, competition, and career advancement opportunities within the firm). This set also
included a free-response question in which they shared their impression of a flatter hierarchy (if in either the “Flatter” or “No Information” condition) or a taller hierarchy (if in the “Taller” condition). The last set of questions inquired regarding the subject’s demographics.

5.4 Measurement

Dependent variable: We measure attraction using Highhouse et al.’s (2003) five-item index of “employee recruitment and organization choice” (for details, see Table 8 in Online Appendix A4). This survey-based measure, rather than actual job applications, was used because Amazon prohibits recruiting subjects for opportunities outside of MTurk. Subjects responded to each of the five items on a five-point Likert scale, ranging from “Strongly Disagree” (1 point) to “Strongly Agree” (5 points). After reverse-coding the second item, we averaged these scores to calculate attraction.

Independent variables: Unlike Study 1, which had two conditions for hierarchical structure, the survey experiment had three conditions: “Flatter,” “No Information,” and “Taller.” Setting the “Taller” condition as the comparison group, we coded these conditions using two binary variables:

\[
\text{NoInformation}_i = \begin{cases} 
1, & \text{if subject } i \text{ received the “No Information” condition} \\
0, & \text{otherwise (i.e., “Flatter” or “Taller” condition).}
\end{cases}
\]

\[
\text{Flatter}_i = \begin{cases} 
1, & \text{if subject } i \text{ received the “Flatter” condition} \\
0, & \text{otherwise (i.e., “No Information” or “Taller” condition).}
\end{cases}
\]

To measure the subject’s perception of how flat the organization’s hierarchy is, we reverse-coded the answers to the following question: “Based on the posting, how many levels of management would you estimate we have between our entry-level employees and our CEO?” Although we assess perceived flatness, the variable is named Flatness (instead of PerceivedFlatness), for simplicity.

Individual attributes: We used the subjects’ responses to the demographic questions to record various individual attributes. These attributes include the subjects’ gender, current employment status, job title, work experience, and highest level of education. Among these attributes, the key variable of interest is FemaleSubject, which equals 1 if the subject is a woman; 0, otherwise.

Mechanisms: As the main mechanism of interest, we examine the subjects’ perception of career advancement opportunities (Opportunity), which was computed by averaging the five-point Likert scales of the two items on the likelihood of getting a pay raise or a promotion (for details, see Table 9
in Online Appendix A4). For the potential micro-mechanisms behind the perception of career advancement opportunities, we assess perceptions of informality, fit, workload, and competition. First, perceived informality (Informality) was measured using the average score for the reverse-coded five-point Likert scales of the four items on how clearly defined procedures would be regarding performance evaluations, compensation, promotion, and conflict resolution. Second, perceived fit (Fit) was operationalized by the five-point Likert scale of whether the subject will fit in well at the organization. Third, perceived workload (Workload) was derived using the five-point Likert scale of whether the subject anticipates the workload to be heavy. Lastly, we measured perceived competition (Competition) by the five-point Likert scale of whether the work environment is competitive.

As for other plausible mechanisms, we explore perceptions of autonomy, fairness, employee voice, and firm quality. First, perceived autonomy was measured using the average score of the five-point Likert scales for Lumpkin et al.’s (2009) five-item index of “autonomy orientation,” which captures the extent to which employees would have the freedom to pursue ideas without their supervisor’s approval. In turn, perceptions of fairness and employee voice were computed by the (average) score of five-point Likert scales for one or more items (for details of each variable measurement, see Table 9). Lastly, for the perception of firm quality, we considered the perceptions of the founder/CEO’s competence and the firm’s likelihood of success, which were measured using the five-point Likert scale of their respective item.

**Boundary condition:** To examine the potential boundary condition of leadership gender, we manipulated the founder/CEO’s gender and coded the variable FemaleFounder as 1 if the subject received the female founder/CEO condition; 0, otherwise.

### 5.5 Summary statistics and randomization balance

In Table 3, the first column presents the summary statistics of subjects in the survey experiment. Among the 9,000 subjects recruited, 8,498 completed the survey and passed the manipulation checks for leadership gender. The majority of these 8,498 subjects were women (58.4%). In terms of the highest level of education, 44.0% completed less than a bachelor’s degree, while 38.6% had a bachelor’s degree and 17.4% a graduate degree. While most had more than six years of work experience (66.4%), only 24.7% held a managerial position in their most recent job. Overall, compared to the Study 1 participants, our survey experiment subjects were more likely to be female,
less educated, and less likely to hold a managerial position. However, compared to the population of U.S. workers, these subjects are more likely to be female, more educated, and more likely to be employed (for more information on their demographics, see Table 7 in Online Appendix A3).

Next, we turn our attention to the second through fourth columns in Table 3, which present the summary statistics of the candidates’ observable characteristics by treatment condition. The differences across these conditions are small and have confidence intervals that substantially overlap with zero, implying that randomization was successful and that regression estimates can be interpreted as causal.

5.6 Estimation

5.6.1 Main Effect

For H0, we first estimate the effect of hierarchy on attraction using the following specification:

\[ \text{Attraction}_i = \beta_0 + \beta_1 \text{NoInformation}_i + \beta_2 \text{Flatter}_i + \varepsilon_i \]  

(3)

where \( \beta_1 \) and \( \beta_2 \) each represent the average treatment effect of the “No Information” and “Flatter” conditions relative to the “Taller” condition.

Because we theorize in Section 2 that a firm’s hierarchy affects its applicant pool through the job seeker’s perceptions of its hierarchy, we use a two-stage least squared (2SLS) approach to test whether the relationship between a firm’s hierarchy and a subject’s attraction is mediated by the subject’s perception of flatness. Specifically, we use the randomly assigned treatments for hierarchical structure (\( \text{NoInformation}, \text{Flatter}, \) and \( \text{Taller} \)) to instrument the subject’s perception of flatness (\( \text{Flatness} \)). This 2SLS approach allows us to not only check whether our manipulations affected the subject’s perceptions of hierarchy but also isolate the changes in perceived flatness caused by our manipulations in the first-stage regression. In the second-stage regression, it provides a more precise, continuous operationalization of our theoretical construct of interest, while maintaining the internal validity that comes from the random assignment of subjects into conditions. Hence, we estimate the following 2SLS regressions, where the first stage is:

\[ \text{Flatness}_i = \gamma_0 + \gamma_1 \text{NoInformation}_i + \gamma_2 \text{Flatter}_i + \epsilon_i \]  

(4)
and the second stage is:

\[ \text{Attraction}_i = \beta_0 + \beta_1 \text{Flatness}_i + \varepsilon_i \]  

(5)

5.6.2 Interaction Effects

For H1, we examine whether the effect of hierarchy on attraction varies by gender using the equation:

\[ \text{Attraction}_i = \beta_0 + \beta_1 \text{NoInformation}_i + \beta_2 \text{Flatter}_i + \beta_3 \text{FemaleSubject}_i \\
+ \beta_4 \text{NoInformation}_i \times \text{FemaleSubject}_i + \beta_5 \text{Flatter}_i \times \text{FemaleSubject}_i + \varepsilon_i \]  

(6)

where \( \beta_4 \) denotes the extent to which attraction to the “No Information” condition relative to the “Taller” condition differs by gender, and \( \beta_5 \) represents the extent to which attraction to the “Flatter” condition compared to the “Taller” condition varies by gender.

We then employ the 2SLS approach to estimate the extent to which the effects of hierarchical structure on attraction are mediated through perceived flatness. For the first stage, \( \text{Flatness}_i \) and \( \text{Flatness}_i \times \text{FemaleSubject}_i \) are instrumented using the following specifications:

\[ \hat{\text{Flatness}}_i = \gamma_0 + \gamma_1 \text{NoInformation}_i + \gamma_2 \text{Flatter}_i + \epsilon_i \]  

(7)

\[ \hat{\text{Flatness}}_i \times \text{FemaleSubject}_i = \gamma_0 + \gamma_1 \text{NoInformation}_i + \gamma_2 \text{Flatter}_i \\
+ \gamma_3 \text{NoInformation}_i \times \text{FemaleSubject}_i \\
+ \gamma_4 \text{Flatter}_i \times \text{FemaleSubject}_i + \epsilon_i \]  

(8)

In turn, the second stage is estimated as follows:

\[ \text{Attraction}_i = \beta_0 + \beta_1 \hat{\text{Flatness}}_i + \beta_2 \text{FemaleSubject}_i + \beta_3 \hat{\text{Flatness}}_i \times \text{FemaleSubject}_i + \varepsilon_i \]  

(9)

where \( \hat{\text{Flatness}}_i \) and \( \hat{\text{Flatness}}_i \times \text{FemaleSubject}_i \) are the predicted values of perceived flatness and its interaction with the gender of the subject, respectively. Here, the parameter of interest, \( \beta_3 \), is the heterogeneous local average treatment effect by gender of the perceived flatness for those subjects that the treatments caused to perceive more or fewer hierarchical levels. Accordingly, for each mechanism, we use its respective measure as the dependent variable (instead of \( \text{Attraction} \)) and apply the above estimation procedures.
5.7 Main results

The main results are displayed in Table 4 (for bar graphs and margins plots, see Figures 6 and 7 in Online Appendix A5). Here, Models 1 and 2 estimate Equations 3 and 6, respectively, using OLS regression. Model 1 demonstrates that, on average, subjects (of all genders) are least attracted to the “Taller” condition and most attracted to the “Flatter” condition ($p < 0.001$). These results cast doubt on the alternative explanation for our Study 1 results that the effects of the “Flatter” condition, relative to the “No Information” condition, were driven by the length of the email or the mere mention of a flatter hierarchy. But, these results differ from the Study 1 results in that they show a significant main effect of a flatter hierarchy on attraction, thus supporting $H_0$. One explanation for this difference is that, given their selection into remote work, MTurk workers likely have a higher baseline preference than our field-study subjects for informality and autonomy associated with flatter hierarchies (as discussed in Section 5.1).

To test $H_1$, we turn to Model 2, which demonstrates that, relative to men, women are somewhat less attracted to the “No Information” condition than to the “Taller” condition ($p = 0.110$). However, compared to men, women are significantly less attracted to the “Flatter” condition than to the “Taller” condition ($p = 0.018$).

Using the 2SLS regressions, we find consistent results in Models 3 and 4. First, for the first-stage regression, Model 3 shows strong evidence that our treatments for hierarchical structure (i.e., “Flatter,” “No Information,” and “Taller” conditions) influenced the perception of flatness ($Flatness$) in their respective directions. Specifically, compared to the “Taller” condition, the “No Information” condition was perceived to have 1.2 fewer levels of management ($p < 0.001$). In turn, the “Flatter” condition was perceived to have 2.5 fewer hierarchical levels than the “Taller” condition ($p < 0.001$). Moreover, the first-stage Kleibergen-Paap rk Wald $F$ statistic is 1,622.6, which suggests the instrument is sufficiently strong. Finally, Hansen’s $J$ statistic is 0.726, which fails to reject the null that the instruments are valid.

Because we instrumented the perceived flatness ($Flatness$) using the randomized variables (i.e., $NoInformation$ and $Flatter$) in the first-stage regression, the coefficient estimate of the variable
Flatness in the second-stage regression can be interpreted as the causal effect of one fewer hierarchical level on attraction. For the second-stage regression, Model 4 indicates that, compared to men, women are disproportionately less attracted to a flatter hierarchy ($p = 0.021$). Taken together, these main results of our survey experiment are consistent with our field-study findings that a flatter hierarchy reduces the female representation in the applicant pool.

5.8 Results for the mechanisms of interest

Given the consistent results for H1 in the field and survey experiments, we now examine the mechanisms underlying the relationship between a flatter hierarchy and the gender diversity in the applicant pool. The results for these mechanism tests are presented in Table 5.

We begin by examining Models 1 to 3, which present the results for the main mechanism: the perception of career advancement opportunities (Opportunity). First, Model 1 indicates that, on average, subjects (of all genders) perceive the fewest career advancement opportunities in the “Flatter” condition, followed by the “No Information” and then “Taller” conditions. As one survey participant pointed out in our free-response question, “[a flat hierarchy] would mean to me that you would most likely stay in the position you were hired into and there would be few opportunities to advance.” Model 2 suggests that, compared to men, however, women perceive that they will have even fewer opportunities in a flatter hierarchy than in a taller one ($p < 0.001$). These results are consistent in Model 3, which reveals that, relative to men, women perceive fewer career advancement opportunities as the perceived flatness increases ($p < 0.001$). In accordance with these results, one female subject mentioned that “I think of [a flatter structure of having] prejudice in the workplace preventing upward mobility of marginalized employees.” Another female participant suggested that a flatter hierarchy has “the potential for decreasing unnecessary bureaucracy. However, that alone will not solve systemic issues like those faced by women in many companies.” These findings support our argument that women, compared to men, find a flatter organization less attractive because they perceive fewer career advancement opportunities in a flatter hierarchy.

We move on to discuss Models 4 to 12, which display the results for the potential micro-mechanisms behind this perception of decreased career advancement opportunities: perceptions
of informality, fit, workload, and competition. First, Models 4 to 6 show the results for perceived informality (Informality). Consistent with our argument, Model 4 illustrates that a flatter structure, compared to a taller one, is perceived to be less formalized ($p < 0.001$). Furthermore, both Models 5 and 6 imply that this perception of informality in a flatter hierarchy does not vary by gender. Accordingly, both male and female subjects acknowledged this informality to a similar degree in the free-response question. For example, male subjects noted that a flatter structure is a “fluid” and “informal workplace with fewer rules.” Likewise, female participants expressed that “[in a flatter structure], there may be less formal processes for work review, promotion, etc.” It is important to note that our theory does not require that women and men perceive different levels of informality. Rather, it suggests that higher levels of informality will differentially disadvantage women, making them less attracted to flatter hierarchies.

In turn, Models 7 to 9 present the results for perceived fit (Fit). Albeit Model 7 shows that subjects, in general, perceive that they fit better in a flatter structure than in a taller one ($p < 0.001$), Model 8 demonstrates that, compared to men, women perceive themselves to fit less in a flatter hierarchy ($p = 0.003$). Similarly, Model 9 implies that women perceive relatively less fit as the perceived flatness increases ($p = 0.003$). In line with these results, female subjects wrote in the free-response question that “a flat hierarchy will run like a boys’ club” and “can be harder to fit in.” Taken together, these findings are consistent with our micro-mechanism that, compared to men, women perceive a lesser degree of fit in a flatter hierarchy.

Next, Models 10 to 12 exhibit the results for perceived workload (Workload). Model 10 indicates that subjects, on average, do not perceive significantly more or less workload in a flatter hierarchy relative to in a taller structure. However, Model 11 suggests that, compared to men, women perceive that they will have significantly more workload in a flatter hierarchy than in a taller one ($p < 0.001$). These results are robust in Model 12, which demonstrates that as the perceived flatness increases, female candidates perceive relatively more workload ($p < 0.001$). Consistent with these results, female subjects in the “Flatter” condition disproportionately raised concerns about an “overwhelming” or “heavy” workload. In contrast, female subjects in the “Taller” condition noted that “what comes to mind is better regulation of workloads” and that “the workload is spread out more fairly” and “very equally distributed.” One of these subjects went as far as to say that “companies that have more levels of management are able to even out the workload to make happier
employees.” In sum, these findings support our micro-mechanism that, compared to men, women perceive that they will be disproportionately burdened with higher workloads.

Lastly, Models 13 to 15 report the results for perceived competition (Competition). In Model 13, we find that subjects generally perceive less competition in a flatter hierarchy than in a taller one ($p < 0.001$). However, Model 14 indicates that women, relative to men, perceive significantly less competition in a flatter hierarchy ($p = 0.002$). This is consistent with Model 15, which shows that women see significantly less competition as perceived flatness increases ($p = 0.002$). In support of these results, a female subject stated in the free-response question that “a flat structure reduces competition.” In contrast, another female participant wrote that a taller hierarchy implies “more opportunities for advancement but more competition.” Hence, we do not find support for the micro-mechanism that, compared to men, women perceive more competition for career advancement opportunities in flatter hierarchies. On the contrary, we find that, relative to men, women prefer taller hierarchies despite the perception that these organizational structures are more competitive.

Overall, our survey experiment results imply that female candidates, compared to their male counterparts, are less attracted to flatter hierarchies not because they perceive more competition over career advancement opportunities, but because they perceive a greater likelihood of not fitting in these informal structures or of being saddled with disproportionately more work.

5.9 Results for the boundary condition and other plausible mechanisms

As a female founder/CEO may attenuate or even reverse our arguments for H1, we test this boundary condition using the three-way interactions between the organization’s hierarchy, the subject’s gender, and the founder/CEO’s gender. The results are presented in Table 10 in the Online Appendix A6. We find no evidence that the founder/CEO’s gender affects the heterogeneous attraction by gender. This finding implies that women in formal leadership may be insufficient to mitigate female job seekers’ concerns regarding the informal structure in flat organizations (Dezső et al. 2022).

We also explore other plausible mechanisms suggested by prior studies. These include perceptions of autonomy (Lee and Edmondson 2017, Raveendran et al. 2022), fairness (Eagly and Carli 2007:139, Schminke et al. 2000), employee voice (Lind et al. 1990, Morrison 2011), the founder/CEO’s competence (Alexy et al. 2021, Baron et al. 1999), and the likelihood of firm success (Lee 2022, Lee and Csaszar 2020). We discuss these mechanisms in detail and report their results in the Online
Appendix A7. These results show little support for these mechanisms.

6 Discussion

Flatter hierarchies have garnered much attention as “egalitarian” alternatives to taller, “authoritarian” hierarchies. Despite the potential benefits in motivating current employees, we argue that flatter hierarchies have a hidden cost of decreasing the diversity in the applicant pool because, relative to men, women may perceive that these informal structures will be difficult for them to fit in, disproportionately burden them with work, and offer them relatively fewer career advancement opportunities. Using a field experiment, we show that a flatter hierarchy does not have a significant effect on the size of the applicant pool but decreases its female representation. Our follow-up survey experiment indicates that this decrease corresponds to cross-gender differences in perceptions of fit, workload, and career opportunities. These findings have several implications for research on organizational design, entrepreneurship, workplace diversity, and human capital.

6.1 Theoretical contributions

First, our research contributes to the literature on organizational design by unveiling the role of organizational structure in attracting prospective employees. To date, this literature has primarily focused on how hierarchical structure motivates and coordinates existing employees (Burton and Obel 2004, Oldham and Hackman 1981, Puranam 2018), neglecting its effects on attracting new employees. Although a flatter hierarchy may benefit firms by motivating current employees and spurring their creativity (Keum and See 2017, Lee and Edmondson 2017, Raveendran et al. 2022, Reitzig 2022), we find that it has an unintended consequence of decreasing the diversity among prospective employees. Given that employee diversity is consequential to firm performance (Dezső and Ross 2012, Hoogendoorn et al. 2013, Østergaard et al. 2011), this finding implies that employee diversity may act as a distinct yet understudied mechanism that ties organizational structure to firm performance. More generally, it indicates the importance of accounting for applicant sorting in examining how organizational structure affects firm performance.

Second, we advance entrepreneurship research by uncovering an unintended, potentially long-lasting consequence of a flat hierarchy for startups. Past research has argued that new ventures should be “organic” with a flatter hierarchy to quickly and flexibly adapt to their turbulent
environment (Burns and Stalker 1961). Adding to the burgeoning stream of work reevaluating this received wisdom (e.g., Colombo and Grilli 2013, Foss and Klein 2022, Keum and See 2017, Lee 2022, Lee and Csaszar 2020, Reitzig 2022), our study reveals that a flatter hierarchy can be detrimental to startup performance by reducing the gender diversity in the applicant pool and thus exacerbating the problem of “diversity debt” (Wu 2017). This finding demonstrates how initial decisions that founders make regarding organizational structure may exert a lasting influence on the firm’s evolution and success (Alexy et al. 2021, Beckman and Burton 2008).

Third, our work adds to the literature on workplace diversity. This literature has discussed the importance of various firm characteristics in determining workplace diversity (e.g., organizational policies on hiring, compensation, and promotion; Abraham and Burbano 2022, Obloj and Zenger 2022, Tsolmon 2022, Wynn 2020). This study complements these studies by revisiting the role of formal hierarchy. Contrary to prior arguments that tall hierarchies are the root cause of gender inequality and segregation (Acker 1990, Ferguson 1984, Ianello 1992, Kalev 2009) and that “non-hierarchical, egalitarian organizations” offer a solution to this problem (Acker 1990:141), we find that, compared to taller ones, flatter hierarchies may exacerbate gender inequality and segregation by attracting fewer female applicants. More broadly, by dissecting how women and men differ in their perceptions of demand-side organizational characteristics (i.e., formal hierarchy) and thus self-select into the labor supply pool, our study answers the recent call to integrate supply- and demand-side perspectives on gender inequality and segregation (Barbulescu and Bidwell 2013, Fernandez-Mateo and Kaplan 2018).

Lastly, this study extends work on human capital by providing an empirical examination of the endogenous process through which individuals self-select into firms. Although extant studies have enumerated various firm characteristics that affect this selection process (e.g., leadership gender, firm size, organizational status, corporate social responsibility; Abraham and Burbano 2022, Bidwell et al. 2015, Burbano 2016, Campero and Kacperczyk 2020, Elfenbein et al. 2010), they have yet to examine the effects of formal hierarchy, which is “one of the defining features of formal organizations” (Sørensen and Sharkey 2014:329). Given the increasing trend of companies highlighting their flat hierarchies in their recruiting efforts, we provide the first study that assesses how perceptions of a firm’s hierarchy shape individuals’ self-selection into firms. We also explore mechanisms underlying this relationship by examining how women and men differ in their perceptions. By doing so, we
shed light on the selection process that determines founding team characteristics (Shah et al. 2019).

6.2 Limitations and future research

Like all research, this study has limitations. First, while our field study is more generalizable than lab-experiment studies, it involves a single firm in a specific industry. Moreover, this experiment considered two job positions, and its subjects were limited to U.S. job candidates with a bachelor’s degree or higher. Although this empirical setting does not seem atypical and the findings are consistent in our follow-up survey experiment on a broader population on MTurk, future work could enhance the generalizability of our findings by examining other empirical settings. Second, as we surveyed an extensive yet finite list of potential mechanisms, future studies could complement our findings by identifying other mechanisms that may explain why women and men react differently to this organizational structure. These studies may provide additional insights on how flatter organizations could mitigate female under-representation in their applicant pool. Third, motivated by the growing trend of companies promoting their flatter hierarchies in their job postings, we examined the vertical dimension of organizational structure (i.e., formal hierarchy). Future research could probe how perceptions of other structural dimensions (e.g., the horizontal division of tasks), work arrangements (e.g., remote work, office design), or human resources policies (e.g., job training) might shape the applicant pool. Lastly, because gender inequality and segregation continue to be crucial problems in today’s workplace, we focused on the gender composition of the applicant pool as our outcome of interest. Future work could further investigate how an organization’s formal hierarchy affects different attributes of its applicant pool (e.g., race, nationality, work experience).

6.3 Conclusion

In sum, our study unveils the hidden cost of flat hierarchies in attracting a diverse applicant pool. Yet, many important insights remain to be uncovered regarding the relationship between organizational structure and human capital acquisition. We hope our study serves as a foundation for future exploration in this area.
References


(a) An example of a job posting that highlights a flat organizational structure with fewer management layers. In this example, the names of the company, the CEO, and the direct manager are redacted.

(b) The growing trend of companies highlighting their flat hierarchies in their job postings. In this figure, the observations are 6,424,963 firms, which posted a total of 156,705,625 jobs in the U.S. between 2010 and 2019. This sample was collected from Burning Glass Technologies, which provides a representative dataset of U.S. job postings (Cammeraat and Squicciarini 2021, Lee and Kim 2022). To assess whether a firm’s job posting features its hierarchical structure, we used the following list of phrases: “chain of command,” “company hierarch*,” “corporate hierarch*,” “firm hierarch*,” “flat* hierarch*,” “flat* manag*,” “flat* organi*,” “flat* structur*,” “flatland,” “formal hierarch*,” “hierarchical organi*,” “hierarchical structur*,” “holacracy,” “layer* of manag*,” “level* of manag*,” “manag* layer*,” “manag* level*,” “no bosses,” “no managers,” “no middle manag*,” “non-hierarchical,” “org* hierarch*,” “org* structur*,” “self-manag*,” “self-organi*,” “tall* organi*,” and “tall* structur*.” In particular, job postings specifically mentioning a flat hierarchy are captured by the phrases containing the word “flat” (e.g., “flat hierarch*”) or their equivalent terms (e.g., “holacracy,” “no bosses,” “no middle manag*,” “non-hierarchical,” “self-manag*”; Foss and Klein 2022, Lee and Edmondson 2017, Lee 2022, Reitzig 2022). Here, the asterisk (*) represents a meta-character for one or more instances of any alphabet letters.

Figure 1: Featuring a flat hierarchical structure in the recruiting efforts.
Dear {Name of job candidate},

We found your profile on ZipRecruiter. Based on your credentials, we invite you to apply for a position in [business development/software engineering] here at {Partner company’s name}.

“Flatter” condition only: We think you’ll be interested to know that we have a flat organizational structure. This means that we have fewer levels of management than similarly sized startups in our industry.

We are a telemedicine company specializing in online therapy. We founded {Partner company’s name} in {Founding year}, and have since expanded to {Number of employees} full-time employees. We are looking to grow significantly.

To learn more about us and to apply for this position, please click this link to our job posting: {Hyperlink to the company’s application web page}.

Figure 2: Email invitation to job candidates in Study 1. The randomized manipulation in the second paragraph is only included for the treatment group (i.e., “Flatter” condition).

---

**Part-Time, Remote Copyeditor**

- We are seeking a part-time, remote copyeditor to help edit our reports and presentations.

**About Us**

- In 2015, we founded Exogeny to provide data-driven HR analytics.

**Job Responsibilities**

- Proofread reports and presentations we prepare for external audiences
- Provide feedback on the clarity of writing style and graphics

**Requirements**

- Exceptional attention to detail
- English fluency
- Ability to meet deadlines

Figure 3: Recruiting material in Study 2 with no mention of the organization’s hierarchical structure or the founder/CEO’s gender.
<table>
<thead>
<tr>
<th>Manipulation</th>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational structure</td>
<td>Flatter</td>
<td>“A defining feature is our flat organizational structure—we have fewer levels of management than similarly sized startups in our industry.”</td>
</tr>
<tr>
<td></td>
<td>Taller</td>
<td>“A defining feature is our tall organizational structure—we have more levels of management than similarly sized startups in our industry.”</td>
</tr>
<tr>
<td>Founder/CEO’s gender</td>
<td>Female</td>
<td>Jessica Chandler</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>Michael Chandler</td>
</tr>
</tbody>
</table>

Figure 4: The interventions in Study 2 inserted in the “About Us” section for the organization’s hierarchical structure or the founder/CEO’s gender.
## Table 1: Summary statistics in Study 1.

<table>
<thead>
<tr>
<th>All</th>
<th>By condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Information</td>
</tr>
<tr>
<td><strong>Click</strong></td>
<td>0.285</td>
</tr>
<tr>
<td><strong>Apply</strong></td>
<td>0.098</td>
</tr>
<tr>
<td><strong>FemaleSubject</strong></td>
<td>0.269</td>
</tr>
<tr>
<td><strong>Manager</strong></td>
<td>0.318</td>
</tr>
<tr>
<td><strong>Experience</strong></td>
<td>134.558</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>Master’s</td>
<td>0.358</td>
</tr>
<tr>
<td>Doctorate</td>
<td>0.017</td>
</tr>
<tr>
<td>Software</td>
<td>0.494</td>
</tr>
<tr>
<td><strong>No. observations</strong></td>
<td>8,171</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. All models are linear probability models.

## Table 2: Results using the linear probability model in Study 1.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Click</th>
<th>Apply</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimation:</strong></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td><strong>Sample:</strong></td>
<td>LPM (All)</td>
<td>LPM (All)</td>
</tr>
<tr>
<td><strong>Flatter</strong></td>
<td>0.003</td>
<td>0.010</td>
</tr>
<tr>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.019)</td>
</tr>
<tr>
<td><strong>FemaleSubject</strong></td>
<td>0.028</td>
<td>0.007</td>
</tr>
<tr>
<td>(0.016)</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td><strong>Flatter × FemaleSubject</strong></td>
<td>-0.050</td>
<td>-0.034</td>
</tr>
<tr>
<td>(0.022)</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td><strong>Software</strong></td>
<td>0.194</td>
<td>0.194</td>
</tr>
<tr>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.019)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>0.191</td>
<td>0.183</td>
</tr>
<tr>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.015)</td>
</tr>
<tr>
<td><strong>No. observations</strong></td>
<td>8,171</td>
<td>8,171</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. All models are linear probability models.
Table 3: Summary statistics in Study 2.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Attraction</th>
<th>Flatness</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>By condition</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flatter</td>
<td>Neutral</td>
</tr>
<tr>
<td>Attraction</td>
<td>3.652</td>
<td>3.782</td>
<td>3.661</td>
</tr>
<tr>
<td>FemaleSubject</td>
<td>0.584</td>
<td>0.576</td>
<td>0.590</td>
</tr>
<tr>
<td>Manager</td>
<td>0.247</td>
<td>0.236</td>
<td>0.248</td>
</tr>
<tr>
<td>Experience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-3 years</td>
<td>0.134</td>
<td>0.148</td>
<td>0.124</td>
</tr>
<tr>
<td>4-6 years</td>
<td>0.142</td>
<td>0.144</td>
<td>0.142</td>
</tr>
<tr>
<td>7-9 years</td>
<td>0.110</td>
<td>0.107</td>
<td>0.107</td>
</tr>
<tr>
<td>10+ years</td>
<td>0.554</td>
<td>0.544</td>
<td>0.568</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>0.386</td>
<td>0.395</td>
<td>0.390</td>
</tr>
<tr>
<td>Graduate</td>
<td>0.174</td>
<td>0.173</td>
<td>0.169</td>
</tr>
<tr>
<td>No. observations</td>
<td>8,498</td>
<td>2,884</td>
<td>2,797</td>
</tr>
</tbody>
</table>

Table 4: Main results in Study 2.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Attraction</th>
<th>Flatness</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1 OLS</td>
<td>Model 2 OLS</td>
<td>Model 3 2SLS: 1st stage</td>
</tr>
<tr>
<td>NoInformation</td>
<td>0.149</td>
<td>0.197</td>
<td>1.199</td>
</tr>
<tr>
<td>Flatter</td>
<td>0.275 (0.025)</td>
<td>0.349 (0.041)</td>
<td>2.500 (0.024)</td>
</tr>
<tr>
<td>FemaleSubject</td>
<td>0.300 (0.038)</td>
<td>0.380 (0.069)</td>
<td></td>
</tr>
<tr>
<td>NoInformation × FemaleSubject</td>
<td>-0.081 (0.051)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flatter × FemaleSubject</td>
<td>-0.123 (0.052)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flatness</td>
<td>3.509 (0.019)</td>
<td>3.333 (0.030)</td>
<td>1.885 (0.018)</td>
</tr>
<tr>
<td>Flatness × FemaleSubject</td>
<td>-0.048 (0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>8,498</td>
<td>8,498</td>
<td>8,498</td>
</tr>
<tr>
<td>No. observations</td>
<td>8,498</td>
<td>8,498</td>
<td>8,498</td>
</tr>
</tbody>
</table>

Note. Robust standard errors in parentheses. Flatter is a binary variable regarding whether the subject received the “Flatter” condition, whereas Flatness is the number of hierarchical levels perceived by the subject.

Table 4: Main results in Study 2.
Table 5: Mechanism tests in Study 2.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Opportunity</th>
<th>Informality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation:</td>
<td>Model 1 OLS</td>
<td>Model 2 OLS 2SLS: 2nd stage</td>
</tr>
<tr>
<td>NoInformation</td>
<td>-0.049 0.016</td>
<td>0.087 0.097</td>
</tr>
<tr>
<td>Flatter</td>
<td>-0.315 -0.213</td>
<td>0.118 0.113</td>
</tr>
<tr>
<td>FemaleSubject</td>
<td>0.050 0.179</td>
<td>0.055 0.037</td>
</tr>
<tr>
<td>NoInformation × FemaleSubject</td>
<td>-0.111</td>
<td>-0.018</td>
</tr>
<tr>
<td>Flatter × FemaleSubject</td>
<td>-0.177</td>
<td>0.011</td>
</tr>
<tr>
<td>Flatness</td>
<td>-0.086</td>
<td>0.044</td>
</tr>
<tr>
<td>Flatness × FemaleSubject</td>
<td>-0.072</td>
<td>0.005</td>
</tr>
<tr>
<td>Constant</td>
<td>3.712 3.68</td>
<td>2.783 2.751</td>
</tr>
</tbody>
</table>

Kleibergen-Paap rk Wald F statistic 1,622.6 1,622.6
No. observations 8,498 8,498 8,498 8,498 8,498 8,498

Note. Standard errors in parentheses. Flatter is a binary variable regarding whether the subject received the “Flatter” condition, whereas Flatness is the number of hierarchical levels perceived by the subject.
Online Appendices

A1 Bar graphs for Study 1

Figure 5: Bar graphs with 95% confidence intervals for each dependent variable in Study 1.
A2 Results using the logistic specification in Study 1

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Estimation:</th>
<th>Sample:</th>
<th>Click</th>
<th>Apply</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 4</td>
</tr>
<tr>
<td></td>
<td>Logit</td>
<td>Logit</td>
<td>Logit</td>
<td>Logit</td>
</tr>
<tr>
<td>Flatter</td>
<td>0.018</td>
<td>0.051</td>
<td>-0.210</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.058)</td>
<td>(0.098)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>FemaleSubject</td>
<td>0.146</td>
<td>0.088</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.119)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flatter × FemaleSubject</td>
<td>-0.261</td>
<td>0.180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software engineer</td>
<td>0.981</td>
<td>0.982</td>
<td>0.947</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.051)</td>
<td>(0.098)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.448</td>
<td>-1.488</td>
<td>-1.323</td>
<td>-1.496</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.052)</td>
<td>(0.086)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>No. observations</td>
<td>8,171</td>
<td>8,171</td>
<td>2,197</td>
<td>5,974</td>
</tr>
</tbody>
</table>

*Note.* Standard errors in parentheses. All models are logistic models.

Table 6: Results using the logistic specification in Study 1.
### A3 Demographic attributes of MTurk workers in Study 2

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>FemaleSubject</td>
<td>0.584</td>
</tr>
<tr>
<td>Manager</td>
<td>0.247</td>
</tr>
<tr>
<td><strong>Experience</strong></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0.060</td>
</tr>
<tr>
<td>1-3 years</td>
<td>0.134</td>
</tr>
<tr>
<td>4-6 years</td>
<td>0.142</td>
</tr>
<tr>
<td>7-9 years</td>
<td>0.110</td>
</tr>
<tr>
<td>10+ years</td>
<td>0.554</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>0.386</td>
</tr>
<tr>
<td>Graduate</td>
<td>0.174</td>
</tr>
<tr>
<td>Others</td>
<td>0.439</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>37.665</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.765</td>
</tr>
<tr>
<td>Black</td>
<td>0.102</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.072</td>
</tr>
<tr>
<td>Asian</td>
<td>0.086</td>
</tr>
<tr>
<td>Others</td>
<td>0.023</td>
</tr>
<tr>
<td><strong>Employment</strong></td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>0.592</td>
</tr>
<tr>
<td>Part-time</td>
<td>0.173</td>
</tr>
<tr>
<td>Unemployed</td>
<td>0.235</td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td></td>
</tr>
<tr>
<td>Less than $40K</td>
<td>0.251</td>
</tr>
<tr>
<td>$40-60K</td>
<td>0.215</td>
</tr>
<tr>
<td>$60-80K</td>
<td>0.146</td>
</tr>
<tr>
<td>$80-100K</td>
<td>0.175</td>
</tr>
<tr>
<td>More than $100</td>
<td>0.213</td>
</tr>
<tr>
<td><strong>Industry</strong></td>
<td></td>
</tr>
<tr>
<td>Healthcare</td>
<td>0.130</td>
</tr>
<tr>
<td>IT</td>
<td>0.101</td>
</tr>
<tr>
<td>Education</td>
<td>0.120</td>
</tr>
<tr>
<td>Retail</td>
<td>0.082</td>
</tr>
<tr>
<td>Others</td>
<td>0.567</td>
</tr>
<tr>
<td><strong>No. observations</strong></td>
<td>8,498</td>
</tr>
</tbody>
</table>

Table 7: Demographic attributes of MTurk workers in Study 2.
### A4 Measurement of variables in Study 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent</strong></td>
<td></td>
</tr>
<tr>
<td>Attraction</td>
<td>Average score of the five-point Likert scales for the following five items:</td>
</tr>
<tr>
<td></td>
<td>· “For me, Exogeny would be a good place to work.”</td>
</tr>
<tr>
<td></td>
<td>· “I would not be interested in Exogeny except as a last resort.” (reverse-coded)</td>
</tr>
<tr>
<td></td>
<td>· “Exogeny is attractive to me as a place of employment.”</td>
</tr>
<tr>
<td></td>
<td>· “I am interested in learning more about Exogeny.”</td>
</tr>
<tr>
<td></td>
<td>· “A job at Exogeny is very appealing to me.”</td>
</tr>
<tr>
<td><strong>Independent</strong></td>
<td></td>
</tr>
<tr>
<td>NoInformation</td>
<td>1 if no mention of the organization’s hierarchical structure; 0 otherwise.</td>
</tr>
<tr>
<td>Flatter</td>
<td>1 if flatter condition; 0 otherwise.</td>
</tr>
<tr>
<td>Flatness</td>
<td>The reverse-coded value for the perceived number of hierarchical levels measured using the question:</td>
</tr>
<tr>
<td></td>
<td>· “Based on the posting, how many levels of management would you estimate we have between our entry-level employees and our CEO?”</td>
</tr>
<tr>
<td><strong>Individual attributes</strong></td>
<td></td>
</tr>
<tr>
<td>FemaleSubject</td>
<td>1 if the subject self-identified as female; 0 otherwise.</td>
</tr>
<tr>
<td><strong>Boundary condition</strong></td>
<td></td>
</tr>
<tr>
<td>FemaleFounder</td>
<td>1 if the subject received the female founder/CEO condition; 0 otherwise.</td>
</tr>
</tbody>
</table>

Table 8: Measurement of primary variables in Study 2. Although the variable Flatness measures job seekers’ perception of how flat the organization’s hierarchy is, its name excludes the term “perceived,” for simplicity.
Table 9: Measurement of mechanisms in Study 2. Although all of the mechanisms are related to the perceptions of job seekers, their variable names exclude the term “perceived,” for simplicity (e.g., we use Opportunity instead of PerceivedOpportunity).
A5  Bar graphs and margins plots for Study 2

Figure 6: Bar graphs with 95% confidence intervals for each dependent variable in Study 2.
Figure 7: Margins plots with the 95% confidence intervals for each dependent variable in Study 2. In each plot, the x-axis specifies the three conditions for hierarchical structure ("Flatter," "No Information," and "Taller"). For each of these conditions, the y-axis shows the marginal effect of FemaleSubject on the respective dependent variable.
Results for the boundary condition of the founder/CEO’s gender in Study 2

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Estimation</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>2SLS: 2\textsuperscript{nd} stage</td>
</tr>
<tr>
<td>( NoInformation \times FemaleSubject \times FemaleFounder )</td>
<td>0.027 (0.105)</td>
<td></td>
</tr>
<tr>
<td>( Flatter \times FemaleSubject \times FemaleFounder )</td>
<td>0.024 (0.107)</td>
<td></td>
</tr>
<tr>
<td>( Flatness \times FemaleSubject \times FemaleFounder )</td>
<td>(-0.021 (0.088))</td>
<td></td>
</tr>
</tbody>
</table>

No. observations 8,498 8,498

*Note.* Standard errors in parentheses. For ease of reading, this table excludes all main effects and lower-level interactions. *Flatter* is a binary variable regarding whether the subject received the “Flatter” condition, whereas *Flatness* is the number of hierarchical levels perceived by the subject.

Table 10: Results for the boundary condition of the founder/CEO’s gender in Study 2.
A7 Results of other plausible mechanisms in Study 2

Perceptions of autonomy: As discussed in Section 2.3, a flatter hierarchy may offer more autonomy (Lee and Edmondson 2017, Lee 2022, Raveendran et al. 2022), the perception of which would, in theory, disproportionately attract more female applicants. Although we find no evidence of this increased female attraction in Study 1, we used the measure of perceived autonomy and examined how autonomy could be perceived differently by gender. The results of these analyses are presented in Models 1 and 2 of Table 11. Consistent with the prior studies, we find that subjects, in general, perceive more autonomy in a flatter structure than in a taller hierarchy ($p < 0.001$). Accordingly, one survey participant stated in the free-response question that, in a flatter structure, “employees have more autonomy in determining work projects and have more decision-making power in the company.” In turn, another subject commented that “tall organizations leave less room for autonomy in the workplace and have levels of redundant leadership that get in the way.” Surprisingly, however, the results indicate that, relative to men, women see substantially less autonomy in a flatter hierarchy than in a taller structure ($p < 0.001$), and that women perceive less autonomy as the perceived flatness increases ($p < 0.001$). In line with these results, while men unanimously showed a negative perception of autonomy in a taller hierarchy in their free-response question, several women shared a more positive view. For instance, one female participant noted that a taller structure offers “more opportunities to be in a supervisory role as well as more autonomy among employees.” Another female subject stated, “I think of [a taller hierarchy as] a semi-autonomous work situation, where employees are allowed to map out their workload allocations.” In sum, these findings counter the argument that, compared to men, women will be more attracted to a flatter hierarchy because they perceive them to offer more autonomy.

Perceptions of fairness and employee voice: Although prior studies have argued that flatter hierarchies would encourage employees to voice their opinions, which can lead to perceptions of fairness (Lind et al. 1990), our qualitative observations suggest that women may, in fact, perceive a flatter hierarchy to hinder employee voice behaviors and have unfair procedures.\footnote{Although (in)formality and fairness may be theoretically intertwined (Lind and Tyler 1988), we distinguish these two constructs as Schminke et al. (2000) show that the former is not related to perceptions of the latter.} If so, the gender difference in attraction could be driven by the perception of fairness or the perceived opportunity to
have one’s voice heard. Hence, we used our measures of perceived fairness and employee voice and estimated Models 3 to 6 in Table 11. However, we find that the effects of hierarchy on perceived fairness and employee voice do not differ by gender.

Perceptions of firm quality: As flatter organizations are more likely to commercially fail (Lee 2022, Lee and Csaszar 2020), prospective employees may have the perception that these firms have a high risk of failure. Moreover, as experienced entrepreneurs tend to draw upon their prior employers to design a more hierarchical structure in their firms (Alexy et al. 2021, Baron et al. 1999), founders of flatter organizations may be perceived to be less experienced and competent. As women tend to be more risk-averse than men (Croson and Gneezy 2009), they may thus be less attracted to flatter organizations. We address these alternative mechanisms by measuring the perceived competency of the founder and the perceived probability of firm success. These explanations are tested in Models 7 to 10 of Table 11. These results demonstrate that women do not differ significantly from men in their perception of the founder/CEO’s competence of flatter organizations or the probability that flatter organizations will succeed.
<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Autonomy</th>
<th>Fairness</th>
<th>Voice</th>
<th>FounderCompetence</th>
<th>FirmSuccess</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 4</td>
<td>Model 5</td>
</tr>
<tr>
<td>OLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2SLS: 2nd stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NoInformation</td>
<td>0.517</td>
<td>0.059</td>
<td>0.051</td>
<td>0.036</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.031)</td>
<td>(0.036)</td>
<td>(0.036)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>Flat</td>
<td>1.053</td>
<td>0.015</td>
<td>0.098</td>
<td>0.167</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.031)</td>
<td>(0.036)</td>
<td>(0.035)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>FemaleSubject</td>
<td>0.149</td>
<td>0.014</td>
<td>0.031</td>
<td>0.136</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.029)</td>
<td>(0.053)</td>
<td>(0.033)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>NoInformation × FemaleSubject</td>
<td>−0.117</td>
<td>−0.023</td>
<td>−0.003</td>
<td>−0.015</td>
<td>−0.025</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.040)</td>
<td>(0.046)</td>
<td>(0.045)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Flat × FemaleSubject</td>
<td>−0.174</td>
<td>−0.026</td>
<td>−0.045</td>
<td>−0.051</td>
<td>−0.062</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.041)</td>
<td>(0.046)</td>
<td>(0.045)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Flatness</td>
<td>0.419</td>
<td>0.005</td>
<td>0.039</td>
<td>0.067</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Flatness × FemaleSubject</td>
<td>−0.065</td>
<td>−0.011</td>
<td>−0.018</td>
<td>−0.020</td>
<td>−0.025</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.016)</td>
<td>(0.018)</td>
<td>(0.019)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.815</td>
<td>3.500</td>
<td>3.369</td>
<td>3.733</td>
<td>3.525</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.022)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Kleibergen-Paap- rk Wald F statistic</td>
<td>1622.6</td>
<td>1622.6</td>
<td>1622.6</td>
<td>1622.6</td>
<td>1622.6</td>
</tr>
<tr>
<td>No. observations</td>
<td>8,498</td>
<td>8,498</td>
<td>8,498</td>
<td>8,498</td>
<td>8,498</td>
</tr>
</tbody>
</table>

*Note.* Robust standard errors in parentheses. Flatter is a binary variable regarding whether the subject received the “Flatter” condition, whereas Flatness is the number of hierarchical levels perceived by the subject.

Table 11: Results for other plausible mechanisms in Study 2: (1) autonomy, (2) fairness and employee voice, and (3) firm quality.