Whom You Know, Whom You Recommend, and the Role of Referrals in Gender Occupational Segregation*

Livia Alfonsi
Harvard Business School

Pedro Ferreira
Harvard Kennedy School

26 March 2026

Preliminary. Please Do Not Circulate.

Abstract

Occupational gender segregation persists even among workers with identical credentials. We study whether and how employee referral decisions contribute to this persistence using a high-stakes field experiment with 750 skilled workers across eleven occupations in Uganda. Our design embeds randomized candidate profiles in real workplace referral decisions, separating two channels: network composition (whether workers can access gender counter-sterotypical candidates) and a referral gender bias (whether they choose to refer them). Gender counter-stereotypical candidates – women in male-dominated occupations and men in female-dominated ones – face a 25 percentage-point referral penalty relative to otherwise identical peers: 29 points when women are the counter-sterotypical candidate and 18 points for when men are. The gap is not explained by network constraints: among workers with contacts of both genders, the penalty persists. These penalties reflect lower perceived likability, expected task ability, anticipated job enjoyment, and second-order beliefs about employer preferences – channels that are particularly pronounced for women trying to work in male-dominated occupations. A calibrated dynamic network model shows that employee referral bias increases steady-state occupational segregation by 8.2–9.2 percent in a conservative baseline; when referrals play a larger role in job finding, the bias can account for more than a quarter of observed segregation. Referral preferences are thus a powerful mechanism sustaining gender gaps in labor markets.

*Alfonsi: Harvard Business School (lalfonsi@hbs.edu). Ferreira: Harvard University (pedro_ferreira@hks.harvard.edu). We thank Lori Beaman, Michal Bauer, Emily Breza, Julie Chytilová, Katie Coffman, Alexia Delfino, Gaia Dossi, Stefano Fiorin, Pat Kline, Gabriel Kreindler, Eliana La Ferrara, Elisa Macchi, Jeremy Magruder, Ted Miguel, Dev Patel, Elisabeth Sadoulet, Christopher Walters and seminar participants for helpful comments. Funding from the J-PAL Gender and Economic Agency Initiative is gratefully acknowledged. AEA RCT Registry: AEARCTR-0012908. We received IRB approval from the Committee on the Use of Human Subjects of the Harvard University Area: IRB23-1074. All errors are our own.

1 Introduction

Occupations remain highly segregated by gender, with large and well-documented consequences for aggregate productivity and growth (Hsieh et al. 2019; Goldberg et al. 2025). Existing explanations emphasize supply-side constraints—such as childcare, safety, norms, and role models—as well as demand-side forces, primarily employer discrimination and gendered preferences.[1] Yet, employers typically hire from a pool of candidates that is itself endogenously formed, often through referrals (Topa, 2011; Beaman and Magruder, 2012; Beaman et al., 2018; Fafchamps and Moradi, 2015; Witte, 2025). Despite the central role referrals play in determining who enters these hiring pools, little is known about how workers' own preferences and beliefs about potential candidates shape referral decisions, and whether these choices systematically contribute to gender occupational segregation.

We study this question using a high-stakes field experiment embedded in urban labor markets in Uganda, spanning eleven occupations that vary widely in gender composition. We recruit 750 workers employed in these occupations and invite them to refer a candidate for a six-week subsidized internship at the firm where they work, with a monetary reward if the referred candidate is retained. The experimental design allows us to separate referrer's beliefs and preferences from the composition of their networks. In Part I, respondents see pairs of candidate profiles that are identical except for gender and labor-market experience, which are cross-randomized, and choose to refer one of the candidates to their employer. This correspondence-style task shuts down the endogenous *network composition channel* (the respondent does not know someone qualified of the 'other' gender) and isolates what we call the *referral gender bias* (they see a qualified candidate of the 'other' gender, but systematically choose not to refer them). In Part II, respondents may nominate real contacts from their own networks, one per gender, and then make a final referral choice among all available candidates (across Part I and Part II).

We document five main findings. First, results from Part I show a large referral gender bias (RGB) exists and is tightly linked to occupational gender dominance. Gender counter-stereotypical (GCS) candidates—women in male-dominated occupations (MDO) and men in female-dominated ones (FDO)—are substantially less likely to be referred than otherwise identical peers. On average, they face a penalty of 25 percentage points. This penalty is asymmetric: women in male-dominated occupations face a referral gap of 29 percentage points, while men in female-dominated occupations face a gap of 18 percentage points. Moreover, the referral gap increases sharply with the degree of occupational

[1] See for example: Bertrand (2011); Bertrand and Duflo (2017); Booth and Leigh (2010); Cortes and Pan (2018); Cortés et al. (2026); Delfino (2024); Fernández and Fogli (2009); Fogli and Veldkamp (2011); Goldin (2014b); Heath and Mobarak (2015); Heath et al. (2024); Ibanez and Riener (2018); Jayachandran (2015, 2021); Kline et al. (2022); Kuhn and Shen (2013).

segregation. Additional experience partially mitigates the penalty for men in female-dominated occupations but does little to offset the disadvantage faced by women in male-dominated occupations.

Second, results from Part II show that network composition is itself highly gendered. Beaman et al. (2018) document gendered referral networks in a less segregated setting, where men are advantaged. We find a similar pattern, but one that closely tracks the gender dominance of the occupation. Among respondents who named at least one potential contact, in female-dominated occupations 85% could name a woman but only 31% could name a man; in male-dominated occupations, 93% could name a man but only 22% could name a woman. This shows that the *network composition channel* plays an important role in both male-dominated and female-dominated occupations.

At the same time, although the relevant sample is small and selected, referral choices conditional on naming contacts of both genders provide suggestive corroboration that the experimental patterns extend to real networks, especially for workers in male-dominated occupations where males were referred more than twice as often as females (53.3 % vs. 23.3 %). Taken together, this evidence suggests that gendered referrals reflect not only network availability, captured by the *network composition channel*, but also differential evaluation of available network candidates, captured by the *referral gender bias*.

Third, we shed light on the mechanisms underlying referral behavior. We model the referrer's decision as the maximization of a single value function over candidates, where the value of a referral depends on two additive components: how much the referrer would enjoy working closely with the candidate (Heilman et al., 2004), and the expected monetary return from retention. In Part II, reciprocity benefits may also matter, but in the experiment they are shut down by design. Both margins display important asymmetries by occupational gender dominance. In male-dominated occupations, women are perceived as less likable and less likely to be retained than otherwise identical men, and these penalties persist even with additional experience. In female-dominated occupations, by contrast, men face no statistically significant likability penalty, and while they also tend to be rated as less likely to be retained, these differences are smaller and estimated less precisely; additional experience largely offsets these disadvantages. While likability captures respondents' own taste for working with a candidate, expected retention is a more composite measure. We posit that it reflects three underlying beliefs: the candidate's perceived ability to complete tasks, second-order beliefs about whether employers will like the candidate, and the candidate's expected enjoyment of the internship. Here too, the asymmetry is clear. Women in male-dominated occupations are seen as less capable, less likely to enjoy the job, and less likely to be liked by employers, while men in female-dominated occupations face weaker penalties and benefit more from added experience. Respondents expect employers to share their own dislikes, but to hold them more strongly. The same qualitative patterns appear when respondents evaluate

real contacts from their own networks, suggesting that network ties do not undo the recommendation-stage penalties faced by gender counter-stereotypical candidates. Free-text descriptions of candidates reinforce this asymmetry in how information is processed. Across 1,506 anonymized profiles, respondents provided nearly 5,000 descriptive adjectives, which we standardize and group into categories capturing competence (e.g., experience, education, cognitive ability), effort and character (e.g., hardworking, caring), professional traits, and demographic descriptors. Women are significantly more likely to be described using demographic terms, including explicitly gendered and age-related descriptors, as well as communal and "grindstone" traits (i.e., descriptors that emphasize effort and industriousness rather than innate ability) that emphasize effort and personality (Eberhardt et al., 2023; Beugnot and Peterlé, 2020), while men are more frequently characterized using competence-coded attributes such as intelligence, skill, education, and strength. These patterns are especially pronounced among less-experienced candidates and sharpen by occupation: in male-dominated occupations, female candidates attract more gendered, age-related, communal, and negative descriptors, whereas competence-coded language tilts more clearly toward men; by contrast, in female-dominated occupations, explicit gendered descriptors largely disappear and negative characterizations are not systematically gendered.

Fourth, we compare employees' beliefs with employers' actual preferences using a smaller sample of firms where we replicate the evaluation exercise with hiring managers. In line with the discrimination literature that looked at employers' choices, employers also display a preference for gender-stereotypical candidates, but the magnitude of this bias is smaller than employees anticipate and smaller than the bias exhibited by employees themselves. This suggests that referral gaps are not merely a passive pass-through of employer discrimination: employees' own preferences and beliefs amplify segregation by discouraging referrals of GCS candidates.

Finally, we embed our experimental estimates in a dynamic network model of job referrals to quantify how much occupational segregation is sustained by referrals, both overall and through the two components of the referral channel: the *network channel* and *referral gender bias*. Building on Calvo-Armengol and Jackson (2004), our model allows occupational segregation to arise from three sources: (i) heterogeneous direct job arrival rates, (ii) homophilous networks, and (iii) gender-biased referral weights calibrated directly from the experiment. Workers receive job signals either directly or through referrals, can forward at most one signal per period to a network contact, and face stochastic job destruction. We estimate the model by simulated method of moments to match observed employment distributions by gender and occupation, and use survey data to calibrate network size and homophily. The model confirms that, when all three forces are shut down, the economy exhibits no gendered occupational segregation. Heterogeneous direct job arrival rates alone generate sizable gender gaps in male- and

female-dominated occupations, capturing non-referral forces such as differential employer demand or access to vacancies. By contrast, homophily alone generates no segregation. Intuitively, if all nodes receive signals at the same rate and do not systematically forward them to particular genders in their network, then gender clustering in networks has no implications for occupational segregation. Referral gender bias, by contrast, generates segregation even in isolation, because it distorts forwarding decisions as a function of the recipient's gender and the type of signal. Intuitively, even if all nodes receive the same signals at the same rates and networks are not clustered by gender, segregation will still arise if nodes are more likely to forward occupation 1 signals to men and occupation 2 signals to women.

We then use the model to study how these channels interact, and in particular to quantify the extent to which referral gender bias is a central driver of occupational segregation. Specifically, we compare a benchmark economy in which occupational segregation arises only from differential direct access to job signals and the network composition channel to an economy that additionally features the experimentally estimated referral gender bias. Under our most conservative calibration, introducing referral gender bias increases steady-state occupational segregation by 8.2 to 9.2 percent, depending on whether occupations are male- or female-dominated. Under alternative calibrations, we proxy heterogeneous direct job arrival rates using measures drawn directly from our data, rather than recovering them through the simulated method of moments. Using either the employer experiment or survey-based measures of job arrival rates, we find an even larger role for referral gender bias, which can account for more than one quarter of observed segregation. More broadly, the model predicts, and the counterfactual exercises confirm, that the contribution of referral gender bias is larger when referrals play a greater role in job allocation, when direct job arrival rates are less gender-segregated, when direct job arrival rates are less relevant for the job switching process, when networks are denser, and when they are less homophilous.

Taken together, our findings show that referral gender biases run in both directions but are largest for women in male-dominated occupations. Employee referrals are not simply a byproduct of network composition or a transmission of employer preferences; they reflect workers' own stereotypes and beliefs and constitute a powerful mechanism sustaining gender occupational segregation by shaping who enters referral-based hiring pipelines.

Contributions First, this paper contributes to the literature on the determinants of gender occupational segregation. Prior work has emphasized supply-side barriers such as childcare constraints, safety concerns, social norms, and limited role models, as well as demand-side forces including employer discrimination and consumer preferences, with much of the most direct evidence coming from hiring decisions in high-income countries (Adda et al., 2017; Agüero and Marks, 2011; Bjerk and Han, 2007; Cristia, 2008;

Dinkelman, 2011; Fernández et al., 2004; Fernández-Kranz and Rodríguez-Planas, 2021; Flory et al., 2015; Fogli and Veldkamp, 2011; Goldin, 2014a; Goldstein et al., 2019; Gradín, 2021; Greenwood et al., 2005; Lommerud et al., 2015; Neumark and Postlewaite, 1998; Nicoletti et al., 2018; Niederle and Vesterlund, 2007; Pan, 2015; Lepage et al., 2025). A growing body of research highlights *access margins*: the channels through which workers are matched to jobs and their qualifications enter, or fail to enter, the hiring funnel (Kuhn and Shen, 2013; Buchmann et al., 2025; Olivetti et al., 2024; Heath et al., 2024; Bohren et al., 2025). We show that referrals are a quantitatively important and largely overlooked access margin, particularly in low-income settings where formal job postings are scarce and many employment relationships originate through personal connections. Critically, referral discrimination excludes qualified candidates precisely at the point of entry. The candidates in our study have completed two-year post-secondary vocational programs and accumulated labor market experience, yet they face systematic referral penalties based on gender. To quantify aggregate implications, we develop and calibrate a dynamic network model that suggests that the referral gender bias increases occupational segregation by 8.2–9.2 percent in a conservative baseline, and can account for between a quarter and more than half of observed segregation under alternative calibrations.

Second, this paper contributes to the emerging literature on gendered referral behavior in the labor market. Some evidence of gendered referral patterns exists (Beaman et al., 2018; Zeltzer, 2020). However, we do not know whether such patterns primarily reflect supply-side constraints in network composition, for example due to homophily that leads individuals to form ties with similar others, such that workers can only refer candidates from within their gender-segregated circles, or whether they also reflect differential evaluation of, and anticipated discrimination against, gender counter-stereotypical candidates. Experimental evidence on these themes remains scarce: Beaman et al. (2018) document pro-male bias in a competitive recruitment drive conducted by a research organization in Malawi, a gender-neutral industry; Beugnot and Peterlé (2020) find female homophily in a laboratory experiment; and Hederos et al. (2024) document gendered-referrals among Swedish students. We extend this evidence in three ways. First, our experiment involves real job opportunities and incumbent employees with meaningful stakes: workers who interact daily with potential coworkers and whose professional reputation depends on whom they recommend, in occupations that differ sharply in gender composition. By embedding randomized profiles in actual referral decisions, we move beyond student samples and laboratory contexts and capture how the identity, incentives, and beliefs of the referrer shape outcomes, something classic correspondence designs cannot do (Azmat and Petrongolo, 2014). Second, our design separates the *network composition channel* from a *referral gender bias*. Referral penalties are comparable in magnitude when network constraints are shut down by design and when respondents draw on their own contacts, implying that the gap reflects differential evaluation rather than a lack of

cross-gender ties. Third, we provide evidence on the mechanisms underlying the referral gender bias. GCS candidates are penalized on perceived competence, likability, and expected job enjoyment, and these penalties are reinforced by employees' second-order beliefs about employer preferences. Additional experience closes these gaps for men in female-dominated occupations but not for women in male-dominated ones, a pattern we trace using nearly 5,000 free-text adjective descriptions that reveal how the same credentials are systematically interpreted differently depending on candidate gender and occupation.

2 Context

2.1 Ugandan Labor Markets

FLFP and Occupational Segregation Even though female labor force participation in Sub-Saharan Africa is relatively high (72% of women engage in some form of work, 36% are employed outside agriculture), this participation is concentrated in low-paying occupations of the economy (Asongu and Odhiambo, 2019). Uganda is no exception: its labor market remains highly segmented by occupation: in nationally representative data, women are disproportionately concentrated in agriculture, while men are more represented in industry and higher-value services. Consistent with these economy-wide patterns, women are disproportionately clustered in the least lucrative vocational occupations such as cosmetology, tailoring, and caregiving, while men dominate higher-paying professions like electrical work and automotive repair (Alfonsi et al., 2023).

Referrals Prevalence In our sample, 62% of respondents report finding their first job through a personal connection, and 25% report that this connection was a referral by an incumbent employee (Figure A.1). Figure A.1 juxtaposes these shares with those documented in Alfonsi et al. (2020), which studies an earlier cohort of vocational graduates in Uganda using a harmonized measure of job-finding channels.[2] The striking similarity across cohorts underscores that personal connections, and in particular referrals, are a persistent feature of early-career job finding in Uganda.

Referring others is also common among relatively young workers: many study participants report having referred at least one person for a job in the previous 12 months. A closer look at the referral module in our survey highlights the nature of these links. Among the 248 referrals for which we collected detailed relationship information, only about 10% involve close or distant relatives. Most referrals are mediated by friendship

[2]Specifically, Alfonsi et al. (2020) tracks vocational graduates from Central Uganda several years prior to our study and elicits the primary channel through which they obtained their first job. We harmonize categories across the two studies into five groups: referrals by non-employees, referrals by incumbent workers, employer connections, walk-ins, and other methods.

ties: roughly 80% are friends, and the remaining 10% are friends-of-friends, relatives of friends, or coworkers' contacts. Referrals therefore operate largely through extended social networks rather than kinship. While most referrals occur between individuals who have interacted recently, in about 3% of cases the referrer reported having had no prior interaction with the referred person.

However, referral behavior is not gender-neutral. In our data, referrals made by male workers disproportionately favor men: 82% of male referrals are for male candidates. Female referrals similarly tilt toward women: 75% of referrals made by women are for female candidates. These patterns are even stronger within gender-stereotypical employment contexts. Among men working in male-dominated occupations, 84% of referrals are for male candidates; among women working in female-dominated occupations, 82% are for female candidates.

These facts motivate the central question of the paper: do gendered referral patterns primarily reflect supply-side constraints in network composition, such that workers can only credibly refer candidates from within their own (gender-segregated) circles, or do they additionally reflect differential evaluation and anticipated discrimination against GCS candidates? The policy implications differ sharply. If network composition is the dominant force, then interventions that increase women's presence in male-dominated occupations, such as scholarships, apprenticeships, mentorship, or targeted recruitment, may be sufficient to narrow gender gaps. If, instead, referral decisions embed systematic penalties for GCS candidates, then closing occupational disparities likely requires more transformative changes that shift workplace norms, culture, and beliefs about fit and performance.

2.2 Study Sample

Our study draws on two waves of an experiment with skilled young workers in urban labor markets in Central and Eastern Uganda (total $N = 1{,}308$). In an initial wave in 2021 ($N = 555$), we documented a large referral gender bias, alongside systematic differences in perceived likability and expected post-internship candidate retention. We then fielded a second wave in 2023 ($N = 753$) with a comparable population, explicitly designed to sharpen identification and probe mechanisms. This second wave expands the experimental variation, including randomization around experience to study how the referral gender bias interacts with experience premia, and adds outcome elicitations aimed at the underlying channels, including second-order beliefs. We treat this second wave as our main sample given its expanded scope and improved design. Appendix Figure A.3 compares the GS–GCS gap in referral probability, likability, and retention probability across the two waves, restricting the 2023 sample to the two treatment arms present in both waves (high vs. low experience). Crucially, the cross-sample comparison shows that

estimates are closely aligned for all three outcomes, whether comparing the first-wave estimates raw or reweighted to match the occupational composition of the second wave.

These young adults completed post-secondary education at well-established vocational training institutes (VTIs) in Central and Eastern Uganda between 2015 and 2021, earning a National Certificate in a range of manufacturing and service occupations. Appendix Table A.1 reports baseline summary statistics and confirms balance across experimental conditions. As of 2023, respondents were on average 24 years old; 37% were women; 55% grew up in a rural area; 52% were wage-employed at the time of the survey; and average monthly income was 108 USD. The first wave of the in-survey experiment was fielded in 2021; wave 1 respondents are slightly older, but otherwise comparable. Online Appendix Table OA.1 reports additional summary statistics for both samples.

Workers in our sample were trained in 11 different occupations, which together employ over 30% of young Ugandans outside of agriculture. As shown in Appendix Table A.2, these occupations display strong gender segregation. Of the workers in our sample, 34% were trained in female-dominated fields. Using census data, we define male-dominated occupations (MDO) and female-dominated occupations (FDO) as follows: MDO include electrical work, motor mechanics, construction, plumbing, business services, machining and fitting, and welding and carpentry, while FDO include food and hospitality, tailoring, teaching (primary and pre-primary), and hairdressing.

Our sample reflects the country's emerging urban working class. Appendix Table A.2 compares our sample to the Ugandan National Household Survey (UNHS): respondents are more educated, earn 37% more than the average young worker, and are more likely to work outside agriculture (72% versus 46%). Compared to a nationally representative sample of vocational graduates, our sample is similarly likely to work outside agriculture and earns comparable income. Women in our sample are similarly likely to be employed as young women nationally (74% versus 72% reporting any work), but far more likely to be in non-agricultural jobs: 72% report non-agricultural work in the two months prior to the survey, versus 36% in nationally representative data. Women in our sample also earn more than the average wage-employed woman in Uganda, although they continue to earn less than men in our sample. Because skilled urban workers are a consequential population for the country's development, understanding how referral decisions are made in this group sheds light on mechanisms that may influence the allocation of labor across jobs and occupations.

3 Study Design and Data

3.1 Experimental Design

Our study uses an incentivized referral experiment (Kessler et al., 2019) that embeds randomized candidate profiles in a high-stakes workplace referral decision while eliciting granular measures of expected likability, perceived productivity, and employers' and candidates' second-order beliefs about fit with the job. The design is shown in Figure 1. Respondents are invited to help fill a vacancy at their workplace by referring a candidate for a six-week subsidized internship. Incumbent workers are informed that the internship placement will be allocated by lottery among eligible referrals, and that the referrer will receive a bonus of 100,000 UGX (approximately five days of wages) if the referred intern is retained after the internship.[3]

The experimental module proceeds in two parts. In Part I, respondents evaluate two candidate profiles and report outcomes for each profile, including perceived competence, likability, expected retention, and anticipated employer assessments.[4] They then choose which of the two candidates they would refer for the internship. Candidate profiles are constructed from templates used in Ugandan labor markets and include four sections: (i) personal information (name, sex, date of birth, nationality, home district), (ii) academic background, (iii) work experience (firm name, role title, duration, and main activities), and (iv) language abilities. Appendix Figure A.2 provides an example profile. Profiles are designed to be realistic and comparable along substantively relevant dimensions, reflecting young vocational graduates with similar training backgrounds. We randomize two focal attributes: candidate gender and post-training work experience. Gender is signaled through the name and the sex field. Experience is randomized in five-month increments ("low" versus "high") and cross-randomized with gender. Other profile fields that are not central to our hypotheses (such as exact date of birth within a narrow age window and high school identifiers drawn from a common quality tier) are varied to preserve realism and avoid duplicate resumes. This implies some limited deception in non-substantive details, while the core qualifications and structure of the profile match the information typically used in referral and hiring decisions in this setting. A full description of profile construction and randomization is provided in the Appendix.[5] At

[3]Respondents who are not currently employed complete the same module in hypothetical terms, allowing us to measure stated referral behavior in a comparable setting while restricting the incentivized referral to those who can credibly connect a candidate to a workplace.

[4]We randomized whether we asked about the employer's and candidate's likability (a) together with the questions on own likability and perceived probability of retention, or (b) after showing both profiles and asking respondents to choose a candidate. We do not find that the timing of the employer's likability question affected subsequent ratings of the profiles.

[5]As a manipulation check, we asked respondents at the end of the profile evaluation module to recall which candidate had more experience (and, in a subset of arms, to report the tenure associated with each profile). Respondents reliably identified the higher-experience profile, with no differential recall by

the end of Part I, respondents choose which of the two profiles to refer for the internship.

In Part II, we elicit potential referrals from respondents' own networks to measure referral choices over real contacts and assess the role of network composition. We first ask whether the respondent knows anyone who could plausibly fill the internship position. For each nominated contact, we collect a structured "network profile" that mirrors the information available for the candidates in Part I, including basic demographics (age and gender) and human capital measures (education and prior work experience). We then elicit the same evaluative outcomes as in Part I, including expected likability, perceived ability to complete the tasks, retention probability, anticipated employer likability, and expected enjoyment of the internship. If the respondent nominates a first contact of a given gender, we subsequently ask whether they can nominate an additional contact of the opposite gender. Thus, respondents evaluate two candidates we provided and, when feasible, up to two candidates from their network, yielding a final choice set of two to four candidates. In the final step, respondents make a single overall selection from this expanded choice set, while always retaining the option to stick with the initial choice they made at the end of Part I, that is, one of the two candidates we originally showed them.

This design is explicitly structured to identify both *whether* referrals are gendered and *why*. Part I provides a clean "lab-in-the-field" measure of evaluators' relative preferences over candidates: respondents make a real referral choice between two experimentally constructed candidate profiles that mirror the information typically used in this setting and are built from real profiles of actual candidates. While some non-substantive details are varied for realism, the profiles hold constant core qualifications, in particular the fact that both candidates are vocational graduates with comparable training backgrounds, and they cross-randomize candidate gender and post-training experience. This cross-randomization allows us to test whether gendered referral patterns reflect beliefs about productivity and experience premia, or instead persist even when credentials and training are comparable.

At the same time, this design moves beyond the binary outcomes typical of correspondence studies by eliciting a rich set of mechanism-relevant measures. Concretely, for each profile we elicit expected post-subsidy retention, likability, perceived competence, and second-order beliefs about whether the candidate would enjoyr the internship and whether the employer would enjoy working with them. These measures allow us to distinguish direct preferences from beliefs about employer or candidate preferences and to trace referral gaps through specific perceived attributes. Finally, by combining randomized-profile choices with referral behavior over network members in Part II, we can distinguish constraints in network composition from differential evaluation at the recommendation stage.

---

candidate gender.

3.2 Empirical Strategy

Our empirical strategy follows directly from the cross-randomized profile design. Each respondent i evaluates two candidate profiles $p \in \{1, 2\}$ for the same occupation. We estimate profile-level treatment effects using the following specification:

$$Y_{p,i} = \alpha + \beta_1 \, \mathbb{1}_p(GCS) + \beta_2 \, \mathbb{1}_p(MoreExp) + \beta_3 \Big[\mathbb{1}_p(GCS) \times \mathbb{1}_p(MoreExp) \Big] + X_i' \gamma + \varepsilon_{p,i}, \quad (1)$$

where $Y_{p,i}$ is a profile-level outcome (e.g., an indicator for being referred, ratings of likability, perceived probability of retention, or other belief measures). The indicator $\mathbb{1}_p(GCS)$ equals one if the profile is *gender counter-stereotypical* for the occupation type (a female profile in a male-dominated occupation or a male profile in a female-dominated occupation), and zero otherwise. The indicator $\mathbb{1}_p(MoreExp)$ equals one if the profile is assigned the higher experience level in the experimental design. X_i includes randomization-strata fixed effects and pre-specified respondent characteristics (including respondent gender and employment status), as well as vocational training institute and training-area fixed effects.[6] The coefficient β_1 captures the counter-stereotypical penalty, β_2 captures the experience premium for gender-stereotypical profiles, and the interaction term allows us to test whether the experience premium varies for counter-stereotypical profiles. Additionally, $\beta_2 + \beta_3$ gives us the experience premium for counter-stereotypical profiles and $\beta_1 + \beta_2 + \beta_3$ the net effect for a counter-stereotypical, more-experienced profile relative to the omitted category of a gender-stereotypical, less-experienced profile. We report these linear combinations in the tables. We estimate equation (1) separately for the full sample and by occupation type (female- and male-dominated). Standard errors are clustered at the respondent level to account for the fact that each respondent contributes two profile observations.

3.3 Outcomes

Main outcomes: probability of being referred, likability, and retention probability Our primary outcomes are the differences between gender-stereotypical (GS) and gender counter-stereotypical (GCS) candidates in three areas: the probability of being referred, their likability score, and the perceived probability of retention.[7] The first is a binary outcome, taking the value of one if the respondent chose the profile. The second outcome is measured on a Likert scale from 0 to 10 while the third, on one

[6] Table 1 reports estimates for the binary referral outcome without respondent-level controls, because the outcome is defined at the profile level and each respondent chooses exactly one of the two profiles, so respondent-level characteristics do not vary within the choice set and are therefore not separately identified.

[7] We adopt the wording of the two core evaluation questions from Heilman et al. (2004): (i) "How much would you like to work with this candidate?" and (ii) "What are the chances this candidate would be retained?".

from 0 to 100. The likability measure was designed to capture respondents' personal preferences. Given that the median business size among our respondents is five employees, they are likely to be acutely aware that any new intern would work closely with them. Given that the median business in our sample employs only five workers, respondents are likely to understand that any new intern would work closely with them. In such small firms, referring a candidate increases the likelihood of daily interaction, making likability a salient consideration with direct implications for workplace harmony. The perceived probability of retention instead captures respondents' broader assessment of the candidate's expected performance and fit for the job. This measure is also tied to real stakes: respondents receive a monetary reward of 100,000 UGX if the candidate they refer is retained after the six-week subsidized internship.

Secondary outcomes: task ability and second-order beliefs Expectations about retention may reflect factors beyond the candidate's competence. Respondents may consider the firm's overall likelihood of retaining a new employee, the employer's preferences, and the candidate's own willingness to stay. To shed light on these factors, we elicit second-order beliefs about both employers and candidates. In particular, we ask respondents how much they think the employer would enjoy having the candidate in the workplace and how much they think the candidate would enjoy the internship. Additionally, we collect more concrete assessments of expected job performance. Respondents first list the three tasks a new intern would most likely perform and then rate, on a 0 to 10 scale, how good each candidate would be at those tasks.

4 Results

4.1 Referral Gender Bias

We begin by examining our primary outcome, the probability of being referred. Table 1 reports estimates of the referral gender bias, where the dependent variable is an indicator for whether a candidate was referred. Column 1 pools all occupations, while columns 2 and 3 restrict the sample to female-dominated and male-dominated occupations, respectively. In the full sample, GCS candidates – men in female-dominated occupations (FDO) and women in male-dominated occupations (MDO) – are 25 percentage points less likely to be referred (column 1) relative to otherwise similar gender-stereotypical candidates. The penalty is smaller in female-dominated occupations (−18 p.p.) and larger in male-dominated ones (−29 pp), although the difference is not statistically significant ($p = .16$). By contrast, more experienced candidates enjoy a large and precisely estimated advantage of about 28–30 p.p. across occupations, with no evidence of differences by occupation type ($p = .73$). Interaction terms are small and statistically insignificant, indicating

that returns to additional experience accrued by counterstereotypical candidates are not statistically different from those accrued to stereotypical ones. Consistent with this, linear combinations imply that counter-stereotypical candidates with additional experience are slightly more likely to be referred than the baseline group of stereotypical, less-experienced candidates (8 pp, $p = .08$). Taken together, these results show that referral decisions strongly penalize counter-stereotypical candidates, while returns to experience are large and similar across occupations and genders.

Appendix Figure A.4 plots occupation-level estimates of the referral penalty against the share of gender counter-stereotypical workers. The relationship is upward sloping and approximately linear: the penalty is most negative in the most gender-segregated occupations and attenuates as the occupation becomes more gender-mixed (slope = 0.416, $p = 0.179$).

These results are consistent with the work of Riach and Rich (2006) and Booth and Leigh (2010), who showed that employers' gender bias correlates with the gender dominance of the occupation. However, unlike Carlsson (2011), we find a significant pro-male bias in MDO. This also expands on Beaman et al. (2018) by providing corroborating evidence of pro-male bias in MDO, while also showing a significant pro-female bias in FDO. In light of evidence on worker displacement (Crépon et al., 2013), this finding suggests that programs aimed at increasing female participation in higher-paying MDO should also consider that, if displaced into FDO, men may face similar barriers to those experienced by women in MDO.

Lastly, we do not observe different referral behavior when we focus specifically on the GCS respondents in our sample. Even among the minority groups, such as males in female-dominated occupations or females in male-dominated occupations, the same referral preferences persist—tending to disadvantage GCS candidates. However, given that GCS respondents are a small subset of the overall sample (9%), this evidence is only suggestive.

Part II then asks whether this experimental penalty persists when respondents consider real contacts from their own networks, and whether any resulting gap reflects limited network availability or differential choice conditional on access. Respondents are asked to name up to two contacts who could plausibly fill the internship position, and for each nominated contact we collect the same observables and elicit the same evaluative outcomes as for the experimental profiles. Overall, 38.2% of respondents name at least one potential candidate (Appendix Table A.3, Panel A). Conditional on naming someone ($N = 285$), network composition is highly gender-segregated: 74.4% name a male contact and 4.7% name a female contact, but only 15.1% are able to name contacts of both genders despite being explicitly prompted to do so. This constraint is particularly stark in male-dominated occupations: conditional on naming someone, 93.0% name a gender-stereotypical contact and only 22.0% name a gender counter-stereotypical contact

(Figure 2), and the GS–GCS gap is larger in male-dominated than in female-dominated occupations by 16.9 percentage points ($p = 0.065$).

Referral choices further amplify these patterns. Conditional on naming at least one contact, respondents are much more likely to refer a gender-stereotypical than a counter-stereotypical contact in both occupation types, with a substantially larger gap in male-dominated occupations (Figure 2). The implied difference-in-differences is 26.4 percentage points ($p = 0.004$). Even in the small and selected subsample that succeeds in naming contacts of both genders ($N = 43$), referral choices remain skewed in male-dominated occupations: 53.3 % refer the male contact versus 23.3 % the female contact (paired t-test $p = 0.059$), and the GS–GCS gap is larger in male-dominated than in female-dominated occupations ($p = 0.097$). Given the small sample size and the highly selected nature of this subsample, we interpret this as suggestive corroboration – rather than confirmation – that the referral gender bias operates beyond the experimental setting. Taken together, Part II delivers two complementary facts: access to potential candidates is constrained by gender-segregated networks, and conditional on access, referral choices remain gendered – especially in male-dominated occupations. [8]

5 Conceptual Framework

We develop a parsimonious framework to discipline the interpretation of referral choices in our experiment. Each respondent i chooses between two candidates $j \in \{1, 2\}$. Referring candidate j yields value

$$V_i(j) \;=\; b \cdot E_i[P(R_j)] \;+\; L_i(j) \;+\; \alpha_i(j), \tag{2}$$

where b is the retention bonus paid to the referrer if the candidate is retained ($b = 100{,}000$ UGX in the incentivized task), $E_i[P(R_j)]$ is respondent i's expectation that candidate j will be retained at the end of the six-week internship, $L_i(j)$ captures the referrer's direct utility from working with the candidate ("own likability"), and $\alpha_i(j)$ captures any additional idiosyncratic social returns to referring the candidate, such as reciprocity, favors, or other network-related benefits. In Part I of our experimental setting, $\alpha_i(j)$ is shut down by design, while in Part II of the study $\alpha_i(j)$ may be nonzero and may differ across candidates.

[8] A final caveat is that gender bias may already operate at the *naming* stage: because respondents are asked to nominate contacts they consider plausible candidates, biased beliefs about GCS workers could prevent some qualified contacts from coming to mind and being listed in the first place, implying that the observed network-composition constraint may understate the extent of bias in access. Our design partially mitigates this concern by explicitly prompting respondents, after an initial nomination, to name an additional contact of the opposite gender. While this prompt may make the study's focus more salient, it also encourages respondents to search beyond their first recall and may elicit qualified contacts who would not otherwise come to mind.

Respondent i refers candidate 1 if $V_i(1) > V_i(2)$. This decomposition highlights the two margins we can measure in the data. First, conditional on beliefs about retention, respondents may prefer to refer candidates they personally like to work with. Second, conditional on own preferences, respondents may prefer candidates they expect to be retained, either because retention generates the monetary bonus or because retention is itself valued.

We model expected retention as a function of three beliefs:

$$E_i[P(R_j)] \;=\; \phi\Big(E_i[Q_j],\, E_i[V_e(j)],\, E_i[J_j]\Big), \tag{3}$$

where Q_j denotes expected task ability (how well the candidate will perform typical intern tasks), $V_e(j)$ denotes the respondent's expectation of the employer's valuation of the candidate, and J_j denotes expected enjoyment of the internship by the candidate. The mapping $\phi(\cdot)$ is left unrestricted. Intuitively, interns who are expected to perform well, fit employer preferences, and enjoy the position should be more likely to complete the internship and to be retained afterwards.

We further interpret expected employer valuation as combining perceived productivity and employer-specific fit:

$$V_e(j) \;=\; g\Big(Q_j,\, L_e(j)\Big), \tag{4}$$

where $L_e(j)$ is the respondent's belief about how much the employer would like the candidate. In the data, we operationalize $E_i[Q_j]$ using respondents' ratings of task performance, $E_i[L_e(j)]$ using beliefs about employer likability, and $E_i[J_j]$ using beliefs about how much the candidate would enjoy the internship. Together, these measures allow us to trace referral gaps through (i) direct preferences $L_i(j)$ and (ii) beliefs that shape expected retention, including second-order beliefs about employers.

A natural additional component of referral value in workplace settings is reputational risk: a referrer may be penalized if a referred candidate underperforms. We do not include a separate reputational-cost term in equation (1) because, in vignette evidence collected in Wave 1, perceived reputational risk (when referred candidates underperform) does not differ systematically between gender-stereotypical and gender counter-stereotypical referrals. Appendix Figure A.7 reports mean perceived reputational risk (0–10 scale) for referrals to gender-stereotypical versus counter-stereotypical candidates, separately by occupation type, and shows small and statistically insignificant differences in both male- and female-dominated occupations.[9] This evidence suggests that the main channels through which gender enters referral choices in our setting operate through expected retention and interpersonal utility rather than differential reputational concerns.

[9]Appendix Figure A.7 presents paired comparisons within respondent of the perceived reputational risk to the referrer if the referred candidate underperforms. The estimated GS–GCS differences are close to zero and not statistically distinguishable from zero in either occupation type.

6 Decomposing Referral Choices

Guided by our framework, we examine data on each of these mechanisms to shed light on the underlying drivers of our main finding— a sizable referral gender bias against GCS candidates in both male- and female-dominated occupations. To understand the mechanisms underpinning our subjects' preferences, we turn to the rating questions that measure profiles' likability ("how much would you like to work with this person") and perceived probability of retention ("how likely do you think this person is to be retained"). We expect the answers to these questions to have guided subjects' referral decisions.

6.1 Likability

Table 2 reports treatment effects on two proposed mechanisms: likability (columns 1–3) and retention probability (columns 4–6). The left panel shows that counter-stereotypical candidates are rated as significantly less likable than stereotypical ones, with a pooled penalty of $-.48$ points (column 1). The gap is more than twice as large in male-dominated occupations ($-.60$, column 3) as in female-dominated ones ($-.26$, column 2), and it is statistically significant only in the former. The implied cross-occupation-type difference is sizeable ($-.34$ points) and in the expected direction, but it is not statistically significant at conventional levels ($p = .12$; Table 2, row "Counter-stereotypical (FDO–MDO)"). More experienced candidates are rated more favorably on likability, with effects of roughly half a point across both occupation types. Interaction terms are small and imprecise, suggesting that experience does not on average differentially improve the likability of counter-stereotypical candidates. However, the linear combinations in row 18 reveal a striking asymmetry. Men in female-dominated jobs can fully overcome the (already smaller) initial likability penalty if they have more experience: once experienced, they are rated more likable than stereotypical, less experienced candidates ($p = .06$). By contrast, women in male-dominated occupations start from a penalty that is more than twice as large, and additional experience does not eliminate it: even when experienced, they are not rated as more likable than less-experienced men ($p = .75$)

6.2 Perceived Retention Probability

The right panel turns to retention probability. Counter-stereotypical candidates are perceived as less likely to remain with the firm: -3.4 percentage points in the pooled sample (column 4). This penalty is concentrated in male-dominated occupations, where women are rated 3.7 p.p. less likely to stay (column 6), while in female-dominated jobs the estimate is statistically indistinguishable from zero (column 5). More experienced candidates are consistently perceived as more likely to stay, with an effect of about 5 p.p. overall and similar magnitudes across occupations. Interaction terms are again

small and statistically insignificant, indicating that the experience premium applies similarly to stereotypical and counter-stereotypical candidates on average. Yet the linear combinations once more highlight asymmetry. In female-dominated jobs, men do not face a meaningful retention penalty to begin with, and with more experience they are rated as likely to stay as their stereotypical peers. By contrast, women in male-dominated jobs face a baseline retention penalty of nearly four percentage points, and need 5 additional months of experience to restore parity. Whereas likability can be interpreted as reflecting respondents' personal taste for or against working with certain candidates, retention probability is a more composite measure. In our theoretical framework, we model the expected probability of retention as a function of the candidate's perceived ability to perform the typical tasks of the job, their enjoyment of the internship (which will determine effort, and willingness to stay) and how likable the candidate is in the eyes of the employer, who ultimately decides whether to retain the worker. Table 3 helps unpack this composite further by separately analyzing perceptions of candidates' ability to complete tasks and of how much they would enjoy the job.

Unpacking the Retention Probability: Perceived Ability to Complete Tasks The first panel of Table 3 shows that counter-stereotypical candidates are systematically judged as less able to perform the core tasks of the job. This pattern is consistent with what (Bohren et al., 2025) describe as inaccurate statistical discrimination, namely differential treatment driven by incorrect beliefs about relevant group characteristics. In our setting, this interpretation is especially plausible because all profiles had completed rigorous vocational training that should have left them well qualified for the relatively basic tasks required of interns. In the pooled sample, they are rated 1.03 points lower than stereotypical peers (column 1, $p < .001$). The penalty is large and precisely estimated in male-dominated occupations (-1.41, column 3, $p < .001$), while in female-dominated jobs it is much smaller ($-.30$, column 2), roughly five times smaller in magnitude.

More experienced candidates are perceived as more able overall ($+.53$, $p < .001$), but the asymmetry across occupation types is stark. In FDO, men with more experience not only erase their initial disadvantage but are rated significantly more able than stereotypical peers (joint effect $+.63$, $p < .01$). In MDO, however, women remain rated significantly less able even when more experienced (joint effect $-.81$, $p < .001$). The tests of equality across occupation types confirm that these differences are statistically meaningful: the counter-stereotypical penalty differs between FDO and MDO at the 1% level ($p = .000$), and the return to experience differs as well ($p = .016$). Taken together, these findings show that respondents expect men in female-typed jobs to benefit disproportionately from additional experience, while women in male-typed jobs remain systematically disadvantaged, consistent with ability perceptions being a key driver of their retention penalty.

Unpacking the Retention Probability: Perceived Candidate Enjoyment The second panel of Table 3 highlights another mechanism behind the retention penalty: respondents' expectations regarding whether candidates would enjoy the job. Research by Buchmann et al. (2025) documents what they term paternalistic discrimination: employers justify discriminatory hiring practices by assuming that certain candidates, particularly women, would not enjoy or fit into specific job roles, even when those candidates have clearly demonstrated interest and qualifications. This projection of assumed preferences onto jobseekers contributes to biased hiring decisions, as employers substitute their own beliefs about "fit" for the candidates' revealed intentions. While Buchmann et al. (2025) document this mechanism in the attitudes of employers themselves, our evidence shows that it also operates among incumbent employees making referral decisions.

In our setting, employees expect counter-stereotypical candidates – and especially women in male-dominated jobs – to derive less satisfaction from the work, despite these women having undertaken two years of specialized post-secondary training and accumulated substantial prior work experience precisely to enter these occupations.

In the pooled sample, counter-stereotypical candidates are rated $-.37$ points lower on a 0–10 enjoyment scale ($p < .001$, column 4). This gap is concentrated entirely in male-dominated occupations, where women are rated $-.48$ points lower than men ($p < .001$, column 6), while in female-dominated jobs the difference is small and statistically insignificant (column 5). More experienced candidates are perceived as enjoying the work significantly more, with an effect of $+.37$ points overall ($p = .01$, column 4) and an even larger premium of $+.67$ points in FDO ($p < .001$, column 5). Yet the asymmetry across occupation types is again stark. In MDO, women remain significantly disadvantaged even with additional experience (joint effect $-.41$, $p = .026$). By contrast, in female-dominated jobs, men with more experience erase their baseline penalty, and the point estimate suggests they may be rated as likely to enjoy the work as stereotypical peers or more – though this effect is imprecisely estimated (joint effect $+.34$, $p = .101$). The equality tests at the bottom of the table confirm that these differences across occupation types are marginally significant ($p = .058$ for the counter-stereotypical penalty).

Unpacking the Retention Probability: Employer Likability (Second-Order Beliefs) The final panel of Table 3 reports second-order beliefs about how much employers themselves would like the candidate. Counter-stereotypical profiles are expected to be rated substantially less likable by employers, with a pooled penalty of $-.54$ ($p < .001$, column 7). As with the other mechanisms, the penalty is larger in male-dominated jobs ($-.62$, $p < .001$, column 9) than in female-dominated ones ($-.37$, $p = .018$, column 8). Experience is positively valued overall ($+.48$, $p = .002$), and in both occupations employers are expected to favor more experienced candidates. However, these gains do not close the gap for women in male-dominated jobs: their joint effect remains close to zero

and statistically insignificant. By contrast, men in female-dominated jobs with more experience more than erase their baseline penalty. These patterns closely mirror the dynamics documented for retention probability in Table 2, suggesting that employees' assumptions about employer likability are a central channel behind the observed retention penalties. At the same time, our findings on ability and enjoyment indicate that retention reflects a composite of multiple mechanisms, with employer likability reinforcing rather than fully substituting for stereotypes about competence and fit. Perceived task ability shows a parallel pattern: counter-stereotypical candidates receive a pooled penalty of -1.02 points ($p = 0.000$, column 1). Once again the penalty is heavily concentrated in male-dominated occupations (-1.40, $p = 0.000$) versus female-dominated ones (-0.30, $p = 0.072$). Experience improves perceived task ability for both stereotypical and counter-stereotypical candidates, but the combined effect for counter-stereotypical profiles with more experience remains negative and sizable (-0.81, $p = 0.000$), indicating that additional experience partially but not fully offsets initial stereotyping in perceived competence for female workers in MDOs.

What do Employers Really Think? For the subset of employed respondents who provided their employers' contact information, we replicated the incentivized resume-rating exercise directly with employers, giving them the opportunity to select an intern for their company. Because participation was low (only 13 % of employers participated in the incentivized hiring decision), and likely subject to selection, these findings should be viewed as suggestive. Still, Appendix Figure A.6 reveals three notable patterns. First, employers give systematically more favorable ratings than employees expect. For example, employers' likability ratings of GS candidates sit well to the right of employees' second-order beliefs, suggesting that employees underestimate how positively employers view these candidates. Similarly, employers rate gender counter-stereotypical (GCS) candidates more positively on ability and retention than employees predict. Second, despite these higher overall ratings, employers continue to favor GS over GCS candidates across nearly all dimensions. Third, this preference becomes strongest when choices carry real consequences: in the incentivized hiring decision, employers are even less likely to select a GCS candidate than their ratings on individual attributes would suggest.

Adjectives To further investigate differences in perceived likability, we examine the adjectives respondents provided to describe candidates after reviewing their profiles. For each candidate they did not personally know, respondents were asked to list up to four adjectives based solely on the profile information, without photographs. This exercise yielded a total of 4,912 adjectives, covering all 1,506 profiles, with 724 unique words. We first standardized the text by cleaning spelling variants, translating rare terms into their closest common equivalent, and collapsing synonyms (e.g., "diligent" and "hardworking")

to ensure comparability. We then classified the resulting set of adjectives into thirteen mutually exclusive categories designed to capture distinct trait domains: (i) experience and skill, (ii) education and training, (iii) cognitive abilities, (iv) grindstone characteristics (e.g., hardworking, diligent), (v) agentic characteristics (e.g., ambitious, confident), (vi) communal characteristics (e.g., caring, cooperative), (vii) professional characteristics (e.g., punctual, organized), (viii) overall qualification, (ix, x, xi) demographic descriptors (age, physical strength, explicit gender references), (xii) negative adjectives (e.g., lazy, incompetent), and (xiii) a residual "other" category. The classification was implemented using GPT-4o, with manual verification to ensure internal consistency. The agentic and grindstone categories in particular follow the definitions developed in prior experimental work on language and labor market discrimination (Eberhardt et al., 2023; Beugnot and Peterlé, 2020). For the analysis reported below, we aggregate to the category level and use these categories as the units of comparison. This approach allows us to study whether gendered referral penalties are systematically associated with the types of traits respondents attribute to candidates, rather than with idiosyncratic word choice.

The left panel of Appendix Figure A.5 plots average marginal effects from regressions of an indicator for "candidate is female" on dummies for each adjective category (one dot per category for all candidates, triangles for more-experienced, circles for less-experienced). Positive values indicate categories more often applied to women; negative values indicate categories more often applied to men. Three patterns stand out. First, demographic descriptors are highly predictive of being female: Gender terms (e.g., "girl," "lady") and Age terms ("young") load strongly positive overall, with the Age signal especially pronounced among less-experienced candidates. Second, women are described disproportionately with Communal and Grindstone traits (e.g., caring, helpful; hardworking, diligent, persistent) echoing the reference-letter evidence that women are praised for effort and dedication rather than innate competence or leadership (Eberhardt et al., 2023). Third, several competence-coded categories are used more for men: Experience & Skill is negative overall, as are Physical Strength, Cognitive Abilities (e.g., "intelligent", "capable"), and Education & Training, with the male tilt clearest among less-experienced candidates. Notably, Negative adjectives are more likely to be applied to women in both experience bins, consistent with backlash dynamics even when women are otherwise strong candidates. Taken together, the adjectives suggest that women are primarily evaluated on character and effort, while men are evaluated on technical and physical ability – patterns that mirror occupational stereotypes and help explain the gendered gaps we document in referrals and downstream ratings.

The right panel shows the same analysis separately for FDO and MDO. The broad patterns are similar, but two differences stand out. In female-dominated jobs, negative adjectives are essentially unrelated to gender, and gendered references themselves are never used. By contrast, in male-dominated jobs, gender terms are strongly predictive

of being female, alongside Age, Negative, Communal, and Grindstone descriptors. In female-dominated jobs, professional characteristics and education/training also emerge as predictors of being female, while competence-coded categories tilt more toward men in male-dominated settings.

Mechanisms Within Networks The mechanism patterns documented in Part I are not specific to evaluations of experimental profiles. Figure 3 replicates the same set of rating questions for the real network nominees elicited in Part II and compares GS and GCS candidates within each occupation type. Across outcomes – likability, expected retention, perceived task ability, employer likability (second-order beliefs), and expected candidate enjoyment – the qualitative asymmetries mirror those in the experimental setting, with the largest differences in male-dominated occupations. This evidence suggests that the channels identified in Part I also shape evaluations of real contacts, reinforcing the interpretation that gender gaps in referrals reflect not only who is available in networks but also how potential candidates are assessed conditional on being considered. These patterns motivate the aggregate exercise that follows, where we embed the experimentally estimated referral bias in a network model to quantify its contribution to occupational segregation.

7 Aggregate Implications: A Model of Gendered Referrals

How sizable is the referral gender bias in furthering gender segregation? To answer this question, we consider a variation of the model of Calvo-Armengol and Jackson (2004).

Setup We assume there is a network of N workers. Individuals in the network can be men or women and can have one of three states: unemployed, employed in occupation 1, and employed in occupation 2. To fix ideas, the reader may consider that occupation 1 is the female-dominated lower-paying occupation and occupation 2 is the male-dominated higher-paying occupation. We denote the state of the economy at time t by a set of binary vectors $e_t^{(0)}, e_t^{(1)}, e_t^{(2)} \in \{0,1\}^N$, where the i-th element of $e_t^{(s)}$ equals 1 if individual i is in state $s \in \{0,1,2\}$, and 0 otherwise.

In every period, individuals engage in job search independently of their current employment status. This can be understood either as direct job search if unemployed, or more broadly as capturing the fact that job opportunities may reach individuals even when they are not actively looking for a job themselves, for example through their employer or through others in their work network who are looking to hire someone new for the shop.

Workers find a job in occupation 1 with probability $\sigma^{(1)}$ and in occupation 2 with probability $\sigma^{(2)}$. If they are unemployed, they will start the new job. If they are unemployed and find a job in both occupations, they will choose one of them randomly, with equal probability. If they are employed, they will refer (forward the job to) someone in their network: at each period, each node may forward a signal to at most one other neighboring node. At the end of each period, jobs are destroyed with probability δ. This process is repeated for T periods.

We incorporate segregation in three ways. First, we assume that men and women have different rates of success in finding a job in occupations 1 and 2, $\sigma_m^{(1)}$, $\sigma_f^{(1)}$, $\sigma_m^{(2)}$, $\sigma_f^{(2)}$. In particular, we assume $\sigma_m^{(2)}/\sigma_m^{(1)} > \sigma_f^{(2)}/\sigma_f^{(1)}$, – i.e., men have relatively better access to jobs in occupation 2 compared to occupation 1, and hence occupation 2 is more male-dominated than occupation 1, even if no jobs are filled through referrals. This reflects the fact that, beyond referrals, GCS candidates face other structural barriers to entering jobs that are nonconforming to traditional gender expectations. These barriers arise on the supply side, through self-selection shaped by preferences, anticipated workplace hostility, and social norms, and on the demand side, through employer screening and discrimination. A large literature shows that such forces help sustain occupational segregation, particularly when women seek entry into male-dominated and often higher-paying roles (Firth, 1982; Riach and Rich, 2006; Booth and Leigh, 2010; Carlsson, 2011; Arceo-Gomez and Campos-Vazquez, 2014; Galarza and Yamada, 2017; Neumark et al., 2019; Campos-Vazquez and Gonzalez, 2020; Bedi et al., 2022). Second, social ties are not formed at random: individuals are more likely to connect with others who resemble them, a phenomenon known as homophily (McPherson et al., 2001). Gender is one important dimension along which homophily arises among individuals (Ibarra, 1992) in general, and workers in particular (Topa, 2011). We can assume a random network comprising the N individuals such that the probability of link formation between two workers of the same gender is p and between workers of different genders is $q < p$. In our setup, we implement this by generating a random graph through a stochastic block model (Holland et al., 1983). The graph comprises $N = 100$ individuals, half of whom are women. Within each gender (block), the probability of link formation is p, while across blocks it is $q < p$. With network homophily, unemployed nodes are more likely to receive referrals from occupations where nodes of the same gender are occupied.

Third, we model the probability of individual i forwarding a signal of occupation $s \in \{1, 2\}$ to a neighbor j is $\theta_{i,j}^{(s)} = \frac{\alpha_{i,j}^{(s)}}{\sum^K \alpha_{i,k}^{(s)}}$, where $\alpha_{i,j}^{(s)}$ as an arbitrary weight associated with the link between i and j for the signal of occupation s and K is the number of neighbors (the degree) of individual i. The larger $\alpha_{i,j}^{(s)}$, the more likely individual i is to forward a signal s to neighbor j – and the more friends an individual has, the less likely any given friend is to learn about the job. In light of our evidence of a referral gap against GCS candidates, we let $\alpha_{i,j}^{(s)}$ depend on the occupation s and the gender of

neighbor j. In the baseline specification, we impose within-gender symmetry, so that for a given occupation s, all male neighbors of i receive the same weight and all female neighbors receive the same weight: $\alpha_{i,j}^{(s)} = \alpha_{i,m}^{(s)}$ if j is a man, and $\alpha_{i,j}^{(s)} = \alpha_{i,f}^{(s)}$ if j is a woman.[10]

We further assume that these gender-specific weights do not depend on the gender of the sender, so that $\alpha_{i,m}^{(s)} = \alpha_m^{(s)}$ and $\alpha_{i,f}^{(s)} = \alpha_f^{(s)}$ for all i.[11] It then follows that referral behavior can be summarized by the four parameters $\alpha_m^{(1)}$, $\alpha_f^{(1)}$, $\alpha_m^{(2)}$, and $\alpha_f^{(2)}$. Any heterogeneity across individuals in the implied referral probabilities comes from differences in the size and gender composition of their realized network, rather than from sender-specific referral weights.[12] It then follows that referral behavior can be summarized by the four parameters $\alpha_m^{(1)}$, $\alpha_f^{(1)}$, $\alpha_m^{(2)}$, and $\alpha_f^{(2)}$. Any heterogeneity across individuals in the implied referral probabilities comes from differences in the size and gender composition of their realized neighborhoods, rather than from sender-specific referral weights. Finally, because each individual can forward a signal to at most one neighbor, the probabilities $\theta_{i,j}^{(s)}$ sum to one over $j \in \mathcal{N}(i)$ whenever i is not isolated.

Referral bias reinforces segregation when these weights favor gender-stereotypical referrals, that is, when $\alpha_m^{(1)} < \alpha_f^{(1)}$ and $\alpha_f^{(2)} < \alpha_m^{(2)}$. Under these inequalities, signals for occupation 1 are more likely to be forwarded to women, while signals for occupation 2 are more likely to be forwarded to men.

Dynamics The process at each period t unfolds as follows. Each individual i of gender g independently receives a first-hand signal from occupation $k \in \{1, 2\}$ with probability $\sigma_g^{(k)}$, with realizations denoted $s_{t,i}^{(k)} \sim \text{Bernoulli}(\sigma_g^{(k)})$. Employed individuals who receive such a signal forward it to exactly one neighbor, drawn via a multinomial:

$$\phi_i^{(k)} \sim \text{Multinomial}\Big(1,\ \big(\theta_{i,j}^{(k)}\big)_{j \in \mathcal{N}(i)}\Big), \tag{5}$$

where $\theta_{i,j}^{(k)}$ is as defined above. Node j receives a second-hand signal (referral) from occupation k if at least one employed neighbor forwards to them: $\tilde{s}_{t,j}^{(k)} = \mathbb{1}\big(\sum_i \phi_{i \to j}^{(k)} > 0\big)$. Unemployed nodes that receive an offer – either first-hand ($s_{t,i}^{(k)}$) or via referral when no first-hand offer arrived – transition into employment in the corresponding occupation; simultaneous offers from both occupations are resolved by a fair coin flip. Referrals

[10] This restriction is imposed for tractability and can be relaxed by allowing $\alpha_{i,j}^{(s)} = \alpha_{g_j}^{(s)} \lambda_{i,j}^{(s)}$, where $\lambda_{i,j}^{(s)}$ captures tie-specific heterogeneity, such as tie strength, salience, or perceived quality. Under a normalization that preserves the total referral weight assigned to male and female neighbors, this extension leaves the direction and interpretation of the gender-specific referral bias unchanged, while potentially affecting its quantitative impact by making referrals more or less concentrated across contacts. We do not pursue this extension quantitatively here.

[11] This is consistent with our suggestive evidence that women in male-dominated occupations and men in female-dominated occupations display broadly similar referral patterns within occupation.

[12] As a robustness exercise reported in the appendix, we extend the model to allow network size and composition to vary by gender and examine how this richer network structure affects the results.

(second-hand signals) cannot be forwarded further. Employed nodes keep their job unless hit by job destruction $d_i \sim \text{Bernoulli}(\delta)$. Denoting by $a_i^{(k)}$ the event that node i receives a net first-hand offer from occupation k (ties broken randomly), and by $b_i^{(k)}$ the analogous event for referrals (see the Online Appendix for precise definitions), employment states evolve as:

$$e_{t,i}^{(0)} = e_{t-1,i}^{(0)} \cdot a_i^{(0)} \cdot b_i^{(0)} \; + \; d_i \cdot \left(e_{t-1,i}^{(1)} + e_{t-1,i}^{(2)}\right) \tag{6}$$

$$e_{t,i}^{(k)} = e_{t-1,i}^{(k)}(1 - d_i) \; + \; e_{t-1,i}^{(0)} \cdot \left(a_i^{(k)} + a_i^{(0)} \cdot b_i^{(k)}\right), \quad k \in \{1, 2\} \tag{7}$$

The steady state is characterized by the fixed point $\mathbb{E}[e_{t,i}^{(k)}] = \mathbb{E}[e_{t-1,i}^{(k)}]$ for all i and k. Because second-hand signals depend non-linearly on the joint forwarding decisions of all neighbors, no closed-form solution exists; we therefore approximate the steady state via Monte Carlo simulation (see the Online Appendix for details).

Parameters We retrieve $\alpha_f^{(1)}$, $\alpha_m^{(1)}$, $\alpha_f^{(2)}$, $\alpha_m^{(2)}$ from the moments of our experiment (i.e. the coefficient for the GCS candidate in the regression of the referral decision). We retrieve $\alpha_f^{(1)}$, $\alpha_m^{(1)}$, $\alpha_f^{(2)}$, and $\alpha_m^{(2)}$ from moments in our experiment, specifically from the coefficient on the GCS candidate in the regression of the referral decision. We again rely on the rich data collected as part of our study. In particular, we elicited information on respondents' networks by asking them to name " people that [they] expected could help [them] in the initial phases of [their] career." We use these data to inform both the degree of each node and the level of homophily in job networks, that is, the extent to which workers' choice sets are constrained by the availability of counter-stereotypical contacts. For this sample, '. Our respondents had on average 5.0 connections, of which 64% were of the same gender as the respondent. As one extension reported in the appendix, we allow network composition and degree to differ by gender and examine how this richer network structure affects our results.

We obtain the values for $\sigma_m^{(1)}$, $\sigma_m^{(2)}$, $\sigma_f^{(1)}$, $\sigma_f^{(2)}$ by inverting our model to match the empirical employment distributions observed in our data. Formally, we employ a simulated method of moments targeting four specific moments: the share of female workers employed in occupation 1, the share of female workers in occupation 2, and the corresponding shares for male workers. Because the model lacks a closed-form solution for the steady-state distribution of employment, for the purpose of the SMM, we simulate the economy for $T = 350$ periods in 300 Monte Carlo simulations of networks with $N = 100$ to approximate the steady state, which is computed by averaging the states in the last 50 periods, across all realized networks. To simulate the steady state after having inverting the parameters, we run 2,000 Monte Carlo simulations over $T = 1,500$ for networks with $N = 100$, averaging the last 50 periods to approximate the steady state. We then estimate the vector of signal parameters by minimizing the squared distance between

the simulated employment shares and the empirical moments. Since we estimate four parameters using four moments, our system is exactly identified.

Steady state To quantify the contribution of the referral gender gap to occupational segregation, we compare the steady state of the model across economies that selectively shut down the three sources of segregation embedded in the model: differential direct access to jobs ("S"), gendered network composition through homophily ("C"), and gender-biased referrals ("RGB"). Formally, "S" corresponds to allowing $\sigma_m^{(s)} \neq \sigma_f^{(s)}$ across occupations, "C" corresponds to homophilous networks with $p > q$, and "RGB" corresponds to referral weights that favor gender-stereotypical candidates, that is, $\alpha_f^{(1)} > \alpha_m^{(1)}$ and $\alpha_m^{(2)} > \alpha_f^{(2)}$. Figure 4 reports the resulting steady-state gender gaps in occupation 1, the low-paying female-dominated occupation, and occupation 2, the high-paying male-dominated occupation.

The figure first shows that when all three forces are shut down ("None"), so that $\sigma_m^{(s)} = \sigma_f^{(s)}$, $p = q$, and $\alpha_f^{(1)} = \alpha_m^{(1)}$ and $\alpha_m^{(2)} = \alpha_f^{(2)}$, the model generates essentially no segregation in steady state. Adding the channels one at a time clarifies their distinct roles. Differential direct job access ("S"), captured by gender differences in $\sigma_g^{(s)}$, is the main source of segregation in the baseline calibration: once introduced on its own, it generates large gender gaps in both occupations. By contrast, network composition alone ("C"), captured by $p > q$, generates essentially no segregation, because in the absence of either unequal direct job access or biased referrals, homophilous ties have little to amplify. Referral bias alone ("RGB"), captured by $\alpha_f^{(1)} > \alpha_m^{(1)}$ and $\alpha_m^{(2)} > \alpha_f^{(2)}$, already generates a non-trivial amount of segregation, because it directly steers job information toward gender-stereotypical candidates.

The figure also shows that network composition and referral bias operate differently once other frictions are present. Holding referral behavior unbiased, adding network composition to the direct-signal economy increases segregation modestly by amplifying the unequal allocation of job information already created by "S" (comparing "S" and "S+C"). Adding referral bias, instead, raises segregation both with and without network composition (comparing "S" to "S+RGB" and "S+C" to "S+C+RGB"). Our main counterfactual compares the benchmark economy with direct job access and network composition only ("S+C") to the economy that additionally incorporates the experimentally estimated referral bias ("S+C+RGB"). Under this comparison, our simulations indicate that the referral bias increases steady-state gender occupational segregation by 9.2 percent in male-dominated occupations and by 8.2 percent in female-dominated occupations, corresponding to roughly 6 and 5 percentage points, respectively.

In the baseline calibration, most segregation is loaded onto "S", which is estimated residually to match observed employment patterns and therefore absorbs a wide range of barriers outside referrals. But once referral bias is introduced into this environment, it

further widens occupational gender gaps by systematically redirecting job information toward stereotypical candidates. In this sense, referral bias is not the main source of segregation in levels, but it is an economically meaningful force that materially amplifies segregation in equilibrium.

Extensions The appendix shows that the contribution of referral bias is larger in environments where referrals play a greater allocative role. In particular, it rises in denser networks, when direct job access is less segregated, and when networks are less homophilous. We also extend the model to allow network size and composition to differ by gender and examine how this richer network structure affects the results. Under alternative calibrations based on employer-side estimates or survey-based job arrival rates, the contribution of referral bias becomes substantially larger than in the baseline calibration.

8 Conclusion

This paper takes seriously a basic but unresolved question about gendered referrals: are gender gaps in referrals primarily a supply phenomenon, driven by the gender segregation of workers' networks, or do they also reflect a distinct *referral gender bias* – the way incumbent workers evaluate and choose among potential candidates conditional on having access to them? We study this question in a high-stakes referral setting in urban Uganda, where personal connections are a dominant route into early-career jobs and where skilled occupations remain sharply gender-segregated.

Our design combines two complementary components. In Part I, respondents make an incentivized referral choice between two realistic candidate profiles whose gender and post-training experience are cross-randomized, holding constant core qualifications – including the fact that both candidates are vocational graduates with comparable training backgrounds. By construction, this task shuts down the network composition channel and isolates differential evaluation at the recommendation stage. In Part II, respondents nominate real contacts from their own networks and make referral choices over these nominees, allowing us to directly measure network availability and examine whether the patterns from Part I persist when choices involve real people the respondent knows.

The evidence points to a clear conclusion. GCS candidates are substantially less likely to be referred in the experimental setting of Part I, and this penalty is closely linked to the degree of occupational gender segregation. In the more lifelike setting of Part II, networks are indeed highly gender-segregated, constraining access to cross-gender candidates. But network composition may not be the whole story. Even among the small subset of respondents who can name contacts of both genders referral choices are tilted toward gender-stereotypical candidates in male-dominated occupations (53.3 %

vs. 23.3 %, $p = 0.059$), though the difference is not statistically significant in the pooled sample. This suggestive evidence is consistent with a referral gender bias operating beyond network constraints, but the small and selected nature of the subsample means that these results should be interpreted with caution.

We use the richness of the design to shed light on why these referral penalties arise. Moving beyond a binary referral outcome, we elicit a battery of mechanism-relevant measures that map naturally into the referrer's objective in small-firm settings. Respondents evaluate candidates on their own likability of working with them, on expected retention (directly tied to the retention bonus), and on the beliefs that underpin retention expectations – including perceived task ability, beliefs about employer likability, and beliefs about whether the candidate would enjoy the internship. These measures reveal a coherent set of gendered beliefs that jointly depress the expected value of referring GCS candidates. Crucially, the penalties are asymmetric across occupation types. For women entering MDO, additional experience does comparatively little to undo negative beliefs about likability, task ability, and candidate enjoyment. In contrast, for men entering FDO, additional experience more often translates into improved evaluations along these same margins. This asymmetry is particularly striking given how selected and credentialed the candidate pool is: candidates have already invested in specialized vocational training and accumulated post-training work experience, yet female candidates in MDO continue to be evaluated as if their credentials convey less information.

Two further pieces of evidence reinforce this interpretation. First, free-text adjective descriptions show that the same observable qualifications are processed through different evaluative lenses depending on candidate gender and occupational context – with women described more in terms of character and effort and men more in terms of competence-coded traits. Second, a natural concern is that employees might simply be transmitting discrimination they anticipate from employers. Our survey with employers suggests otherwise: employers exhibit a preference for GS candidates, but the bias is smaller than employees anticipate and smaller than the bias embedded in employees' own referral choices and beliefs. While we interpret this evidence cautiously given the small and selected sample of participating employers, it points toward an important implication: incumbent workers are not merely conduits of employer discrimination. Their own stereotypes and expectations can amplify gender segregation by shaping who enters referral-based hiring pipelines in the first place.

Finally, we quantify the aggregate implications of the referral gender bias we document. We embed the experimentally estimated referral bias in a dynamic network model of job finding in which workers receive job signals directly and through referrals, with referral probabilities depending on gendered forwarding weights. The model separates baseline segregation – driven by differential job arrival rates and homophilous networks – from the additional segregation generated by gender-biased referral weights. Under a conservative

baseline calibration, introducing the RGB increases steady-state occupational segregation by about 8.2–9.2 percent. Under alternative calibrations that assign a larger role to referrals in job finding, the contribution of the referral bias rises substantially and can account for a large share of observed occupational sorting. These exercises underscore the paper's main message: even when network segregation is severe, gendered evaluation and choice at the recommendation stage can be quantitatively important in sustaining occupational segregation.

Taken together, our findings imply that policies focused solely on expanding women's access to MDO – through scholarships, training slots, or recruitment – may be necessary but not sufficient if referral pipelines remain a central entry point into employment. When referral decisions embed systematic penalties for GCS candidates, reducing segregation likely also requires shifting beliefs about fit, competence, and workplace interactions among incumbent workers, not only among employers. More broadly, the paper highlights that understanding who enters hiring funnels, and why, requires taking seriously the behavior of intermediaries. In many labor markets, incumbent workers are among the most important of these intermediaries.

Main Tables and Figures

Table 1: Referral Gender Bias

	Referred		
	Full Sample (1)	Female Dominated Occupations (2)	Male Dominated Occupations (3)
Counter-stereotypical	-.250***	-.178***	-.288***
	(.037)	(.064)	(.045)
More Exp.	.285***	.301***	.276***
	(.035)	(.062)	(.042)
Counter-stereotypical × More Exp.	.053	-.002	.082
	(.045)	(.079)	(.055)
Mean in control	.549	.515	.566
More Exp. + Counter-stereotypical × More Exp.	.338***	.300***	.358***
	[.000]	[.000]	[.000]
More Exp. + Counter-stereotypical × More Exp. + Counter-stereotypical	.088*	.121	.070
	[.079]	[.154]	[.257]
Counter-stereotypical (FDO-MDO)		-.109	
		[.164]	
More Exp. (FDO-MDO)		-.026	
		[.730]	
Counter-stereotypical × More Exp. (FDO-MDO)		.084	
		[.382]	
Controls	No	No	No
Observations	1494	512	982

Notes: Each column reports estimates from regressions of an indicator for referring a candidate (=1 Referred) on indicators for counter-stereotypical candidates, more-experienced candidates, and their interaction. Columns correspond to the full sample, female-dominated occupations, and male-dominated occupations. Rows with linear combinations report sums or differences of coefficients; p-values are in brackets on the line below the estimates. Standard errors are clustered at the trainee level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 2: Likability and Retention Probability

	Likability			Retention Probability		
	Full Sample (1)	Female Dominated Occupations (2)	Male Dominated Occupations (3)	Full Sample (4)	Female Dominated Occupations (5)	Male Dominated Occupations (6)
Counter-stereotypical	-.481***	-.259	-.597***	-.034***	-.028	-.037***
	(.106)	(.173)	(.132)	(.010)	(.017)	(.012)
More Exp.	.501***	.554**	.471**	.051***	.041	.056***
	(.149)	(.250)	(.184)	(.015)	(.026)	(.018)
Counter-stereotypical × More Exp.	.106	.181	.063	.010	.037	-.005
	(.223)	(.370)	(.277)	(.023)	(.038)	(.030)
Mean in control	7.000	6.862	7.073	.659	.661	.658
More Exp. + Counter-stereotypical × More Exp.	.607***	.735***	.534***	.061***	.078***	.051**
	[.000]	[.005]	[.007]	[.000]	[.003]	[.013]
More Exp. + Counter-stereotypical × More Exp. + Counter-stereotypical	.126	.476*	-.063	.026*	.050*	.014
	[.419]	[.064]	[.748]	[.087]	[.063]	[.466]
Counter-stereotypical (FDO-MDO)			-.338			-.009
			[.121]			[.662]
More Exp. (FDO-MDO)			-.083			.015
			[.790]			[.631]
Counter-stereotypical × More Exp. (FDO-MDO)			-.118			-.042
			[.798]			[.383]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1494	512	982	1493	512	981

Notes: Columns 1–3 report regressions of likability scores on indicators for counter-stereotypical candidates, more-experienced candidates, and their interaction. Columns 4–6 report analogous regressions for the probability the candidate is retained. Within each panel, columns correspond to the full sample, female-dominated occupations, and male-dominated occupations. Rows with linear combinations report sums of coefficients; bracketed values on the line below are p-values. Rows labeled (2–3; 5–6) report differences in coefficients between female- and male-dominated occupations for the likability and retention outcomes, respectively. Standard errors are clustered at the trainee level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 3: Task Ability, Perceived Enjoyment, and Employer Likability

	Perceived Candidate Ability to Complete Tasks			Perceived Candidate Enjoyment			Employer's Likability		
	Full Sample (1)	Female Dominated Occupations (2)	Male Dominated Occupations (3)	Full Sample (4)	Female Dominated Occupations (5)	Male Dominated Occupations (6)	Full Sample (7)	Female Dominated Occupations (8)	Male Dominated Occupations (9)
Counter-stereotypical	-1.025***	-.302*	-1.405***	-.367***	-.145	-.484***	-.535***	-.367**	-.621***
	(.107)	(.167)	(.134)	(.086)	(.141)	(.108)	(.099)	(.157)	(.125)
More Exp.	.554***	1.042***	.295	.370**	.665***	.218	.482***	.493*	.470***
	(.157)	(.240)	(.199)	(.150)	(.249)	(.188)	(.150)	(.277)	(.179)
Counter-stereotypical × More Exp.	.162	-.112	.303	-.152	-.183	-.141	.367	.534	.283
	(.235)	(.367)	(.297)	(.231)	(.371)	(.292)	(.233)	(.390)	(.291)
Mean in control	7.662	7.541	7.726	8.023	7.789	8.147	6.685	6.792	6.629
More Exp. + Counter-stereotypical × More Exp.	.716***	.930***	.598***	.218	.483**	.077	.849***	1.027***	.753***
	[.000]	[.000]	[.005]	[.168]	[.047]	[.705]	[.000]	[.000]	[.000]
More Exp. + Counter-stereotypical × More Exp. + Counter-stereotypical	-.309*	.628***	-.806***	-.149	.337	-.407**	.314**	.660***	.132
	[.054]	[.006]	[.000]	[.286]	[.101]	[.026]	[.033]	[.002]	[.493]
Counter-stereotypical (FDO-MDO)			-1.103***			-.339*			-.255
			[.000]			[.058]			[.205]
More Exp. (FDO-MDO)			-.747**			-.447			-.023
			[.016]			[.152]			[.945]
Counter-stereotypical × More Exp. (FDO-MDO)			.415			.041			-.251
			[.378]			[.931]			[.606]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1372	468	904	1369	467	902	1364	461	903

Notes: Columns 1–3 report regressions of perceived candidate ability to complete tasks on indicators for counter-stereotypical candidates, more-experienced candidates, and their interaction. Columns 4–6 report analogous regressions for perceived candidate enjoyment, and columns 7–9 for the employer's own likability rating of the candidate. Within each panel, columns correspond to all occupations, female-dominated occupations, and male-dominated occupations, respectively. Rows with linear combinations report sums of coefficients; bracketed values on the line below are p-values. Rows labeled (2–3, 5–6, 8.2–9.2) report differences between female- and male-dominated occupations within each panel. Standard errors are clustered at the trainee level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Figure 1: Experimental Design

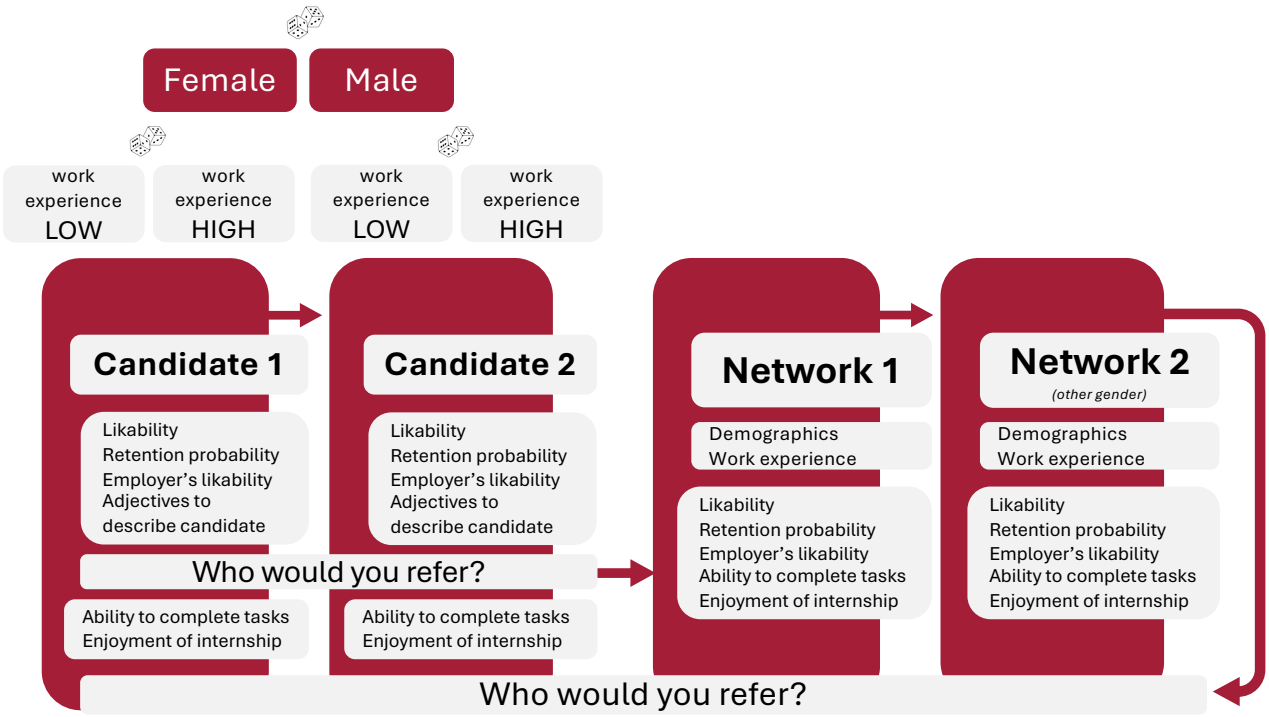
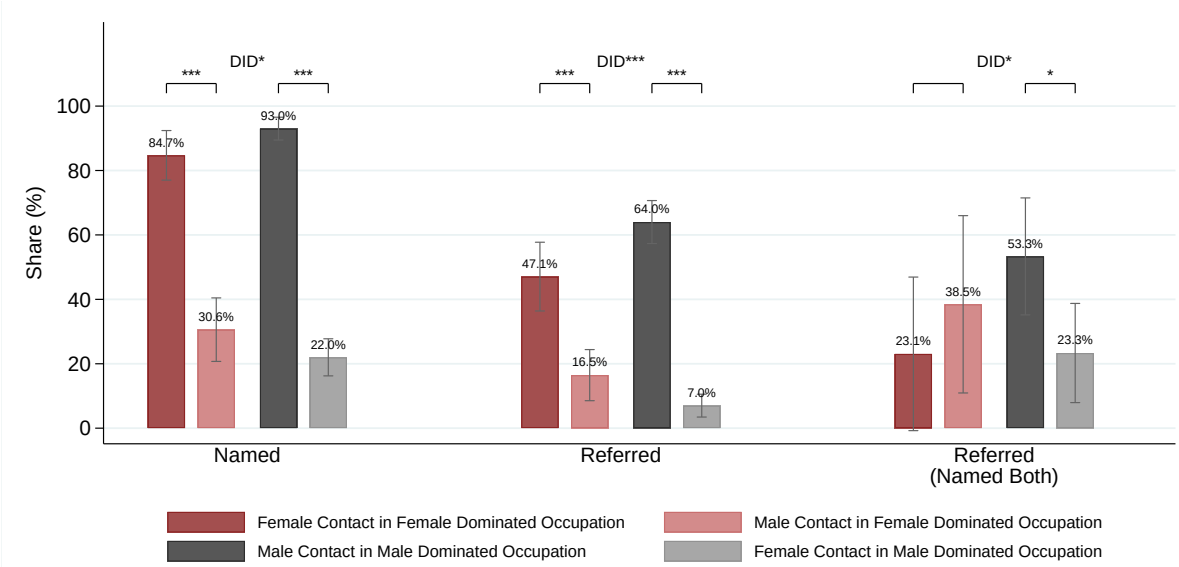
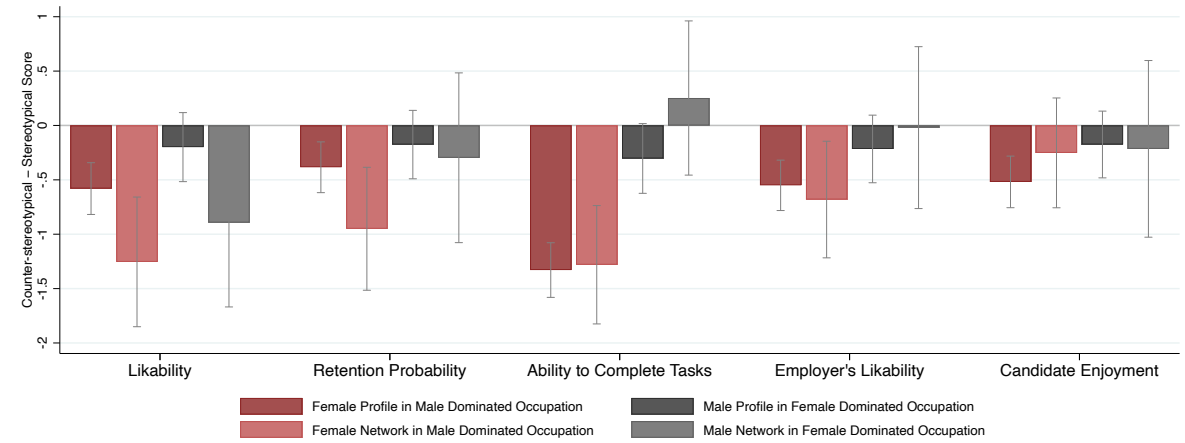


Figure 2: Network Composition: Named and Referred Contacts



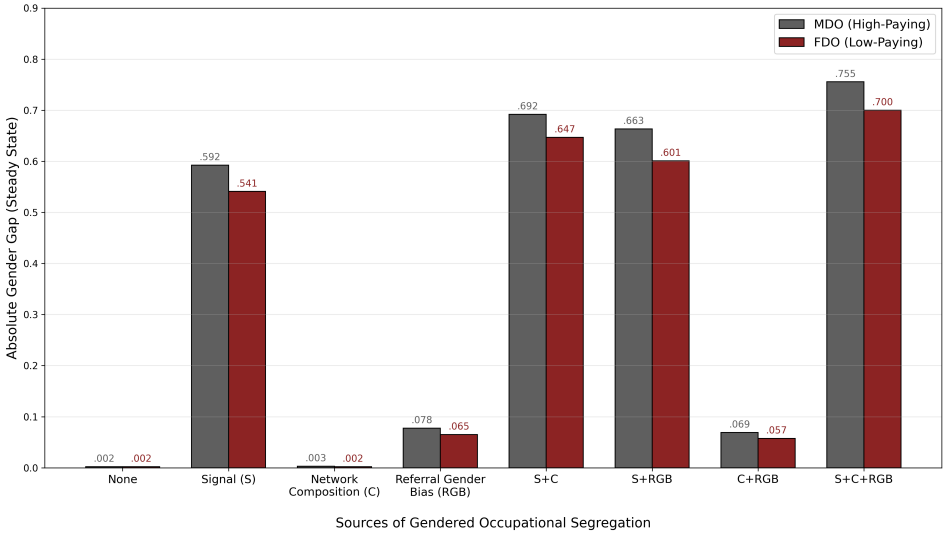
Notes: The three column groups show the share of workers who (i) *named* a gender-stereotypical (GS) or gender-counter-stereotypical (GCS) contact, conditioning on naming at least one contact ($N = 285$); (ii) *referred* a GS or GCS contact, conditioning on naming at least one contact ($N = 285$); and (iii) *referred* a GS or GCS contact, conditioning on having named contacts of both genders ($N = 43$). Whiskers denote 95% confidence intervals. Brackets indicate within-occupation-type differences between GS and GCS contacts (paired t-test); stars above each bracket show significance. "DID" reports the difference-in-differences: the GS–GCS gap in MDO minus the GS–GCS gap in FDO, estimated as the OLS coefficient on an MDO indicator in a regression of the individual-level GS–GCS gap. DID estimates: 16.9 pp (Named), 26.4 pp (Referred), and 45.4 pp (Referred, Named Both). $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Figure 3: Counter-Stereotypical Penalty Across Outcomes



Notes: The figure plots differences in outcomes between counter-stereotypical and stereotypical candidates within each occupation type. Bars show estimated gaps in likability, retention probability, ability to complete tasks, employer's likability, and candidate enjoyment. Red bars correspond to female profiles and networks in MDOs; grey bars correspond to male profiles and networks in FDOs. Vertical lines denote 95% confidence intervals. Negative values indicate the degree to which stereotypical candidates are evaluated more favorably than counter-stereotypical ones on the given outcome, by occupation type.

Figure 4: Steady State Gender Occupational Segregation



Note: Gender gap is defined as the difference between the share of men and the share of women in an occupation. The figure plots the average gender gap across 2,000 Monte Carlo simulations and $T = 1,500$. Steady states are approximated by averaging the last 50 periods. Each pair of columns represents the steady state gender gap under different combinations of sources of segregation. "S" stands for differential job arrival rates (signals). "C" stands for network composition (homophily). "RGB" stands for referral gender bias. Parameters for "S+C+RGB" were: $N = 100$, $p = .064$, $q = .035$, $\delta = .03$, $\sigma_m^{(1)} = .0072$, $\sigma_m^{(2)} = .1536$, $\sigma_f^{(1)} = .1792$, $\sigma_f^{(2)} = .0072$, $\alpha_S^{(1)} = .356$, $\alpha_{NS}^{(1)} = .644$, $\alpha_S^{(2)} = .589$, $\alpha_{NS}^{(2)} = .411$. When "S" is shut off, $\sigma_m^{(1)} = \sigma_f^{(1)} = .0932$, $\sigma_m^{(2)} = \sigma_f^{(2)} = .0804$. When "C" is shut off, $p = .0509$, $q = .0498$. When "RGB" is shut off, $\alpha_S^{(s)} = \alpha_{NS}^{(s)} = 1$. **??** plots the evolution of segregation across periods, under different scenarios.

References

Adda, J., Dustmann, C., and Stevens, K. (2017). The Career Costs of Children. *Journal of Political Economy*, 125(2):293–337.

Agüero, J. M. and Marks, M. S. (2011). Motherhood and Female Labor Supply in the Developing World: Evidence from Infertility Shocks. *Journal of Human Resources*, 46(4):800–826.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., and Vitali, A. (2020). Tackling Youth Unemployment: Evidence from a Labor Market Experiment in UGanda. *Econometrica*, 88(6):2369–2414.

Alfonsi, L., Namubiru, M., and Spaziani, S. (2023). Gender Gaps: Back and Here to Stay? Evidence from Skilled Ugandan Workers during COVID-19. *Review of Economics of the Household*, 22(3):999–1046.

Arceo-Gomez, E. O. and Campos-Vazquez, R. M. (2014). Race and Marriage in the Labor Market: A Discrimination Correspondence Study in a Developing Country. *American Economic Review*, 104(5):376–80.

Asongu, S. and Odhiambo, N. (2019). Inequality and the Economic Participation of Women in Sub-SAharan AFrica: An Empirical Investigation. *African Journal of Economic and Management Studies*, 11(2):193–206.

Azmat, G. and Petrongolo, B. (2014). Gender and the Labor Market: What Have We Learned from Field and Lab Experiments? *Labour Economics*, 30:32–40.

Beaman, L., Keleher, N., and Magruder, J. (2018). Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor Economics*, 36(1):121–157.

Beaman, L. and Magruder, J. (2012). Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review*, 102(7):3574–93.

Bedi, A., Majilla, T., and Rieger, M. (2022). Does Signaling Childcare Support on Job Applications Reduce the Motherhood Penalty? *Review of Economics of the Household*, 20(2):373–387.

Bertrand, M. (2011). New Perspectives on Gender. In *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier.

Bertrand, M. and Duflo, E. (2017). Field Experiments on Discrimination. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Economic Field Experiments*, volume 1, pages 309–393. North-Holland.

Beugnot, J. and Peterlé, E. (2020). Gender Bias in Job Referrals: An Experimental Test. *Journal of Economic Psychology*, 76:102209.

Bjerk, D. and Han, S. (2007). Assortative Marriage and the Effects of Government Homecare Subsidy Programs on Gender Wage and Participation Inequality. *Journal of Public Economics*, 91(5–6):1135–1150.

Bohren, J. A., Hull, P., and Imas, A. (2025). Systemic Discrimination: Theory and Measurement. *The Quarterly Journal of Economics*, 140(3):1743–1799.

Booth, A. and Leigh, A. (2010). Do Employers Discriminate by Gender? A Field Experiment in Female-Dominated Occupations. *Economics Letters*, 107(2):236–238.

Buchmann, N., Sullivan, C., and Meyer, C. (2025). Paternalism and Discrimination. Working Paper, Revised and Resubmitted to *Econometrica*.

Calvo-Armengol, A. and Jackson, M. O. (2004). The Effects of Social Networks on Employment and Inequality. *American economic review*, 94(3):426–454.

Campos-Vazquez, R. M. and Gonzalez, E. (2020). Obesity and Hiring Discrimination. *Economics & Human Biology*, 37:100850.

Carlsson, M. (2011). Does Hiring Discrimination Cause Gender Segregation in the Swedish Labor Market? *Feminist Economics*, 17(3):71–102.

Cortés, P., Hwang, J., Pan, J., and Schönberg, U. (2026). Gender Norms and the Labor Market. Working Paper 34716, National Bureau of Economic Research. Forthcoming in Econometric Society Monographs.

Cortes, P. and Pan, J. (2018). Occupation and Gender. In Averett, S. L., Argys, L. M., and Hoffman, S. D., editors, *The Oxford Handbook of Women and the Economy*, pages 425–452. Oxford University Press.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2013). Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment. *Quarterly Journal of Economics*, 128(2):531–580.

Cristia, J. P. (2008). The Effect of a First Child on Female Labor Supply: Evidence from Women Seeking Fertility Services. *Journal of Human Resources*, 43(3):487–510.

Delfino, A. (2024). Breaking Gender Barriers: Experimental Evidence on Men in Pink-Collar Jobs. *American Economic Review*, 114(6):1816–1853.

Dinkelman, T. (2011). The Effects of Rural Electrification on Employment: New Evidence from South Africa. *American Economic Review*, 101(7):3078–3108.

Eberhardt, M., Facchini, G., and Rueda, V. (2023). Gender Differences in Reference Letters: Evidence from the EConomics Job Market. *The Economic Journal*, 133(655):2676–2698.

Fafchamps, M. and Moradi, A. (2015). Referral and Job Performance: Evidence from the Ghana Colonial Army. *Economic Development and Cultural Change*, 63(4):715–751.

Fernández, R. and Fogli, A. (2009). Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177.

Fernández, R., Fogli, A., and Olivetti, C. (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *The Quarterly Journal of Economics*, 119(4):1249–1299.

Fernández-Kranz, D. and Rodríguez-Planas, N. (2021). Too Family Friendly? The Consequences of Parent Part-Time Working Rights. *Journal of Public Economics*, 197:104407.

Firth, M. (1982). Sex Discrimination in Job Opportunities for Women. *Sex Roles*, 8(8):891–901.

Flory, J. A., Leibbrandt, A., and List, J. A. (2015). Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions. *The Review of Economic Studies*, 82(1):122–155.

Fogli, A. and Veldkamp, L. (2011). Nature or Nurture? Learning and the Geography of Female Labor Force Participation. *Econometrica*, 79(4):1103–1138.

Galarza, F. B. and Yamada, G. (2017). Triple Penalty in Employment Access: The Role of Beauty, Race, and Sex. *Journal of applied economics*, 20(1):29–47.

Goldberg, P. K., Gottlieb, C. T., Lall, S., Mehta, M., Peters, M., and Ratan, A. L. (2025). The Global Gender Distortions Index (GGDI). Working Paper 34142, National Bureau of Economic Research.

Goldin, C. (2014a). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.

Goldin, C. (2014b). A Pollution Theory of Discrimination: Male and Female Differences in Occupations and Earnings. In *Human Capital in History: The American Record*, pages 313–348. National Bureau of Economic Research.

Goldstein, M., Gonzalez Martinez, P., and Papineni, S. (2019). Tackling the Global Profitarchy: Gender and the Choice of Business Sector. Policy Research Working Paper 8865, The World Bank.

Gradín, C. (2021). Occupational Gender Segregation in Post-APartheid SOuth AFrica. *Feminist Economics*, 27(3):102–133.

Greenwood, J., Seshadri, A., and Yorukoglu, M. (2005). Engines of Liberation. *The Review of Economic Studies*, 72(1):109–133.

Heath, R., Bernhardt, A., Borker, G., Fitzpatrick, A., Keats, A., McKelway, M., Menzel, A., Molina, T., and Sharma, G. (2024). Female Labour Force Participation. *VoxDevLit*, 11(1).

Heath, R. and Mobarak, A. M. (2015). Manufacturing Growth and the Lives of Bangladeshi Women. *Journal of Development Economics*, 115:1–15.

Hederos, K., Sandberg, A., Kvissberg, L., and Polano, E. (2024). Gender Homophily in Job Referrals: Evidence from a Field Study among University Students. *Labour Economics*.

Heilman, M. E., Wallen, A. S., Fuchs, D., and Tamkins, M. M. (2004). Penalties for Success: Reactions to Women Who Succeed at Male Gender-Typed Tasks. *Journal of applied psychology*, 89(3):416.

Holland, P. W., Laskey, K. B., and Leinhardt, S. (1983). Stochastic Blockmodels: First Steps. *Social Networks*, 5(2):109–137.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica*, 87(5):1439–1474.

Ibanez, M. and Riener, G. (2018). Sorting through Affirmative Action: Three Field Experiments in COlombia. *Journal of Labor Economics*, 36(2):437–478.

Ibarra, H. (1992). Homophily and Differential Returns: Sex Differences in Network Structure and Access in an Advertising Firm. *Administrative Science Quarterly*, 37(3):422–447.

Jayachandran, S. (2015). The Roots of Gender Inequality in Developing Countries. *Annual Review of Economics*, 7:63–88.

Jayachandran, S. (2021). Social Norms as a Barrier to Women's Employment in Developing Countries. *IMF Economic Review*, 69(3):576–595.

Kessler, J. B., Low, C., and Sullivan, C. D. (2019). Incentivized Resume Rating: Eliciting Employer Preferences Without Deception. *American Economic Review*, 109(11):3713–44.

Kline, P., Rose, E. K., and Walters, C. R. (2022). Systemic Discrimination among Large U.S. Employers. *The Quarterly Journal of Economics*, 137(4):1963–2036.

Kuhn, P. and Shen, K. (2013). Gender Discrimination in Job Ads: Evidence from CHina. *The Quarterly Journal of Economics*, 128(1):287–336.

Lepage, L.-P., Li, X., and Zafar, B. (2025). Anticipated Discrimination and Major Choice. Working Paper 33680, National Bureau of Economic Research.

Lommerud, K. E., Straume, O. R., and Vagstad, S. (2015). Mommy Tracks and Public Policy: On Self-Fulfilling Prophecies and Gender Gaps in Hiring and Promotion. *Journal of Economic Behavior & Organization*, 116:540–554.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27:415–444.

Neumark, D., Burn, I., and Button, P. (2019). Is It Harder for Older Workers to Find Jobs? New and Improved Evidence from a Field Experiment. *Journal of Political Economy*, 127(2):922–970.

Neumark, D. and Postlewaite, A. (1998). Relative Income Concerns and the Rise in Married Women's Employment. *Journal of Public Economics*, page 27.

Nicoletti, C., Salvanes, K. G., and Tominey, E. (2018). The Family Peer Effect on Mothers' Labor Supply. *American Economic Journal: Applied Economics*, 10(3):206–234.

Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away from Competition? Do Men Compete Too Much? *The quarterly journal of economics*, 122(3):1067–1101.

Olivetti, C., Pan, J., and Petrongolo, B. (2024). The Evolution of Gender in the Labor Market. In *Handbook of Labor Economics*. Elsevier. NBER Working Paper 33153.

Pan, J. (2015). Gender Segregation in Occupations: The Role of Tipping and Social Interactions. *Journal of Labor Economics*, 33(2):365–408.

Riach, P. A. and Rich, J. (2006). An Experimental Investigation of Sexual Discrimination in Hiring in the English Labor Market. *Advances in Economic Analysis & Policy*, 5(2).

Topa, G. (2011). Labor Markets and Referrals. In *Handbook of social economics*, volume 1, pages 1193–1221. Elsevier.

Witte, M. (2025). Why Do Workers Make Job Referrals? Experimental Evidence from Ethiopia. Discussion Paper 18258, Institute of Labor Economics (IZA).

Zeltzer, D. (2020). Gender Homophily in Referral Networks: Consequences for the Medicare Physician Earnings Gap. *American Economic Journal: Applied Economics*, 12(2):169–197.

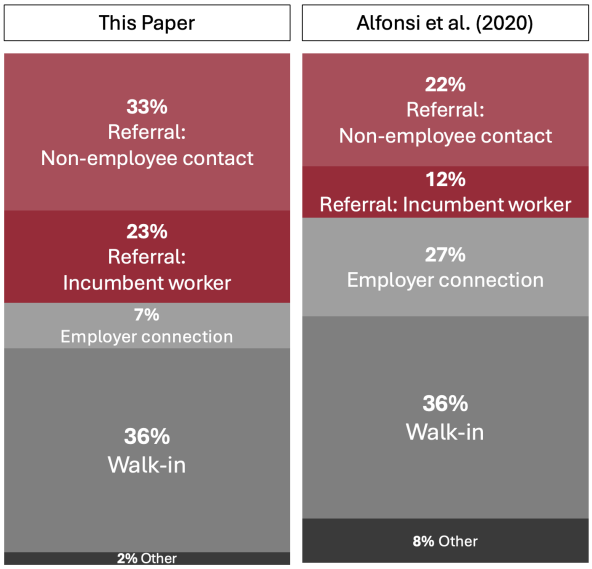
Appendix Tables and Figures

Table A.1: Balance Table

	Same Experience (1)	More Experienced Female (2)	More Experienced Male (3)	Test of equality [1=2] (4)	Test of equality [1=3] (5)	Test of equality [2=3] (6)
Demographics						
Age	23.787	23.972	24.204	[.278]	[.018**]	[.279]
	(.096)	(.145)	(.157)			
Sex (=1 female)	.376	.365	.371	[.790]	[.902]	[.900]
	(.025)	(.036)	(.036)			
Rural household	.553	.525	.575	[.538]	[.611]	[.333]
	(.026)	(.037)	(.036)			
District of origin: central	.289	.287	.263	[.958]	[.518]	[.610]
	(.023)	(.034)	(.032)			
Labor market history						
Any work in last month	.771	.740	.790	[.426]	[.605]	[.260]
	(.022)	(.033)	(.030)			
Wage employed	.524	.464	.575	[.188]	[.248]	[.033**]
	(.026)	(.037)	(.036)			
Monthly earnings (USD)	111.887	101.314	107.974	[.215]	[.644]	[.491]
	(4.880)	(6.112)	(7.089)			
Num. of Referrals	.937	.895	.909	[.842]	[.900]	[.957]
	(.124)	(.159)	(.197)			
Female dominated occupation	.342	.331	.355	[.804]	[.765]	[.639]
	(.024)	(.035)	(.035)			
Heavily Segregated Occupation	.782	.773	.769	[.829]	[.732]	[.916]
	(.021)	(.031)	(.031)			
Number of observations	380	181	186	561	566	367

Notes: This table compares baseline characteristics across the four treatment groups. "Same Experience" is a composite of both the "Both Low Experience" and "Both High Experience" treatment groups. For each variable, the table reports the mean and standard error, as well as the number of observations. The final three columns report pairwise t-tests of equality in means between treatment groups. The sample for earnings is half because of missing values.

Figure A.1: Referral Prevalence in Multiple Samples



Notes: Bars show the distribution of main job-finding channels for workers in our sample ("This Paper") and workers from Alfonsi et al. (2020). Channels are harmonized into five groups: (i) Referral: Non-employee contact, (ii) Referral: Incumbent workers, (iii) Employer connection, (iv) Walk-in, and (v) Other methods (media, internet, middlemen, NGO placements). Overall, 63% and 61% of jobs are obtained through a personal connection across both studies.

Figure A.2: Example of Profile Shown to Participants

CANDIDATE ID: 7484179
PERSONAL PROFILE
Name: Paul
Date of birth: 2nd July 2004
Sex: Male
Nationality: Ugandan
Home district: Jinja
ACADEMIC BACKGROUND
2021-2022: Nile Vocational Institute: National Certificate in Welding
2017-2020: Pilkington College Muguluka: Uganda Certificate of Education (UCE)
EXPERIENCE
April - August 2023: Bawumaye Metal Fabrication: Full-time worker. Main activities: performing welding repairs and modifications, aligning parts prior to assembly, and melting and soldering metal plates.
January - March 2023: Bawumaye Metal Fabrication: Trainee. Main activities: performing weld inspections and quality control checks, cutting metal plates, and aligning parts prior to assembly.
January - March 2021: J&K Market: Part-time job. Main activities: selling phone accessories, receiving payments, and stock management.
LANGUAGE
Excellent written and spoken English, Luganda and Lusoga.

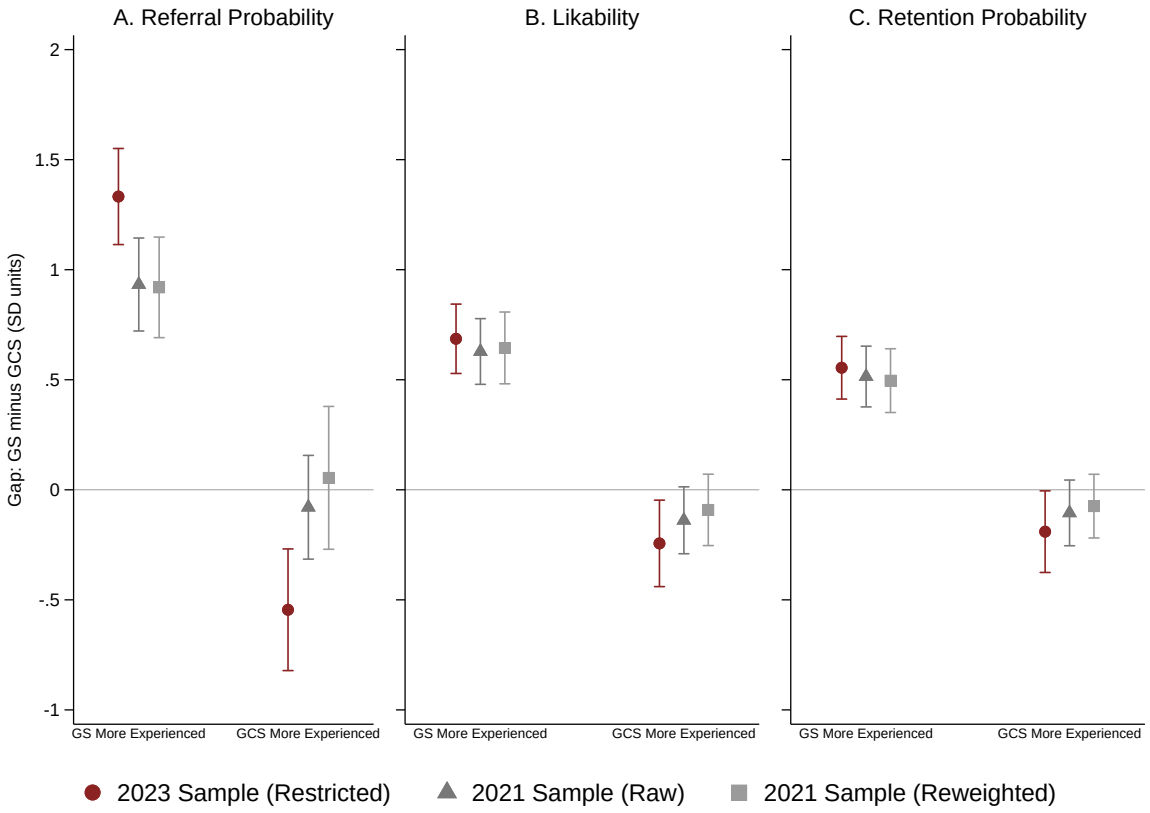
Notes: This figure presents an example of the candidate profile shown to participants during the experiment. Each profile contained standardized sections covering personal details, academic qualifications, work experience with specific tasks performed, and language abilities. All profiles followed the same structure and formatting, with only candidate-specific information varying.

Table A.2: External Validity: Study Sample vs UNHS Sectors

	UNHS Sample			Study Sample			
Occupation	% All	Max % (gender)	Dominant	% All (employment)	% All (training)	Max % (gender)	Dominant
Food and hospitality	9.0	53.5	Male	7.5	8.2	62.5	Female
Tailoring	1.1	57.2	Female	9.0	12.1	95.8	Female
Electrical work	0.3	96.3	Male	13.8	20.0	97.3	Male
Motor mechanics	2.4	95.2	Male	9.2	17.9	100.0	Male
Construction	8.3	99.9	Male	10.3	8.0	92.7	Male
Plumbing	0.1	100.0	Male	7.9	12.5	83.3	Male
Retail	28.4	60.8	Male	12.3	0.0	51.5	Female
Business services	1.1	54.9	Male	2.8	3.7	53.3	Male
Teaching (primary and pre-primary)	7.2	62.8	Male	9.5	9.2	94.1	Female
Hairdressing	2.8	62.8	Male	2.2	2.2	91.7	Female
Health services	2.0	60.8	Male	0.2	0.0	100.0	Male
Machining and fitting	1.6	96.0	Male	1.9	2.2	100.0	Male
Welding and carpentry	1.8	91.8	Male	3.6	3.7	100.0	Male
Transport	12.2	99.4	Male	1.9	0.0	100.0	Male
Public administration	0.9	86.1	Male	3.4	0.0	61.1	Male
Communication	1.0	70.1	Male	0.0	0.0	0.0	-
Arts and recreation	0.7	82.7	Male	0.0	0.0	0.0	-
Domestic/household services	2.4	60.7	Female	1.5	0.0	87.5	Female
Other unskilled	10.0	74.3	Male	2.8	0.0	86.7	Male
Other skilled	6.9	77.3	Male	0.4	0.0	100.0	Male
N (observations)		7131			535		

Notes: "Max % (gender)" is the larger of the female and male shares within a sector; "Dominant" is that gender. Agriculture is excluded from both samples. *Other unskilled* includes painting, street food vending, street hawkers, market vending, gate keepers/guards, factory work, and car/boda washing; it also absorbs UNHS categories other personal activities, waste management, other support activities, and other service activity. *Other skilled* includes mining and quarrying, other manufacturing, international organizations, and other skilled manufacturing.

Figure A.3: GS–GCS Gaps Across Samples



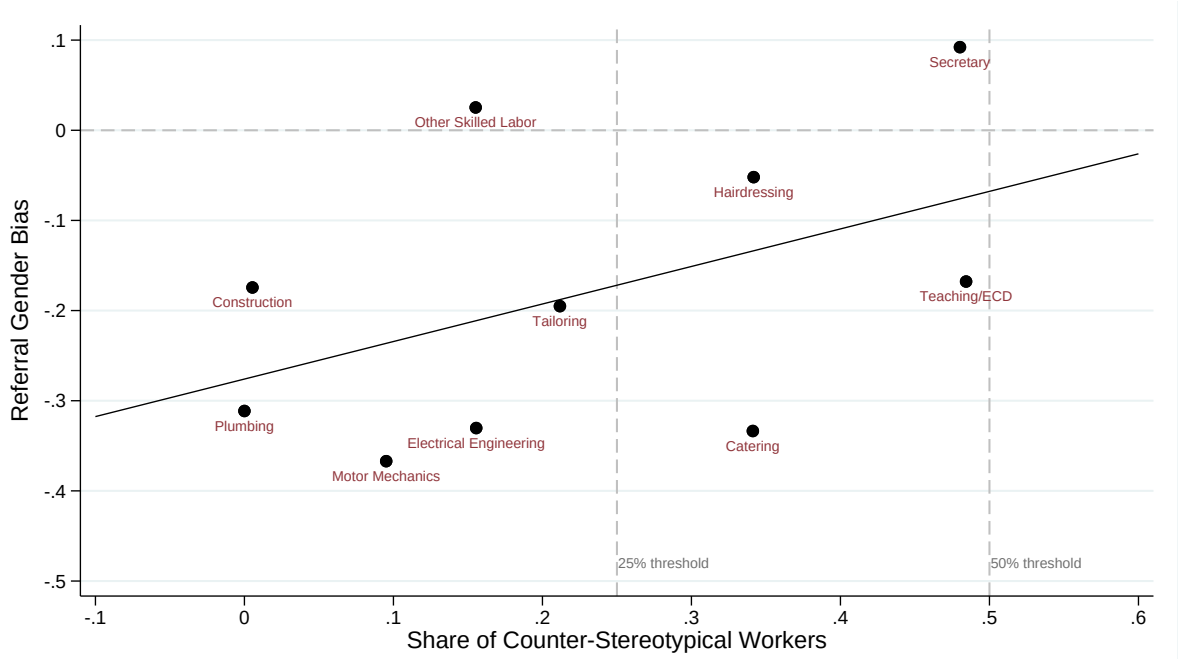
Notes: Each point is a linear combination from the specification $Y_{p,i} = \alpha + \beta_1 \cdot \text{nonstereo} + \beta_2 \cdot \text{plus5} + \text{controls} + \varepsilon_{p,i}$, estimated separately for each sample on all occupations (female- and male-dominated pooled). "GS More Experienced" plots $\beta_2 - \beta_1$; "GCS More Experienced" plots $-(\beta_1 + \beta_2)$. The 2023 sample is restricted to the two treatment arms present in both waves (high-low experience only). The 2021 sample is shown both with raw estimates and reweighted by training-area composition to match the occupational mix of the 2023 sample. All estimates are standardized by the standard deviation of each outcome in the restricted 2023 sample to permit cross-outcome comparison; the y-axis is in standard-deviation units. Controls for likability and retention: respondent gender, an indicator for female-dominated occupation, baseline work experience, interview status at the previous survey wave, VTI and training-area fixed effects. Referral regressions include no controls. Whiskers denote 95% confidence intervals; standard errors clustered at the respondent level.

Table A.3: Network Segregation

	Sample of reference when conditioning	Share in Full Sample %	Female Dominated Occupations %	Male Dominated Occupations %
Panel A: Gender Composition of Named and Referred Workers				
Network Availability				
Named a network	747	38.2	33.2	40.7
Named female network	285	40.7	84.7	22.0
Named male network	285	74.4	30.6	93.0
Named both genders network	285	15.1	15.3	15.0
Network Referred				
Referred female network	285	18.9	47.1	7.0
Referred male network	285	49.8	16.5	64.0
Referred female network (cond. named both)	43	23.3	23.1	23.3
Referred male network (cond. named both)	43	48.8	38.5	53.3
Panel B: Network Segregation by Network Constraint				
Constrained Network				
Named a network	747	38.2	33.2	40.7
Named female network only	285	25.6	69.4	7.0
Named male network only	285	59.3	15.3	78.0
Unconstrained Network				
Named both genders network	43	100.0	100.0	100.0
Referred female network (cond. named both)	43	23.3	23.1	23.3
Referred male network (cond. named both)	43	48.8	38.5	53.3

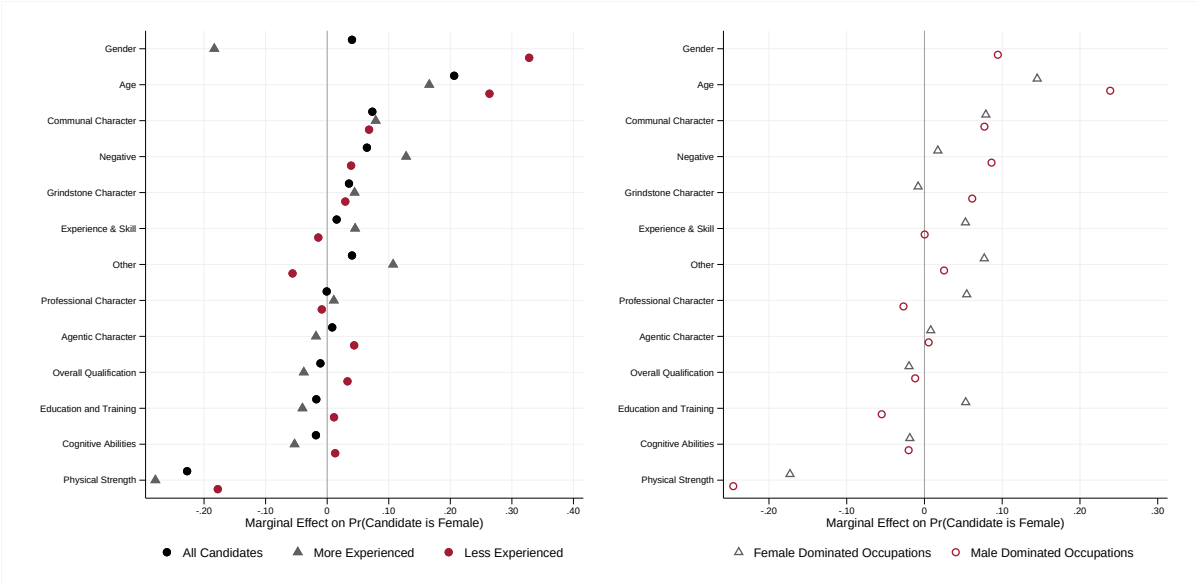
Notes: Each row reports the conditioning sample size (column 1) and the share satisfying the condition (columns 2–4). Columns 2–4 correspond to the full sample, female-dominated occupations, and male-dominated occupations. *Panel A*: "Network Availability" refers to named contacts; "Network Referred" refers to contacts chosen when asked to refer. *Panel B*: "Constrained Network" names only one gender of contacts; "Unconstrained Network" names both genders.

Figure A.4: Referral Gender Bias and Counter-Stereotypical Employment Share



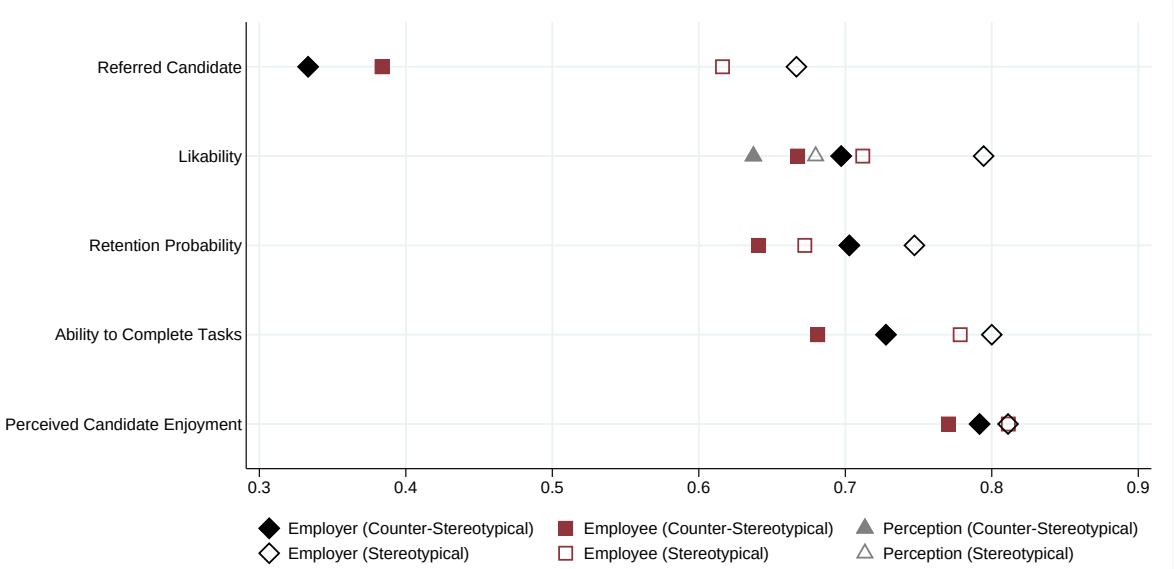
Notes: Each point represents an occupation-level referral gender bias, where negative values indicate a penalty against counter-stereotypical workers. The occupation-level penalty is plotted against the share of counter-stereotypical workers in that occupation. The fitted line has slope = 0.416 ($p = 0.179$). Vertical dashed lines mark the 25% and 50% thresholds for the share of counter-stereotypical workers.

Figure A.5: Gender Predictors in LLM-Assigned Adjective Categories



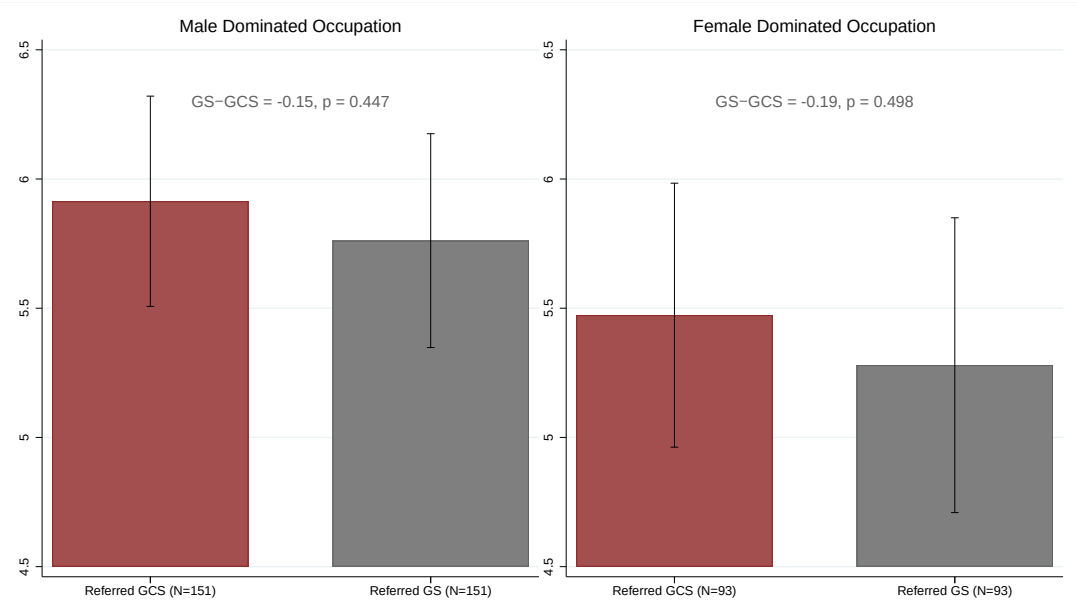
Notes: This figure reports marginal effects from logistic regressions of the probability that a CV belongs to a female candidate on indicator variables for each adjective category assigned by the LLM. Adjectives are grouped into 13 mutually exclusive categories. The left panel shows marginal effects estimated separately by experience level. The right panel estimates marginal effects separately for CVs in FDOs and MDOs. All specifications cluster standard errors at the respondent level. Positive coefficients indicate categories more commonly assigned to female CVs; viceversa for negative coefficients.

Figure A.6: What Do Employers Really Think?



Notes: This figure summarizes how employers and employees evaluate gender stereotypical and counter-stereotypical candidates. Each marker plots the average probability or rating for a given outcome on the vertical axis—referral probability, likability, retention probability, perceived ability to complete tasks, and perceived candidate enjoyment—against its level on the horizontal axis (0–1 scale). Diamonds show employer assessments, squares show employee assessments, and triangles show employees' perception of employer assessments; filled symbols correspond to counter-stereotypical candidates and hollow symbols to stereotypical candidates.

Figure A.7: Reputational Risks: Vignette Analysis



Notes: This figure reports mean perceived reputational risk (0–10 scale) to the referrer if a referred candidate underperforms, by occupation type (MDO or FDO). Data come from Wave 1. Within each panel, bars compare referrals to gender stereotypical (GS) and counter-stereotypical (GCS) candidates, where GCS candidates are the gender minority in the occupation; whiskers show 95% confidence intervals. Each panel reports the GS–GCS difference in means and p-value from a paired equality test.

Material for Online Appendix

A A Model of Referrals: Counterfactuals and Robustness

In this section, we present additional counterfactual exercises with different model calibrations to gauge the importance of the referral gender bias in driving labor market segregation. Overall, all exercises indicate that, under our model, the relevance of the network channel and of the referral gender bias depends on three factors: (1) the relative importance of first-hand signals in the job search process, (2) network density, and (3) network homophily.

First-hand signals In our baseline calibration, first-hand signals ($\sigma_g^{(k)}$) are retrieved through a simulated method of moments, assuming that labor market segregation is currently in equilibrium. However, our estimated parameters indicate differences of more than one order of magnitude between stereotypical and non-stereotypical candidates.[13] To assess robustness of our results relative to the importance of RGB, we propose several alternative calibrations.

First, we use estimates from an audit study we carried out with a subsample of our respondents' employers, which allows us to retrieve employers' gender preferences for hiring. Results from this audit study indicate a much smaller gap between stereotypical and non-stereotypical candidates than our SMM estimates indicate – specifically, a .33 pp penalty to non-stereotypical candidates across all sectors. Using this information, we recalibrate the model to set $\sigma_m^{(1)} = .0615$, $\sigma_m^{(2)} = .1078$, $\sigma_f^{(1)} = .1249$, $\sigma_f^{(2)} = .0531$.

Second, we use survey data from our respondents to approximate the rate at which workers receive job offers. In our survey, we collect information on job search behavior of all 753 respondents, asking whether they looked for a job in the last 30 days, time spent in job search, number of positions applied, number of interviews/callbacks received, and number of offers received. In our model, we assume that workers may receive job offers even if they are currently employed and not looking for a job – which we deem realistic. However, in our survey, we only collected information about offers received from those who were looking for a job. Here, we assume that the success rate of job seekers (probability of getting an offer conditional on job search in the last 30 days) is an approximation of $\sigma_g^{(k)}$. In other words, we extrapolate the number of offers received from job seekers to workers who are not looking for a job. Though this might seem unrealistic for the employed worker, in our model, $\sigma_g^{(k)}$ need not be interpreted strictly

[13] Estimated parameters are: $\sigma_m^{(1)} = .0072$, $\sigma_m^{(2)} = .1536$, $\sigma_f^{(1)} = .1792$, $\sigma_f^{(2)} = .0072$.

as job offers received by them, but also as *job openings heard of in their company or sector*. Using this information, we recalibrate the model to set $\sigma_m^{(1)} = .0476, \sigma_m^{(2)} = .0686$, $\sigma_f^{(1)} = .0227, \sigma_f^{(2)} = .0297$.

Figure OA.1 presents the results of these counterfactual exercises, using reweighted parameters from the correspondence study with employers (Panel a) and using proxies from survey data (Panel b). These results indicate that the importance of RGB in driving labor market segregation is larger when first-hand signals are either less relevant or less segregated. Reweighing the first-hand signals based on the audit study with employers (Figure OA.1, Panel a) increases the importance of RGB to 23.9-26.3% ("S+C+RGB" vs. "S+C"). Using survey data on job offers (Figure OA.1, Panel b) increases the importance of RGB to 46.9%-105.8% ("S+C+RGB" vs. "S+C").

Network density In our baseline calibration, we use survey data to approximate the average degree of our respondents' networks, which is around 5 connections per individual. Then, based on the size of our simulated economy ($N = 100$), we calibrate the stochastic block model to match this average degree. However, it is possible that our survey underestimates the true size of individuals' networks. To gauge the importance of network density in driving segregation, we carry out a counterfactual exercise where we double the average degree of individuals' networks (from 5 to 10 connections per individual). Figure OA.2 shows the steady state estimates. Under a denser network (average degree 10), RGB increases segregation by 9.9%-11.2% ("S+C+RGB" vs. "S+C").

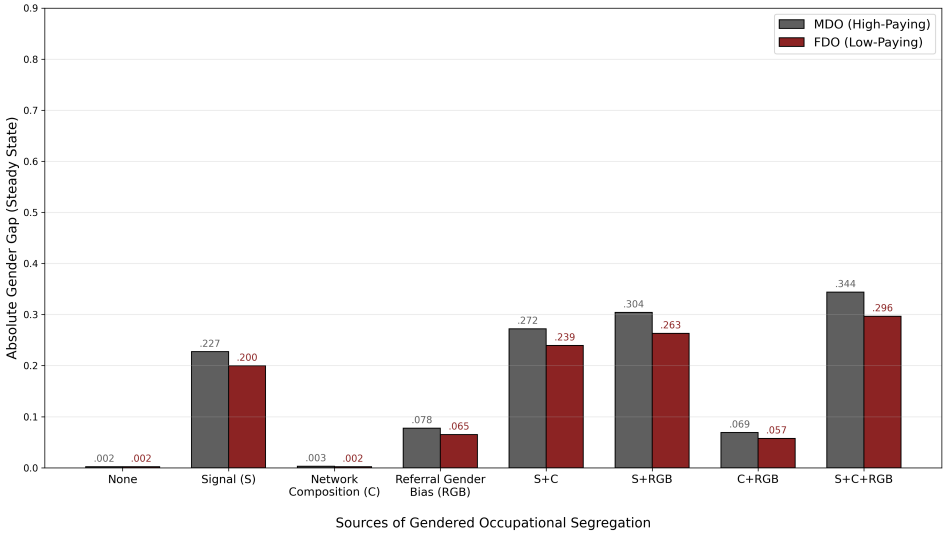
Labor market gender composition In addition to these recalibrations, we also assess whether the gender composition of the economy affects the importance of RGB in driving segregation. We simulate a different stochastic block model with $p_m = .0615, p_f = .0667, q = .0368$ in a network with $N_m = 62, N_f = 38$. These values yield a network that matches our survey data, with men having on average 5.15 connections out of which 3.76 are men, whereas women have on average 4.75 connections out of which 2.28 are men. Figure OA.3 shows the steady states for this scenario, under the different scenarios. In our baseline comparison ("S+C+RGB" vs. "S+RGB"), this change in labor market composition increases the importance of RGB to 9.0-1.2%.

Overall, the results indicate that the importance of RGB in driving labor market segregation is larger when: (1) first-hand signals are less relevant – or less segregated, (2) when networks are denser, and (3) when networks are less homophilic.

Figure OA.1: Steady State Segregation, Under Different $\sigma_g^{(s)}$

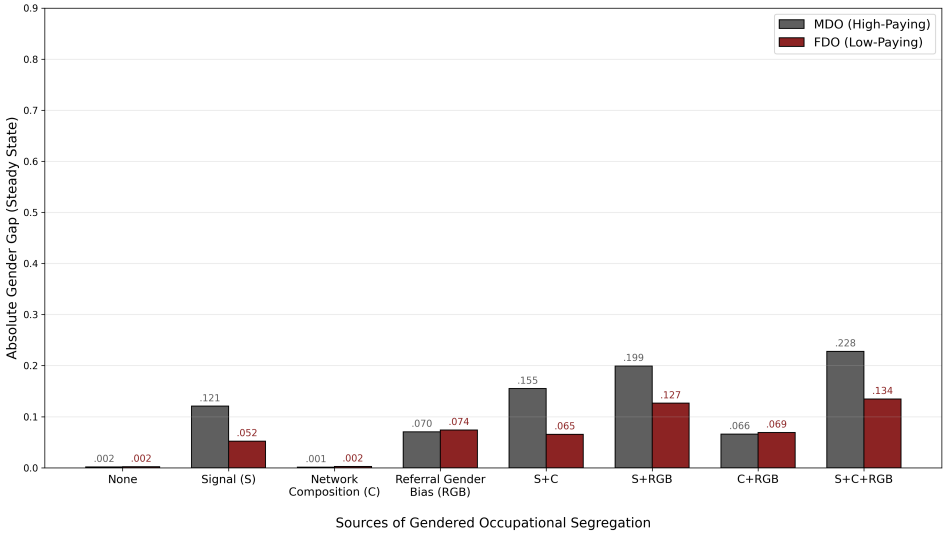
(a) Reweighed estimates

$\sigma_m^{(1)} = .0615$, $\sigma_m^{(2)} = .1078$, $\sigma_f^{(1)} = .1249$, $\sigma_f^{(2)} = .0531$



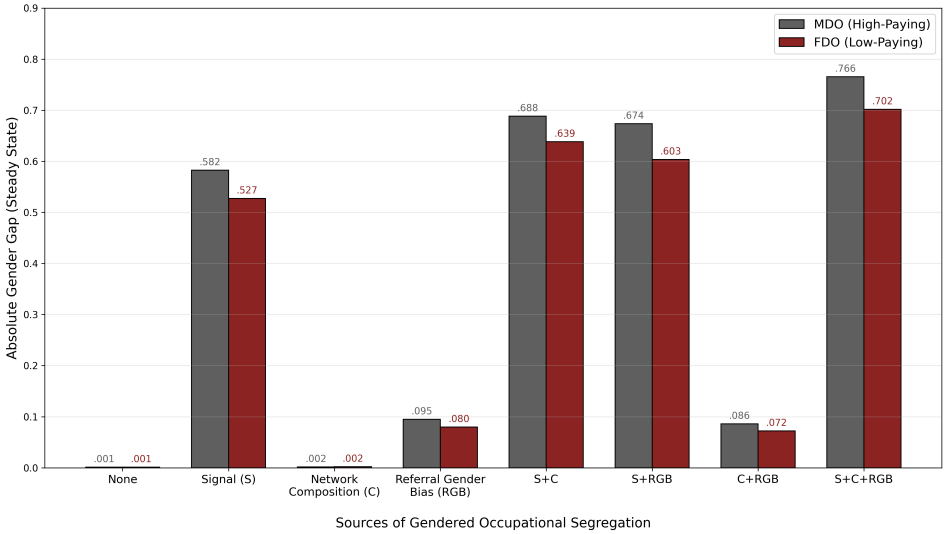
(b) Survey data

$\sigma_m^{(1)} = .0476, \sigma_m^{(2)} = .0686$, $\sigma_f^{(1)} = .0227, \sigma_f^{(2)} = .0297$



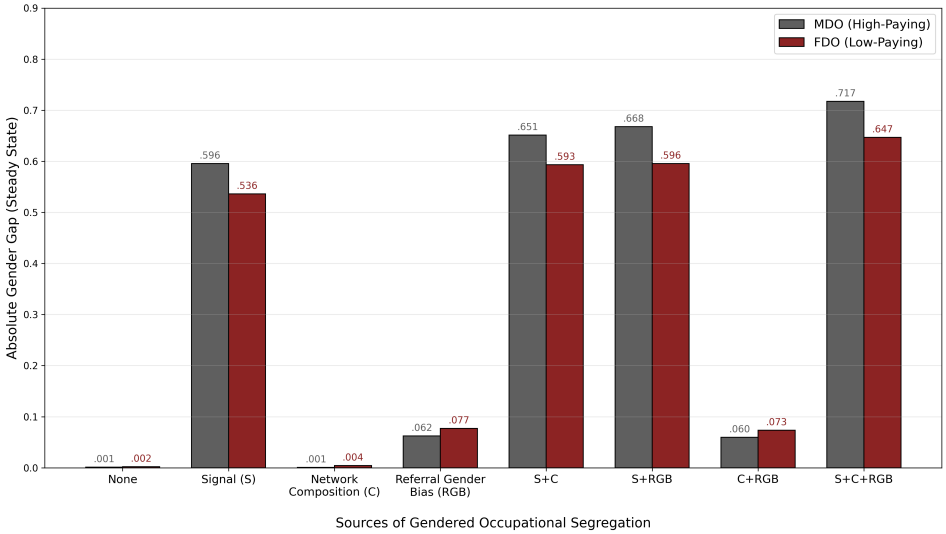
Note: Gender gap is defined as the difference between the share of men and the share of women in an occupation. The figure plots the average gender gap across 2,000 Monte Carlo simulations and $T = 1,500$. Steady states are approximated by averaging the last 50 periods. Each pair of columns represents the steady state gender gap under different combinations of sources of segregation. "S" stands for differential job arrival rates (signals). "C" stands for network composition (homophily). "RGB" stands for referral gender bias. Parameters for "S+C+RGB" were: $N = 100$, $p = .064$, $q = .035$, $\delta = .03$, $\alpha_S^{(1)} = .356$, $\alpha_{NS}^{(1)} = .644$, $\alpha_S^{(2)} = .589$, $\alpha_{NS}^{(2)} = .411$. When "S" is shut off, $\sigma_m^{(1)} = \sigma_f^{(1)}$, $\sigma_m^{(2)} = \sigma_f^{(2)}$. When "C" is shut off, $p = .0509$, $q = .0498$. When "RGB" is shut off, $\alpha_S^{(s)} = \alpha_{NS}^{(s)} = 1$.

Figure OA.2: Steady State Segregation, Under Denser Network



Note: Gender gap is defined as the difference between the share of men and the share of women in an occupation. The figure plots the average gender gap across 2,000 Monte Carlo simulations and $T = 1,500$. Steady states are approximated by averaging the last 50 periods. Each pair of columns represents the steady state gender gap under different combinations of sources of segregation. "S" stands for differential job arrival rates (signals). "C" stands for network composition (homophily). "RGB" stands for referral gender bias. Parameters for "S+C+RGB" were: $N = 100$, $p = .128$, $q = .070$, $\delta = .03$, $\sigma_m^{(1)} = .0072$, $\sigma_m^{(2)} = .1536$, $\sigma_f^{(1)} = .1792$, $\sigma_f^{(2)} = .0072$, $\alpha_S^{(1)} = .356$, $\alpha_{NS}^{(1)} = .644$, $\alpha_S^{(2)} = .589$, $\alpha_{NS}^{(2)} = .411$. When "S" is shut off, $\sigma_m^{(1)} = \sigma_f^{(1)} = .0932$, $\sigma_m^{(2)} = \sigma_f^{(2)} = .0804$. When "C" is shut off, $p = .0509$, $q = .0498$. When "RGB" is shut off, $\alpha_S^{(s)} = \alpha_{NS}^{(s)} = 1$.

Figure OA.3: Steady State Segregation, Under Different Gender Composition



Note: Gender gap is defined as the difference between the share of men and the share of women in an occupation. The figure plots the average gender gap across 2,000 Monte Carlo simulations. Each pair of columns represents the steady state gender gap under different combinations of sources of segregation. "S" stands for differential job arrival rates (signals). "C" stands for network composition (homophily). "RGB" stands for referral gender bias. Parameters for "S+C+RGB" were: $N_m = 62$, $N_f = 38$ $p_m = .0615$, $p_f = .0667$, $q = .0368$, $\delta = .03$, $\sigma_m^{(1)} = .0072$, $\sigma_m^{(2)} = .1536$, $\sigma_f^{(1)} = .1792$, $\sigma_f^{(2)} = .0072$, $\alpha_S^{(1)} = .356$, $\alpha_{NS}^{(1)} = .644$, $\alpha_S^{(2)} = .589$, $\alpha_{NS}^{(2)} = .411$. When "S" is shut off, $\sigma_m^{(1)} = \sigma_f^{(1)} = .0726$, $\sigma_m^{(2)} = \sigma_f^{(2)} = .0980$. When "C" is shut off, $p_m = .0539$, $p_f = .0464$, $q = .0489$. When "RGB" is shut off, $\alpha_S^{(s)} = \alpha_{NS}^{(s)} = 1$.

B Matrix Notation

Denote $e_t^{(0)}, e_t^{(1)}, e_t^{(2)} \in \{0,1\}^N$ as the vectors representing whether, at period t, each of the N individuals is unemployed, employed in sector 1, and employed in sector 2, respectively. Individuals are ordered by gender in the vector, so that the first rN are female nodes. At period $t = 1$, the probability that individual of gender g is in a state k is given by $\eta_g^{(k)}$, so that $e_0^{(k)}$ is a realization of a multinomial distribution with parameters $\eta_g^{(0)}, \eta_g^{(1)}, \eta_g^{(2)}$. Therefore:

$$\mathbb{E}[e_0^{(0)}] = \begin{bmatrix} \eta_f^{(0)} \cdot \mathbb{1}_{Nr} \\ \eta_m^{(0)} \cdot \mathbb{1}_{N(1-r)} \end{bmatrix} \mathbb{E}[e_0^{(1)}] = \begin{bmatrix} \eta_f^{(1)} \cdot \mathbb{1}_{Nr} \\ \eta_m^{(1)} \cdot \mathbb{1}_{N(1-r)} \end{bmatrix} \mathbb{E}[e_0^{(2)}] = \begin{bmatrix} \eta_f^{(2)} \cdot \mathbb{1}_{Nr} \\ \eta_m^{(2)} \cdot \mathbb{1}_{N(1-r)} \end{bmatrix}$$

where r is the share of nodes that are female.

Next, denote p the $N \times N$ **connection probability matrix** of the network. Assuming a stochastic block model with within-block connection probability p and across-block connection probability $q < p$, the adjacency matrix is $P_{ij} = p$ if nodes i and j are of the same gender and $P_{ij} = q < p$ if nodes i and j are of different genders. Importantly, $P_{ii} = 0$ for all i. Therefore:

$$P = \begin{bmatrix} p \cdot (J_{Nr} - I_{Nr}) & q \cdot J_{Nr \times N(1-r)} \\ q \cdot J_{N(1-r) \times Nr} & p \cdot (J_{N(1-r)} - I_{N(1-r)}) \end{bmatrix}$$

where r is the share of nodes that are female, J_k is a $k \times k$ matrix of ones, and $J_{k \times l}$ is a $k \times l$ matrix of ones. Denote also the **adjacency matrix** $A \in \{0,1\}^{N \times N}$ such that $A_{ij} = 1$ if there is a link between nodes i and j and $A_{ij} = 0$ otherwise. Note that A is a random matrix generated according to the probabilities in P.

Denote now the **weight vectors** $w^{(1)}, w^{(2)} \in \mathbb{R}^N$ such that $w_i^{(k)}$ is the weight assigned to forwarding a signal of sector k to individual i. Therefore:

$$w^{(1)} = \begin{bmatrix} \alpha_S^{(1)} \cdot \mathbb{1}_{Nr} \\ \alpha_{NS}^{(1)} \cdot \mathbb{1}_{N(1-r)} \end{bmatrix} w^{(2)} = \begin{bmatrix} \alpha_{NS}^{(2)} \cdot \mathbb{1}_{Nr} \\ \alpha_S^{(2)} \cdot \mathbb{1}_{N(1-r)} \end{bmatrix}$$

where $\alpha_S^{(k)}$ is the weight assigned to forwarding a signal to a stereotypical candidate and $\alpha_{NS}^{(k)}$ is the weight assigned to forwarding a signal to a non-stereotypical candidate.

Denote now the **weight matrices** $W^{(1)}, W^{(2)} \in \mathbb{R}^{N \times N}$ such that $W_{ij}^{(k)}$ is the weight assigned by individual i to forwarding a signal of sector k to individual j – so that $W^{(k)} = \mathbb{1}_N \cdot w^{(k)T}$ and all rows of $W^{(k)}$ are equal to $w^{(k)}$. Therefore:

$$W^{(1)} = \begin{bmatrix} \alpha_S \cdot J_{Nr} & \alpha_{NS} \cdot J_{Nr \times N(1-r)} \\ \alpha_S \cdot J_{N(1-r) \times Nr} & \alpha_{NS} \cdot J_{N(1-r)} \end{bmatrix}$$
$$W^{(2)} = \begin{bmatrix} \alpha_{NS} \cdot J_{Nr} & \alpha_S \cdot J_{Nr \times N(1-r)} \\ \alpha_{NS} \cdot J_{N(1-r) \times Nr} & \alpha_S \cdot J_{N(1-r)} \end{bmatrix}$$

Denote the **transmission probability matrix** $\Theta^{(k)} \in \mathbb{R}^{N \times N}$ such that:

$$\Theta^{(k)} = S^{(k)^{-1}}(A \circ W^{(k)}) \quad \text{where} \quad S_{ii}^{(k)} = \begin{cases} \sum_{j=1}^{N} A_{ij} \cdot W_{ij}^{(k)} & \text{if } \sum_{j=1}^{N} A_{ij} \cdot W_{ij}^{(k)} > 0 \\ 1 & \text{otherwise} \end{cases}$$

where $\circ$ is the Hadamard product and $S^{(k)^{-1}}$ is the inverse of $S^{(k)}$, a diagonal matrix with the sum of the weights assigned by individual i to forwarding a signal of sector k to all her neighbors in the diagonal (with a normalization for isolated nodes). In other words, $\Theta_{ij}^{(k)}$ is the probability that individual i forwards a signal of sector k to individual j weighted (1) by the weights assigned to stereotypical and non-stereotypical candidates and (2) by the network structure.

Dynamics. In our process, each node i picks exactly one neighbor j to forward the signal if they have received one and are not unemployed. Denote the **first-hand signal probability vectors** $\sigma^{(1)}, \sigma^{(2)} \in [0,1]^N$ such that $\sigma_i^{(k)}$ is the probability that individual i receives a first-hand signal from sector k. By construction, these probabilities depend on the gender of the individual. Denote also the realization **first-hand signal vectors** $s_t^{(1)}, s_t^{(2)} \in \{0,1\}^N$ such that $s_{t,i}^{(k)}$ is the Bernoulli realization of whether individual i received a first-hand signal from sector k at period t. Therefore:

$$\mathbb{E}[s_t^{(k)}] = \begin{bmatrix} \sigma_f^{(k)} \cdot \mathbb{1}_{Nr} \\ \sigma_m^{(k)} \cdot \mathbb{1}_{N(1-r)} \end{bmatrix}$$

Now denote the **forwarder vectors** $\hat{s}_t^{(1)}, \hat{s}_t^{(2)} \in \{0,1\}^N$ such that $\hat{s}^{(k)} = s^{(k)} \circ (\mathbb{1}_N - e_t^{(0)})$ – i.e., $\hat{s}_i^{(k)}$ equals 1 if individual i can forward a signal of sector k at period t (it is the realization of whether node i is capable of forwarding a signal of type k at period t) – which happens when (1) i received a first-hand signal from sector k and (2) i is employed.

Now we describe the forwarding process through which nodes receive referrals. Conditional on receiving a first-hand signal and being employed, a node becomes a potential forwarder. If $\hat{s}_{t,i}^{(k)} = 1$, node i selects exactly one neighbor j to pass the referral to. This selection process is probabilistic and depends on the weights defined in the transmission matrix $\Theta^{(k)}$. Let $\phi_{i \to j}^{(k)} \in \{0,1\}$ be a binary random variable indicating whether i forwards

the signal to j, if they can forward one. The vector of forwarding decisions for node i, denoted $\phi_i^{(k)} = [\phi_{i\to 1}^{(k)}, \ldots, \phi_{i\to N}^{(k)}]$, is drawn from a multinomial distribution:

$$\phi_i^{(k)} \sim \text{Multinomial}(1, \Theta_{i,\cdot}^{(k)})$$

where $\Theta_{i,\cdot}^{(k)}$ is the i-th row of the transmission matrix representing the probabilities of choosing each neighbor. By definition, $\sum_{j=1}^{N} \phi_{i\to j}^{(k)} = 1$ for nodes that are connected.

We define the **forwarding matrix** $\Phi_t^{(k)} \in \{0,1\}^{N\times N}$ such that $\Phi_{t,ij}^{(k)} = \phi_{i\to j}^{(k)}$ if $\hat{s}_{t,i}^{(k)} = 1$, and 0 otherwise. In other words, $\Phi_{t,ij}^{(k)}$ indicates whether node i forwards a signal of type k to node j at period t.

A node j receives a second-hand signal of type k if at least one neighbor forwards it to them. Let $\tilde{s}_t^{(k)} \in \{0,1\}^N$ be the **vector of realized second-hand signals**. The element for node j is given by:

$$\tilde{s}_{t,j}^{(k)} = \mathbb{1}\left(\sum_{i=1}^{N} \Phi_{t,ij}^{(k)} > 0\right)$$

where $\mathbb{1}(\cdot)$ is the indicator function.

Updating states. Finally, the state vectors are updated. To resolve cases where an agent receives offers from both sectors simultaneously, we introduce a **tie-breaking vector** $\gamma_t \in \{0,1\}^N$, where each element is an independent Bernoulli realization with $E[\gamma_t] = .5$. Define the **offer vectors** $a^{(0)}, a^{(1)}, a^{(2)} \in \{0,1\}^N$ as:

$$\begin{aligned}
a^{(0)} &= (1 - s^{(1)}) \circ (1 - s^{(2)}) \\
a^{(1)} &= s^{(1)} \circ (1 - s^{(2)}) + \gamma_t \circ s^{(1)} \circ s^{(2)} \\
a^{(2)} &= (1 - s^{(1)}) \circ s^{(2)} + (1 - \gamma_t) \circ s^{(1)} \circ s^{(2)}
\end{aligned}$$

Define the **referral vectors** $b_t^{(0)}, b_t^{(1)}, b_t^{(2)} \in \{0,1\}^N$ as:

$$\begin{aligned}
b_t^{(0)} &= (1 - \tilde{s}_t^{(1)}) \circ (1 - \tilde{s}_t^{(2)}) \\
b_t^{(1)} &= \tilde{s}_t^{(1)} \circ (1 - \tilde{s}_t^{(2)}) + \gamma_t \circ \tilde{s}_t^{(1)} \circ \tilde{s}_t^{(2)} \\
b_t^{(2)} &= (1 - \tilde{s}_t^{(1)}) \circ \tilde{s}_t^{(2)} + (1 - \gamma_t) \circ \tilde{s}_t^{(1)} \circ \tilde{s}_t^{(2)}
\end{aligned}$$

With a **job destruction vector** d defined as the realization of a Bernoulli random variable with parameter δ, the state vectors are updated as:

$$
\begin{aligned}
e_t^{(0)} &= (e_{t-1}^{(0)} \circ a^{(0)}) \circ b_{t-1}^{(0)} + d \circ (e_{t-1}^{(1)} + e_{t-1}^{(2)}) \\
e_t^{(1)} &= (e_{t-1}^{(1)} + e_{t-1}^{(0)} \circ a^{(1)}) + (e_{t-1}^{(0)} \circ a^{(0)}) \circ b_{t-1}^{(1)} - d \circ e_{t-1}^{(1)} \\
e_t^{(2)} &= (e_{t-1}^{(2)} + e_{t-1}^{(0)} \circ a^{(2)}) + (e_{t-1}^{(0)} \circ a^{(0)}) \circ b_{t-1}^{(2)} - d \circ e_{t-1}^{(2)}
\end{aligned}
$$

Untractability of the model. There are two main challenges that prevent us from deriving a closed-form solution for the steady-state distribution of employment in this model. First, the presence of network effects complicates the analysis of the evolution of the steady states: to track the evolution of $\mathbb{E}[e_t^{(k)}]$, we would need to keep track of higher-order moments, such as $\mathbb{E}[e_t^{(k)} \circ e_t^{(l)}]$ for $k, l \in \{0, 1, 2\}$, which leads to an infinite hierarchy of equations. Second, the multinomial nature of the forwarding process introduces additional non-linearities that further complicate the analysis. Therefore, we resort to Monte Carlo simulations to analyze the long-term behavior of the system.

C Additional Tables and Figures

Table OA.1: Wave 1 and Wave 2 Demographics and Labor Market Outcomes

	Earlier Sample (2021)			Current Study Sample (2023)		
Variable	N	Mean	SE	N	Mean	SE
Age	546	27.075	0.140	747	23.936	0.072
Female	547	0.389	0.021	747	0.372	0.018
Employed during survey	547	0.786	0.018	747	0.768	0.015
Female dominated occupation	547	0.377	0.021	747	0.343	0.017
Rural	474	0.519	0.023	747	0.552	0.018
Monthly earnings (USD), Employed	214	114.369	4.729	389	108.538	3.419
Num. of referrals in the last 12 mo.	0	.	.	747	0.920	0.089

Notes: This table compares baseline characteristics between an earlier sample (2021) and the current study's sample (2023). For each variable, the table reports the sample size, mean, and standard errors. Variables include age, gender, current employment status, whether the trainee works in a female-dominated occupation, whether the trainee is originally from a rural area, monthly earnings among those employed, and the number of referrals made in the past 12 months. Number of referrals not collected in 2021 sample.

Table OA.2: External Validity: Socio-economic and Labor Market Characteristics

	(1) UNHS Sample		(2) UNHS VTI Sample		(3) Study Sample		T-test P-value	
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(3)-(1)	(3)-(2)
Female	20409	0.55	785	0.52	747	0.37	0.00***	0.00***
		(0.00)		(0.03)		(0.02)		
Age	20409	27.18	785	29.37	747	23.94	0.00***	0.00***
		(0.06)		(0.26)		(0.07)		
Completed primary school	17216	0.59	785	1.00	747	1.00	0.00***	N/A
		(0.01)		(0.00)		(0.00)		
Completed secondary school	17216	0.16	785	1.00	747	1.00	0.00***	0.16
		(0.00)		(0.00)		(0.00)		
Completed vocational training	16499	0.05	785	1.00	747	0.92	0.00***	0.00***
		(0.00)		(0.00)		(0.01)		
Any work in last 7 days - no agri.	20409	0.46	785	0.75	747	0.77	0.00***	0.59
		(0.00)		(0.02)		(0.02)		
Any work in last 7 days - agri. included	20088	0.77	785	0.84	747	0.79	0.28	0.02**
		(0.00)		(0.02)		(0.01)		
Monthly earnings (USD) - wage employed	3850	79.46	434	106.66	389	108.54	0.00***	0.80
		(2.36)		(6.69)		(3.42)		
Female sample								
Age	11209	27.09	391	28.67	278	23.31	0.00***	0.00***
		(0.08)		(0.39)		(0.10)		
Completed primary school	9462	0.56	391	1.00	278	1.00	0.00***	N/A
		(0.01)		(0.00)		(0.00)		
Completed secondary school	9462	0.14	391	1.00	278	0.99	0.00***	0.16
		(0.01)		(0.00)		(0.01)		
Completed vocational training	9117	0.05	391	1.00	278	0.90	0.00***	0.00***
		(0.00)		(0.00)		(0.02)		
Any work in last 7 days - no agri.	11209	0.36	391	0.69	278	0.73	0.00***	0.34
		(0.01)		(0.03)		(0.03)		
Any work in last 7 days - agri. included	11040	0.72	391	0.76	278	0.75	0.23	0.75
		(0.01)		(0.03)		(0.03)		
Monthly earnings (USD) - wage employed	1398	63.06	191	83.06	147	87.87	0.00***	0.60
		(3.17)		(7.85)		(4.91)		
Male sample								
Age	9200	27.29	394	30.12	469	24.31	0.00***	0.00***
		(0.09)		(0.34)		(0.09)		
Completed primary school	7754	0.62	394	1.00	469	1.00	0.00***	N/A
		(0.01)		(0.00)		(0.00)		
Completed secondary school	7754	0.19	394	1.00	469	1.00	0.00***	N/A
		(0.01)		(0.00)		(0.00)		
Completed vocational training	7382	0.05	394	1.00	469	0.93	0.00***	0.00***
		(0.00)		(0.00)		(0.01)		
Any work in last 7 days - no agri.	9200	0.58	394	0.82	469	0.79	0.00***	0.39
		(0.01)		(0.03)		(0.02)		
Any work in last 7 days - agri. included	9048	0.84	394	0.93	469	0.81	0.16	0.00***
		(0.01)		(0.01)		(0.02)		
Monthly earnings (USD) - wage employed	2452	88.59	243	125.91	242	121.09	0.00***	0.66
		(3.18)		(10.12)		(4.43)		

Notes: This table compares socio-economic and labor market characteristics across three samples: the Uganda National Household Survey (UNHS) population, the UNHS subsample with vocational training (VTI), and the study sample. Columns report means and standard errors for each group, with pairwise t-tests comparing each UNHS sample to the study sample. The UNHS sample is reweighted to match the study sample on age and gender using entropy balancing. Asterisks denote * $p < .10$, ** $p < .05$, and *** $p < .01$.

Table OA.3: Network Characteristics

Variable	N	Mean/(SE)	N	Mean/(SE)	N	Mean/(SE)	N	P-value
Age	285	23.737 (.204)	90	22.367 (.259)	195	24.369 (.261)	285	.000***
Friend	285	.874 (.020)	90	.878 (.035)	195	.872 (.024)	285	.888
Sibling	285	.060 (.014)	90	.078 (.028)	195	.051 (.016)	285	.382
Completed post-secondary education	282	.727 (.027)	90	.678 (.050)	192	.750 (.031)	282	.206
Wage-employed	267	.449 (.030)	84	.369 (.053)	183	.486 (.037)	267	.074*
Employed	267	.551 (.030)	84	.488 (.055)	183	.579 (.037)	267	.166
Work experience (mo.)	254	34.453 (3.335)	75	22.773 (2.462)	179	39.346 (4.574)	254	.023**
Work experience in training area (mo.)	257	28.681 (3.209)	75	17.600 (2.294)	182	33.247 (4.391)	257	.026**

Notes: This table reports characteristics of network members in the study sample. Columns (1)–(3) present, respectively, the number of observations, means, and standard errors for the full sample, female network members, and male network members. The final two columns report sample size and p-values from pairwise t-tests comparing female and male network members. Work experience variables are measured in months. Asterisks denote * $p < .10$, ** $p < .05$, and *** $p < .01$.

Table OA.4: Candidate Profile Balance

Variable	N	Mean/(SE)	N	Mean/(SE)	N	Mean/(SE)	N	P-value
Demographics								
Jinja (not Kampala)	1494	.778	747	.778	747	.778	1494	1.000
		(.011)		(.015)		(.015)		
High Experience	1494	.501	747	.505	747	.498	1494	.796
		(.013)		(.018)		(.018)		
High Schools								
Makerere College School (Kampala)	1494	.114	747	.118	747	.111	1494	.685
		(.008)		(.012)		(.012)		
Mengo Senior School (Kampala)	1494	.108	747	.104	747	.111	1494	.677
		(.008)		(.011)		(.012)		
St. Gonzaga Gonza SSS (Jinja)	1494	.368	747	.369	747	.367	1494	.915
		(.012)		(.018)		(.018)		
Pilkington College Muguluka (Jinja)	1494	.410	747	.408	747	.411	1494	.916
		(.013)		(.018)		(.018)		
Firms								
Aunt Hannah Fashions	1494	.062	747	.064	747	.060	1494	.748
		(.006)		(.009)		(.009)		
Bawumaye Metal Fabrication	1494	.022	747	.024	747	.020	1494	.598
		(.004)		(.006)		(.005)		
Bawumaye Plumbers	1494	.066	747	.060	747	.071	1494	.403
		(.006)		(.009)		(.009)		
Bosco Electricals	1494	.107	747	.110	747	.104	1494	.738
		(.008)		(.011)		(.011)		
Bosco Workshop	1494	.012	747	.013	747	.011	1494	.636
		(.003)		(.004)		(.004)		
Escape and Rest Hotel	1494	.042	747	.044	747	.040	1494	.700
		(.005)		(.008)		(.007)		
Gateway Hotel	1494	.044	747	.041	747	.046	1494	.704
		(.005)		(.007)		(.008)		
Gifted Hands School College	1494	.016	747	.017	747	.015	1494	.681
		(.003)		(.005)		(.004)		
Joanna's Salon	1494	.011	747	.009	747	.013	1494	.465
		(.003)		(.004)		(.004)		
Kalongosi Seed School	1494	.025	747	.024	747	.027	1494	.743
		(.004)		(.006)		(.006)		
Kataba Garage	1494	.086	747	.086	747	.086	1494	1.000
		(.007)		(.010)		(.010)		
Kataba Metal Works Fabricators	1494	.009	747	.007	747	.011	1494	.404
		(.002)		(.003)		(.004)		
Kataba Plumbing Solutions	1494	.064	747	.070	747	.059	1494	.399
		(.006)		(.009)		(.009)		
Katana and Sons Wiring Contractors	1494	.102	747	.099	747	.104	1494	.732
		(.008)		(.011)		(.011)		
Ketty's Tailors	1494	.056	747	.054	747	.058	1494	.735
		(.006)		(.008)		(.009)		
Kitagata General Workshop	1494	.009	747	.008	747	.011	1494	.592
		(.002)		(.003)		(.004)		
M&A Auto Garage	1494	.091	747	.091	747	.091	1494	1.000
		(.007)		(.011)		(.011)		
New Kabale Junior School	1494	.042	747	.046	747	.039	1494	.520
		(.005)		(.008)		(.007)		
Roko construction company	1494	.044	747	.040	747	.047	1494	.526
		(.005)		(.007)		(.008)		
Soleco Construction Company	1494	.046	747	.050	747	.043	1494	.538
		(.005)		(.008)		(.007)		
Southridge Primary School	1494	.031	747	.028	747	.035	1494	.459
		(.005)		(.006)		(.007)		
Tina's Beauty Salon	1494	.013	747	.015	747	.011	1494	.489
		(.003)		(.004)		(.004)		

Notes: This table reports balance in observable characteristics across all CV profiles. Column (1) shows, for the full CV sample, the share of profiles in each category (with standard errors in parentheses); columns (2) and (3) report the corresponding shares separately for male and female CVs. The "Pairwise t-test" column reports the number of observations and the p-value from a test of equality in these shares between male and female CVs. Demographic rows indicate CV-level characteristics, the "High Schools" panel lists whether each school appears on the CV, and the "Firms" panel indicates whether each firm is referenced as a current or previous employer.