Who Gets the Job Offer?

Manager Gender and Job Skill Requirements*

Liat Eldor Michal Hodor Abigail Hurwitz

Abstract

This study examines gender bias at the job-offer stage, focusing on how the gender of hiring managers shapes recruitment decisions across a wide range of occupations with varying skill requirements. We develop a conceptual model of job-offer selection that allows managerial preferences and beliefs to vary by applicant gender and position skill level, and use it to characterize potential sources of bias. We estimate the model using proprietary data on over 30,000 job applications submitted between 2015 and 2018 of a large retail chain. Our identification strategy leverages the quasi-random assignment of applicants to male and female managers and the unique institutional setting where managers and applicants never interact directly throughout the selection process. We find no evidence of gender bias among female managers. In contrast, male managers are 12 percentage points more likely to extend offers to male applicants than female applicants in high-skill positions, while displaying no significant bias in low-skill roles. These results identify the gender of the hiring evaluator as a critical driver of the persistent under-representation of women in high-skill occupations. Furthermore, we present evidence that these findings are consistent with statistical discrimination driven by gender-skill stereotypes, as the hiring penalty for women is concentrated in regions with more traditional cultural norms.

*We thank conference participants at Queen's University for helpful suggestions.

1 Introduction

Persistent gender disparities characterize labor markets worldwide. Women participate in the labor force at substantially lower rates than men, 47% globally compared to 72% for men (Sayeh et al., 2023), and are markedly underrepresented in high-skill occupations, comprising only 28.2% of the STEM workforce (Friedmann and Efrat-Treister, 2023). These gaps emerge at the initial margin of labor market entry. Research indicates that 15% of women, compared to 10.5% of men, encounter recruitment biases and occupational segregation at this stage, making the hiring decision a key margin at which gender inequality arises and persists (International Labour Organization, 2023).[1]

Despite this evidence, little is known about the mechanisms through which gender disparities aggregate into disparate job-offer outcomes. Prior studies have documented gender bias in early employment stages, but remain confined to idiosyncratic contexts or specific occupations. For example, Goldin and Rouse (2000) show that blind auditions in orchestras significantly improve female musicians' hiring prospects. While this seminal work provides a proof-of-concept for gender-blind screening, the results from this unique setting may not generalize to representative labor markets. More recent evidence from online gig platforms further documents gender differences in employer contact rates, reflecting occupational sorting and gender-typed role preferences(Chan and Wang, 2018; Hangartner et al., 2021). Complementary audit and correspondence experiments, which vary candidate's gender across otherwise identical fictitious applications, find that evaluators often rate female candidates less favorably at the screening stage, although the magnitude and direction of this effect depend on the context and the type of job (Ashnai et al., 2020; Campero and Fernandez, 2019; Friedmann and Efrat-Treister, 2023; Ruffle and Shtudiner, 2015).

Empirically identifying the sources of gender bias at the job-offer stage remains a central challenge, as offer outcomes typically depend on subjective assessments of candidate ability and interpersonal interactions, both of which are endogenous to the decision. Crucially, the gender of the decision-maker is likely pivotal in shaping these interactions, both independently and relative to

[1]The 25-percentage-point gap in global labor force participation (LFP) reflects both voluntary non-participation and involuntary exclusion. To isolate the involuntary component, International Labour Organization (2023) utilizes the "Jobs Gap" metric, which captures individuals who are available for work but face binding structural constraints. The fact that this unmet desire to work is significantly higher for women than for men suggests that the observed LFP gap is driven by systemic frictions rather than differential preferences for home production.

candidate's gender. Yet, prior research has abstracted from the gender of the decision-maker, reflecting both data constraints and the inherent difficulty of isolating manager-specific effects. This omission is notable given evidence of systematic gender differences in managerial preferences (see, e.g., Blais and Weber, 2001; Borghans et al., 2009; Byrnes et al., 1999; Sapienza et al., 2009). This paper overcomes these identification challenges by leveraging administrative data from a national retail chain to estimate the first causal effect of manager gender on the probability of receiving a job offer across a broad range of occupations.

The empirical strategy exploits a unique institutional feature: the complete absence of personal interaction between hiring managers and applicants throughout the selection process, from initial application to the final decision. The retail chain recruits for a wide range of occupations, spanning high-skill positions such as optometrists and pharmacists, as well as lower-skill logistics and service roles. The data cover the universe of applicants between 2015 and 2018 and record information on job offer outcomes. Importantly, the dataset includes the gender of both applicants and managers, allowing us to identify gender-matched effects in a high-stakes, real-world setting.

Job applications are submitted through the firm's centralized online platform and routed to human resources, which assigns applicants to stores based on staffing needs. Following this, suitable applicants undergo a set of standardized assessment tests at an external evaluation center, covering both cognitive and non-cognitive skills. The center produces a summary report that serves as an objective proxy for candidate's personality and ability level. Subsequently, store managers make job-offer decisions solely based on this report. This process eliminates interpersonal interactions and visual cues, thereby overcoming a major source of endogeneity in the literature on hiring bias.

To formalize potential biases in the job-offer decision, we develop a conceptual framework in which managers evaluate candidates based on ability, gender, and the preferences and beliefs of the decision-maker. Within this framework, we construct the ability-adjusted average gender gap in job-offer probabilities across positions. The model illustrates how this gap can arise through (i) disparate treatment, where the manager's decision rule depends explicitly on the candidate's gender, and (ii) disparate impact, where outcomes differ because gender is correlated with other characteristics or because candidates are sorted to managers with systematically different hiring

tendencies.

We map this framework to our data to estimate empirically the ability-adjusted average gender gap. Our identification strategy exploits the firm's centralized hiring process, which precludes subjective impressions and interpersonal interactions ensuring that an applicant's gender is effectively independent of the manager's gender. Furthermore, the quasi-random assignment of applicants to stores ensures that candidate's gender is independent of other managerial characteristics, thereby neutralizing the disparate impact driven by non-random sorting. This unique setting enables us to address two core questions: first, conditional on measured ability and personality traits, are equally qualified male and female applicants treated equivalently? and second, does this treatment vary systematically with the gender of the decision-maker and the occupational skill level?

In support our identification strategy, we show that applicants are allocated equally to male and female managers. We further conduct balance tests showing that applicants' abilities and personality scores are uncorrelated with manager gender across applicant genders and skill levels. These results support the assumption that manager assignment is quasi-random and orthogonal to applicant characteristics.

Our main findings reveal a significant systematic gender bias against female applicants. On average, equally qualified male applicants are 3 percentage points more likely to receive a job offer than their female counterparts. When decomposed by skill level, the bias is concentrated in high-skill occupations, where male applicants are 8.6 percentage points more likely to receive a job-offer than comparable female applicants. In lower-skill occupations, the bias persists but is substantially attenuated, at 1.9 percentage points. Disaggregating the results by decision-maker gender shows that in high-skill occupations, the bias is driven by male managers, who are 12 percentage points more likely to extend job offers to men, whereas female decision makers exhibit no measurable bias. A similar pattern emerges in lower-skill positions, where the male job-offer advantage is again driven entirely by decisions made by male managers.

We further exploit heterogeneity within lower-skill occupations to examine gender differences across distinct occupational categories: pink-collar roles (e.g., sales and service associates), blue-collar roles (e.g., inventory associates), white-collar support roles (e.g., pharmacist assistants),

and middle-managerial positions. We find that male managers' preference for male applicants is concentrated in white-collar support positions. In blue-collar roles, however, male managers favor female applicants. Female managers, by contrast, exhibit a preference for female applicants in middle-managerial roles while remaining neutral in other occupation categories.

The results indicate that male managers' job-offer decisions are consistent with statistical discrimination driven by gender-skill stereotypes. Male managers perceive men as better suited for high-skill positions, systematically favoring male applicants even when they possess identical measured traits and positions are gender-balanced. Female managers do not exhibit consistent gender bias in their job-offer outcomes. However, their higher likelihood of extending job offers to female middle managers may reflect positive in-group homophily, potentially facilitating female representation within managerial tracks. Notably, the heterogeneity across lower-skill occupations suggests that gender bias is context-dependent, varying with occupational tasks rather than operating as a uniform penalty against female applicants.

Our paper contributes to the growing body of research on persistent gender disparities in the labor market in several ways. First, we identify the decision maker's gender as critical determinant of gender bias at the job offer stage, providing robust evidence that male managers favor male applicants in high-skill positions. In doing so, we establish a clear causal link between applicant gender, manager gender, and job-offer outcomes. These findings highlight a foundational mechanism underlying the under-representation of women in high-skill occupations, whereby male decision-makers disproportionately favor men, leading to disparities that compound over time.

Second, our setting provides a unique opportunity to overcome endogeneity concerns and long-standing challenges in identifying gender bias in job offer decisions. While the literature has traditionally addressed these challenges using audit and correspondence studies, these methods typically capture stylized or partial responses in restricted information environments. Reflecting these limitations, recent meta-analyses of correspondence studies report inconclusive effects, with discrimination varying significantly by context (Lippens et al., 2023; Park and Oh, 2025). Similarly, laboratory studies show that the direction and magnitude of gender bias depend on occupation and evaluation context (Booth and Leigh, 2010; Carlsson, 2011; Moss-Racusin et al., 2012; Reuben

et al., 2014). By contrast, our setting allows us to analyze actual job offer decisions made by professional managers in a high-stakes business environment, leveraging rich information on both applicants and evaluators. This advances the literature in three distinct ways. First, it provides a clear identification of gender bias at the job-offer stage, offering clear evidence where previous field and lab results have been inconsistent. Second, it enables us to disentangle the sources of this bias by isolating the gender of the decision-maker. Finally, we assess how such bias varies systematically across a broad spectrum of occupations and skill levels. By providing precise estimates across these dimensions, we show that the mixed results in the literature are largely explained by the interaction between evaluator gender and occupational complexity.

Third, our findings demonstrate that gender disparities in the labor-market arise not only from structural and cultural barriers experienced on the job, but originate at the initial point of entry: the job-offer decision. Existing evidence highlights that structural barriers impede women's career progression, resulting in limited representation in leadership positions (Eagly and Karau, 2002; Eckel et al., 2021). In academic settings, research shows that women face unequal professional treatment, including less favorable reference letters (Eberhardt et al., 2023), more frequent and aggressive interruptions in seminars (Dupas et al., 2021), and harsher evaluations of research proposals (Wu, 2020), factors that contribute to their persistent underrepresentation in the labor market(Dupas et al., 2021). By showing that biased decisions emerge at the threshold of the job offer, our study identifies a "leaky pipeline" at its source. This suggests a cumulative disadvantage that not only impacts women's initial placement but also restricts access to high-skill tracks and reinforces gender inequality over time.

Lastly, our study contributes to the literature on gender disparities by identifying a behavioral bottleneck in what has traditionally been considered a model occupation for labor equality. While Goldin (2014) identifies the pharmacy profession as a benchmark for temporal linearity, where the absence of a part-time penalty and high worker substitutability should theoretically preclude gender gaps, our findings show otherwise. Leveraging a high-skill dataset primarily composed of pharmacists and optometrists, we find that structural linearity alone does not eliminate gender bias at the job offer stage. Rather, we reveal that managerial discretion introduces bias into

the decision process, demonstrating that male managers exhibit a preference for male applicants despite identical qualifications. In doing so, we demonstrate that even when job design eliminates the motherhood penalty, gender-based frictions in job-offer outcomes persist. This suggests that the labor-force participation gap is not merely a reflection of women's choice for flexibility, but is compounded by a demand-side constraint where bias acts as an involuntary barrier to entry.

The remainder of the paper proceeds as follows: We introduce our theoretical framework in Section 2. In Section 3 we describe our data and settings followed by a discussion of our empirical design in Section 4. Section 5 presents our main findings on gender bias in job-offering and Section 6 presents supplemental analysis. In Section 7, we conclude and discuss policy implications.

2 Theoretical Framework

We develop a conceptual framework that models job-offering decisions, allowing for managers' heterogeneous preferences and beliefs about applicant quality.

Conceptual framework of the gender gap in job offers. Each applicant i for position p is characterized by ability A_i and gender F_i, where $F_i \in \{0, 1\}$ denotes male and female, respectively. Applicant i is assigned to store $j \in \{1, ..., J\}$ via the mapping $J(i)$, where $Z_{ij} = \mathbb{I}\{J(i) = j\}$ denotes store assignment.

Within store j, the job offer decision for position p is made by the store manager with the objective of assessing applicant quality based on observable application information, according to the decision rule:

$$H_j^p(a, f, u),$$

where a denotes applicant ability, f gender, and u represents store-specific contextual factors, such as local labor market conditions or customer demographics.

For applicant i, the realized job offer outcome is the potential outcome associated with store assignment $J(i)$:

$$H_i^p \equiv \sum_{j=1}^{\mathcal{J}} Z_{ij} H_j^p(A_i, F_i, U_{ij}) = H_{J(i)}^p(A_i, F_i, U_{iJ(i)}).$$

The firm-level hiring propensity for applicants of type (a, f) is:

$$\bar{H}^p(a, f) \equiv \mathbb{E}\left[\sum_{j=1}^{\mathcal{J}} Z_{ij} H_j^p(a, f, U_{ij}) \,\middle|\, A_i = a,\ F_i = f\right],$$

which averages job offer decisions over both the stochastic assignment of applicants to stores and store-specific heterogeneity in offers decision rules, including manager-specific factors and contextual shocks captured by U_{ij}.

Ability-adjusted gender gap in job offers. The primary parameter of interest is the ability-adjusted average gender gap in job offers,

$$\beta = \int \omega(a) \left(\bar{H}^p(a, 0) - \bar{H}^p(a, 1)\right) \, dG(a),$$

where $G(a)$ denotes the distribution of applicant ability, the weighting function $\omega(a)$ aggregates conditional gaps across the ability distribution, and the firm-level hiring propensity is defined as:

$$\bar{H}^p(a, f) = \mathbb{E}[H_i^p \mid A_i = a, F_i = f],$$

where H_i^p denotes the realized job offer outcome after assignment to a store. Consequently, β corresponds to a weighted average of conditional gender gaps in realized outcomes across the ability distribution. If the gender gap is constant across ability levels, i.e., if $\mathbb{E}[H_i^p \mid F_i = 0, A_i = a] - \mathbb{E}[H_i^p \mid F_i = 1, A_i = a]$ is constant for all a, then β β is invariant to the choice of $\omega(a)$.

Potential sources of gender gaps. Suppose each applicant is characterized by a latent quality, $q_i \in Q \subseteq \mathbb{R}$, not directly observed by the manager. Instead, the manager observes a set of signals, including ability a, gender f, contextual factors u, and unobservables ε capturing idiosyncratic biases, interpersonal effects, or other manager- or applicant-specific factors. These unobservables may reflect documented sources of bias, including differential treatment by age or ethnicity, reliance on heuristics under time pressure, or situational shocks affecting attention or decision quality (e.g., Agan and Starr, 2018; Bertrand and Mullainathan, 2004; Sherman et al., 2025; Veit et al., 2022).

A manager of gender $g \in \{m, f\}$ maximizes expected utility derived from the decision $d \in \{0, 1\}$:

$$v_j^g(d, q, f, \varepsilon).$$

. Manager gender may shift the parameters governing preferences or beliefs without restricting the functional form of the utility function itself. Because quality q_i is unobserved, managers form subjective beliefs conditional on observed signals. Let $\tilde{F}_j^g(q \mid a, f, u)$ denote manager j's subjective posterior distribution of applicant quality after observing a, f, and u. Differences across managers in $\tilde{F}_j^g(\cdot)$ may reflect heterogeneous beliefs, stereotypes, or information-processing constraints.

The manager's job offer decisions are made to maximize expected utility:

$$H_j^p(a, f, u) = \arg \max_{d \in \{0,1\}} \int v_j^g(d, q, f, \varepsilon) d\tilde{F}_j^g(q \mid a, f, u).$$

This formulation accommodates departures from subjective beliefs $\tilde{F}_j^g(\cdot)$ and the true quality distribution of applicant quality, capturing biased or inaccurate inference.

Disparate treatment. Disparate treatment arises when hiring decisions depend directly on applicant gender, holding fixed all other observed characteristics and the store to which the applicant is assigned. Formally, there exist values of a and u such that

$$H_j(a, 1, u) \neq H_j(a, 0, u),$$

meaning that two applicants who are identical in all observable characteristics except gender may receive different job offer outcomes.

Such differences may reflect taste-based discrimination, in which gender enters the manager's utility function directly (Becker, 1957), or statistical discrimination, whereby managers hold gender-specific beliefs about the distribution of latent quality conditional on observed signals (Aigner and Cain, 1977). More generally, disparate treatment may arise from distorted, stereotyped, or biased beliefs about applicant quality, even when gender is not intrinsically payoff-relevant (Bordalo et al., 2016).

Disparate impact. Disparate impact arises when job offer decisions depend on characteristics correlated with applicant gender, even if the decision rule H_j is invariant to f. In this case, gender gaps emerge indirectly through the interaction of hiring criteria, contextual factors, and institutional features of the hiring process.

For instance, if store-specific factors U_{ij}, such as local labor market conditions or customer demographics, affect hiring decisions and are unevenly distributed across applicants by gender, gaps may arise indirectly. Similarly, heterogeneity in store-level hiring thresholds or managerial tendencies, combined with systematic differences in the assignment of applicants across stores through Z_{ij}, can generate firm-level gender gaps in the absence of direct disparate treatment.

3 Data and Settings

3.1 Data

We exploit comprehensive individual-level data on job applicants to a national retail chain comprising 265 stores from 2015 to 2018, covering over 32,000 applications.[2] Positions in the firm include low-skill roles such as cashier, sales associate, administrative positions, service roles, shift manager, and assistant pharmacists, as well as high-skill roles such as pharmacist and optometrist. We classify applicants into low- and high-skill categories based on the educational requirements associated with each position.[3] The dataset contains detailed information on job-offer outcomes, specifically whether an applicant received a job offer, as well as demographic characteristics, including age and gender.

A unique feature of the data is the inclusion of extensive assessment scores, which serve as the primary source of information managers use to inform their job offer decisions. These assessments measure a range of personality traits, including emotional stability, warmth and sociability, conscientiousness, energy, persistence, and social desirability, and are administered to applicants across

[2] The data were obtained from a Western country, but the identity of the firm is confidential.

[3] Although the titles "shift manager" and "assistant pharmacist" may suggest managerial or high-skill responsibilities, these positions do not require post-secondary education and are classified as low-skill roles. An assistant pharmacist primarily performs sales-related tasks within the pharmacy, and a shift manager is typically a cashier with additional responsibilities for shift scheduling under the close supervision of the store manager. The classification into high- and low-skill positions reflects our categorization rather than the firm's internal definitions.

all occupations.[4] The data also include measures of analytical and cognitive abilities, including numeracy, matrix reasoning, feature recognition speed, general processing speed, mathematical ability, and comprehension of instructions, which vary by occupation. Each test score is standardized on a 1 to 9 scale within each occupation, and a summary report of the scores, along with a brief psychologist assessment, is provided to managers. This report provides objective proxies for applicant personality traits and abilities, facilitating standardized cross-applicant comparisons. We use these measures to capture the multidimensional vector of applicant ability A_i observed by the manager during the hiring process.

3.2 Job application and offer process

The hiring process operates as follows. Applicants first view job postings on the retail chain's website, which provides detailed descriptions of available positions by city or region without disclosing the specific assigned store. The human resources (HR) department, located at the headquarters of the retail chain, manages the entire recruiting process until the final offer decision, which is made exclusively at the discretion of the manager of the assigned store. Centralizing the recruitment process is deliberately designed to reduce the administrative load for store managers and to ensure a standardized, systematic hiring process, as turnover and burnout rates are high in retail environments Ton and Huckman (2008).

Through the retail chain website, applicants submit personal information, including name, gender, and education. Applicants are assigned to stores' open positions based on the order in which stores flag staffing needs and vacancies. For each position, an HR associate identifies suitable candidates after an initial screening process and invites the selected candidates to a short phone interview. Applicants who pass this stage undergo a one-day evaluation at an external psychometric assessment center, where they complete a set of assessment tests.

The set of assessment tests conducted by each applicant is tailored to the requirements of each position. Our data include only applicants who were tested. In the final stage of the screening process, store managers receive a form summarizing the standardized assessment scores along with

[4] The personality measures correspond to standard Big Five personality dimensions (McCrae and Costa, 1987; McCrae and John, 1992).

a brief written evaluation by a psychologist regarding the applicant's suitability. The summary form also includes the applicant's name, age, and gender. The store manager decides whether to make a job offer based solely on this form. An example of such a form is presented in Appendix Figure B3. Crucially, applicants have no direct interaction with the store manager during the hiring process.

TABLE 1

Panel A of Table 1 reports summary statistics for the sample of 32,087 applicants. Approximately 72.8% of applicants are female, comprising 73.4% of applicants in low-skill occupations and 57% in high-skill occupations. Consistent with the post-secondary education requirements of high-skill roles, applicants to these positions are significantly older.[5] Moreover, the job offer likelihood for high-skill positions is 61%, compared to 33% for low-skill applicants. These patterns likely reflect documented labor shortages for pharmacists and optometrists, reported in the OECD Health Statistics for 2015-2018 (OECD, 2018).

The table also reports standardized personality and ability scores. While personality tests are uniform for all applicants, the ability score is an average of role-specific tests. For any given position, the set of ability tests is identical for all candidates. Panel B reports information on store managers, showing a relatively balanced gender distribution with 57% female managers. On average, managers have 6.6 years of tenure in their current store and 17.7 years of tenure within the retail chain, indicating that managers move between stores over the course of their careers.

4 Research Design and Empirical Model

4.1 Identification strategy

The hiring process described above provides a unique opportunity to test for gender bias in job offer decisions. In the next section, we draw on the conceptual framework developed in Section 2

[5] Applicants younger than 18 are excluded from the sample. These account for 1,108 applicants (3.5% of the full applicant pool).

to discuss the potential sources of such bias, their implications for identification, and how our empirical setting addresses these concerns.

First, one potential source of bias could arise from the assignment of applicants to stores, denoted by Z_{ij}. In our setting, applicants do not apply to specific stores, and they do not know the rule governing store assignment, as they observe only the geographic region or city of the posting (e.g., "north district of CITY" or "CITY"). At the same time, store managers do not select which applicants to evaluate, meaning they have no influence on Z_{ij}, which is entirely administered by the HR department. In practice, Z_{ij} is determined by the temporal sequence of store-level vacancies, which are orthogonal to applicant characteristics A_i and F_i, rendering manager assignment exogenous. Therefore, the centralized mechanism ensures independence from the discretion of HR associates, store managers, and applicants. In empirical terms, this means that the process of assigning an applicant to a specific manager is effectively as good as random with respect to the applicant's characteristics, eliminating selection bias stemming from applicant self-selection, manager-driven screening, or geographic sorting.

Second, the setting mitigates manager-specific situational factors such as time pressure, attention interruptions, multitasking, or fatigue during application evaluation. In our conceptual framework, these factors are captured by store- and manager-specific contextual variables U_{ij}, which can shift hiring outcomes by affecting how observed information is processed.[6] In the context of our setting, the assessment summary mitigates such situational shocks by providing standardized measures of applicant ability and personality, which facilitate comparable evaluations and reduce the influence of unstructured discretion. By reducing reliance on unstructured discretion and informal screening practices, this institutional feature limits the scope for situational constraints to systematically affect hiring decisions.

Finally, the centralized screening design eliminates a key source of bias in traditional job screening, namely subjective impressions and interpersonal interactions. Job-offer decisions may be influenced by rapport, physical appearance, accent, vocal tone, or other subjective cues, which are

[6]That is, for a fixed set of applicant characteristics, variation in the evaluation environment can lead to different hiring decisions:

$$H_j^p(a, f, u) \neq H_j^p(a, f, u').$$

captured in our framework by the unobservable component ε (Barrick et al., 2012; Purkiss et al., 2006; Ruffle and Shtudiner, 2015). Here, the screening process is fully blinded, as store managers and applicants never interact. All decisions are based exclusively on standardized assessment results for both personality traits and ability, applied uniformly across candidates. This design removes interpersonal and appearance-based components from ε, narrowing the set of unobserved factors to those that are orthogonal to applicant gender and strengthening the identification of gender-related bias.

The unique features of our data allow us to rule out applicant sorting and interpersonal interactions as sources of bias in job offer decisions. Managers do not select certain types of applicants (e.g., higher ability or a specific gender), and applicants are not steered toward specific stores based on those traits or the manager's own traits. Moreover, the design enables us to control for personality and ability measures, as well as age, which are key determinants of managerial evaluations (Carlsson and Eriksson, 2019; Neumark, 2024). After accounting for these factors, applicant gender remains the primary observable characteristic through which differential treatment is assessed. This allows us to claim that any variation in job-offer decisions is due to how managers respond to applicant gender, rather than the underlying qualities of the applicants. Accordingly, this setting enables a direct test of whether, conditional on age and qualifications, managers' decisions systematically favor one gender over the other, and whether the gender of the decision maker shapes this pattern.

4.2 Research design validation

To test the validity of the identification assumption, we first examine the patterns of applicant assignment to stores by store manager gender. The data indicate that applicants are not systematically assigned to managers of a specific gender. Specifically, the distribution of applicants is well-balanced across manager types: male applicants are assigned nearly equally to stores managed by women (28%) and men (27%), while female applicants are similarly distributed across male-managed (73%) and female-managed (72%) stores.

Moreover, Table 2 presents a formal balance test of whether applicant ability and personality

assessment scores vary by manager gender. The estimates confirm that neither measure is associated with the manager's gender. If applicants were assigned to managers non-randomly, one would expect higher-ability applicants to be disproportionately matched with a specific manager gender. The balance test alleviates this concern. The results reported separately by skill level in Panels (a) and (b), strongly support the identification assumption. They indicate that manager gender is orthogonal to applicant ability and personality scores for both skill levels, consistent with an effectively random assignment process.

TABLE 2

FIGURE 1

As a final step in our descriptive analysis, we examine the distributions of applicant age, ability, and personality measures by job-offer outcome, as shown in Figure 1. This analysis assesses how job-offer outcomes vary with the observable characteristics available to decision makers. Panels (a) and (b) plot the distributions of age at test for high- and low-skill occupations, respectively, by job offer decisions. For low-skill roles, managers show a preference for younger applicants. These patterns align with prior evidence on age-related hiring (Neumark, 2024; Neumark et al., 2016), which show that age is a significant factor in retail hiring. Specifically, younger applicants for low-skill roles such as sales associates and cashiers receive more favorable outcomes. This pattern is often attributed to statistical or taste-based discrimination, regarding customer fit, lack of flexibility, or technological orientation. However, this age preference disappears in high-skill occupations, where the age distributions for successful and unsuccessful applicants are nearly identical.

Panels (c) through (f) of Figure 1 present the corresponding distributions for ability and personality scores. For low-skill occupations, applicants with higher scores in both dimensions are significantly more likely to receive a job offer. In contrast, for high-skill occupations, the differences are less pronounced. While personality scores continue to play a role in job offer decisions, the ability measure distributions for high-skill roles are effectively indistinguishable across offer outcomes. One possible explanation is that formal education and professional credentials already provide substantial information about applicant quality in high-skill roles. Consequently, decision

makers may place greater weight on these external tests. This is consistent with research showing that richer information environments reduce the reliance of decision makers on specific direct measures or group-based priors (Altonji and Pierret, 2001).

4.3 Empirical model

Our data allow us to compare applicants who are observationally equivalent across all relevant dimensions except gender, enabling an assessment of differential treatment in job-offer outcomes. Under this design, any systematic variation in hiring outcomes by applicant gender can be attributed to manager bias. We estimate this bias using the following linear probability model:

$$\begin{aligned} \text{OFFER}_{ij} = \alpha + \beta_1 \cdot \text{Female}_i + \beta_2 \cdot \text{Female Manager}_j + \beta_3(\text{Female}_i \times \text{Female Manager}_j) \\ + \mu_1 \cdot \text{Ability}_i + \mu_2 \cdot \text{Personality}_i + \gamma X_{ij} + \varepsilon_{ij}, \end{aligned} \tag{1}$$

where OFFER_{ij} is an indicator equal to one if applicant i received a job offer from store manager j. Female_i denotes the applicant's gender, Female Manager_j indicates whether the applicant is assigned to a store j managed by a female, and the interaction term captures potential differences in gender bias across manager types. The variables Ability_i and Personality_i proxy the applicant's latent traits. Since these scores are correlated across multiple tests, we summarize them using principal component analysis (PCA), denoted by $\text{Ability}_i \equiv \text{Ability}_i(T) \in \mathbb{R}^k$ and $\text{Personality}_i \equiv \text{Personality}_i(T) \in \mathbb{R}^k$, where T is the set of assessment tests relevant for the occupation to which applicant i applied. We retain the first two components for each measure to capture the primary variation in the assessment battery.

The vector X_{ij} includes a rich set of covariates to account for temporal, managerial, and branch-level characteristics. These include month-year fixed effects and day-of-week fixed effects, applicant age, manager age and tenure, store years of operation, and fixed effects for geographic region and store type (e.g., mall, neighborhood, or downtown). We also include a measure of the store manager's empowering leadership style, ranging from 1 to 5, as evaluated by the manager's current subordinates in the year of the application.[7] Standard errors are clustered at the store–month–year

[7]The empowering leadership measure captures the extent to which store managers exhibit behaviors that empower

level to account for within-store correlation over time in unobserved hiring shocks.

5 Gender Bias in Job-Offering

Table 3 presents estimates from Eq. (1). Panel A reports pooled estimates across all applicants, and Panel B reports results separately by manager gender, based on interactions between applicant gender and manager gender. Column (1) of Panel A reveals a statistically significant bias in favor of male applicants across all occupations. Female candidates are approximately 2 percentage points less likely to receive a job offer than otherwise comparable male candidates. This estimate represents the ability- and personality-adjusted gender gap, capturing the average effect of applicant gender on job-offering decisions.

These results naturally raise the question of whether differential treatment varies by occupational skill level. Within the framework of our theoretical model, we examine heterogeneity in the job-offer function $H_j^p(a, f, u)$ across occupational skill level. We re-estimate Eq. (1) by interacting applicant and manager gender with a high-skill indicator, High Skill_i, which equals one for pharmacists and optometrists and zero for low-skill roles.[8]

Disaggregating the gender bias by occupational skill level, as shown in Columns 2 and 3 of Table 3, reveals substantial heterogeneity. The aggregate gender gap is entirely driven by high-skill occupations. Male applicants in high-skill roles are 8.2 percentage points more likely to receive a job offer than similarly qualified female applicants. In low-skill occupations, the estimated gender gap is negligible and statistically indistinguishable from zero. The table further shows that the difference between high- and low-skill roles is both statistically and economically significant, with the gender gap 7.3 percentage points larger in high-skill occupations.

These findings are consistent with models of statistical discrimination, in which gender serves as an informative signal for unobserved productivity or ability (Arrow, 1971; Phelps, 1972). In high-skill roles, where individual performance is more difficult to infer from standardized assessments alone, managers may rely more heavily on group-level averages. Within this framework of statistical

and guide employees, with higher scores indicating a greater prevalence of this leadership style.

[8] Pharmacists and optometrists are classified as high-skill due to their post-secondary educational requirements.

discrimination, such behavior represents a rational, albeit biased, response to higher costs of a "bad hire" in complex roles (Aigner and Cain, 1977; Altonji and Blank, 1999). The concentration of bias in high-skill roles also points to the presence of statistical stereotypes, which are preconceived beliefs about the distribution of ability or technical competence (Bohnet et al., 2016). These stereotypes disproportionately penalize women in professional tiers while remaining dormant in routine, low-skill tasks (Bordalo et al., 2019).

TABLE 3

High-skill workers in our sample are primarily pharmacists, an occupation described by Goldin (2014) and Goldin and Katz (2016) as exhibiting high "linearity." In linear occupations, earnings are proportional to hours worked because workers are highly substitutable; consequently, gender gaps are expected to be minimal. The evidence in Panel (a) challenges this notion, suggesting that structural linearity alone does not eliminate gender bias at the job-offer stage. Even in such roles, if managers hold statistical stereotypes about female pharmacist's inherent productivity (Kleven et al., 2019; Marianne, 2011), they may discount women's high assessment scores relative to those of comparable male peers. This suggests that while linearity may reduce the penalty associated with temporal flexibility, it does not necessarily mitigate the impact of subjective manager evaluations or biased priors (Blau and Kahn, 2017).

Panel (b) reports estimates by manager gender. The results indicate that the bias against female applicants is driven primarily by male managers. Row (a) shows that male managers exhibit a significant bias against women: female applicants evaluated by men are 3.6 percentage points less likely to receive a job-offer overall, with the gap reaching 10.3 percentage points in high-skill roles and 2.1 percentage points in low-skill roles. In contrast, row (b) shows that female managers do not exhibit statistically significant bias in either category.

This divergence suggests that the "informational value" of gender is not objective, but instead depends on manager's prior beliefs. If male managers hold more rigid statistical stereotypes regarding gendered aptitude in high-skill domains, they will reach different hiring conclusions than female managers, even when faced with identical applicant information. This points toward "belief-based" statistical discrimination (Coate and Loury, 1993). Under this framework, bias stems not from

actual differences in group productivity, but from inaccurate or exaggerated perceptions of those differences, particularly among male evaluators (Bohren et al., 2019; Bordalo et al., 2019; Hentschel et al., 2019).

The magnitude of these effects is substantial. Differences of 2 to 5 percentage points in job-offer rates, and especially gaps of 8 to 10 percentage points in high-skill roles, imply significant exclusion of qualified female applicants. The institutional setting of our data, in which applicants are quasi-randomly assigned to stores and evaluated using identical standardized assessments, allows for precise identification of these disparities within observational data.[9]

Low-skill occupations. We explore how contextual factors captured by U_{ij} are related to the job-offering decision. Table 4 decomposes the low-skill occupations into four categories: pink-collar, blue-collar, white-collar, and managerial roles. We then examine whether gender disparities differ across these categories, which reflect meaningful distinctions within low-skill jobs. Pink-collar roles (e.g., customer-facing or service-oriented positions) involve substantial interpersonal interaction. Blue-collar positions involve physical or logistics-related tasks. White-collar roles include administrative or office-based work, and managerial positions refer to roles such as shift managers that do not require post-secondary education.

TABLE 4

Panel (a) of Table 4 shows, on average, a negative bias against women in pink-collar occupations, although the estimates are not statistically significant.In contrast, blue-collar positions show a positive and statistically significant coefficient, indicating that the likelihood of offering a job-offer to a woman is 11 percentage points higher than for man in these roles.

Panel (b) provides the results separately for male and female managers. For male managers, the estimates indicate bias against women in pink-collar and white-collar occupations, though these differences do not reach statistical significance in these subsamples, likely due to reduced sample size. In blue-collar positions, male managers are 19 percentage points more likely to provide a job-offer to a female applicant. This pattern is consistent with the psychological mechanism of moral

[9]Moreover, these estimates may represent a lower bound of the bias if managers perceive female applicants in high-skill positions as a positively selected subgroup.

cleansing, where initial discriminatory decisions against an outgroup member (i.e., female in our case could threaten one's moral organizational-image (i.e., store manager), thus prompting them to perform later on compensatory, pro-social actions in favor of the initially biased outgroup in other contexts. Prior research on moral cleansing shows how prior immoral decision can motivate and be followed by moral compensatory actions (Jordan et al., 2011; Sachdeva et al., 2009; Zhong and Liljenquist, 2006).

Panel (b) also presents female managers' job-offer patterns in low-skill occupations. The estimates for pink- and blue-collar occupations are statistically insignificant. In white-collar roles, the estimate is positive but does not reach statistical significance, presumably due to the small sample size. In managerial positions, the estimates are large and significant, showing that female managers are 5.3 percentage points more likely to provide a job-offer to a female applicant than to an equally qualified male.

These results aligns with the broader literature on homophily, which suggest that similarity increases communication quality, trust, and interaction frequency, particularly in roles requiring close coordination. For example, Fortin et al. (2022) document homophily among female managers that leads to larger pay increases for female employees than for male employees. Thus, female managers may be more likely to support female candidates in managerial roles that involve close interaction and communication where they anticipate better fit, work environment, and productivity with female shift managers, factors that are particularly salient in roles characterized by intensive day-to-day interaction (Cohen and Huffman, 2007; Dickinson et al., 2018; Ibarra, 1992).

Interestingly, Cullen and Perez-Truglia (2023) emphasize the role of social interaction in shaping gender gaps, noting that men are more likely to form professional connections through channels that are less accessible to women. They further find that women are not necessarily more successful under female management, suggesting that institutional context, such as the financial-sector setting they study, may shape these dynamics. Our results indicate that variations in occupational context provide a potential explanation for these disparate findings.

6 Supplemental Analysis

6.1 Quasi-experimental design

In this section, we use a quasi-experimental research design, leveraging incoming and outgoing managers. We exploit changes in the gender of store managers to examine differential treatment of male and female applicants by manager gender. Specifically, we assess whether job-offers change when a male (female) manager is replaced by a female (male) manager, using same-gender manager transitions as a control group.

Table 5 presents the estimates, showing that the only statistically significant change occurs when a female manager is replaced by a male manager. All other types of manager transitions, fail to yield statistically significant changes in the likelihood that female applicants receive a job offer in either high- or low-skill positions.

While we do not observe statistically significant effects for low-skill positions, the results for high-skill jobs are striking. When leadership shifts from a female to a male manager, the estimated bias against female applicants increases sharply: female applicants are 70.1 percentage points less likely to receive a job-offer under the new male manager. Although only 7 stores experienced a transition from a female to a male manager during the study period, this subsample accounts for 888 applications, providing robust support for our primary finding

TABLE 5

6.2 Suggestive mechanisms

To further explore the drivers of the observed hiring disparities, we examine heterogeneity in the differential treatment of female applicants across local cultural and geographic contexts. Employment opportunities are often concentrated in specific geographic hubs, which in our setting are primarily located in the center of the country. At the same time, women tend to exhibit lower geographic mobility. As a result, women may be less able to access these high-opportunity labor markets, potentially leading to less favorable employment outcomes (e.g., Le Barbanchon et al., 2021; Petrongolo and Ronchi, 2020). Under a framework of statistical discrimination, such con-

straints may also shape employers' beliefs. When women are perceived as having fewer outside options, employers may be more likely to rely on gender-based stereotypes leading to lower job-offer rates in periphery areas. If these stereotypes are socially constructed, the magnitude of the hiring penalty should vary significantly across socio-geographic environments that harbor distinct cultural norms and in cases that female have less opportunities.

Table 6 provides clear evidence of this heterogeneity. In Central Areas (Panel A), characterized by more progressive gender norms and a higher density of professional opportunities for women, we find no statistically significant hiring penalty for women in either high-skill or low-skill roles. The coefficients for both male and female managers are economically small and statistically indistinguishable from zero, suggesting that in these "modern" hubs, gender does not serve as a primary heuristic for skill.

In sharp contrast, Peripheral Areas (Panel B) exhibit a pervasive and significant penalty against female applicants. For high-skill positions in these regions, male managers are 13 percentage points less likely to offer a position to a woman than to a man. Notably, while female managers in the periphery also exhibit a negative coefficient for high-skill roles, they are significantly more equitable in low-skill hiring than their male counterparts, providing a 6 percentage point advantage relative to male managers. Finally, Panel C (Other Areas) serves as a baseline, with estimates consistent with those of the pooled sample.

This geographic divergence suggests that the hiring bias we document is not a fixed feature of the labor market, but varies with the local cultural environment. The concentration of the penalty in peripheral regions, where traditional gender roles remain more entrenched, strongly supports the hypothesis that gender-skill stereotypes serves as a key mechanism for statistical discrimination. This pattern is particularly salient in high-skill roles, where greater complexity of the position increases employer's reliance on subjective, location-specific priors.

TABLE 6

7 Discussion and Conclusions

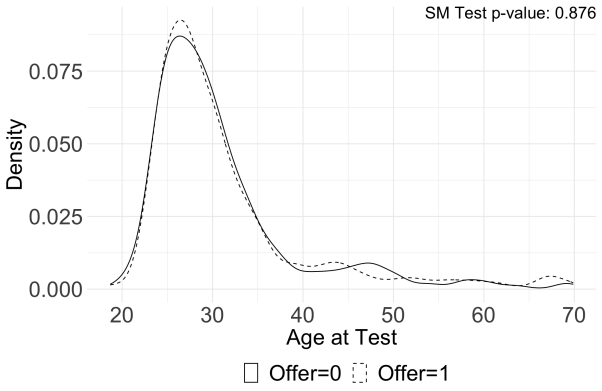
This study examines gender bias at a critical stage of the hiring process, the job-offer decision. We leverage a unique and rich dataset that combines detailed measures of applicant ability and personality with information on the gender of both applicants and managers across a wide range of occupations and skill levels. Our results document a persistent bias against female applicants during the initial evaluation. Even absent workplace interactions, equally qualified women are less likely to receive job-offers.

The institutional setting of our analysis is characterized by a blind application process, where candidates are quasi-randomly assigned to stores and evaluated without prior contact. This allows us to isolate the role of applicant gender and its interaction with the gender of the decision-maker. Within this setting, we document that the bias against female applicants is concentrated primarily in high-skill occupations and driven almost entirely by male managers.

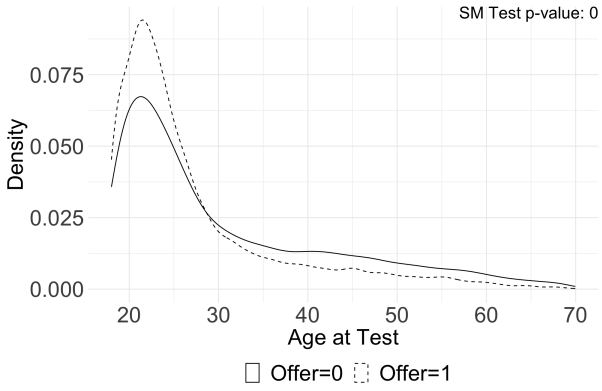
While the existing literature often documents bias within specific markets or stylized experimental settings, there is limited evidence on how such biases vary across occupational skill levels and manager gender. Our results show that gender bias is not uniform. Specifically, it is strongest in high-skill roles and when job-offer decisions are made by male managers. This means that the identity of the evaluator is as consequential as the credentials of the candidate in shaping early-career opportunities.

We acknowledge that our data are derived from a single firm in one country. However, the firm operates across diverse geographical markets and skill levels, providing a unified institutional setting to study heterogeneity across job types.

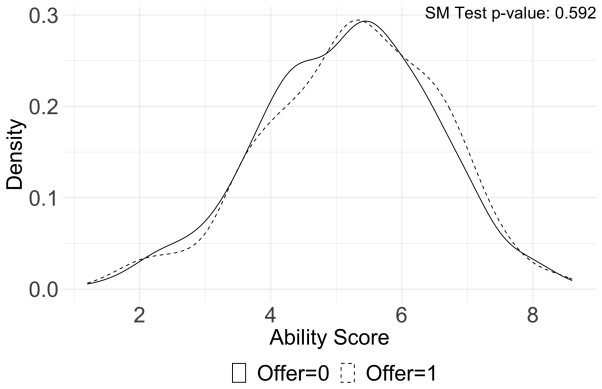
Our findings suggest several policy interventions to mitigate gender disparities. First, our results support the expansion of blind evaluation practices that prioritize objective assessments over gender-identifying information. Alternatively, implementing gender-diverse selection panels rather than relying on individual managers may dilute the impact of personal bias. Firms should also consider regular internal monitoring of job-offer outcomes by applicant and manager gender to detect and address disparate treatment in real time.



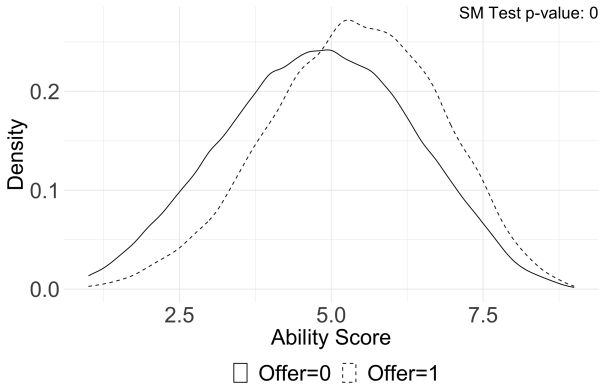
(a) Age at Test, High Skill Occupations



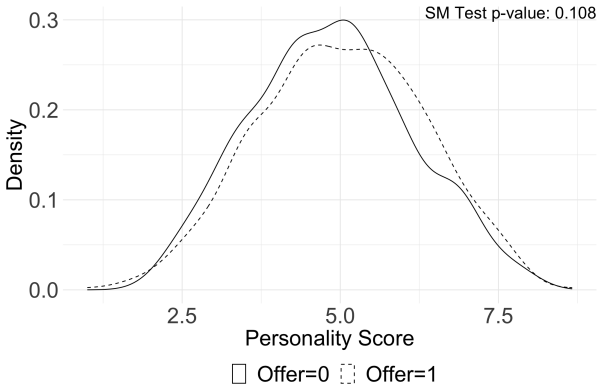
(b) Age at Test, Low Skill Occupations



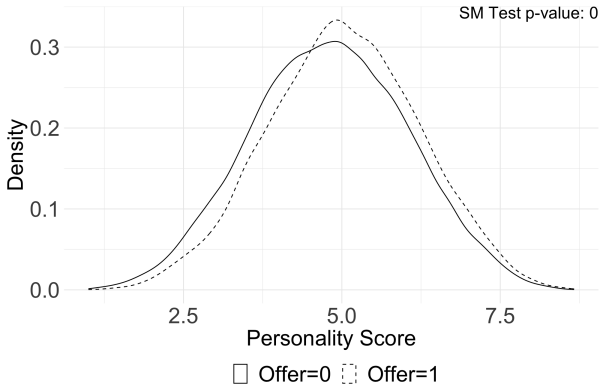
(c) Ability Score, High Skill Occupations



(d) Ability Score, Low Skill Occupations



(e) Personality Score, High Skill Occupations



(f) Personality Score, Low Skill Occupations

Figure 1: Distributions of applicants' age and assessment test scores, separately for applicants who were offered a position and those who were not.

Notes: Each panel reports kernel density estimates by hiring outcome.

Table 1: Summary Statistics

	Overall (1)	High (2)	Low (3)	*t-test* (3)-(2)
Panel A: *Applicant Characteristics*				
Female share	0.728	0.571	0.734	-10.99
	(0.445)	(0.495)	(0.442)	
Age at test	29.636	31.161	29.579	5.88
	(11.447)	(8.843)	(11.529)	
Offer=1	0.336	0.610	0.326	19.44
	(0.472)	(0.488)	(0.469)	
Standardized personality score	4.876	4.961	4.873	2.27
	(1.227)	(1.292)	(1.224)	
Standardized ability score	4.981	5.213	4.973	5.97
	(1.512)	(1.332)	(1.518)	
N	32087	1151	30936	
Panel B: *Manager Characteristics*				
Female share	0.57			
	(0.496)			
Age	43.203			
	(8.391)			
Tenure in the store	6.605			
	(5.599)			
Tenure in the Org.	17.751			
	(6.89)			
N	300			

Table 2: Balance Tests of Applicant Ability and Personality Scores by Manager Gender

	Male Applicants	Female Applicants
Panel A: *High Skill Occupations: Female vs. Male Manager*		
Ability Score	-0.096	0.138
	(0.135)	(0.107)
Personality Score	-0.101	-0.094
	(0.128)	(0.101)
N	494	657
Panel B: *Low Skill Occupations: Female vs. Male Manager*		
Ability Score	-0.022	0.003
	(0.037)	(0.020)
Personality Score	0.008	-0.020
	(0.029)	(0.016)
N	8241	22686

Notes: Each coefficient reports the regression-adjusted difference in applicant scores between female- and male-managed stores. Regressions include geographic cluster fixed effects to account for geographically constrained assignment of applicants to stores. Robust standard errors are reported in parentheses. Standard errors in parentheses. $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 3: The Likelihood of Offering a Position to Female Relative to Male Applicants

			Position Skill Level		
		All	High Skill	Low Skill	Diff
		(1)	(2)	(3)	(3)-(2)
Panel A: *Offer Likelihood: Female vs. Male Applicants*					
	All applicants	-0.020**	-0.082**	-0.009	0.073**
		(0.006)	(0.028)	(0.006)	(0.028)
Panel B: *Offer Likelihood by Manager Gender: Female vs. Male Applicants*					
(a)	Male manager	-0.036**	-0.103**	-0.021**	
		(0.009)	(0.039)	(0.009)	
(b)	Female manager	-0.007	-0.058	0.001	
		(0.008)	(0.039)	(0.008)	
(b)-(a)	Diff	0.029**	0.045	0.022*	
		(0.012)	(0.055)	(0.012)	
	N	32078	1151	30927	

Notes: Standard errors in parentheses. $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 4: The Likelihood of Offering a Position to Female Relative to Male Applicants in Low-Skill Occupations

	Position Type			
	Pink Collar	Blue Collar	White Collar	Managerial
Panel A: *Offer Likelihood: Female vs. Male Applicants*				
All applicants	-0.004	0.110**	0.011	0.035**
	(0.008)	(0.055)	(0.021)	(0.015)
Panel B: *Offer Likelihood by Manager Gender: Female vs. Male Applicants*				
Male managers	-0.014	0.190**	-0.035	0.011
	(0.012)	(0.081)	(0.034)	(0.024)
Female managers	0.006	0.022	0.039	0.053**
	(0.011)	(0.077)	(0.027)	(0.020)
N	21105	2697	2839	4286

Notes: $^{**}p < 0.05$.

Table 5: Manager Change Quasi-Experiment: The Effect on Position Offer Likelihood for Female vs. Male Applicants

		Position Type			
		High skill	Low skill	N	Unique Stores
(1)	Female to Female	-0.126	0.019	3393	24
		(0.166)	(0.036)		
(2)	Female to Male	-0.701**	0.015	888	7
		(0.277)	(0.072)		
(2)-(1)	Diff	-0.576*	-0.004		
		(0.323)	(0.081)		
(3)	Male to Male	0.080	-0.043	2658	17
		(0.167)	(0.041)		
(4)	Male to Female	0.217	0.018	558	5
		(0.199)	(0.090)		
(4)-(3)	Diff	0.138	-0.060		
		(0.260)	(0.099)		

Notes: $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 6: The Likelihood of Offering a Position to Female Relative to Male Applicants By Geographical Areas

		Position Type	
		High Skill	Low Skill
Panel A: *Central Areas*			
(a)	Male manager	-0.017	0.027
		(0.102)	(0.016)
(b)	Female manager	0.005	0.005
		(0.072)	(0.014)
	N	265	8405
Panel B: *Peripheral Areas*			
(a)	Male manager	-0.130**	-0.062**
		(0.057)	(0.016)
(b)	Female manager	-0.102	-0.002
		(0.063)	(0.015)
	N	488	9701
Panel C: *Other Areas*			
(a)	Male manager	-0.109*	-0.031**
		(0.063)	(0.013)
(b)	Female manager	-0.095	-0.014
		(0.070)	(0.012)
	N	398	12821

References

Agan, A. and Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics*, 133(1):191–235.

Aigner, D. J. and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30(2):175–187.

Altonji, J. G. and Blank, R. M. (1999). Race and gender in the labor market. *Handbook of labor economics*, 3:3143–3259.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The quarterly journal of economics*, 116(1):313–350.

Arrow, K. (1971). The theory of discrimination.

Ashnai, B., Mani, S., Kothandaraman, P., and Shekari, S. (2020). Gender bias in the recruitment of entry-level b2b salespeople. *Journal of Business & Industrial Marketing*, 35(8):1335–1344.

Barrick, M. R., Dustin, S. L., Giluk, T. L., Stewart, G. L., Shaffer, J. A., and Swider, B. W. (2012). Candidate characteristics driving initial impressions during rapport building: Implications for employment interview validity. *Journal of Occupational and Organizational Psychology*, 85(2):330–352.

Becker, G. (1957). (1971),"The economics of discrimination".

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4):991–1013.

Blais, A.-R. and Weber, E. U. (2001). Domain-specificity and gender differences in decision making. *Risk, Decision and Policy*, 6(1):47–69.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Bohnet, I., Van Geen, A., and Bazerman, M. (2016). When performance trumps gender bias: Joint vs. separate evaluation. *Management Science*, 62(5):1225–1234.

Bohren, J. A., Imas, A., and Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. *American economic review*, 109(10):3395–3436.

Booth, A. and Leigh, A. (2010). Do employers discriminate by gender? A field experiment in female-dominated occupations. *Economics Letters*, 107(2):236–238.

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

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2019). Beliefs about gender. *American Economic Review*, 109(3):739–773.

Borghans, L., Heckman, J. J., Golsteyn, B. H., and Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3):649–658.

Byrnes, J. P., Miller, D. C., and Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125(3):367.

Campero, S. and Fernandez, R. M. (2019). Gender composition of labor queues and gender disparities in hiring. *Social Forces*, 97(4):1487–1516.

Carlsson, M. (2011). Does hiring discrimination cause gender segregation in the swedish labor market? *Feminist Economics*, 17(3):71–102.

Carlsson, M. and Eriksson, S. (2019). Age discrimination in hiring decisions: Evidence from a field experiment in the labor market. *Labour Economics*, 59:173–183.

Chan, J. and Wang, J. (2018). Hiring preferences in online labor markets: Evidence of a female hiring bias. *Management Science*, 64(7):2973–2994.

Coate, S. and Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, pages 1220–1240.

Cohen, P. N. and Huffman, M. L. (2007). Working for the woman? female managers and the gender wage gap. *American sociological review*, 72(5):681–704.

Cullen, Z. and Perez-Truglia, R. (2023). The old boys' club: Schmoozing and the gender gap. *American Economic Review*, 113(7):1703–1740.

Dickinson, D. L., Masclet, D., and Peterle, E. (2018). Discrimination as favoritism: The private benefits and social costs of in-group favoritism in an experimental labor market. *European Economic Review*, 104:220–236.

Dupas, P., Modestino, A. S., Niederle, M., Wolfers, J., et al. (2021). Gender and the dynamics of economics seminars. Technical report, National Bureau of Economic Research.

Eagly, A. H. and Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3):573.

Eberhardt, M., Facchini, G., and Rueda, V. (2023). Gender differences in reference letters: Evidence from the economics job market. *The Economic Journal*, 133(655):2676–2708.

Eckel, C., Gangadharan, L., Grossman, P. J., and Xue, N. (2021). The gender leadership gap: Insights from experiments. In *A Research Agenda for Experimental Economics*, pages 137–162. Edward Elgar Publishing.

Fortin, N. M., Markevych, M., and Rehavi, M. (2022). Closing the gender pay gap in the us federal service: the role of new managers. Technical report, Working Paper.

Friedmann, E. and Efrat-Treister, D. (2023). Gender bias in stem hiring: Implicit in-group gender favoritism among men managers. *Gender & Society*, 37(1):32–64.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American economic review*, 104(4):1091–1119.

Goldin, C. and Katz, L. F. (2016). A most egalitarian profession: pharmacy and the evolution of a family-friendly occupation. *Journal of Labor Economics*, 34(3):705–746.

Goldin, C. and Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4):715–741.

Hangartner, D., Kopp, D., and Siegenthaler, M. (2021). Monitoring hiring discrimination through online recruitment platforms. *Nature*, 589(7843):572–576.

Hentschel, T., Heilman, M. E., and Peus, C. V. (2019). The multiple dimensions of gender stereotypes: A current look at men's and women's characterizations of others and themselves. *Frontiers in psychology*, 10:376558.

Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative science quarterly*, pages 422–447.

International Labour Organization (2023). New data shine light on gender gaps in the labour market. Ilo brief, ILO, Geneva.

Jordan, J., Mullen, E., and Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, 37(5):701–713.

Kleven, H., Landais, C., and Søgaard, J. E. (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender differences in job search: Trading

off commute against wage. *The Quarterly Journal of Economics*, 136(1):381–426.

Lippens, L., Vermeiren, S., and Baert, S. (2023). The state of hiring discrimination: A meta-analysis of (almost) all recent correspondence experiments. *European Economic Review*, 151:104315.

Marianne, B. (2011). New perspectives on gender. In *Handbook of labor economics*, volume 4, pages 1543–1590. Elsevier.

McCrae, R. R. and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, 52(1):81.

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, 60(2):175–215.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41):16474–16479.

Neumark, D. (2024). Age discrimination in hiring: Evidence from age-blind versus non-age-blind hiring procedures. *Journal of Human Resources*, 59(1):1–34.

Neumark, D., Burn, I., and Button, P. (2016). Experimental age discrimination evidence and the heckman critique. *American Economic Review*, 106(5):303–308.

OECD (2018). *Health at a Glance 2015: OECD Indicators*. OECD Publishing, Paris. 2015–2018.

Park, S. Y. and Oh, E. (2025). Getting a foot in the door: A meta-analysis of us audit studies of gender bias in hiring. *Sociological Science*, 12:26–50.

Petrongolo, B. and Ronchi, M. (2020). Gender gaps and the structure of local labor markets. *Labour Economics*, 64:101819.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, 62(4):659–661.

Purkiss, S. L. S., Perrewé, P. L., Gillespie, T. L., Mayes, B. T., and Ferris, G. R. (2006). Implicit sources of bias in employment interview judgments and decisions. *Organizational Behavior and Human Decision Processes*, 101(2):152–167.

Reuben, E., Sapienza, P., and Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, 111(12):4403–4408.

Ruffle, B. J. and Shtudiner, Z. (2015). Are good-looking people more employable? *Management Science*, 61(8):1760–1776.

Sachdeva, S., Iliev, R., and Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, 20(4):523–528.

Sapienza, P., Zingales, L., and Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences*, 106(36):15268–15273.

Sayeh, A. M., Badel, A., and Goyal, R. (2023). Countries that close gender gaps see substantial growth returns. *IMF Blog*.

Sherman, A., Atad, E., and Shtudiner, Z. (2025). Are employers happy to hire happy candidates? happiness and employability sources. *Journal of Behavioral and Experimental Economics*, 115:102334.

Ton, Z. and Huckman, R. S. (2008). Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science*, 19(1):56–68.

Veit, S., Arnu, H., Di Stasio, V., Yemane, R., and Coenders, M. (2022). The "big two"' in hiring discrimination: Evidence from a cross-national field experiment. *Personality and Social Psychology Bulletin*, 48(2):167–182.

Wu, A. H. (2020). Gender bias among professionals: An identity-based interpretation. *Review of Economics and Statistics*, 102(5):867–880.

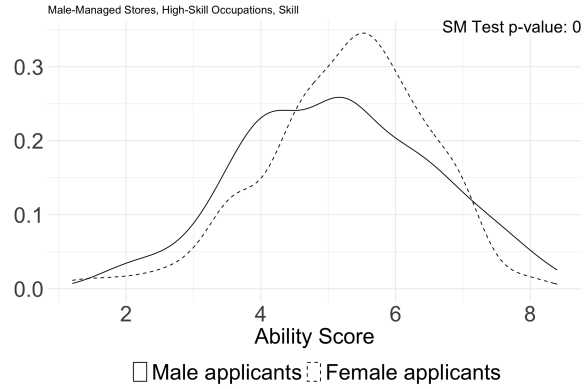
Zhong, C.-B. and Liljenquist, K. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, 313(5792):1451–1452.

Appendices

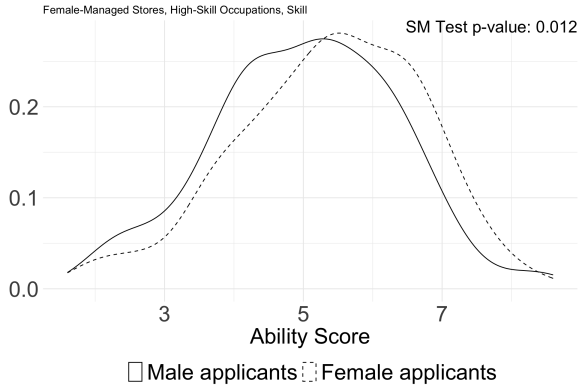
Table A1: Gender Share by Occupation

	Applicants Data		**Firm Data**	
Occupation	Female Share	Total	Female Share	Total
Administrative Position	0.98	224	0.99	375
Assistant pharmacist	0.78	2839	0.80	973
Pharmacist	0.55	1008	0.47	929
Optometrist	0.76	143	0.80	69
Sellers & service role	0.82	20889	0.86	3754
Store Inventory Keeper	0.03	2698	0.05	358
Shift Manager	0.70	4286		

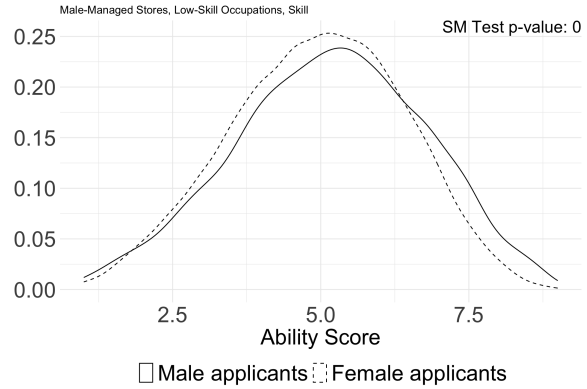
Notes: "Applicants Data" contains all workers used in the main analysis. The "Firm Data" is from 2015 and includes only workers not hired from the "Applicants Data." The Shift Manager role is missing for "Firm Data" because these employees also hold other positions; the firm data only lists the "other position" alongside the Shift Manager role.



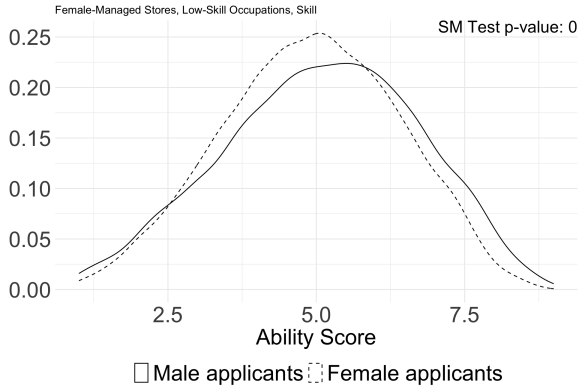
(a) Male-Managed Stores, High Skill Occupations



(b) Female-Managed Stores, High Skill Occupations



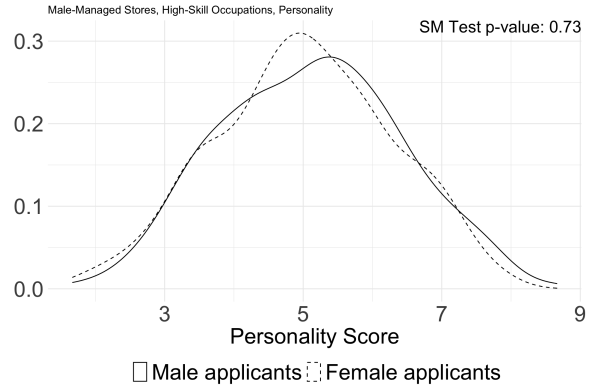
(c) Male-Managed Stores, Low Skill Occupations



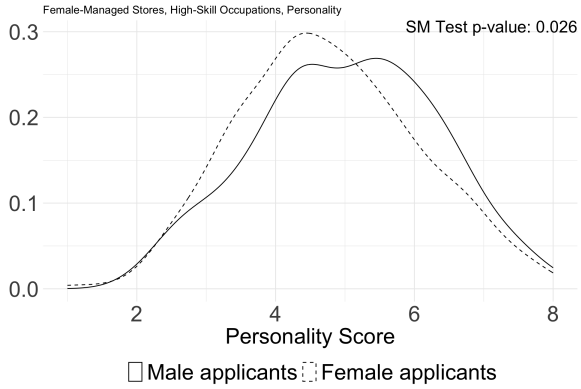
(d) Female-Managed Stores, Low Skill Occupations

Figure B1: Standardized Ability Score Distributions Across Applicant Gender and Occupation Skill Level.

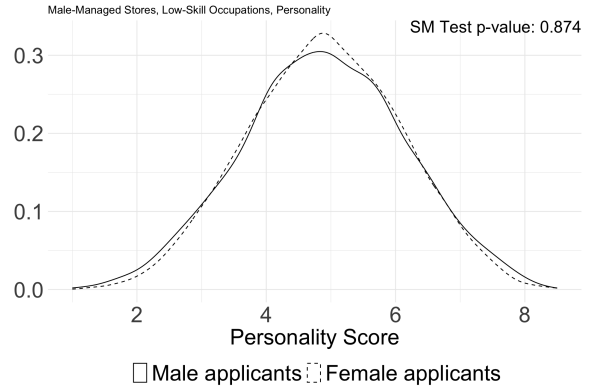
Notes: This figure plots kernel density estimates of the aggregated assessment test score for male and female applicants separately, conditional on whether they were assigned to male- or female-managed stores. p-values from the nonparametric two-sample test for equality of distributions are reported, indicating no statistically significant differences across manager gender.



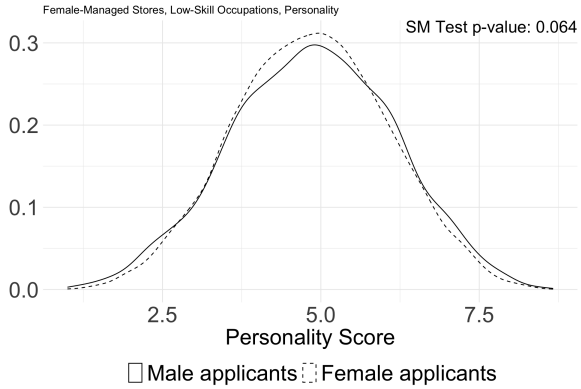
(a) Male-Managed Stores, High Skill Occupations



(b) Female-Managed Stores, High Skill Occupations



(c) Male-Managed Stores, Low Skill Occupations



(d) Female-Managed Stores, Low Skill Occupations

Figure B2: Standardized Personality Score Distributions Across Applicant Gender and Occupation Skill Level.

Notes: This figure plots kernel density estimates of the aggregated assessment test score for male and female applicants separately, conditional on whether they were assigned to male- or female-managed stores. p-values from the nonparametric two-sample test for equality of distributions are reported, indicating no statistically significant differences across manager gender.

Test Results Report
Confidential – Personal

This report contains personal, sensitive, and confidential information. The recipient is responsible for securing the confidentiality of the information and preventing any use other than for the sole purpose of assessing the candidate's suitability for the designated position. Disclosure of information from this report to the candidate without the professional judgment of a licensed psychologist may cause harm to the candidate.

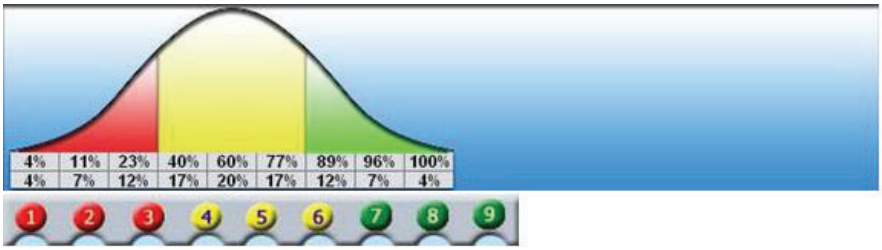
Candidate's Personal Details

Identification Number	
Examinee Name	
Gender	Female
Store Name	
Examination Date	06/11/2018
Examination Language	
Country of Birth	
Date of Birth	
Education	High School
Years of Education	12
Occupation	Service Team Member
Extra Time Accommodation	No
Scoring Norm	Cashier
Address	
Telephone Numbers	

Figure B3: Assessment test scores - Example - Page 1

Assessment Test Scores- A Summary Report

Cognitive Ability Scores



Test	Score (Out of 1-9)
Overall Impression	8
Visual Memory	5
Attention to Detail	6
Detail Recognition and Difference Identification	7

Overall Impression – Distribution

1	2	3	4	5	6	7	8	9

Visual Memory – Distribution

1	2	3	4	5	6	7	8	9

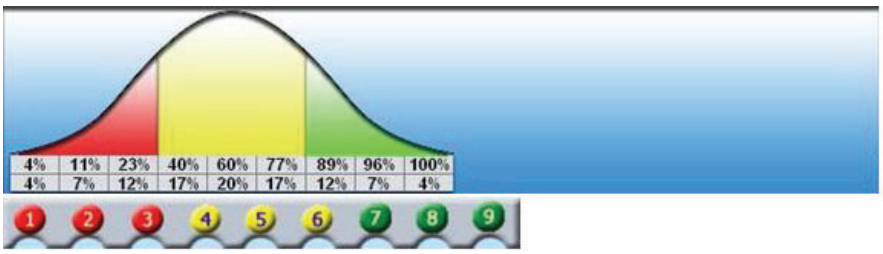
Attention to Detail – Distribution

1	2	3	4	5	6	7	8	9

Detail Recognition and Difference Identification – Distribution

1	2	3	4	5	6	7	8	9

Figure B3: Assessment test scores - Example - Page 2



Personality Test Scores

Test	Score (between 1-9)
Service Provision (Agreeableness)	4
Stress Management	3
Independence and Self Esteem	4
Openness and Extraversion	5
Organizational Commitment	6
Social Desirability	5

Service Provision (Agreeableness)– Distribution

1	2	3	4	5	6	7	8	9

Stress Management – Distribution

1	2	3	4	5	6	7	8	9

Independence and Self Esteem – Distribution

1	2	3	4	5	6	7	8	9

Figure B3: Assessment test scores - Example - Page 3

Openness and Extraversion – Distribution

1	2	3	4	5	6	7	8	9

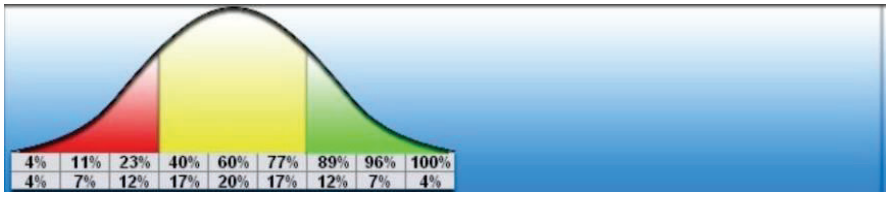
Organizational Commitment – Distribution

1	2	3	4	5	6	7	8	9

Social Desirability – Distribution

1	2	3	4	5	6	7	8	9

Overall Evaluation



	Score (between 1-9)
Overall Evaluation Score	6
Promotion Potential	5

Overall Evaluation Score – Distribution

1	2	3	4	5	6	7	8	9

Promotion Potential – Distribution

1	2	3	4	5	6	7	8	9

Figure B3: Assessment test scores - Example - Page 4

Assessment Summery

(score in parentheses between 1-9)

- **Overall Level (8) -** The candidate's analytical cognitive ability is very high. The candidate is capable of absorbing and processing various sources of information, creating the connections between them, and reaching correct conclusions. The candidate possesses a high intellectual level and a strong ability to comprehend new subjects and implement them into the work routine.
- **Visual Memory (5) -** Average memory ability. The candidate absorbs and recalls information with average accuracy and speed.
- **Attention to Detail (6) -** Average ability to identify details accurately, efficiently, and quickly.
- **Detail Recognition and Difference Identification (7) -** High ability to detect and identify graphical signs and symbols and to compare between them.

- **Service Provision - Agreeableness (4) -** Moderate in people-oriented approach. Friendly, outgoing, and warm at the average level, with the ability to enjoy working and interacting with people. The candidate can be attentive and kind when required. Many individuals with this score are reasonably suitable for service roles.
- **Stress Management (3) -** Slightly sensitive and emotional, which is noticeable when facing difficult or frustrating situations. May respond with embarrassment or stress that can impair work efficiency.
- **Independence and Self Esteem (4) -** Displays normal self-confidence with an average level of assertiveness and need to stand out. Capable of asserting his opinion when necessary and appears able to exercise formal authority when given. Individuals with this score are reasonably suitable for management roles.
- **Openness and Extraversion (5) -** Balanced, moderate, and composed personality. Open to change and new experiences to a reasonable extent, but not adventurous or highly spontaneous. Tends to evaluate and consider before making decisions.
- **Organizational Commitment (6)** - The candidate appears to have a normal level of responsibility, diligence, perseverance, and loyalty to the work systems and operates. Fairly conforming, respects social norms and follows moral values, at the average level without standing out more than others in these aspects.

Figure B3: Assessment test scores - Example - Page 5