

Do Job Networks Disadvantage Women?

Evidence from a Recruitment Experiment in Malawi *

Lori Beaman[†], Niall Keleher[‡], and Jeremy Magruder[§]

August 2012

Abstract

Using a field experiment in Malawi where applicants took part in a competitive selection process to become qualified for future surveyor positions with a local firm, we find that highly skilled women are systematically disadvantaged through the use of referrals. This happens both because men choose to recommend other men, and because women on average refer fewer candidates who qualify for the position. Our experimental design allows us to shed light on the network characteristics which lead to this disadvantage. Both men and women are connected to men and women at equal rates, so that the biases in referral behavior are not caused by segregated networks. By developing a model of how individuals choose referrals from among their social network and exploiting random variation in the terms of the referral contract, we document that both men's and women's biases result from social incentives rather than expectations of performance. Moreover, we also document that screening is maximized when men refer men. Therefore, employers who use networks to grant non-pecuniary benefits to employees or to screen new hires would have an incentive to encourage this disadvantage. This paper suggests that the use of social networks in hiring creates an additional channel through which women are disadvantaged in the labor market.

1 Introduction

While the gender gap in labor force participation has declined sharply in the last 30 years, women continue to earn less than men in countries around the world (World Bank Group and

*We thank IPA-Malawi field staff and Sam Arenberg for research assistance. We also thank participants at Brown University's Mini-conference on Networks, Northwestern University, SITE, the North America Summer Meetings of the Econometric Society, and the University of New South Wales for helpful comments. All errors are our own.

[†]Northwestern University. Email: l-beaman@northwestern.edu

[‡]Innovations for Poverty Action. Email: nkeleher@poverty-action.org

[§]University of California, Berkeley. Email: jmagruder@berkeley.edu

others, 2011). In Malawi, women are significantly under-represented in the formal sector (World Bank Group and others, 2010) as is common in many developing countries (Bell and Reich, 1988). A large portion of the literature in economics has focused on labor market discrimination (taste-based or statistical) or differences in human capital accumulation as reasons for the gender gap in earnings (Altonji and Blank, 1999).¹ Another possibility is that hiring processes themselves disadvantage women. We conduct a natural field experiment recruiting employees for a job in which men and women regularly compete in order to ask whether the use of referrals inherently disadvantage women in the labor market.

A large fraction of jobs - up to 50% - are attained through informal channels, including employee referrals (Bewley, 1999; Ioannides and Loury, 2004). While there is relatively little empirical evidence on the distributional consequences of referral systems, the potential of referral systems to create inequality between groups has been described theoretically (Calvo-Armengol and Jackson, 2004).² Mortensen and Vishwanath (1994) also show theoretically that network-based job information dissemination can disadvantage women, even if men and women are equally productive but men have a higher contact probability. Moreover, occupational segregation is commonly cited as a source of income disparity across gender (Blau and Kahn, 2000; Arbache, Kolev, and Filipiak, 2010). The use of employee referrals may be one of the mechanisms creating this segregation (Fernandez and Sosa, 2005; Tassier and Menczer, 2008). Observational data seems to support the hypothesis that women benefit less from job networks than men. Ioannides and Loury (2004) document stylized facts that women are less likely to

¹Additional explanations include the role of technology (Goldin and Katz, 2002), deregulation and globalization (Black and Strahan, 2001; Black and Brainerd, 2004), and differences in psychological attributes and preferences such as risk preferences, attitudes towards competition, other-regarding preferences, and negotiation (Niederle and Vesterlund, 2007; Bertrand, 2011).

²For the Calvo-Armengol and Jackson (2004) mechanism to be a relevant source of long-run inequality between men and women, job networks would need to be characterized by gender homophily. A large literature in Sociology (reviewed in McPherson, Smith-Lovin, and Cook, 2001) suggests that gender homophily in networks begins at early ages and is particularly strong in workforce networks.

be hired through a referral and that unemployed women are less likely than unemployed men to search using family and friends. Of course, these stylized facts alone do not show that women are disadvantaged by the use of networks in the labor market: women may work in occupations where networks are less relevant, or they may be less likely to report network help for the same hiring procedure. Therefore, it remains an open question whether women are made worse off by firms using employee referrals.

We used a competitive recruitment drive conducted by a research organization in Malawi, Innovations for Poverty Action (IPA-Malawi), as an opportunity to test how job referrals affect the recruitment of men and women in an experimental setting. IPA-Malawi was particularly interested in exploring whether referrals could reveal an otherwise untapped pool of qualified female applicants specifically, and qualified applicants in general³. Not only are women applicants relatively scarce, it may be particularly challenging to identify good women in the applicant pool. This could make referral hiring particularly attractive, if individuals are able and willing to screen for their employers (Montgomery, 1991; Beaman and Magruder, 2011). The position was advertised using the traditional method of posting flyers. Initial applicants attended a half-day interview process which included a written exam and a mock interview, where the candidate surveyed an actor playing the role of a typical respondent. At the conclusion of the interview, candidates were asked to refer a friend or relative to apply for the position and were offered a finder's fee. The referral process was cross-randomized along two dimensions: candidates were either told that they must refer a woman, that they must refer a man, or that they may refer someone of either gender, and their finder's fee was randomly

³Often times the gender of the enumerator is important: for example, a survey of women about their family planning practices should be implemented by female enumerators. Therefore, IPA wanted to recruit both men and women, and qualified women are particularly difficult to attract. The job often requires travelling to remote areas of the country. While men and women may be equally qualified, we may think that men are more open to the job. In the case of IPA-Malawi, qualified women candidates often state that they would prefer to take a job that is close to their home in order to attend to domestic tasks. Thus, jobs that require less travel (i.e. data entry) tend to see more women employees.

selected to be a fixed fee of either 1000 or 1500 Malawi Kwacha (MWK) or a performance incentive (a guaranteed 500 MWK with the potential to earn an additional 1300 MWK if the referral attained a certain threshold).

We find that qualified female candidates are strongly disadvantaged by the use of social networks in the hiring process. Among the conventional applicants (CAs) who were allowed to choose either gender for a referral, only 30% of referrals are women. This is significantly lower than the fraction of women who apply through traditional recruitment channels (40%). The low number of women referred is driven largely by male candidates. When given the choice, men systematically refer men; 77% of men's referrals are other men. Women refer women at approximately the same rate at which they apply through the traditional recruitment method. However, women systematically refer people who are less likely to qualify: a female candidate is nearly 20 percentage points less likely to refer someone who qualifies. These two effects combine to create a scenario where very few people ultimately refer qualified women when given a choice over which gender to refer: only 14% of men and 17% of women refer qualified women, which compares to 42% of men and 21% of women who refer qualified men. This is consistent with the finding from observational data from a call-center in Fernandez and Sosa (2005) and is the first experimental evidence that we know of that supports the large literature in sociology arguing that informal referral processes are one of the drivers of segregation of jobs (Doeringer and Piore, 1971; Mouw, 2006; Rubineau and Fernandez, 2010). This disadvantage for women, however, does not appear due to men (or women) being unconnected to women that they deem suitable: men and women make referrals at identical rates when required to refer women as when they are required to refer men under all contracts.

Since men and women are capable of identifying suitable female referrals, this disadvantage for women appears due to some aspect of the referral choice problem. We propose a

simple model of who individuals choose to refer under different types of referral contracts. In the model, individuals receive a social payment from referring a particular network member, in addition to any payment provided by the firm. They also receive a noisy signal of each network members' quality. A network members' social payment is decreasing in expected ability.⁴ We allow there to be several key gender differences in this decision process, which could generate the main empirical finding that qualified women receive different opportunities from their networks: (i) for any type of CA, networks of men and women may differ in the magnitude of top social payments ("closest gender"); (ii) the observable quality of friends who provide top social payments ("quality"); (iii) the extent of the tradeoffs between social payments and observable quality ("network shallowness"); and (iv) and the accuracy of the signal of observed quality relative to actual quality ("information").

Because of random variation in incentives, we are able to observe the characteristics of both potential referrals when CAs are motivated in their choice only by network incentives and CAs when firms provide an additional incentive to find a person who is high ability. This facet allows several tests of the various sources of heterogeneity. First, the model suggests that referrals selected under fixed fee payments should reflect the people closest to CAs, so that both the quality of closest people and any gender difference in top social payments will be reflected in referral choices under fixed fees. From this framework, we identify that men systematically receive the highest social payments from other men. We also identify that both the men and women who give highest social payments in mens' networks are of similar ability, which is slightly below the average CA's ability. Women are not systematically closer to one gender over the other, but women CAs prefer to refer women who are less likely to qualify than their closest men (and the average CA).

⁴We justify this assumption through a selection rule: if we consider only the subset of the overall network which is not dominated (so that we only consider potential referrals who are not dominated in both social payments and observed ability) then this result is immediate.

Second, we derive that if tradeoffs between social payments and observed quality are higher in one gender, or if information is relatively worse for that gender, we should observe a smaller performance premium. In other words, incentivizing CAs to find the best possible referral will be less effective in the presence of either bad information or high tradeoffs, which may have implications for firm motivation to address any gender disparities which exist for reasons of social preference. Taking this prediction to the data, we find that men exhibit a large performance premium when referring men, but no performance premium when referring women: this factor allows us to conclude that men’s networks of women have either worse information or greater social tradeoffs than men’s networks of men. Further examining the distribution of test scores when men refer women under performance pay suggests that worse information about women is likely to be at least part of the explanation. By contrast, women in performance pay treatments are not more likely to make referrals who qualify when referring men or women, but performance incentives do change women’s referrals of both men and women in some productive dimensions.

Taken together, the use of social networks in the labor market lead to further reductions in the number of women who gained access to this particular formal sector job. This paper suggests that an additional factor leading to the observed gender gap in women’s earnings is the way employees are recruited, even in the absence of any intended (taste-based or statistical) discrimination on the part of the firm. The results also highlight that firms must be willing to generate incentives to refer only the best people within their employees’ networks, and that there may be important gender differences in underlying social networks.

The paper is organized as follows. The experiment design and data are described in section 2. The main results are discussed in section 3. The theoretical framework which highlights potential mechanisms for why women are disadvantaged by referral hires is elaborated

in section 5. The empirical evidence for each of the three hypotheses are presented in section 6, followed by the conclusion.

2 Experimental Design

2.1 Setting and Overview

Gender-based difference in employment is common in many developing countries (Bell and Reich, 1988). Women in Africa are more likely to be in the informal sector and the proportion of women in the formal employment is less than half that of men. A survey of 14 African countries found that women are more likely to be employed (formally or informally) in agricultural jobs and less likely to hold jobs in the manufacturing and services (Arbache, Kolev, and Filipiak, 2010).

In Malawi, women are significantly under-represented in the formal sector. A recent survey of Malawian households suggests that less than one-third of women participate in the formal labor force, while nearly 58% of men do so (World Bank Group and others, 2010). Among urban women, 38.2% had not been employed in the preceding twelve months; this rate is more than double that found among urban men (18.6%) (National Statistics Office (NSO) and ICF Macro, 2011).

IPA-Malawi hires enumerators to conduct interviews of farmers, business owners, and households in rural and urban Malawi. In the 12 months following the recruitment drive (our experiment), IPA-Malawi projected hiring a minimum of 200 enumerators for its survey activities. IPA-Malawi had an explicit motivation to hire more female enumerators than their usual recruitment methods allow. Typically, only 15 to 20 percent of enumerators hired by IPA-Malawi are women, and some survey tasks require same-gendered enumerators (for ex-

ample, same-gendered enumerators are sometimes important for sensitive questionnaires). For this experiment, we introduced incentives for job applicants to make referrals during IPA's recruitment sessions in the two main Malawian cities, Blantyre and Lilongwe.

The standard method for recruiting enumerators is to post announcements at community centers, technical schools, and government offices. An initial screening session is open to all applicants with minimum qualifications. Minimum requirements to be hired for an enumerator position are: a) a secondary certificate; b) fluency in the local language (Chichewa), and c) English reading and oral comprehension. Candidates with data collection experience, good math skills, and basic computer skills are given preferential review.

The standard IPA-Malawi screening session consists of submitting a CV to IPA and sitting for a written test. The written test assesses reading comprehension, hand writing, math ability, and computer literacy (via self-assessment). Following the screening session, applicants deemed to be qualified may be invited for a survey-specific training of enumerators.⁵ At the end of the training, job offers are made to a group of individuals deemed to be adequate for work on the survey.

In this experiment, IPA posted fliers indicating a hiring drive at a number of visible places in urban areas. The posters included information on the minimum requirements for IPA enumerators, the dates and times of the recruitment sessions, and a solicitation to bring a CV and certificate of secondary school completion (MSCE). Participants then attended an interview session, where they submitted their CV and were registered with a unique applicant number. Participants were limited to those individuals who had never worked for IPA. Each day, two sessions were conducted by IPA staff. At the start of each session, participants were

⁵These trainings consist of a multiple-day workshop on proper technique and procedures for conducting paper-assisted or computer-assisted personal interviews. Each training is tailored to a specific survey; however, interview techniques for facilitating and documenting interviews is rather standardized. Also, during a training workshop, practical skills are assessed through a field pilot of the given survey.

introduced to IPA and the role of an enumerator was described.

2.2 Quality Assessment

The screening session included a written test similar to the one IPA had previously used, and a practical test which served as a condensed version of the training that IPA had previously used to select enumerators. Participants were given one of two distinct written tests. The two tests were distributed at random to limit cheating. Each test consisted of several math problems, ravens matrices, English skills assessment, job comprehension component, and a computer skills assessment. For richer survey data, our screening session integrated the practical test.

For the practical test, the participant played the role of the enumerator for a computer assisted personal interview.⁶ An experienced IPA enumerator played a scripted role of the interview respondent. The respondent scripts included implausible or inconsistent answers (i.e. age, household size, household acreage) to survey questions. These false answers were used as checks on the participant’s ability to pay attention to detail and verify inaccuracies in responses. When the participant pressed the respondent for a correction, the respondent gave a plausible answer. Among the respondents, two sets of implausible answers were used in order to limit any ability to predict the practical test.

Scores were calculated for all participants on a 0-to-100 scale. The total score was a combination of the CV score, written test score and practical test score.

⁶All participants were required to go through a short self-administered training with a computer-assisted personal interviewing (CAPI) software in order to ensure a consistent level of familiarity with the computer program. (The CAPI software used was Blaise. To our knowledge, no participant had used this particular software prior to the recruitment session.) Once finished with the self-administered CAPI training, participants moved to the practical test.

2.3 Referral Instructions

The setting offered an opportunity to test several potential channels through which a firm can influence the type and quality of applicants generated through a referral process. Prior to leaving the recruitment session, participants had a one-on-one conversation with the recruitment manager. During this conversation, a letter was provided to the applicant inviting the applicant to identify another individual to refer to IPA for consideration as an enumerator. The message provided to the participant was the crux of this experiment. All original participant letters described a specific set of instructions about the referral process. We randomly varied the content of the letters.

Each letter included an instruction about the gender requirement of the referral who could be invited to attend a future recruitment session. The letter instructed the original participants that their referral had to be male, had to be female, or could be either gender. The referral needed to be someone who had not worked for or been tested by IPA in the past. The letter also said that the referral should be highly qualified for the enumerator position and given a suggestive guide about what this would entail. Namely, the letters stated that a strong enumerator should have a secondary school certificate, fluency in Chichewa, excellent comprehension of English, data collection experience, and good math and computer skills. The CA was told that the referral would need to complete the same written and practical assessments as done by the CA.

Conventional applicants were randomly assigned into one of three pay categories (cross randomized with the gender treatments): a fixed fee of 1000 Malawi Kwacha, a fixed fee of 1500 MWK, or a performance incentive of 500 MWK if their referral does not qualify or 1800 MWK if their referral does qualify. All treatments were fully blind from the perspective of all evaluators. All conventional applicants were eligible to receive payment (fixed fee or base pay,

if in the incentive group) if their referral attended and completed a recruitment session.

Referrals typically participated in recruitment sessions 3 to 4 days after the conventional applicant’s session. The screening session, including the written and practical test components, were the same as for conventional applicants.

Conventional applicants were asked to complete an anonymous questionnaire as an assessment of their referral’s quality and whether or not they shared any of their payment with the referral. In addition, CAs were contacted by an IPA staff member to ask how the conventional applicant identified the referral and how the payment was used.

Each week, a list of qualified applicants was posted at the recruitment venue, and qualified applicants were told that they would be considered for future job opportunities with IPA-Malawi. Any original applicant who qualified for a payment was informed and given payment in a sealed envelope.

Appendix Table 1 displays summary statistics for the sample of CAs, for men and women separately. It also shows that the randomization lead to balance along most characteristics. The p value for the joint test of all the treatment variables, and their interactions, is displayed in column (2) for male CAs and column (5) for female CAs. No characteristic is significantly different at the 5% level, and only MSCE Math Score is significantly different at the 10% level for female CAs (though the Practical Component Z-score is almost significant at the 10% level for both men and women CAs).

3 Are Qualified Women Disadvantaged?

Figure 1 plots kernel densities of CA overall test score⁷ separately for men and women, and confirms that men and women who respond to the traditional recruitment method on average

⁷The test score measure used here was corrected for a few data entry mistakes (which was done in real time), after qualification had been determined.

have similar distributions of test scores. There is some evidence that women are slightly more represented in the left tail of the distribution, but the right tails are identical, and overall they are very similar. Table 1 shows estimates from a regression of qualification rates on CA gender. There is some evidence that men slightly outperform women across the CAs, but these differences are visibly small and statistically indistinguishable. While there are some characteristics which systematically vary across men and women, the likelihood of qualifying is similar.

Panels A and B of Table 2 document the primary result of this paper. While 39% of applicants themselves were women, and those applicants performed as strongly as the male applicants did, only 30% of referrals are women when we allow CAs to choose which gender to refer (difference significant at the 5% level). This difference in application rates happens entirely because men systematically do not refer women: women refer women at approximately the rate by which women apply themselves (43% of the time), while men refer women only 23% of the time when given the choice. The difference between male and female CAs is significant at the 1% level, as shown in column (4).

However, qualified women can also be disadvantaged if unqualified women are being referred more so than unqualified men regardless of any gender preference. As it turns out, while there is no gender difference in the qualification rate of CAs, there is a large gender difference in the qualification rate of referrals: while men make references who are more likely to qualify than CAs are on average, women make references who are almost twenty percentage points less likely to qualify (37% versus 56%) when given an unrestricted choice of genders. Rows 3 and 4 reveal that this difference is held up by each gender of CA, regardless of which gender they choose to refer: while male CAs very rarely refer women, the women who they do refer are fairly likely to qualify, while female CAs often refer women but rarely refer anyone who

IPA would actually hire. These two differences together put qualified women at a substantial disadvantage: most men seem to respond to an unrestricted referral situation by identifying men, while most women seem to respond to such a situation by finding unqualified people of either gender. Overall, we conclude that the use of referral systems strongly disadvantages women (and particularly qualified women) in this context.

4 Are Men Connected to Women (and Women Connected to Men)?

One explanation for why men refer so few women is that it may not be a choice: men may simply not be connected to women. Indeed, one proposed cause of gender segregation in the labor market is segregated social networks (Tassier and Menczer, 2008). Based on this explanation, referrals serve to perpetuate job segregation due to the limited overlap of groups from which referrals are drawn.

Our view is that a sensible definition of connectedness would reflect contract terms: clearly, any of our male CAs would be successful at finding a female referral at a sufficiently high price, particularly in fixed fee treatments where the CA need not be concerned with referral quality. For now, suppose simply that each CA i receives a number of draws of friends, who may be male or female. Each friend j is characterized by two characteristics: a social payment α_j (net of costs of recruiting that person) which (s)he will give to i if (s)he is offered the referral, and some probability of qualifying, λ_j as well as their gender. Thus, when CA i is offered a contract with fixed component F_i and performance component P_i , if i refers j , then i receives in expectation

$$F_i + \alpha_j + \lambda_j P_i \tag{1}$$

If we assume that CAs do not make referrals if they cannot receive positive payments in expectation, then it allows an easy definition of connectedness.

Definition 1 *CA i is **connected** to gender g if $\max_{j \in g} F_i + \alpha_j + \lambda_j P_i > 0$*

Given this definition of connectedness, it is straightforward that being unconnected under fixed fees would indicate that the largest possible social payment is less than $-F_i$, while being unconnected under performance pay would require a low expected α_j coupled with a low probability of qualifying among all referrals. Clearly, if male CAs are less connected to women at our contracting terms, it could generate the disadvantage that women face in referral systems.

We can analyze this in a straightforward way: define an indicator $R_i = 1$ if the CA makes a referral, and $R_i = 0$ if the CA does not. Since we randomly restricted some CAs to referring only women, and other CAs to referring only men, we can test whether CAs are more or less likely to be connected to women or men at our contracting terms. Moreover, because some CAs were allowed to refer either women or men, we will additionally be able to test whether CAs who are unconnected to men at a particular contract are likely to be connected to women at that contract: if so, it would suggest that CAs are receiving a number of draws of both men and women, so that even if all the draws of an CA's own gender fail to make the participation threshold, there is a strong chance that the other gender exceeds it. As a test, then, we simply regress

$$R_i = \sum_k \alpha_k T_{ik} + \delta_t + u_i$$

Where T_{ik} is the exogenously assigned treatment in terms of referral gender and contract

payment and δ_t are time trends. Table 3 presents this analysis, where restricted male treatments (or male fixed fee treatments in specifications which disaggregate by contract terms) are the excluded group. Overall, neither men nor women are significantly less likely to make a reference when assigned to refer women than men, and point estimates on any gender differences are small in magnitude. When we disaggregate by treatment, we observe that men are statistically significantly less likely to make a reference when they are given performance pay than when they are given fixed fees if they are required to refer either men or women. The mean referral rate under fixed fees for men in restricted treatments is 90%; point estimates suggest that if these men are instead given the performance contract return rates fall to 75%. However, if men are given the choice of referring either men or women, the return rate rises back to 90% - taken as a whole, this suggests that there are 15% of men who know either a man or a woman who is worth referring under performance pay, but not both. For female CAs, there is a similar trend, though the point estimate is smaller and it is not statistically significant.

We find these results striking in several ways. First, they reject the hypothesis that the trend of men referring men noted in section 3 occurs because of men being unconnected to women. Most male applicants are connected to suitable women, and they are as likely to be connected to women as they are to be connected to men under either contract structure. There are also a sizeable number of men who are only connected to women, when the performance of the referral matters. Second, they suggest that mean returns under performance pay are lower than under fixed fees. CA return rates under fixed fees remain at 90% for both genders of CAs assigned to both levels of fixed fees; this suggests that the expected return to performance pay is lower than 1000 MWK, the lower fixed fee level. Given that the performance pay contract featured a guaranteed fee of 500 MWK and a performance premium of 1300 MWK, this suggests that the person they choose to refer under fixed fees (who they could have also referred under

performance pay) has an expected qualification likelihood below $5/13$, which we return to below. Finally, there are no differences in referral rates for CAs of either gender when we restrict them to men versus when we restrict them to women. There are, however, differences in return rates between contract terms for gender-restricted referrals. Given that attrition rates are in any event low relative to most panel studies and uncorrelated with the gender treatments which are the focus of this study, we abstract from them in the main analysis, though when we analyze performance premia we will discuss the potential role of attrition in biasing our estimates.

Given that CAs are connected to both men and women, it seems likely that some other factor leads to the disadvantage women face from referral systems. In the next section, we add some structure to the referral problem in the interest of identifying key differences between an individual CA's networks of men and women.

5 Model and Mechanisms

Building on equation 1, retain the notation α_j as a social incentive supplied by friend j and now suppose that CA i expects j to score Q_j points on the skills assessment. If j belongs to gender g , suppose that his (her) actual performance is $Y_j = Q_j + \varepsilon_j$, where ε_j is distributed $N\left(0, (\sigma_\varepsilon^g)^2\right)$, and the referral qualifies if $Y_j > 60$. Note that σ_ε^g may be different between men and women. Using the language from the previous section, this suggests that $\lambda_j = \left(1 - \Phi\left(\frac{60 - Q_j}{\sigma_\varepsilon^g}\right)\right)$ where $\Phi(\cdot)$ is the cdf of the standard normal distribution. As before, CA i is given a contract (F_i, P_i) and is restricted to make a referral out of individuals who belong to set \mathcal{G} , where \mathcal{G} could be $\{male\}, \{female\}$ or $\{everyone\}$.

While i knows a number of people in each gender specific network, we focus on the subset of those draws who could be optimal referrals under various contracting conditions. In particular, individual j will only get chosen under some contract (F_i, P_i) if $Q_j \in \arg \max_{k \in \mathcal{G}} Q_k | \alpha_k \geq$

α_j , that is, j will only get chosen if his or her observed quality is the best among eligible referrals who offer at least as much in social payments. For each gender g , define $h^g(\alpha_j) = Q_j$ to be the mapping between α_j and Q_j in this set, where $h^g(\alpha_j)$ is decreasing in α_j by the selection rule. Denote $\alpha_1^g = \max_{j \in g} \alpha_j$, where $j \in g$ if j is of gender g . To make analysis tractable, approximate $h^g(\alpha_j) = Q_1^g + \gamma^g(\alpha_1 - \alpha_j)$. CA i therefore solves

$$\pi_i(\alpha_j, P_i, F_i) = \max_{j \in \mathcal{G}} P_i \left(1 - \Phi \left(\frac{60 - Q_1^g - \gamma^g(\alpha_1^g - \alpha_j)}{\sigma_\varepsilon^g} \right) \right) + F_i + \alpha_j \quad (2)$$

Gender-specific networks can be heterogeneous, therefore, in 4 different ways: they may differ in $\alpha_1^g, Q_1^g, \gamma^g$, and σ_ε^g . The following set of definitions characterize these differences

Definition 2 CA i is **closer** to gender g than to gender g' if $\alpha_1^g > \alpha_1^{g'}$

Definition 3 CA i 's network of gender g is **higher quality** than his network of gender g' if $Q_1^g > Q_1^{g'}$ ⁸

Definition 4 CA i faces a **shallower** network of gender g' than of gender g if $\gamma^g > \gamma^{g'}$

Definition 5 CA i has **better information** about gender g than about gender g' if $\sigma_\varepsilon^g < \sigma_\varepsilon^{g'}$

Our interest is to test whether gender-differences in these four characteristics can contribute to the observed differences in referral choices. To more carefully consider how referral choice behavior interacts with differences along these dimension, we consider separately optimal behavior under fixed fee contracts of the form $(F_i, 0)$ and performance pay contracts of the form (F'_i, P_i) where $P_i > 0$.

⁸This is a rather special definition of higher quality and does not indicate that all members of gender g are higher quality than gender g' or even that members of gender g are on average higher quality than members of gender g' .

5.1 Behavior under Fixed Fees

In interpreting our data in the context of the model, we start with a description of what happens when we provide contracts of the form $(F_i, 0)$. Examining the optimization condition leads to the following proposition.

Proposition 1 *Under fixed fee contracts of the form $(F_i, 0)$, CAs always refer the closest person of the eligible gender, friend 1. Differences in α_1^g can lead to different return rates between genders, but differences in Q_1^g , γ^g , and σ_ε^g will not result in different return rates between genders or differences in which gender is selected when the CA has the option of referring either gender.*

Proposition 1 reveals that if we examine referrals who were recruited under fixed fees when restricted to a particular gender, then we are directly examining friend 1 of the appropriate gender. As a result, the quality of referrals recruited under fixed fee treatments restricted to referring a particular gender can be taken as an estimate of Q_1^g , and the fraction of CAs who refer men when allowed a choice of genders can be interpreted as the fraction of CAs for whom $\alpha_1^1 > \alpha_1^2$. If CAs systematically refer men under fixed fee treatments, proposition 1 suggests that the only interpretation of this fact is that men are systematically closer to CAs. Our analysis therefore begins by examining fixed fee treatments to estimate both Q_1^1, Q_1^2 and the relationship between α_1^1 and α_1^2 for male and female CAs and determine whether gender differences in networks quality or in closeness could be related to the disadvantage women face.

5.2 Behavior under Performance Payments

While behavior under fixed fee payments allows clean tests of two dimensions of heterogeneity which may lead to overall advantages for one gender or the other, one facet of network heterogeneity which should be of particular interest to employers is not who men and women are

closest to but rather what the potential of networks are, given proper (financial) motivation. To describe how network heterogeneity affects incentivized referrals, we commence by defining a performance premium in this context.

Definition 6 *There is a performance premium if CA i selects referral p with social incentive α_p and observed quality q_p , $p > 1$ when $P_i > 0$ ⁹.*

This definition of a performance premium is essentially whether the CA selects an internal solution or a corner solution to the referral choice equation. It is without loss of generality to suppose that the performance premium is positive regardless of which network member p is chosen and easy to verify that a larger payment P_i both leads to a more likely interior solution and a higher valued interior solution. However, the linearity assumption on $h(\cdot)$ allows us to say something more specific about the quality of person p who would be selected and also puts a restriction on the quality of friend 1, depending on differences in referral rates.

Lemma 1 *If CAs optimally select referral p , $p > 1$, then person p has an expected performance of above 60 on the skills assessment $\left(60 < Q_1^g + \gamma_g \left(\alpha_1^g - \alpha_j^g\right)\right)$*

Given the result from section 4 that person 1 has a qualification probability below 5/13, we can conclude that person p has a substantially higher expected performance which both motivates the definition of performance premium and will be used in some results below.

Proposition 2 *CAs are more likely to select referrals which feature a performance premium in gender g if*

⁹This definition focuses on the extensive margin of the performance premium as, by definition, whenever the person chooses a referral other than the closest (and lowest ability) friend, the performance premium will rise. Marginal effects on the magnitude of the performance premium are theoretically ambiguous with respect to most key parameters due to the non-linearity of the normal distribution.

A): the gender g network is less shallow

B): CAs have more information about gender g and Lemma 1 holds ¹⁰

C): CAs have a higher (lower) quality network of gender g than gender g' and $|60 - Q_1^g| <$

$(>) |60 - Q_1^g - \gamma^g (a_1^g - \alpha_p^g)|$ (for person p who solves equation 2)

Corollary 1 *The likelihood of selecting a referral with a performance premium is unaffected by differences in α_1^g .*

In other words, as networks become more shallow, the return of someone other than the closest friend declines relative to the closest friend. If Lemma 1 holds, then the same is true for better information: in that case, the index partner is not expected to qualify and so better information increases the chance that the CA is correct about the index partners (expected) failure. In terms of network quality, the non-linearity of the normal distribution causes the distinction in effects: as quality increases, the returns to recruiting the index partner increases relative to other potential partners if the index partner is at a point of higher density in the normal distribution, or if (s)he is closer to the mean. In this case, the added quality of the index partner increases the probability that the index partner qualifies, making referring someone else less desirable.

Proposition 3 *If there is no Performance Premium for gender g , and Lemma 1 holds, then the overall return under performance pay is: increasing in α_1^g ; increasing in Q_1^g ; increasing in σ_ε^g ; and unrelated to γ^g .*

Proposition 4 *If there is a Performance Premium for gender g , and Lemma 1 holds, then the overall return under performance pay is increasing in α_1^g ; increasing in Q_1^g ; decreasing in σ_ε^g ;*

¹⁰Strictly, the condition is weaker than this. What needs to hold is $\left(\frac{60-Q_1^g}{\sigma_\varepsilon^2}\right)\phi\left(\frac{60-Q^g}{\sigma_\varepsilon}\right) > \left(\frac{60-Q^g-\gamma^g(\alpha_1-\alpha_j)}{\sigma_\varepsilon^2}\right)\phi\left(\frac{60-Q^g-\gamma^g(\alpha_1-\alpha_j)}{\sigma_\varepsilon}\right)$.

Table 4: Estimated directions of network differences

Parameter	$\pi_i(F_i, 0)$	$\pi_i(F'_i, P_i)$		$E[Y_i P_i > 0]$
		$E[Y_i P_i > 0] = Y_i$	$E[Y_i P_i > 0] > Y_i$	
α_1^g	\uparrow	\uparrow	\uparrow	0
Q_1^g	0	\uparrow	\uparrow	Ambiguous
γ^g	0	0	\downarrow	\downarrow
σ_ε^g	0	\uparrow	\downarrow	\downarrow

and decreasing in γ^g .

Table 4 summarizes Propositions 1, 2, 3, and 4 and documents that heterogeneity in any of the underlying parameters could put men at an advantage under performance pay. In particular, if one gender (say, men) is closer than the other, then men should be referred more under performance pay just as they would be in fixed fees, all else equal. Unlike in fixed fee treatments, if men (or women) are higher quality, then they also will be referred preferentially under performance pay. This result holds whether CAs intend to return with the index partner or intend to choose a partner with a performance premium, as either will be higher quality (with higher expected return) when the network's quality shifts upwards. Therefore, if we conclude (from our tests on fixed fee CAs) that gender-specific networks are different in closeness or quality, we should expect those results to translate to gender preference under performance pay.

Shallowness and bad information affect gender preference in ambiguous ways under performance pay. If one gender network is shallower than the other, then it affects gender preference under performance pay only if a referral other than the index partner was going to be selected. Bad information can go either way: if the optimal referral under performance pay was the index partner, then bad information should make that gender more likely to be selected in unrestricted treatments as the chance that the index partner actually qualifies is increased (conditional on the signal). On the other hand, if the CA intended to select someone who he

expects to qualify, worse information about that gender can only hurt his payoff to selecting that gender. Shallowness and bad information are similar, however, in how they affect the CA's choice of optimal partner: if one gender-specific network is more shallow, or if information is worse, then the CA will unambiguously be more likely to select the index partner rather than a partner who he expects to perform well on the recruitment exercise. This leads to a test: if there is more of a performance premium in one gender than the other, and if there is no difference in quality of the index partner, then we can conclude that the gender with a larger performance premium has a network which is either less shallow or better informed. Given that employers are no doubt interested in using referral systems to screen, this may suggest that employers would want to encourage referrals of the less shallow or better informed network.

While bad information and shallowness are similarly negative characteristics from an employer's perspective, it remains interesting to disentangle them. One way to test whether the data seem more consistent with bad information or shallowness is determining whether low quality referrals are less likely to be referred under performance pay than under fixed fees. With good information, the return to referring a very low quality referral is substantially lower under fixed fees, and if networks are shallow (but well-informed) we may expect these particular individuals to be likely to attrit under performance pay. In contrast, if CAs are simply poorly informed as to their referral's capabilities, then in expectation their payments from making that referral are reduced much less.

5.3 Empirical Tests

The model above outlines how differences in underlying heterogeneity may lead to different referral choices. This section summarizes how these differences may disadvantage one gender vis-a-vis the other when referrals are given the option to refer either gender. As our focus in this

paper is on understanding potential mechanisms which could disadvantage women, discussion in this section will presume that women are the disadvantaged gender under any hypothesized mechanism though of course algebraically either gender could be the disadvantaged g' . Under “either” treatments, CAs will exhibit a preference to refer men over women if the overall returns to referring men are higher. Table 4 highlights four potential reasons that returns may be higher under different contracts: men could be closer than women to CAs; men could be higher quality than women; men could be in less shallow networks than women; and CAs could have different information about men than about women (which in principle could have ambiguous sign as to the advantage).

Ultimately, we present 4 tests: first, using the choice of gender under fixed fee, unrestricted treatments, we infer which gender (male or female) is closer to CAs. If one gender is systematically closer, we should also observe that gender being chosen more under performance pay. Second, we estimate the quality of each gender-specific network for CAs by examining the quality of references under fixed fees. Third, we test for the presence of a performance premium. If there is no performance premium, it will be consistent with bad information or shallowness. To separate these, we compare distributions of gender-specific referral scores under fixed fees and performance pay to evaluate whether very low quality referrals are less likely to take place under performance pay, which would be consistent with well-informed shallowness, or not, which would be consistent with bad information being at least part of the story.

6 Results

6.1 Do Social Payments disadvantage women?

As we discuss above, referrals under fixed fee treatments provide consistent estimates of characteristics of the index partner of each gender, partner 1. This allows us to draw several inferences: first, if we repeat Panel B of table 2 using only CAs in fixed fee treatments, we can infer whether male and female CAs are closer to men or women. That analysis is presented in Panel C of table 2, which indicates that 75% of male CAs are closest to men, which contrasts to 57% of female CAs. From proposition 1, we can conclude directly that male CAs get their highest social payments from other men, which can lead to across the board disadvantages to women coming from male referrals. Thus, while we rejected the idea that networks were sufficiently homophilic so that men were unconnected to women (and vice versa), we can conclude that there is homophily in the social incentives which live over the network, at least for men: the closest people for men tend to be men.

6.2 Are there quality differences in men’s and women’s networks of men and women?

A second reason why women may be disadvantaged by social networks would be that men (or women) did not have high quality women in their network,¹¹ or, that men or women are very distant from high quality women. The model provides a clear formulation of network quality, being defined as the expected quality of the person who an CA is closest to. As we detail above, if women are lower quality in men’s networks, it could lead to lower returns under performance

¹¹The purest test for the highest potential quality man or woman that an CA could bring in would put extremely large performance premium attached to a high threshold. In the interests both of maintaining a somewhat normal employment scenario, we were unable to include extremely high stakes treatments. Of course, those treatments also seem unlikely to be reflective of the referral incentives which exist in other employment settings.

pay and an advantage for men. Here, we examine the quality of mens' and womens' networks of men and women. Figure 2 presents kernel densities of the ability of men's male and female networks recruited under fixed fees. The two distributions overlap, and a Kolmogorov-Smirnov test does not statistically differentiate them. If anything, it appears that the quality of mens' networks of women dominates that of mens' networks of men. We conclude, therefore, that differences of the quality of women in men's networks as opposed to men in men's networks does not contribute to men's preference for referring men.

For women's networks, in contrast, there is a sharp difference. Figure 3 also presents kernel densities for the women and men who are closest to female CAs. The ability distribution of men who are closest to women clearly stochastically dominates the distribution of women who are closest to women, with the Kolmogorov-Smirnov test rejecting the distributions being the same at the 5% level. In terms of means, women who are closest to women perform 0.42 of a standard deviation below the CA mean, on average, while men who are closest to women perform 0.08 standard deviations below the CA mean, which is higher than the men (or women) who are closest to men though not statistically different. Our results indicate that women are closest to women who are particularly low ability.

6.3 Are there gender differences in bad information or shallowness?

Results from fixed fees are informative about the characteristics of closest people within networks, but cannot speak to other characteristics of gender-specific networks. What is potentially most interesting to employers is not the ability of the closest individuals to men and women, but rather the ability of people that men and women can identify, when properly motivated. In the theoretical model, we identified that CAs would not change their referral choices if either information was too poor, or if networks of a particular gender were too shallow. To

test whether men and women are willing to identify high quality men and women, we regress

$$Y_i = \sum_k \alpha_k T_k + \delta_t + v_i$$

as before, where Y_i could alternately be an indicator for referring a qualified referral or a normalized z-score of overall performance, T_k are the treatment categories in terms of gender and contract structure, and δ_t are time trends. Once again, CAs in restricted male, fixed fee treatments are used as the excluded group.

Table 5 presents the results of this analysis. Male CAs experience a substantial performance premium when referring men: when restricted to refer men, male CAs refer someone who is about 24 percentage points more likely to qualify when assigned to the performance pay treatment. However, they do not experience any performance premium when restricted to refer women. Female CAs show positive coefficients on performance pay for all gender treatments, but they are never significant and always small in magnitude. We therefore conclude that male CAs have useful information for employers about men, and that tradeoffs are not too high to prohibit using it.

However, men do not seem to choose female referrals with a performance premium, which could be consistent with bad information about women or with shallow networks of women. These two explanations are fairly similar: in one, men would observe a similar signal for all women that was quite accurate; in the other, men would observe different quality signals about women but know to place little weight on these noisy signals. We can, however, provide some evidence that male CAs' information about women is not very accurate. For male CAs who were asked to refer women, their referrals are similarly likely to score more than a standard deviation below the CA mean under performance pay and fixed fee treatments (18% versus

20%). These male CAs are also more likely to refer women who score more than $1/2$ of a standard deviation below the CA mean under performance pay than under fixed fee treatments (40% versus 30%). This suggests that, at least, male CAs are not declining to make a referral under performance pay if they have highly accurate but homogeneous signals of low ability. This does not rule out, of course, that there are also more stark tradeoffs among women network members than among men.

Women, in contrast, do not demonstrate a performance premium under any treatments. In tables available upon request, we document that women are nonetheless changing their optimal referral choices of both men and women: when we provide performance pay, they refer women with better English skills and who perform more strongly on math questions and ravens matrices (though only the language result is significant), and they refer men who are more likely to have worked for a survey firm in the past.¹² These suggest that women are responding to performance pay and do have some useful information for employers, particularly about other women (as cognitive ability is likely harder to observe in a resume than past experience) but that they do not have enough information or face networks which are too shallow overall to find women or men who are able to qualify.

Finally, a striking result from table 5 is that performance premia are in fact lower for men under unrestricted treatments than under male restricted treatments. Returning to the model, we observe that there is no guarantee that allowing the choice of either genders will lead to a larger performance premium, which leads to the following proposition.

Proposition 5 *Allowing the choice of either gender under performance pay leads to an unambiguously higher return rate. However, the effect on the performance premium is ambiguous if there are different performance premia for gender g and g' .*

¹²The results in table 5 are not driven by conventional applicants with previous survey experience.

Based on our definition of performance premia, if CAs refer someone for both genders who is not person 1 (recall person 1 is the individual who provides the largest social payments) when they receive performance pay, then there will be a performance premium for all CAs under both restricted treatments and under the unrestricted either gender treatment. If CAs always refer person 1 when they receive performance pay under either restricted treatment, then there will be no performance premia under either either of the restricted treatments or the unrestricted either gender treatment. However, when some CAs optimize for one gender with a performance premium, and optimize for the other gender without a performance premium, the same CAs may end up referring the performance premium candidate under the “either” option while others refer the index candidate. To better understand when we expect a performance premium, consider an CA who would opt to refer a person with a performance premium in gender g but not in gender g' . In this event, the CA will prefer to refer his premium candidate, candidate p (of gender g) over the close candidate of gender g' if the differential probability of qualification is large enough relative to the difference in social incentives, or if

$$P_i \left(\Phi \left(\frac{60 - Q_1^{g'}}{\sigma_\varepsilon^{g'}} \right) - \Phi \left(\frac{60 - Q_1^g - \gamma^g (\alpha_1^g - \alpha_p^g)}{\sigma_\varepsilon^g} \right) \right) > \alpha_1^{g'} - \alpha_p^g$$

Obviously, all of the dimensions of heterogeneity we have discussed before enter here. What is interesting for prediction is to consider how different characteristics which may be related to performance premia interact with which candidate gets referred. More specifically, if we increase Q_1 , α_p or decrease γ for the gender with the performance premium, we anticipate that CAs with the option of referring either gender will be more likely to refer the performance premium candidate. The same is true if we decrease σ_ε for either gender, or if we decrease Q_1 or α_1 for the candidate with no performance premium. In contrast, if the opposite is true,

so that we increase the quality (Q_1), social payments (α_1) of the low-performance premium gender, or if information gets worse for either gender, then we anticipate that allowing the opportunity to refer either gender would be associated with a lower performance premium than restricting referrals to gender g . Given the preceding evidence that men's information about women appears poor, the lack of a performance premium observed when men have unrestricted choices would be quite consistent with men having some women who are close and who might qualify (since information is poor, any women would have a strong chance of qualifying) and sometimes optimizing by choosing these women over the men who they know will perform well but who give them low social payments.

In section 4, we made note of the fact that there was strong evidence that male CAs were more likely to make a referral in the presence of fixed fees than performance pay, and weaker evidence that female CAs responded similarly. In principle, these differential return rates could influence our estimates of the performance premium, though the fact that we rely on differences between restricted-gender treatments (where return rates were identical) does ameliorate this concern. Still, for example, one interpretation which would be consistent with presented results is that all CAs will only refer person 1, but CAs will just attrit rather than refer person 1 under performance pay if they are in a restricted male treatment and person 1 is low quality. Interestingly, the gender differences here still require differences in information: clearly CAs have good information if they attrit because they know that their optimal referral is low quality, and so they would have to have poor information about women's capabilities if they do not attrit in the same way when required to refer women.

7 Conclusion

There is a large literature in economics and sociology which has used observational data to suggest that women benefit less from job networks than men do, so much so that two of the stylized facts presented in Ioannides and Loury (2004) survey emphasize gender differences in network usage. Using an experiment designed around a recruitment drive for real-world jobs, we provide the first experimental evidence that the use of referral systems puts women at a disadvantage. We find that qualified women tend not to be referred by networks for two reasons: first, men exhibit a preference for referring men, and second, women exhibit a preference for referring unsuitable candidates. This result suggests that the ubiquity of job networks as a hiring system could contribute to persistent gender gaps in wages.

Our experiment allows us to say several additional things about the structure of networks which could lead to a disadvantage for women. First, we confirm that women and men are present in each other's networks. Thus, while we cannot describe the relative gender homophily of networks, we can conclude that women are not being left out simply because they are fully absent from men's networks. Using variation in contract structure, instead, we find that several network characteristics work against women. First, men's networks are characterized by other men being systematically the closest members, who give the highest social incentives for a referral. However, the women who are closest to men tend to be high ability. This contrasts with women's networks, which are characterized by low quality people and especially low quality women providing the highest social incentives.

We also find that men have a greater potential to screen men than to screen women, and some weaker evidence that women may be able to screen both genders in different ways. Men also exhibited a preference for performance premium when we required them to refer men, but the preference eroded when they were allowed to refer women as well. Our framework

allowed us to say little about why job referral systems are used, but a frequent hypothesis is that they are used to help employers in screening (Montgomery, 1991; Beaman and Magruder, 2011). If employers use referral systems because they hope to screen, then this result suggests that employers may want to emphasize men referring men when hiring through networks. In other words, if screening is a productive use of networks, then this result suggests that a profit-maximizing employer may be encouraged to discriminate when using referrals to hire new workers. This result suggests that in order to prevent discrimination against women, careful hiring procedures may need to be followed. One such procedure which these results suggest would work well is a quota-based referral system, where people making references are required to make references of either gender - given that male CAs can both screen men and tend to be close to high quality women, these estimates suggest that a quota based system may be effective at identifying high quality workers of both genders.

References

- Altonji, J. and R. Blank (1999). Race and gender in the labor market. *Handbook of Labor Economics* 3, 3143–4259.
- Arbache, J. S., A. Kolev, and E. Filipiak (Eds.) (2010). *Gender Disparities in Africa’s Labor Market*. World Bank Publications.
- Beaman, L. and J. Magruder (2011). Reliability in referral networks: Experimental evidence from India. *Mimeo, University of California, Berkeley*.
- Bell, D. and M. Reich (Eds.) (1988). *Health, Nutrition, and Economic Crises: Approaches to Policy in the Third World*. Auburn House Publishing Company.
- Bertrand, M. (2011). New perspectives on gender. *Handbook of Labor Economics* 4, 1543–1590.
- Bewley, T. F. (1999). *Why Wages Don’t Fall During a Recession*. Harvard University Press.
- Black, S. E. and E. Brainerd (2004). Importing equality? The impact of globalization on gender discrimination. *Industrial and Labor Relations Review* 57, 540–559.
- Black, S. E. and P. E. Strahan (2001). The division of spoils: Rent-sharing and discrimination in a regulated industry. *American Economic Review* 91, 814–831.
- Blau, F. D. and L. M. Kahn (2000). Gender differences in pay. *Journal of Economic Perspectives* 14, 75–99.

- Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *American Economic Review* 94, 426–454.
- Doeringer, P. B. and M. J. Piore (1971). *Internal Labor Markets and Manpower Analysis*. M.E. Sharpe, Inc.
- Fernandez, R. M. and M. L. Sosa (2005). Gendering the job: Networks and recruitment at a call center. *American Journal of Sociology* 111, 859–904.
- Goldin, C. and L. F. Katz (2002). The power of the pill: Oral contraceptives and women’s career and marriage decisions. *Journal of Political Economy* 110, 730–770.
- Ioannides, Y. M. and L. D. Loury (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature* 42, 1056–1093.
- McPherson, M., L. Smith-Lovin, and J. M. Cook (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. *American Economic Review* 81, 1408–1418.
- Mortensen, D. and T. Vishwanath (1994). Personal contacts and earnings: It is who you know! *Labour Economics* 1, 187–201.
- Mouw, T. (2006). Estimating the causal effect of social capital: A review of recent research. *Annual Review of Sociology* 32, 79–102.
- National Statistics Office (NSO) and ICF Macro (2011). *Malawi Demographic and Health Survey 2010*. NSO and ICF Macro.
- Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122, 1067–1101.
- Rubineau, B. and R. M. Fernandez (2010). Missing links: Referrer behavior and job segregation. *MIT Sloan School Working Paper No. 4784-10*.
- Tassier, T. and F. Menczer (2008). Social network structure, equality, and segregation in a labor market with referral hiring. *Journal of Economic Behavior and Organization* 66, 514–528.
- World Bank Group and others (2010). *The Education System in Malawi*. World Bank Publications.
- World Bank Group and others (2011). *World Development Report 2012: Gender Equality and Development*. World Bank Publications.

Figure 1: CA Ability by Gender

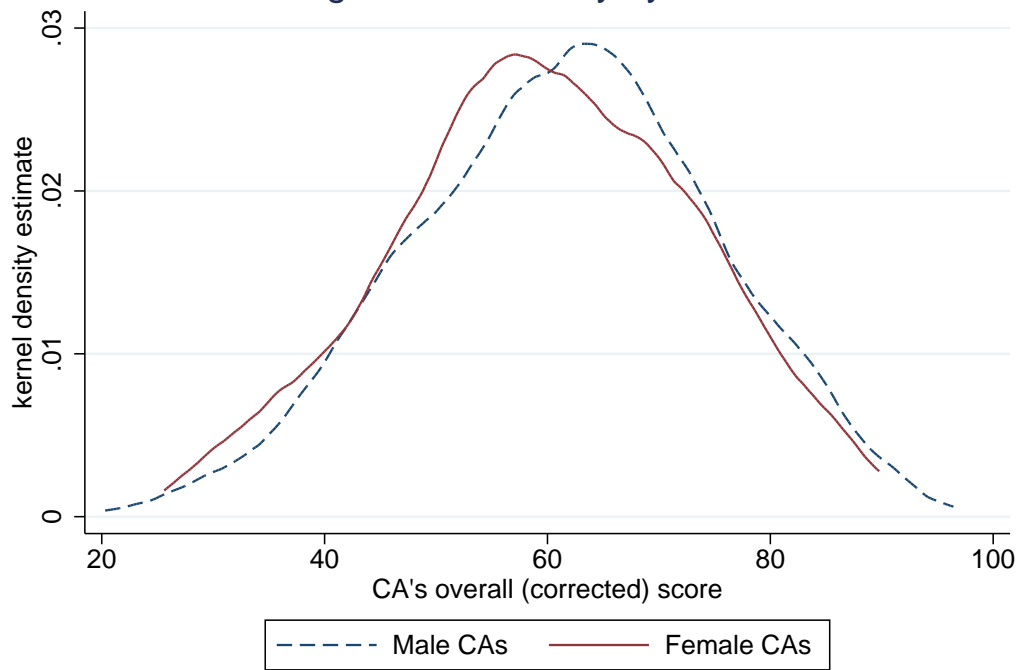


Figure 2: Men's Fixed Fee Referrals

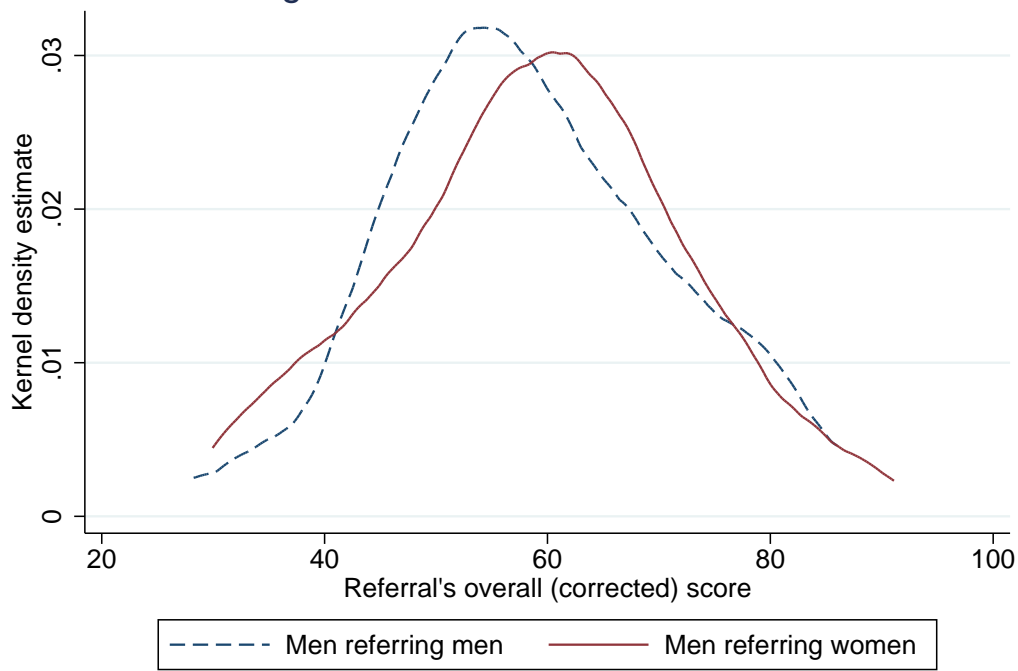


Figure 3: Women's Fixed Fee Referrals

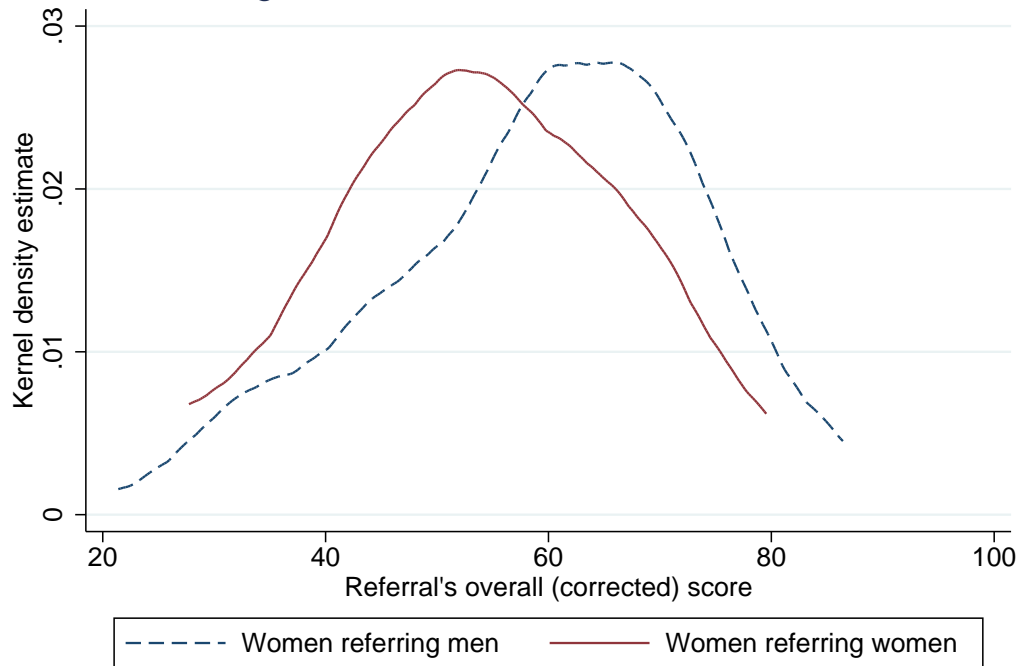


Table 1: Gender differences in CA qualification

	(1)	(2)
Dependent Variable	CA Qualifies	CA Qualifies
CA is female	-0.048 (0.038)	-0.048 (0.041)
Sample Mean - Male CAs	0.548	0.551
SD - Male Cas	0.498	0.498
Observations	768	657
Sample	All CAs	CAs who make a referral

Notes

1. The dependent variable is an indicator for the CA qualifying.

2. All specifications include indicators for the day that the CA was interviewed.

Table 2: Gender Distributions of OPs and Referrals

	(1)	(2)	(3)	(4)
	All CAs	Male CAs	Female CAs	Diff: <i>p</i> value
<u>A. OP Characteristics</u>				
Fraction of OPs	100%	61%	39%	
OP is qualified	56%	58%	53%	0.418
N	217	130	87	
<u>B. Referral Characteristics</u>				
Referral is Female	30%	23%	43%	0.002
Referral is Qualified	48%	56%	37%	0.012
Referral is Qualified Male	33%	42%	21%	0.002
Referral is Qualified Female	15%	14%	17%	0.568
N	195	117	78	
<u>C. Referral Characteristics, Fixed Fee Treatments</u>				
2. Referral is Female	32%	25%	43%	0.042
Referral is Qualified	48%	57%	35%	0.015
Referral is Qualified Male	31%	40%	18%	0.013
Referral is Qualified Female	17%	18%	16%	0.853
N	117	68	49	

Notes

1 Includes only CAs in either gender-unrestricted treatments who make referrals.

Table 3: Probability of Making a Referral

	(1)	(2)	(3)	(4)
Female Treatment	-0.004 (0.038)	-0.055 (0.054)	-0.004 (0.050)	-0.042 (0.074)
Either Gender Treatment	0.014 (0.040)	0.017 (0.055)	-0.052 (0.052)	-0.024 (0.071)
Performance Pay			-0.148 *** (0.056)	-0.113 (0.080)
Perf Pay * Female Treatment			0.004 (0.076)	-0.013 (0.111)
Perf Pay * Either Treatment			0.152 * (0.079)	0.086 (0.110)
Observations	506	310	506	310
CA Gender	Men	Women	Men	Women

Notes

1. The dependent variable is an indicator for whether the CA makes a referral.
2. All specifications include indicators for the day that the CA was interviewed.

Table 5: Referral Performance

	Referral Qualifies			
	(1)	(2)	(3)	(4)
Female Referral Treatment	-0.020 (0.063)	-0.109 (0.082)	0.082 (0.082)	-0.122 (0.112)
Either Gender Treatment	0.070 (0.066)	-0.203 (0.081)	** 0.178 (0.085)	** -0.202 (0.106)
Performance Pay			0.235 (0.094)	** 0.062 (0.121)
Perf Pay * Female Treatment			-0.254 (0.128)	** 0.029 (0.169)
Perf Pay * Either Treatment			-0.267 (0.134)	** 0.006 (0.167)
Observations	390	227	390	227
CA Gender	Men	Women	Men	Women

Notes

- 1 The dependent variable is an indicator for the referral qualifying.
- 2 All specifications include CA visit day dummies.

Appendix Table 1: Summary Statistics and Randomization Check

Dependent Variable	Mean and SD: Male	p value of joint test of treatments	N	Mean and SD: Female	p value of joint test of treatments	N
	(1)	(2)	(3)	(4)	(5)	(6)
CA Age	25.52 [3.88]	0.441	445	24.61 [4.62]	0.787	271
CA qualified	0.55 [0.50]	0.783	480	0.51 [0.50]	0.180	288
CA Overall Test Score (corrected)	61.31 [13.52]	0.567	506	59.68 [13.38]	0.121	309
CA Has Previous Survey Experience	0.31 [0.46]	0.410	480	0.26 [0.44]	0.189	288
CA Has Tertiary Education	0.69 [0.46]	0.367	480	0.78 [0.42]	0.186	287
CA MSCE Math Score	5.65 [2.30]	0.867	419	6.84 [1.80]	0.061	242
CA MSCE English Score	5.68 [1.49]	0.651	435	5.75 [1.41]	0.594	256
CA Job Comprehension Score	0.80 [0.40]	0.894	480	0.81 [0.39]	0.573	288
CA Math Score	0.21 [0.10]	0.245	480	0.18 [0.09]	0.351	288
CA Ravens Score	0.61 [0.40]	0.146	480	0.56 [0.39]	0.460	288
CA Language Score	0.15 [0.03]	0.302	480	0.14 [0.03]	0.602	288
CA Practical Component Z-score	-0.10 [1.03]	0.102	476	0.17 [0.90]	0.101	284
CA Computer Score	0.44 [0.21]	0.533	480	0.43 [0.20]	0.523	288
CA Feedback Points	25.90 [7.28]	0.037	474	27.92 [6.31]	0.252	284

Notes

- 1 The displayed p value is from the joint test of all the treatment variables and their interactions from a regression of the dependent variable listed at left on indicators for each treatment and CA visit day controls. The regressions are done separately for men and women.
- 2 All specifications include indicators for the day that the CA was interviewed.