Learning About the Employer-Employee Match When Workers Refer Job Candidates: Referrals and Search Efficiency*

Tavis Barr† Raicho Bojilov‡ Lalith Munasinghe§

March 2016

Abstract

This paper investigates the effect of information transmitted by referrals on search efficiency. We test several implications of our model using data from a call center company that contain rich information on applicants, employees, and referrers. The joint estimation of job offers, acceptances, turnover, and performance allows us to identify the contribution of referrals to search efficiency from the differences in expected performance between referred and non-referred applicants. The results show that referrals induce selection on unobservables at the job offer stage, which in turn drives referred employees to perform better, receive early promotions and stay longer on the job.

KEYWORDS: Referrals, employer learning, learning about match quality

JEL CODE: J31, J33, M51

---

*An earlier version of the paper was circulated under the title "When Workers Refer Job Candidates, Who Learns What? Referrals and Search Efficiency." We would like to thank P. Cahuc, L. Gee, M. Hoffman, G. Hollard, H. Kleven, F. Kramarz, O. Nordström Skans, B. Salanié and many seminar participants for their helpful comments and suggestions. Bojilov's work is supported by a grant of the French National Research Agency (ANR), "Investissements d'Avenir" (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

†TalenTeck. E-mail: tavis.barr@talenteck.com

‡Department of Economics, Ecole Polytechnique; CREST. E-mail: raicho.bojilov@polytechnique.edu.

§Department of Economics, Barnard College; Columbia University. E-mail: lm25@columbia.edu.

1 Introduction

Employment match quality often depends on idiosyncratic characteristics of firms and prospective employees that are typically difficult, if not impossible, to verify through commonly observable metrics. This is especially the case prior to the start of an employment relationship. The extreme version of match uncertainty is articulated in the theory of search by experience, where information about match quality is revealed only on the job. In reality of course, firms and job candidates learn about their potential match quality both during the hiring process and, if the candidate is hired, over the duration of their employment relationship. The information signals revealed during the recruiting and hiring process are implicated not only in job offers and acceptances but also in subsequent employment outcomes, such as performance and turnover dynamics. In an environment of information uncertainty, a match-maker - for example, a current employee who knows both the firm and the candidate well - can provide information to resolve some of the uncertainty about match quality prior to job start. Perhaps for this reason, referrals are a common feature in the labor market: Topa (2011) reports that about 70 percent of all U.S. firms have programs to encourage hiring through referrals, while more than half of all new jobs in the US are found through informal networks.

Recently, there has been a renewed interest in understanding the role of referrals in employment outcomes based on experimental data and the personnel records of firms that implement formal referral programs. This literature, including Brown, Setren, and Topa (2016) and Burks, Cowgill, Hoffman, and Hausman (2015), extensively documents that referred candidates are more likely to be hired, have longer tenure, and higher starting wages than their non-referred counterparts. Hansvik and Skans (2016) also establish that referrals have worse observable characteristics. Moreover, Pallais and Sands (2016) perform a random-hiring experiments and show that referrals contain positive information about productivity that cannot be inferred from observable characteristics. Still, important questions related to the channels through which referrals impact search dynamics and efficiency remain unexplored. For example, what specific match-related information, if any, do referrers transmit to employers and referred job candidates? Is the associated referral signal indicative of match quality that can be learned only on the job? Or is this signal verifiable during the meet-and-greet of the hiring process? Furthermore, at what stages of the hiring process do referrals have the greatest impact on search efficiency? And finally, to what extent is the quality of the referral signal impacted by the relationship between the match-maker and the referral on the one hand, and between the match-maker and the company, on the other?

We model the hiring process and the role of referrals within a standard search theoretic framework. Match quality between a firm and a potential job applicant is initially unknown. The meet-and-greet of the recruiting and hiring processes resolves some of this match uncertainty for both parties. On the basis of this updated assessment of match quality the firm decides whether to extend a job offer and, subsequent to such an offer, the applicant decides on whether to accept the offer or not. If the applicant accepts and is hired, then the true match quality is revealed over the course of the employment relationship. The extent to which match uncertainty is resolved prior to employment on account of the information signals revealed during the hiring process not only determines the offer and accept decisions, but also the dynamics of on the job performance and subsequent turnover. The key distinction between a referred and non-referred applicant is that in the case of the former, additional information signals are brought to bear in the hiring process. Referrals may simply indicate that a candidate comes from a pool of applicants whose distribution of productivity is stochastically superior to the distribution of non-referred employees. They may also provide information about the quality of each specific potential employee-employer match. In particular, if the referrer shares such information that is not easily observed by either the firm or the applicant, then a referred employee is more likely to perform better and stay longer on the job compared to her observationally equivalent non-referred counterpart. This is the case even when their actual observable and unobservable characteristics (at least partially unknown to them) are the same.

Our model encompasses several limit cases of interest when the only differences between referred and non-referred applicants are informational. In one extreme, if the formal screening process does not provide any information on the difficult to observe dimensions, then the entry decision of non-referred candidates will only depend on observable characteristics. However, if a referrer provides information about some of the unobservable aspects of the match, then the hiring and employment outcomes of a referred applicant will differ from those of an observationally identical non-referred candidate. In the other extreme, if true match quality is fully revealed during the screening process and thus entry decisions are based on actual match quality (the case of search by inspection), then hiring and employment outcomes will not differ between referred and non-referred individuals. In the general case, as long as the referral process generates additional informative signals about the idiosyncratic characteristics of a potential match, referred individuals will sort more strongly than their non-referred counterparts into the following states: employment, stay, performance, and promotions. We further examine whether the referred candidate and the firm share the same information during the hiring process and, if not, which party is better informed.

We investigate these issues using the personnel records of US contact centers of a large global customer services company. Our dataset provides a multi-stage breakdown of the hiring process, including job offer and acceptance decisions, as well as post-hire information such as termination dates, a performance measure for those who stay long enough, and information on early promotions. The data also include standard demographic controls, information on work experience, educational attainment, cognitive and non-cognitive skills, and a set of local labor market variables for each candidate on county and zip code level. Moreover, we have detailed information on whether a candidate was referred by a current employee through a formal company-provided referral program. To investigate the possibility of sorting on unobservables and for any related differences between referred and non-referred candidates, we estimate a multi-stage model of entry, stay, and performance. In our empirical analysis, we allow for differences not only in the error process between referred and non-referred applicants, but also for differences in the impact of observable characteristics.

By explicitly modelling how referrals impact sorting on both observable and unobservable dimensions, we are able to recover the distributions of performance of referred and non-referred candidates and how these evolve at each stage of the hiring process and early employment.

The results show that referrals contribute to search efficiency in several ways. Referred applicants have superior performance compared to their non-referred counterparts controlling for selection effects. These differences in performance are largely due to the fact that referrals allow the employer to identify pools of candidates who on average turn out to be of high quality on difficult to observe dimensions, either specific to the employer-employee match or related to general abilities and qualifications, that otherwise become known only post-hire. At the same time, the results also show that the same observable characteristics may have differential impact: for example, the effect of educational attainment and related prior-experience on performance are stronger (and more precise) for the referred compared to the non-referred.

Moreover, we also find evidence that the referral signal speeds up learning about match quality. Sorting during the hiring process and early employment reduces the differences between the performance of referred and non-referred long-term employees relative to the differences between the performance of referred and non-referred candidates. All employees eventually find out their true match quality with the firm, but referred employees complete much of the associated sorting on unobservables during the hiring process. For the subsample of referred applicants, we find a strong positive association between the errors, i. e. unobserved components, in the entry, stay and performance equations. That is, controlling for their observable characteristics and particularly for the fact that they are referred, applicants who are unexpectedly more likely to get hired are also more likely to stay longer and perform better. In contrast, for the subsample of non-referred candidates we find no such evidence of selection on unobservables during the hiring process.

To investigate whether referrers – i.e. match-makers – provide the same information to both parties, we model separately the firm's offer decision and the candidate's acceptance decision within a multi-stage choice model. Our estimates show that there is a strong positive correlation between the errors in the offer, stay, and performance equations for referred individuals. At the same time, the decision of a referred applicant to accept an offer is uncorrelated with the likelihood of staying longer on the job or high performance. In addition, the correlation between the error terms in the offer and acceptance equations is positive but not statistically significant. These patterns contrast sharply with the results for the non-referred applicants, for whom we do not find any significant selection effects on unobservables at any stage of the hiring process.

Our estimates suggest that prior to employment neither the firm nor the job candidates are fully informed about their potential match quality, but that the referral process provides some of this missing information. More specifically, referrers transmit signals that help the firm make job offers to candidates who eventually turn out to have high performance, presumably due to some initially unobserved characteristics. In contrast, whatever the information referrers may or may not have communicated to the applicants whom they have referred, it does not seem to induce a similar selection on unobservables at the time of job acceptance. Finally, while we show that the selection on unobservables is not sensitive to the characteristics of the relationship of the matchmaker, i.e. the referrer, with the firm and with the referred job applicant, it also turns out that referrals provided by experienced former coworkers who have known the referred candidate for more than five years lead to better matches.

The rest of the paper is organized as follows. Section 2 relates the results of our paper to the rest of the literature. Section 3 investigates the econometric implications of informative referral signals within a model of learning about match quality. Section 4 presents the data and conducts some basic descriptive analysis. Section 5 reviews our estimation approach, while Section 6 presents our empirical results. Section 7 concludes.

2 Related Literature

Our work contributes to a large body of experimental and non-experimental literature on referrals. Recently, Burks et al. (2015) review referral practices across companies within several industries and show that referrals boost profits by cutting hiring and turnover costs. Along with Brown et al. (2016), they confirm a widely documented pattern: referred candidates are more likely to be hired, enjoy higher starting wage, and have longer tenure. Hansvik and Skans (2016) provide strong evidence, in addition to the descriptive analysis in Burks et al. (2015), that referrals have better hard-to-observe but worse observable characteristics. A related earlier strand of the literature has investigated referrals from a sociological perspective. Fernandez and Weinberg (1997), Fernandez and Castilla (2000, 2001) and Castilla (2005) use firm-level data from a retail bank and a call center to study the role of referral networks on turnover, compensation, and performance for low to moderate skilled jobs. In particular, Castilla (2005) tests formally but rejects the hypothesis that selection on unobservables affects entry, stay, and performance. In contrast, Pallais and Sands (2016) perform a set of field experiments that show that referrals contain positive information about performance that cannot be inferred from observables. They warn that the associated selection on unobservable dimensions may account for the previously documented fact that referrals have large impact on entry and stay, but the differences in performance of referred and non-referred workers while persistent are relatively small. Yet, the ability of Pallais and Sands (2016) to explore the associated implications for search dynamics and efficiency is limited, since by experimental design they abstract away from the actual hiring process and turnover dynamics. In their setting Pallais and Sands (2016) rely only on implicit incentives to motivate individuals to provide referrals, while Beaman and Magruder (2012) show that the referral of high quality performers is highly sensitive to pay incentives. Thus, the bias due to selection on unobservables may be larger when firms give an explicit bonus for successful referrals.

By jointly estimating the decisions of the employer and job candidates, we contribute to the literature by identifying in a non-experimental setting the distributions of performance of referred and of non-referred applicants from selection effects induced by referrals. Our estimation approach relies on a flexible functional form specification that allows us to quantify how the information transmitted by referrals contribute at each stage of the hiring process to sorting in and out of employment. In this context, the results suggest that unwarranted functional form restrictions may account for the puzzling fact that, while there is much recent evidence that referrals are related to hard-to-observe characteristics of the employer-employee match, previous formal tests reject the hypothesis of selection on unobservables. We point out that, except in the very special case when the only transmitted information is that all referred candidates are expected to be on average better than their non-referred counterparts, the empirical analysis of referrals requires the explicit modelling of selection on unobservables.

This paper also contributes to a literature that has directly tested predictions of models in which referrals transmit information about match quality. Such research relates closely to the theoretical literature on search by experience, starting with Jovanovic (1984). Simon and Warner (1992) present an early application of this theoretical framework to the analysis of referral networks, while more recently Galenianos (2013, 2016) has explored the equilibrium implications of referrals when workers learn about match quality. Montgomery (1991) offers an alternative but closely related framework of analysis of referrals based on employer learning about the quality of candidates from those of their contacts in the firm. The observation that information asymmetries induce selection on unobservables also relates to Munasinghe (2006), which shows the impact of different priors on labor market outcomes such as turnover, wage dynamics, and promotions. Using firm level data, Brown et al. (2016) report results that are consistent with the bulk of the predictions generated by these models: conditional on being hired, referred employees have higher initial wages and longer tenure than non-referred employees; still the differences in the compensation of referred and non-referred individuals decrease over time. Moreover, Hansvik and Skans (2016) test the hypothesis that referrals convey information that would otherwise remain unknown to employers until after the beginning of the employment relation. Their work relies on administrative data about Swedish military test scores which the econometrician observes but the potential employers do not.

Other papers that explore testable implications of referral models include Dustmann et al. (2016), which however relies only on an indirect proxy for referral networks. Similarly, Kramarz and Skans (2014) investigate how family networks affect employment outcomes, while Oyer and Schaefer (2012) consider the impact of educational institutions. Other related papers in the style of Bayer et al. (2008) have also explored the implications of referral networks based on place of residence. Last but not least, recent contributions on the impact of social networks on employment outcomes include Gee et al. (2016) and Galenianos (2014), which investigates theoretically the endogenous formation of referral networks. This paper also abstracts away from issues related to the formation of social networks and their equilibrium implications, central to research in the tradition of Calvo-Armengol and Jackson (2004), but our results are consistent with the equilibrium predictions in Galenianos (2013).

Relative to Hansvik and Skans (2016) and Brown et al. (2016), we explore how the transmission of hard-to-observe information through referrals affects the dynamics of the hiring process and early employment. Specifically, we show that learning about match quality implies that referred individuals select in and out of employment on unobservables more strongly than non-referred individuals. In contrast, learning by inspection implies that the subsamples of referred and non-referred applicants and employees exhibit the same correlation structure for the errors in the entry, stay, and performance equations. Indeed, we show that referrals induce selection on unobservables that would not be present otherwise. These results are consistent with the hypothesis that referrals speed up the process of learning about match quality. We also explicitly model the firm's decision to extend job offers and the candidate's decision to accept such an offer and document that the performance of referred individuals is highest when experienced employees refer former coworkers whom they have known for more than five years. Consequently, the information transmission hypothesis generates strong predictions that can be tested even when the information set of the econometrician is not superior to that of the potential employers and job candidates.

3 Model

The employment match between a firm and a worker generates a production surplus that is not necessarily known precisely at the time of hiring. In the course of the hiring process, employers and candidates may observe signals about the quality of their potential employment relationship. In addition, third parties, such as existing employees, may refer specific candidates, which also provides information about the quality of the employment match. The notion that referrals transmit information that is not otherwise available implies the presence of heterogeneity between the information sets of referred candidates, non-referred candidates, and the potential employer. The actually communicated signal is not observed by the econometrician and may even remain partially or fully unknown to some of the market participants.

To investigate this issue formally, we revisit the theoretical framework of standard models of search in the tradition of Jovanovic (1984). Our setting is very close to the one investigated in Simon and Werner (1992), Galenianos (2013), Brown, Setren, and Topa (2016), and Dustmann et al. (2016). Unlike these papers, we focus on the testable implications of referral signals for the dynamics of the hiring process and early turnover. Our main result relates to the observation that, if referral signals improve the precision of beliefs, referred candidates sort more efficiently in and out of employment than non-referred individuals. Thus, the model predicts that referrals boost the stochastic dependence between entry, stay, and actual performance.

3.1 Environment

Some job candidates have a referral, $r = 1$, while others do not, $r = 0$. Production surplus depends on characteristics x, known to everyone, and firm-worker specific match quality θ, which is unknown to the firm and the candidate at the beginning of the hiring process. The firm and the individual learn about this match quality through the hiring process and post-hire performance (productivity). In addition, their beliefs are influenced by referral information that may or may not be the same for the two sides in the market. In the literature, referrals are associated with two distinct phenomena: on average referred applicants may be of superior quality, which market participants are likely to know; and referrals may also provide informative signals about match quality. We focus on this latter issue below.

Match quality θ_r of both referred and non-referred candidates is drawn from a common normal distribution, $N\left(\mu, \sigma^2\right)$, which is independent from x and its actual realization is revealed after the candidate is hired. The firm and each applicant share a common prior which coincides with the distribution of θ. They learn about the specific match quality through Bayesian updating. All candidates go through the same hiring process which generates a signal that may or may not be the same for the different sides of the market. The candidate receives a signal $\theta_r + \xi_c$, where $\xi_c \sim N\left(0, \sigma^2_{\xi c}\right)$, while the firm receives a signal $\theta_r + \xi_f$, where $\xi_f \sim N\left(0, \sigma^2_{\xi f}\right)$. Referrals provide additional information to the firm and the candidate, which also may or may not be the same: the referred candidate receives a signal $\theta_1 + \zeta_c$, where $\zeta_c \sim N\left(0, \sigma^2_{\zeta c}\right)$, while the firm receives a signal $\theta_1 + \zeta_f$, where $\zeta_f \sim N\left(0, \sigma^2_{\zeta f}\right)$.

Based on the available information, the firm decides whether to make an offer or not. If the candidate obtains an offer, she decides to accept it or not. For simplicity, match quality becomes known after hiring, and the econometrician observes a performance signal for those who stay upon learning their true match quality:

$$y = \theta_r + f(x) + \epsilon \quad \text{for} \quad r = 0, 1$$

where ϵ and x are independent from match quality and from each other. Profits and individual utility are linear functions of output. Finally, the firm and its candidates have an outside option, normalized to zero.

Our model specification allows for information asymmetries between the firm and the candidates. We assume that the signals and match quality are normally distributed in order to simplify the exposition and preserve the closed-form formulas of the posterior beliefs. We also consider an additive technology in match quality, since this specification conforms to the statistical properties of our data. It also happens to be the most commonly used specification in the preceding empirical literature. The rest of the assumptions are standard for such models and allow us to identify the effect of referrals on the net value of employment in a particular firm relative to alternative jobs.

To complete the setting, we maintain that the econometrician observes whether a candidate is referred or not, whether she receives an offer or not, whether she accepts it, if such is extended, how long she stays in the firm, and her performance, conditional on sufficiently long tenure. In particular, we allow for the possibility that during the hiring process, the employer and the applicant have some information about the quality of their potential employment match that remains unobserved to the econometrician.

3.2 Testable Predictions

This environment has already generated a number of testable predictions summarized in Brown, Serten, and Topa (2015). Our key new insight is that the informational content of referrals generates strong testable predictions for the dynamics of the hiring process. Individuals and the firm update their beliefs about match quality following Bayes' rule. For non-referred candidates, the posterior belief after observing the signal from the hiring process is $N\left(\mu_{0c}, \sigma_{0c}^2\right)$, where

$$\sigma_{0c}^2 = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_{\xi c}^2}\right)^{-1} \text{ and } \mu_{0c} = \left(\frac{\mu}{\sigma^2} + \frac{\theta + \xi_c}{\sigma_{\xi c}^2}\right)\sigma_{0c}^2$$

In contrast, referred applicants have an additional signal and form posterior beliefs about the match quality $\theta_{1c} \sim N\left(\mu_{1c}, \sigma_{1c}^2\right)$, where

$$\sigma_{1c}^2 = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_{\xi c}^2} + \frac{1}{\sigma_{\zeta c}^2}\right)^{-1} \text{ and } \mu_{1c} = \left(\frac{\mu}{\sigma^2} + \frac{\theta_1 + \xi_c}{\sigma_{\xi c}^2} + \frac{\theta_1 + \zeta_c}{\sigma_{\zeta c}^2}\right)\sigma_{1c}^2$$

The posterior mean is just the weighted average of the prior and the informative signals about match quality. Since referred candidates observe strictly more informative signals, their posterior means are more strongly correlated with actual match quality than the posterior means of non-referred candidates. The variance of the posterior beliefs of referred individuals is also lower than the variance of posterior beliefs of non-referred individuals. This observation plays a crucial role in the derivation of the testable implications below. The same argument extends to the posterior beliefs of the employer, $N\left(\mu_{rf}, \sigma_{rf}^2\right)$, $r = 0, 1$. If they share the same information during the hiring process, the firm and the candidate have the same posterior beliefs, $N\left(\mu_r, \sigma_r^2\right)$, and agree on the value of employment.

Suppose that after all signals are observed the value to the firm of employing a candidate of type r is $v\left(\mu_{rf}, \sigma_{rf}^2, x\right)$. The applicant receives an offer if $v\left(\mu_{rf}, \sigma_{rf}^2, x\right) > 0$ and not otherwise. Following Jovanovic (1984), we solve for the threshold posterior mean $\underline{\mu}_{rf}$ that makes the firm indifferent between the two alternatives:

$$v\left(\underline{\mu}_{rf}, \sigma^2_{rf}, x\right) = 0 \;\Rightarrow \underline{\mu}_{rf} = \underline{\mu}_{rf}\left(x\right)$$

where σ^2_{rf} is absorbed into the functional form of $\underline{\mu}_{rf}$, since the posterior variance does not depend on the specific signals. As the precision of beliefs increases, the option value of employment decreases which, pushes up the threshold posterior mean. As a result, when the firm has more precise posterior beliefs for referred than for non-referred candidates, $\sigma^2_{0f} > \sigma^2_{1f}$, it requires higher posterior mean for entry from the referred than from the non-referred, $\underline{\mu}_{0f}\left(x\right) < \underline{\mu}_{1f}\left(x\right)$. In a similar way, we obtain the thresholds for the acceptance and stay decisions, $\underline{\mu}_{rc}\left(x\right)$ and $\underline{\theta}\left(x\right)$.[1] For simplicity, we maintain that candidates are 'shortsighted' in the sense that they do not update their beliefs after receiving an offer.

Thus, the optimal decision rules give rise to a multistage model of offer, acceptance, stay and performance for both referred and non-referred candidates, $r = 0, 1$:

$$\begin{aligned} o &= 1\left[\mu_{rf} > \underline{\mu}_{rf}\left(x\right)\right] \\ a &= 1\left[\mu_{rc} > \underline{\mu}_{rc}\left(x\right)\right] \\ s &= 1\left[\theta_r > \underline{\theta}\left(x\right)\right] \\ y &= f\left(x\right) + \theta_r + \epsilon \end{aligned}$$

where o is a binary indicator for offer, a for acceptance, s for stay, and y denotes performance. The acceptance decision is observed if $o = 1$, the stay decision is observed if $a = 1$, and performance is observed if $s = 1$. The functions $\underline{\mu}_{rk}\left(x\right)$ and $\underline{\theta}\left(x\right)$ decrease in x, $\underline{\theta}\left(x\right) > \underline{\mu}_{rk}\left(x\right)$, and $\underline{\mu}_{1k}\left(x\right) > \underline{\mu}_{0k}\left(x\right)$, where $k = f, c$ and $r = 0, 1$.

This representation of the optimal decision rules highlights how the accumulation and transmission of information leads to dependence in the outcomes observed by the econometrician. It also emphasizes that the hiring dynamics of referred and non-referred candidates differ

[1] Most testable predictions in the existing literature are derived from these threshold conditions.

completely, which implies that only a sufficiently flexible specification can hope to recover the effect of referrals on offer, acceptance, stay, and performance. In such a context, the fact that some observationally equivalent candidates are hired, while others are not, provides information to the econometrician about future performance, stay, and promotions, even conditional on observed characteristics. Since referred applicants have more precise beliefs about their match, their hiring has greater predictive power about stay and performance than the hiring of a non-referred candidate, even conditional on all observable characteristics. The following proposition states these observations formally.

Proposition 1 *Under the assumptions of the model, if referrals provide an informative signal about match quality, there is stronger positive dependence between offer, acceptance, stay, and performance for the referred than for the non-referred candidates:*

$$Corr\left(\mu_{1f}, \theta_1\right) \geq Corr\left(\mu_{0f}, \theta_0\right)$$
$$Corr\left(\mu_{1c}, \theta_1\right) \geq Corr\left(\mu_{0c}, \theta_0\right)$$

These relations hold with equality when the signal during the hiring process becomes perfectly informative, i.e. as $\sigma^2_{\xi k} \rightarrow 0$ *and* $\mu_{rk} \rightarrow \theta_r$ *for* $k = c, f$. *Also, if the hiring signals are informative,* $Corr\left(\mu_{0k}, \theta_0\right) > 0$ *for* $k = c, f$ *. When the candidate and the firm observe the same signal during the hiring process,* $Corr\left(\mu_{0f}, \mu_{0c}\right) = 1$. *When they also observe the same referral signal,* $Corr\left(\mu_{1c}, \mu_{1f}\right) = 1$.

This proposition incorporates several special cases of interest. Suppose that referral signals are informative but those associated with the hiring process are not, i.e. $\sigma^2_{\xi k} \rightarrow \infty$ for $k = f, c$. Then controlling for the observable characteristics, offer, acceptance, and stay decisions of non-referred individuals are completely independent: entry has nothing to do with the underlying match quality on the difficult to verify or observe dimension. In contrast, the decision of a referred candidate has predictive power for the decision to stay, even after controlling for observable characteristics and the referral status. This difference in the stochastic dependence of entry and stay decisions between the referred and non-referred individuals extends to the more general case when the hiring process provides at least some information about the potential value of the employment relation. In another extreme case, when the hiring process resolves all uncertainty, $\sigma_\xi^2 \to 0$, all candidates are hired on the basis of their true match quality. Thus, the association between the entry, stay, and performance is exactly the same for both referred and non-referred individuals.

To conclude, referral signals induce selection in and out of employment on dimensions that are unobservable to the econometrician and may remain partially observable to market participants. Therefore, to evaluate their contribution to search efficiency, one must recover their effect at each stage of the hiring and employment process.

4 Data

Our empirical environment provides several advantages: there is a formal referral system with verifiable information, both the referred and non-referred candidates go through the same hiring process, and we capture detailed information on all candidates and on the quality of the referral relationship. The descriptive analysis reveals that the referred candidates are more likely to proceed to the following stages of the hiring process and more likely to stay at any tenure horizon. At the same time, at any stage of the hiring process and early employment, the non-referred candidates have superior observable qualifications to the referred candidates. Moreover, conditional on observable characteristics, referred employees still have a slightly higher long-term retention rate, but not a perceptibly different long-term performance. We interpret these findings as evidence for selection on observable and unobservable characteristics during the hiring process.

4.1 Environment and Referral Program

The data come from ten US-based call centers of a large multinational company, whose main activity is debt collection. The personnel records allow us to track hiring and employment outcomes for any individual who applied for a job at the company. The company instituted its formal referral system in August 2011. While all candidates go through the same hiring process, some of them may be referred by current employees in the company. The referrers complete a detailed form in which they indicate the name of the referred individual, the context in which they met, and the duration of their acquaintance. If a referred candidate is hired and stays for more than one year, the referrer obtains a sizable bonus: for example, when someone is hired through a referral at the lowest hierarchy level and stays for one year, the cash reward is 1000 US dollars. The bonus is higher if the referred candidate enters the company at a higher hierarchical level.

Potential employees are requested to fill in an online survey of around 50 questions on background, employment and education history, reading comprehension, math, logic, response to work situations, language, and non-cognitive skills. The individual answers are then aggregated into a score to determine whether a candidate should be rejected. Crucially, the scores remain for internal use only and are never communicates to the candidates or the newly-hired employees. Each location determines its threshold score, which fluctuates monthly, largely due to demand shocks for the services of the company. Those who pass the threshold score go through an interview. The approved candidates receive job offers, mostly within two weeks from the interview. Those who accept the job offer attend a two week training program. They get introduced to the work, take some obligatory courses related to their and the company's legal obligations and rights, and pass an exam to certify that they have a good understanding of these matters. Most new employees enter the company at the bottom of the hierarchy, but some deemed to be of high quality start on higher hierarchical levels. These early promotions are finalized after the completion of the training. More than 95% of all employees work full time. Each workstation consists of a computer, a telephone, and a recording device. Importantly, only one worker handles a given call and an automatic switchboard assigns inbound and outbound calls by matching the call at the top of the queue with the longest waiting operator.

We use as our performance measure an indicator designed by the firm that aggregates productivity signals on multiple dimensions, such as total debt collection, average handling time of a call, adherence to schedule, quality of customer service, etc. Individual performance is evaluated and recorded regularly, once every three months. In our empirical work, we focus on average performance during the first six months of employment, which also coincides with the tenure horizon at which the hazard rate of quitting levels off. Performance is recorded on a scale from zero to four, where higher numbers correspond to superior outcomes and the spread of outcomes over its fine scale allows us to treat average performance as a continuous variable in the following empirical work. Thus, our measure of performance shares common features with schooling grades. It also has some major advantages including multiple observations on the performance of most individuals at a given task and a great similarity in the tasks assigned to different individuals. We take as given that the firm is using the correct formula to arrive at its indicator.

Promotions are closely related to the performance measure: those who receive 3 or more are considered for promotion to the next hierarchical level. Compensation consists largely of a base pay linked to the hierarchy level and it is relatively high for low-skilled workers and comparable to that in the manufacturing sector: workers of tenure longer than six months receive between 14 and 21 dollars per hour. Importantly, the performance measures are completely unrelated to the recruitment process or the people involved in it, and the promotion decisions are made by a committee of supervisors who are generally not involved in the interview and hiring decisions of potential subordinates.

4.2 Descriptive Analysis

We limit our analysis to the candidates and employees engaged in the main activity of the firm, debt collection. Furthermore, we restrict our attention to referrals made by regular call operators, which account for more than 92 percent of all referrals. These restrictions are imposed in order to limit concerns about favoritism and heterogeneity in tasks across the workforce. The resulting sample includes information on 145 730 candidates who have applied in the period from August 2011 to July 2013.

Table 1 reveals that referred candidates are much more likely to proceed to the following stages of the hiring process than non-referred candidates. Approximately 56 percent of those referred are interviewed and 30 percent of them receive an offer. In contrast, about 30 percent of the non-referred are invited for an interview and a total of 14 percent end up with an offer. In other words, conditional on having an interview, the probability that a referred candidate receives a job offer is 0.57, while the corresponding probability for the non-referred is 0.45. We also find that 77 percent of the referred candidates who receive an offer accept it, while 74 percent of the non-referred candidates accept an extended offer. Combining all effects, the probability that a referred candidate starts work at the end of the hiring process is approximately 0.19 compared to 0.07 for the non-referred. The comparison between the absolute numbers is even more revealing. While the firm considers 115, 893 non-referred and only 29,837 referred applicants, it actually hires 8,906 non-referred and 5,844 referred individuals. Finally, we observe that the difference between the probabilities of progressing to the next stage of the hiring process of referred and non-referred candidates decreases with each successive stage. The bottom part of Table 1 presents the stay rates for referred and non-referred workers at various tenure horizons. Its striking feature is that there appears to be little difference between the referred and non-referred individuals conditional on being hired.

At first sight, this finding seems to suggest that referrals play a role only at the hiring stage due to the successful screening practices of the firm. Yet, if referrals provide a positive signal about the quality of the potential employment relation that is otherwise difficult to observe, then non-referred candidates start with a disadvantage. Consequently, if they are to pass the hiring criteria of the firm, they would have to have an advantage on other dimensions. In other words, referred and non-referred workers may have the same value to the firm but for different reasons: the referred because of, say, specific difficult to observe quality, while the non-referred because of a mixture of other, easier to verify, qualifications. As a result, screening during the hiring process acts to diminish the differences in employment outcomes between the different types of workers actually hired compared to the corresponding differences in the general pool of applicants.

We explore this issue further in Table 2 which summarizes the observable characteristics of referred and non-referred individuals at different stages of the recruitment and the employment relation. The table shows that as the hiring process progresses the characteristics of the remaining non-referred candidates are superior to those of the remaining referred candidates. While the differences at each stage may not be statistically significant each time, they persist consistently across the various hiring stages and during the employment relation itself. The fraction of non-referred applicants with only high school education drops from 45 percent in the pool of all candidates to 34 percent among those who receive an offer and remains stable thereafter. In contrast, the referred applicants have slightly worse educational credentials with 47 percent holding only a high-school degree. The share of referred individuals with only a high school degree drops to about 40 percent among new hires and reaches 37 percent among those who stay employed for at least 6 months.

Similarly, non-referred candidates are more likely to have prior call center experience: the share of non-referred individuals with such a qualification at all stages of the hiring process and the employment relation ranges between 58 and 62 percent, while the corresponding range for the referred is 56-58 percent. We also note that the same pattern persists for general work experience. Throughout the hiring process referred individuals are more likely to be with little or no work experience than non-referred candidates. Similarly, while all applicants are equally likely to have at least 5 years of work experience, the screening during the hiring process introduces a wedge: the share of non-referred hires with at least 5 years of prior experience is 55 percent, while for referred workers it is 51 percent. Interestingly, referrals also allow the firm to attract candidates who live further away from the premises of the call center. Among the applicants, those with an offer, the new hires, and the stayers, the distance from home to work for the referred is about 2-3km greater than the corresponding distance for the non-referred. Finally, we note that the discrepancies between the observable qualifications of the referred and non-referred pools of hires persists during the employment relation itself.

We complete the descriptive analysis by investigating outcomes conditional on observed characteristics in Table 3. In this way, we can evaluate whether the differences reported above are driven by compositional differences between the observable characteristics of referred and non-referred individuals. The summary statistics show that even conditional on having similar observable qualifications, referred applicants are much more likely to receive an offer and to accept it than their non-referred counterparts. While referrals make workers more likely to stay at any tenure horizon, the effect dissipates over time. To investigate how the hiring and turnover dynamics relate to underlying differences in ability, we also compare average performance of referred and non-referred 'long-run' stayers in the firm. In particular, Table 3 reports average performance for the first six months of employment of workers who stay more than six months. The results show that there is virtually no difference in the performance of the remaining referred and non-referred workers. We also find that having only high school education has negative impact on entry, stay, and performance. Furthermore, prior experience and, in particular, prior call center experience increase the probability of being hired, stay, and performance. Finally, we also document that individuals who live close by are more likely to be hired and remain employed.

Together, Tables 1, 2 and 3 lead to several observations that motivate the following empirical work. Clearly referrals have a major impact on the hiring process, but the differences in turnover between the referred and the non-referred after that are smaller. Crucially, the observable qualifications of the non-referred individuals are consistently superior to those of the referred. Finally, there appear to be no differences in the performance of referred and non-referred stayers in the long run. In combination, these facts suggest that referred and non-referred long-run stayers may end up with different mixes of observable and unobservable characteristics that still yield similar performance. These are the classical signs for selection. Therefore, controlling and quantifying the effect of selection on observables and unobservables during the hiring process and the first months of employment is crucial to evaluating the effect of referrals. Moreover, the differences in the hiring and turnover dynamics between referred and non-referred individuals suggest that referrals provide relevant information about the quality of the employment match.

5 Estimation

This section starts by introducing the specification taken to the data that allows us to investigate the testable implications of the model from section 3. It continues with the review of our estimation approach.

5.1 Specification

Similarly to other choice models, one can use empirical probabilities to recover the underlying parameters. Consistent with the theoretical model, we maintain that the error processes for the referred and non-referred differ and that the observable characteristics have differential impact for referred and non-referred individuals. However, we do not impose additional restrictions on the model, since our primary interest lies in controlling for and quantifying the afore-mentioned effect of unobserved heterogeneity induced by referral-based hiring. We model the outcomes from the hiring process and the employment relationship by conditioning on information available at the time someone applies for a job at the company.[2] Let subscript i denote observations associated with candidate i. As before, subscript $r = 0, 1$ indicates whether a candidate is not referred or referred, respectively. We define X_{ir}^{k} to be a vector of characteristics observable to the econometrician that impact outcome k, where $k = o, a, s, y$ stand for offer, acceptance of an offer, stay, and performance.

The following sequential choice model is taken to the data. First, the firm decides to make an offer to candidate i who is referred or not, $r = 0, 1$:

$$o_i = 1\left[F_{or}\left(X_{ir}^{o}\right) + \varepsilon_{ir}^{o} > 0\right] \tag{1}$$

If the candidate receives an offer, she accepts it or not:

$$a_i = 1\left[F_{ar}\left(X_{ir}^{a}\right) + \varepsilon_{ir}^{a} > 0\right] \tag{2}$$

[2] In this sense, our specification follows the approach of Pakes and Ericson (1999) to identifying the presence of Bayesian learning about time-invariant productivity parameter.

If the offer is accepted, the worker decides whether to stay at the firm sufficiently long that her performance is:

$$s_i = 1\left[F_{sr}\left(X_{ir}^{s}\right) + \varepsilon_{ir}^{s} > 0\right] \tag{3}$$

If that is the case, her performance is observed:

$$y_i = F_{yr}\left(X_{ir}^{y}\right) + \varepsilon_{ir}^{y} \tag{4}$$

Formal discussion of identification conditions can be found in Maddala (1983). As usual in choice models, it is impossible to identify scale and location parameters, so the standard normalization for offer, acceptance, and stay applies: $\varepsilon_{ir}^{k} \sim N(0,1)$ for $k = a,\ o,\ s$. To achieve nonparametric identification, there must be at least one variable that affects offer but not the subsequent outcomes, another variable that affects acceptance but not stay and performance, and yet another that affects stay but not performance. In other words, X_{ir}^{y} is a strict subset of X_{ir}^{s}, which is a strict subset of X_{ir}^{a}, which is a strict subset of X_{ir}^{o}. We discuss our exclusion restrictions at the beginning of the next section.

Below, we present the estimation method in the context of the model of offer, acceptance, stay and performance. However, as a first step in our empirical analysis, we lump together decisions (1) and (2) into a joint entry decision:

$$e_i = 1\left[F_{er}\left(X_{ir}^{e}\right) + \varepsilon_{ir}^{e}\right] \tag{5}$$

Eventually, we test the restriction and reject it.

5.2 Estimation Method

To estimate the sequential choice model, we use simulated maximum likelihood (SML) based on the Geweke-Hajivassiliou-Keane smooth recursive conditioning simulator. To save computing time, we generate Halton draws for the SML. Our motivation for doing so is that, as discussed in Train (2003), Halton draws provide the same accuracy with fewer draws. Let Λ

be a set that contains the parameters of the model. Following Maddala (1983), the conditional likelihood for individual i with observed characteristics X_i is:

$$l_i\left(\Lambda|X_i, o_i, a_i, s_i, y_i\right) = P\left(o_i|X_{ir}^o\right) P\left(a_i|o_i, X_{ir}^a\right) P\left(s_i|a_i, o_i, X_{ir}^s\right) \left[\varphi_{y|o,a,s}\left(\frac{y_i - F_{yr}\left(X_{ir}^y\right)}{\sigma}\right)\right]^{o_i a_i s_i}$$

The expressions for each of the first three terms are as follows:

$$\begin{aligned} P\left(o_i|X_{ir}^o\right) &= \Phi_o\left(F_{or}\left(X_{ir}^o\right)\right)^{o_i}\left(1-\Phi_o\left(F_{or}\left(X_{ir}^o\right)\right)\right)^{1-o_i} \\ P\left(a_i|o_i, X_{ir}^a\right) &= \left[\Phi_{a|o}\left(F_{ar}\left(X_{ir}^a\right)\right)^{a_i}\left(1-\Phi_{a|o}\left(F_{ar}\left(X_{ir}^a\right)\right)\right)^{1-a_{ir}}\right]^{o_i} \\ P\left(s_i|a_i, o_i, X_{ir}^s\right) &= \left[\Phi_{s|o,a}\left(F_{sr}\left(X_{ir}^s\right)\right)^{s_i}\left(1-\Phi_{s|o,a}\left(F_{sr}\left(X_{ir}^s\right)\right)\right)^{1-s_{ir}}\right]^{o_i a_i} \end{aligned}$$

where $\Phi_o\left(F_{or}\left(X_{ir}^o\right)\right)$ is the normal CDF , $\Phi_{a|o}\left(F_{ar}\left(X_{ir}^a\right)\right)$ is the normal CDF conditional on $o_i = 1$, and $\Phi_{s|o,a}\left(F_{sr}\left(X_{ir}^s\right)\right)$ is the normal CDF conditional on $o_i, a_i = 1$. The final piece is the conditional density of the disturbance term for the performance signal. Combining the contributions of all candidates, we obtain the conditional likelihood:

$$l\left(\Lambda|X\right) = \prod_{i=1}^{n} l_i\left(\Lambda|X_i\right)$$

where X is a collection of the vectors of characteristics, X_i, of all candidates.

6 Results

The notion that referrals transmit information to employers and job candidates that is not otherwise available to them implies differences in the information sets of referred, non-referred candidates, and the potential employer. The actually communicated signal is not observed by the econometrician and may even remain unknown to some of the market participants. Still, as discussed in Section 3, the transmission of informative signals through referrals induces selection on unobservables that would not be present otherwise. This section focuses on how our empirical results relate to this central hypothesis about the informational content of referrals.

We start by showing that entry and stay decisions are affected by sorting on dimensions that remain unobservable to the econometrician. This sorting on unobservables is much stronger for the referred than for the non-referred candidates and employees. Moreover, the unobserved component positively influences not only entry and stay but also performance. Next, we investigate whether referrals convey the same information to both the employer and the candidates. The results reject this hypothesis. Instead, it turns out that only the employer benefits from the transmission of information when making job offers. As robustness checks, we conclude with a discussion on early promotions and on the relationship between the unobserved heterogeneity and observable measures of referral quality.

6.1 Specifications

We consider three stay horizons: three months, six months and one year. The hazard rate of quitting levels off by the sixth month of employment, suggesting that any transitional dynamics associated with learning ends by then. For this reason, the six month tenure horizon is our benchmark case. If employees stay for at least six months, we observe their average past performance and study how it relates to the unobserved heterogeneity. Another dependent variable is a binary indicator for early promotions that equals one if a candidate's starting job title is different from the one associated with the standard entry-level position. The management makes these early promotions by the time a newly hired employee finishes the training program. Since subsequent promotions are tightly linked to meeting certain performance standards and compensation largely depends on job titles, we prefer to focus on the analysis of performance itself rather than its derivatives. To address possible concerns about favoritism, we exclude from the sample referrals made by someone from the management, which account for only about 8 percent of all referrals.

We also estimate the effect of both observable and unobservable characteristics separately for referred and non-referred candidates. The explanatory variables include years of past work experience and of past call center experience, educational attainment, age, race, gender, and distance between work and home in kilometers. In addition, we try to control extensively for differences in the local labor markets by including county and zip code level median income, shares of men and women below 25 in the labor force, shares of women and men below 25 with at least some college education, total labor force, and unemployment rate. Finally, we include the entry-level test score used by the firm to decide on interview invitations and offers, as well as the associated ranking relative to the other job candidates.

6.2 Identification

The identification of the sequential decisions in our model explores the nature and timing of the hiring process. The intuition for the identification of selection on unobservables is that, controlling for their characteristics, individuals who get unexpectedly hired also turn our to have (unexpectedly) high performance. The multistage choice model is identified nonparametrically through exclusion restrictions.

The first exclusion restriction relates to the way the firm determines who is to be interviewed. Soon after the entry test, the employer decides who receives an invitation to an interview. This decision is based on a comparison between the score of a candidate and a historically determined threshold. The technical explanation is that the distribution of scores and the associated rankings for each wave of candidates is aggregated and becomes known to the recruiters only by the time the firm makes offer decisions or, sometimes, even post-hire. For this reason, we include in our model two variables that affect the offer decision but not the subsequent stages of the hiring process: the average test score of recently interviewed individuals in the same location in the past 90 days before the candidate applied and the associated ranking of the candidate relative to the distribution of these past test scores.[3] Undoubtedly one's test score and ranking relative to the other current candidates affect decisions during the hiring process and we include them as controls wherever necessary. Still, we believe that, once we control for their impact on the offer decision, the test scores of past interviewed candidates

[3] In practical terms, we explore several time horizons: interviewed candidates in the preceding 30, 60, and 90 days before a specific individual applies for a job.

should have no residual effect on subsequent outcomes.[4]

The second exclusion restriction is based on the time between the date of completing the test and the date on which the offer is generated. This waiting period depends on the day of the hiring cycle on which a given candidate takes the test, and the schedule of the hiring cycle is not public. Since the average waiting time is around two weeks, the time to offer likely has an impact on extending and accepting a job offer through the likelihood that a candidate is still looking for a job, but not on the prospects of remaining employed in the firm in the long run.

To identify performance from the hiring and stay decisions, we explore the fact that the test score is used internally by the HR teams throughout the hiring process, but it is never communicated to the candidates. Obviously, the test score affects the likelihood of an interview, job offer, and possibly separation decisions in the first months of employment. At the same time, as it remains unknown to candidates and employees, the hiring test score has no residual impact on acceptance and performance once we control for the direct effect on of the variables that are its constituent elements. Finally, we also explore an argument similar to those used by Heckman and Honoré (1989) to achieve identification in the classical Roy model. Specifically, we verify empirically that local labor market conditions affect both the hiring outcomes and the stay decisions, but not performance. This fact suggests that these control variables capture in a reduced-form the alternative employment possibilities available to a particular individual without an indirect effect on performance.

[4] In our preliminary work, we did not find that the characteristics of the other employees in the company have any effect on the probability that a particular individual receives an offer, accepts it, stays long, and performs well. Consequenty, we did not include them as controls. These findings conform to what we know about the operations of the company and the industry. Similarly to the rest of the industry, the biggest concern of the company is filling in its positions and reducing turnover. As a result, the entry-level questionnaire and the associated score have been primarily motivated by the need of the company to evaluate how likely a particular candidate is to remain employed for a long time. For these reasons, the firm also does not engage in relative performance evaluation in order to put pressure on low performers to quit.

6.3 Entry and Stay

Our first objective is to test whether the unobservable factors that make it more likely that a referred candidate becomes employed also make it more likely that she stays employed. We consider three different tenure horizons for the stay decision, three, six and twelve months, in order to investigate whether the impact of the unobservables at entry dissipates over time. If the effect does not diminish over time, then one may conclude that there are persistent difference between candidates on dimensions that remain unobserved to the econometrician. The differences between the candidates on these dimensions may also remain partially observable by potential employers and even the candidates themselves. The hypothesis that referrals provide information on these difficult to observe dimensions implies that the dependence between entry and stay decisions is stronger for the referred than for the non-referred candidates.

Table 4 presents the estimates of our choice model. We discuss the results for the referred candidates first and then contrast them with those for the non-referred. Controlling for their observable characteristics, referred individuals who enter the firm are also more likely to stay at any tenure horizon. We find a correlation of 0.59 between the errors in the entry decision and the decision to remain employed in the firm for more than three months. This correlation increases to 0.72 as we consider the likelihood of remaining employed for more than six months and more than twelve months. All estimates are significantly different from zero at the one percent significance level.

In contrast, the dependence between entry and stay for non-referred candidates is weak. In the specifications for the three and six month tenure horizon, the correlations between the errors in the entry and stay decisions are not statistically significant, while the point estimates are only half the magnitude of their counterparts for the referred candidates. We only recover a positive and significant, but relatively small, dependence between the random components in the entry and stay decisions at the one year tenure horizon. After performing Welch's unequal variance t-test, we verify that the estimated correlation coefficients for the referred subsample are significantly different from their counterparts for the non-referred sample at the one percent significance level.

With respect to the observable characteristics, we find that having more general work experience increases the chances that a candidate, referred or non-referred, enters the firm and stays employed for a long time. Not surprisingly, all candidates who previously worked at call centers are more likely to become employed. However, once hired, non-referred individuals with prior call center experience are also more likely to quit sooner rather than later. In contrast, referred employees with prior call center experience are more, not less, likely to stay with the firm in the long run. These findings highlight the need for interactions between the referral indicator and observable individual characteristics, since the same observable characteristics may have a very different interpretation and effect depending on whether a candidate is referred or not. Applicants with long prior experience in the industry are desirable to all potential call centers, so attracting and retaining them is difficult. However, someone may have many years of experience as a call operator because he is really good at talking on the phone or because he is not employable in any other industry. Referrals may help employers find out whether it is one way or the other, i.e. they may convey additional and more precise information about what stands behind an observable public signal.

6.4 Performance

There exist many possible explanations for the statistical dependence between entry and stay. To shed some light on the issue we investigate how the referral signals and the dependence between entry and stay relate to performance. The hypothesis that referrers provide information to the potential employer and the candidates implies that there is a positive correlation between the unobservable component in the average performance during the first six months of employment and the unobserved components in the entry and stay decisions. To investigate this issue, we estimate performance within a model with two selection stages.

Table 5 summarizes our related results. For both referred and non-referred employees, we find that even after controlling for observable characteristics, individuals with high average performance are less likely to have quit in the first six months of employment. Specifically, the correlations between the errors in the stay decision and in the performance equation are highly statistically significant with point estimates of 0.87 and 0.84 for the referred and non-referred employees respectively. These correlation coefficients are statistically different from each other at the five percent significance level according to Welch's unequal variance t-test, but economically they are very similar. Moreover, after controlling for observable characteristics, referred individuals who are more likely to become employed also turn out to have high average performance. The correlation between the associated errors is statistically significant at the five percent level with a point estimate of 0.4. In contrast, we do not find such a relationship between performance and the entry decision for the subsample of non-referred employees. As before, the correlation between the errors in the entry and performance equations for the subsample of non-referred candidates is statistically different from its counterpart for the subsample of referred candidates. Turning to the errors in the entry and stay equations, we find a significant positive correlation for both referred and non-referred individuals. Again, the effect is stronger and more statistically significant for the subsample of referred than for the subsample of non-referred individuals.

With respect to the observable characteristics, the results reveal that referred employees with longer work experience have higher chances of staying and substantially better performance than those with little or no experience. Work experience has a similarly positive effect on entry, stay, and performance of non-referred employees, but its impact is of smaller magnitude and less precisely estimated. Interestingly, we find that the positive effect of prior call center experience on entry and stay for referred individuals does not extend also to their performance. This finding suggests that referrals address problems of turnover by communicating specific information about the labor mobility of candidates with prior call center experience. In addition, gender, race, and ethnic background also play some role during the hiring process, but they do not affect stay or performance. The estimates also show that educational attainment and age have a positive effect on entry, stay, and performance for both types of candidates.

Finally, Table 6 relates our estimation approach to specifications used in the preceding literature. It incorporates the restriction that all difference between referred and non-referred individuals can be captured through a referral dummy, which amounts to a restriction that the error processes, as well as the impact of observable characteristics, of the referred and non-referred individuals are the same. The first three columns of the table report our estimates for entry, stay, and performance without controlling for selection on observables and unobservables. The following three columns present the results for these equations, when we explicitly model and control for selection. The patterns in the first three columns are broadly consistent with what is reported by the preceding literature in similar settings: the referral dummy has a positive and statistically significant effect on entry, stay, but not on performance. However, the results for the jointly-estimated equations strongly reject the hypothesis that there is no selection on unobservables. In contrast to previous findings, they show that controlling for selection on unobservables referrals are associated with a strong positive effect on performance. Thus, we provide evidence in support of the hypothesis of Pallais and Sands (2016) that if firms use referral signals when hiring, referred and non-referred workers may not differ in their productivity even though referred candidates may have superior productivity.

6.5 Offer and Acceptance

Until now we have abstracted away from distinguishing between the various stages of the hiring process and focused on the relationship between the recruitment outcome, stay, and performance. In what follows, we relax this restriction and investigate how referrals contribute to the likelihood of making and accepting job offers separately. We consider that an offer is accepted if the candidate starts working at the company.

Table 7 presents the estimates of our model of offer, acceptance, stay, and average performance in the first six months of employment. For both referred and non-referred employees, we find a very strong dependence between the unobservable components in the stay and performance equations. As before, the correlations between the errors in the two equations are highly statistically significant with point estimates of 0.84 and 0.87 for the referred and non-referred employees, respectively. Our main finding is that referrals appear to provide informative signals at the job offer stage but not at the job acceptance stage. Namely, controlling for observable characteristics, referred individuals who are more likely to receive an offer also turn out to have high average performance. The correlation between the associated errors is significant at the five percent level with a point estimate of 0.35. In contrast, the results do not reveal a similar relationship for the subsample of non-referred employees. A formal unequal variance t-test rejects the hypothesis that the correlation coefficient for the subsample of referred individuals is the same as the correlation coefficient for the subsample of non-referred individuals. Similarly, referred applicants who receive a job offer are also less likely to quit, but we do not find such a relation for the non-referred candidates. Again, the Welch's unequal variance t-test rejects the null hypothesis that the correlation coefficients between the errors in the offer and stay equations for the subsamples of referred and non-referred candidates are the same. For all types of applicants, the unobserved component in the acceptance decision appears unrelated to stay and performance. Interestingly, the correlation between the errors in the job offer and acceptance equations is positive for all candidates but only borderline statistically significant only for the non-referred.

With respect to the observable characteristics, the estimates reveal that both referred and non-referred applicants with more general and specific work experience are more likely to receive a job offer, but only the referred are more likely to accept it. As before, race and ethnic background play some role during the hiring process, but they do not affect stay or performance. The results also show that educational attainment has a positive effect on job offer, stay, and performance but not on acceptances. The time between the submission of an application and the offer decision has a statistically significant but economically small negative effect on the probability of receiving and accepting an offer. Finally, an increase in quality of recent job candidates has a significant but small negative effect on the probability of receiving a job offer.

6.6 Referrals and Search Efficiency

The preceding results are consistent with the interpretation of referrals as signals within a model of learning about match quality. As the value of the match is eventually revealed, both referred and non-referred employees select strongly along the unobserved dimensions into long run employment. However, the referred individuals have the advantage of completing much of the sorting during the hiring process and the early stages of employment. Consequently, our results allow us to quantify the contribution of referrals to search efficiency. In particular, we distinguish between their contributions to more informed job offers and to more informed job acceptances. As a result, we can also identify whether before the hiring process starts the pool of referred candidates has superior observable and unobservable characteristics than the pool of non-referred candidates.

Figure 1 presents by referral status the quality mix of job candidates and of workers who remain employed for more than six months. Specifically, it plots the kernel densities associated with performance of referred and non-referred individuals, respectively. The figure reveals that referred applicants are superior to non-referred applicants in terms of their performance in the sense of first order stochastic dominance. However, by the sixth month of employment, the differences between the two distributions virtually disappear. Referred stayers appear to be slightly better on average but the difference is much smaller than the one in the pools of candidates before the hiring starts. Figure 2 shows that unobserved heterogeneity accounts for large part of the variation in performance and also explains most of the positive selection into employment. In combination, these results indicate that due to strong positive selection into employment long term employees have similar performance, independent of the recruitment channel through which they come.

Next, we investigate the evolution of the unobserved heterogeneity in performance through the different stages of the hiring process. Figure 3 shows how referred candidates and workers positively sort into employment on the basis of performance. It plots the distributions of performance before recruitment, after offers, after acceptances, and after six months of employment. The results reveal that the firm tends to make offers to referred applicants who eventually turn out to have high performance. Figure 5 reveals that much of this selection is again driven by unobserved heterogeneity. We observe that the firm uses its referral signals to make offers to candidates with high unobserved component of performance leading to a shift to the right in the associated distribution. In contrast, there exists no evidence for selection on unobservables at the acceptance stage. The figure also shows that despite the contribution of referrals, a substantial part of the sorting on unobservables still takes place on the job in the form of turnover. In contrast, Figure 4 and Figure 6 reveal that, for the subsample of non-referred candidates, both individuals and the firm are not influenced in their choices at the recruitment stage by unobserved heterogeneity. Thus, they appear to be symmetrically ignorant of any specific or difficult to observe aspects of their potential match. Moreover, it also appears that without referrals the employer's ability to make inference about one's performance from observable characteristics is also limited.

To summarize, our results show that referral signals are informative for both stay and performance. Controlling for observable characteristics, we find that referred candidates who receive a job offer are also more likely to stay and perform well. However, we do not observe a similar relationship between employment outcomes and job acceptance. These findings suggest that the informational content of referrals differs between candidates and potential employers. The low correlation between the errors in the job offer and acceptance equations lends further support to this interpretation of our results.

6.7 Alternative Explanations and Robustness Checks

The preceding results lend support to the hypothesis that referrals provide informative signals to employers about difficult to observe dimensions of the employment relationship. The positive association between offer, stay and performance of referred candidates casts doubt on the alternative hypothesis that referrals simply stand for a preference for well-connected 'insiders.' If that were the case, we should have found a positive association between offer, acceptance, and stay but no positive relation between entry and performance, or stay and performance. In addition, our results do not provide evidence for adverse selection into employment, i.e. referred and non-referred candidates of low unobserved performance do not self-select into employment.

An interesting alternative explanation is based on the hypothesis that, while all candidates are ex ante identical, the referrers provide help and training on the job to their newly-hired friends. This preferential treatment eventually leads to the observed differences in performance and stay. However, such a story does not appear likely, since referred candidates do not select into employment on unobservables when they decide to accept a job offer. Indeed, we find that the firm uses informative referral signals to make job offers to candidates of high unobserved performance, while the candidates in question appear oblivious to the underlying motivations. The only case in which the hypothesis may be consistent with the estimated error structure is when the referred individuals systematically do not know that their friends will help them, which is hard to believe.

Yet another alternative explanation is based on moral hazard. Its underlying hypothesis is that the firm uses the referral link as a disciplining device to provide implicit incentives to exert effort. Unfortunately, this incentive mechanism does not generate unequivocal predictions for the statistical properties of the errors. To strengthen the case that referrals transmit information about the quality of the potential employment relation, we consider next how referrals affect early promotions.

6.7.1 Early Promotions

Compensation in the firm is tightly linked to job titles, while promotions depend heavily on workers passing certain performance thresholds.[5] However, at the beginning of the employment relationship, the management may make the decision to assign a newly-hired worker to a more senior job title than the one associated with the entry-level position. Promotions at that stage are most likely based on the existing characteristics of an employee rather than expectations about future help from the referring friend or considerations about moral hazard. These decisions are usually made during the hiring process or during the training program for new arrivals. For sure, they are finalized by the time employees actually start their work duties. We limit the room for favoritism of friends and relatives by focusing on referrals by regular employees who are not involved in any way in promotion decisions.

[5] While bonuses exist, they account on average for just five to ten percent of total compensation.

Table 8 summarizes the related estimates. We find that conditional on their observable characteristics, referred candidates with an offer are more likely to also get an early promotion. The estimated correlation between the associated errors is positive and significant at the five percent level. In contrast, such a relationship does not exist for the non-referred candidates. Interestingly, referred individuals who accept an offer are less likely to get an early promotion, while the corresponding relation for non-referred individuals is positive but not statistically significant. The correlations between offer and acceptance are similar to the ones reported before. With respect to the observable characteristics, all candidates with more general and specific experience are more likely to receive an offer and an early promotion. Low educational attainment decreases both the probability of job offer and early promotion, but not the probability of acceptance.

These results conform with findings in the existing literature that document how referred employees start with higher wages than their non-referred counterparts. However, we highlight a specific mechanism that leads to the observed heterogeneity in compensation: while the firm's policy is to pay the same to equally situated workers, it is more likely to hire qualified referred individuals at hierarchical levels that are higher than the one associated with the entry level position. Moreover, we find that at least partially such early promotions are driven by superior idiosyncratic characteristics rather than simply the threat that a competitor may poach a referred employee with attractive observable characteristics.

6.7.2 Referral Quality

The results reported in Tables 4-7 remain agnostic about the nature and origins of the referral signals. It may be that the firm learns about the candidates from the people who refer them as suggested by Montgomery (1991). Alternatively, it may be that conditional on the available information about its employees, referrals provide additional information about the applicants. In this context, it is important both how well the referrer knows the firm and how well she knows the candidate. Fortunately, our dataset contains very detailed information on both dimensions. Thus, we can evaluate how the estimates of the model for the referred candidates change as we include these controls for the quality of the referral relations. Moreover, the re-estimation of the model with these observable controls constitutes a test in the tradition of Altonji, Elder, and Taber (2005) for selection on unobservables that may be correlated with the already included observable characteristics. Table 9 reports the associated results for the model of offer, acceptance, stay, and performance.

We introduce as controls job title and job tenure, which should be positively correlated with the quality of the employment relationship between the firm and the referrer. Tenure may also reflect the precision of the information about the firm that the referrer can communicate to the candidate. To control for the quality of the relationship between the candidate and the referrer, we include in our specification the time the referrer and the candidate have known each other and the context in which they formed their acquaintance. The estimates indicate that individuals referred by medium and senior level call operators are more likely to receive an offer, stay for more than six months, and perform well, but these hierarchy categories appear to have no impact on job acceptance. In contrast, job tenure of referrers has no significant effect on any of the hiring decisions, stay, or performance. With respect to the relationship between the referrer and the candidate, we find that referrals made by employees who have known the applicant for more than five years tend to be associated with superior chances of both receiving a job offer and accepting it. Interestingly, these positive and significant effects do not extend to stay and performance. Furthermore, referrals by a former coworker are positively related to offer, acceptance, and stay, but not performance. We do not find that close family links have any impact on recruitment and employment outcomes. In this sense, our results corroborate the hypothesis that referral quality depends positively on the duration and relevance of the referral network

Last but not least, we do not find that the introduction of the controls for referral quality significantly alters our estimates of the correlation structure between the errors in the offer, acceptance, stay, and performance equations. Similarly, there are no substantial changes in the estimated coefficients of the other observable variables. Thus, referral quality seems to play a significant role for some employment outcomes, but it also appears to be largely orthogonal to observable and unobservable characteristics. This finding lends support to the intuition that the referral network that we study is not formed explicitly to secure employment at the call center company. We find very similar results for the re-estimated model of early promotions, reported in Table 10.

7 Conclusion

In this paper, we investigate how referral signals about difficult to observe dimensions of match quality induce statistical dependence between job offers, acceptances, early performance and stay decisions. Apart from the obvious econometric implications, the resulting selection on unobservables is crucial to quantifying the contribution of referral signals to search efficiency at each of the various stages of the hiring process. We find that the distribution of unobserved heterogeneity in performance of referred candidates statistically dominates the distribution of unobserved heterogeneity of non-referred candidates. However, due to selective hiring and early turnover the differences between the distributions of the referred and non-referred virtually disappear by the sixth month of employment. The key role of referrals in the labor market is, therefore, the provision of informative signals that improve search efficiency. In particular, we find that the referral process provides informative signals to employers making job offers but not to referral candidates deciding whether to accept or reject such offers. Of course, it is possible that the referral process leads to pre-selection of candidates who apply. The point, however, is that referred employees complete much of the sorting on unobservables during the hiring process. In contrast, non-referred individuals start their employment relationship with the firm considering only observable characteristics. In terms of external validity, the descriptive analysis and the estimates reported in Table 6 show that the patterns in our data are similar to those reported in the recent literature. Moreover, our framework for the identification of the informational content of referrals can be applied to any of the recently investigated referral settings.

If we consider non-referred workers as some sort of socially disadvantaged group, then our model can be interpreted as a model of statistical discrimination. Thus, we believe that our methodology can be extended to the analysis of other environments of information uncertainty, such as employer learning, communication, and statistical discrimination. In contexts such as these, market participants have to learn on-the-job about their match prospects, and typically neither share nor access the same information signals. A number of related topics are left for future research. For example, referral signals may induce selection on unobservables in social networks when the matchmakers are not current employees. The distinction between internal (current employee) and external (non-employee) referrers raises further questions about the incentives that underpin individual behavior. Most internal referral systems, like the one investigated in this paper, rely on explicit incentives in the form of fixed payments whenever the referral remains employed for a certain duration with the company. Although such an incentive structure suggests that employees might recommend as many candidates as possible, the empirical evidence, including our finding of a strong informational content of referring, points to a more nuanced set of facts. The distribution of referrals across employees seem to suggests the existence of implicit (reputational or other) costs associated with a strategy of maximizing the number of referrals. An analysis of the quality and quantity of referrals between different referrers is likely to provide further insight about true match-makers in the labor market.

References

[1] Abraham, Katherine and Henry Farber. 1987. Job duration, seniority, and earnings. *American Economic Review* **77** (3): 278 - 97.

[2] Altonji, Joseph, Todd Elder , and Christopher Taber. 2005. Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. *Journal of Political Economy* **113**: 151-184.

[3] Altonji, Joseph, and Charles Pierret. 2001. Employer learning and statistical discrimination. *Quarterly Journal of Economics* **116**: 313-350

[4] Bayer, Patrick, Stephen Ross, and Giorgio Topa. 2008. Place of work and place of residence: informal hiring networks and labor market outcomes. *Journal of Political Economy* **116 (6)**: 1150-1196.

[5] Brown, Meta, Elizabeth Setren, and Giorgio Topa. 2016. Do informal referrals lead to better matches? evidence from a firm's employee referral system. *Journal of Labor Economics*, forthcoming

[6] Burks, Stephen. Bo Cowgill, Mitchell Hoffman, and Michael Hausman. 2015. The value of hiring through employee referrals. *Quarterly Journal of Economics*, **130(2)**: 805-839

[7] Calvo-Armengol, Antoni, Matthew O. Jackson. 2004. The effects of social networks on employment and inequality. *American Economic Review* **94(3)**: 426-454.

[8] Castilla, Emilio. 2005. Social networks and employee performance in a call center. *American Journal of Sociology* **110 (5)**: 1243-1283.

[9] Dustmann, Christian, Albrecht Glitz, Uta Schoenberg, and Herbert Brücker. 2016. Referral-based job search networks. *Review of Economic Studies* **83 (2): 514-546.**

[10] Fernandez, Roberto, and Emilio Castilla. 2001. How much is that network worth? Social capital returns for referring prospective hires. In *Social Capital: Theory Research*, ed. K. Cook, N. Lin, and R. Burt. Hawthorne, N.Y.: Aldine De Gruyter.

[11] Fernandez, Roberto, Emilio Castilla, and Paul Moore. 2000. Social capital at work: networks and employment at a phone center. *American Journal of Sociology* **105 (5)**: 1288–1356.

[12] Fernandez, Roberto, and Nancy Weinberg. 1997. Sifting and sorting: personal contacts and hiring in a retail bank. *American Sociological Review* **62**: 883-902

[13] Galenianos, Manolis. 2013. Learning about match quality and the use of referrals. *Review of Economic Dynamics* **16:** 1239-1269.

[14] Galenianos, Manolis. 2014. Hiring through referrals. *Journal of Economic Theory* **152**: 304-323.

[15] Gee, Laura, Moira Burke, and Jason Jones. 2016. Social networks and labor markets: how strong ties relate to job transmission using facebook's social network. *Journal of Labor Economics*, forthcoming.

[16] Heath, Rachel. 2011. Why do firms hire using referrals? Evidence from Bangladeshi garment factories. Manuscript, University of Washington.

[17] Hensvik, Lena, and Oskar Nordström Skans. 2016. Social networks, employee selection and labor market outcomes. *Journal of Labor Economics* **34(4)**.

[18] Heckman, James, and Bo Honoré. 1990. The empirical content of the Roy model. *Econometrica* **58(5)**: 1121-1149.

[19] Jovanovic, Boyan. 1984. Matching, turnover, and unemployment. *Journal of Political Economy* **92**: 108-22.

[20] Keane, Michael. 1994. A computationally practical simulation estimator for panel data. *Econometrica* **62**: 95–116.

[21] Kramarz, Francis, and Oskar Nordström Skans. 2014. When strong ties are strong: networks and youth labour market entry. *Review of Economic Studies* **81(3)**: 1164-1200.

[22] Maddala, Gangadharrao. 1983. *Limited-Dependent and Qualitative Variables in Economics*. Cambridge University Press.

[23] Montgomery, James. 1991. Social networks and labor market outcomes: Toward an Economic Analysis. *American Economic Review* **81(5)**:1407-18.

[24] Munasinghe, Lalith. 2006. Expectations matter: job prospects and turnover dynamics. *Labour Economics* **13(5):** 589-609.

[25] Oyer, Paul, and Scott Schaefer. 2012. Firm/Employee matching: an industry study of American lawyers. NBER Working Paper 18620.

[26] Pakes, Ariel, Richard Ericson. 1999. Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory* **79**: 1-45.

[27] Pallais Amanda, and Emily Sands. (2016). Why the referential treatment? Evidence from field experiments on referrals. *Journal of Political Economy*. Forthcoming.

[28] Topa, Giorgio. 2011. Labor markets and referrals. *Handbook of Social Economics* **1**: 1193-1221.

[29] Simon, Curtis, and John Warner. 1992. Matchmaker, matchmaker: the effect of old boy networks on job match quality, earnings, and tenure. *Journal of Labor Economics* **10(3)**: 306-330

[30] Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*, Cambridge University Press.

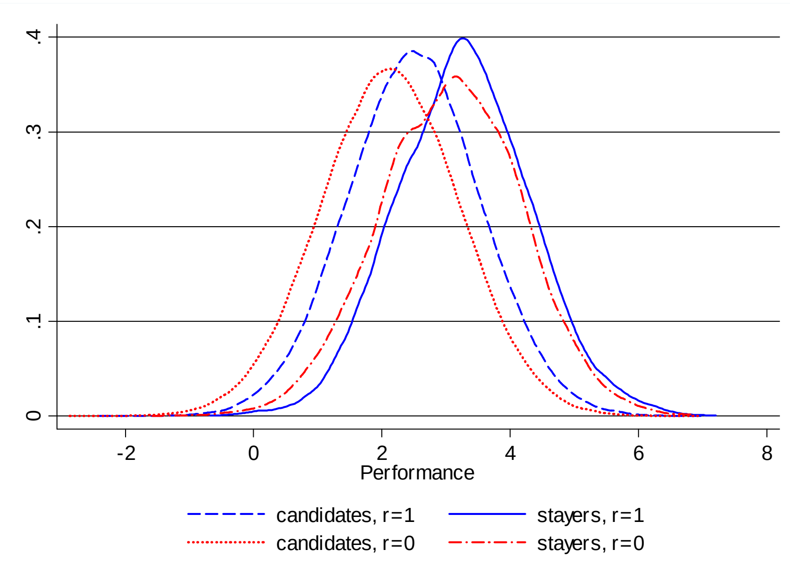

Figure 1: Distributions of performance of candidates and workers who remain employed for more than six months: by referral status. The figure plots the kernel densities associated with performance of referred and non-referred individuals, $r = 0, 1$.

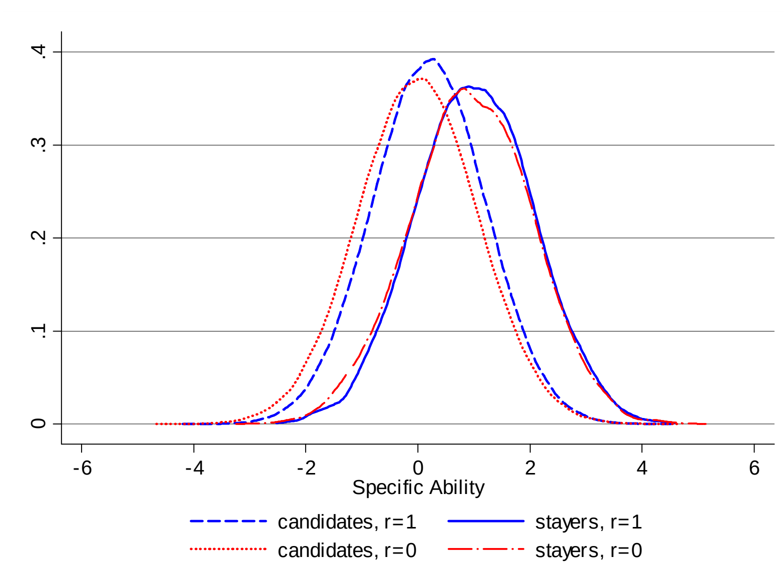

Figure 2: Distributions of unobserved heterogeneity in performance, denoted specific ability, of candidates and workers who remain employed for more than six months: by referral status. The figure plots the kernel densities associated with the specific ability of referred and non-referred individuals, $r = 0, 1$.

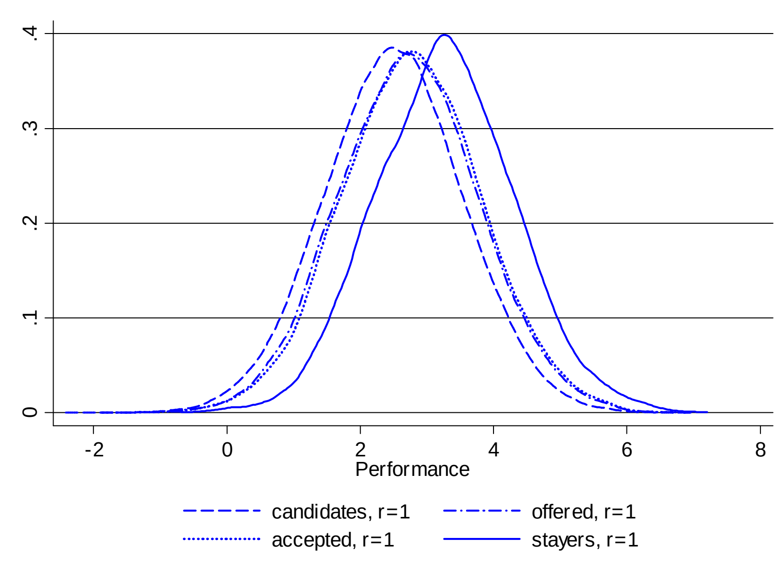

Figure 3: Sorting of referred candidates and workers on performance. The figure plots the kernel densities associated with performance of referred individuals before recruitment, after offers, after acceptances, and after six months of employment.

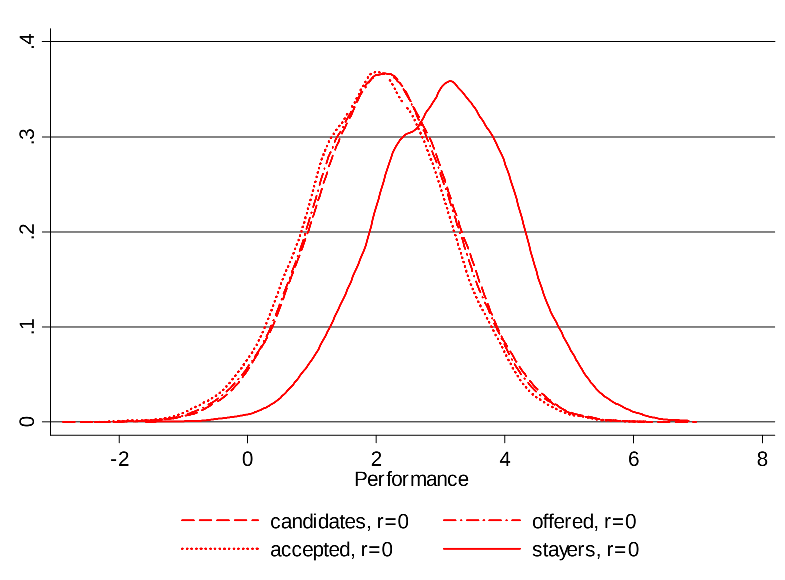

Figure 4: Sorting of non-referred candidates and workers on performance. The figure plots the kernel densities associated with performance of non-referred individuals before recruitment, after offers, after acceptances, and after six months of employment.

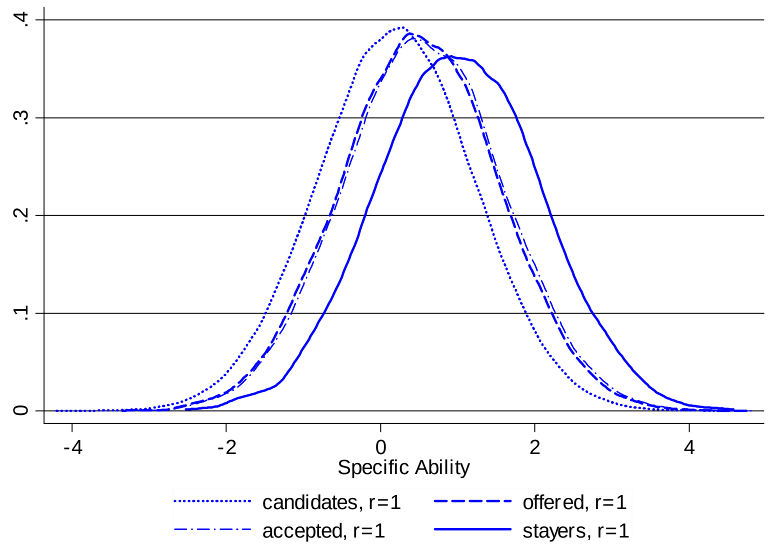

Figure 5: Sorting of referred candidates and workers on unobserved heterogeneity in performance, denoted specific ability. The figure plots the kernel densities associated with the specific ability of referred individuals before recruitment, after offers, after acceptances, and after six months of employment.

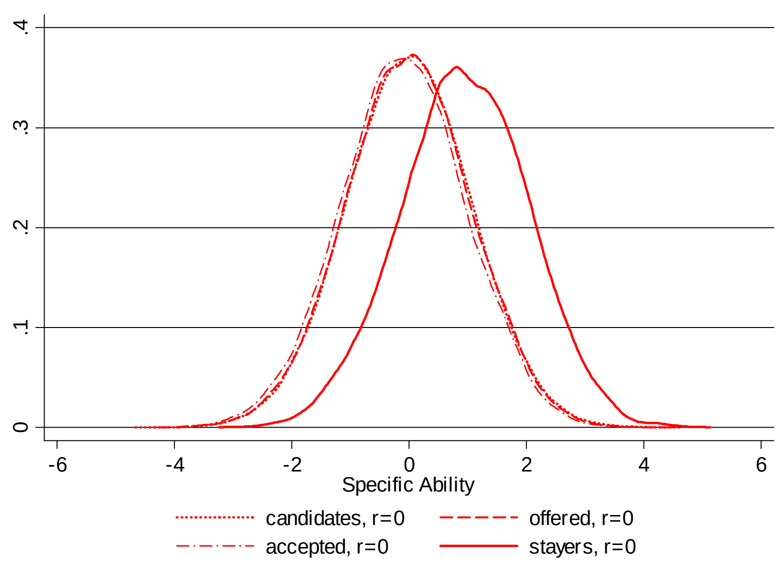

Figure 6: Sorting of non-referred candidates and workers on unobserved heterogeneity in performance, denoted specific ability. The figure plots the kernel densities associated with the specific ability of non-referred individuals before recruitment, after offers, after acceptances, and after six months of employment.

Table 1: Hiring and employment outcomes of referred and non-referred candidates.

Entry	r=1	r=0
Interviewed	0.562	0.298
	(0.50)	(0.46)
Offered	0.319	0.136
	(0.47)	(0.34)
Accepted	0.253	0.103
	(0.43)	(0.30)
Started	0.196	0.077
	(0.40)	(0.27)
Observations.	29837	115893
Stay	**r=1**	**r=0**
Passed Training	0.887	0.879
	(0.32)	(0.33)
Stay $>$ 1 mo.	0.795	0.793
	(0.40)	(0.41)
Stay $>$ 2 mo.	0.674	0.661
	(0.47)	(0.47)
Stay $>$ 3 mo.	0.574	0.572
	(0.49)	(0.49)
Stay $>$ 6 mo.	0.400	0.389
	(0.49)	(0.49)
Obs.	5844	8906

Note: The table reports the fraction of initial candidates who reach successive stages of the hiring process and proportion of hired workers who stay until various tenure horizons. Standard errors are reported in parentheses.

Table 2: Observable characteristics of referred and non-referred candidates

	Candidates		Offered		Accepted		Stay > 180 days	
Variable	r=1	r=0	r=1	r=0	r=1	r=0	r=1	r=0
HSD only	0.468	0.453	0.388	0.339	0.392	0.346	0.371	0.335
	(0.50)	(0.50)	(0.49)	(0.47)	(0.49)	(0.48)	(0.48)	(0.47)
Distance	25.552	24.959	24.568	22.914	23.925	21.937	24.288	22.009
	(24.38)	(26.03)	(21.30)	(21.52)	(19.81)	(19.91)	(20.40)	(21.04)
Call exp.	0.562	0.592	0.584	0.627	0.570	0.602	0.560	0.583
	(0.50)	(0.49)	(0.49)	(0.48)	(0.50)	(0.49)	(0.50)	(0.49)
Low exp.	0.155	0.152	0.145	0.120	0.150	0.128	0.139	0.129
	(0.36)	(0.36)	(0.35)	(0.32)	(0.36)	(0.33)	(0.35)	(0.33)
Exp. > 5 yr.	0.558	0.559	0.530	0.572	0.515	0.551	0.538	0.550
	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)
Black	0.518	0.563	0.464	0.503	0.498	0.530	0.512	0.543
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Hispanic	0.099	0.082	0.127	0.121	0.121	0.127	0.121	0.137
	(0.30)	(0.27)	(0.33)	(0.33)	(0.33)	(0.33)	(0.33)	(0.34)
Female	0.661	0.700	0.636	0.667	0.648	0.679	0.669	0.682
	(0.47)	(0.46)	(0.48)	(0.47)	(0.48)	(0.47)	(0.47)	(0.47)
Observations	29837	115893	9510	15728	5844	8906	2339	3464

Note: The table contains observable characteristics of referred and non-referred individuals at each stage of the hiring process and among those who stay for at least 6 months after hiring. Standard errors are reported in parentheses.

Table 3: Hiring and employment outcomes, conditional on observable characteristics.

Variable	Ref.	Offered	Accepted	Stay>0.5mo.	Stay>3mo.	Stay>6mo.	Perf.
HSD only =1	r=1	0.27	0.17	0.87	0.56	0.38	2.86
		(0.44)	(0.38)	(0.33)	(0.50)	(0.48)	(0.78)
	r=0	0.10	0.06	0.87	0.56	0.37	2.99
		(0.31)	(0.24)	(0.34)	(0.50)	(0.48)	(0.76)
HSD only =0	r=1	0.37	0.23	0.90	0.58	0.41	3.02
		(0.48)	(0.42)	(0.31)	(0.49)	(0.49)	(0.78)
	r=0	0.17	0.10	0.89	0.57	0.39	3.04
		(0.37)	(0.30)	(0.32)	(0.49)	(0.49)	(0.78)
Low exp. =0	r=1	0.33	0.21	0.89	0.58	0.40	3.02
		(0.47)	(0.40)	(0.32)	(0.49)	(0.49)	(0.78)
	r=0	0.14	0.08	0.88	0.57	0.39	3.02
		(0.35)	(0.28)	(0.33)	(0.50)	(0.49)	(0.78)
Low exp. =1	r=1	0.31	0.20	0.89	0.54	0.37	2.93
		(0.46)	(0.40)	(0.32)	(0.50)	(0.48)	(0.79)
	r=0	0.11	0.07	0.90	0.57	0.39	3.04
		(0.31)	(0.25)	(0.30)	(0.49)	(0.49)	(0.75)
Exp.<5yr.	r=1	0.32	0.21	0.89	0.58	0.40	3.03
		(0.47)	(0.40)	(0.31)	(0.49)	(0.49)	(0.77)
	r=0	0.14	0.08	0.88	0.58	0.39	3.05
		(0.34)	(0.27)	(0.32)	(0.49)	(0.49)	(0.77)
Exp≥5yr.	r=1	0.33	0.20	0.88	0.55	0.40	2.87
		(0.47	(0.40)	(0.33)	(0.50)	(0.49)	(0.82)
	r=0	0.15	0.08	0.86	0.52	0.35	2.88
		(0.36)	(0.28)	(0.35)	(0.50)	(0.48)	(0.77)
Call exp. =0	r=1	0.32	0.20	0.88	0.57	0.39	3.02
		(0.47)	(0.40)	(0.32)	(0.50)	(0.49)	(0.77)
	r=0	0.14	0.08	0.88	0.58	0.39	3.04
		(0.34)	(0.27)	(0.32)	(0.49)	(0.49)	(0.77)
Call exp. =1	r=1	0.38	0.23	0.91	0.63	0.49	2.89
		(0.48)	(0.42)	(0.28)	(0.48)	(0.50)	(0.82)
	r=0	0.16	0.08	0.85	0.52	0.35	2.90
		(0.37)	(0.28)	(0.36)	(0.50)	(0.48)	(0.79)
Distance<4km	r=1	0.32	0.21	0.90	0.60	0.41	2.97
		(0.46)	(0.41)	(0.30	(0.49)	(0.49)	(0.74)
	r=0	0.15	0.09	0.89	0.59	0.41	3.04
		(0.36)	(0.29)	(0.31)	(0.49)	(0.49)	(0.77)
Distance≥4km	r=1	0.33	0.20	0.88	0.56	0.39	3.02
		(0.47)	(0.40)	(0.32)	(0.50)	(0.49)	(0.79)
	r=0	0.13	0.08	0.87	0.56	0.37	3.01
		(0.34)	(0.27)	(0.33)	(0.50)	(0.48)	(0.78)

Standard errors are reported in parentheses.

Table 4: Estimates of Model of Entry and Stay

	Stay>3mo.		Stay>6mo.		Stay>1yr.	
	r=1	r=0	r=1	r=0	r=1	r=0
Stay|Entry	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
Call exp.>0	-0.034	-0.008	-0.038*	-0.015	-0.038	-0.035*
	(0.018)	(0.015)	(0.019)	(0.016)	(0.020)	(0.016)
2yr.<Call exp.<5yr.	0.003	-0.017	0.009	-0.014	-0.003	0.025
	(0.019)	(0.016)	(0.020)	(0.016)	(0.022)	(0.017)
Call exp. >5yr.	0.114**	-0.020	0.139**	-0.010	0.079**	0.018
	(0.028)	(0.023)	(0.028)	(0.023)	(0.029)	(0.024)
2yr.<Exp.<5yr.	0.038	0.024	0.014	-0.001	0.030	0.004
	(0.021)	(0.018)	(0.023)	(0.018)	(0.022)	(0.019)
5yr.<Exp.<10yr.	0.054*	0.039*	0.039	0.023	0.064*	0.012
	(0.024)	(0.019)	(0.025)	(0.028)	(0.026)	(0.021)
10yr.<Exp.<15yr.	0.092**	0.049*	0.075*	0.035	0.092**	0.032
	(0.033)	(0.023)	(0.034)	(0.029)	(0.035)	(0.029)
HSD only	-0.045**	-0.033**	-0.046**	-0.028**	-0.056**	-0.019**
	(0.014)	(0.012)	(0.015)	(0.012)	(0.015)	(0.001)
Entry	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
Call exp.>0	0.006	-0.003	0.007	-0.003	0.003	-0.005*
	(0.007)	(0.002)	(0.008)	(0.003)	(0.007)	(0.002)
2yr.<Call exp.<5yr.	0.012	0.005*	0.012	0.005	0.011	0.005*
	(0.008)	(0.002)	(0.008)	(0.003)	(0.008)	(0.002)
Call exp. >5yr.	0.060**	0.013**	0.060**	0.013**	0.052**	0.010**
	(0.012)	(0.004)	(0.012)	(0.004)	(0.011)	(0.004)
2yr.<Exp.<5yr.	0.018*	0.004	0.018*	0.004	0.010	0.004
	(0.009)	(0.003)	(0.008)	(0.003)	(0.008)	(0.003)
5yr.<Exp.<10yr.	-0.004	0.008**	-0.004	0.008*	-0.010	0.006*
	(0.010)	(0.004)	(0.010)	(0.004)	(0.013)	(0.003)
10yr.<Exp.<15yr.	0.017	0.012**	0.016	0.012**	0.001	0.007
	(0.014)	(0.005)	(0.014)	(0.005)	(0.013)	(0.004)
HSD only	-0.018**	-0.013**	-0.019**	-0.013**	-0.019**	-0.014**
	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)
Time to offer	-0.003*	-0.001*	-0.003**	-0.001*	-0.001*	-0.001*
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Corr (ϵ_e, ϵ_s):	0.587**	0.248	0.719**	0.258	0.722**	0.695**
	(0.173)	(0.200)	(0.145)	(0.197)	(0.200)	(0.142)
Obs.	21693	84829	21693	84829	20804	83370

Note: All specifications include location and month dummies, gender, race, distance from home, age, age^2, the score used by the firm to make offers, the associated score rank relative to other candidates, and local labor market controls on county and zip code level, such as median income, shares of men and women below 25 in labor force, shares of college educated men and women below 25, labor force, and unemployment rate. Standard errors are reported in parentheses. Stars indicate significance level of estimates (*=5%, **=1%). Average marginal effects: dy/dx.

Table 5: Estimates of Model of Entry, Stay and Performance

	Perf. | Stay>6mo.		Stay>6mo. | Entry		Entry	
	r=1	r=0	r=1	r=0	r=1	r=0
	Coef.	Coef.	dy/dx	dy/dx	dy/dx	dy/dx
Call exp.>0	-0.035	-0.003	-0.036*	-0.015	0.007	-0.003
	(0.058)	(0.050)	(0.018)	(0.015)	(0.008)	(0.003)
2yr.<Call exp.<5yr.	0.003	-0.066	0.013	-0.011	0.012	0.005*
	(0.063)	(0.055)	(0.020)	(0.016)	(0.008)	(0.002)
Call exp. >5yr.	0.162	-0.073	0.160**	-0.015	0.061**	0.013**
	(0.092)	(0.077)	(0.031)	(0.023)	(0.012)	(0.004)
2yr.<Exp.<5yr.	0.149*	0.059	0.028	-0.002	0.019*	0.004
	(0.066)	(0.059)	(0.022)	(0.018)	(0.009)	(0.003)
5yr.<Exp.<10yr.	0.175*	0.074	0.044	0.011	-0.003	0.008*
	(0.076)	(0.068)	(0.025)	(0.020)	(0.010)	(0.004)
10yr.<Exp.<15yr.	0.237*	0.106	0.089**	0.029	0.017	0.012**
	(0.103)	(0.094)	(0.033)	(0.028)	(0.014)	(0.005)
Black	-0.084	-0.048	0.020	0.011	0.006	0.015**
	(0.049)	(0.047)	(0.016)	(0.014)	(0.006)	(0.002)
Hispanic	-0.029	0.046	0.040	0.032	0.049**	0.030**
	(0.075)	(0.064)	(0.026)	(0.020)	(0.010)	(0.003)
Female	0.134**	0.043	0.040**	-0.010	0.002	0.005*
	(0.044)	(0.039)	(0.014)	(0.012)	(0.006)	(0.002)
Age	0.012*	0.013*	0.004*	0.005*	-0.011**	-0.007*
	(0.006)	(0.006)	(0.002)	(0.002)	(0.004)	(0.003)
HSD only	-0.064	-0.125**	-0.050**	-0.024*	-0.018**	-0.013**
	(0.047)	(0.041)	(0.016)	(0.012)	(0.006)	(0.002)
Score			0.011*	-0.012*	0.013**	0.083**
			(0.005)	(0.003)	(0.002)	(0.004)
Time to offer					-0.003*	-0.001*
					(0.001)	(0.000)
Interview rank					-0.023	-0.005
					(0.012)	(0.004)
$\overline{\text{Score}}$, past 90d.					-0.003*	-0.001**
					(0.001)	(0.000)
Error Correlations:	r=1			r=0		
	ϵ_y, ϵ_e	ϵ_e, ϵ_s	ϵ_y, ϵ_s	ϵ_y, ϵ_e	ϵ_e, ϵ_s	ϵ_y, ϵ_s
	0.400*	0.650*	0.837**	0.011	0.236*	0.887**
	(0.186)	(0.250)	(0.036)	(0.115)	(0.113)	(0.018)

Note: All specifications also include location and month dummies, distance from home, the score used by the firm to make offers, the associated rank relative to other candidates, age^2 and local labor market controls on county and zip code level, such as median income, shares of men and women below 25 in labor force, shares of college-educated men and women below 25, labor force, and unemployment rate. Standard errors are reported in parentheses. Stars denote significance level of estimates (*=5%, **=1%). Obs., r=1: 21693. Obs., r=0: 84829. Average marginal effects: dy/dx.

Table 6: Benchmark Estimates of Entry, Stay and Performance

	Perf. Coef.	Stay>6mo. dy/dx	Entry dy/dx	Perf.|Stay>6mo. Coef.	Stay>6mo.|Entry dy/dx	Entry dy/dx
Referred	-0.004	0.028^{*}	0.052^{**}	0.102^{*}	0.059^{**}	0.052^{**}
	(0.021)	(0.009)	(0.002)	(0.039)	(0.016)	(0.002)
Call exp.>0	0.013	-0.028^{*}	-0.001	-0.024	-0.029^{*}	-0.001
	(0.028)	(0.011)	(0.003)	(0.038)	(0.012)	(0.003)
2yr.<Call exp.<5yr.	-0.032	-0.001	0.007^{*}	-0.035	0.003	0.007^{*}
	(0.030)	(0.012)	(0.003)	(0.041)	(0.013)	(0.003)
Call exp. >5yr.	-0.061	0.046^{*}	0.022^{**}	0.025	0.055^{*}	0.022^{**}
	(0.043)	(0.017)	(0.004)	(0.058)	(0.018)	(0.004)
2yr.<Exp.<5yr.	0.061^{*}	0.004	0.008^{*}	0.097^{*}	0.012	0.008^{*}
	(0.032)	(0.013)	(0.003)	(0.044)	(0.014)	(0.003)
5yr.<Exp.<10yr.	0.072	0.026	0.007^{*}	0.122^{*}	0.031	0.007^{*}
	(0.037)	(0.015)	(0.003)	(0.050)	(0.016)	(0.003)
10yr.<Exp.<15yr.	0.063	0.050^{*}	0.015^{*}	0.164^{*}	0.061^{*}	0.015^{*}
	(0.051)	(0.021)	(0.005)	(0.069)	(0.022)	(0.005)
Black	-0.092^{**}	0.017	0.014^{**}	-0.050	0.024^{*}	0.014^{**}
	(0.025)	(0.010)	(0.002)	(0.034)	(0.011)	(0.002)
Hispanic	-0.032	0.035^{*}	0.034^{**}	0.014	0.045^{*}	0.034^{**}
	(0.034)	(0.014)	(0.003)	(0.049)	(0.016)	(0.003)
Female	0.062^{*}	0.015	0.004^{*}	0.080^{*}	0.014^{*}	0.004^{*}
	(0.022)	(0.009)	(0.002)	(0.029)	(0.009)	(0.002)
Age	0.012	0.005^{**}	-0.008^{**}	0.013^{*}	0.006^{**}	-0.007^{**}
	(0.010)	(0.001)	(0.001)	(0.03)	(0.002)	(0.002)
HSD only	-0.046^{*}	-0.030^{**}	-0.013^{**}	-0.104^{*}	-0.037^{**}	-0.013^{**}
	(0.022)	(0.009)	(0.002)	(0.030)	(0.010)	(0.002)
Score		0.017^{**}	0.020^{**}		0.019^{**}	0.021^{*}
		(0.003)	(0.001)		(0.002)	(0.001)
Time to offer			-0.002^{*}			-0.002^{*}
			(0.000)			(0.000)
Interview rank			-0.007			-0.007
			(0.004)			(0.004)
$\overline{\text{Score}}$, past 90d.			-0.002^{*}			-0.002^{*}
			(0.001)			(0.001)
Error Correlations:						
				ϵ_y, ϵ_e	ϵ_s, ϵ_e	ϵ_y, ϵ_s
				0.175	0.319^{*}	0.861^{**}
				(0.112)	(0.149)	(0.070)

Note: All specifications also include location and month dummies, distance from home, the score used by the firm to make offers, the associated rank relative to other candidates, and local labor market controls on county and zip code level, such as median income, shares of men and women below 25 in labor force, shares of college-educated men and women below 25, labor force, and unemployment rate. Standard errors are reported in parentheses. Stars indicate significance level of estimates (*=5%, **=1%).

Table 7: Estimates of Model of Offer, Acceptance, Stay and Performance

	Perf.|Stay>6mo.		Stay>6mo.|Accept		Accept|Offer		Offer	
	r=1	r=0	r=1	r=0	r=1	r=0	r=1	r=0
	Coef.	Coef.	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
Call exp.>0	-0.027	0.000	-0.046*	-0.018	-0.025	-0.006	0.022**	-0.003
	(0.059)	(0.050)	(0.023)	(0.016)	(0.014)	(0.012)	(0.008)	(0.003)
2yr<Call exp.<5yr	-0.007	-0.068	0.026	-0.007	0.034*	-0.007	0.001	0.010**
	(0.065)	(0.054)	(0.029)	(0.016)	(0.015)	(0.012)	(0.009)	(0.003)
Call exp. >5yr.	0.145	-0.073	0.173**	-0.008	0.041*	-0.010	0.063**	0.022**
	(0.092)	(0.076)	(0.044)	(0.023)	(0.021)	(0.016)	(0.012)	(0.004)
2yr.<Exp.<5yr.	0.142*	0.058	0.031	0.000	0.007	-0.002	0.020**	0.006
	(0.067)	(0.059)	(0.023)	(0.018)	(0.017)	(0.014)	(0.009)	(0.004)
5yr.<Exp.<10yr.	0.174*	0.081	0.040	0.018	-0.016	-0.032*	0.002	0.019**
	(0.076)	(0.068)	(0.026)	(0.023)	(0.019)	(0.016)	(0.011)	(0.004)
10yr.<Exp.<15yr.	0.224*	0.110	0.088*	0.037	0.002	-0.010	0.016	0.020**
	(0.103)	(0.093)	(0.034)	(0.029)	(0.025)	(0.021)	(0.014)	(0.006)
Black	-0.092	-0.058	0.025	0.011	0.017	0.041**	-0.002	0.013**
	(0.050)	(0.047)	(0.018)	(0.024)	(0.012)	(0.010)	(0.007)	(0.003)
Hispanic	-0.025	0.031	0.038	0.038	-0.007	0.054**	0.074**	0.032**
	(0.075)	(0.065)	(0.026)	(0.033)	(0.017)	(0.014)	(0.010)	(0.004)
Female	0.130**	0.039	0.040**	-0.009	0.005	0.012	-0.003	0.005*
	(0.044)	(0.039)	(0.014)	(0.013)	(0.011)	(0.009)	(0.006)	(0.002)
Age	0.013	0.014*	0.003	0.005	-0.003*	-0.006**	0.000	0.000
	(0.008)	(0.007)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
HSD only	-0.059	-0.119**	-0.054**	-0.025*	-0.011	-0.003	-0.021**	-0.018**
	(0.046)	(0.040)	(0.018)	(0.013)	(0.011)	(0.009)	(0.006)	(0.002)
Score			0.009**	-0.004*			0.022**	0.019**
			(0.004)	(0.002)			(0.002)	(0.001)
Time to offer					-0.001*	-0.001*	-0.004**	-0.002**
					(0.000)	(0.00)	(0.001)	(0.000)
Interview rank							-0.013	0.000
							(0.013)	(0.005)
$\overline{\text{Score}}$, past 90d.							-0.006**	-0.002*
							(0.002)	(0.001)
Error Correlations:	r=1				r=0			
		ϵ_s	ϵ_a	ϵ_o		ϵ_s	ϵ_a	ϵ_o
ϵ_y		0.843**	0.100	0.345*		0.873**	-0.138	-0.010
		(0.034)	(0.268)	(0.155)		(0.021)	(0.235)	(0.105)
ϵ_s			0.133	0.577*			-0.211	0.028
			(1.105)	(0.252)			(0.770)	(0.172)
ϵ_a				0.253				0.261*
				(0.176)				(0.118)

Note: All specifications also include location and month dummies, distance from home, the score used by the firm to make offers, the associated rank relative to other candidates, age^2, county/zip code median income, shares of men and women below 25 in labor force, shares of college-educated men and women below 25, labor force, and unemployment. Standard errors are reported in parentheses. Stars indicate significance level of estimates (*=5%, **=1%). Obs., r=1: 21693. Obs., r=0: 84829. Average marginal effects: dy/dx.

Table 8: Estimates of Model of Offer, Acceptance, and Early Promotion

	Promotion|Accept		Accept|Offer		Offer	
	r=1	r=0	r=1	r=0	r=1	r=0
	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
Call exp.>0	0.032*	0.032**	-0.033*	-0.029**	0.021**	-0.003
	(0.013)	(0.013)	(0.016)	(0.011)	(0.008)	(0.003)
2yr.<Call exp.<5yr.	0.060**	0.067**	0.034*	0.009	0.001	0.010**
	(0.012)	(0.011)	(0.017)	(0.012)	(0.009)	(0.003)
Call exp. >5yr.	0.139**	0.141**	0.052*	0.027	0.063**	0.022**
	(0.014)	(0.013)	(0.024)	(0.016)	(0.012)	(0.004)
2yr.<Exp.<5yr.	-0.014	-0.019	0.008	-0.011	0.021**	0.006
	(0.018)	(0.018)	(0.019)	(0.013)	(0.009)	(0.004)
5yr.<Exp.<10yr.	0.007	0.017	-0.018	-0.035*	0.002	0.019**
	(0.018)	(0.018)	(0.021)	(0.015)	(0.011)	(0.004)
10yr.<Exp.<15yr.	0.035	0.041*	0.019	0.009	0.016	0.020**
	(0.020)	(0.020)	(0.028)	(0.019)	(0.014)	(0.006)
Black	-0.049**	-0.055**	0.015	0.024*	-0.002	0.013**
	(0.009)	(0.008)	(0.013)	(0.009)	(0.007)	(0.003)
Hispanic	-0.010	-0.024	0.007	0.049**	0.075**	0.032**
	(0.016)	(0.014)	(0.021)	(0.014)	(0.010)	(0.004)
Female	-0.017*	-0.019*	0.001	-0.006	-0.002	0.005**
	(0.008)	(0.007)	(0.012)	(0.008)	(0.006)	(0.002)
Age	0.001	0.001	-0.002	-0.003**	0.000	0.000
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)
HSD only	-0.016	-0.019*	-0.020	-0.006	-0.021**	-0.018**
	(0.010)	(0.008)	(0.013)	(0.009)	(0.006)	(0.002)
Score	0.003	0.002*			0.023**	0.019**
	(0.002)	(0.001)			(0.002)	(0.004)
Time to offer			-0.005**	-0.003**	-0.004**	-0.002**
			(0.001)	(0.000)	(0.001)	(0.000)
Interview rank					-0.014	0.000
					(0.013)	(0.005)
$\overline{\text{Score}}$, past 90d.					-0.006**	-0.002*
					(0.002)	(0.001)
Error Correlations:	r=1			r=0		
	ϵ_p, ϵ_o	ϵ_p, ϵ_a	ϵ_o, ϵ_a	ϵ_p, ϵ_o	ϵ_p, ϵ_a	ϵ_o, ϵ_a
	0.622*	-0.431*	0.362*	0.140	0.841	0.325*
	(0.283)	(0.171)	(0.180)	(0.121)	(0.666)	(0.124)

Note: All specifications also include location and month dummies, distance from home, the score used by the firm to make offers, the associated rank relative to other candidates, age^2, county or zip code median income, shares of men and women below 25 in labor force, shares of college-educated men and women below 25, labor force, and unemployment. Std. errors are reported in parentheses. Stars indicate significance level of estimates (*=5%, **=1%). Obs., r=1: 21693. Obs., r=0: 84829. Average marginal effects: dy/dx.

Table 9: Estimates of Model with Referral Quality of Offer, Acceptance, Stay and Performance

	r=1			
	Perf.|Stay>6mo.	Stay>6mo.|Accept	Accept|Offer	Offer
	dy/dx	dy/dx	dy/dx	dy/dx
Call exp.>0	-0.034	-0.036	-0.023	0.020**
	(0.059)	(0.022)	(0.014)	(0.008)
2yr.<Call exp.<5yr.	-0.023	0.003	0.035*	-0.004
	(0.065)	(0.028)	(0.015)	(0.009)
Call exp. >5yr.	0.113	0.133**	0.041*	0.050**
	(0.091)	(0.041)	(0.021)	(0.012)
2yr.<Exp.<5yr.	0.141*	0.026	0.008	0.020**
	(0.066)	(0.023)	(0.017)	(0.009)
5yr.<Exp.<10yr.	0.165*	0.044	-0.014	-0.004
	(0.076)	(0.026)	(0.019)	(0.011)
10yr.<Exp.<15yr.	0.207*	0.081*	0.004	0.008
	(0.103)	(0.034)	(0.025)	(0.014)
Female	0.131**	0.038*	0.005	-0.002
	(0.044)	(0.014)	(0.011)	(0.006)
HSD only	-0.060	-0.049**	-0.010	-0.025**
	(0.047)	(0.018)	(0.011)	(0.006)
Referral by coworker	0.112	0.112*	0.059*	0.043*
	(0.102)	(0.051)	(0.029)	(0.016)
Referred known >5yr.	-0.027	0.054	0.051*	0.064**
	(0.070)	(0.042)	(0.018)	(0.010)
Job title of referrer: low	-0.130	-0.054*	0.023	-0.091**
	(0.072)	(0.025)	(0.016)	(0.010)
Score		0.008*		0.023**
		(0.005)		(0.002)
Time to offer			-0.001*	-0.004**
			(0.000)	(0.001)
$\overline{\text{Score}}$, past 90d.				-0.006**
				(0.002)
Error Correlations:	r=1			
		ϵ_s	ϵ_a	ϵ_o
ϵ_y		0.841**	0.098	0.331*
		(0.035)	(0.264)	(0.158)
ϵ_s			0.156	0.475
			(0.816)	(0.251)
ϵ_a				0.260
				(0.186)

Note: All specifications also include location and month dummies, distance from home, the score used by the firm to make offers, the associated rank relative to other candidates, race, age, age^2,county/zip code median, income, shares of men and women below 25 in labor force, shares of college-educated men and women below 25, labor force, and unemployment. Std. err. are reported in parentheses. Stars denote significance level (*=5%, **=1%). Obs., r=1: 21693. Obs., r=0: 84829. Avg. marg. effects: dy/dx.

Table 10: Estimates of Model with Referral Quality of Offer, Acceptance, and Promotions

	r=1		
	Promotion|Accept	Accept|Offer	Offer
	dy/dx	dy/dx	dy/dx
Call experience>0	0.026*	-0.029	0.020**
	(0.013)	(0.016)	(0.008)
2yr.<Call exp<5yr.	0.047**	0.036*	-0.004
	(0.012)	(0.017)	(0.009)
Call exp0. >5yr.	0.119**	0.052*	0.050**
	(0.014)	(0.024)	(0.012)
2yr.<Exp<5yr.	-0.014	0.009	0.020**
	(0.018)	(0.018)	(0.009)
5yr.<Exp<10yr.	0.000	-0.016	-0.004
	(0.017)	(0.021)	(0.011)
10yr.<Exp<15yr.	0.029	0.022	0.009
	(0.020)	(0.028)	(0.014)
Female	-0.017*	0.000	-0.002
	(0.008)	(0.012)	(0.006)
Only High School	-0.017	-0.019	-0.025**
	(0.010)	(0.013)	(0.006)
Referral by coworker	0.084**	0.099**	0.042**
	(0.017)	(0.032)	(0.016)
Referred known >5yr.	0.011	0.098**	0.064**
	(0.015)	(0.021)	(0.010)
Job title of referrer: low	-0.116**	0.025	-0.091**
	(0.015)	(0.022)	(0.010)
Score	0.003		0.023**
	(0.003)		(0.002)
Time to offer		-0.005**	-0.004**
		(0.001)	(0.001)
$\overline{\text{Score}}$, past 90d.			-0.006**
			(0.002)
Error Correlations:	ϵ_p, ϵ_o	ϵ_p, ϵ_a	ϵ_o, ϵ_a
	0.615*	-0.406*	0.362*
	(0.213)	(0.181)	(0.184)

Note: All specifications also include location and month dummies, distance from home, the score used by the firm to make offers, the associated rank relative to other candidates, race, age, age^2, county/zip code median income, share of men and women below 25 in labor force, share of college-educated men and women below 25, labor force, and unemployment. Std. errors are reported in parentheses. Stars indicate significance level of estimates (*=5%, **=1%). Obs., r=1: 21693. Obs., r=0: 84829. Average marginal effects: dy/dx.

8 Theoretical Appendix

We revisit the theoretical framework of standard models of search by experience and by inspection in the tradition of Jovanovic (1984). Some job candidates have a referral, $r = 1$, while others do not, $r = 0$. Production surplus depends on characteristics x, known to everyone, and firm-worker specific match quality θ which is unknown to the firm and the candidate at the beginning of the hiring process. The assumptions that complete the description of the environment are presented below.

Assumption 1: Match Quality *Match quality θ_r of both referred and non-referred candidates is drawn from $N\left(\mu, \sigma^2\right)$. The firm and the candidates share a common prior which coincides with the distribution of θ. θ is independent from x and its actual realization is revealed after the candidate is hired.*

We start with the assumption that the quality of referred and non-referred candidates is the same in order to highlight the econometric implications of the informational content of referrals. In line with the overwhelming empirical evidence, we eventually relax this assumption and consider the case when the match quality of referred candidates stochastically dominates that of non-referred candidates.

Assumption 2: Signals *The firm and the candidate learn about match quality through Bayesian updating. The firm and the candidates start with the same prior beliefs. Both types of candidates, $r = 0, 1$, go through the same hiring process which generates a signal which may or may not be the same: the candidate receives a signal $\theta_r + \xi_c$, where $\xi_c \sim N\left(0, \sigma^2_{\xi c}\right)$, while the firm receives a signal $\theta_r + \xi_f$, where $\xi_f \sim N\left(0, \sigma^2_{\xi f}\right)$. Referrals provide additional information to the firm and the candidate, which also may or may not be the same: the referred candidate receives a signal $\theta_1 + \zeta_c$, where $\zeta_c \sim N\left(0, \sigma^2_{\zeta c}\right)$, while the firm receives a signal $\theta_1 + \zeta_f$, where $\zeta_f \sim N\left(0, \sigma^2_{\zeta f}\right)$.*

Assumption 2 formalizes the intuitive notion that referrals may provide additional information about the quality of potential employment relations. It also allows for asymmetries of information between the firm and the candidates. Moreover, Assumption 2 capture a distinct aspect of the referral process: in a more general setting all parties may know that referred candidates have superior match quality on average, but they do not necessarily also have more precise information about each specific match. To simplify the exposition and preserve the closed-form formulas of the posterior beliefs, we maintain that the signals are normally distributed and that the participants in the labor market do not learn from the decisions of the other party. This simplifying assumption can be relaxed in a more general setting without affecting the spirit of the main results. The following Assumption 3 relates match quality to observed performance (productivity). We consider an additive technology in match quality, since this specification conforms to the statistical properties of our data. It also happens to be the most commonly used specification in the preceding empirical literature.

Assumption 3: Performance *The production function is defined by $y_r = f(x) + \theta_r + \epsilon$ for $r = 0, 1$, where ϵ and x are independent from match quality and from each other.*

The following Assumption 4 completes the setting by imposing some structure on the outside options of referred and non-referred candidates. As in most empirical settings, we do not have detailed information about the referral network and can practically identify only the net value of employment with the firm. Moreover, our focus is on the ability of referrals to shed light on the specific firm-worker match quality, so we also maintain the standard assumptions in the search literature on utility and profits.

Assumption 4: Utility, Profits, and Outside Option *Profits and individual utility are linear functions of output. The firm and its candidates have an outside option, normalized to* 0.

Finally, the following Assumption 5 summarizes the information available to the econometrician. It is included to highlight how the theoretical framework relates to the following empirical work. Realistically, the econometrician is not privy to all relevant information contained in a referral and, for this reason, we believe that our Assumption 5 captures salient aspects of the empirical environment. In particular, the econometrician usually does not observe the specific signals about match quality that the firm and the referred candidate may receive.

Assumption 5: Observable Information *The econometrician observes individual characteristics x, whether a candidate is referred or not, whether she receives an offer, whether she accepts it if such is extended, then how long she remains employed in the firm and, conditional on sufficiently long tenure, performance y.*

Individuals and the firm update their beliefs about match quality following Bayes rule. For those who come through the regular general pool of applicants, the individual posterior belief after observing the signal from the hiring process is $N\left(\mu_{0c}, \sigma_{0c}^2\right)$, where

$$\sigma_{0c}^2 = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_{c\xi}^2}\right)^{-1} \text{ and } \mu_{0c} = \left(\frac{\mu}{\sigma^2} + \frac{\theta + \xi_c}{\sigma_{\xi c}^2}\right)\sigma_{0c}^2$$

In contrast, referred applicants come with an additional signal and form posterior beliefs about the match quality $\theta_{1c} \sim N\left(\mu_{1c}, \sigma_{1c}^2\right)$, where

$$\sigma_{1c}^2 = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_{\xi c}^2} + \frac{1}{\sigma_{\zeta c}^2}\right)^{-1} \text{ and } \mu_{1c} = \left(\frac{\mu}{\sigma^2} + \frac{\theta_1 + \xi_c}{\sigma_{\xi c}^2} + \frac{\theta_1 + \zeta_c}{\sigma_{\zeta c}^2}\right)\sigma_{1c}^2$$

Similarly, the firm forms posterior beliefs $N\left(\mu_{rf}, \sigma_{rf}^2\right)$, $r = 0, 1$. If they share the same information during the hiring process, the firm and the candidate have the same posterior beliefs, $N\left(\mu_r, \sigma_r^2\right)$. In such a case, the candidate and the firm agree on the value of their potential relation.

Suppose that after all signals are observed the value to the firm of employing a candidate of type r is $v\left(\mu_{rf}, \sigma_{rf}^2, x\right)$. The candidate is hired if $v\left(\mu_{rf}, \sigma_{rf}^2, x\right) > 0$ and not otherwise. Following Jovanovic (1984) and Mortensen (1988), we solve for the threshold posterior mean $\underline{\mu}_{rf}$ that makes the firm indifferent between the two alternatives:

$$v\left(\underline{\mu}_{rf}, \sigma_{rf}^2, x\right) = 0 \;\Rightarrow \underline{\mu}_{rf} = \underline{\mu}_{rf}\left(x\right)$$

where σ_{rf}^2 is absorbed into the functional form of $\underline{\mu}_{rf}(.)$ since the posterior variance does not depend on the specific signals. As the precision of beliefs increases, the option value of employment decreases which pushes up the threshold posterior mean. As a result, when the firm has more precise posterior beliefs for referred than for non-referred candidates, $\sigma_{0f}^2 > \sigma_{1f}^2$, it requires higher posterior mean for entry from the referred than from the non-referred, $\underline{\mu}_{0f}(x) < \underline{\mu}_{1f}(x)$. In a similar way, we obtain the thresholds for the acceptance and stay decisions, $\underline{\mu}_{rc}(x)$ and $\underline{\theta}(x)$. For simplicity, we maintain that candidates are 'shortsighted' in the sense that they do not update their beliefs after receiving an offer, so that they make their acceptance decision based only on the individual posterior $N\left(\mu_{1c}, \sigma_{1c}^2\right)$. As in the case of the firm, even under the same common outside option and the same posterior mean, the threshold for referred candidates, $\underline{\mu}_{1c}(x)$, is higher than the threshold for non-referred candidates, $\underline{\mu}_{0c}(x)$. Finally, after they learn the actual match quality with the firm, both the referred and non-referred workers face the same problem, so the stay threshold $\underline{\theta}(x)$ is the same.

Our key new insight is that the informational content of referrals generates strong testable predictions for the dynamics of the hiring process. We start by exploring the relation between observed probabilities and underlying match quality. The following proposition summarizes the optimal hiring and separation rules in a form that is easy to take to the data.

Proposition 1. *Under Assumptions 1-4, the optimal decision rules can be represented by a simple multistage choice model for both referred and non-referred candidates,* $r = 0, 1$:

$$\begin{aligned}
o &= 1\left[\mu_{rf} > \underline{\mu}_{rf}(x)\right] \\
a &= 1\left[\mu_{rc} > \underline{\mu}_{rc}(x)\right] \\
s &= 1\left[\theta_r > \underline{\theta}(x)\right] \\
y &= f(x) + \theta_r + \epsilon
\end{aligned}$$

where a *is observed if* $o = 1$, s *is observed if* $a = 1$, *and* y *is observed if* $s = 1$. *The functions* $\underline{\mu}_{rk}(x)$ *and* $\underline{\theta}(x)$ *decrease in* x, $\underline{\theta}(x) > \underline{\mu}_{rk}(x)$, *and* $\underline{\mu}_{1k}(x) > \underline{\mu}_{0k}(x)$, *where* $k = f, c$ *and* $r = 0, 1$.

We allow referrals and interviews to transmit different informations to the different parties, so that we can test explicitly whether such is the case. If they provide the same information to all, the posterior means in the offer and acceptance decisions are strongly positively correlated with each other and with the stay decision. We interpret a rejection of this prediction as evidence that referrals provide different information to the candidate and the firm. The following corollary presents the claim formally.

Corollary 1 *If their posterior beliefs coincide,* $\theta_{rc} = \theta_{rf}$, *and Assumptions 1-4 hold, the firm and the candidate agree on whether they should enter in an employment relation. Thus, the first two stages can be combined into one entry decision:* $e = 1\left[\mu_r > \underline{\mu}_r(x)\right]$, where $\underline{\mu}_1(x) > \underline{\mu}_0(x)$.

When information available before, during, and after the hiring process is correlated with the actual quality of the employment relation, its presence introduces dependence in the hiring, stay and performance decisions. While it may or may not be available to both the firm and the candidates, such information remains unobservable to the econometrician. In such a context, the fact that some observationally equivalent candidates are hired, while others are not, provides information to the econometrician about future performance, stay, and promotions, even conditional on observed characteristics. When referred candidates have relatively more precise beliefs about their match with the firm than the non-referred candidates, their hiring has greater predictive power about stay and performance than the hiring of a non-referred candidate. The relation between match quality and productivity can be established by studying the dependence between the distributions of entry, stay and performance. If both productivity and stay decisions depend positively on match quality, positive selection into employment also implies higher likelihood of retention and higher expected performance. The following proposition presents formally these predictions.

Proposition 2. *Suppose that Assumptions 1-4 hold, and that, for simplicity, the posterior beliefs of the firm and the candidate coincide. Then:*

1. *There is stronger positive dependence between entry, stay, and performance for the referred than for the non-referred candidates:* $Corr\left(\mu_1, \theta_1\right) \geq Corr\left(\mu_0, \theta_0\right)$*, implying that for given thresholds* k_s *and* k_e

$$\Pr\left(\theta_1 > k_s | x, \mu_1 > k_e\right) \geq \Pr\left(\theta_0 > k_s | x, \mu_0 > k_e\right)$$

The inequalities hold strictly unless match quality becomes known during the hiring process, $\sigma_\xi^2 \to 0$*, or signals are perfectly uninformative,* $\sigma_\varepsilon^2 \to \infty$ *and* $\sigma_\xi^2 \to \infty$*.. Also,* $Corr\left(\mu_0, \theta_0\right) \geq 0$*, with equality attained when the hiring process is non-informative,* $\sigma_\xi^2 \to \infty$.

2. *Even if the distributions of match quality of referred and non-referred candidates are the same, conditional on entry the referred are more likely to stay and perform better than the non-referred:*

(i). $\Pr\left(\theta_1 > \underline{\theta}\left(x\right) | x, \mu_1 > \underline{\mu}_1\left(x\right)\right) \geq \Pr\left(\theta_0 > \underline{\theta}\left(x\right) | x, \mu_0 > \underline{\mu}_0\left(x\right)\right)$;

(ii). $E\left(y_1 | x, \mu_1 > \underline{\mu}_1\left(x\right)\right) \geq E\left(y_0 | x, \mu_0 > \underline{\mu}_0\left(x\right)\right)$.

3. *If a referred candidate and non-referred candidate have the same probability of entry or stay, or the same performance, then the observable characteristics of the referred cannot dominate those of the non-referred. Formally, if* $\Pr\left(\mu_1 > \underline{\mu}_1\left(x\right) | x\right) = \Pr\left(\mu_0 > \underline{\mu}_0\left(x'\right) | x'\right)$*, then it is not possible that* $x > x'$*. Similar statement holds for the conditional probabilities of stay and for performance.*

Proof of Proposition 2: Observe that for σ_ξ^2, σ_ε^2 finite

$$\mu_1 = \frac{\sigma^2\left(\sigma_\xi^2 + \sigma_\varepsilon^2\right)}{\sigma_\xi^2\sigma_\varepsilon^2 + \sigma^2\left(\sigma_\xi^2 + \sigma_\varepsilon^2\right)}\theta_1 + \lambda_1$$

$$\mu_0 = \frac{\sigma^2}{\left(\sigma_\xi^2 + \sigma^2\right)}\theta + \lambda_0$$

where λ_0 and λ_1 contain the remaining terms in the posterior means. Note that

$$\frac{\sigma^2\left(\sigma_\xi^2+\sigma_\varepsilon^2\right)}{\sigma_\xi^2\sigma_\varepsilon^2+\sigma^2\left(\sigma_\xi^2+\sigma_\varepsilon^2\right)}=\frac{\sigma^2\left(\frac{\sigma_\varepsilon^2}{\sigma_\xi^2}+1\right)}{\sigma_\varepsilon^2+\sigma^2\left(\frac{\sigma_\varepsilon^2}{\sigma_\xi^2}+1\right)}$$

Since $\left(\frac{\sigma_\varepsilon^2}{\sigma_\xi^2}+1\right)>1$, it follows that $Corr\left(\mu_1,\theta_1\right)\geq Corr\left(\mu_0,\theta_0\right)$. Note that in the limit,

$$Lim_{\sigma_\xi^2\to\infty}\frac{\sigma^2\left(\sigma_\xi^2+\sigma_\varepsilon^2\right)}{\sigma_\xi^2\sigma_\varepsilon^2+\sigma^2\left(\sigma_\xi^2+\sigma_\varepsilon^2\right)}=\frac{\sigma^2}{\sigma_\xi^2+\sigma^2}>0 \text{ and } Lim_{\sigma_\xi^2\to\infty}\frac{\sigma^2}{\sigma_\xi^2+\sigma^2}=0$$

while

$$Lim_{\sigma_\xi^2\to 0}\frac{\sigma^2\left(\sigma_\xi^2+\sigma_\varepsilon^2\right)}{\sigma_\xi^2\sigma_\varepsilon^2+\sigma^2\left(\sigma_\xi^2+\sigma_\varepsilon^2\right)}=1 \text{ and } Lim_{\sigma_\xi^2\to 0}\frac{\sigma^2}{\left(\sigma_\xi^2+\sigma^2\right)}=1$$

which implies the limit results. Given the properties of truncated normal distribution and $Corr\left(\mu_1,\theta_1\right)\geq Corr\left(\mu_0,\theta_0\right)$,

$$\Pr\left(\theta_1>k_s|\mu_1>k_e\right)\geq\Pr\left(\theta_0>k_s|\mu_0>k_e\right)$$

Note that from the characterization of the search problem, $\underline{\mu}_1\left(x\right)>\underline{\mu}_0\left(x\right)$. This observation, in combination with the properties of truncated normal distributions and $Corr\left(\mu_1,\theta_1\right)\geq Corr\left(\mu_0,\theta_0\right)$, implies:

$$\Pr\left(\theta_1>\underline{\theta}\left(x\right)|x,\mu_1>\underline{\mu}_1\left(x\right)\right)-\Pr\left(\theta_0>\underline{\theta}\left(x\right)|x,\mu_0>\underline{\mu}_0\left(x\right)\right)>0$$

On the other hand, we have that

$$\Pr\left(\theta_1>\underline{\theta}\left(x\right)|x\right)-\Pr\left(\theta_0>\underline{\theta}\left(x\right)|x\right)=0$$

Similarly, we obtain the result in part 2 (ii). Inverting the probability to solve for the observable x delivers part 2 (iii). ■

These results can be generalized in several ways. One important extension relates to the empirically plausible case in which referred candidates have higher match quality than non-referred candidates.

Assumption 1': Differences in Match Quality *The original Assumption 1 holds except that* θ_0 *is drawn from* $N\left(\mu,\sigma^2\right)$, *while* θ_1 *is drawn from* $N\left(\mu_*,\sigma^2\right)$, *where* $\mu_*>\mu$.

This assumption ensures that referred candidates come from a stochastically dominant distribution of match quality. Proposition 8 obviously survives. In fact, a version of it still holds for the de-meaned signals and match quality. Moreover, this new environment introduces additional dimensions in the analysis of the effects of referrals: the implications of referrals for persistent differences in performance and compensation among employees who remain employed in the long run.

Proposition 2' *If Assumption 1' and Assumptions 2-4 hold, then the differences in performance of referred and non-referred employees who stay employed long enough to learn their true match quality are smaller than the differences in performance of referred and non-referred candidates:*

$$E\left(y|x,\theta_1>\underline{\theta}\left(x\right)\right)-E\left(y|x,\theta_0>\underline{\theta}\left(x\right)\right)<E\left(y|x\right)-E\left(y|x\right)$$

Proof of Proposition 2': Note that

$$\begin{aligned}
&E\left(y|x,\theta_1>\underline{\theta}\left(x\right)\right)-E\left(y|x,\theta_0>\underline{\theta}\left(x\right)\right)\\
&=\mu_*-\mu+\sigma\left(\frac{\phi\left(\frac{\underline{\theta}(x)-\mu_*}{\sigma}\right)}{\Phi\left(\frac{\underline{\theta}(x)-\mu_*}{\sigma}\right)}-\frac{\phi\left(\frac{\underline{\theta}(x)-\mu}{\sigma}\right)}{\Phi\left(\frac{\underline{\theta}(x)-\mu}{\sigma}\right)}\right)\\
&\leq\mu_*-\mu
\end{aligned}$$

since the inverse Mill's ratio is an increasing function. The proposition presents the main implication of the model: referral signals increase the dependence between entry, stay and performance for the referred candidates relative to the non-referred candidates. Thus, the informational content of referrals induces selection on an unobserved firm-specific characteristic. If not controlled for, this selection leads to biased estimates of the effect of referrals on entry, turnover, performance, and promotions because

$$E\left(\theta_r|\mu_r>\underline{\mu}_r\left(x\right)\right)>E\left(\theta_r\right)$$

Moreover, without explicitly controlling for the dependence induced by referrals, one cannot identify underlying differences in the quality of referred and non-referred individuals from the effects of referrals on sorting during the hiring process.■