LIVIA ALFONSI

Assistant Professor, Harvard Business School
Soldiers Field, Boston, MA 02163 | lalfonsi@hbs.edu

August 23, 2026

Emily Breza, Davide Viviano, and Arun Chandrasekhar
Workshop Organizers
NBER Workshop on Development Economics and Networks for Early-Career Researchers

Dear Arun, Davide and Emily:

I am writing to apply to the Workshop on Development Economics and Networks for Early-Career Researchers, to be held at the NBER on December 5, 2026. I am an Assistant Professor at Harvard Business School, and my research examines how social connections, beliefs, and workplace institutions shape labor-market matching and gender inequality in low- and middle-income countries.

I would be excited to present my working paper, "Who You Know, Who You Recommend: Referrals and Occupational Gender Segregation," coauthored with Pedro Ferreira. The paper combines a high-stakes referral experiment in Uganda with a dynamic network model of job finding. It separates two mechanisms through which referral systems can reproduce occupational segregation: gendered access to qualified candidates and gendered referral choices conditional on access. We find a substantial referral penalty for gender-counter-stereotypical candidates and show that this choice margin meaningfully amplifies segregation in steady state.

The workshop is especially well matched to the direction of my research. In complementary work, "Meet Your Future," I study mentorship as a deliberately created weak tie that changes how young workers interpret labor-market information and make early-career decisions. Taken together, these projects motivate a broader agenda on when networks improve matching and when they instead transmit biased beliefs or reinforce exclusion. I am particularly interested in developing interventions that preserve the informational advantages of referrals while reducing the distortions they create.

I would value the opportunity to receive feedback from researchers working at the frontier of empirical network research and to engage closely with the other early-career participants. If I am shortlisted, Professor Jeremy Magruder (jmagruder@berkeley.edu) can provide a recommendation.

Thank you for considering my application.

Sincerely,

Livia Alfonsi

MOTIVATION AND RESEARCH INTERESTS

Livia Alfonsi

My research studies how networks, beliefs, and workplace institutions shape labor-market matching in low- and middle-income countries, with particular attention to young workers and women. Across my work, I ask when social connections help workers navigate thin and uncertain labor markets, when they instead reproduce existing inequalities, and how network-based institutions can be redesigned to allocate talent more efficiently. I address these questions through large-scale field experiments, original longitudinal data, and partnerships with workers, firms, training institutions, and online labor-market platforms.

My paper "Meet Your Future: Experimental Evidence on the Labor Market Effects of Mentors", currently R&R at the Journal of Political Economy, illustrates the potential of a deliberately created weak tie. In Uganda, we matched recent vocational graduates with somewhat older workers who had successfully navigated the same transition into employment. These mentors were socially close enough to provide credible, context-specific advice but sufficiently removed from participants' immediate circles to bring a different perspective on the labor market. Mentorship increased labor-force participation by 27 percent and earnings by 18 percent one year later, while also improving retention and promotion. The mechanism was not simply access to more contacts: mentors corrected overoptimistic expectations about early jobs and encouraged young workers to accept viable opportunities and persist through the first stages of their careers. The project led me to think of networks not only as conduits for vacancies, but also as structures that determine whose information is credible and how workers form and revise beliefs.

My current working paper, "Who You Know, Who You Recommend: Referrals and Occupational Gender Segregation," coauthored with Pedro Ferreira, turns to a setting in which networks can reinforce inequality. Employee referrals are a major source of jobs, yet gendered referral patterns may reflect either whom workers know or whom they choose to recommend. We combine a high-stakes referral experiment with direct measurement of workers' networks to separate these margins. Among 753 skilled workers in Uganda, gender-noncongruent candidates are 23.4 percentage points less likely to be referred than otherwise identical candidates. The penalty reflects not only anticipated employer preferences, but also workers' own beliefs about candidates' ability, fit, and desirability as coworkers. We embed these estimates in a dynamic network model with homophilous ties and direct job arrivals and show that biased referral choices amplify steady-state occupational segregation by 7.8 to 9.2 percent, even under a conservative calibration.

I am now thinking about the next set of questions and interventions to pursue in this space. The workshop would be an ideal setting both to receive feedback on the referral paper and to think with other researchers about how this agenda could develop, including how referral systems might be redesigned to preserve their informational value while reducing biased screening and occupational segregation.

Who You Know, Who You Recommend: Referrals and Occupational Gender Segregation*

Livia Alfonsi
Harvard Business School

Pedro Ferreira
Harvard Kennedy School

14 Aug 2026

Abstract

Employee referrals are central to hiring, but the choice of whom to recommend is difficult to study because referral opportunities, feasible candidate sets, and workers' private information are rarely observed. We create a real referral opportunity and adapt audit and incentivized resume-rating methods to incumbent workers, who help determine which candidates reach employers. Across two experimental rounds with 1,308 skilled workers in Uganda, respondents refer candidates for six-week subsidized trials in their workplaces. Randomized profiles hold candidate availability fixed, while a second module elicits plausible referrals from respondents' own networks. This design separates whom workers know from whom they choose to recommend. In the main round, candidates whose gender does not match the occupation's majority are 23.4 percentage points less likely to be referred than otherwise comparable candidates. The penalty appears for both women entering male-dominated occupations and men entering female-dominated occupations. Network composition is also sharply gendered, yet unequal referral choices persist among workers who can name contacts of both genders. The largest evaluative penalties fall on incumbents' own willingness to work with candidates and perceived task ability, rather than only on anticipated employer preferences. Experience improves evaluations of men entering female-dominated occupations more than those of women entering male-dominated occupations. A dynamic network model disciplined by the experiment shows that this referral choice margin increases steady-state occupational segregation by 7.8 to 9.2 percent. Incumbents' preferences and beliefs therefore create an upstream filter, beyond network composition, that can exclude qualified GNC candidates before employer evaluation and slow the translation of occupational training into job access.

*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, Thomas Le Barbanchon, Elisa Macchi, Jeremy Magruder, Ted Miguel, Dev Patel, Imran Rasul, 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 Harvard and UC Berkeley. We used Claude Code (Anthropic) and ChatGPT (OpenAI) to assist with code drafting and debugging, manuscript organization, and editorial refinement. All analyses, interpretations, errors, and omissions are our own.

1 Introduction

Employee referrals are one of the principal ways workers and firms find one another (Topa, 2011a; Beaman and Magruder, 2012; Beaman et al., 2018; Fafchamps and Moradi, 2015; Witte, 2025). They can improve matching by transmitting information that is difficult to obtain through formal hiring channels. Incumbent workers may know who can perform the job, while job seekers may learn through their contacts which firms are hiring and what the work entails (Jovanovic, 1979; Montgomery, 1991; Calvo-Armengol and Jackson, 2004; Mortensen and Vishwanath, 1994; Galenianos, 2021). Referrers may also have reputational incentives to recommend candidates who will succeed (Pallais and Sands, 2016; Heath, 2018). But referrals do more than transmit information. They place an economically important allocation decision in the hands of incumbent workers: whom to bring to the employer's attention in the first place.

This upstream decision differs from the employer's hiring decision. An incumbent may care about whether the candidate will perform well and be retained, but also about whether the candidate would be a desirable coworker, how the employer will react to the recommendation, and whether referring a friend or relative creates social benefits. Incumbents therefore possess information and face incentives that differ from those of employers, while exercising substantial influence over the pool from which employers choose. A candidate whom no worker recommends may never reach the stage at which employer preferences can be observed.

Despite its importance, the referral choice is exceptionally difficult to study. Referral opportunities often arise outside posted hiring processes, so researchers may not know that a decision occurred. Even when a referral is observed, the feasible choice set is typically hidden. Observational data rarely reveal everyone the worker could plausibly have recommended, what the worker knew about each candidate, how each candidate was evaluated, or why one was selected over another. Referred and non-referred candidates may also differ in private information, job fit, social proximity, and willingness to accept the job. These features make it difficult to determine whether observed referral patterns reflect who workers know, what they know about those people, or how they choose among available candidates. This paper develops a high-stakes experimental approach to studying the causes and consequences of referral decisions. We create a real referral opportunity and adapt audit and incentivized resume-rating methods to a new decision-maker, the incumbent worker, and a new choice, the referral. Across two experimental rounds, 1,308 skilled urban workers in Uganda were invited to help fill a six-week subsidized trial at a workplace in their occupation. The choice carried meaningful stakes. Employers and candidates received subsidies if the placement was implemented, and the incumbent received UGX 100,000, approximately five days of wages, if the referred candidate was retained after the trial.

We use this design to study occupational gender segregation, a setting in which the identities of those who enter referral-based hiring pools may be particularly consequential. Occupations remain sharply segregated by gender, with implications for earnings, productivity, and the allocation of talent (Hsieh et al., 2019; Goldberg et al., 2025). Existing work largely examines why men and women enter different educational and occupational pipelines

or how employers evaluate applicants once they reach the hiring pool.[1] Referrals create an additional, earlier margin: incumbent workers help determine whose qualifications enter the pool at all. A referral system can disadvantage workers whose gender does not match the occupation's majority through two distinct channels. We call candidates whose gender matches the occupation's majority gender-congruent (GC), and those whose gender does not gender non-congruent (GNC). First, incumbents may not know qualified GNC candidates. This is a network composition channel. Second, even when qualified GNC candidates are available, incumbents may choose not to recommend them. This is a referral choice margin. The distinction matters because the two mechanisms imply different barriers. Expanding training pipelines and cross-gender professional ties may ease network constraints, but access alone will not eliminate a referral gap generated by differential evaluation among candidates already available to the worker.

Uganda's urban vocational labor market provides a revealing setting in which to study these channels. Women are concentrated in female-dominated occupations (FDOs), including tailoring, hairdressing, food and hospitality, teaching, and secretarial work. Men dominate male-dominated occupations (MDOs), including electrical work, motor mechanics, plumbing, welding, construction, and carpentry. Personal connections are central to job finding: 63 percent of respondents obtained their first job through a personal connection, including 23 percent through an incumbent worker. Past referrals also track occupational gender composition. Among workers in MDOs, 83 percent of prior referrals were to men; among workers in FDOs, 77 percent were to women. The experiment proceeds in two parts designed to separate the two referral margins.

In Part I, each respondent evaluates two experimentally constructed candidate profiles for the relevant occupation. The profiles represent qualified young vocational graduates and are comparable on the credentials central to this labor market. We cross-randomize candidate gender and post-training work experience. Because both candidates are presented to the respondent, Part I holds candidate availability fixed. It identifies whether GNC candidates are evaluated and recommended differently even when a qualified candidate is placed directly in the worker's choice set.

In Part II, respondents nominate real contacts from their own networks who could plausibly fill the same position, with an explicit prompt to name a contact of the other gender when possible. We collect a structured profile for each nominee, elicit the same evaluations as in Part I, and ask respondents to make a final choice from the expanded set of experimental and network candidates. Part II is descriptive because real contacts differ in observable and unobservable ways. Its value is to measure the network composition constraint, observe choices among people respondents actually know, and assess whether the evaluative logic isolated experimentally persists when social ties and candidate-specific information are restored. Together, the two parts distinguish whom workers know from whom they choose to recommend.

We also elicit the considerations entering the referral choice. For each experimental and network candidate, respondents report how much they would like to work with the

[1]For examples across this literature, see 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).

candidate and the probability that the candidate would be retained. We unpack perceived retention into three beliefs: whether the candidate can perform the relevant tasks, how much the employer would value the candidate, and how much the candidate would enjoy the job. Randomized experience allows us to examine how respondents update these evaluations when given a stronger signal of occupation-specific preparation. A vignette experiment and a randomization of referral confidentiality test whether GNC candidates are avoided because recommending them creates greater reputational or conformity costs. We also measure confidence in retention predictions to assess whether respondents penalize GNC candidates because they are less certain about their likely performance. The design therefore moves beyond observing the final referral and opens up the decision that produces it.

The first result is that both referral margins matter. In Part I, GNC candidates are 23.4 percentage points less likely to be referred than otherwise comparable GC candidates. The same penalty appears in an independently elicited ranking later in the survey. Because candidate gender is randomized and qualified candidates are made available by design, this cannot be explained incumbents lacking access to a candidate of the underrepresented gender. The penalty also becomes more pronounced as occupations become more gender-segregated.

Part II shows that access is nevertheless an important constraint in practice. Among respondents in FDOs who name at least one contact, 85 percent can name a woman but only 31 percent can name a man. In MDOs, 93 percent can name a man but only 22 percent can name a woman. Yet network composition is not the whole explanation. Among respondents able to name contacts of both genders, GNC contacts are 26.2 percentage points less likely to be referred and 30.8 percentage points less likely to be ranked above the GC contact. GNC workers therefore face a double barrier: they are less likely to be represented in incumbents' networks and less likely to be selected when available.

The mechanism measures show that the referral penalty contains a substantial incumbent-side component. GNC profiles receive lower ratings on both top-level margins: respondents' willingness to work with the candidate and perceived retention. When retention is unpacked, the largest gap is perceived task ability. GNC candidates are rated .980 points lower on expected task performance, compared with .438 points lower on anticipated employer valuation and .402 points lower on anticipated candidate enjoyment. The largest penalties therefore arise not only in respondents' beliefs about what employers or candidates will prefer, but in respondents' own willingness to work with the candidate and their own assessment of whether the candidate can do the job. Incumbents are not merely passing anticipated employer preferences through the referral process; they add an evaluative filter of their own. The evidence also provides little support for strategic competition: respondents view GNC candidates as less able, rather than unusually strong, and the penalty is no larger among shorter-tenured incumbents.

The same evaluations predict referral choices when real contacts enter the choice set. Likability and perceived retention strongly predict choices between experimental profiles, and likability remains strongly associated with choice among network candidates. The interactions between candidate ratings and network-contact status are small and imprecise, so we cannot reject that respondents place similar weight on these dimensions when evaluating experimental profiles and people they actually know. Relationship-specific information and social considerations matter in Part II, but do not appear to replace the

evaluative margins isolated experimentally in Part I.

The design deliberately spans both directions of occupational non-congruence. Men considered for positions in FDOs face a 17.2 percentage-point referral penalty, while women considered for positions in MDOs face a 26.7 percentage-point penalty. Both estimates are large, although their difference is not statistically significant. The evaluative content of the two penalties, however, differs sharply. Women in MDOs face broad penalties in likability, expected retention, task ability, anticipated employer valuation, and expected enjoyment. Men in FDOs face a sizable referral penalty but much smaller differences on the measured preference and belief margins, including no statistically significant likability penalty. This comparison requires an important qualification. The occupations in our sample are strongly gender-dominated, and 91 percent of respondents are themselves GC, meaning women in FDOs or men in MDOs. Candidate gender is randomized within respondent, so the GC–GNC penalty within each occupational context is experimentally identified and cannot be attributed to respondent characteristics. The difference between FDO and MDO estimates, however, also largely compares female and male incumbents and therefore cannot cleanly separate occupational context from respondent gender. Among the 9 percent of respondents who are themselves GNC, the point estimates do not indicate a systematic reversal or consistently different referral behavior. Because this group is small, however, the corresponding estimates are imprecise and should be treated as suggestive.

Randomized experience further reveals that similar referral penalties can rest on different beliefs. For male candidates in FDOs, additional experience raises perceived ability and anticipated employer valuation to or above the less-experienced GC baseline. Likability and perceived retention also move above that benchmark. For female candidates in MDOs, additional experience brings anticipated employer valuation approximately to parity, but leaves respondents' own assessments substantially behind. Experienced women remain rated .806 points less able to perform the core tasks and .407 points less likely to enjoy the work than less-experienced GC men. Respondents therefore update their beliefs about how employers will view experienced women in MDOs more than they update their own beliefs about whether those women can and want to do the job.

Open-ended candidate descriptions provide a complementary view of this asymmetry. Respondents supplied nearly 5,000 adjectives describing the experimental profiles. In MDOs, female candidates are more often described through gender and age markers, negative descriptors, and communal or grindstone terms emphasizing social qualities and effort. Competence-related language, including skill, cognitive ability, education, and strength, tilts toward male candidates. These patterns are much weaker in FDOs and do not simply reverse. The closed-ended rating gaps therefore reflect a broader difference in which attributes incumbents make salient when evaluating otherwise comparable candidates.

Several pieces of evidence make it difficult to interpret the perceived-ability penalty as an accurate response to a large underlying productivity gap. In a separate longitudinal sample of Ugandan vocational graduates, GC and GNC workers are statistically indistinguishable on purpose-built theoretical and practical skill assessments, wage employment, earnings conditional on working in their training sector, and career satisfaction. Nor are earnings gaps concentrated in the MDOs where physical demands should be most binding. These comparisons cannot eliminate the possibility of productivity differences on unmeasured

dimensions, but they provide little affirmative evidence that the perceived-ability penalty reflects measured differences in skill or subsequent performance. Nor does the pattern appear to be driven by greater uncertainty about GNC candidates. Respondents report nearly identical confidence in retention predictions for GC and GNC profiles, and the small confidence difference appears only for men considered for FDO positions rather than women considered for MDO positions, for whom the ability penalty is largest.

Reputational and conformity concerns explain only a limited portion of the penalty. A vignette experiment finds no larger perceived reputational cost when a GNC rather than a GC referral underperforms. The confidentiality randomization directly tests whether observability drives the GNC penalty. Making referrals private only modestly reduces the penalty, which remains large and significant, and the attenuation is not statistically significant within either occupation type. Being observed may therefore affect referral behavior at the margin, but does not account for most of the referral gap.

Finally, we examine the consequences of referral choice for aggregate occupational segregation. We embed the experimentally estimated GNC penalty in a dynamic network model of job finding, building on Calvo-Armengol and Jackson (2004). The model allows segregation to arise through three channels: gender- and occupation-specific direct access to job signals, gender-homophilous networks, and gendered referral weights disciplined by the experiment. We calibrate network size and homophily using our survey data and estimate direct job-arrival rates by simulated method of moments to match observed employment shares. The model separates the role of referral choice from other forces producing occupational segregation. Differential direct job access generates substantial segregation even without referrals, absorbing supply- and demand-side barriers outside the modeled referral process. Homophily amplifies existing imbalances by concentrating job information within gendered networks. The GNC referral penalty creates an additional force because it directs signals for FDO jobs disproportionately toward women and signals for MDO jobs disproportionately toward men. Our main counterfactual adds the experimentally estimated referral penalty to an economy with differential direct job access and gendered network composition. In the conservative baseline calibration, the penalty increases steady-state occupational segregation by 9.2 percent in MDOs and 7.8 percent in FDOs, corresponding to approximately 6.3 and 5.1 percentage points. Its contribution rises when referrals play a larger role in job allocation, including when networks are denser or less homophilous and when direct job access is less segregated. Referral choice is therefore not the sole or dominant source of occupational segregation in the baseline calibration, but it is an economically meaningful amplifier.

Taken together, the findings show why the referral decision warrants study as an economic choice in its own right. Referrals can improve matching by transmitting valuable information, while simultaneously restricting access through gendered networks and incumbent evaluations. Workers are not passive conduits between candidates and employers. They are upstream gatekeepers with their own information, incentives, preferences, and beliefs. In the setting we study, those choices systematically direct job information toward GC candidates and help reproduce occupational segregation over time.

Contributions. First, this paper contributes to the literature on referrals and network-based job finding. Existing research documents the prevalence of referrals, the structure of

referral networks, who receives referrals, and how referred workers perform (Granovetter, 1995; Topa, 2011a; Beaman and Magruder, 2012; Beaman et al., 2018; Fafchamps and Moradi, 2015; Kramarz and Skans, 2014; Galenianos, 2021; Witte, 2025). We study the choice that precedes those outcomes: whom an incumbent recommends when a referral opportunity arises. Our design creates a real and consequential referral decision, experimentally constructs part of the candidate set, measures the composition of the worker's actual network, and elicits the preferences and beliefs associated with each candidate. In doing so, we adapt tools developed to study employer evaluation to an intermediary whose information and incentives differ from those of the employer, but whose choice determines who enters the employer's consideration set. Our setting also differs from work on strategic information sharing among competing job seekers (Chiplunkar et al., 2024): our referrers are already employed, evaluate recent graduates for an additional position, with a reward tied to retention.

Second, the paper contributes to the literature on occupational gender segregation. Prior work emphasizes supply-side barriers such as childcare, safety, social norms, preferences, and limited role models, and demand-side barriers such as employer discrimination and customer preferences (Adda et al., 2017; Agüero and Marks, 2011; Fogli and Veldkamp, 2011; Goldin, 2014a; Heath and Mobarak, 2015; Heath et al., 2024; Delfino, 2024; Bar and Zussman, 2017; Lepage et al., 2025). A growing literature studies access to hiring funnels and the channels through which qualifications reach employers (Kuhn and Shen, 2013; Buchmann et al., 2025; Olivetti et al., 2024; Bohren et al., 2025b). We identify referrals as an upstream access margin that can exclude qualified workers after they acquire relevant training and experience but before employers evaluate them. The bidirectional design shows that this margin operates for both men entering FDOs and women entering MDOs, while the mechanism evidence reveals that the content of the penalty is especially broad for women in MDOs. The dynamic model then quantifies how this individual referral choice can amplify occupational segregation in equilibrium. Third, the paper contributes to the emerging experimental literature on gendered referral behavior (Beaman et al., 2018; Beugnot and Peterlé, 2020; Hedegaard and Tyran, 2018b; Hederos et al., 2024; Zeltzer, 2020). Prior studies generally cannot observe both the potential referral set and the worker's choice within it, or focus on students, laboratory participants, generic short-term tasks, or gender-neutral occupations. We study incumbent workers making consequential recommendations into workplaces and occupations they know. By combining randomized candidate profiles, real network contacts, mechanism measures, randomized experience and privacy, and open-ended descriptions, we show not only that referral choices are gendered, but how the preferences and beliefs supporting those choices vary across occupational contexts.

2 Conceptual Framework

This section first reviews the economic functions and distributional consequences of referrals, then develops a parsimonious framework linking the referrer's decision to the measures elicited in our experiment.

2.1 Functions and Distributional Consequences of Referrals

A referral is a hiring channel in which access to an employment opportunity is mediated by a recommendation from someone connected to the employer, such as an incumbent employee, former colleague, or other professional contact.

Referral theory centers on two-sided information frictions. Firms cannot directly observe applicant productivity, and conventional hiring channels often provide noisy signals. Referrals can reduce this uncertainty by drawing on incumbents' private information (Montgomery, 1991), while referrers' workplace reputational stakes may further align their incentives with employers (Pallais and Sands, 2016; Heath, 2018).[2] Together, these information and incentive channels lower the firm's expected cost of a bad hire. Job seekers face parallel search frictions because vacancies and match quality are costly to discover. Personal connections reduce these frictions by transmitting information about openings, hiring timing, and match quality (Calvo-Armengol and Jackson, 2004; Mortensen and Vishwanath, 1994; Galenianos, 2021).[3]

These two-sided gains help explain why referrals are ubiquitous (Granovetter, 1995). In the United States, roughly half of workers report learning about jobs through friends or relatives, and estimates suggest that 30 to 60 percent of jobs are found through such contacts (Bewley, 1999; Ioannides and Loury, 2004). Many U.S. firms also monetarily incentivize employee referrals (CareerBuilder, 2012). Referrals may be especially important in low- and middle-income labor markets, where formal postings are scarce and personal connections often shape worker–firm matching. In our sample, 63 percent of vocational graduates obtained their first job through a personal connection (Section 3). These gains need not be equally distributed. Access can depend on a worker's network position, perceived ability, and the social identities associated with them (Calvo-Armengol and Jackson, 2004; Beaman et al., 2018; Galenianos, 2021; Bolte et al., 2026; Kramarz and Skans, 2014).

Yet, most of this existing work largely asks who receives referrals or how referred workers perform. We instead put the upstream decision that generates those outcomes at the center of the analysis: whom an incumbent chooses to recommend. Referral inequality may reflect network composition, because workers know different people, but it may also arise at the point of choice, when workers evaluate available candidates differently. Conditional on access to a plausible candidate, referral decisions may reflect the incumbent's own preferences, beliefs about the candidate's ability and fit, expectations about how the employer will evaluate the candidate, and, for network contacts, social benefits from the relationship. This distinction matters because incumbents help determine who enters the hiring funnel; their choices can therefore exclude qualified candidates before employers ever evaluate them. The framework below formalizes these considerations and maps them into the measures

[2]Burks et al. (2015), using personnel data from nine large firms in three industries, document that referred workers have similar baseline skills but lower turnover, fewer accidents in trucking, and more patents in high-tech, generating higher profits through lower recruiting and turnover costs. Related studies document productivity gains, lower attrition, and longer tenure, especially early in the employment relationship (Castilla, 2005; Hensvik and Skans, 2016; Brown et al., 2016; Dustmann et al., 2016; Friebel et al., 2023; Barr et al., 2019).

[3]Early within-firm studies find that referred applicants are more likely to be hired and have better post-hire outcomes (Fernandez and Weinberg, 1997; Fernandez et al., 2000; Petersen et al., 2000; Yakubovich and Lup, 2006). Topa (2011b) and Ioannides and Loury (2004) review the broader evidence.

elicited in our experiment.

2.2 Inside the Referrer's Decision

The worker's decision is whom to refer. We organize it around three sources of value. First, the referrer may value recommending a candidate who is expected to perform well during the trial and be retained afterward, because a successful referral may generate both monetary rewards and reputational returns with the employer (Pallais and Sands, 2016). Second, because the candidate may become a coworker the referrer may care about the workplace benefits of working with the candidate. The referrer may therefore value bringing in someone familiar, cooperative, that they would like to work with (Donald and Grosset, 2024; Emanuel et al., 2023). Third, referring a socially close candidate may generate benefits through altruism, reciprocity, repayment of favors, or future exchange (Beaman and Magruder, 2012).

We formalize these considerations as follows. Each respondent i chooses between candidates $j \in \{1, 2\}$. The value of referring candidate j is

$$V_i(j) \;=\; b \cdot E_i[P(R_j)] \;+\; L_i(j) \;+\; \alpha_i(j), \tag{1}$$

where b is the payoff from a successful referral, including any financial reward tied to referral success and nonmonetary returns such as employer trust and future consideration. $E_i[P(R_j)]$ is respondent i's perceived probability that candidate j is retained after the relevant probationary period.[4],[5] $L_i(j)$ is the direct amenity value from having j as a coworker; and $\alpha_i(j)$ captures social payoffs such as altruism, reciprocity, favors, and network obligations. Respondent i refers candidate 1 if $V_i(1) > V_i(2)$.

We model expected referral success as a function of three perceived candidate attributes:

$$E_i[P(R_j)] \;=\; \phi\Big(E_i[Q_j],\, E_i[V_e(j)],\, E_i[J_j]\Big), \tag{2}$$

where $E_i[Q_j]$ is perceived task ability to complete the required tasks, $E_i[V_e(j)]$ anticipated employer valuation, and $E_i[J_j]$ anticipated candidate enjoyment. We assume $\phi(\cdot)$ is weakly increasing in each argument and otherwise unrestricted: candidates expected to perform well, be valued by the employer, and enjoy the job should be perceived as more likely to complete the trial and be retained.[6] A richer model could allow referrers to perceive different degrees of uncertainty across candidates and to respond to that uncertainty when making referrals, even when expected retention probabilities are identical. We abstract from this channel because we find no systematic differences across candidate types in neither profile recall nor respondents' reported confidence in their retention predictions (see Online

[4]If the reputational gains (costs) from successful (failed) referrals varied by candidate type, the value function would require an additional candidate-type-specific term. An in-survey vignette experiment we run finds no evidence that referring an underperforming GNC candidate carries greater reputational costs than referring an underperforming GC one. See Online Appendix Figure A.7 for more detail.

[5]In our experiment, incumbents were offered a UGX 100,000 bonus if the referred candidate was retained after the six-week subsidized trial.

[6]These objects are conceptually distinct but may covary; for example, a more capable candidate may also be a more desirable coworker. We interpret each measure as the dominant signal in its category.

Appendix E for more details). Finally, we interpret employer valuation as combining the employer's productivity assessment and likability of candidate j: $V_e(j) = g\Big(Q_j, L_e(j)\Big)$.

3 Context and Sample

3.1 Gender Segregation in Uganda's Labor Market

Women's labor force participation in Sub-Saharan Africa is relatively high, at 57 percent, compared with 39 percent in South Asia and 52 percent in Latin America. Yet women remain concentrated in lower-paying segments of the economy (World Bank, 2024; Asongu and Odhiambo, 2019). Uganda reflects this broader pattern. Data from the 2021 National Labor Force Survey show substantial occupational gender segregation: women are more concentrated in agriculture, while men are more represented in higher-paying industry and services (Uganda Bureau of Statistics, 2021). Segregation is also pronounced within vocational occupations, which account for more than half of urban main employment.[7] Women are concentrated in lower-paying occupations such as cosmetology, tailoring, and caregiving, whereas men dominate higher-paying fields such as electrical work and automotive repair (Alfonsi et al., 2020, 2023, 2026a).

3.2 A Sample of Urban Workers

Our study draws on two separate rounds of the same field experiment, implemented with 1308 urban workers across Central and Eastern Uganda.[8] Round 1 ($N = 555$) included the core experimental design and outcome measures. Round 2 ($N = 753$), fielded with a comparable population, retained these elements while adding random variation in candidate experience and richer mechanism measures, including second-order beliefs. We use the Round 2 sample as our main analysis sample because of its broader scope and Round 1 to assess the robustness of the main patterns across rounds. Respondents completed post-secondary vocational education at established vocational training institutes (VTIs), earning National Certificates in manufacturing and service occupations. Appendix Table A.1 and Online Appendix Table OA.7 report summary statistics for both samples. Using the female employment share in the corresponding occupation in the 2016/17 Uganda National Household Survey, we classify occupations with shares below 50 percent as MDOs and those with shares at or above 50 percent as FDOs. MDOs include electrical work, motor mechanics, construction, plumbing, welding, carpentry, and machining and fitting; FDOs include food and hospitality, tailoring, primary and pre-primary teaching, hairdressing, and secretarial/accounting work. In our sample, 34 percent of workers were trained in FDOs. Our sample reflects Uganda's skilled urban workforce. Relative to young workers

[7] Although "vocational occupations" are not a single category in national labor-force statistics, the broad occupational groups closest to our setting are service and sales workers, craft and related trades workers, and plant and machine operators and assemblers. In the 2021 NLFS, these groups accounted for 56.6 percent of urban main employment. Their combined share was similarly large in earlier survey rounds.

[8] Both rounds were pre-registered in detail on the AEA RCT Registry under separate entries (AEARCTR-0008036; AEARCTR-0012908). Online Appendix E reports pre-specified analyses not included in the main paper.

nationally, respondents are more educated, earn more, and are more likely to work outside agriculture; relative to nationally representative vocational graduates, they have similar non-agricultural employment and earnings.[9] The sample is therefore well suited to studying access to better urban jobs and the role referrals may play in shaping mobility into non-agricultural employment. For a small set of ancillary analyses we also draw on a separate longitudinal sample from Alfonsi et al. (2020), henceforth the ECMA sample. The study follows an earlier cohort of vocational graduates from Central Uganda for roughly a decade and contains harmonized information on job-finding channels, skills, labor-market outcomes, and the tasks workers perform.

3.3 Referral Use and Patterns

In our sample, 63 percent of respondents obtained their first job through a personal connection: 33 percent through a nonemployee referral, 23 percent through an incumbent worker, and 7 percent through a direct employer connection. Another 36 percent found their first job through walk-ins and 2 percent through other channels, including media, the internet, middlemen, and NGO placements. Referral-based job finding is similarly prevalent in the ECMA sample: 61 percent of first jobs are obtained through a personal connection, with a nearly identical split across referral channels (22 percent non-employee, 12 percent incumbent-worker) and comparable rates of walk-ins (36 percent) and other channels (8 percent). The consistency across cohorts underscores that referrals are a persistent feature of early-career job finding in Uganda. Referring others is also common among our workers: about 34 percent of our respondents report making at least one job referral in the previous twelve months. About 10 percent involved relatives, 80 percent friends, and the remaining 10 percent friends-of-friends, relatives of friends, or coworkers' contacts. Referrals therefore operate mainly through extended social networks rather than kinship. Past referral behavior is also strongly aligned with occupational gender composition. 83 percent of referrals made by workers in MDOs were to men, and 77 percent of those made by workers in FDOs were to women.

These facts motivate our central question: do these referral patterns reflect constraints in network composition, differential evaluation of GNC candidates, or both? The distinction matters because the remedies differ: expanding cross-gender occupational networks may ease composition constraints, but if disparities arise from differential evaluation, access alone will not eliminate the referral gap and interventions targeting how referrers assess available candidates may also be necessary.

4 Experimental Design

4.1 The Incentivized Referral Task

The referral choice is empirically difficult to study. Referral opportunities often arise outside posted hiring processes, so researchers rarely observe either that a decision occurred or

[9] Online Appendix Table OA.8 compares our sample with nationally representative young workers and vocational graduates from the Uganda National Household Survey.

the relevant choice set. Fully understanding the decision would require reconstructing the worker's complete set of plausible candidates and, for each one, observing measurable characteristics, relationship-specific information, and beliefs about the candidate's ability and fit, preferences over the job, and likely employer valuation.

We address these challenges with two design choices. First, we create a real referral opportunity: respondents were invited to refer a candidate for a six-week subsidized trial at their workplace, allocated by lottery. If selected, the employer would receive UGX 200,000 to host the trial, the candidate UGX 300,000, and, importantly, the incumbent employee UGX 100,000 if the candidate was retained afterward. These amounts correspond to approximately ten days of wages for the employer subsidy, six weeks of pay for the candidate, and five days of wages for the retention bonus.[10] Second, we adapt audit and incentivized resume-rating methods to a new decision-maker, the incumbent worker, and a new choice, the referral.[11] Referral choices were high-stakes and linked a lottery through which we attempted to create 178 placements; Online Appendix A.7 reports the implementation outcomes.

The task proceeds in two parts (Figure 1). In Part I, each respondent evaluates two experimentally constructed profiles on perceived task ability, $E_i[Q_j]$; likability, $L_i(j)$; perceived retention probability, $E_i[P(R_j)]$; anticipated employer valuation, $E_i[V_e(j)]$; and anticipated candidate enjoyment, $E_i[J_j]$. After evaluating both profiles, the respondent chooses one to refer.[12] Part II adds respondents' own contacts to the choice set, allowing us to study referrals of actual network members and measure network composition. Respondents first nominate someone who could plausibly fill the trial position. For each contact, we collect a structured profile containing demographics, education, and prior work experience, and elicit the same five evaluations as in Part I. If the first nominee is of one gender, we ask whether the respondent can name an additional contact of the other gender. Respondents therefore evaluate two experimental profiles and, when feasible, up to two network contacts, before making a final referral from the expanded set of two to four candidates.

4.2 Candidate Profiles and Randomization

Part I profiles follow CV templates commonly used in Ugandan labor markets and collected through our prior work in the setting. Each contains personal information (name, sex, date of birth, nationality, and home district), academic background, work experience (firm, role, duration, and main activities), and language abilities. Appendix Figure A.1 shows a real CV and an experimental profile. Profiles are designed to be realistic and comparable, representing young vocational graduates with similar training backgrounds.

[10] A minority of respondents were unemployed at the time of the survey. Although not currently employed, they were skilled vocational graduates with substantial prior labor-market experience and remained attached to the labor market. They completed a parallel version of the module in which they selected a candidate for a vacancy at a firm that typically hires graduates from their training track. Online Appendix F.1 and Table OA.11 show that excluding these respondents does not affect the results.

[11] Online Appendix A.2 reviews these literatures and relates our methodology to the classic resume-correspondence design.

[12] As an additional robustness check, we randomized whether anticipated employer valuation and candidate enjoyment were elicited alongside likability and retention or after respondents had evaluated both profiles and made their referral choice. Question placement did not affect subsequent profile ratings.

Like respondents, all candidates are described as holding a two-year post-secondary National Certificate, the nationally accredited vocational credential in this setting. We cross-randomize candidate gender and post-training experience, signaling gender through name and the sex field and varying experience in five-month increments. This design, combined with a sample deliberately spanning both female- and male-dominated occupations, allows us to test whether GNC candidates face a referral penalty, whether it differs between women in MDOs and men in FDOs, and whether experience attenuates it differently across occupation types. To preserve realism and avoid duplicate profiles, we vary non-focal fields such as exact date of birth within a narrow range, high school within a common quality tier, prior-job tasks, and firm names. Online Appendix A describes profile construction and randomization in full.[13]

Profile framing differed across rounds. Round 1 followed the deception-free incentivized resume-rating design of Kessler et al. (2019): respondents were told that the profiles were hypothetical but closely resembled candidates in the relevant labor market, and that their responses would be used to identify a real candidate with similar characteristics. In Round 2, respondents were told that the profiles corresponded to real candidates, although the research team experimentally assigned some non-substantive details. Results are materially identical across the two framings.[14]

4.3 Outcomes

We map the objects in Section 2.2 to the outcomes elicited in the experiment.

Primary outcomes: referral choice, likability, and perceived retention probability. We study three primary outcomes: referral choice, likability, and perceived retention.[15] The referral outcome equals one for the profile selected. At the end of the survey, respondents also rank all experimental and nominated network candidates in case their first choice is unavailable, yielding an ordinal preference measure that complements the binary referral outcome. Likability, $L_i(j)$, is measured on a 0–10 scale and captures the respondent's willingness to work with the candidate. This margin is especially relevant in our small-firm setting: the median business employs five workers, so respondents can expect substantial day-to-day interaction with a new hire. Perceived retention, $E_i[P(R_j)]$, is measured on a 0–100 scale and captures the candidate's expected success in the job.

[13]Online Appendix A.6 reports profile-delivery and recall checks. Without reviewing the profiles, respondents were asked to recall each candidate's months of experience. They reliably identified the more-experienced profile, with no differential recall by candidate gender or congruence status.

[14]We changed the framing because the hypothetical-profile formulation confused some respondents and lengthened the survey, and because the rarity of GNC profiles could lead respondents to doubt whether such candidates existed or were available. Round 2 therefore stated that the candidates existed, were interested, and were available to work at the respondent's shop. We viewed the remaining deception as limited because assigned details, such as exact age within a narrow range, high school within the same quality tier, and firm name, were non-substantive. Core characteristics, including completion of two years of vocational education at a specific institution, occupation of training, and the existence of post-training experience in that occupation, were real. These are also the dimensions employers in this setting weigh most heavily (Alfonsi et al., 2026b), which motivated the randomization of experience in our design.

[15]The two core evaluation questions follow Heilman et al. (2004): "How much would you like to work with this candidate?" and "What are the chances this candidate would be retained?"

Secondary outcomes: components of perceived retention. As outlined in Section 2.2, perceived retention probability may reflect the candidate's perceived task ability, $E_i[Q_j]$, as well as beliefs about the employer's valuation of the candidate, $E_i[V_e(j)]$, and the candidate's own expected enjoyment of the position, $E_i[J_j]$. We therefore use secondary outcomes to unpack the components of perceived retention probability. To unpack perceived retention, we measure perceived task ability, $E_i[Q_j]$; anticipated employer valuation, $E_i[V_e(j)]$; and anticipated candidate enjoyment, $E_i[J_j]$. Respondents first list the three tasks a candidate would most likely perform during a trial at their workplace and then rate, on a 0–10 scale, how well each candidate would perform them. They then report their second-order beliefs about how much the employer would value the candidate and how much the candidate would enjoy the job.

Together, these outcomes move beyond the binary choice typical of correspondence studies, which is especially important for referrals because employees recommend candidates while employers ultimately decide whether to retain them. They distinguish respondents' own willingness to work with a candidate from beliefs about task ability, employer valuation, and candidate enjoyment. Combined with the two-part design, they allow us to separate network composition from differential evaluation and assess how these margins respond to candidate experience.

4.4 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 first estimate the baseline gender non-congruent (GNC) referral penalty:

$$Y_{p,i} = \alpha + \beta_1 \, \mathbf{1}_p(\mathit{GNC}) + X_i'\gamma + \varepsilon_{p,i}, \tag{3}$$

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); $\mathbf{1}_p(\mathit{GNC})$ equals one if the profile is *gender non-congruent* for the occupation type (a female profile in a male-dominated occupation or a male profile in a female-dominated one), and zero otherwise; and X_i includes randomization-strata fixed effects and pre-specified respondent characteristics (including gender and employment status), as well as vocational training institute and training-area fixed effects.[16] In this baseline specification, β_1 captures the average GNC referral penalty.

We then introduce the experimental variation in experience to ask whether additional experience differentially attenuates or reinforces the penalty:

$$Y_{p,i} = \alpha + \beta_1 \, \mathbf{1}_p(\mathit{GNC}) + \beta_2 \, \mathbf{1}_p(\mathit{MoreExp}) + \beta_3 \Big[\mathbf{1}_p(\mathit{GNC}) \times \mathbf{1}_p(\mathit{MoreExp})\Big] + X_i'\gamma + \varepsilon_{p,i}, \tag{4}$$

where $\mathbf{1}_p(\mathit{MoreExp})$ equals one if the profile is assigned the higher experience level. In this specification, β_1 is the GNC penalty among less-experienced profiles, β_2 the experience premium for GC profiles, and β_3 the differential experience premium for GNC profiles. Two linear combinations summarize the experience margin: $\beta_2 + \beta_3$, the experience premium

[16] Tables 1 and 5 report binary-referral estimates without respondent-level controls, which are not separately identified because respondent characteristics are constant within each two-profile choice set.

for GNC profiles, and $\beta_1 + \beta_2 + \beta_3$, the net effect of a more-experienced GNC profile relative to a less-experienced GC profile. We report these linear combinations in the tables. We estimate equations (3) and (4) for the full sample and by occupation type (female- and male-dominated), clustering standard errors by respondent because each respondent evaluates two profiles. For Part II, we estimate analogous linear-probability specifications.

Demand effects. A potential concern is that the design may have increased the salience of gender: each respondent evaluated one male and one female experimental profile, and in these highly gender-segregated occupations one of the two candidates was necessarily gender non-congruent. We view this concern as limited for several reasons. First, profiles were presented sequentially in randomized order, and respondents completed all evaluations of the first candidate before seeing the second; we find no differences in evaluations by profile order. Second, any desire to appear nondiscriminatory would likely attenuate, rather than generate, the GNC penalty. The large penalties across referral choices and evaluation outcomes are therefore unlikely to be mechanically driven by experimenter demand. Third, profiles were presented sequentially in randomized order, and respondents completed all evaluations of the first candidate before seeing the second; we find no evidence of a difference in evaluations by profile order.

5 Results

5.1 The Gender Non-Congruent Referral Penalty

Table 1 reports our central empirical fact. Columns 1–2 use the Part I data, in which each respondent evaluates two hypothetical candidate profiles that are virtually identical on observables except for gender and randomized experience. Column 1 reports linear-probability estimates of the primary referral choice: GNC candidates are 23.4 percentage points less likely to be referred than otherwise identical GC candidates ($p < .001$). The same pattern appears in a soft-preference ranking elicited after Part II (column 2), which agrees with the primary referral choice on more than 95 percent of profiles: on this independent later measure the GNC candidate is ranked below the GC candidate at a rate close to that of column 1. That the same penalty appears in an independent later elicitation, with different framing, is a first indication that the referral choice is not an artifact of the specific wording of the primary elicitation.

Columns 3–5 corroborate the same fact using real, self-nominated network members. Each respondent is asked to name up to two candidates from their own network who could plausibly fill the internship position, prompted after an initial nomination to also name a contact of the opposite gender. Column 3 reports the naming margin: respondents are substantially less likely to name a GNC contact than a GC contact.[17] Columns 4–5 restrict

[17]The two rounds cannot be pooled symmetrically across Parts I and II. Part I in Round 1 used a simpler design with only two of the four treatment arms fielded in Round 2, so the Part I samples are not stackable; on the two arms common to both rounds, the GC–GNC gaps are essentially identical (Figure A.3, Panel A). Part II, by contrast, was administered identically in Round 1 and Round 2, which is why we pool the two rounds. We pool for power: not every respondent named a network, and even fewer named contacts of both genders, so the both-genders subsample used in columns 4–5 is small. Appendix Table A.2 reports,

attention to the subsample of respondents who succeed in naming contacts of both genders and therefore have an unambiguous binary choice. Conditional on knowing at least one contact of each gender, GNC contacts are 26.2 percentage points less likely to be referred (column 4) and 30.8 percentage points less likely to be ranked above the GC contact (column 5). Because Part II holds constant, at the level of the respondent, the set of people the respondent knows, the column 4 and column 5 gaps are not driven by differences in who is available: they are choices made within the respondent's own network.[18]

Appendix Figure A.2 plots the occupation-level Part I referral penalty against the share of GNC workers in the occupation. The relationship is upward sloping and approximately linear. The penalty is most negative in the most gender-segregated occupations, such as motor mechanics and welding at the male-dominated end and hairdressing and tailoring at the female-dominated end, and attenuates as occupation composition becomes more mixed. This dose-response pattern is consistent with the audit-based evidence in Riach and Rich (2006) and Booth and Leigh (2010) that employer gender bias tracks the gender dominance of the occupation.

5.2 Preferences, Beliefs, and Referral Choice

Preference- and belief-based evaluations in Part I. Guided by the framework in Section 2, Table 2 decomposes the Part I evaluation into two top-level channels: the respondent's own preference for working with the candidate (likability, L_i) and the respondent's belief about the candidate's expected success as a referral (perceived retention probability, $E_i[P(R_j)]$). We then unpack retention into its three belief subcomponents: perceived task ability ($E_i[Q_j]$), anticipated candidate enjoyment ($E_i[J_j]$), and anticipated employer valuation ($E_i[V_e(j)]$). Both top-level channels move against GNC candidates. Likability drops by .450 points on a 0–10 scale ($p < .001$), while perceived retention falls by 3.12 percentage points ($p < .001$). The two penalties are therefore of comparable standardized magnitude, indicating that the referral gap is not driven only by beliefs about whether GNC candidates will ultimately be retained; it also reflects respondents' own willingness to work with them.[19] Within the retention channel, perceived task ability shows the largest penalty (.980 points; $p < .001$), followed by anticipated enjoyment (.402 points; $p < .001$) and anticipated employer valuation (.438 points; $p < .001$). In standardized units, perceived task ability is the dominant retention subchannel. Appendix Figure A.4 complements these estimates by plotting the GNC-candidate coefficient from a sequence of referral OLS specifications that progressively add the elicited ratings, together with each specification's R^2. Two features stand out. First, the full set of ratings explains a substantial share of

separately for Round 1 and Round 2, the rates at which respondents named a network, named a GNC contact, and referred a GNC contact; the sign and rough magnitude of the gaps are preserved in each round.

[18]Gender bias may also operate at the naming stage: because respondents nominate contacts they consider plausible candidates, biased beliefs about GNC 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. The opposite-gender prompt after an initial nomination partially mitigates this concern by encouraging respondents to search beyond their first recall.

[19]Appendix Figure A.3 shows that the results for the two top-level channels are highly similar across Rounds 1 and 2. Because Round 1 did not separately elicit the three components of perceived retention, these subchannels cannot be compared across rounds.

variation in referral choice. Second, the GNC coefficient is attenuated unevenly across channels. Among the two top-level measures, likability appears to absorb slighly more of the gap than perceived retention and yields a higher R^2. In the second block, we hold likability fixed and add the three retention subcomponents separately. In this sequence, task ability produces the largest gain in R^2 and reduces the GNC coefficient to statistical insignificance. Adding anticipated enjoyment and employer valuation produces little further attenuation. The largest reductions in the GNC coefficient therefore occur after accounting for likability and perceived task ability, the two measures most closely reflecting the respondent's own evaluation. Employer valuation and candidate enjoyment account for less of the gap. This suggests that respondents' own preferences and beliefs, rather than only their expectations about others' preferences, contribute importantly to the GNC penalty.

Evidence of a common evaluative logic across Parts I and II. Part I isolates the evaluative component of the referral decision experimentally. Because the candidates are strangers to respondents, systematic differences in referral behavior cannot reflect private information about the specific individuals being evaluated. In the language of the conceptual framework, Part I shuts down the social-payoff term $\alpha_i(j)$ and greatly limits candidate-specific private information entering beliefs about Q, J, and V_e. Part II reintroduces these elements: respondents evaluate real network contacts, about whom they may hold relationship-specific information and with whom referrals may generate social benefits. Table 3 asks whether the same elicited evaluations that predict referral choice in Part I also remain associated with choice once real network contacts enter the choice set. We estimate conditional logits of referral choice on the elicited ratings, with respondent choice-set fixed effects. Columns 1–4 report choices between the two experimental profiles, first for the full Part I sample and then for the subset of respondents who also nominated at least one network contact. Likability and perceived retention both strongly predict referral choice (column 1: .732 and .426, respectively; both $p < .01$). When perceived retention is replaced by its three subcomponents, task ability, anticipated enjoyment, and anticipated employer valuation are each positively associated with referral choice, with task ability carrying the largest coefficient (.327, $p < .01$; column 3). Columns 5–6 repeat the exercise for Part II, when real network candidates enter the choice set. Likability remains positively and precisely associated with referral choice (.354, $p < .01$; column 5). Among the three retention subcomponents, anticipated enjoyment is also significantly associated with choice (.261, $p < .05$; column 6), while the other coefficients are smaller and less precisely estimated. The rating-by-network-status interactions are individually small and imprecise, consistent with respondents applying the same evaluative logic to both experimental and network candidates. Though noisier, Part II estimates are consistent with a common evaluative logic across the two parts: the attributes that matter for choosing between experimental profiles continue to matter when respondents choose among people they actually know. Part II adds relationship-specific information and social considerations, rather than replacing the evaluative margins identified experimentally in Part I.

Strategic competition. A possible alternative interpretation is that the respondent-side component of the GNC penalty reflects strategic competition: incumbent workers may avoid

referring GNC candidates if they view them as threats to their own positions. Chiplunkar et al. (2024) document that this channel can be economically meaningful — jobseekers in their large-scale referral experiment share less job information with their high-ability peers when the position is competitive, reducing the quality of the applicant pool. In our setting, however, the referred candidates are new labor-market entrants (fresh vocational graduates on a six-week trial), while incumbents are established workers with a track record at the firm; a fresh graduate is unlikely to be seen as a direct replacement risk. Consistent with this, the elicited ratings provide little initial support for the strategic-competition account, since respondents perceive GNC candidates as less able and expect employers to value them less rather than viewing them as unusually strong competitors. Appendix Table A.4 provides an additional, though indirect, test by interacting the GNC-candidate indicator with an indicator for above-median tenure at the respondent's current firm. If competitive threat were an important driver, the GNC penalty should be larger among shorter-tenured workers, who are plausibly less securely established at the firm. We find no such pattern: across all six outcomes, the interaction estimates are small and imprecise, and the point estimates do not indicate systematically larger penalties among shorter-tenured respondents.

5.3 Occupational Context: Female- and Male-Dominated Occupations

The pooled estimates establish that GNC candidates face a referral penalty on average. We now exploit a distinctive feature of the design: the same experimental task captures both directions of occupational non-congruence, men considered for positions in FDOs and women considered for positions in MDOs. This bidirectional comparison is built into the experiment rather than generated by a post hoc sample split. It reveals a central distinction: the referral penalty appears in both contexts, but the evaluations accompanying it are substantially more negative in MDOs.

This comparison requires one qualification. Respondent gender and occupation type are highly correlated in our sample: 91 percent of respondents are themselves GC (women in FDOs or men in MDOs). Because candidate gender is randomized within respondent, the GC–GNC contrast within each occupation type is not confounded by respondent characteristics. The comparison of those contrasts across FDOs and MDOs, however, combines differences in occupational context with differences in respondent gender. Appendix Table A.3 explores this issue by testing whether the GNC penalty differs for respondents who are themselves GNC. The point estimates provide no consistent indication that GNC respondents behave differently, but this group constitutes only 9 percent of the sample and the relevant interactions are imprecisely estimated. We therefore treat this evidence as suggestive rather than as resolving the distinction between occupational context and respondent gender.

A bidirectional referral penalty. Table 4 reports the Part I estimates separately for FDOs and MDOs. GNC men considered for positions in FDOs are 17.2 percentage points less likely to be referred than otherwise identical GC women ($p = .005$). GNC women

considered for positions in MDOs are 26.7 percentage points less likely to be referred than otherwise identical GC men ($p < .001$). Both penalties are large and statistically distinguishable from zero. Although the point estimate is larger in MDOs, the difference between the two coefficients is not statistically significant ($p = .209$). The referral-choice outcome therefore establishes a penalty in both directions, but not a statistically detectable difference in its magnitude across occupational contexts.

Asymmetric evaluative content. Similar referral-choice gaps mask a pronounced difference in the evaluations accompanying them. First, the respondent's own willingness to work with the candidate moves sharply against women in MDOs: their likability rating falls by .580 points ($p < .001$). The corresponding estimate for men in FDOs is smaller and statistically insignificant (.199; $p = .188$). The overall perceived-retention penalty is also more negative in MDOs, although the FDO–MDO difference in this top-level outcome is not statistically significant. When perceived retention is unpacked, however, the cross-context asymmetry is pronounced across all three subcomponents. In MDOs, women receive a 1.33 point penalty on perceived task ability ($p < .001$), a .519 point penalty on anticipated enjoyment ($p < .001$), and a .551 point penalty on anticipated employer valuation ($p < .001$). The corresponding estimates in FDOs are uniformly smaller. Moreover, the FDO–MDO difference is statistically significant for all three subcomponents. The asymmetry is especially pronounced for perceived task ability: the penalty for women in MDOs is more than four times the corresponding penalty for men in FDOs. The Part I estimates therefore distinguish the existence of a referral penalty from its evaluative content. Men entering FDOs face a sizable referral penalty, but it is accompanied by comparatively modest differences on the measured preference and belief margins. Women entering MDOs, by contrast, are penalized across the respondent's own willingness to work with them and across all three beliefs underlying expected retention. Thus, similar reduced-form referral gaps coexist with different measured evaluative patterns.

The language of evaluation. The closed-ended ratings document the dimensions on which GNC candidates are penalized, but they do not reveal which attributes respondents spontaneously make salient when evaluating otherwise comparable profiles. The open-ended descriptions collected in Part I provide complementary evidence. Because candidate gender and experience are randomly assigned and respondents have no prior knowledge of the profiles, systematic differences in the language used cannot reflect candidate-specific private information. Appendix Figure A.6 classifies the nearly 5,000 elicited adjectives into thirteen mutually exclusive trait categories using an LLM-assisted procedure, with all classifications manually reviewed.[20] The language qualitatively mirrors the occupational

[20]We first standardized spelling and translated rare terms into their nearest common equivalents. GPT-4o then assigned each word to one of thirteen mutually exclusive trait categories: (i) experience and skill, (ii) education and training, (iii) cognitive abilities, (iv) agentic, (v) grindstone, (vi) communal, (vii) professional, (viii) overall qualification, (ix)–(xi) demographic descriptors (age, physical strength, explicit gender references), (xii) negative, and (xiii) a residual "other" category. The agentic and grindstone definitions follow prior work on language and labor-market discrimination (Eberhardt et al., 2023; Beugnot and Peterlé, 2020). Our team manually validated 99.6–99.8 percent of assignments; human–LLM agreement was 80.1 percent on unique standardized adjectives and 89.5 percent when weighted by frequency of occurrence.

asymmetry in the elicited ratings. In MDOs, where female profiles are GNC, women are more frequently described using gender and age markers, negative descriptors, and communal or grindstone terms—attributes such as caring, cooperative, hardworking, and diligent that emphasize interpersonal qualities or effort rather than innate ability, echoing the reference-letter evidence in Eberhardt et al. (2023). Competence-related language, by contrast, tilts toward male profiles. In FDOs, where female profiles are GC, these contrasts are substantially weaker and do not simply reverse: women are not disproportionately described using gender or negative terms, while professional and education-and-training descriptors tilt toward female profiles. The pattern is therefore not merely a general tendency to describe women and men differently. Occupational context shapes which candidate attributes respondents make salient, with the sharpest competence-related contrast arising when women are evaluated for positions in MDOs. This free-text evidence reinforces the rating decomposition. In MDOs, the large perceived-ability and likability penalties are accompanied by spontaneous language that associates competence-related attributes more readily with men and describes women more often in terms of social identity, effort, and interpersonal traits. The asymmetry is therefore not confined to the particular closed-ended scales used in the survey.

Parallel patterns among network contacts. Part II shows a similar occupational asymmetry when respondents consider people from their own networks. Because real contacts are not experimentally assigned and may differ in both observable and unobservable characteristics, we treat these comparisons as descriptive. Panel A of Figure 2 shows that the GC–GNC gap at the naming margin is larger in MDOs than in FDOs by 16.9 percentage points ($p = .065$). The asymmetry is sharper at the referral margin: conditional on naming at least one contact, the GC–GNC referral gap is 26.4 percentage points larger in MDOs than in FDOs ($p = .004$). Panel B reports the evaluations of the nominated contacts. Across likability, perceived retention, task ability, anticipated enjoyment, and anticipated employer valuation, the GC–GNC differences qualitatively mirror those found in Part I, with the largest gaps again appearing in MDOs. Because the contacts are selected from respondents' networks, these differences may reflect either differential evaluation or private information about the particular contacts nominated. We therefore do not interpret Part II as a second causal estimate. Its value is that the same cross-context asymmetry recurs when respondents move from otherwise comparable experimental profiles to people they actually know.

Taken together, the bidirectional design yields two distinct conclusions. Occupational non-congruence reduces referral access in both directions. But the evaluative content of that penalty is asymmetric: women entering MDOs face a broad penalty centered on respondents' own willingness to work with them and on perceived task ability, while the corresponding evaluations of men entering FDOs are markedly less negative. The next section uses the randomized experience signal to ask whether stronger occupation-specific credentials can overcome these different evaluative barriers.

5.4 What Attenuates the Penalty? Experience, Privacy, and Exposure

Having documented the penalty and its channels, we ask what moves them. Three sources of variation are informative: randomized candidate experience, randomized privacy of the referral choice, and observational variation in respondents' exposure to GNC coworkers and classmates.

Experience. Table 5 adds a randomized "more experience" indicator, equal to one if the candidate was assigned an additional five-month post-training job spell, nearly doubling the experience shown on the profile, and interacts it with GNC status. The direct GNC-by-experience interaction is small and statistically insignificant on all outcomes; on its own the interaction would suggest that experience does not differentially attenuate the GNC penalty. We therefore focus on the joint effect $\beta_1 + \beta_2 + \beta_3$, which compares an experienced GNC candidate to a less-experienced GC candidate. This linear combination asks a specific question: how does the evaluation of a more experienced GNC candidate compare to the evaluation of a less experienced GC candidate, the baseline used elsewhere in the paper?

Estimated separately by occupation type (Appendix Table A.5), the joint effects reveal a structurally different response to experience across FDOs and MDOs. For men in FDOs, additional experience moves respondents' first-order beliefs about task ability decisively above the less-experienced GC baseline (.628, $p = .006$) and moves anticipated employer valuation above baseline as well (.660, $p = .002$), with likability and retention also above baseline (.476, $p = .064$; .050, $p = .063$). For women in MDOs, the joint effect on anticipated employer valuation reaches parity with the less-experienced GC baseline but does not exceed it (.132, $p = .493$), so respondents update their guess about the employer's view toward equal treatment. What does not move as far are the two objects that capture respondents' own first-order beliefs about the candidate: experienced women in MDOs are still rated $-.806$ points less able on the core tasks of the job than less-experienced GC men ($p < .001$) and $-.407$ points less likely to enjoy the work ($p = .026$). The referral joint effect for experienced MDO women is correspondingly small (.070, $p = .257$).

We interpret this as evidence of a first-order versus second-order belief split rather than as evidence that experience closes the referral gap for FDO men but not for MDO women: for men in FDOs, additional experience moves both first-order beliefs (task ability) and second-order beliefs (anticipated employer valuation) toward parity (there was no penalty in candidate enjoyment to begin with); for women in MDOs additional experience moves the second-order object but not the first-order objects. Comparable reduced-form referral penalties across the two directions of gender-line crossing therefore mask a substantively different pattern of belief updating, a sharper mechanism finding our bidirectional design supports.[21]

[21] Bohren et al. (2019) develop a framework distinguishing three sources of discrimination (belief-based with correctly specified priors, belief-based with biased priors, and preference-based). One implication is the subjectivity-of-judgment prediction that belief-based discrimination attenuates as evaluators observe more precise signals of candidate quality, while preference-based discrimination is invariant to signal precision. Randomized experience is one such signal in our setting. An alternative reading, also in Bohren et al. (2019), is that biased evaluators may not update priors on signals from minority-group candidates; the

The muted first-order enjoyment response for MDO women constrains an alternative explanation of the enjoyment penalty. One might argue that respondents anticipate that women in MDOs will enjoy the job less because respondents know their own workplace is hostile to women, so what looks like a belief about the candidate is really a report about the workplace. That reading has some difficulty absorbing the absence of an experience response. A woman with additional experience in motor mechanics, welding, or another male-dominated trade has completed additional training and accumulated additional work experience in a similar Ugandan labor-market environment that, by the same logic, would also have likely been hostile. A respondent who anticipated such hostility could treat her prior experience as demonstrated persistence in similar settings. For the workplace-knowledge story to rationalize the muted response, respondents would have to hold, simultaneously, that the candidate's prior workplaces were unusually friendly and that their own is unusually hostile. The extension of the paternalistic-discrimination mechanism of Buchmann et al. (2025) from employers to incumbent employees is a natural reading of the residual enjoyment penalty.

The discussion so far has focused on Part I, where candidate experience is experimentally manipulated. In Part II we cannot repeat that exercise: network contacts' experience is not randomized, and can only be examined observationally. Despite this observational selection, when we regress network referral (and network ranking) choices on candidate GNC status and control for observables including experience, the GNC gap on the network side barely moves (Table 1, columns 3 and 5). The residual referral gap in Part II is therefore not a mechanical consequence of GNC contacts being systematically less experienced.

Privacy of the referral choice. Table 6 reports a within-Part-I randomization of whether the referral is public (identifiable to the employer) or private (confidential and transmitted through BRAC Uganda, a well-known NGO and our implementing partner for the study). The GNC-by-private interaction is positive and marginally significant in the pooled sample (.141, $p = .083$), meaning that the GNC penalty is somewhat smaller under privacy. However, the penalty under private referral remains large and highly significant ($-.270$, $p < .001$). We read this as evidence that reputational or conformity costs from being observed making a non-congruent recommendation contribute to the referral penalty at the margin but cannot account for the bulk of it.

Exposure to GNC coworkers and classmates. Appendix Table A.4 reports observational heterogeneity in the referral penalty by the respondent's own exposure to GNC individuals. Respondents whose current shop includes a GNC coworker exhibit a smaller task-ability penalty (.505, $p < .05$) and a smaller anticipated employer-valuation penalty (.380, $p < .10$) toward GNC candidates in Part I, while the corresponding referral, likability, retention, and enjoyment attenuations are close to zero. On the other hand, respondents who had a GNC classmate during their vocational training show no systematic attenuation across outcomes. These are observational heterogeneities and we do not read them as causal: the composition of one's shop and classmates is not randomly assigned, and workers with

muted first-order response to experience for MDO women is consistent with either rigid priors or selective under-weighting of the signal.

different unobserved priors may sort into or out of shops that already employ GNC workers. What we take from the table is that direct professional contact with GNC coworkers is associated with a specific and interpretable attenuation of the first-order and second-order belief channels most implicated in the penalty, while more distant historical exposure through a classmate is not, a pattern consistent with contact-based updating that acts through belief channels rather than through preferences over coworkers.[22]

Taken together, the three levers we test suggest that the GNC penalty is difficult to close. Productivity-linked rewards, which can offset coworker prejudice in other settings (Hedegaard and Tyran, 2018a), are unlikely to work here because the largest penalty is itself a belief that GNC candidates are less productive. A different policy margin is who gets to refer. In an Indian manufacturing firm, Srivastava (2026) reallocates referral invitations toward lower-caste incumbents, increasing minority employment and output without reducing cohesion; cohesion falls instead when minority workers enter as outsiders. Although caste and gender differ deeply, this result suggests that diversifying the pool of incumbents invited to refer may complement quotas and evaluation redesign, whose effects depend on context (Bohnet, 2016; Matsa and Miller, 2013). Whether the same approach can expand women's entry into gender-segregated occupations is a promising avenue for future research.

5.5 Are firms missing out on productive workers?

The dominant belief-based penalty in MDOs, perceived task ability, admits two readings. Respondents could hold accurate beliefs about a genuine skill gap between GC and GNC workers, or inaccurate beliefs about group-level productivity, what Bohren et al. (2025a) describe as inaccurate statistical discrimination. A leading model of the latter is the representativeness-based stereotypes framework of Bordalo et al. (2016), which Bordalo et al. (2019) apply to gender to show that beliefs about relative male–female performance systematically overshoot the true gap in a way that would rationalize a perceived-ability penalty against GNC candidates. Our experiment does not adjudicate between accurate and inaccurate beliefs. Adjacent evidence, however, does not point to a real skill gap. In the ECMA sample, GC and GNC graduates are statistically indistinguishable on a purpose-built theoretical and practical skill assessment, on wage-employment rates, on earnings conditional on employment in the training sector, and on self-reported career satisfaction (Online Appendix Section D.1 and Table A.6). Macchi and Raisaro (2026) report the same pattern in three of the MDOs we study using an independent skill measure.

One concern that may arises is whether the perceived-ability penalty reflects the physical demands of some MDOs. GC and GNC workers may have similar measured skills and labor-market outcomes, while some jobs still require physical strength not captured by those measures. Although many of the trades in our setting are heavily mechanized, we examine this possibility using the same task-level ECMA data described in Table A.6, which record the activities workers routinely perform and allow us to construct a more granular

[22]Although our exposure measures are observational, the concentration of attenuation in task ability and anticipated employer valuation, rather than likability, is consistent with prior evidence that exposure to counter-stereotypical women changes beliefs about women's competence and effectiveness without necessarily changing underlying tastes (Beaman et al., 2009; Dasgupta and Asgari, 2004).

occupation-level measure of physical demands. As expected, MDOs are more physically demanding on average than FDOs. Within MDOs, however, the pattern runs opposite to the threshold prediction. In the least physically demanding MDOs, the GNC referral penalty is greater. In the more physically demanding MDOs, plumbing, welding, and construction, the GNC referral penalty is smaller (Figure A.2). This comparison is purely descriptive and imprecisely estimated. Even so, the GNC referral gap is not concentrated in the occupations where physical strength should be most binding. The convergence across cohorts and measurement tools, together with the fact that the GNC referral gap is largest where physical strength is least likely to bind, makes an accurate-beliefs interpretation of the perceived-ability penalty harder to sustain, though we cannot rule it out.[23]

6 Aggregate Implications: A Model of Gendered Referrals

So far, we have shown that referrals can contribute to occupational segregation through two margins: gender-segregated networks limit access to GNC candidates, and incumbents are less likely to recommend them even when they are available. As shown in Section 5, this recommendation-stage penalty reflects incumbents' own preferences and beliefs about candidate ability and fit, together with their beliefs about employer preferences and candidate enjoyment. We now ask how much this second margin contributes to aggregate segregation.

We embed the experimentally estimated GNC referral penalty in a dynamic network model of job finding, following Calvo-Armengol and Jackson (2004). The model does not seek to explain separately all sources of occupational segregation. Instead, it asks a disciplined counterfactual question: holding fixed observed employment and referral-network structure, how much additional segregation arises when workers forward job information using the referral weights implied by the experiment? We proceed in three steps. The setup describes the network of workers and their possible employment states and the three ways in which segregation can arise in our model. The dynamics describe what happens in each period: workers receive job signals directly or via referrals from employed neighbors, jobs are destroyed stochastically, and employment evolves. The steady-state and counterfactual exercises selectively shut down these channels to isolate the contribution of referral bias.

6.1 Setup

Consider a network of N workers, each unemployed or employed in one of two occupations, with occupation 1 female-dominated and occupation 2 male-dominated. In each period, workers may receive job signals directly.[24] An unemployed recipient may take the job; an

[23]Following Bertrand and Mullainathan (2004), we also do not claim that this evidence identifies whether the residual penalty is taste-based, statistical, or a mix; our narrower claim is that GNC candidates are evaluated as if their credentials conveyed less information than the ex-post record from a similar population would justify.

[24]Unemployed workers may search directly, while job opportunities may also reach employed workers who are not searching, as in models of on-the-job search (Burdett and Mortensen, 1998), for example when their employer is hiring.

employed recipient cannot take it but may forward the signal to one network contact. Jobs are destroyed stochastically, and the economy evolves toward a steady state.[25]

The model incorporates three channels that can generate or amplify occupational segregation. First, men and women may differ in their direct access to jobs outside the referral process. Let $\sigma_g^{(k)}$ denote the probability that a worker of gender g receives a direct signal for occupation k. We allow these probabilities to differ by gender and occupation. In particular, we assume

$$\frac{\sigma_m^{(2)}}{\sigma_m^{(1)}} > \frac{\sigma_f^{(2)}}{\sigma_f^{(1)}},$$

so men have relatively better direct access to occupation 2 than to occupation 1, compared with women.[26] This reflects the fact that, beyond referrals, GNC candidates face other structural barriers to entering jobs that do not conform to traditional gender expectations. These barriers may arise on the supply side, through self-selection shaped by constraints and preferences, including anticipated workplace hostility and social norms, and on the demand side, through employer screening and discrimination. We do not model these forces separately; instead, the gender- and occupation-specific signal-arrival rates summarize their combined effect on access outside the referral process.[27] We refer to this channel as differential direct job access, or S.

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

Second, social networks may be gender-homophilous, a well-documented feature of social and professional ties (Ibarra, 1992; McPherson et al., 2001). We implement this by generating a random graph using a stochastic block model (Holland et al., 1983). For the simulations, we set $N = 100$, half of whom are women, in which the probability of a same-gender link, p, exceeds the probability of a cross-gender link, q. We refer to it as network composition, or NC.

Third, workers may favor GC over GNC contacts when forwarding job signals. When an employed worker received a job signal for occupation k, she forwards it to one neighbor. We model the probability that worker i forwards the signal to neighbor j as

$$\theta_{ij}^{(k)} = \frac{\omega_{g_j}^{(k)}}{\sum_{\ell \in \mathcal{N}(i)} \omega_{g_\ell}^{(k)}}, \qquad j \in \mathcal{N}(i), \tag{5}$$

where g_j is the gender of neighbor j, and $\omega_{g_j}^{(k)}$ is the referral weight assigned to a neighbor

[25] We denote the economy's state at time t by binary vectors $e_t^{(0)}, e_t^{(1)}, e_t^{(2)} \in \{0,1\}^N$, where $e_{it}^{(k)} = 1$ if worker i is in state $k \in \{0,1,2\}$.

[26] We do not impose this inequality in estimation; rather, it emerges from the estimated parameters.

[27] Modeling occupational preferences explicitly would require forward-looking offer-acceptance decisions, because unemployment may have option value, as well as job-to-job transitions when employed workers receive offers in preferred occupations. We abstract from these dynamics for tractability and allow $\sigma_g^{(k)}$ to absorb worker-side search and acceptance behavior together with demand-side forces.

of gender g_j for occupation k. In the baseline specification, these weights depend on the recipient's gender and the occupation, but not on the sender's gender.[28] We also impose within-gender symmetry, so that, for a given occupation, all male neighbors of i receive the same weight and all female neighbors receive the same weight.[29]

Referral behavior is therefore parameterized by four weights: $\omega_f^{(1)}$, $\omega_m^{(1)}$, $\omega_f^{(2)}$, and $\omega_m^{(2)}$. A larger $\omega_g^{(k)}$ makes a neighbor of gender g more likely to receive a signal for occupation k, conditional on the size and gender composition of the sender's network. Because the weights are normalized over the sender's realized neighbors, having more contacts reduces the probability that any particular contact receives the signal.[30]

The GNC referral penalty corresponds to weights that favor GC recipients:

$$\omega_f^{(1)} > \omega_m^{(1)} \qquad \text{and} \qquad \omega_m^{(2)} > \omega_f^{(2)}. \tag{6}$$

Under these inequalities, signals for the female-dominated occupation are more likely to be forwarded to women, while signals for the male-dominated occupation are more likely to be forwarded to men. For any worker with at least one network neighbor, the probabilities $\theta_{ij}^{(k)}$ sum to one over $j \in \mathcal{N}(i)$, so the worker forwards the selected signal to exactly one neighbor. An isolated worker cannot forward the signal. These referral weights determine how direct job signals arriving at rates $\sigma_g^{(k)}$ propagate through the network.

Dynamics In each period, 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 $s_{t,i}^{(k)} \sim$ Bernoulli$(\sigma_g^{(k)})$. Unemployed individuals who receive a signal transition into employment in the corresponding occupation. If they receive signals from both occupations, the offer accepted is chosen with a fair coin flip. Conditional on the selected signal being from occupation k, an employed individual forwards it to one neighbor with probability $\theta_{ij}^{(k)}$, as defined above. Equivalently, the recipient is drawn from a multinomial distribution over i's neighbors:

$$\phi_i^{(k)} \sim \text{Multinomial}\Big(1,\ \big(\theta_{i,j}^{(k)}\big)_{j \in \mathcal{N}(i)}\Big), \tag{7}$$

Worker j receives a referral from occupation k if at least one employed neighbor forwards that signal to j:

$$\tilde{s}_{t,j}^{(k)} = \mathbf{1}\big(\sum_i \phi_{i \to j}^{(k)} > 0\big) \tag{8}$$

[28]This parsimonious restriction is consistent with suggestive evidence on the absence of a systematic reversal of the GNC referral pattern among respondents who are themselves GNC.

[29]This restriction is imposed for tractability and can be relaxed by allowing $\omega_{ij}^{(k)} = \omega_{g_j}^{(k)} \lambda_{ij}^{(k)}$, where $\lambda_{ij}^{(k)}$ 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 GNC penalty unchanged, while potentially affecting its quantitative impact by making referrals more or less concentrated across contacts. We do not pursue this extension quantitatively.

[30]Under the baseline specification, conditional on receiving a signal, heterogeneity in forwarding probabilities across senders arises from differences in the size and gender composition of their realized networks. Realized referral flows also vary with workers' employment states and signal arrivals. As a robustness exercise reported in the Online Appendix, we allow network size and composition to vary by gender and examine how this richer network structure affects the results.

Referral signals cannot be forwarded further. At the end of each period, employed workers keep their job unless hit by job destruction $d_i \sim \text{Bernoulli}(\delta)$. Let $a_i^{(k)}$ denote the event that worker i receives a net first-hand offer from occupation k, after resolving ties, and let $b_i^{(k)}$ denote the analogous event for referrals. Employment states evolve according to:

$$e_{t,i}^{(0)} = e_{t-1,i}^{(0)} \cdot a_i^{(0)} \cdot b_i^{(0)} \; + \; d_i \cdot \left(e_{t-1,i}^{(1)} + e_{t-1,i}^{(2)}\right) \tag{9}$$

$$e_{t,i}^{(k)} = e_{t-1,i}^{(k)}(1 - d_i) \; + \; e_{t-1,i}^{(0)} \cdot \left(a_i^{(k)} + a_i^{(0)} \cdot b_i^{(k)}\right), \quad k \in \{1, 2\} \tag{10}$$

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 employed neighbors, the model has no closed-form solution. We therefore approximate the steady state using Monte Carlo simulations, with additional details reported in the Online Appendix.

Parameters We discipline the model using three sources of information from the data. First, we recover the referral weights $\omega_f^{(1)}$, $\omega_m^{(1)}$, $\omega_f^{(2)}$, and $\omega_m^{(2)}$ from the experimental estimates of the GNC referral penalty. The coefficient on the GNC candidate in the referral-choice regression pins down the relative probability that job signals are forwarded to gender-congruent rather than gender non-congruent candidates.

Second, we use our network data to discipline network size and homophily. Respondents were asked to name "people [they] expected could help [them] in the initial phases of [their] career." These data imply an average degree of 5.0 connections, 64 percent of whom are of the same gender as the respondent. In the baseline model, we use these moments to set the average degree and the extent of gender homophily in the stochastic block model. An appendix extension allows degree and composition to differ by gender.

Third, we estimate the four direct job-arrival rates by simulated method of moments, matching the shares of women and men employed in each occupation.[31]

6.2 Steady State and Counterfactuals

To quantify the contribution of the GNC referral penalty, we compare steady-state segregation across economies that selectively shut down the three channels: differential direct job access S, network composition NC, and the GNC referral penalty $GNCp$. Figure 3 contains the results of this quantification exercise. Figure 3 quantifies the contribution of the GNC referral penalty by comparing steady-state segregation across economies that selectively shut down differential direct job access, S, network composition, NC, and the GNC referral penalty, $GNCp$. When all three channels are shut down ("None"), direct arrival rates are equal across genders, networks are not homophilous, and referral weights are

[31] For estimation 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. We then estimate the vector of signal parameters by minimizing the squared distance between the simulated employment shares and the empirical moments. Four parameters and four moments make the system exactly identified. Table OA.4 reports the SMM parameter estimates, standard errors, and moment fit. To simulate the steady state at the estimated parameters, we run 2,000 Monte Carlo simulations over $T = 1{,}500$ for networks with $N = 100$, averaging the last 50 periods.

gender-neutral. The model then generates essentially no occupational gender segregation. Adding the channels one at a time clarifies their distinct roles. Differential direct job access ("S") is the main source of segregation in the baseline calibration, generating large gender gaps in both occupations even on its own. Network composition alone generates little segregation because homophily has little to amplify when direct access is equal and referrals are unbiased.[32] The GNC referral penalty, by contrast, generates segregation even on its own because it directly steers job information toward gender-congruent candidates.

Our main counterfactual adds the experimentally estimated GNC referral penalty to an economy with differential direct job access and gendered network composition. The GNC penalty increases steady-state occupational segregation by 9.2 percent in male-dominated occupations and 7.8 percent in female-dominated occupations, corresponding to approximately 6.3 and 5.1 percentage points, respectively. Most baseline segregation is loaded onto S, which is estimated residually to match observed employment patterns and therefore absorbs a wide range of barriers outside referrals. The GNC penalty is thus not the main source of segregation in levels, but it materially amplifies segregation by redirecting job information toward gender-congruent candidates. The Online Appendix shows that the contribution of the GNC penalty rises when referrals play a greater allocative role: in denser or less homophilous networks and when direct job access is less segregated. We also allow network size and composition to differ by gender. The contribution of referral bias is substantially larger under alternative calibrations using employer-side estimates. These exercises suggest that the baseline is a conservative benchmark and that the GNC referral penalty has quantitatively meaningful aggregate effects.

7 Conclusion

WIP

[32]This result relies on equal direct arrival rates and gender-neutral referral weights. When existing employment imbalances are allowed to propagate through referrals, homophily amplifies segregation (Figure 3, "S" versus "S+NC"), as in Bolte et al. (2026).

Main Tables and Figures

Table 1: The GNC Referral Penalty

	Part I (Experimental Candidates)		**Part II** (Network Candidates)		
	Referred	Ranked Higher	Named	Referred	Ranked Higher
	(1)	(2)	(3)	(4)	(5)
GNC Experimental Candidate	−.234***	−.205***			
	(.036)	(.036)			
GNC Network Contact			−.607***	−.262***	−.308***
			(.023)	(.064)	(.071)
GC mean	.617	.602	.941	.529	.654
Observations	1,494	1,494	1,360	374	370

Notes: Column 1 reports the Part I hypothetical referral (full sample). Column 2 reports a Part I "ranked higher" outcome, constructed from a separate survey elicitation administered *after* Part II: respondents were asked to rank their preferences across all candidates in case their referred candidate proved unavailable. Between the two Part I candidates, we assign "ranked higher" = 1 to whichever profile the respondent placed above the other in this later elicitation, with fall-backs to `other_pref_cand` ranks, the `preference_referral` top choice, and finally the primary referral (`refer_fp`/`refer_sp`) to resolve zero-ranks and skipped rows. The measure agrees with the primary referral choice on 95.4% of profiles; the remainder are cases where the respondent reversed their ordering when reconsidering. Column 2 therefore gives an independent confirmation of the referral penalty from a later, softer preference elicitation in which the respondent had time to reconsider. Columns 3–5 report Part II network outcomes: probability of being named as a network contact (col. 3), being the final referred contact conditional on the respondent naming contacts of both genders (col. 4), and being ranked higher than the other named contact (col. 5). Part II columns correspond to columns 1, 2, and 4 of the Network Choice table; specifications pool the Round 1 and Round 2 samples (Part II was fielded identically in both rounds) and therefore include wave and training-area fixed effects, but no observable candidate controls. "Control / GC mean" is the outcome mean among GC candidates (Part I) and among GC network contacts (Part II). Standard errors clustered at the respondent level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Table 2: Forces Behind the Referral Choice in Part I

	Top-Level Channels		**Unpacking Perceived Retention Probability**		
	Likability	Perceived Retention Probability	Tasks Ability	Candidate Enjoyment	Employer Valuation
	(1)	(2)	(3)	(4)	(5)
GNC Experimental Candidate	-.450***	-.031***	-.980***	-.402***	-.438***
	(.090)	(.008)	(.091)	(.067)	(.082)
GC Mean	7.115	.671	7.786	8.107	6.796
Controls	Yes	Yes	Yes	Yes	Yes
Observations	1,494	1,493	1,372	1,369	1,364

Notes: Full-sample estimates from Part I (experimental candidates). Columns 1–2 report the two top-level channels: own likability of the candidate and the perceived probability the candidate is retained. Columns 3–5 unpack perceived retention into first-order beliefs about task ability, candidate enjoyment, and employer valuation. Standard errors clustered at the respondent level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Table 3: Forces Behind the Referral Choice in Part II

	Part I				Part II	
	Referred	Referred	Referred	Referred	Referred	Referred
	(1)	(2)	(3)	(4)	(5)	(6)
Likability	.732***	.759***	.649***	.687***	.354***	.253***
	(.093)	(.119)	(.102)	(.121)	(.071)	(.092)
Likability × Is Network					.079	.165
					(.103)	(.125)
Perceived Retention Probability	.426***	.377***			.083	
	(.068)	(.111)			(.077)	
Perceived Retention Probability × Is Network					.068	
					(.111)	
Employer Valuation			.240***	.404***		.146
			(.092)	(.144)		(.108)
Employer Valuation × Is Network						-.127
						(.115)
Perceived Tasks Ability			.327***	.301***		.076
			(.066)	(.099)		(.082)
Tasks Ability × Is Network						.080
						(.114)
Candidate Enjoyment			.277***	.304*		.261**
			(.088)	(.163)		(.106)
Candidate Enjoyment × Is Network						.000
						(.145)
Is Network					-.308	-.209
					(.903)	(1.011)
Pseudo R-Squared	.383	.395	.430	.475	.290	.317
Sample	All	Named Network	All	Named Network	Named Network	Named Network
Observations	1,492	570	1,354	518	898	818

Notes: Conditional logit estimates of the referral probability, with respondent (choice-set) fixed effects. Unit of observation is a candidate-profile row within a respondent's choice set. Columns (1)–(4) use Part I (experimental candidates) hypothetical referrals; columns (5)–(6) use Part II (network candidates) final choice, restricted to respondents who named a network contact, and include an indicator for whether the candidate is a network contact together with interactions between the network indicator and each rating. Within Part I, columns (1) and (3) use all respondents while columns (2) and (4) restrict to respondents who named at least one network contact. Columns (1)–(2) and (5) use likability and perceived retention probability as predictors; columns (3)–(4) and (6) instead use employer valuation, perceived task ability, and candidate enjoyment. Evaluated at overall sample means of the rating variables, the marginal effect of network membership on the logit is .76 in column (5) and .77 in column (6), both highly significant ($p < .001$). Standard errors clustered by respondent. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Table 4: Heterogeneity by Occupation Type in Part I

			Top-Level Channels				**Unpacking Perceived Retention Probability**					
	Referred		Likability		Perceived Retention Probability		Tasks Ability		Candidate Enjoyment		Employer Valuation	
	FDO	MDO	FDO	MDO	FDO	MDO	FDO	MDO	FDO	MDO	FDO	MDO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
GNC Experimental Candidate	-.172***	-.267***	-.199	-.580***	-.018	-.038***	-.303**	-1.330***	-.175	-.519***	-.216*	-.551***
	(.062)	(.044)	(.151)	(.112)	(.015)	(.010)	(.139)	(.115)	(.106)	(.086)	(.130)	(.104)
GC Mean	.586	.633	6.980	7.185	.670	.672	7.778	7.790	7.931	8.197	6.896	6.746
Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	512	982	512	982	512	981	468	904	467	902	461	903
GNC Candidate (FDO − MDO)	.095		.381**		.021		1.026***		.344**		.335**	
	[.209]		[.043]		[.239]		[.000]		[.012]		[.045]	

Notes: Heterogeneity in the GNC penalty by occupation type, using Part I (experimental candidate) data. Each outcome spans a pair of columns reporting FDO and MDO occupations: referral (cols. 1–2), likability (cols. 3–4), perceived retention probability (cols. 5–6), task ability (cols. 7–8), candidate enjoyment (cols. 9–10), and employer valuation (cols. 11–12). "GNC Candidate (FDO − MDO)" reports the between-occupation-type difference in the GNC coefficient with the p-value in brackets. Standard errors clustered at the respondent level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Table 5: The Experience Effect on the GNC Penalty

	Part I						Part II
	Referred	Likability	Retention Probability	Tasks Ability	Candidate Enjoyment	Employer Valuation	Referred
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GNC Experimental Candidate	−.250***	−.481***	−.034***	−1.025***	−.367***	−.535***	
	(.037)	(.106)	(.010)	(.107)	(.086)	(.099)	
More Experience	.285***	.501***	.051***	.554***	.370**	.482***	
	(.035)	(.149)	(.015)	(.157)	(.150)	(.150)	
GNC × More Experience	.053	.106	.010	.162	−.152	.367	
	(.045)	(.223)	(.023)	(.235)	(.231)	(.233)	
GNC Network Contact							−.243***
							(.066)
Linear combinations (Part I)							
More Exp. + GNC × More Exp.	.338***	.607***	.061***	.716***	.218	.849***	
	[.000]	[.000]	[.000]	[.000]	[.168]	[.000]	
More Exp. + GNC × More Exp. + GNC Experimental Candidate	.088*	.126	.026*	−.309*	−.149	.314**	
	[.079]	[.419]	[.087]	[.054]	[.286]	[.033]	
GC mean	.549	7.000	.659	7.662	8.023	6.685	.527
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,494	1,494	1,493	1,372	1,369	1,364	368

Notes: Full-sample estimates. Columns 1–6 (Part I, experimental candidates) regress the corresponding outcome on an indicator for a GNC candidate, an indicator equal to one if the candidate was randomly assigned five additional months of post-training work experience ("More Exp."), and their interaction. "More Exp. + GNC × More Exp." is the joint effect of experience for a GNC candidate; adding "+ GNC Experimental Candidate" is the level for an experienced GNC candidate relative to a less-experienced GC candidate; bracketed values are p-values. Column 7 (Part II, network candidates) reports the GNC gap on the referral margin for the respondent's own network, conditional on having named contacts of both genders, with observable controls including candidate experience: with experience controlled, the network GNC gap barely moves. Standard errors clustered at the respondent level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Table 6: Heterogeneity by Referral Confidentiality

	Referred		
	All Occupations (1)	FDO (2)	MDO (3)
GNC Experimental Candidate	-.341***	-.344***	-.339***
	(.070)	(.121)	(.086)
Private Referral	-.070*	-.113	-.048
	(.041)	(.070)	(.050)
GNC × Private Referral	.141*	.226	.096
	(.081)	(.140)	(.100)
Linear Combinations			
Private Referral + GNC × Private Referral	.070*	.113	.048
	[.083]	[.107]	[.337]
Private Referral + GNC × Private Referral + GNC Candidate	-.270***	-.231***	-.291***
	[.000]	[.001]	[.000]
Differences Across Occupation Types			
GNC Candidate (FDO − MDO)		-.005	
		[.971]	
Private Referral (FDO − MDO)		-.065	
		[.447]	
GNC × Private Referral (FDO − MDO)		.131	
		[.447]	
GC Candidate / Public	.670	.672	.669
Observations	1494	512	982

Notes: GNC penalty and effect of referral confidentiality on whether respondents refer a hypothetical candidate. Referral confidentiality is randomized into three arms: (1) public, where the employer will know the referrer's identity; (2) confidential, where we do not disclose the referrer's identity; and (3) NGO, where we tell the employer that BRAC made the referral. Arms (2) and (3) are pooled as "private." Columns report regressions of an indicator for referring a candidate on indicators for non-congruent candidates, private referral, and their interaction. Columns correspond to the full sample, female-dominated occupations, and male-dominated occupations. Standard errors clustered at the respondent level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Figure 1: Experimental Design

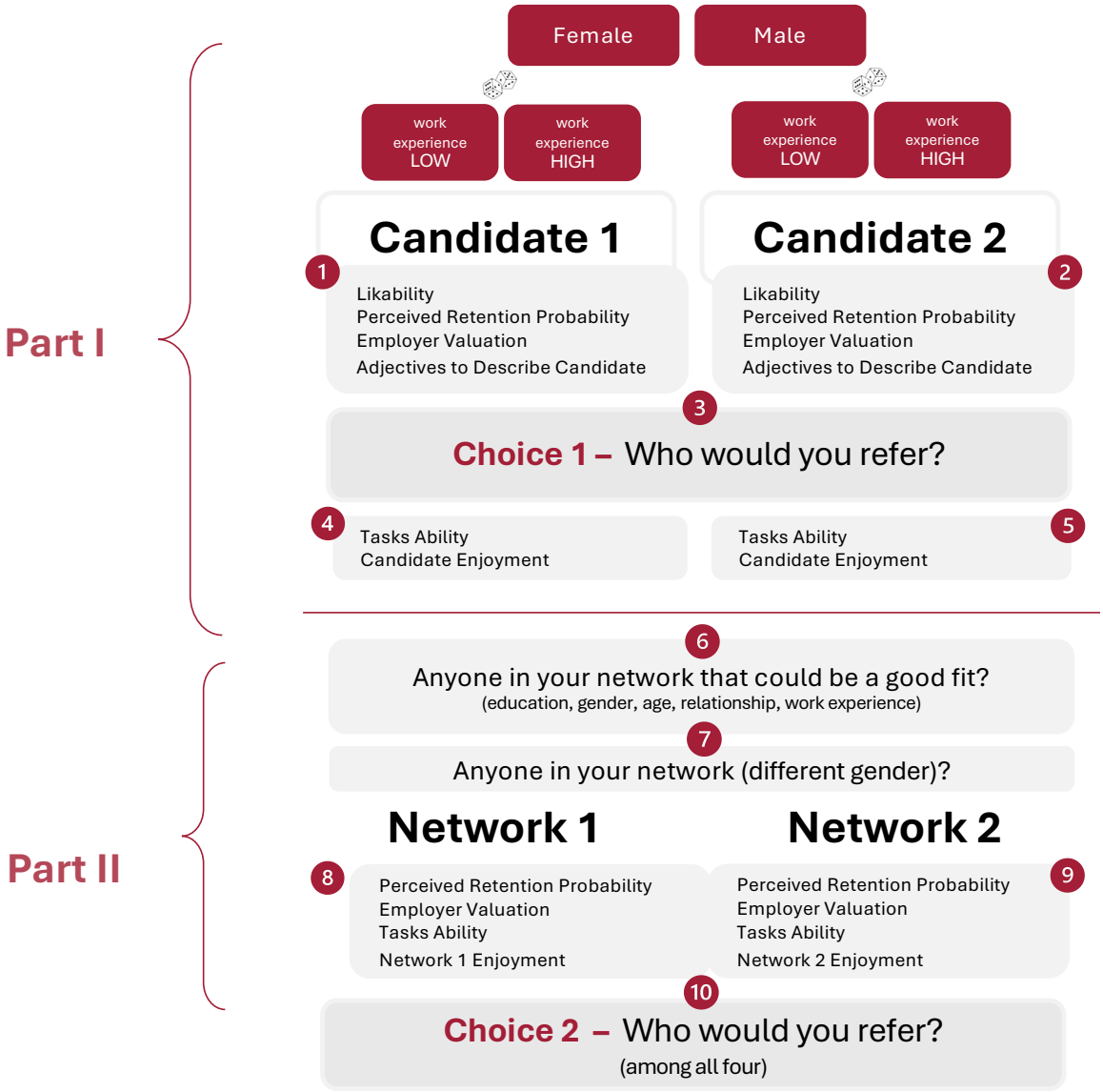
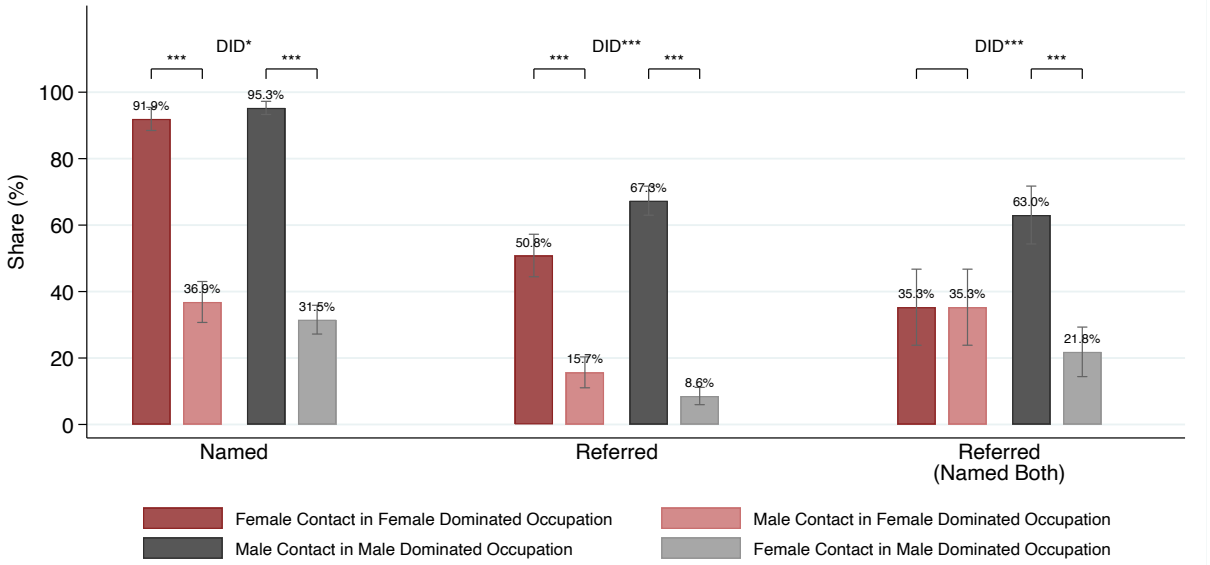
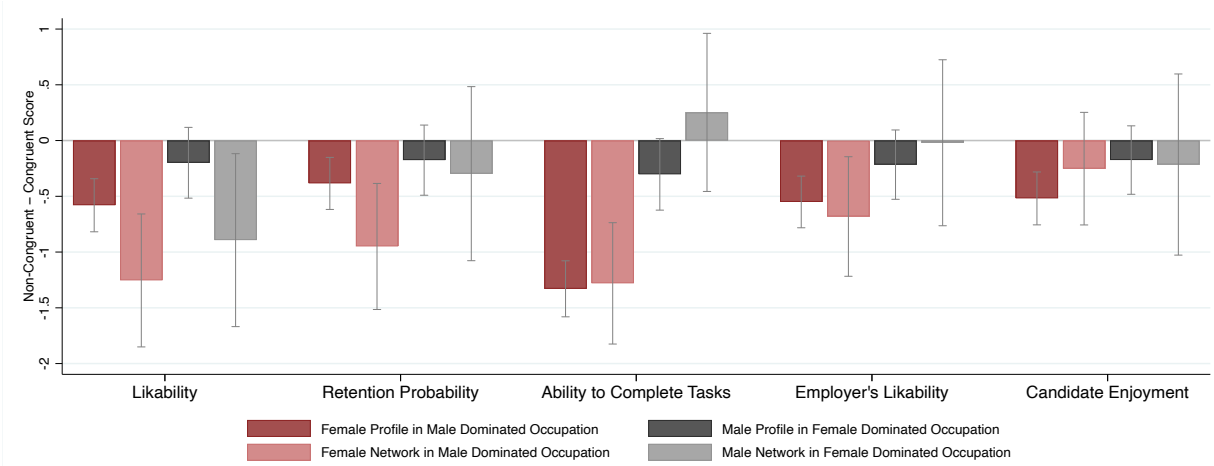


Figure 2: Heterogeneity by Occupation Type in Part II

Panel A: Network Composition and Referral Choice by Occupation Type

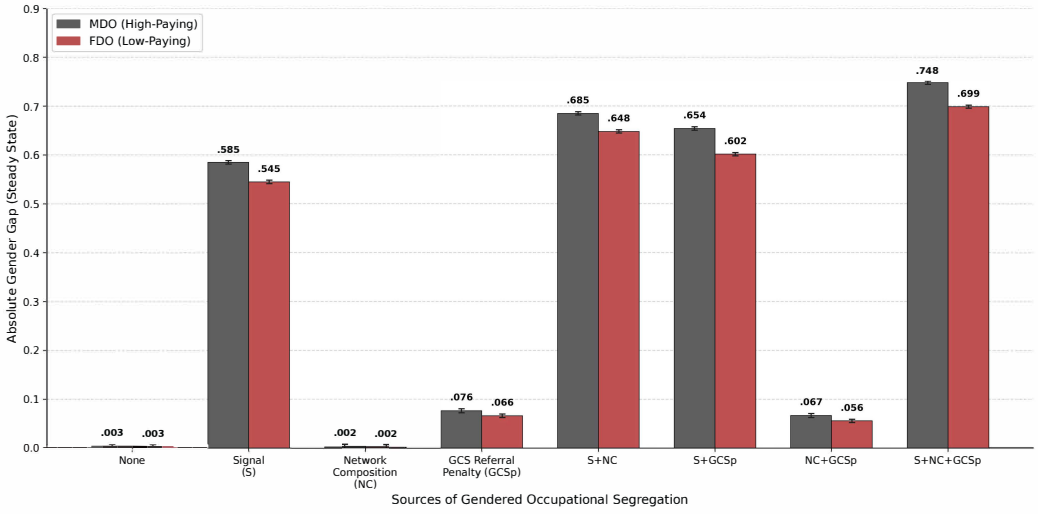


Panel B: Forces Behind the Referral Choice in Part II by Occupation Type



Notes: **Panel A:** share of workers who (i) *named* a gender-congruent (GC) or gender-non-congruent (GNC) contact, conditional on naming at least one contact; (ii) *referred* a GC or GNC contact, conditional on naming at least one contact; (iii) *referred* a GC or GNC contact, conditional on having named contacts of both genders. Whiskers show 95% confidence intervals; brackets show within-occupation-type differences (paired t-test). "DID" reports the difference-in-differences (MDO − FDO). **Panel B:** estimated gaps in ratings (non-congruent − congruent, in points on the 0–10 scale) on likability, retention probability, ability to complete tasks, employer's likability, and candidate enjoyment, by occupation type and by candidate type. Within each outcome, four bars: red shades show MDO occupations (dark red = female Part I experimental profiles; light red = female Part II network contacts) and grey shades show FDO occupations (dark grey = male Part I experimental profiles; light grey = male Part II network contacts). Vertical lines: 95% confidence intervals. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Figure 3: Steady State Occupational Gender Segregation



Notes: Gender gap is the difference between the share of men and share of women in an occupation. Averages over 2,000 Monte Carlo simulations, $T = 1{,}500$; steady states approximated by averaging the last 50 periods. "S" = differential job arrival rates (signals); "NC" = network composition (homophily); "GNCp" = GNC referral penalty. Baseline parameters: $N = 100$, $p = .064$, $q = .035$, $\delta = .03$, $\sigma_m^{(1)} = .0075$, $\sigma_m^{(2)} = .1558$, $\sigma_f^{(1)} = .1763$, $\sigma_f^{(2)} = .0074$, $\omega_C^{(1)} = .586$, $\omega_{NC}^{(1)} = .414$, $\omega_C^{(2)} = .6335$, $\omega_{NC}^{(2)} = .3665$. When "S" is shut off, $\sigma_m^{(1)} = \sigma_f^{(1)} = .0919$ and $\sigma_m^{(2)} = \sigma_f^{(2)} = .0816$. When "NC" is shut off, $p = .0509$, $q = .0498$. When "GNCp" is shut off, $\omega_C^{(s)} = \omega_{NC}^{(s)} = 1$. Confidence intervals are computed by holding signal parameters $\sigma_g^{(s)}$ fixed at their SMM point estimates and attributing all variation to simulation randomness.

References

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

Agan, A. and Starr, S. (2018). Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment. *The Quarterly Journal of Economics*, 133(1):191–235.

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., Bassi, V., Rasul, I., and Spadini, E. (2026a). Firm responses to uncertainty and implications for workers: Experimental evidence from uganda during the pandemic. Working Paper.

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.

Alfonsi, L., Namubiru, M., and Spaziani, S. (2026b). Meet Your Future: Experimental Evidence on the Labor Market Effects of Mentors. Working Paper.

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.

Banerjee, A., Bertrand, M., Datta, S., and Mullainathan, S. (2009). Labor Market Discrimination in Delhi: Evidence from a Field Experiment. *Journal of comparative Economics*, 37(1):14–27.

Bar, R. and Zussman, A. (2017). Customer Discrimination: Evidence from Israel. *Journal of Labor Economics*, 35(4):1031–1059.

Barr, T., Bojilov, R., and Munasinghe, L. (2019). Referrals and search efficiency: who learns what and when? *Journal of Labor Economics*, 37(4):1267–1300.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: Does exposure reduce bias? *The Quarterly Journal of Economics*, 124(4):1497–1540.

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.

Becker, S. O., Fernandes, A., and Weichselbaumer, D. (2019). Discrimination in hiring based on potential and realized fertility: Evidence from a large-scale field experiment. *Labour Economics*, 59:139–152.

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.

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American economic review*, 94(4):991–1013.

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

Bewley, T. F. (1999). *Why wages don't fall during a recession*. Harvard university press.

Bohnet, I. (2016). *What Works: Gender Equality by Design*. Harvard University Press, Cambridge, MA.

Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2025a). Inaccurate statistical discrimination: An identification problem. *Review of Economics and Statistics*, 107(3):605–620.

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

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

Bolte, L., Immorlica, N., and Jackson, M. O. (2026). The role of referrals in immobility, inequality, and inefficiency in labor markets. *Journal of Labor Economics*, 44(2):447–479.

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

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

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

Brown, M., Setren, E., and Topa, G. (2016). Do informal referrals lead to better matches? evidence from a firm's employee referral system. *Journal of Labor Economics*, 34(1):161–209.

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

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, pages 257–273.

Burks, S. V., Cowgill, B., Hoffman, M., and Housman, M. (2015). The Value of Hiring through Employee Referrals. *The Quarterly Journal of Economics*, 130(2):805–839.

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.

CareerBuilder (2012). *Referral madness: How employee referral programs turn good employees into great recruiters and grow your bottom line*. CareerBuilder e-book.

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

Castilla, E. J. (2005). Social networks and employee performance in a call center. *American journal of sociology*, 110(5):1243–1283.

Chiplunkar, G., Kelley, E., and Lane, G. (2024). Competitive job seekers: When sharing less leaves firms at a loss. *NBER Working Paper*, (32171).

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.

Dasgupta, N. and Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, 40(5):642–658.

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

Deming, D. J., Yuchtman, N., Abulafi, A., Goldin, C., and Katz, L. F. (2016). The Value of Postsecondary Credentials in the Labor Market: An Experimental Study. *American Economic Review*, 106(3):778–806.

Donald, A. and Grosset, F. (2024). Complementarities in labor supply: How joint commuting shapes work decisions. Working Paper.

Dustmann, C., Glitz, A., Schönberg, U., and Brücker, H. (2016). Referral-based job search networks. *The Review of Economic Studies*, 83(2):514–546.

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.

Edelman, B., Luca, M., and Svirsky, D. (2017). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American economic journal: applied economics*, 9(2):1–22.

Emanuel, N., Harrington, E., and Pallais, A. (2023). The power of proximity to coworkers: Training for tomorrow or productivity today? Technical report, National Bureau of Economic Research.

Eriksson, S. and Rooth, D.-O. (2014). Do Employers Use Unemployment as a Sorting Criterion When Hiring? Evidence from a Field Experiment. *American economic review*, 104(3):1014–39.

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.

Fernandez, R. M., Castilla, E. J., and Moore, P. (2000). Social capital at work: Networks and employment at a phone center. *American journal of sociology*, 105(5):1288–1356.

Fernandez, R. M. and Weinberg, N. (1997). Sifting and sorting: Personal contacts and hiring in a retail bank. *American Sociological Review*, pages 883–902.

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

Friebel, G., Heinz, M., Hoffman, M., and Zubanov, N. (2023). What do employee referral programs do? measuring the direct and overall effects of a management practice. *Journal of Political Economy*, 131(3):633–686.

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.

Galenianos, M. (2021). Referral networks and inequality. *The Economic Journal*, 131(633):271–301.

Ge, Y., Knittel, C. R., MacKenzie, D., and Zoepf, S. (2020). Racial Discrimination in Transportation Network Companies. *Journal of Public Economics*, 190:104205.

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.

Granovetter, M. (1995). *Getting a job: A study of contacts and careers*. University of Chicago press, 2 edition.

Heath, R. (2018). Why Do Firms Hire Using Referrals? Evidence from Bangladeshi Garment Factories. *Journal of Political Economy*, 126(4):1691–1746.

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.

Hedegaard, M. S. and Tyran, J.-R. (2018a). The price of prejudice. *American Economic Journal: Applied Economics*, 10(1):40–63.

Hedegaard, M. S. and Tyran, J.-R. (2018b). The Price of Prejudice. *American Economic Journal: Applied Economics*, 10(1):40–63.

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.

Hensvik, L. and Skans, O. N. (2016). Social networks, employee selection, and labor market outcomes. *Journal of labor economics*, 34(4):825–867.

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.

Ioannides, Y. M. and Loury, L. D. (2004). Job information networks, neighborhood effects, and inequality. *Journal of economic literature*, 42(4):1056–1093.

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.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87(5, Part 1):972–990.

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.

Kline, P. and Walters, C. (2021). Reasonable doubt: Experimental detection of job-level employment discrimination. *Econometrica*, 89(2):765–792.

Kramarz, F. and Skans, O. N. (2014). When strong ties are strong: Networks and youth labour market entry. *The Review of Economic Studies*, 81(3):1164–1200.

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

Lahey, J. N. (2008). Age, Women, and Hiring: An Experimental Study. *The Journal of Human Resources*, 43(1):30–56.

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

Macchi, E. (2023). Worth Your Weight: Experimental Evidence on the Benefits of Obesity in Low-Income Countries. *American Economic Review*, 113(9):2287–2322.

Macchi, E. and Raisaro, C. (2026). Hidden discrimination in frictional labor markets. Working Paper.

Matsa, D. A. and Miller, A. R. (2013). A female style in corporate leadership? evidence from quotas. *American Economic Journal: Applied Economics*, 5(3):136–169.

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.

Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *The American economic review*, 81(5):1408–1418.

Mortensen, D. T. and Vishwanath, T. (1994). Personal contacts and earnings: It is who you know! *Labour economics*, 1(2):187–201.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human resources*, 47(4):1128–1157.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, 56(3):799–866.

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.

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.

Oreopoulos, P. (2011). Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes. *American Economic Journal: Economic Policy*, 3(4):148–71.

Pallais, A. and Sands, E. G. (2016). Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, 124(6):1793–1828.

Petersen, T., Saporta, I., and Seidel, M.-D. L. (2000). Offering a job: Meritocracy and social networks. *American Journal of Sociology*, 106(3):763–816.

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).

Srivastava, D. (2026). Referrals, diversity, and firm productivity: Evidence from an indian manufacturing firm. Working paper.

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

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

Uganda Bureau of Statistics (2021). The national labour force survey 2021 – main report. Technical report, Uganda Bureau of Statistics, Kampala, Uganda.

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

World Bank (2024). *Labor force participation rate, female (% of female population ages 15+) (modeled ILO estimate) - Uganda*. Indicator code: SL.TLF.CACT.FE.ZS. Data source: ILO Modelled Estimates database (ILOEST), International Labour Organization. License: CC BY-4.0. Accessed May 11, 2026.

Yakubovich, V. and Lup, D. (2006). Stages of the recruitment process and the referrer's performance effect. *Organization science*, 17(6):710–723.

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 of Respondent Characteristics Across Candidate-Pair Configurations

	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)			
Observations	380	181	186	561	566	367

Notes: Columns 1–3 correspond to the three randomized experience configurations of the pair of candidate profiles shown to each respondent: (1) both candidates have the same experience; (2) the female candidate has more experience than the male candidate; (3) the male candidate has more experience than the female candidate. Each cell reports the mean of the variable across respondents in that configuration, with the standard error in parentheses. Columns 4–6 report p-values from pairwise t-tests of equality across configurations.

Table A.2: Network Availability

	Round 1 Sample				Round 2 (Main) Sample			
	Sample of Reference When Conditioning	Share in Full Sample %	Share in FDO %	Share in MDO %	Sample of Reference When Conditioning	Share in Full Sample %	Share in FDO %	Share in MDO %
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Gender Composition of Named and Referred Workers								
Network Availability								
Named a Network	547	72.2	73.3	71.6	747	38.2	33.2	40.7
Named Female Network	395	61.0	96.0	39.3	285	40.7	84.7	22.0
Named Male Network	395	75.4	40.4	97.1	285	74.4	30.6	93.0
Named Both Genders Network	395	36.5	36.4	36.5	285	15.1	15.3	15.0
Network Referred								
Referred Female Network	395	26.3	53.0	9.8	285	18.9	47.1	7.0
Referred Male Network	395	49.1	15.2	70.1	285	49.8	16.5	64.0
Referred Female Network (Cond. Named Both)	144	27.8	38.2	21.3	43	23.3	23.1	23.3
Referred Male Network (Cond. Named Both)	144	54.2	34.5	66.3	43	48.8	38.5	53.3
Panel B: Network Segregation by Network Constraint								
Constrained Network								
Named a Network	547	72.2	73.3	71.6	747	38.2	33.2	40.7
Named Female Network Only	395	24.6	59.6	2.9	285	25.6	69.4	7.0
Named Male Network Only	395	39.0	4.0	60.7	285	59.3	15.3	78.0
Unconstrained Network								
Named Both Genders Network	144	100.0	100.0	100.0	43	100.0	100.0	100.0
Referred Female Network (Cond. Named Both)	144	27.8	38.2	21.3	43	23.3	23.1	23.3
Referred Male Network (Cond. Named Both)	144	54.2	34.5	66.3	43	48.8	38.5	53.3
Panel C: Network Characteristics								
Age (Years)	532	24.8			299	23.8		
Experience in Training Sector (Months)	537	31.4			224	28.3		
Relationship: Relative	539	9.6			299	11.0		
Relationship: Friend	539	82.4			299	87.6		
Education: Primary or Less	539	3.7			296	6.1		
Education: Secondary or Tertiary	539	21.9			296	26.7		

Notes: Panels A and B report 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. *Panel C* reports descriptive characteristics of the named network contacts.

Table A.3: Respondents' Own GNC Status: Alternative Specification (Referral, Likability, Retention)

	Referred			Likability			Perceived Retention Probability		
	All Occupations (1)	FDO (2)	MDO (3)	All Occupations (4)	FDO (5)	MDO (6)	All Occupations (7)	FDO (8)	MDO (9)
GNC Candidate	-.239***	-.149**	-.286***	-.473***	-.204	-.614***	-.028***	-.014	-.035***
	(.037)	(.065)	(.045)	(.095)	(.163)	(.118)	(.009)	(.015)	(.011)
Respondent GNC	-.026	.140	-.108	-.094	.206	-.269	.025	.064	.004
	(.064)	(.104)	(.080)	(.218)	(.381)	(.268)	(.023)	(.040)	(.027)
GNC Candidate × Respondent GNC	.051	-.280	.216	.270	.061	.381	-.039	-.043	-.037
	(.128)	(.208)	(.159)	(.276)	(.380)	(.367)	(.028)	(.051)	(.034)
Linear Combinations									
GNC Candidate + GNC Candidate × Respondent GNC	-.188	-.429**	-.070	-.203	-.143	-.233	-.067**	-.057	-.072**
	[.128]	[.031]	[.648]	[.433]	[.678]	[.504]	[.012]	[.236]	[.026]
Differences Across Occupation Types									
GNC Candidate (FDO − MDO)		.137*			.410**			.021	
		[.084]			[.041]			[.255]	
Respondent GNC (FDO − MDO)		.248*			.566			.064	
		[.059]			[.222]			[.187]	
GNC Candidate × Respondent GNC (FDO − MDO)		-.496*			-.320			-.006	
		[.059]			[.544]			[.919]	
GC Mean	.619	.574	.643	7.120	6.974	7.196	.669	.665	.671
Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1494	512	982	1494	512	982	1493	512	981

Notes: Alternative specification of the heterogeneity by respondent GNC status on referral, likability, and perceived retention probability. Standard errors clustered at the respondent level. $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

Table A.4: Heterogeneity by Tenure and GNC Exposure — Summary

	Referred (1)	Likability (2)	Perceived Retention Probability (3)	Task Ability (4)	Employer Valuation (5)	Candidate Enjoyment (6)
GNC × Had GNC Classmate	.037	-.105	-.014	.138	.017	.028
	(.073)	(.179)	(.017)	(.188)	(.167)	(.139)
Observations	1,494	1,494	1,493	1,372	1,364	1,369
GNC × Shop Has GNC Coworker	-.023	.347	.027	.505**	.380*	.232
	(.098)	(.246)	(.024)	(.248)	(.220)	(.181)
Observations	808	808	808	772	771	770
GNC × Above-Median Tenure	-.009	-.232	-.025	-.160	.039	-.101
	(.087)	(.216)	(.021)	(.223)	(.194)	(.159)
Observations	1,016	1,016	1,016	942	940	940

Notes: Each row reports the coefficient on the interaction GNC Candidate × H from a full-sample OLS regression $Y = \alpha + \beta_1 \text{GNC} + \beta_2 H + \beta_3(\text{GNC} \times H) + \text{controls} + \varepsilon$, where H is the heterogeneity indicator listed in the row label. Three H variables are shown: whether the respondent had at least one gender-non-congruent classmate during vocational training; whether the respondent's current shop has at least one gender-non-congruent coworker; and whether the respondent's tenure at the current shop is above the sample median. Sample: Round 2 (our main sample). Standard errors clustered at the respondent level. Controls follow the paper's convention: no controls for the referral outcome (col. 1), full controls for the five rating outcomes (cols. 2–6; respondent gender, occupation type, baseline work status, endline follow-up status, VTI and training-area fixed effects). Observation counts vary across rows because the "shop GNC coworker" and "high tenure" variables are only defined for wage-employed respondents. Positive coefficients indicate that the GNC penalty is smaller (less negative) when $H = 1$; negative coefficients indicate the GNC penalty is larger. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

Table A.5: Heterogeneity by Occupation Type in Part I — Experience Decomposition

			Top-Level Channels				**Unpacking Perceived Retention Probability**					
	Referred		Likability		Perceived Retention Prob.		Tasks Ability		Candidate Enjoyment		Employer Valuation	
	FDO	MDO	FDO	MDO	FDO	MDO	FDO	MDO	FDO	MDO	FDO	MDO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
GNC Experimental Candidate	−.178***	−.288***	−.259	−.597***	−.028	−.037***	−.302*	−1.405***	−.145	−.484***	−.367**	−.621***
	(.064)	(.045)	(.173)	(.132)	(.017)	(.012)	(.167)	(.134)	(.141)	(.108)	(.157)	(.125)
More Experience	.301***	.276***	.554**	.471**	.041	.056***	1.042***	.295	.665***	.218	.493*	.470***
	(.062)	(.042)	(.250)	(.184)	(.026)	(.018)	(.240)	(.199)	(.249)	(.188)	(.277)	(.179)
GNC × More Experience	−.002	.082	.181	.063	.037	−.005	−.112	.303	−.183	−.141	.534	.283
	(.079)	(.055)	(.370)	(.277)	(.038)	(.030)	(.367)	(.297)	(.371)	(.292)	(.390)	(.291)
Linear Combinations												
More Exp. + GNC × More Exp.	.300***	.358***	.735***	.534***	.078***	.051**	.930***	.598***	.483**	.077	1.027***	.753***
	[.000]	[.000]	[.005]	[.007]	[.003]	[.013]	[.000]	[.005]	[.047]	[.705]	[.000]	[.000]
More Exp. + GNC × More Exp. + GNC Experimental Candidate	.121	.070	.476*	−.063	.050*	.014	.628***	−.806***	.337	−.407**	.660***	.132
	[.154]	[.257]	[.064]	[.748]	[.063]	[.466]	[.006]	[.000]	[.101]	[.026]	[.002]	[.493]
GNC Candidate (FDO − MDO)	.109		.338		.009		1.103***		.339*		.255	
	[.164]		[.121]		[.662]		[.000]		[.058]		[.205]	
GC Candidate / Baseline Experience	.515	.566	6.862	7.073	.661	.658	7.541	7.726	7.789	8.147	6.792	6.629
Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	512	982	512	982	512	981	468	904	467	902	461	903

Notes: Heterogeneity in the experience effect by occupation type using Part I (experimental candidate) data. Both panels regress the corresponding outcome on indicators for a GNC candidate, more experience, and their interaction. Panel A reports main outcomes (referral, likability, retention probability); Panel B unpacks retention into perceived task ability, candidate enjoyment, and employer valuation. Linear combinations are the joint effect of experience for a GNC candidate; adding "+ GNC Candidate" gives the level for an experienced GNC candidate relative to a less-experienced GC candidate; bracketed values are p-values. "GNC Candidate (FDO − MDO)" reports the between-occupation-type difference in the GNC main effect. Standard errors clustered at the respondent level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

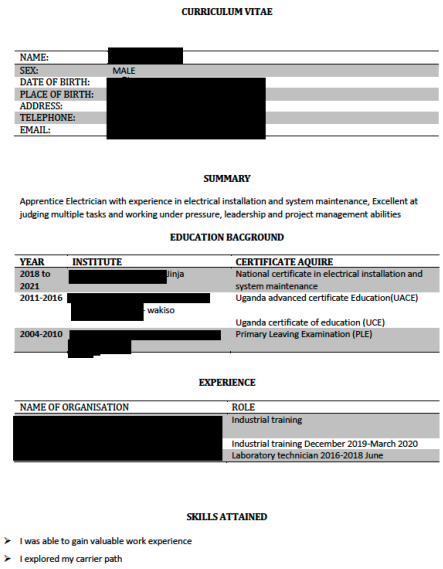
Table A.6: GC vs GNC Graduates in the ECMA Sample — Career Outcomes and Skills

	MDO					FDO				
	GC		GNC		p	GC		GNC		p
	Mean	N	Mean	N		Mean	N	Mean	N	
% Wage Employed	26.6	2758	29.2	373	.365	23.3	1813	30.4	171	.077
	(44.2)		(45.5)			(42.3)		(46.1)		
SAT Score (0-100)	32.6	430	33.4	63	.824	37.1	310	36.9	26	.972
	(19.8)		(25.2)			(24.1)		(20.6)		
Earnings | WE (USD)	97.6	694	71.6	105	.000	58.3	402	96.0	49	.002
	(68.4)		(61.7)			(48.3)		(82.2)		
Earnings | Training Sector (USD)	90.8	317	92.7	24	.916	53.2	242	77.7	28	.016
	(76.4)		(97.9)			(48.7)		(43.8)		
Career Satisfaction (0-10)	4.1	2216	3.7	302	.095	3.6	1472	3.7	139	.973
	(2.8)		(2.9)			(2.9)		(2.9)		
% Performed Physical Work	9.7	734	7.3	109	.363	7.3	423	5.8	52	.638
	(29.6)		(26.2)			(26.1)		(23.5)		

Notes: The unit of observation is the worker-wave; the ECMA data (Alfonsi et al. 2020) form a panel of 1,140 vocational graduates followed over five survey rounds (5,700 worker-waves). Each cell reports the group mean, the standard deviation in parentheses, and the number of non-missing worker-waves (N). Sample sizes differ across rows because items were collected in different survey rounds or conditioned on employment. p is the p-value for the GC-versus-GNC difference within occupation type, from a two-sample t-test that treats each worker-wave as an independent observation.

Figure A.1: Example of Real and Experimental Profiles

(a) Real Profile, Page 1



CURRICULUM VITAE

NAME:	[redacted]
SEX:	MALE
DATE OF BIRTH:	[redacted]
PLACE OF BIRTH:	
ADDRESS:	
TELEPHONE:	
EMAIL:	

SUMMARY

Apprentice Electrician with experience in electrical installation and system maintenance, Excellent at judging multiple tasks and working under pressure, leadership and project management abilities

EDUCATION BACGROUND

YEAR	INSTITUTE	CERTIFICATE AQUIRE
2018 to 2021	[redacted]linja	National certificate in electrical installation and system maintenance
2011-2016	[redacted] wakiso	Uganda advanced certificate Education(UACE) Uganda certificate of education (UCE)
2004-2010	[redacted]	Primary Leaving Examination (PLE)

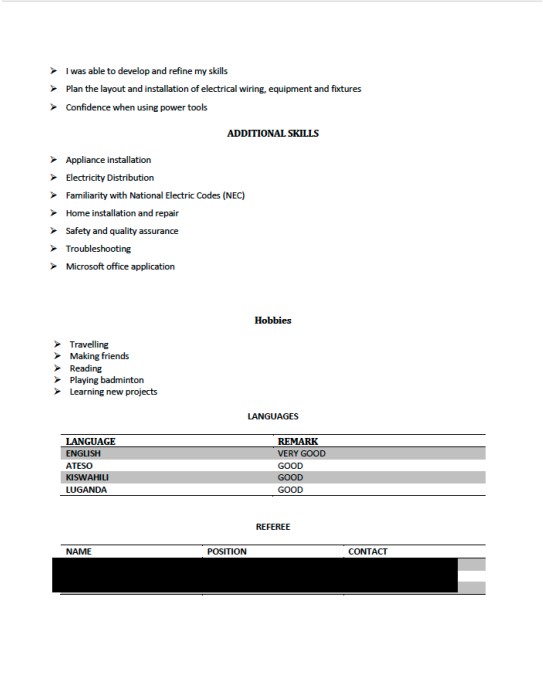
EXPERIENCE

NAME OF ORGANISATION	ROLE
[redacted]	Industrial training
	Industrial training December 2019-March 2020
	Laboratory technician 2016-2018 June

SKILLS ATTAINED

- I was able to gain valuable work experience
- I explored my carrier path

(b) Real Profile, Page 2



- I was able to develop and refine my skills
- Plan the layout and installation of electrical wiring, equipment and fixtures
- Confidence when using power tools

ADDITIONAL SKILLS

- Appliance installation
- Electricity Distribution
- Familiarity with National Electric Codes (NEC)
- Home installation and repair
- Safety and quality assurance
- Troubleshooting
- Microsoft office application

Hobbies

- Travelling
- Making friends
- Reading
- Playing badminton
- Learning new projects

LANGUAGES

LANGUAGE	REMARK
ENGLISH	VERY GOOD
ATESO	GOOD
KISWAHILI	GOOD
LUGANDA	GOOD

REFEREE

NAME	POSITION	CONTACT
[redacted]		

(c) Experimental Profile



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

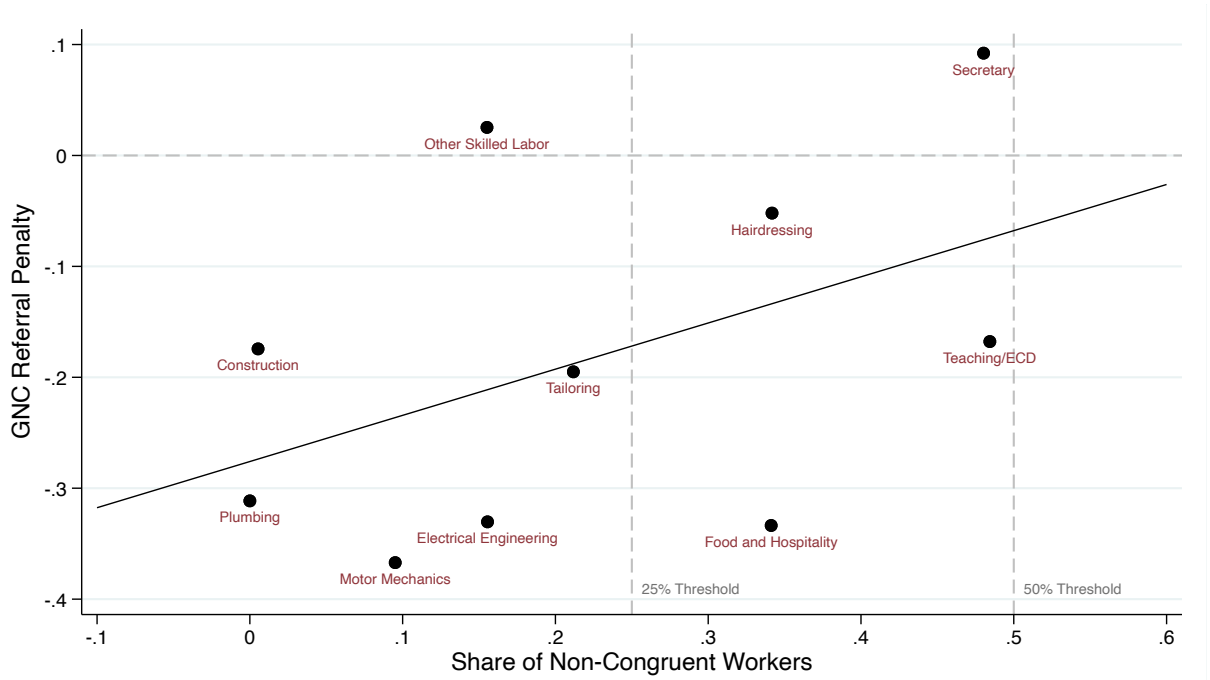
April - August 2023: Bawumaye Metal Fabrication: Full-time worker. Main activities: performing welding repairs and modifications, aligning parts prior to assembly, and melting and soldering metal plates.
January - March 2023: Bawumaye Metal Fabrication: Trainee. Main activities: performing weld inspections 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.

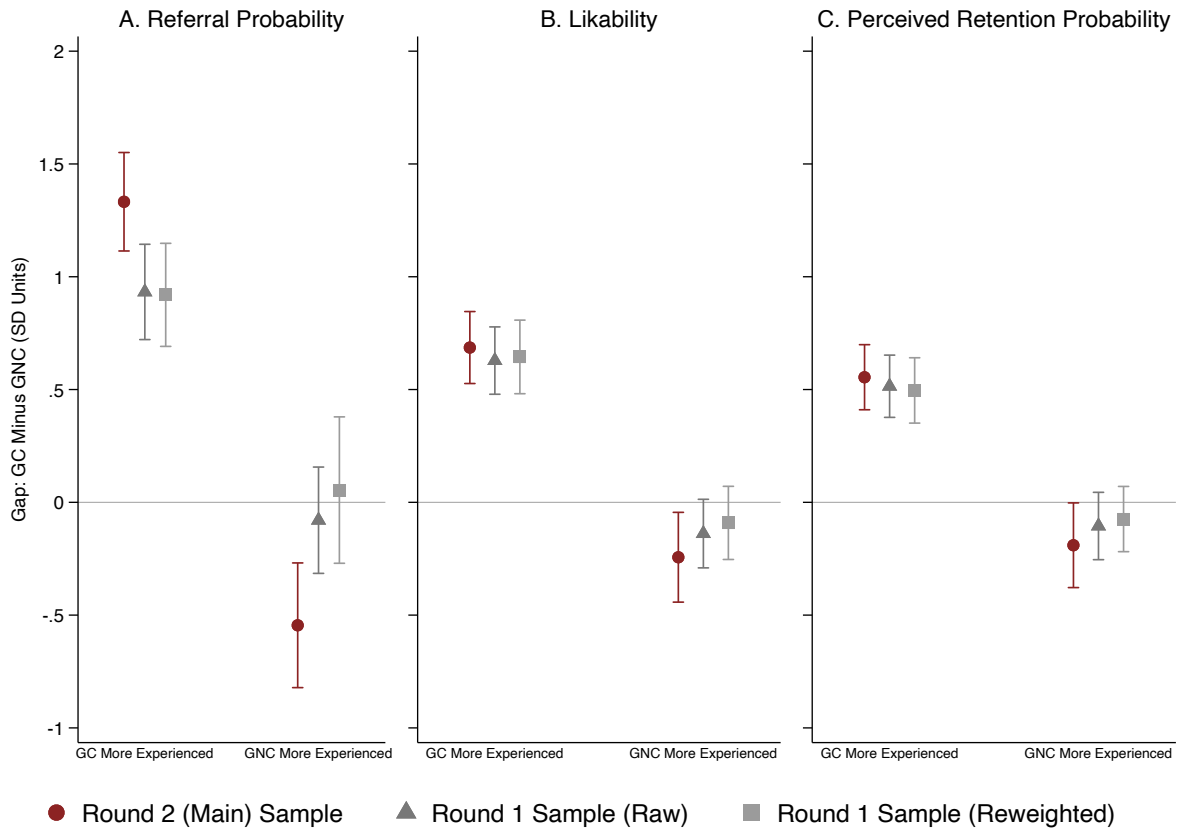
Notes: Panels A and B present an example of a real CV from a VTI graduate. Panel C shows an example of the experimental candidate profile shown to participants.

Figure A.2: Gender Non-Congruent Referral Penalty and Employment Share



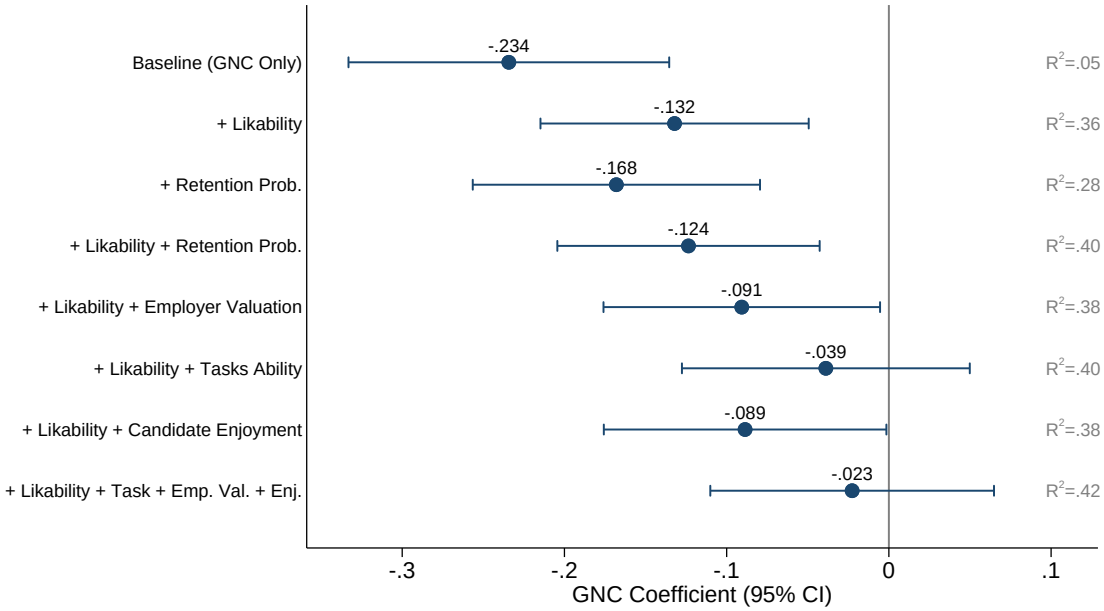
Notes: Each point represents an occupation-level referral gender bias, where negative values indicate a penalty against non-congruent workers. The occupation-level penalty is plotted against the share of non-congruent workers in that occupation. Fitted line slope = .416 ($p = .179$). Vertical dashed lines mark the 25% and 50% thresholds for the share of non-congruent workers.

Figure A.3: GC–GNC Gaps Across Samples



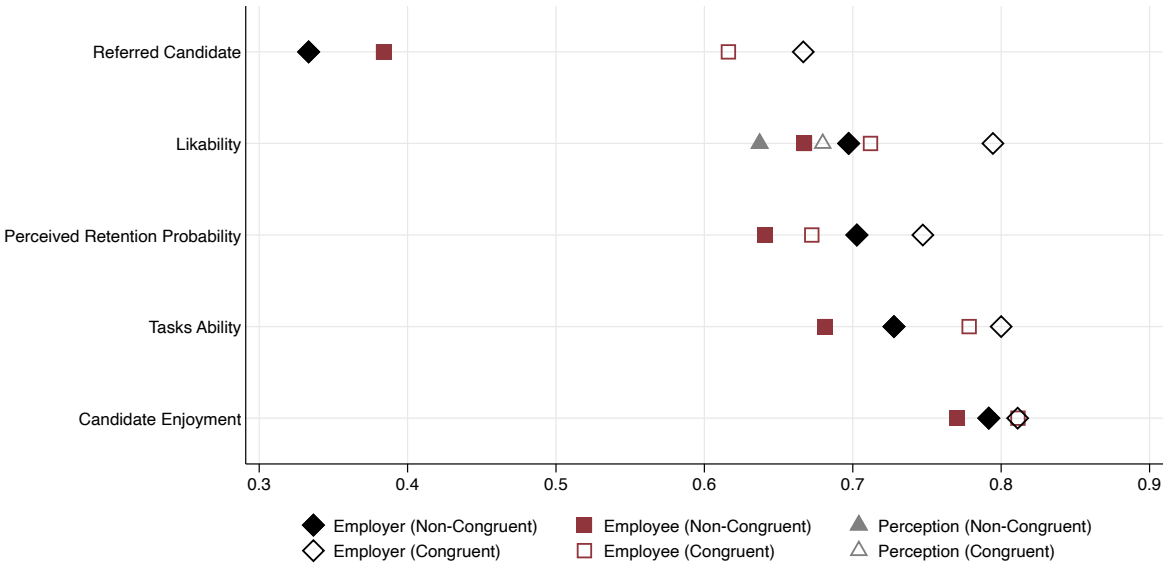
Notes: Each point is a linear combination from the specification $Y_{p,i} = \alpha + \beta_1 \cdot \text{non-congruent} + \beta_2 \cdot \text{more experience} + \text{controls} + \varepsilon_{p,i}$, estimated separately for each sample on all occupations (female- and male-dominated pooled). "GC More Experienced" plots $\beta_2 - \beta_1$; "GNC More Experienced" plots $-(\beta_1 + \beta_2)$. The Round 2 sample (our main sample) is restricted to the two treatment arms present in both waves, i.e., the asymmetric-experience configurations in which one profile has more experience than the other, (High, Low) and (Low, High). The Round 1 sample is shown both with raw estimates and reweighted by training-area composition to match the occupational mix of the Round 2 sample. All estimates are standardized by the standard deviation of each outcome in the restricted Round 2 sample.

Figure A.4: Mechanism Decomposition — GNC Coefficient with Progressively Added Ratings



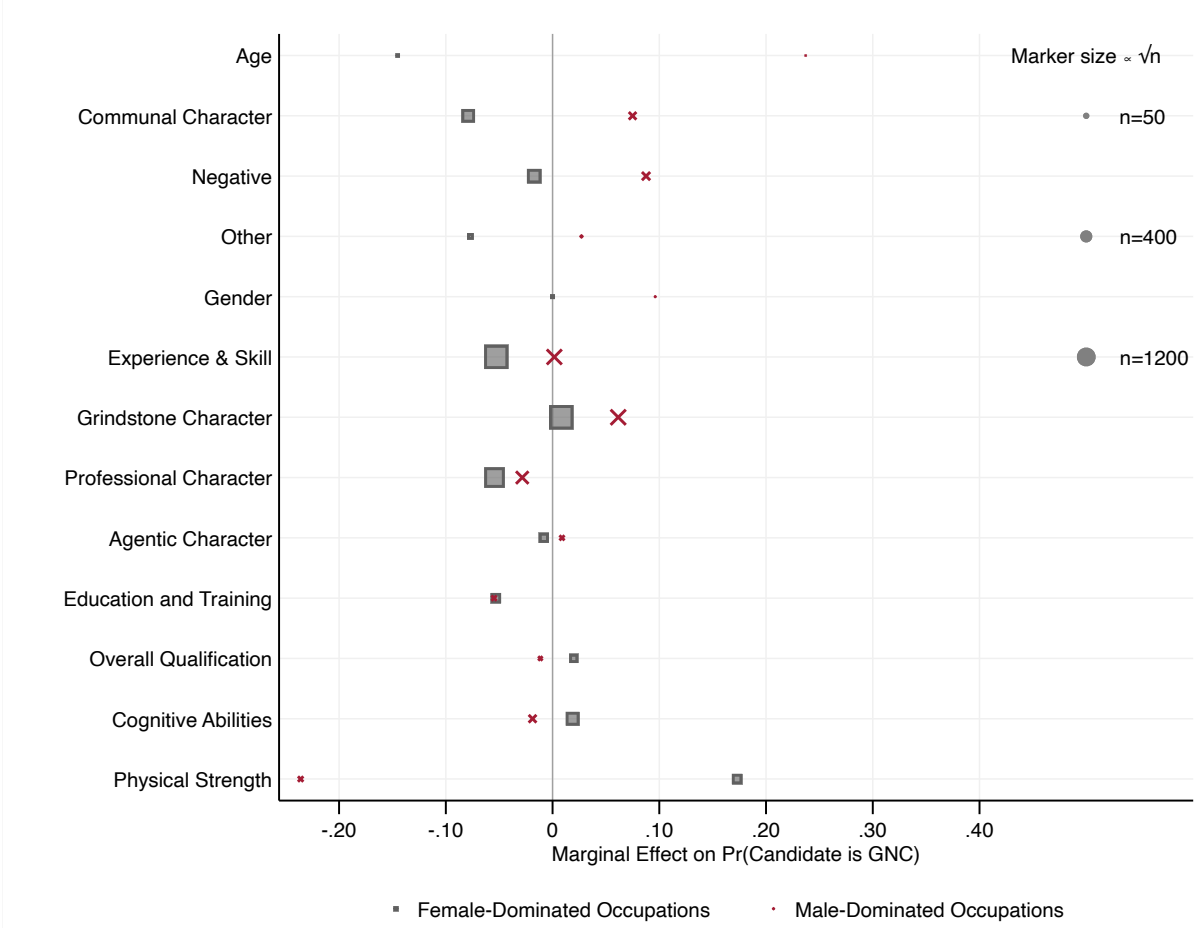
Notes: Each row plots the OLS coefficient on the GNC Candidate indicator from a regression of the Part I referral outcome on GNC Candidate and the ratings listed on the left, with respondent fixed effects absorbed and standard errors clustered at the respondent level. Whiskers show 95% confidence intervals; the numeric coefficient is printed above each point. Respondent fixed effects are the linear analogue of a conditional logit at the choice-set level, respecting the paired within-respondent design. The first block (rows 2–4) introduces likability and perceived retention separately and then jointly. The second block (rows 5–7) holds likability fixed and adds one retention subcomponent at a time — anticipated employer valuation, perceived task ability, and anticipated candidate enjoyment. The final row includes likability together with all three retention subcomponents. The rightmost annotation reports the regression R^2: it climbs from about .06 in the baseline row (GNC alone) to about .42 once likability and all three retention subcomponents are conditioned on, indicating that ratings are highly predictive of the referral choice. The residual GNC coefficient shrinks toward zero as ratings are added and is not statistically distinguishable from zero once likability is combined with perceived task ability, or with all four ratings. Sample: Round 2 (our main sample) Part I candidate-profile rows (N = 1,494; 747 respondents). Table OA.11 reports the full regression output.

Figure A.5: What Do Employers Really Think?



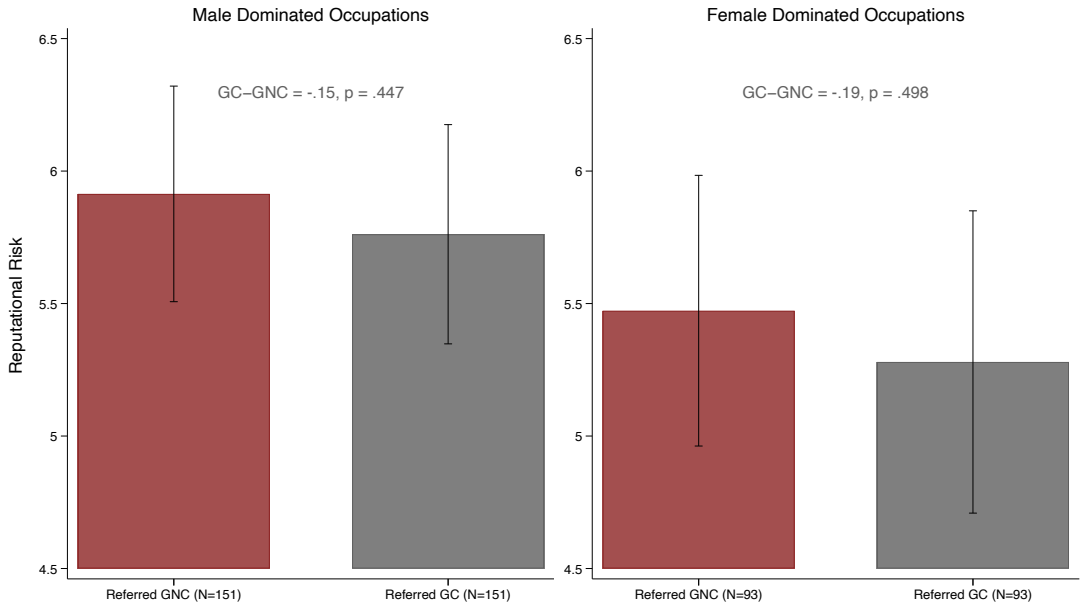
Notes: This figure summarizes how employers and employees evaluate gender-congruent and non-congruent candidates. Each marker plots the average probability or rating for a given outcome on the vertical axis (referral probability, likability, retention probability, perceived task ability, perceived candidate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

Figure A.6: GNC Predictors of LLM-Assigned (Human-Validated) Adjective Categories



Notes: The figure reports average marginal effects from logistic regressions of the probability that a CV belongs to a gender-non-congruent (GNC) candidate on indicators for LLM-assigned adjective categories, estimated separately for female-dominated occupations (grey squares) and male-dominated occupations (red crosses). Categories are ordered by the signed difference between the male-dominated-occupation and female-dominated-occupation estimates, with the largest positive differences at the top and the most negative at the bottom. Positive estimates indicate categories more commonly assigned to GNC candidates; negative estimates indicate categories more commonly assigned to gender-congruent candidates. Marker size is proportional to $\sqrt{n}$, where n is the number of adjective mentions underlying each estimate, so categories used more often appear as larger markers. Adjective-mention counts and shares underlying each estimate are, in descending order: Grindstone Character (1,254; 25.5%), Experience & Skill (1,220; 24.8%), Professional Character (840; 17.1%), Negative (360; 7.3%), Cognitive Abilities (324; 6.6%), Communal Character (309; 6.3%), Physical Strength (158; 3.2%), Education and Training (152; 3.1%), Agentic Character (138; 2.8%), Overall Qualification (103; 2.1%), Other (40; 0.8%), Gender (11; 0.2%), and Age (6; 0.1%), for a total of 4,915 mentions elicited across the 1,506 CVs in our main (Round 2) sample. Adjectives flagged as negative are classified in the Negative category.

Figure A.7: Reputational Risks — Vignette Experiment



Notes: Mean perceived reputational risk to a referrer in a hypothetical referral scenario. Respondents were presented with a vignette in which an incumbent employee of the respondent's same gender, working in the respondent's occupation, refers a new worker for a job in that same occupation. The gender of the referred candidate is randomized; the gender of the vignette's protagonist (the referrer) matches the respondent's own gender. Respondents were asked, on a 0–10 scale, how much they thought the referrer's employer would think worse of the referrer if the referred candidate underperformed. Higher values indicate greater perceived reputational risk. Bars compare referrals to gender-congruent (GC) and non-congruent (GNC) candidates; whiskers show 95% confidence intervals.

Material for Online Appendix

A Methodological Details for the Referral Experiment

This appendix provides additional details on the implementation of the referral experiment. The design had three goals. First, it created a real referral opportunity with meaningful stakes for respondents, employers, and candidates. Second, it allowed us to compare gender-congruent (GC) and gender non-congruent (GNC) candidates while holding constant the core qualifications relevant for the job. Third, it separated the network-composition margin from the referral-choice margin by presenting each respondent with qualified candidates of both genders before eliciting nominations from the respondent's own network.

A.1 The Talent Lift Referral Program

The experiment was implemented through a program we created with BRAC Uganda called the Talent Lift Referral Program. The program offered a six-week subsidized internship to young vocational graduates in the respondent's field of training. For wage-employed respondents, the internship was attached to the business where the respondent worked. If selected, the employer would receive UGX 200,000 to host the trial, the candidate UGX 300,000 over the six-week period, and the incumbent employee UGX 100,000 (about five days of wage) if the candidate was retained afterward. The selected profile entered a lottery for implementation, so respondents' choices could affect which candidate was offered a real internship. Respondents who were not currently employed completed a parallel version of the module. Because they are vocationally trained and active in the labor market, they were asked to help select a candidate for a vacancy at a firm of the type that typically hires graduates from their training track. They were told their choices would be considered when assessing candidates for the program. Online Appendix F.1 reports a robustness check showing that the main results are unchanged when these respondents are excluded.

A.2 Relation to Audit Studies and IRR Designs

The labor-market resume-correspondence tradition was formalized in economics by Bertrand and Mullainathan (2004) and has since been applied to discrimination on race, immigrant status, gender, age, parenthood, criminal record, unemployment history, and educational credentials across a wide range of settings (e.g. Riach and Rich, 2006; Booth and Leigh, 2010; Carlsson, 2011; Oreopoulos, 2011; Eriksson and Rooth, 2014; Deming et al., 2016; Agan and Starr, 2018; Becker et al., 2019; Neumark et al., 2019), including in low- and middle-income labor markets (Arceo-Gomez and Campos-Vazquez, 2014; Banerjee et al., 2009; Galarza and Yamada, 2017; Lahey, 2008; Campos-Vazquez and Gonzalez, 2020; Bedi et al., 2022). Recent methodological work has refined inference and identification within this design (Neumark, 2012; Kline and Walters, 2021), and large-scale applications have documented substantial between-firm variation in discrimination (Kline et al., 2022). Bertrand and Duflo (2017) and Neumark (2018) provide comprehensive surveys. The incentivized resume-rating methodology that refines the classic audit-study approach is due to Kessler et al. (2019). Closely related designs have been deployed in other markets, including housing and the sharing economy (Edelman et al., 2017), credit access (Macchi, 2023), ride-sharing platforms (Ge et al., 2020), and more. Our design adapts these toolkits to a new decision-maker —

the incumbent employee — and a new choice: the referral. Whereas classic audit studies estimate discrimination in employer callback decisions, we present each respondent with a within-choice-set contrast of gender-congruent and gender non-congruent candidates and elicit rich evaluative outcomes at the incumbent-worker margin, spanning own likability, perceived retention probability, and beliefs about task ability, candidate enjoyment, and employer valuation.

A.3 Experimental Candidate Profiles: Structure and Source

In Part I, each respondent evaluated two standardized candidate profiles for the respondent's training area. The profiles were designed to resemble CVs used by young vocational graduates in Ugandan labor markets. We built the profile structure using real CV templates collected by the field team and information common among recent vocational graduates. Each profile included four sections: personal information (name, sex, date of birth, nationality, home district); academic background (vocational training institute, high school); work experience (firm name, role title, duration, main activities); and language abilities.

The profiles matched the respondent's area of specialization. Respondents were presented with candidates graduating from their exact same program (occupation and type of course) and vocational training institute. Home district was conditional on the respondent's training institute and could be either Kampala or Jinja. Nationality and languages were held constant across profiles (Ugandan; English, Luganda, and Lusoga). The candidates' first names were drawn from a small pool of six common Ugandan names balanced by gender: Joel, Paul, and George for male profiles, and Robina, Rachel, and Rose for female profiles. The name of the high school was conditional on the district and randomly drawn from a pool of two names per district. The name of the firm shown on the profile was conditional on the occupation and randomly drawn from a pool of two names per training area. Task descriptions were conditional on the training area and randomly drawn from a pool of six tasks per training area. If the candidate was in the higher-experience condition, two tasks from the first job were randomly kept and a third new task was added. Profiles were displayed in the survey software and sent to respondents through SMS or WhatsApp. Enumerators confirmed receipt of the correct profile before proceeding with the evaluation. Appendix Figure A.1 shows an example real CV alongside an example experimental profile.

A.4 What Was Real, Masked, and Randomized

The experimental profiles should be read as standardized representations of the underlying candidate types rather than as records of any single, named individual. We follow the logic of incentivized resume rating (Kessler et al., 2019): each profile is defined by the substantive characteristics of the pool of vocational graduates available for placement in the respondent's field of training, and any profile the respondent selects can be matched, via lottery, to a specific graduate in that pool sharing those substantive characteristics. "Real," in what follows, describes the candidate type, not a specific individual.

The substantive matching characteristics — accredited vocational training, occupation of training, and post-training experience (approximately) in that occupation — were drawn from the actual candidate pool. These are the dimensions employers in this setting weigh

most heavily (Alfonsi et al., 2026b). Names, the exact date of birth within a narrow age window, high-school identifiers, firm names, and task descriptions were randomized within realistic occupation- or district-specific pools to protect candidate privacy, avoid duplicate resumes, and keep non-focal traits balanced across gender conditions. Candidate gender and experience were cross-randomized across the two profiles each respondent evaluated. Table OA.1 summarizes each component.

Table OA.1: Candidate Profile Components: Real, Randomized, and Masked

Component	**"Real"**	**How it was constructed**	**Why**
Occupation / training area	Yes	Matched to the respondent's training area	Ensures the candidate is relevant for the referral opportunity
Vocational credential	Yes	Two-year National Certificate in the relevant field, accredited	Holds constant the core qualification required for the internship
Candidate gender	Yes	Randomized through name and sex field	Identifies the GNC referral penalty
Post-training experience	∼Yes	Randomized in five-month increments (low vs. high)	Tests whether experience attenuates the GNC penalty
Vocational training institute	Yes	Matched to respondent's VTI	Approximates a plausible referral setting where respondent and candidate share training background
Candidate first name	No	Drawn from a small pool of gender-balanced common Ugandan names	Signals gender and protects candidate identity
High school, firm name, exact date of birth (within narrow age window), task descriptions	No	Randomized within occupation- or district-specific pools of realistic names/tasks	Preserves realism, avoids duplicate profiles, keeps non-focal traits balanced
Nationality, languages, home district (given VTI)	Yes	Held constant across profiles	Prevents confounding from non-focal characteristics

Respondent framing differed across the two rounds of the experiment. In Round 1, we adopted the fully deception-free IRR framing of Kessler et al. (2019): respondents were told that the profiles were hypothetical but drawn to be close in characteristics to the candidates with whom they would later be matched. In Round 2 (our main sample), we told respondents that the profiles corresponded to real candidates who wanted to work at their shop and whose names had been masked for privacy; the underlying matching procedure was the same as in Round 1. We changed the framing for two reasons. First, the fully hypothetical framing was harder to explain in the field and lengthened survey time. Second, and more substantively, because the design rotates rare profile types (women in male-dominated occupations and men in female-dominated occupations), we did not want preferences over these profiles to be confounded with respondents' beliefs about whether

such candidates could plausibly be located. Unlike in Kessler et al.'s setting, here beliefs about the very existence of gender non-congruent candidates could shape responses. The main patterns are materially similar across the two rounds.

A.5 Respondent Instructions

Below we reproduce the core text read to respondents at the start of the referral module implemented with Round 2 (Main Sample). Bracketed text indicates experimentally randomized or respondent-specific inserts.

Employed respondents. We are launching the third round of the Talent Lift Referral Program, which offers a six-week subsidized internship to highly competitive young candidates in your field of training. In the previous year, we successfully created 50 internship positions, and this year we aim to achieve more than that. The business in which you currently work as an employee is eligible for the program. The business will receive UGX 200,000 for training the intern, and the selected candidate will be paid UGX 300,000. We want you to indicate which candidate is best for the program, leveraging your knowledge of the business. We will show you two profiles of competitive candidates who may be a good fit for this internship. The names shown in the profiles are hypothetical, but the profiles correspond to real candidates who have the same substantive characteristics as the profiles shown, except for the name. This is done to protect candidates' privacy. We will run a lottery, and the candidate you select may become eligible to be hired as an intern in your company. Your referred candidate has a 50 percent chance of being selected for the internship. If the candidate is selected, their name will be revealed. If the candidate is retained in your company after the internship, you will receive UGX 100,000 as a reward for making a good referral. [Confidentiality condition inserted here.]

Respondents not currently wage-employed. We are launching a program offering six-week subsidized internships to highly competitive young candidates in businesses in your training area. The business will receive UGX 200,000 for training the intern, and the selected candidate will be paid UGX 300,000. We want you to indicate which candidate is best for the program in a business in your training area, leveraging your knowledge of the area. We will show you two profiles of competitive candidates who may be a good fit for this internship. The names shown in the profiles are hypothetical, but the profiles correspond to real candidates with the same substantive characteristics as the profiles shown, except for the name. This is done to protect candidates' privacy. Please give your sincere choices and opinions. Very important: in this exercise, imagine you are an employee in a company in your training area and you have to choose someone to work with you. Keep in mind that this exercise is purely hypothetical, but your choices will be considered when we assess candidates. [Confidentiality condition inserted here.]

Confidentiality conditions. Respondents were randomly assigned to one of four conditions describing how the employer would be informed of the referral.

- *Public:* The employer (either business owner or high manager) will know that you referred the candidate. Specifically, we will explain that we offer a subsidized internship position to the company and that you referred your chosen candidate.
- *Anonymous worker (intermediate):* We will not inform the employer that you specifically referred the candidate and will never mention your name, but we will tell the employer that the referral came from a worker in the company.
- *Private:* The business owner will not know that you referred the candidate.
- *BRAC selected:* The employer will be told that the matched candidate was selected by BRAC, with no mention of any worker in the company.

A.6 Comprehension and Recall of the Experience Signal

We implemented direct checks to assess whether respondents understood and retained the experimentally assigned experience signal. After respondents had evaluated the two experimental profiles, they were asked to recall, without looking back at the message, how many months of work experience Candidate 1 and Candidate 2 had. The two questions, phrased in parallel, asked: "Without looking at the text message that we sent to you, do you remember how many months of experience Candidate [1/2] (`[name]`) had?" Respondents were explicitly told that they would not be penalized for incorrect answers and that, if they did not remember, they should give their best estimate. When respondents did not report exact months for at least one candidate, we asked a follow-up ordinal recall question: which of the two candidates had more work experience? Respondents reliably identified the higher-experience profile, and recall did not differ by candidate gender or by whether the candidate was gender-congruent or gender non-congruent.

A.7 Implementation of Referrals and Internship Placements

Referral choices were linked to a real implementation effort. We attempted to create 178 placements through the lottery. Of these attempts, 12 percent could not be matched to an available and willing vocational trainee in the relevant district and sector; 10 percent involved an identified candidate who refused, became unavailable, or could not be replaced; 30 percent failed on the referrer side — the referring worker could not be reached, had left the relevant job, or no longer wished to proceed; and 34 percent failed on the employer side — the firm refused, did not respond, had no vacancy or training capacity, was not a suitable placement site, or imposed requirements our candidates did not meet. Ultimately, 27 internships were started and completed. The low conversion rate reflects the operational difficulty of creating short-term placements in small-firm labor markets. In retrospect, we underestimated this burden: the process required substantial coordination across referrers, employers, and candidates, and was handled by a small implementation team, so several weeks elapsed before all placement attempts could be completed, likely increasing attrition on both sides. The conversion rate is consistent with evidence from related worker-firm matching and firm-training interventions in low-income labor markets; in Uganda, for example, Alfonsi et al. (2020) find that only 24 percent of workers assigned to firm-provided

training were ultimately hired and trained by the matched firm. The implementation difficulty does not affect the interpretation of the referral-choice estimates, which are based on respondents' choices under the program incentives, but it is relevant for interpreting implementation feasibility.

A.8 Profile Balance and Descriptive Statistics

Tables OA.2 and OA.3 report the descriptive statistics for the 1,494 profiles in the main sample, listing each component that was randomized and its realized probability.

Table OA.2: CVs Characteristics: Skills

Variable	(1) Female Mean/SE	(2) Male Mean/SE	(3) Total Mean/SE	T-test P-value (1)-(2)
Area 1: Motor-mechanics: fixing engines	.110	.118	.114	.625
	(.011)	(.012)	(.008)	
Area 2: Plumbing: fixing pipes	.092	.076	.084	.264
	(.011)	(.010)	(.007)	
Area 3: Food: preparing ingredients	.060	.055	.058	.657
	(.009)	(.008)	(.006)	
Area 4: Tailoring: repairing defective garments	.079	.078	.078	.923
	(.010)	(.010)	(.007)	
Area 5: Hairdressing: shampooing customers' hair	.016	.021	.019	.446
	(.005)	(.005)	(.004)	
Area 6: Construction: mixing and leveling concrete	.058	.050	.054	.491
	(.009)	(.008)	(.006)	
Area 7: Electrical Work: inspecting and fixing malfunctioning electrical element	.124	.124	.124	1.000
	(.012)	(.012)	(.009)	
Area 8: Welding: cutting metal plates	.027	.019	.023	.298
	(.006)	(.005)	(.004)	
Area 9: Carpentry: joining wooden pieces together with glue, dowels, and nails	.001	.003	.002	.564
	(.001)	(.002)	(.001)	
Area 10: Teaching: preparing lesson plans	.046	.050	.048	.715
	(.008)	(.008)	(.006)	
Area 11: Secretary: answering calls and taking messages	.028	.023	.025	.511
	(.006)	(.005)	(.004)	
Area 12: Agriculture: planting and cultivating crops	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: threading and drilling	.011	.011	.011	1.000
	(.004)	(.004)	(.003)	
Area Retail: shelf replenishment	.493	.503	.498	.679
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: aligning wheels	.104	.118	.111	.411
	(.011)	(.012)	(.008)	
Area 2: Plumbing: replacing defective parts	.070	.084	.077	.286
	(.009)	(.010)	(.007)	
Area 3: Food: cutting vegetables	.056	.060	.058	.741
	(.008)	(.009)	(.006)	
Area 4: Tailoring: fixing pockets, zippers, and other parts	.064	.076	.070	.363
	(.009)	(.010)	(.007)	
Area 5: Hairdressing: cleaning cutting stations	.013	.015	.014	.826
	(.004)	(.004)	(.003)	
Area 6: Construction: loading and unloading building materials	.056	.054	.055	.820
	(.008)	(.008)	(.006)	
Area 7: Electrical Work: installing electrical systems	.134	.143	.139	.600
	(.012)	(.013)	(.009)	
Area 8: Welding: applying solder along edges of workpieces	.021	.019	.020	.712
	(.005)	(.005)	(.004)	
Area 9: Carpentry: building and installing cabinets	.001	.003	.002	.564
	(.001)	(.002)	(.001)	
Area 10: Teaching: delivering lectures and class activities	.044	.046	.045	.901
	(.008)	(.008)	(.005)	
Area 11: Secretary: handling emails and correspondence	.023	.021	.022	.860
	(.005)	(.005)	(.004)	
Area 12: Agriculture: irrigating and fertilizing fields	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: operating lathes and milling machines	.009	.009	.009	1.000
	(.004)	(.004)	(.002)	
Area Retail: packing customers' goods	.499	.501	.500	.959
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: replacing carburators and generators	.119	.111	.115	.627
	(.012)	(.012)	(.008)	
Area 2: Plumbing: mending pipes	.080	.083	.082	.850
	(.010)	(.010)	(.007)	
Area 3: Food: washing dishes	.058	.055	.056	.822
	(.009)	(.008)	(.006)	
Area 4: Tailoring: hemming clothes	.074	.072	.073	.921
	(.010)	(.009)	(.007)	
Area 5: Hairdressing: coloring and highlighting hair	.012	.017	.015	.391
	(.004)	(.005)	(.003)	
Area 6: Construction: installing electrical wiring	.059	.046	.052	.245
	(.009)	(.008)	(.006)	
Area 7: Electrical Work: troubleshooting electrical issues	.123	.126	.124	.876
	(.012)	(.012)	(.009)	
Area 8: Welding: melting and soldering metal plates	.015	.019	.017	.545
	(.004)	(.005)	(.003)	
Area 9: Carpentry: crafting custom furniture	.003	.003	.003	1.000
	(.002)	(.002)	(.001)	
Area 10: Teaching: assessing student performance	.055	.036	.046	.082*
	(.008)	(.007)	(.005)	
Area 11: Secretary: managing appointment schedules	.024	.027	.025	.743
	(.006)	(.006)	(.004)	
Area 12: Agriculture: harvesting and storing produce	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: fabricating precision parts	.012	.011	.011	.807
	(.004)	(.004)	(.003)	
Area Retail: keeping vegetables fresh	.498	.501	.499	.918
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: changing oil	.118	.122	.120	.811
	(.012)	(.012)	(.008)	
Area 2: Plumbing: helping install pipe sections	.079	.067	.073	.371
	(.010)	(.009)	(.007)	
Area 3: Food: taking orders	.056	.056	.056	1.000
	(.008)	(.008)	(.006)	
Area 4: Tailoring: sewing buttons and fasteners	.070	.074	.072	.764
	(.009)	(.010)	(.007)	
Area 5: Hairdressing: performing scalp treatments	.015	.012	.013	.653
	(.004)	(.004)	(.003)	
Area 6: Construction: framing and erecting structures	.062	.047	.054	.209
	(.009)	(.008)	(.006)	
Area 7: Electrical Work: upgrading electrical panels	.143	.137	.140	.709
	(.013)	(.013)	(.009)	
Area 8: Welding: aligning parts prior to assembly	.023	.024	.023	.864
	(.005)	(.006)	(.004)	
Area 9: Carpentry: installing wooden flooring	.004	.003	.003	.654
	(.002)	(.002)	(.001)	
Area 10: Teaching: providing individualized instruction	.050	.046	.048	.715
	(.008)	(.008)	(.006)	
Area 11: Secretary: maintaining filing systems	.032	.028	.030	.650
	(.006)	(.006)	(.004)	
Area 12: Agriculture: caring for livestock	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: assembling mechanical components	.015	.009	.012	.343
	(.004)	(.004)	(.003)	
Area Retail: receiving payments	.495	.498	.497	.918
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: inspecting tires	.116	.116	.116	1.000
	(.012)	(.012)	(.008)	
Area 2: Plumbing: clearing clogged drains	.080	.084	.082	.778
	(.010)	(.010)	(.007)	
Area 3: Food: serving customers	.055	.056	.056	.910
	(.008)	(.008)	(.006)	
Area 4: Tailoring: altering clothes	.078	.076	.077	.923
	(.010)	(.010)	(.007)	
Area 5: Hairdressing: styling and setting hair	.015	.015	.015	1.000
	(.004)	(.004)	(.003)	
Area 6: Construction: painting and finishing surfaces	.064	.066	.065	.916
	(.009)	(.009)	(.006)	
Area 7: Electrical Work: wiring new construction projects	.137	.143	.140	.709
	(.013)	(.013)	(.009)	
Area 8: Welding: performing weld inspections and quality control checks	.015	.023	.019	.253
	(.004)	(.005)	(.004)	
Area 9: Carpentry: framing structures	.004	.001	.003	.317
	(.002)	(.001)	(.001)	
Area 10: Teaching: organizing educational materials	.048	.060	.054	.304
	(.008)	(.009)	(.006)	
Area 11: Secretary: assisting with bookkeeping tasks	.024	.029	.027	.522
	(.006)	(.006)	(.004)	
Area 12: Agriculture: operating farm machinery	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: performing quality control checks	.009	.011	.010	.795
	(.004)	(.004)	(.003)	
Area Retail: selling phone accessories	.514	.506	.510	.756
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: brake system repairs and replacements	.103	.116	.110	.408
	(.011)	(.012)	(.008)	
Area 2: Plumbing: repairing leaky faucets	.080	.083	.082	.850
	(.010)	(.010)	(.007)	
Area 3: Food: cleaning tables and dining areas	.055	.052	.054	.818
	(.008)	(.008)	(.006)	
Area 4: Tailoring: pressing and ironing garments	.078	.076	.077	.923
	(.010)	(.010)	(.007)	
Area 5: Hairdressing: providing hair consultations	.013	.015	.014	.826
	(.004)	(.004)	(.003)	
Area 6: Construction: roofing installation and repairs	.060	.064	.062	.748
	(.009)	(.009)	(.006)	
Area 7: Electrical Work: installing and maintaining lighting fixtures	.124	.130	.127	.756
	(.012)	(.012)	(.009)	
N	747	747	1494	

Notes: The value displayed for t-tests are p-values. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table OA.3: CVs Characteristics: Skills (Continued)

Variable	(1) Female Mean/SE	(2) Male Mean/SE	(3) Total Mean/SE	T-test P-value (1)-(2)
Area 8: Welding: cutting metal plates	.027	.019	.023	.298
	(.006)	(.005)	(.004)	
Area 9: Carpentry: joining wooden pieces together with glue, dowels, and nails	.001	.003	.002	.564
	(.001)	(.002)	(.001)	
Area 10: Teaching: preparing lesson plans	.046	.050	.048	.715
	(.008)	(.008)	(.006)	
Area 11: Secretary: answering calls and taking messages	.028	.023	.025	.511
	(.006)	(.005)	(.004)	
Area 12: Agriculture: planting and cultivating crops	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: threading and drilling	.011	.011	.011	1.000
	(.004)	(.004)	(.003)	
Area Retail: shelf replenishment	.493	.503	.498	.679
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: aligning wheels	.104	.118	.111	.411
	(.011)	(.012)	(.008)	
Area 2: Plumbing: replacing defective parts	.070	.084	.077	.286
	(.009)	(.010)	(.007)	
Area 3: Food: cutting vegetables	.056	.060	.058	.741
	(.008)	(.009)	(.006)	
Area 4: Tailoring: fixing pockets, zippers, and other parts	.064	.076	.070	.363
	(.009)	(.010)	(.007)	
Area 5: Hairdressing: cleaning cutting stations	.013	.015	.014	.826
	(.004)	(.004)	(.003)	
Area 6: Construction: loading and unloading building materials	.056	.054	.055	.820
	(.008)	(.008)	(.006)	
Area 7: Electrical Work: installing electrical systems	.134	.143	.139	.600
	(.012)	(.013)	(.009)	
Area 8: Welding: applying solder along edges of workpieces	.021	.019	.020	.712
	(.005)	(.005)	(.004)	
Area 9: Carpentry: building and installing cabinets	.001	.003	.002	.564
	(.001)	(.002)	(.001)	
Area 10: Teaching: delivering lectures and class activities	.044	.046	.045	.901
	(.008)	(.008)	(.005)	
Area 11: Secretary: handling emails and correspondence	.023	.021	.022	.860
	(.005)	(.005)	(.004)	
Area 12: Agriculture: irrigating and fertilizing fields	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: operating lathes and milling machines	.009	.009	.009	1.000
	(.004)	(.004)	(.002)	
Area Retail: packing customers' goods	.499	.501	.500	.959
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: replacing carburators and generators	.119	.111	.115	.627
	(.012)	(.012)	(.008)	
Area 2: Plumbing: mending pipes	.080	.083	.082	.850
	(.010)	(.010)	(.007)	
Area 3: Food: washing dishes	.058	.055	.056	.822
	(.009)	(.008)	(.006)	
Area 4: Tailoring: hemming clothes	.074	.072	.073	.921
	(.010)	(.009)	(.007)	
Area 5: Hairdressing: coloring and highlighting hair	.012	.017	.015	.391
	(.004)	(.005)	(.003)	
Area 6: Construction: installing electrical wiring	.059	.046	.052	.245
	(.009)	(.008)	(.006)	
Area 7: Electrical Work: troubleshooting electrical issues	.123	.126	.124	.876
	(.012)	(.012)	(.009)	
Area 8: Welding: melting and soldering metal plates	.015	.019	.017	.545
	(.004)	(.005)	(.003)	
Area 9: Carpentry: crafting custom furniture	.003	.003	.003	1.000
	(.002)	(.002)	(.001)	
Area 10: Teaching: assessing student performance	.055	.036	.046	.082*
	(.008)	(.007)	(.005)	
Area 11: Secretary: managing appointment schedules	.024	.027	.025	.743
	(.006)	(.006)	(.004)	
Area 12: Agriculture: harvesting and storing produce	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: fabricating precision parts	.012	.011	.011	.807
	(.004)	(.004)	(.003)	
Area Retail: keeping vegetables fresh	.498	.501	.499	.918
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: changing oil	.118	.122	.120	.811
	(.012)	(.012)	(.008)	
Area 2: Plumbing: helping install pipe sections	.079	.067	.073	.371
	(.010)	(.009)	(.007)	
Area 3: Food: taking orders	.056	.056	.056	1.000
	(.008)	(.008)	(.006)	
Area 4: Tailoring: sewing buttons and fasteners	.070	.074	.072	.764
	(.009)	(.010)	(.007)	
Area 5: Hairdressing: performing scalp treatments	.015	.012	.013	.653
	(.004)	(.004)	(.003)	
Area 6: Construction: framing and erecting structures	.062	.047	.054	.209
	(.009)	(.008)	(.006)	
Area 7: Electrical Work: upgrading electrical panels	.143	.137	.140	.709
	(.013)	(.013)	(.009)	
Area 8: Welding: aligning parts prior to assembly	.023	.024	.023	.864
	(.005)	(.006)	(.004)	
Area 9: Carpentry: installing wooden flooring	.004	.003	.003	.654
	(.002)	(.002)	(.001)	
Area 10: Teaching: providing individualized instruction	.050	.046	.048	.715
	(.008)	(.008)	(.006)	
Area 11: Secretary: maintaining filing systems	.032	.028	.030	.650
	(.006)	(.006)	(.004)	
Area 12: Agriculture: caring for livestock	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: assembling mechanical components	.015	.009	.012	.343
	(.004)	(.004)	(.003)	
Area Retail: receiving payments	.495	.498	.497	.918
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: inspecting tires	.116	.116	.116	1.000
	(.012)	(.012)	(.008)	
Area 2: Plumbing: clearing clogged drains	.080	.084	.082	.778
	(.010)	(.010)	(.007)	
Area 3: Food: serving customers	.055	.056	.056	.910
	(.008)	(.008)	(.006)	
Area 4: Tailoring: altering clothes	.078	.076	.077	.923
	(.010)	(.010)	(.007)	
Area 5: Hairdressing: styling and setting hair	.015	.015	.015	1.000
	(.004)	(.004)	(.003)	
Area 6: Construction: painting and finishing surfaces	.064	.066	.065	.916
	(.009)	(.009)	(.006)	
Area 7: Electrical Work: wiring new construction projects	.137	.143	.140	.709
	(.013)	(.013)	(.009)	
Area 8: Welding: performing weld inspections and quality control checks	.015	.023	.019	.253
	(.004)	(.005)	(.004)	
Area 9: Carpentry: framing structures	.004	.001	.003	.317
	(.002)	(.001)	(.001)	
Area 10: Teaching: organizing educational materials	.048	.060	.054	.304
	(.008)	(.009)	(.006)	
Area 11: Secretary: assisting with bookkeeping tasks	.024	.029	.027	.522
	(.006)	(.006)	(.004)	
Area 12: Agriculture: operating farm machinery	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: performing quality control checks	.009	.011	.010	.795
	(.004)	(.004)	(.003)	
Area Retail: selling phone accessories	.514	.506	.510	.756
	(.018)	(.018)	(.013)	
Area 1: Motor-mechanics: brake system repairs and replacements	.103	.116	.110	.408
	(.011)	(.012)	(.008)	
Area 2: Plumbing: repairing leaky faucets	.080	.083	.082	.850
	(.010)	(.010)	(.007)	
Area 3: Food: cleaning tables and dining areas	.055	.052	.054	.818
	(.008)	(.008)	(.006)	
Area 4: Tailoring: pressing and ironing garments	.078	.076	.077	.923
	(.010)	(.010)	(.007)	
Area 5: Hairdressing: providing hair consultations	.013	.015	.014	.826
	(.004)	(.004)	(.003)	
Area 6: Construction: roofing installation and repairs	.060	.064	.062	.748
	(.009)	(.009)	(.006)	
Area 7: Electrical Work: installing and maintaining lighting fixtures	.124	.130	.127	.756
	(.012)	(.012)	(.009)	
Area 8: Welding: performing welding repairs and modifications	.021	.020	.021	.856
	(.005)	(.005)	(.004)	
Area 9: Carpentry: repairing damaged woodwork	.004	.003	.003	.654
	(.002)	(.002)	(.001)	
Area 10: Teaching: collaborating with parents and colleagues	.041	.041	.041	1.000
	(.007)	(.007)	(.005)	
Area 11: Secretary: drafting and proofreading documents	.023	.027	.025	.618
	(.005)	(.006)	(.004)	
Area 12: Agriculture: monitoring and controlling pests	.000	.000	.000	N/A
	(.000)	(.000)	(.000)	
Area 13: Machining: maintaining and repairing machinery	.008	.011	.009	.592
	(.003)	(.004)	(.002)	
Area Retail: stock management	.501	.491	.496	.717
	(.018)	(.018)	(.013)	
N	747	747	1494	

Notes: The value displayed for t-tests are p-values. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

B 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 GNC referral penalty in driving labor market segregation. Overall, all exercises indicate that, under our model, the relevance of network composition and of the GNC referral penalty 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. Our estimated parameters indicate differences of more than one order of magnitude between gender-congruent and gender non-congruent candidates.[33] To assess robustness of our results relative to the importance of GNCp, we propose an alternative calibration where we use estimates from the 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 smaller gap between gender-congruent and gender non-congruent candidates than our SMM estimates indicate – specifically, a 33 pp penalty to gender non-congruent candidates across all sectors. Using this information, we recalibrate the model to set $\sigma_m^{(1)} = .0607$, $\sigma_m^{(2)} = .1094$, $\sigma_f^{(1)} = .1232$, $\sigma_f^{(2)} = .0539$.

Figure OA.1 presents the results of this counterfactual exercise, using reweighted parameters from the correspondence study with employers. Reweighing the first-hand signals based on the audit study with employers (Figure OA.1, Panel a) increases the importance of GNCp to 23.2–25.9% ("S+NC+GNCp" vs. "S+NC"), showing that the importance of GNCp in driving labor market segregation is larger when first-hand signals are either less relevant and/or less segregated.

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.1b shows the steady state estimates. Under a denser network (average degree 10), GNCp increases segregation by 9.5–10.5% ("S+NC+GNCp" vs. "S+NC").

Labor market gender composition In addition to these recalibrations, we also assess whether the gender composition of the economy affects the importance of GNCp in driving segregation. We simulate a different stochastic block model with $p_m = .0615, p_f = .0667, q =$

[33] Estimated parameters are: $\sigma_m^{(1)} = .0075$, $\sigma_m^{(2)} = .1558$, $\sigma_f^{(1)} = .1763$, $\sigma_f^{(2)} = .0074$. Table OA.4 reports the SMM parameter estimates and moment fit for the baseline calibration.

.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.1c shows the steady states for this scenario, under the different scenarios. In our baseline comparison ("S+NC+GNCp" vs. "S+NC"), this change in labor market composition increases the importance of GNCp to 9.2–10.4%.

Overall, the results indicate that the importance of GNCp 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.

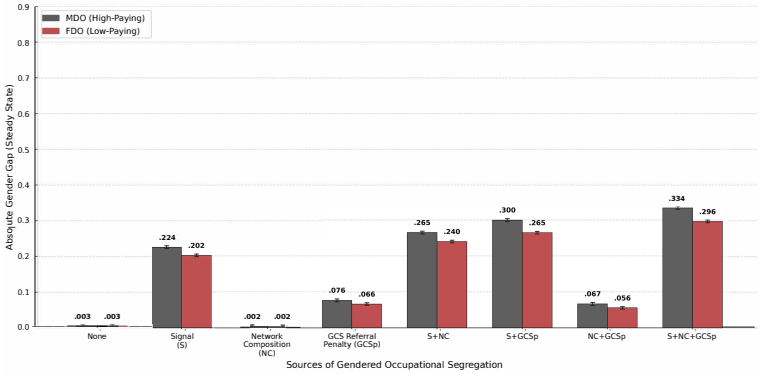
Table OA.4: SMM Parameter Estimates and Moment Fit

	Simulated	Observed
Signal parameters		
$\sigma_m^{(1)}$	.0075	—
	[.000, .028]	
$\sigma_m^{(2)}$	.1558	—
	[.103, .203]	
$\sigma_f^{(1)}$	.1763	—
	[.075, .263]	
$\sigma_f^{(2)}$	.0074	—
	[.000, .028]	
Employment shares (moments)		
Share f in type 1	.791	.797
Share f in type 2	.116	.108
Share m in type 1	.149	.147
Share m in type 2	.752	.757

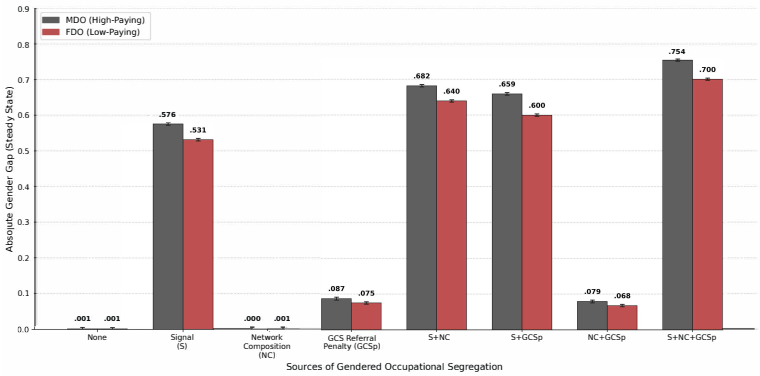
Notes: Signal parameters estimated by SMM. Brackets report 95% Bayesian bootstrap confidence intervals ($B = 1000$). Observed shares are sample proportions ($n_f = 222$, $n_m = 374$). Simulated shares are steady-state averages under S+C+GNCp.

Figure OA.1: Steady State Segregation — Robustness Counterfactuals

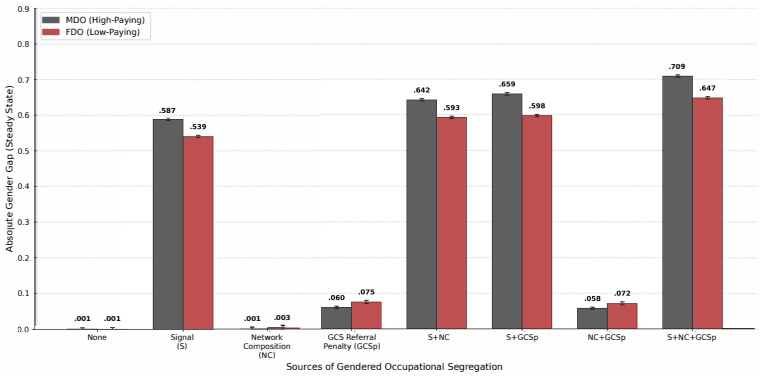
(a) Estimates from Audit Study



(b) Denser Network



(c) Different Gender Composition



Notes: Gender gap is the difference between the share of men and the share of women in an occupation. Each panel plots the average gender gap across 2,000 Monte Carlo simulations with $T = 1{,}500$; steady states are approximated by averaging the last 50 periods. Each column set reports the steady state under a different combination of segregation sources: "S" = differential job arrival rates (signals); "NC" = network composition (homophily); "GNCp" = GNC referral penalty. When "GNCp" is shut off, $\omega_C^{(s)} = \omega_{NC}^{(s)} = 1$; when "S" is shut off, $\sigma_m^{(s)} = \sigma_f^{(s)}$; when "NC" is shut off, cross-gender and same-gender link probabilities are set equal. Referral weights are constant across all panels ($\delta = .03$, $\omega_C^{(1)} = .586$, $\omega_{NC}^{(1)} = .414$, $\omega_C^{(2)} = .6335$, $\omega_{NC}^{(2)} = .3665$). **Panel (a)** (audit-study calibration): $N = 100$, $p = .064$, $q = .035$; signals reweighted from an audit study with employers to $\sigma_m^{(1)} = .0607$, $\sigma_m^{(2)} = .1094$, $\sigma_f^{(1)} = .1232$, $\sigma_f^{(2)} = .0539$. **Panels (b) and (c)** use baseline SMM signal estimates $\sigma_m^{(1)} = .0075$, $\sigma_m^{(2)} = .1558$, $\sigma_f^{(1)} = .1763$, $\sigma_f^{(2)} = .0074$. **Panel (b)** doubles average network degree to 10, so $N = 100$, $p = .128$, $q = .070$ (S off: $\sigma^{(1)} = .0919$, $\sigma^{(2)} = .0816$). **Panel (c)** calibrates the network to match observed gender composition: $N_m = 62$, $N_f = 38$, $p_m = .0615$, $p_f = .0667$, $q = .0368$ (S off: $\sigma^{(1)} = .0716$, $\sigma^{(2)} = .0994$; NC off: $p_m = .0539$, $p_f = .0464$, $q = .0489$). In Panels (a) and (b), NC off gives $p = .0509$, $q = .0498$.

Table OA.5: Model Extensions

		Robustness	Recalibration	
	Baseline	Audit Study	Denser Network	Gender Composition
	(1)	(2)	(3)	(4)
Benchmark segregation (FDO)	.648	.240	.640	.593
% Δ segregation from adding GNCp (FDO)	7.8%	23.2%	9.5%	9.2%
Benchmark segregation (MDO)	.685	.265	.682	.642
% Δ segregation from adding GNCp (MDO)	9.2%	25.9%	10.5%	10.4%
Parameter				
$\sigma_m^{(1)}$	.008	.061	.008	.008
$\sigma_m^{(2)}$	.156	.109	.156	.156
$\sigma_f^{(1)}$	.176	.123	.176	.176
$\sigma_f^{(2)}$	.007	.054	.007	.007
p_m	.066	.066	.131	.062
p_f	.066	.066	.131	.067
q	.035	.035	.071	.037
N_m / N_f	50/50	50/50	50/50	62/38
Avg. degree	5	5	10	5

Notes: Parameter values used in each model specification and the resulting contribution of the GNC referral penalty (GNCp) to steady-state occupational segregation. Column (1) is the baseline; column (2) ("Robustness") recalibrates signal parameters using audit-study estimates; columns (3)–(4) ("Recalibration") vary network parameters (denser network; gender-specific network parameters). "Benchmark segregation" is the steady-state segregation level under S+NC (signals and network, without the GNC referral penalty). "% Δ segregation from adding GNCp" is the percentage increase in steady-state segregation from adding the GNC referral penalty.

C 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:

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

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:

$$w^{(1)} = \begin{bmatrix} \omega_C^{(1)} \cdot \mathbf{1}_{Nr} \\ \omega_{NC}^{(1)} \cdot \mathbf{1}_{N(1-r)} \end{bmatrix} w^{(2)} = \begin{bmatrix} \omega_{NC}^{(2)} \cdot \mathbf{1}_{Nr} \\ \omega_C^{(2)} \cdot \mathbf{1}_{N(1-r)} \end{bmatrix}$$

where $\omega_C^{(k)}$ is the weight assigned to forwarding a signal to a gender-congruent candidate and $\omega_{NC}^{(k)}$ is the weight assigned to forwarding a signal to a gender non-congruent 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 $W^{(k)} = \mathbf{1}_N \cdot w^{(k)T}$ and all rows of $W^{(k)}$ are equal to $w^{(k)}$. Therefore:

$$W^{(1)} = \begin{bmatrix} \omega_C \cdot J_{Nr} & \omega_{NC} \cdot J_{Nr \times N(1-r)} \\ \omega_C \cdot J_{N(1-r) \times Nr} & \omega_{NC} \cdot J_{N(1-r)} \end{bmatrix}$$

$$W^{(2)} = \begin{bmatrix} \omega_{NC} \cdot J_{Nr} & \omega_C \cdot J_{Nr \times N(1-r)} \\ \omega_{NC} \cdot J_{N(1-r) \times Nr} & \omega_C \cdot J_{N(1-r)} \end{bmatrix}$$

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 gender-congruent and gender non-congruent 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:

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

Now denote the **forwarder vectors** $\hat{s}_t^{(1)}, \hat{s}_t^{(2)} \in \{0,1\}^N$ such that $\hat{s}^{(k)} = s^{(k)} \circ (\mathbf{1}_N - e_t^{(0)})$ – 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:

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

where $\mathbf{1}(\cdot)$ 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 $k, l \in \{0,1,2\}$. 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.

D Additional Analyses and Heterogeneities

Table OA.6: Classifying Occupations as Male- or Female-Dominated: UNHS Benchmarks

	UNHS Sample			Study Sample		
	% All	% Female	% Male	% All	% Female	% Male
Food and hospitality	4.82	65.88	34.12	7.42	62.50	37.50
Tailoring	.64	78.83	21.17	8.91	95.83	4.17
Electrical work	.11	15.56	84.44	13.73	2.70	97.30
Motor mechanics	.81	9.51	90.49	9.09	.00	100.00
Construction	3.17	.54	99.46	10.20	7.27	92.73
Plumbing	.07	.00	100.00	7.79	16.67	83.33
Retail	15.57	58.98	41.02	12.24	51.52	48.48
Secretary/accounting	.77	51.98	48.02	5.94	43.75	56.25
Teaching (primary and pre-primary education)	2.32	58.20	41.80	10.20	87.27	12.73
Hairdressing	1.21	65.82	34.18	2.60	78.57	21.43
Agriculture	54.06	57.91	42.09	1.30	.00	100.00
Machining and Fitting	.48	4.35	95.65	2.04	.00	100.00
Other unskilled	9.73	30.41	69.59	3.53	47.37	52.63
Other skilled	6.23	33.63	66.37	5.01	.00	100.00
Total	1.00	.52	.48	1.00	.37	.63

Notes: Each row reports an occupation's share of workers overall and by gender in the 2016/17 Uganda National Household Survey (UNHS) and in our study sample. We classify an occupation as female-dominated (FDO) if women account for at least half of workers in that occupation in the UNHS, and as male-dominated (MDO) otherwise. In our sample, most respondents on this track trained in secretarial work, so we classify the pooled sector as FDO.

Table OA.7: Round 1 and Round 2 Demographics and Labor Market Outcomes

	Round 1 Sample			Round 2 (Main) Sample		
Variable	N	Mean	SE	N	Mean	SE
Age	546	27.075	0.140	747	23.936	0.072
Sex (1=Female)	547	0.389	0.021	747	0.372	0.018
Rural household	474	0.519	0.023	747	0.552	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
Monthly Earnings (USD), Employed	214	114.369	4.729	389	108.538	3.419

Notes: Comparison of baseline characteristics between the Round 1 sample and the Round 2 sample (our main sample). For each variable, the table reports the sample size, mean, and standard errors. Variables include age, sex, whether the respondent is from a rural household, current employment status, whether the respondent works in a female-dominated occupation, and monthly earnings among those employed.

Table OA.8: 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)		

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.9: Mechanism Decomposition — OLS with Respondent Fixed Effects

	Referred (Part I)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GNC Experimental Candidate	-.234***	-.132***	-.168***	-.124***	-.091**	-.039	-.089**	-.023
	(.050)	(.042)	(.045)	(.041)	(.043)	(.045)	(.044)	(.045)
Likability		.227***		.173***	.186***	.175***	.197***	.143***
		(.020)		(.025)	(.023)	(.025)	(.025)	(.026)
Perceived Retention Probability			2.117***	1.030***				
			(.164)	(.226)				
Anticipated Employer Valuation					.089***			.056**
					(.027)			(.028)
Perceived Tasks Ability						.103***		.076***
						(.020)		(.021)
Anticipated Candidate Enjoyment							.098***	.061**
							(.029)	(.029)
R-Squared	.055	.363	.279	.399	.384	.401	.382	.422
Observations	1,494	1,494	1,493	1,493	1,364	1,372	1,369	1,361

Notes: OLS regressions of the Part I referral indicator on GNC Candidate and progressively added rating channels, with respondent (`trainee_id`) fixed effects absorbed via `areg` and standard errors clustered at the respondent level. Column 1 is the baseline (GNC alone, respondent FE). Column 2 adds likability; column 3 adds retention probability; column 4 adds both. Columns 5–7 add each of the three other rating channels individually (employer valuation, task ability, candidate enjoyment). Column 8 adds all four rating channels simultaneously. Respondent fixed effects are the linear analogue of a conditional-logit-at-the-choice-set specification, matching the design's within-pair identification. Sample: Round 2 Part I candidate-profile rows. This table reports the full regression output visualized in Figure A.4. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

D.1 Are Firms Missing Out on Productive Workers?

A natural concern about our results is whether the referral penalty imposes real productivity costs on firms: are incumbents systematically passing over more productive candidates? Our experiment is not designed to answer this question directly, since candidate productivity is neither manipulated nor measured on the job. However, we can draw on an earlier cohort of vocational-graduate workers in the same Ugandan sectors studied here, tracked over roughly a decade in Alfonsi et al. (2020).

The ECMA sample comes from the randomized evaluation of a vocational-training intervention implemented in partnership with BRAC in 2013. The evaluation used an oversubscription design, so individuals were randomly assigned to receive six months of vocational training (Short-VT) or assigned as controls (no-VT). The program was advertised using standard recruitment channels, with no requirement to participate in other BRAC programs to be eligible. Trainees were followed with detailed longitudinal surveys over more than a decade, and for each worker we observe earnings over time, the specific tasks performed at each job, and self-reported career satisfaction. All trainees also took a purpose-built Skill Assessment Task (SAT) shortly after graduating, measuring both theoretical and practical occupation-specific skills at entry level.[34]

Appendix Table A.6 compares GC and GNC graduates from the ECMA sample. GC and GNC graduates are statistically indistinguishable on wage-employment rates, SAT scores, career satisfaction, and, importantly, earnings conditional on employment in the training sector, in both female- and male-dominated occupations. The last row uses task-level survey data: every worker in the ECMA sample was asked to enumerate the tasks they routinely perform at work, and we classified each reported task as involving strictly physical work or not. The share of workers who perform strictly physical tasks is very small overall (under 4% in both occupation types) and does not differ by gender: GC and GNC graduates report performing strictly physical tasks at essentially identical, and uniformly low, rates. Together, these data cannot rule out productivity differences on unobserved dimensions, but they suggest that the referral penalty we document is unlikely to reflect firms optimizing away from less-productive candidates, and that concerns about physical fit are unlikely to rationalize the penalty in this setting.

E PAP WIP

E.1 Heterogeneity by Past Referral History

We ask whether the GNC referral penalty varies with the respondent's recent referral behavior. About 35% of respondents report making more than zero referrals in the twelve months preceding the survey, with most of those who do reporting between one and ten referrals. We split the sample by whether the respondent has ever referred a GNC contact in this period and re-estimate equation (4). Table OA.10 reports the results. Among workers who have ever referred a GNC contact, the GNC referral penalty in male-dominated

[34]The SAT is a standardized skill test co-designed with vocational-training instructors and industry practitioners. Macchi and Raisaro (2026) independently develop their own skill assessment for the same Ugandan occupations studied here, with which our SAT overlaps substantively.

occupations is statistically indistinguishable from zero (point estimate near zero), compared with a penalty of roughly −0.29 for workers who have never referred a GNC contact. This pattern is consistent with the broader interpretation that the penalty reflects something about referrers' priors or comfort with gender non-congruent choices, rather than a uniform feature of the choice environment.

Table OA.10: Heterogeneity by Referral History: Referral, Likability, Retention

	Referred			Likability			Retention Prob.		
	All	FDO	MDO	All	FDO	MDO	All	FDO	MDO
Panel A — Never Referred GNC									
GNC Candidate	-.244***	-.148**	-.293***	-.458***	-.178	-.602***	-.030***	-.017	-.037***
	(.037)	(.065)	(.045)	(.095)	(.163)	(.116)	(.009)	(.016)	(.011)
FDO−MDO		.145* [.068]			.424** [.034]			.020 [.282]	
Observations	1354	460	894	1354	460	894	1353	460	893
Panel B — Ever Referred GNC									
GNC Candidate	-.143	-.385**	.000	-.371	-.385	-.364	-.044*	-.027	-.055*
	(.120)	(.186)	(.153)	(.315)	(.471)	(.438)	(.025)	(.047)	(.030)
FDO−MDO		-.385 [.113]			-.021 [.973]			.028 [.602]	
GC Mean	.571	.692	.500	6.957	6.846	7.023	.681	.692	.675
Observations	140	52	88	140	52	88	140	52	88
Difference: Panel A − Panel B									
GNC Candidate	-.101	.237	-.293*	-.086	.206	-.238	.014	.010	.018
	(.124)	(.193)	(.158)	(.307)	(.435)	(.416)	(.024)	(.043)	(.029)
FDO−MDO		.530** [.033]			.445 [.460]			-.007 [.888]	

Notes: Estimates of equation (4) interacted with an indicator for whether the respondent ever referred a GNC contact in the twelve months preceding the survey. Respondents who named no network contact are coded as referral = 0. Standard errors in parentheses; FDO−MDO difference p-values in brackets. $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

F Robustness

F.1 Excluding Unemployed Respondents

Table OA.11: The GNC Referral Penalty, Excluding Unemployed Respondents

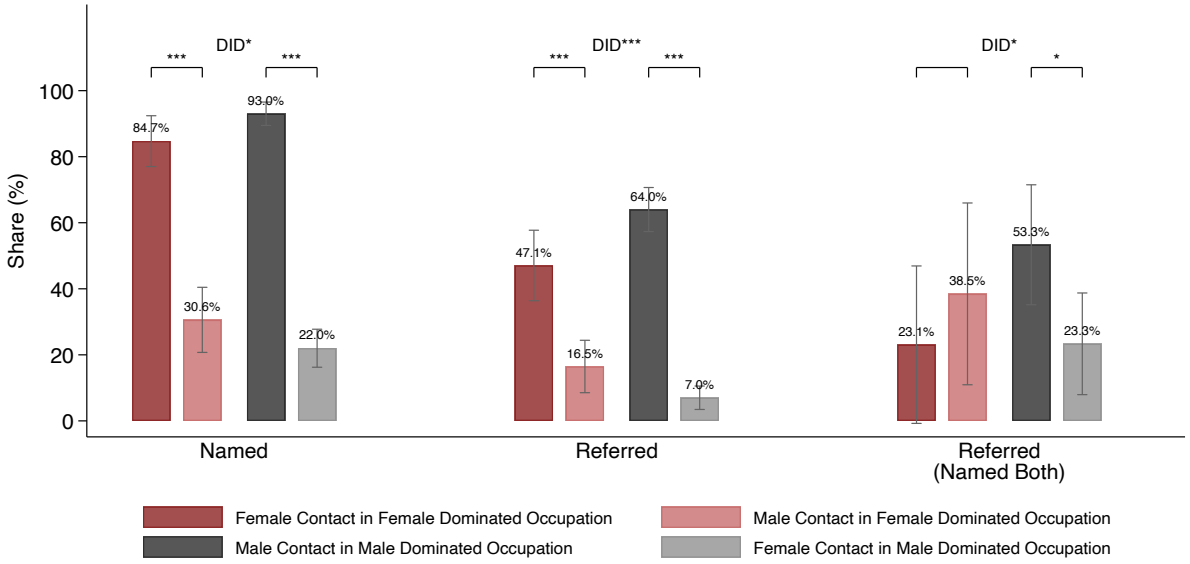
Panel A - Including Unemployed									
	Referred			**Likability**			**Perceived Retention Probability**		
	All Occupations (1)	FDO (2)	MDO (3)	All Occupations (4)	FDO (5)	MDO (6)	All Occupations (7)	FDO (8)	MDO (9)
GNC Candidate	-.234***	-.172***	-.267***	-.450***	-.199	-.580***	-.031***	-.018	-.038***
	(.036)	(.062)	(.044)	(.090)	(.151)	(.112)	(.008)	(.015)	(.010)
Differences Across Occupation Types									
GNC Candidate (FDO − MDO)		.095			.381**			.021	
		[.209]			[.043]			[.239]	
GC Mean	.617	.586	.633	7.115	6.980	7.185	.671	.670	.672
Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1494	512	982	1494	512	982	1493	512	981
Panel B - Excluding Unemployed									
GNC Candidate	-.228***	-.120*	-.287***	-.496***	-.251	-.632***	-.038***	-.012	-.053***
	(.042)	(.072)	(.052)	(.103)	(.179)	(.126)	(.010)	(.017)	(.012)
Differences Across Occupation Types									
GNC Candidate (FDO − MDO)		.167*			.381*			.041*	
		[.061]			[.083]			[.051]	
GC Mean	.614	.560	.643	7.101	6.963	7.177	.676	.669	.679
Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1072	382	690	1072	382	690	1072	382	690
Difference: Panel A - Panel B									
GNC Candidate	-.007	-.051	.020	.046	.052	.051	.007	-.006	.015**
	(.022)	(.035)	(.029)	(.059)	(.084)	(.077)	(.005)	(.008)	(.007)
Differences Across Occupation Types									
GNC Candidate (FDO − MDO)		-.072			.001			-.020*	
		[.115]			[.995]			[.051]	

Notes: Estimates on the referral outcome, comparing the full sample and the sample excluding respondents who were unemployed at the time of the survey. Standard errors clustered at the respondent level. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

F.2 Network Composition in the Round 2 (Main) Sample

Figure 2 in the main text combines the Round 1 and Round 2 samples in order to maximize precision on the network-naming and network-referral patterns. Figure OA.2 restricts the figure to the Round 2 (Main) sample only. The patterns are consistent.

Figure OA.2: Network Composition: Round 2 (Main) Sample



Notes: Replication of Figure 2 restricted to the 2023 sample only. Sample sizes and within-cell tests reflect the smaller sample ($N = 285$ at the naming margin; $N = 43$ for the referred-conditional-on-both-genders margin). Whiskers denote 95% confidence intervals; brackets and stars as in Figure 2. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.