Job Market Signalling Via Social Ties

Shahir Safi*

August 27, 2020

Abstract

When there is asymmetric information about abilities of applicants, employees can "signal"—recommendation letter or put in a good word—some of this information to firms by recommending applicants they personally know. I consider how employees strategically transmit information to firms when there are both gratitude benefits for recommendations and reputation costs for providing inaccurate information. Unlike the classic setting of job market signalling by applicants, senders (employees) can now have imperfect information about abilities and my analysis builds upon this feature. I then develop two applications of this model which consider the strategic "selections" of ties and hiring channels by applicants. This allows me to help explain mixed evidences about the returns to using different types of ties (weak vs strong) and hiring channels (formal market vs social tie), as well as the relatively higher use of social ties by low skill/ability applicants.

Keywords: Signalling, Social Ties, Employee Referrals, and Hiring Channels.

JEL Codes: D82, D83, D85, J31.

*Aix-Marseille University (Aix-Marseille School of Economics). E-mail: shahir.safi@univ-amu.fr. This work was supported by French National Research Agency Grants ANR-17-EURE-0020. I also want to acknowledge funding that I received during my PhD at Universitat Autònoma de Barcelona from various sources, including Severo Ochoa Program for Centers of Excellence in R&D of the Spanish Ministry of Economy and Competitiveness, Fundación Areces Excellence Distinction, RecerCaixa project on "Inequality, Social Mobility, Effort, and Education", and Formación de Personal Investigador fellowship from the Ministerio de Economía, Industria y Competitividad. All errors are mine.

1 Introduction

Economists are increasingly becoming interested in how social networks based mechanisms influence economic outcomes. This paper considers one such mechanism, the employee referrals, i.e. when firms are looking to hire new workers, they ask their employees to recommend potential candidates. The use of employee referrals is widespread.[1] For instance, Careerbuilder.com is a leading online job site where firms can register and post vacancies, and according to their internal data, about 70% of firms have employee referral programs (Careerbuilder (2012)). So referral-based hiring is not something firms do at random, they have a clear agenda laid out by these programs. Scholars have long recognized firms' interest in the use of referrals; in particular, economists often argue that this is because employees can help screen job applicants.[2] The underlying idea here is that when there is asymmetric information about the abilities of applicants, employees can "signal" (recommendation letter or put in a good word) some of this information to firms by recommending applicants they personally know. Even though the use of employee referrals has gained increasing attention from scholars, this literature has mainly focused on the perspectives of job seeking workers and firms.[3] Studies examining the importance of networks for job seeking workers are numerous and span over a half-century, studies adopting the hiring firm's viewpoint have emerged over the past two decades, and yet, very few studies consider the employee's (or intermediary's) side of the story. As a result, the existing models of referrals either do not consider an active role for employees, or attribute to firms and employees the same objective—a simplifying assumption.

In this paper, I propose a model of referrals which considers how employees strategically transmit information about applicants to firms. To fix ideas, imagine the following scenario. A firm faces a job application, and it has an employee who is partially informed about the applicant's ability. The applicant's own ability, θ, is his private information but the employee receives a noisy signal about it: $s = \theta + \epsilon$ where $\epsilon \sim N(0, \sigma_\epsilon^2)$. I will refer to this noisy signal as the "exo-signal" hereafter because it comes from an exogenously given signalling process. Given this information, the employee sends a recommendation message—an "endogenous job market signal"—to the firm, which then offers a wage to the applicant. Even though the firm does not receive an exo-signal, it knows how well the employee knows the applicant (precision of exo-signals $1/\sigma_\epsilon^2$), and this allows the firm to form rational expectations. I allow the firm

[1]Studies covering both the US and other countries report that roughly half (or more) of all jobs are found through social ties (friends, neighbors, acquaintances, etc). See Topa (2011) for a recent survey of studies documenting the use of referrals across a variety of occupations, skill levels, and socioeconomic backgrounds.

[2]This idea dates back to 1966 in a paper by the labor economist Albert Rees.

[3]See Ioannides & Loury (2004) and Topa (2011) for two surveys of the economics literature, and Marsden and Gorman (2001) for a survey of the sociology literature.

and its employee to have different interests: the employee faces reputation costs of providing inaccurate information, but he also gets social benefits from recommending the applicant, and the value of these benefits are increasing in the wage offered to the applicant. So the employee faces a trade-off—recommendations that lead to higher wages for the applicant are those that can hurt his reputation. This communication game is similar to the classic signalling setting of Spence (1973) but with the innovation that the sender (employee) can have imperfect information, and the quality of his information is given by the precision of random exo-signals. While Spence was concerned with job market signalling (education level) by the applicant, I consider job market signalling (recommendation level) by the employee, who has incentives to provide such a costly (reputation costs) service due to gratitude benefits that come with having a social tie. Unlike education level, which is represented as an unbounded "effort" variable, the recommendation level is a bounded "report" variable; moreover, since the employee is reporting about the ability level, his recommendations are bounded by the values of ability.[4]

After presenting the main features of this model, I build two applications.

First, I consider the case with two types of employees who differ in their knowledge of applicants' abilities—those with strong ties (such as family members and close friends) have perfect information, while those with weak ties (such as acquaintances) have imperfect information. This allows me to examine how the tie strength (weak vs strong) influences job market outcomes, and the decision of job seeking workers (applicants) to choose different types of ties based on their returns in the labor market.

The empirical evidence on the labor market returns to weak versus strong ties is mixed. In the pioneering works of Granovetter (1973, 1995), he found that a large proportion of jobs are found through weak rather than strong ties. He argued that weak ties—individuals with whom one has few common friends—are most useful for job search, because they provide access to otherwise unobtainable information about job openings. This finding led to the coining of the well-known phrase, "strength of weak ties", implying that tie strength matters for determining labor market outcomes, and that weak ties lead to better outcomes. However, subsequent studies have also found the frequent use of strong ties, and overall, the empirical evidence on the relative use of weak versus strong ties is mixed.[5] In addition, Bridges & Villemez (1986) and Marsden & Hurlbert (1988) found no significant relationship between tie

[4]Bounded message space in costly signalling setting has been considered before. See Cho & Sobel (1990) for an early reference and Kartik (2009) for a more recent work. I add to this literature by allowing the sender to have imperfect information, and much of my analysis has to do with this feature.

[5]Lin et al (1981), Lai et al (1998), and Yakubovich (2005) find further empirical support for Granovetter's strength of weak ties hypothesis, while Murray, Rankin, and Magill (1981), Watanabe (1987), and Bian (1997, 1999) find evidence against his hypothesis.

strength and wages after controlling for worker characteristics. As a result, some researchers have argued against the "strength of weak ties" hypothesis.

My analysis shows that a model of referrals with strategic recommendations by employees and tie selections by job seeking workers (applicants) can help reconcile Granovetter's "strength of weak ties" theory with the opposing empirical findings. Now consider an environment where a fraction of applicants have access to both weak and strong ties, and so these applicants can choose which type of tie/employee to use for the referral. Even though the firm does not know whether an applicant had the option to choose between the two types of employees, it knows that a fraction of applicants have access to both types of employees, and it also knows the type of employee chosen to provide the referral, which then allows it to form rational expectations.

The starting point of my analysis is to consider how much information firms can infer from the recommendations of its employees. If firms could always infer exo-signals of its employees, then indeed returns to weak ties would be higher, as predicted by Granovetter. This follows from the fact that exo-signals of weak ties are imprecise. Consequently, some applicants get "lucky", they send high exo-signals to their weak ties, which then allows them to pool with a fraction of higher ability applicants that don't have strong ties and got "unlucky" by sending low exo-signals to their weak ties. Since weak ties allow applicants to pool with those above them, returns to weak ties would always be higher. However, I find that firms cannot infer sufficiently high exo-signals, which then lowers the relative returns to using weak ties for higher ability applicants—because such applicants can also use strong ties to pool with those above them via "pool at the top" and returns from pooling with weak ties via "pool at the top" can be lower for them. This follows from equilibria of the communication game; in particular, I characterize the Perfect Bayesian Equilibria of this game, satisfying a standard belief based refinement (criterion D1). At such equilibria, an employee will always recommend higher than his expected value of applicant's ability. This is due to the presence of gratitude benefits, which incentivizes employees to recommend higher than their valuations. Since recommendation messages are bounded by ability values, employees will eventually run out of recommendation values that they can send (above some value of exo-signals s^*), firms are unable to distinguish among applicants that gave sufficiently high exo-signals to employees (above the threshold s^*), and these applicants segment into a pool.

The low exo-signals separate and high exo-signals pool (LSHP) communication structure then provides a basis for understanding the empirical findings regarding tie strength. The relationship between tie strength and wages is negative for sufficiently low ability workers (applicants) but can be positive higher ability workers, which in turn can help explain why some researchers found no relationship between tie strength and wages—when considering workers of all abilities. The extent to which workers use weak and strong ties depends on the access to strong ties, which determines the returns from pooling with weak ties. Contrary to the existing explanation, the frequent use of weak ties may not be due to its efficiency in the matching of workers and firms. When the access to strong ties is really scarce for workers, many workers—even those with access to strong ties—use weak ties to pool with even higher ability workers. Since the relative use of weak and strong ties vary with the access to strong ties, this can then help explain the mixed evidence about the use of different types of ties. Although these predictions do not emerge in the existing models of employee referrals, they are consistent with the empirical evidence, suggesting that strategic recommendations by employees and tie selections by workers are important aspects of employee referrals.

In addition, I find that more informative employees are not necessarily more informative to firms. Even though exo-signals of more informative employees are more precise, firms may infer fewer exo-signals from them. In particular, it can be shown that firms infer more exo-signals from weak ties—less informative employees—when reputation costs are sufficiently high. In such case, returns to weak versus strong ties vary non-monotonically with abilities—returns to weak ties are higher for sufficiently low and high ability applicants, whereas returns to strong ties are higher for medium ability applicants. Since one would expect reputation costs to be sufficiently high for referrals to be informative to firms, one should also expect its novel implication (about returns to weak and strong ties) to be present in the data, which provides a new testable finding for future research.

Second, I consider the case in which applicants can apply for a job through social ties or directly through the formal market channel. As before, an applicant's own ability is his private information but now both employees and firms receive exo-signals about his ability. To fix ideas, noisy signals received by firms represent the presence of formal screening mechanisms, such as aptitude tests and other attribute measurements, which are typically used in a formal hiring process to acquire some information about the abilities of applicants. Conversely, referral is the screening mechanism used in an "informal" (via social tie) hiring. I then examine how hiring channels (formal market vs social tie) and their corresponding screening methods (formal screening vs referral) influence job market outcomes, and the decision of workers to choose different types of hiring channels based on their returns in the labor market.

Although one might think that returns to using social tie over formal market channel are always positive, the empirical evidence on this is mixed. Some studies show that workers who found their jobs through family, friends, and acquaintances earned more than those using formal and other informal job-search methods (Rosenbaum et al. (1999), Marmaros and Sacerdote (2002)). Others indicate that the initial wage advantage declined over time (Corcoran, Datcher, and Duncan (1980), Simon and Warner (1992)). Some analysts found no general initial or persistent wage effects (Bridges and Villemez (1986), Holzer (1987), Marsden and Gorman (2001)). In fact, some studies (Elliott (1999), Green, Tigges, and Diaz (1999)) show that those using contacts earned less than those using formal methods. My model can help explain these opposing empirical findings because returns to using social ties vary with the abilities of applicants—they can be positive for lower ability applicants but are negative for higher ability applicants.

In fact, my analysis shows that for sufficiently high ability applicants, returns to using the formal market channel is always higher than social ties—regardless of how bad the formal screening mechanism may be—, which is consistent with one of the most robust empirical findings in the referrals literature. At this point, there is an empirical consensus that referrals are used more by "low skill" workers—those with lower education levels and lower socio-economic status, and for "lower status" jobs.[6] Ornstein (1976), Corcoran et al (1980), Datcher (1983), Marx and Leicht (1992), all report higher usage for less educated job seeking workers. Elliot (1999) finds that social ties are more frequently used in high-poverty neighborhoods than in low-poverty ones. Similarly, Green et al. (1995) report that poor job seeking workers in Atlanta were more likely to use friends and relatives than non-poor ones. Rees and Schultz (1970) and Corcoran et al (1980) both find that referrals are used more often for blue-collar than for white-collar occupations. In my model, referrals limit firms' capacity to infer sufficiently high exo-signals (due to LSHP communication structure), whereas firms can always infer exo-signals through formal screening mechanisms. Even if social ties receive more precise information (less noise in exo-signals) than what firms receives through formal screening mechanisms, the inability of firms to infer those exo-signals from the recommendations of its employees makes formal screening more informative to firms and more desirable for higher ability applicants.

In summary, this is the first model of referrals which considers how employees strategically transmit information to firms, and its influence on the "selection" of ties and hiring channels by applicants. My analysis shows that referrals are not effective in transmitting information about high ability applicants. Since the effectiveness of referrals vary with the ability of applicants, so will the returns to using them. This allows me to help explain mixed evidences about the returns to using different types of ties (weak and strong) and hiring channels (formal and informal/social tie), as well as the relatively higher use of social ties by low skill/ability applicants.

The rest of the paper is organized as follows. I will end this section by discussing some

[6] Note that skill is synonymous with ability in my model.

closely related literatures. Section 2 introduces the referrals model. In section 3, I characterize equilibria—recommendation strategies and wage offers—, and examine comparative statics with respect to the precision of exo-signals (noisy signals). Here I assume exo-signals come from a normal distribution, whereas they come from a uniform distribution in section 4. In section 4, I develop two applications of the referrals model which consider the strategic "selections" of ties and hiring channels by applicants. Both the precision of exo-signals and "selections" (of ties and hiring channels) have effects on recommendation strategies. However, when there is a lack of prior knowledge—which is the case with a uniform distribution—, there are no effects of the precision of exo-signals on recommendations; hence, in section 4 I consider a uniform distribution in order to isolate the effects of "selections" (of ties and hiring channels) on recommendation strategies. In section 5, I conclude and suggest some avenues for future research. All proofs are available in the appendix.

1.1 Related Literature

Before proceeding to the model, let me mention some closely related literatures.

My model helps explain some empirical findings, which contributes to three strands of the empirical literature on employee referrals: the mixed evidence on the returns to using weak versus strong ties, the mixed evidence on the returns to using social tie versus formal market channel, and the robust empirical finding that referrals are used more by "low skill" workers.

In addition, this paper also contributes to the theoretical literature of employee referrals along two dimensions.

First, there are two main classes of referral models—(i) those in which workers find information about new job vacancies through their social contacts (reducing search frictions), and (ii) those in which employees transmit information about the ability of workers (helping screen applicants). To name a few, Montgomery (1994), Calvó-Armengol (2004), Calvó-Armengol and Jackson (2004, 2007), Calvó-Armengol and Zenou (2005), and Gaelotti and Merlino (2014) belong to the first class of models. In such models, information about new job vacancies arrive randomly to workers, and already employed workers pass their information to one of their social contacts at random. So employees are non-strategic in these models, and their decisions are purely mechanical. On the other hand, Saloner (1985), Montgomery (1991), Simon and Warner (1992), Arrow and Borzekowski (2004), Galenianos (2013), and Dustmann et al (2016) are some examples of the second class of models. In Saloner's (1985) model, firms and employees have the same objective–they both want to hire applicants with the highest abilities–, whereas in all the other models, employees do not play an active role. For instance, in Montgomery's (1991) model, each employee is connected with at most one job seeking worker and each employee must provide referral for his job seeking contact. Employees are more likely to be connected with people like themselves (homophily by ability), and firms can use this information to learn about applicants abilities. Similarly, in the models of Simon and Warner (1992), Arrow and Borzekowski (2004), and Galenianos (2013), employees are non-strategic. Firms receive noisy signals about applicants, and these signals are assumed to be more precise for referred applicants. Dustmann et al (2011) extend the model of Simon and Warner (1992) to consider the case where the probability of obtaining a job through a referral is endogenous, and it depends on the workforce composition of the firm. Galenianos (2013) also consider an extension where firms can choose the precision of signals for non-referred applicants at a cost. My work adds to the second class of referral models by considering an active role for employees and allowing them to have different objectives than their firms. As a result, in my model employees are strategic in their transmission of information to firms.

Second, the issue of weak versus strong ties have been considered before. Following the seminal paper of Granovetter (1973), the existing literature has only considered tie strength when the role of referrals is to reduce search frictions. For instance, Montgomery (1994) added a simple social structure and pattern of social interaction to a Markov model of employment transitions. In his model, society is composed of many small (two-person) groups. Unemployed individuals find jobs through strong ties (intra-group social interaction), weak ties (random intergroup interaction), and formal channels. Holding constant the total level of social interaction, the author examines how a change in the composition of social interaction (in particular, the use of weak ties) affects the steady-state equilibrium. He finds that an increase in weak-tie interactions reduces inequality, thereby creating a more equitable distribution of employment across groups. Furthermore, Calvó-Armengol et al (2007) and Zenou (2013, 2015) extended his model to consider the effects of unemployment benefits on crime and the effects of weak-tie interactions on the employment rate, respectively. My work is complementary to theirs in that they considers tie strength when referrals reduce search frictions, whereas I consider tie strength when referrals can help screen job applicants. In addition, job seeking workers are non-strategic in these models in that their decisions to interact with different types of ties is purely random.[7] Thus, I also add to this literature by allowing tie selection to be endogenous, that is job seeking workers can choose their preferred

[7]The only exception is an early paper by Boorman (1975), in which he examines the incentives of workers to form weak versus strong ties. Like the rest of the existing literature, he also focuses on the role of referrals in reducing search frictions. However, his model is hard to solve and can only provide results through simulations. My contribution here is to provide a feasible framework in which one can analytically solve for the decisions of workers to choose between weak and strong ties.

type of tie.

And finally, this paper contributes to the literature on costly signalling. Although a sender with imperfect information has been considered before in communication games, such as the cheap talk setting, this is the first paper to introduce a sender with imperfect information in the costly signalling setting.[8] Indeed, employee referrals provide a natural environment in which senders (employees) have imperfect information because employees only partially know about the abilities of applicants, and signals (recommendations) are costly because there are reputation costs of providing inaccurate information.

2 Model

A firm faces a job application, and it has an employee who is partially informed about the applicant's ability. The applicant's own ability θ is his private information, and it is drawn from a normal distribution with mean μ and variance σ_θ^2. The realization of θ is not observed by the employee; instead, he observes a noisy signal $s = \theta + \epsilon$ where ϵ is independent of θ and is drawn from a normal distribution with mean 0 and variance σ_ε^2. I will refer to this noisy signal as the "exo-signal" hereafter because it comes from an exogenously given process. Thus, the posterior for θ given an exo-signal s is normal with mean $z = E[\theta|s] = \lambda s + (1-\lambda)\mu$ and variance $\sigma_z^2 = \left(\sigma_\epsilon^{-2} + \sigma_\theta^{-2}\right)^{-1}$, where $\lambda = \frac{\sigma_\epsilon^{-2}}{\left(\sigma_\epsilon^{-2}+\sigma_\theta^{-2}\right)}$. Given the exo-signal s, the employee sends a recommendation message m to the firm, which is an "endogenous job market signal" and the focus of this paper. After observing the recommendation message m, the firm offers a wage equal to its valuation $w = E\left[\theta|m\right]$.

The job is suitable for applicants with abilities in the interval $[0,1]$, and applicants are underqualified if $\theta < 0$ or overqualified if $\theta > 1$. The applicant's payoff is given by $U^A(w) = w$. The employee gets a payoff of zero when he does not refer, and his payoff from providing a referral is negative if he believes the applicant is underqualified ($s < s_{min}$ with $E[\theta|s_{min}] = 0$) or overqualified ($s > s_{max}$ with $E[\theta|s_{max}] = 1$). Therefore, the employee does not recommend an applicant if he believes the applicant is unqualified $s \notin [s_{min}, s_{max}]$. When the employee believes the applicant is qualified for the job $s \in [s_{min}, s_{max}]$, his payoff is given by

$$U^E(w, m, s) = \gamma w - (1-\gamma) E\left[(m-\theta)^2 |s\right],$$

where γw represents the value of gratitude benefits from the applicant, $(1-\gamma)E\left[(m-\theta)^2 |s\right]$

[8]See Battaglini (2004), Ambrus & Lu (2014), and Lu (2017) for some examples of cheap talk games in which senders have imperfect information. See Riley (2001) and Horner (2006) for two surveys of the costly signalling literature.

represents the reputation costs of providing inaccurate information to the firm, and $0 < \gamma < 1$ indicates the relative weight that an employee puts on gratitude benefits over reputation costs. Here, reputation costs take the form of a quadratic loss function $(m - \theta)^2$—loss is measured by the difference between the recommendation and the ability values—, and they are in expectation terms (conditional on the exo-signal s) because the employee does not perfectly observe the ability value. So the employee faces a trade-off because recommendations that lead to higher wages for the applicant are those that can hurt his reputation.

Recommendation messages are reports about what the employee believes, and so without loss of generality, I will assume that $m \in [0, 1]$. Then, the employee's recommendation m carries the literal meaning "I believe the applicant's ability is m". All aspects of the game are common knowledge, except the values of θ and s—the applicant knows his own ability θ and the exo-signal s, the employee only knows the exo-signal s, and the firm knows neither.

2.1 Discussion

Here are some comments on the model's interpretations and assumptions.

2.1.1 Interpretations

So far, I have described the model in a way that abstracts from all but the bare essentials necessary to illustrate the motivating ideas. I will now discuss how it can be adapted to consider more sophisticated situations.

Wage Determination Process. One can consider a more realistic wage setting process, where wage is determined by a "Nash" bargaining between the firm and the applicant. In particular, the outcome of bargaining is determined by choosing the wage that maximizes the Nash product

$$E\left[(\theta - w)\,|m\right]^{\delta}\left[w\right]^{1-\delta},$$

where $E\left[(\theta - w)\,|m\right]$ is the expected payoff (net profit) of the firm given recommendation m, w is the payoff of the applicant, and δ is the bargaining power of the firm.[9] By taking the first order condition, one can immediately see that the firm offers a wage equal to a fraction of its valuation $w = \delta E[\theta|m]$. As a result, the assumption that the firm offers a wage equal to its valuation is innocuous. Indeed, one can relax this assumption by assuming that the firm offers a wage equal to some fraction of its valuation; nothing of substance will change.

[9]If the firm and the applicant cannot agree on a wage, then they both get nothing. In other words, their "outside options" are zero production and unemployment, respectively. Here I make these assumptions for a clean illustration, and of course, these assumptions can be further relaxed.

Labor Market. The model can easily be extended from a single referral environment to that of many referrals, and it can represent the aggregate labor market as follows. There are infinitely many job seeking workers characterized by their ability and social ties (θ, T), where ability is distributed $\theta \sim N(\mu, \sigma_\theta^2)$ independently from T. Each job seeking worker can either have a social tie t or not, that is $T \in \{t, \emptyset\}$. To fix ideas, there are infinitely many firms, each firm has exactly one employee, an employee can have at most one social tie, and δ portion of job seeking workers have a social tie with exactly one employee. In section 4, I further extend this to consider the case with two types of ties, weak and strong, denoted as t_{weak} and t_{strong} respectively.

2.1.2 Assumptions

Strategic Employees. Employees are strategic under $0 < \gamma < 1$. If $\gamma = 0$, the model reduces to the case where employees and firms have the same objective. In such case, the employee will always report his exo-signal truthfully, and the model is synonymous with the one in which the firm directly receives noisy signals about the applicant. On the other hand, when $\gamma = 1$, the employee only cares about gratitude benefits, and referrals become completely uninformative to the firm.

Job Qualification. Even though the employee's payoff from providing a referral is negative when he believes the applicant is unqualified for the job, the underlying reason for the case of an underqualified applicant is distinct from that of an overqualified applicant. In particular, following ideas are implicitly behind this assumption. On the one hand, a firm does not want to hire underqualified workers, and so the employee faces large reputation costs when he believes the applicant is underqualified. On the other hand, an overqualified applicant does not want to get paid below his value, and so the employee expects little gratitude benefits when he believes the applicant is overqualified.

2.2 Strategies and Equilibrium

A pure strategy for an employee is a function $\rho : [s_{min}, s_{max}] \to [0, 1]$, where $\rho(s)$ is the message that the employee sends when he receives the exo-signal s. Denote the posterior beliefs of the employee given exo-signal s by the CDF $F(\theta|s)$ (which is a normal distribution with mean z and variance σ_z^2, as specified above), and the posterior beliefs of the firm given recommendation m by the CDF $F(\theta|m)$. The firm offers a wage equal to its valuation $E[\theta|m]$, which is formed using its posterior beliefs after every message. The solution concept is Perfect Bayesian Equilibrium (PBE), which requires the employee to maximize utility for each exo-signal given that the firm offers a wage equal to its valuation, and beliefs to satisfy

Bayes' rule wherever it is well defined.

I consider pure strategy PBE, which are referred to as just "equilibria" hereafter. Moreover, I restrict attention to equilibria where the employee's strategy is weakly increasing, in the sense that if $s > s'$, then $\rho(s) \geq \rho(s')$. This property is called *message monotonicity*, and an equilibrium that satisfies the property is a *monotone equilibrium*. In a monotone equilibrium, the employee makes a weakly higher recommendation when he receives a higher exo-signal. Attention to such equilibria seems reasonable because employees with higher exo-signals prefer higher recommendations due to the reputation cost structure. In particular, observe that the recommendation message that minimizes reputation costs conditional on exo-signal s sets $m = E[\theta|s]$. By the assumptions on distributions of ability and exo-signal, $E[\theta|s]$ is increasing in exo-signal, and so the employees preferred recommendations are also increasing in exo-signal.

3 Strategic Recommendations

3.1 Impossibility of Full Separation

The basic signalling problem is that the sender (employee) wants the receiver (firm) to infer that he is a higher type (exo-signal) than he actually is. Say that a type s is *separating* if $\{s' : \rho(s') = \rho(s)\} = \{s\}$, i.e. the recommendation message sent by employee with exo-signal s reveals that his exo-signal is s. A *separating equilibrium* is one where all types are separating. Naturally, the first question is whether such equilibria exist. The following result specifies a key property of ρ in any separating region of types.

LEMMA 1. *If types (s_l, s_h) are separating in a monotone equilibrium, then for each $s \in (s_l, s_h)$, $\rho(s) > E[\theta|s]$ and*

$$\rho'(s) = \frac{\gamma\lambda}{2\,(1-\gamma)\,(\rho(s) - E[\theta|s])}. \tag{DE}$$

The lemma indicates that in a monotone equilibrium, ρ must solve (DE) on any open interval of types that are separating. Conditional on differentiability of ρ, this result is immediate from the first order condition of the following maximization problem for each $s \in (s_l, s_h)$:

$$argmax_{\widetilde{s}\in(s_l,s_h)} \left\{\gamma w(\widetilde{s}) - (1-\gamma)\, E\left[(\rho(\widetilde{s}) - \theta)^2\,|s\right]\right\}. \tag{1}$$

I then prove that ρ is indeed differentiable using arguments from Mailath (1987). Observe that the sign of ρ' in (DE) is determined by the sign of $(\rho(s) - E[\theta|s])$, as all the other terms on the right-hand side are positive. As a monotone equilibrium requires ρ to be weakly increasing, (DE) implies that all types in the interior of a separating interval must recommend higher than their valuations $\rho(s) > E[\theta|s]$. This is due to gratitude benefits, which incentivize employees to recommend higher than their valuations.

Next, I consider whether there can be complete separation in a monotone equilibrium. If so, Lemma 1 implies that all types in (s_{min}, s_{max}) must be reporting higher than their valuations. However, as recommendation messages are bounded by ability values $m \in [0, 1]$, the highest type s_{max} cannot be reporting higher than his valuation. This intuition underlies the following result.

THEOREM 1. *There is no separating monotone equilibrium.*

The formal logic is as follows. A separating monotone equilibrium requires a continuous function to solve an initial value problem defined by the (DE) together with some boundary condition $\rho(s_{min}) \geq 0$. The proof of Theorem 1 shows that there is no solution on the domain $s \in (s_{min}, s_{max})$, regardless of the choice of boundary condition. This is because each type needs to recommend higher than his valuation in order to separate itself from lower types, but one eventually "runs out" of recommendation claims that can be made. In particular, the solution to the initial value problem hits the upper bound of 1 at some $s < s_{max}$.

3.2 Low Types Separate and High Types Pool (LSHP)

Theorem 1 implies that any monotone equilibrium involves some pooling. The fact that the employee eventually runs out of types to mimic for separation suggests a class of pooling equilibria to study, where *low types separate and high types pool on the highest message*. I will start by defining this class precisely, then discuss the structure and existence of such equilibria, and finally provide a justification for focusing on this class.

A *separating function* is a continuous and increasing function that solves (DE) with the initial value condition $\rho(s_{min}) = 0$. This choice of initial value condition is motivated by the usual "Riley" condition of least costly separation in signalling games. The Appendix (Lemma A1) shows that there is a unique separating function, denoted ρ^* hereafter, and it is well defined on some interval $[0, \overline{s}]$, where $\overline{s} < s_{max}$ is such that $\rho^*(\overline{s}) = 1$.

DEFINITION 1. An employee's strategy ρ is an *LSHP (Low Types Separate and High Types Pool)* strategy if there exists a $\underline{s} \in [0, \overline{s}]$ such that:

(1) $\rho(s) = \rho^*(s)$ for all $s < \underline{s}$, and
(2) $\rho(s) = 1$ for all $s \geq \underline{s}$.

An equilibrium (ρ, w) is an LSHP equilibrium if ρ is an LSHP strategy.

An LSHP equilibrium features a cut-off type $\underline{s}$ such that any type $s < \underline{s}$ separates by playing $\rho^*(s)$, whereas all types above $\underline{s}$ send the highest message of one $\rho(s) = 1$. The following observations will help clarify this structure. It is clear what the firm must play on equilibrium path: $w(\rho^*(s)) = E[\theta|s]$ for all $s < \underline{s}$, and $w(1) = E[\theta|s \in [\underline{s}, s_{max}]]$ for all $s \geq \underline{s}$. As usual, the optimality of employee's strategy places some constraints on what responses the firm can take off the equilibrium path. On the employee's side, the following indifference condition must hold if $\underline{s} > s_{min}$:

$$\gamma E[\theta|s \in [\underline{s}, s_{max}]] - (1-\gamma)\, E\left[(1-\theta)^2\,|\underline{s}\right] = \gamma E[\theta|\underline{s}] - (1-\gamma)\, E\left[(\rho^*(\underline{s})-\theta)^2\,|\underline{s}\right]. \quad (2)$$

That is, type $\underline{s} > s_{min}$ must be indifferent between sending the highest message $\rho(\underline{s}) = 1$ and inducing wage $w(1) = E[\theta|s \in [\underline{s}, s_{max}]]$ vs. sending any $\rho^*(\underline{s})$ and inducing wage $E[\theta|\underline{s}]$. This follows from the fact that any $s < \underline{s}$ plays $\rho^*(s)$ and induces wage $E[\theta|s]$, combined with the equilibrium requirement that type $\underline{s}$ should not prefer to mimic some type $s < \underline{s}$, and conversely, no $s < \underline{s}$ should prefer to mimic $\underline{s}$. Equation (2) imposes a restriction on which types can be sending $\rho(s) = 1$. As $\rho(s) = 1$ for all $s \geq \underline{s}$, it is equally expensive in terms of direct message costs for a given type to mimic any type $s \geq \underline{s}$.

The preceding remarks have identified a set of necessary condition for an LSHP equilibrium with $\underline{s} > s_{min}$. The following result establishes that they are also sufficient, and provides an existence result taking into account the possibility that there may be LSHP equilibria where no types separate ($\underline{s} = s_{min}$).

THEOREM 2. *In any LSHP equilibrium, there is a cut-off type* $\underline{s} \in [s_{min}, \overline{s}]$ *such that*

$$\begin{aligned} &\gamma E[\theta|s \in [\underline{s}, s_{max}]] - (1-\gamma)\, E\left[(1-\theta)^2\,|\underline{s}\right] \\ &= \gamma E[\theta|\underline{s}] - (1-\gamma)\, E\left[(\rho^*(\underline{s})-\theta)^2\,|\underline{s}\right] \qquad \text{if } \underline{s} > s_{min}. \end{aligned} \quad (3)$$

Conversely, given any cut-off type that satisfy (3) and

$$\begin{aligned} &\gamma E[\theta|s \in [s_{min}, s_{max}]] - (1-\gamma)\, E\left[(1-\theta)^2\,|s_{min}\right] \\ &\geq -(1-\gamma)\, E[\theta^2|s_{min}] \qquad \text{if } \underline{s} = s_{min}, \end{aligned} \quad (4)$$

there is a corresponding LSHP equilibrium.

For any $0 < \gamma < 1$, *there is an LSHP equilibrium that satisfies (3) and (4).*

If $\gamma < \frac{1}{2}$, *there is an LSHP equilibrium with* $\underline{s} > s_{min}$.

Inequality (4) says that if the cut-off type is s_{min}, then type s_{min} weakly prefers pooling with $[s_{min}, s_{max}]$ by claiming to be the highest type s_{max} to revealing itself (thus receiving wage $E[\theta|s_{min}] = 0$) with a least costly message $\rho(s_{min}) = 0$. The existence of LSHP equilibrium is established by showing that a cut-off type $\underline{s}$ satisfying (3) and (4) always exists. The last part of the theorem is intuitive—it says that if the intensity of reputation cost $(1-\gamma)$ is sufficiently large, then at least some types can be separating in equilibrium.

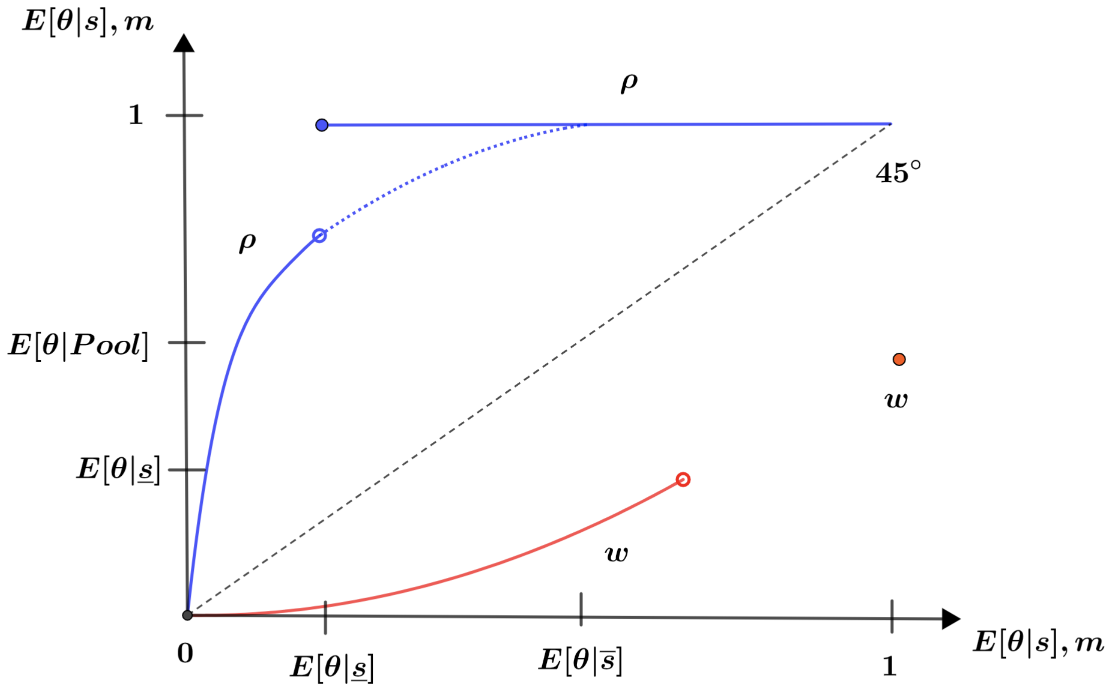

FIGURE 1. *LSHP Equilibrium*

Blue solid curve represents employee's recommendation strategy ρ, dotted curve is the separating function ρ*, and red solid curve represents firm's wage offer w.

Figure 1 illustrates an LSHP equilibrium. The employee's recommendation strategy, as a function of his valuation $E\,[\theta|s]$, is represented by the solid blue curve. The separating function, ρ^*, coincides with ρ on $[0, E\,[\theta|\underline{s}])$, and then extends as the dotted curve up to $E\,[\theta|\overline{s}]$. The firm's wage offer w, as a function of employee's recommendation message, is depicted by the solid red curve. Note that up until $\rho(\underline{s})$, w is the mirror image of ρ around the 45° line, which verifies that for the region of separating types, the firm inverts the employee's message. As the separating function ρ^* has $\rho^*(s) > E[\theta|s]$ for all $s \in (s_{min}, \overline{s}]$, such types are revealing themselves in a costly manner. However, telling the truth is not profitable. This is because it would lead to adverse inferences from the firm, who expects the employee to recommend higher than his valuation, and thus rationally deflates (seen in Figure 1 by w being strictly below the 45° line).

LSHP equilibria are an intuitive class of partially-pooling equilibria to focus on—they feature message monotonicity, which is appealing given the preference of higher types for higher messages; low types separate in the cheapest way possible that is consistent with message monotonicity; and all pooling occurs on the highest message, which is effectively the barrier to complete separation.

For these reasons, the structure of LSHP equilibria is reminiscent of the "D1" equilibria of Cho & Sobel (1990) and the "monotonic D1" equilibria of Kartik (2009). Indeed, Cho and Sobel (1990) established that in standard signalling games where complete separation is not possible owing to a bound on the signal space, the D1 refinement of Cho and Kreps (1987) select equilibria where low types separate in the least costly manner and high types pool on the maximal costly signal. Although I consider an environment that is different from theirs, under message monotonicity, the D1 refinement can be shown to rule out any non-LSHP equilibrium.[10] Moreover, there is always an LSHP equilibrium that satisfies the refinement criterion. Therefore, focusing on this class of equilibria can be formally justified. See appendix B (available online) for complete details.

3.3 Precision of Exo-Signals

Referral's environment is similar to the classic costly signalling setting of Spence (1973), in that employees can "signal" (recommendations) some information about the abilities of applicants, but with the innovation that senders (employees) can have imperfect information. I will now consider how the precision of employee's information (exo-signal) influences the LSHP equilibria. To this end, I will make the referrals environment explicit by indicating variables as a function of $t = (\sigma_\epsilon^2, \sigma_\theta^2, \mu)$. In particular, I now denote $\rho\,(s, t)$ as the employee's recommendation strategy, $F\,(\theta|s, t)$ as the employee's posterior beliefs, and $F\,(\theta|m, t)$ as the firm's posterior beliefs. In addition, I will denote $\theta(s \in [\underline{s}, s_{max}]\,, t)$ as the posterior distribution of θ at the pool $[\underline{s}, s_{max}]$.

> DEFINITION 2. An LSHP strategy ρ under $t = (\sigma_\epsilon^2, \sigma_\theta^2, \mu)$ is more *informative to the firm* than another LSHP strategy ρ' under $t' = (\sigma_\epsilon^{2'}, \sigma_\theta^2, \mu)$ if $\sigma_\epsilon^2 < \sigma_\epsilon^{2'}$ and $|\underline{s}(t) - s_{min}(t)| > |\underline{s}(t') - s_{min}(t')|$.

The above definition is appropriate because it considers not only the precision of employee's exo-signals but also how many exo-signals the firm can infer from his recommendation strategy. Note that even though $\underline{s}(t)$ determines the threshold below which the firm can infer

[10] In addition to message monotonicity, Kartik (2009) also considers action monotonicity to establish a similar result. This is because his model has a cheap-talk payoff structure in which the utility of sender is not always increasing with the action of receiver. Importantly, my model differs from his and the standard signalling games in that I allow the sender to have imperfect information, and much of my analysis has to do with this feature.

signals in an LSHP equilibrium, the interval of signals that the firm can infer is $[s_{min}(t), \underline{s}(t))$. For this reason, one needs to compare the length of such intervals, which measures how many exo-signals the firm can infer, under different LSHP strategies.

Naturally, the first question is whether the "more informative employee", the one with "more precise information", is more *informative to the firm*. The following theorem shows that this is not necessarily the case.

THEOREM 3. *Fix $t = (\sigma_\epsilon^2, \sigma_\theta^2, \mu)$ and $t' = (\sigma_\epsilon^{2\prime}, \sigma_\theta^2, \mu)$ such that $\sigma_\epsilon^{2\prime} > \sigma_\epsilon^2$.*

(1) For $\mu \leq 0$, $s_{min}(t) \leq s_{min}(t')$ and $\rho_s^*(s,t) > \rho_s^*(s,t')$.
(2) For $\mu \geq 1$, $s_{min}(t) \geq s_{min}(t')$ and $\rho_s^*(s,t) < \rho_s^*(s,t')$.
(3) For $\mu \in (0,1)$, $s_{min}(t) \geq s_{min}(t')$ and if $min\left\{\underline{s}(t), \underline{s}(t')\right\} > \mu$, then there is a pair (s_l, s_r) such that (i) $\rho_s^*(s,t) < \rho_s^*(s,t')$ when $s < s_l$, and (ii) $\rho_s^*(s,t) > \rho_s^*(s,t')$ when $s > s_r$.

For γ sufficiently low, $\underline{s}(t) < \underline{s}(t')$ when $\mu \leq 0$ and $\underline{s}(t) > \underline{s}(t')$ when $\mu \geq 1$. Moreover, as $\sigma_\theta^2 \to \infty$ (lack of prior), (i) $\rho(s,t) = \rho(s,t')$ for all s, t, $t' \neq t$, and (ii) $\theta(s \in [\underline{s}, 1], t) \succ_{SOSD} \theta(s \in [\underline{s}, 1], t')$.[11]

The first part of this theorem ((1) – (3)) shows how the precision of an employee's information influences his recommendation strategy and the LSHP equilibrium. For sufficiently low prior mean $\mu \leq 0$, the firm can end up inferring fewer exo-signals from the more informative employee. This is because the precision of an employee's information can influence his recommendation strategy through both gratitude benefits and reputation costs, which can be seen in Lemma 1 where the right-hand side of the (DE) is increasing in both $\lambda(t)$ (influence through gratitude benefits) and the posterior mean $E[\theta|s,t]$ (influence through reputation costs). Observe that the posterior mean also depends on $\lambda(t)$, $E[\theta|s,t] = \mu + \lambda(t)(s - \mu)$, and it can be shown that $\frac{\partial \lambda(t)}{\partial \sigma_\epsilon^2} < 0$. As a result, the recommendation strategy varies faster for the employee with more precise information $\rho_s^*(s,t) > \rho_s^*(s,t')$ if $s > \mu$, which is always the case when $\mu \leq 0$. Consequently, one can show that the threshold $\underline{s}(t)$, below which the firm can infer signals in an LSHP equilibrium, is lower for the employee with more precise information when $\mu \leq 0$ and γ is sufficiently low. If $s > \mu$, then there are two opposing forces, as $\lambda(t)$ is decreasing in σ_ϵ^2 but the posterior mean $E[\theta|s,t]$ is increasing in σ_ϵ^2. For $\mu \geq 1$, it can be shown that the impact of posterior mean dominates and the recommendation strategy varies faster for the employee with less precise information $\rho_s^*(s,t) > \rho_s^*(s,t')$. For $\mu \in (0,1)$, whether the recommendation strategy varies faster for the employee with more or

[11]The remaining extreme cases are trivial. As $\sigma_\epsilon^2 \to 0$ (or $\sigma_\epsilon^2 \to \infty$), all ties are equally informative (either completely precise or imprecise); therefore, $\rho(s,t) = \rho(s,t')$ for all s, t, $t' \neq t$. As $\sigma_\theta^2 \to 0$, there is only one ability level, and so there is no uncertainty in the model.

less precise information depends on the values of exo-signal—if $min\left\{\underline{s}(t), \underline{s}(t')\right\} > \mu$, then $\rho_s^*(s,t) < \rho_s^*(s,t')$ when $s < s_l$, and $\rho_s^*(s,t) > \rho_s^*(s,t')$ when $s > s_r$. For this reason, one cannot determine whether the threshold $\underline{s}(t)$, below which the firm can infer signals in an LSHP equilibrium, is lower or higher for the employee with more precise information. Thus, the firm's relative capacity to infer exo-signals from two employees with different precision levels is ambiguous in such case.

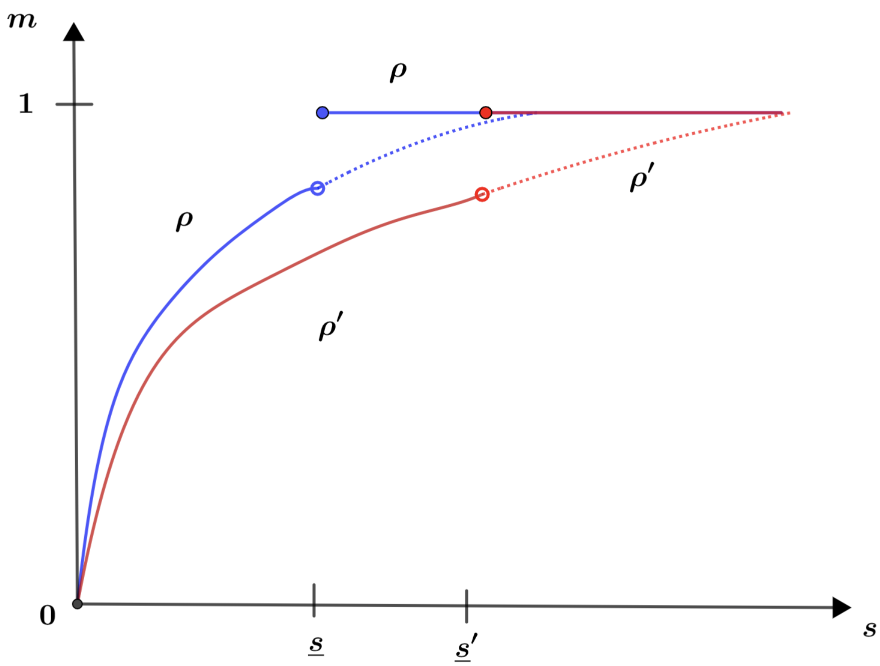

FIGURE 2. *Precision of Exo-Signals*

Blue (red) solid curve represents the recommendation strategy ρ (ρ') of a more (less) informative employee, the one with more (less) precise exo-signals. For both employees, dotted curves are the separating functions.

Even though exo-signals of the more informative employee are more precise, Theorem 3 shows that the firm can infer fewer exo-signals from him, and so the less informative employee may end up providing more information to the firm. As an illustration, consider Figure 2 above. For $\mu = 0$, recommendation strategies of two employees, as a function of their exo-signals, is represented by two solid curves. Blue (red) solid curve represents the recommendation strategy ρ (ρ') of a more (less) informative employee, the one with more (less) precise exo-signals. Observe the separating function ρ^* coincides with the recommendation strategy ρ on $[0, \underline{s})$, and then extends as the dotted curve up to $\overline{s}$.[12] Similar structure follows for the less informative employee with recommendation strategy ρ' and the interval $[0, \underline{s}')$. Since the threshold value for the more informative employee is lower $\underline{s} < \underline{s}'$ (for γ sufficiently low), the interval of signals that the firm can infer from the more informative

[12]Note that by definition $E\left[\theta|s_{min}, t\right] = 0$, which implies $s_{min}(t) = -\frac{(1-\lambda(t))}{\lambda(t)}\mu = 0$.

employee is strictly included in the interval of signals that the firm can infer from the less informative employee $[0, \underline{s}) \subset [0, \underline{s}')$.

The second part of Theorem 3 shows that when there is a lack of prior $\sigma_\theta^2 \to \infty$, the recommendation strategy does not depend on the precision of employee's signal, but the distribution of the pool at the top does. In particular, the posterior distribution of θ at the pool $[\underline{s}, 1]$ for the employee with more precise exo-signal second-order stochastically dominates (SOSD) the one for the employee with less precise exo-signals. This follows from the fact that the precision of employee's exo-signals do not affect the posterior mean $E[\theta|s,t] = s$ when there is a lack of prior, and so the pool at the top only reflects their relative precision levels.

4 Strategic Selections & Applications

I now develop two applications of the referrals model which consider the strategic "selections" of ties and hiring channels by applicants.

4.1 Tie Selection: Weak Versus Strong Ties

Consider an environment with two types of ties, $t \in \{t_{weak}, t_{strong}\}$, where t_{weak} and t_{strong} denote weak and strong ties, respectively. An applicant (or a job seeking worker) is now characterized by both his ability and his set of social ties (θ, T). There is a continuum of abilities uniformly distributed on $[0 - 2\alpha, 1 + 2\alpha]$ with $\alpha > 0$. An employee observes a noisy signal $s_{weak} \sim U[\theta - \alpha, \theta + \alpha]$ when his tie is weak, whereas he observes a precise signal $s_{strong} = \theta$ when his tie is strong. The set of ties that an applicant has, T, is distributed independently from abilities,

$$T = \begin{cases} \{t_{weak}, t_{strong}\} & \text{with prob } \delta \\ \{t_{weak}\} & \text{with prob } 1 - \delta \end{cases},$$

i.e. all applicants have a weak tie but only a $\delta \in [0, 1]$ proportion of them have a strong tie as well.

Each referral involves three stages. Given the set of ties T, an applicant chooses the type of tie $t \in T$ that maximizes his wage. After observing an exo-signal s_t, the employee with the selected type of tie sends a recommendation message m to the firm. The firm observes the recommendation message and the type of tie selected; then, it offers a wage equal to its valuation $w = E[\theta|m, t]$. All aspects of the game are common knowledge, except the values of θ, s_t, and T—the applicant knows his own ability θ, the exo-signal s_t, and the set of ties he has T; the employee only knows the exo-signal s_t; the firm knows none of them. Even though the firm does not know whether an applicant had the option to choose between the two types of ties, it knows the tie strength between its employee and the applicant, and it also knows that a fraction of applicants have access to both types of ties, which then allows it to form rational expectations.

Remarks

Both the precision of exo-signals and "selections" (of ties and hiring channels) have effects on the recommendation strategies. An important implication of Theorem 3 is that recommendation strategies do not vary with the precision of exo-signals when there is a lack of prior knowledge about the abilities, which is the case with a uniform distribution, and so I now consider a uniform distribution of abilities in order to isolate the effects of "selections" on recommendation strategies. For consistency, I assume that exo-signals are also uniformly distributed and that abilities are distributed on $[0 - 2\alpha, 1 + 2\alpha]$. Then, it can be shown that $s_{min}(t_{weak}) \geq 0$, and so for any exo-signal $s_{weak} \geq s_{min}(t_{weak})$, the employee can always suppose that abilities are on the interval $[s_{weak} - \alpha, s_{weak} + \alpha]$.

To keep the presentation neat, I assume that everyone has weak ties. In addition, since the set of ties T are distributed independently from abilities, its distribution is characterized by a single parameter δ.

Strength of Weak Ties Hypothesis

I will now examine the "strength of weak ties" hypothesis, which implies that weak ties lead to better job market outcomes. To this end, I will define the wage differential between weak and strong ties at θ level of ability as $WG_t(\theta) = [w(t_{weak}, \theta) - w(t_{strong}, \theta)]$, where $w(t, \theta)$ is the average wage at the given ability level when the type of tie is t.

The starting point of my analysis is to consider how much information firms can infer from weak versus strong ties. If firms could always infer exo-signals of both ties, then indeed returns to weak ties would be higher, as predicted by Granovetter. This follows from the fact that exo-signals of weak ties are imprecise. Consequently, some applicants get "lucky", they send high exo-signals to their weak ties, which then allows them to pool with a fraction of higher ability applicants that don't have strong ties and got "unlucky" by sending low exo-signals to their weak ties. Since weak ties allow applicants to pool with those above them, returns to weak ties would always be higher. However, firms cannot infer sufficiently high exo-signals by the structure of LSHP equilibria, which then lowers the relative returns to using weak ties for higher ability applicants—because such applicants can also use strong ties to pool with those above them via "pool at the top" and returns from pooling with weak ties via "pool at the top" can be lower for them.

As a result, there are two cases to consider. If firms could infer more exo-signals from strong ties, then the wage differential between weak and strong ties would vary monotonically with abilities—for low ability applicants, returns to weak ties are higher because this leads to firms pooling them with some higher ability applicants (those that don't have strong ties), whereas for sufficiently high ability applicants, returns to strong ties are higher because because this leads to the firm inferring that they have higher abilities. On the other hand, if firms could infer more exo-signals from weak ties, then the wage differential between weak and strong ties would vary non-monotonically with abilities—for sufficiently low and high ability applicants, returns to weak ties are higher, whereas for medium ability applicants returns to strong ties are higher. For sufficiently low ability applicants, the argument is same as before, firms can infer exo-signals but applicants can use weak ties to pool with those above them because exo-signals of weak ties are imprecise. Since firms can infer fewer exo-signals from strong ties, applicants can now also use strong ties to pool with those above them using the "pool at the top" (which is given by the structure of LSHP equilibria). For medium ability applicants, it can be shown that the returns to pooling with strong ties (via "pool at the top") are higher than pooling with weak ties (via imprecision of exo-signals). Sufficiently high ability applicants select ties based on the relative returns to "pools at the top". Since firms can infer fewer exo-signals from strong ties, it can be shown that applicants get higher returns from "pool at the top" of weak ties than the one for strong ties.

Moreover, it can be shown that firms infer more exo-signals from weak ties when reputation costs are sufficiently high (γ is sufficiently low). This is surprising because one would expect reputation costs to be sufficiently high for referrals to be informative to firms; consequently, one should also expect the implications of the second case to be present in data—returns to weak ties are higher for sufficiently low and high ability applicants, whereas returns to strong ties are higher for medium ability applicants—, which provides a new testable finding. The following proposition summarizes how recommendation strategies of employees and tie selection strategies of applicants influence each other, as well as their influence on LSHP equilibria and labor market returns to ties.

PROPOSITION 1. *There are two distinct cases.*

(1) If $\underline{s}(t_{weak}) < \underline{s}(t_{strong})$*, then* $E\left[\theta | s_{weak}, t_{weak}\right] = s_{weak} - \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right]$ *and there is a cut-off ability* θ^* *such that*

(i) $WG_t(\theta) > 0$ *if* $\theta < \theta^*$*, whereas*

(ii) $WG_t(\theta) < 0$ *if* $\theta > \theta^*$.

(2) *If* $\underline{s}(t_{weak}) > \underline{s}(t_{strong})$, *then*

$$E[\theta|s_{weak}, t_{weak}] = \begin{cases} s_{weak} - \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right] & \text{if } s_{weak} < \underline{s}(t_{strong}) \\ s_{weak} & \text{otherwise} \end{cases}$$

and there is a cut-off pair (θ_l^*, θ_r^*) *such that*
(i) $WG_t(\theta) > 0$ *if* $\theta < \theta_l^*$ *and* $\theta > \theta_r^*$, *whereas*
(ii) $WG_t(\theta) < 0$ *if* $\theta \in [\theta_l^*, \theta_r^*]$.

For γ *sufficiently low,* $\underline{s}(t_{weak}) > \underline{s}(t_{strong})$.

Observe that both recommendation (via (DE)) and tie selection strategies (via wage offers) depend on $E[\theta|s,t]$. Since $E[\theta|s_{strong}, t_{strong}]$ is always equal to s_{strong}, the relationship between recommendation and tie selection strategies ultimately depend on $E[\theta|s_{weak}, t_{weak}]$. The discussion preceding Proposition 1 describes how recommendation strategies can influence returns to ties, which in turn determines the tie selection strategies in that applicants select the type of tie that gives them higher returns—conditional on having access to such type of tie. Conversely, I will now describe how tie selections can influence recommendation strategies. Consider Figure 3 below, as it accompanies the discussion that follows. For γ sufficiently low, recommendation strategies of two employees, as a function of their exo-signals, is represented by two solid curves. Red and blue solid curves represent the recommendation strategies of weak and strong ties, respectively.[13] For signals below $\underline{s}$, employees know that applicants are using weak ties to pool with those above them, and so employees with weak ties will recommend lower than otherwise. This can be seen by the fact that employees deflate their valuations $E[\theta|s_{weak}, t_{weak}] = s_{weak} - \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right]$, which in turn leads to recommendation strategies that varies slower than those of strong ties. For signals above $\underline{s}$, employees know that applicants will use strong ties to pool (via "pool at the top") with those above them; consequently, employees with weak ties will not deflate their valuations $E[\theta|s_{weak}, t_{weak}] = s_{weak}$, and their recommend strategies will vary faster than before. This sudden change in recommendation strategy of weak ties is solely due to "tie selection" aspect; this can be seen by comparing it to figure 2 in section 3, which also compares recommendation strategies of two employees with different levels of informativeness (precision of exo-signals) but with no strategic selections of ties by applicants.

[13] Note that by definition $E[\theta|s_{min}, t] = 0$, which implies that $s_{min}(t_{strong}) = 0$ and $s_{min}(t_{weak}) = \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right]$.

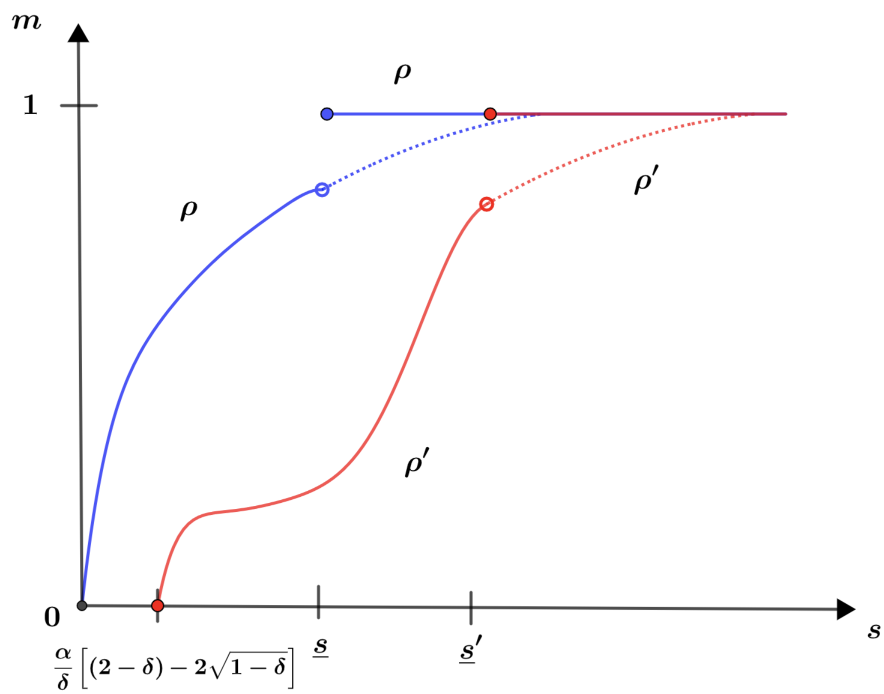

FIGURE 3. *Tie Selection – Weak vs Strong*

Red and blue solid curves represent the recommendation strategies of weak and strong ties, respectively. For both types, dotted curves are the separating functions.

Proposition 1 shows that returns to weak versus strong ties vary with the ability of applicants, which in turn can help reconcile Granovetter's "strength of weak ties" theory with the opposing empirical findings. The extent to which applicants use weak and strong ties ultimately depends on the access to strong ties δ, which determines the returns from pooling with weak ties. Contrary to the existing explanation, the frequent use of weak ties may not be due to its efficiency in the matching of workers and firms. When the access to strong ties is really scarce for workers, many workers—even those with access to strong ties—use weak ties to pool with even higher ability workers. Since the relative use of weak and strong ties vary with the access to strong ties, this can then help explain the mixed evidence about the use of different types of ties. In addition, the relationship between tie strength and wages can be negative for sufficiently low and high ability workers but positive for medium ability workers, which can help explain why some researchers found no relationship between tie strength and wages (when considering workers of all abilities). Although these predictions do not emerge in the existing models of employee referrals, they are consistent with the empirical evidence, suggesting that strategic recommendations by employees and tie selections by workers are important aspects of employee referrals.

4.2 Hiring Channel Selection: Formal Versus Informal Channels

Consider an environment with two types of hiring channels, $h \in \{h_t, h_0\}$, where h_t and h_0 denote social ties and formal market channels, respectively. An applicant (or a job seeking worker) is now characterized by both his ability and his set of hiring channels (θ, H). As before, there is a continuum of abilities uniformly distributed on $[0 - 2\alpha, 1 + 2\alpha]$ with $\alpha > 0$. However, now both employees and firms can observe signals about the abilities. Employees observe precise signals $s_t = \theta$, whereas firms observe noisy signals $s_0 \sim U[\theta - \alpha, \theta + \alpha]$. The set of channels that an applicant has, H, is distributed independently from abilities,

$$H = \begin{cases} \{h_t, h_0\} & \text{with prob } \delta \\ \{h_0\} & \text{with prob } 1 - \delta \end{cases},$$

i.e. all applicants have access to the formal market channel but only a $\delta \in [0, 1]$ proportion of them have a social tie as well.

Each job market application involves three stages. Given the set of hiring channels H, an applicant chooses the type of channel $h \in H$ that maximizes his wage. If the formal market channel is selected, then the firm observes an exo-signal s_0 and offers a wage equal to its valuation $w = E[\theta|s_0]$. Conversely, if the referral channel is selected, then the employee observes an exo-signal s_t and sends a recommendation message m to the firm. The firm observes the recommendation message and offers a wage equal to its valuation $w = E[\theta|m]$. All aspects of the game are common knowledge, except the values of θ, s_t, s_0, and H—the applicant knows his own ability θ, observes the exo-signal received by the selected channel (s_t if $h = h_t$ or s_0 if $h = h_0$), and his set of hiring channels H; the employee observes the exo-signal s_t if $h = h_t$; the firm observes the exo-signal s_0, if $h = h_0$.

Remarks

To fix ideas, noisy signals received by firms represent the presence of formal screening mechanisms, such as aptitude tests and other attribute measurements, which are typically used in a formal hiring process to acquire some information about the abilities of applicants. Conversely, recommendation messages received by firms represent referrals, which is the screening mechanism used in an "informal hiring" (via social ties).

Recall, this application was built to show why referrals are used less by high ability applicants, which is one of the most robust empirical findings in the referrals literature. Since sufficiently high ability applicants want to reveal their abilities, if high ability applicants prefer to not use referrals even when employees receive precise signals, then surely they will not use them when employees receive noisy signals. As a result, the assumption that employees observe precise signals is innocuous. Indeed, one can relax this assumption by considering the case where both employees and firms receive noisy signals; nothing of substance will change.

One can consider more involved firms that use formal screening mechanisms even during the informal hiring process. Then, its valuation is formed using both exo-signal s_0 and recommendation message m when the referral channel is selected, $w = E[\theta|m, s_0]$. Such valuation will be a weighted average of $E[\theta|m]$ and $E[\theta|s_0]$; therefore, if $E[\theta|m] < E[\theta|s_0]$ for sufficiently high ability applicants, then surely $E[\theta|m, s_0] < E[\theta|s_0]$ for them as well. Consequently, nothing of substance will change under this more sophisticated environment.

Following the same logic as the first application, I assume that abilities and exo-signals are uniformly distributed and that the set of hiring channels H is distributed independently from abilities. However, the two applications differ in that firms can always observe exo-signals s_0 and there is no pooling at the top with such exo-signals.

Screening Methods: Formal Screening vs Referrals

I will now examine how hiring channels (formal market vs social tie) and their corresponding screening methods (formal screening vs referral) influence job market outcomes, and the decision of workers to choose different types of hiring channels based on their returns in the labor market. To this end, I will define the wage differential between social tie and formal market channels at θ level of ability as $WG_h(\theta) = [w(h_t, \theta) - w(h_0, \theta)]$, where $w(h, \theta)$ is the average wage at the given ability level when the hiring channel is h.

The wage differential between social tie and formal market channels vary non-monotonically with abilities—returns to ties are weakly negative for sufficiently low ability applicants, positive for medium ability applicants, and negative for sufficiently high ability applicants. Sufficiently low ability applicants can use the formal market channel to pool with those above them and get higher returns. This is because exo-signals of formal screening mechanisms are imprecise, whereas exo-signals of social ties are precise and firms can also infer them (through the recommendations of its employees) for sufficiently low abilities. Note that this is purely driven by the relative precision of the two screening mechanisms; in fact, returns to ties will be zero when formal screening and social ties are equally precise ($\alpha = 0$). However, the following results are completely robust to the relative precision of the two screening mechanisms. By the structure of LSHP equilibria, firms cannot infer exo-signals of its employees when applicants have higher abilities. Consequently, applicants can now also use social ties to pool with those above them using the "pool at the top". For medium ability applicants, it can be shown that the returns to pooling with ties (via "pool at the top") are higher than pooling with the formal market channel (via imprecision of exo-signals). Since sufficiently high ability applicants want to reveal their abilities, they prefer to avoid the "pool at the top" and get higher returns from using the formal market channel. The following proposition summarizes these results for the two hiring channels (formal market vs social tie) and their corresponding screening methods (formal screening vs referral).

PROPOSITION 2. *There is a cut-off pair* $(\theta_l^{**}, \theta_r^{**})$ *and* $E[\theta|s_0, h_0] = s_0 - \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right]$ *such that*

(1) $WG_h(\theta) = -2\frac{\alpha}{\delta}\sqrt{1-\delta}$ *if* $\theta < \theta_l^{**}$,

(2) $WG_h(\theta) > 0$ *if* $\theta \in [\theta_l^{**}, \theta_r^{**}]$, *and*

(3) $WG_h(\theta) < 0$ *if* $\theta > \theta_r^{**}$.

For δ *sufficiently low,* $\theta_r^{**} < 1$.

Proposition 2 shows that for sufficiently high ability applicants, returns to using the formal market channel is always higher than social ties—this is consistent with one of the most robust empirical findings in the referrals literature, which is that referrals are used relatively less by high ability applicants. In my model, even though social ties receive more precise information than what firms receive through formal screening mechanisms, the inability of firms to infer exo-signals from the recommendations of its employees makes formal screening mechanisms more informative to firms and more desirable for sufficiently high ability applicants. Observe that high ability applicants want to reveal their abilities; since they prefer to not use referrals even when employees receive precise signals, they will surely not use them when employees receive noisy signals. Thus, this result holds regardless of the relative precisions of the two screening mechanisms.

In addition, the empirical evidence on the relative returns to using these two channels (formal market vs social tie) is mixed. In my model, this result follows from the fact that returns to using social ties vary with the ability of applicants. For sufficiently high ability applicants, returns to using social ties are always lower (due to LSHP communication structure), but the returns to using social ties are higher for some of the lower ability applicants.

5 Conclusion

To recap, when there is asymmetric information about abilities of applicants, employees can "signal" some of this information to firms by recommending applicants they personally know. I consider how employees strategically transmit information to firms when there are both gratitude benefits for recommendations and reputation costs for providing inaccurate information. My analysis shows that referrals are not effective in transmitting information about high ability applicants. Since the effectiveness of referrals vary with the ability of applicants, so will the returns to using them. This allows me to help explain mixed evidences about the returns to using different types of ties (weak vs strong) and hiring channels (formal market vs social tie), as well as the relatively higher use of social ties by low ability applicants.

I will now conclude by discussing some ideas for future research.

My analysis shows that the wage differential between weak and strong ties vary non-monotonically with abilities—returns to weak ties are higher for sufficiently low and high ability applicants, whereas returns to strong ties are higher for medium ability applicants—, which is a novel and testable insight for future empirical work. As discussed before, there is an interesting debate about the "strength of weak ties" hypothesis, and an empirical test of my finding would add to this literature.

Although I consider the strategic role of employees in transmitting information about applicants to firms (helping screen applicants), employees can also be strategic in how they spread information about new job vacancies to job seeking workers (reducing search frictions). In the existing literature, employed workers pass information about new job vacancies to one of their social contacts at random. In reality, individuals have preferences over their social contacts—partly because who they pass their information to can effect their opportunities in the future. To this end, one could look deeper at the pattern of relationships between workers (employed and unemployed). See Calvo (2004) and Calvo-Armengol & Jackson (2004) for models that consider the complex network structure of job contacts—but employees are non-strategic. In particular, the former considers the formation of such networks and the latter considers their influence on economic outcomes. Following their seminal works, the existing literature has only considered the network structure when considering the role of referrals as a search frictions reducing mechanism. At this point, little is understood about the formation of job contact networks and how they influence economic outcomes when considering the role of referrals as a screening mechanism. It seems to me that exploring these two aspects—network structures and strategic employees—further are fruitful avenues for future research.

Finally, I would like to point out that although my focus is on "employee referrals", in my model referees do not need to be employees per se. For instance, advisors giving referrals for their students in academia would also face similar trade-offs. Indeed, my model encompasses any situation in which referees face both gratitude benefits for recommendations and reputation costs for providing inaccurate information.

References

[1] Ambrus, Attila, and Shih En Lu, "Almost Fully Revealing Cheap Talk with Imperfectly Informed Senders," *Games and Economic Behavior*, 88 (2014), 174–189.

[2] Arrow, Kenneth J. and Ron Borzekowski, "Limited Network Connections and the Distribution of Wages," *FEDS Working Paper,* 2004.

[3] Battaglini, Marco, "Policy Advice with Imperfectly Informed Experts," *Advances in Theoretical Economics*, 4 (2004).

[4] Bian, Yanjie, "Bringing Strong Ties Back-In: Indirect Ties, Network Bridges, and Job Searches in China," *American Sociological Review*, 62 (1997), 366–385.

[5] Bian, Yanjie, "Getting a Job through a Web of Guanxi in China," *In Networks in the Global Village: Life in Contemporary Communities*, B. Wellman, ed. (Boulder, CO: Westview Press, 1999).

[6] Boorman, Scott A, "A Combinatorial Optimization Model for Transmission of Job Information Through Contact Networks," *Bell Journal of Economics*, 6 (1975), 216–49.

[7] Bridges, William P., and Wayne J. Villemez, "Informal Hiring and Income in the Labor Market," *American Sociological Review*, 51 (1986), 574–582.

[8] Calvó-Armengol, Antoni, "Job Contact Networks," *Journal of Economic Theory*, 115 (2004), 191–206.

[9] Calvó-Armengol, Antoni and Matthew O. Jackson, "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 94 (2004), 426–454.

[10] Calvó-Armengol, Antoni and Yves Zenou, "Job Matching, Social Network and Word-of-Mouth Communication," *Journal of Urban Economics*, 57 (2005), 500–522.

[11] Calvó-Armengol, Antoni and Matthew O. Jackson, "Networks in Labor Markets: Wage and Employment Dynamics and Inequality," *Journal of Economic Theory*, 132 (2007), 27–46.

[12] Calvó-Armengol, Antoni, Thierry A. Verdier, and Yves Zenou, "Strong and Weak Ties in Employment and Crime," *Journal of Public Economics,* 91 (2007), 203–233.

[13] CareerBuilder, *Referral Madness: How Employee Referral Programs Turn Good Employees into Great Recruiters and Grow Your Bottom Line* (CareerBuilder e-Book, 2012).

[14] Cho, In-Koo, and David M. Kreps, "Signalling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102 (1987), 179–221.

[15] Cho, In-Koo, and Joel Sobel, "Strategic Stability and Uniqueness in Signalling Games," *Journal of Economic Theory*, 50 (1990), 381–418.

[16] Corcoran, Mary, Linda Datcher, and Greg J. Duncan, "Information and Influence Networks in Labor Markets," *In Five Thousand American Families: Patterns of Economic Progress*, Greg J. Duncan, and James N. Morgan, ed. (Ann Arbor, MI: Institute For Social Research, 1980).

[17] Datcher, Linda, "The Impact of Informal Networks on Quit Behavior," *Review of Economics and Statistics*, 65 (1983), 491–495.

[18] Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker, "Referral-Based Job Search Networks," *Review of Economic Studies*, 83 (2016), 514–546.

[19] Elliott, James R, "Social Isolation and Labor Market Insulation: Network and Neighborhood Effects on Less-Educated Urban Workers," *Sociological Quarterly*, 40 (1999), 199–216.

[20] Galenianos, Manolis, "Learning About Match Quality and the Use of Referrals," *Review of Economic Dynamics*, 16 (2013), 668–690.

[21] Galeotti, Andrea and Luca Paolo Merlino, "Endogenous Job Contact Networks," *International Economic Review*, 55 (2014), 1201–1226.

[22] Green, Gary P., Leann M. Tigges, and Irene Browne, "Social Resources, Job Search, and Poverty in Atlanta," *Research in Community Sociology*, 5 (1995), 161–182.

[23] Green, Gary, Leann Tigges, and Daniel Diaz, "Racial and Ethnic Differences in Job Search Strategies in Atlanta, Boston, and Los Angeles," *Social Science Quarterly,* 80 (1999), 559–77.

[24] Granovetter, Mark, "The Strength of Weak Ties," *American Journal of Sociology*, 78 (1973), 1360–80.

[25] Granovetter, Mark, *Getting a Job: A Study of Contacts and Careers.* (Chicago, Illinois: The University of Chicago Press, 1995).

[26] Hörner, Johannes, "Signalling And Screening," *The New Palgrave Dictionary of Economics*, 2nd ed. (Basingstoke, UK: Palgrave Macmillan, 2006).

[27] Holzer, Harry, "Job Search by Employed and Unemployed Youth," *Industrial and Labor Relations Review,* 40 (1987), 601–11.

[28] Ioannides, Yannis M., and Linda Datcher Loury, "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42 (2004), 1056–93.

[29] Kartik, Navin, "Strategic Communication with Lying Costs," *Review of Economic Studies*, 76 (2009), 1359–1395.

[30] Lai, Gina, Nan Lin, Shu-Yin Leung, "Network Resources, Contact Resources, and Status Attainment," *Social Networks*, 20 (1998), 159–178.

[31] Lin, Nan, Walter M. Ensel, and John C. Vaughn, "Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment," *American Sociological Review,* 46 (1981), 393–405.

[32] Lu, Shih En, "Coordination-Free Equilibria in Cheap Talk Games," *Journal of Economic Theory*, 168 (2017), 177–208.

[33] Marmaros, David, and Bruce Sacerdote, "Peer and Social Networks in Job Search," *European Economic Review,* 46 (2002), 870–79.

[34] Marsden, Peter V., and Jeanne S. Hurlbert, "Social Resources and Mobility Outcomes: A Replication and Extension." *Social Forces*, 66 (1988), 1038–1059.

[35] Marsden, Peter V., and Elizabeth H. Gorman, "Social Networks, Job Changes and Recruitment," *In Sourcebook on labor markets: Evolving structures and processes*, Ivar Berg and Arne L. Kalleberg, ed. (New York: Kluwer Academic/Plenum, 2001).

[36] Marx, Jonathan, and Kevin Leicht, "Formality of Recruitment to 229 Jobs: Variation by Race, Sex and Job Characteristics," *Sociology and Social Research*, 76 (1992), 190–196.

[37] Montgomery, James D, "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review*, 81 (1991), 1408–1418.

[38] Montgomery, James D, "Weak Ties, Employment, and Inequality: An Equilibrium Analysis," *American Journal of Sociology*, 99 (1994), 1212–1236.

[39] Murray, Stephen O., Joseph H. Rankin, and Dennis W. Magill, "Strong Ties and Job Information," *Sociology of Work and Occupation*, 8 (1981), 119–136.

[40] Ornstein, Michael D, *Entry into the American Labor Force*. (New York, NY: Academic Press, 1976).

[41] Rees, Albert, "Information Networks in Labor Markets," *American Economic Review*, 56 (1966), 559–566.

[42] Rees, Albert, and George Pratt Schultz, *Workers and Wages in an Urban Labor Market.* (Chicago: University of Chicago Press, 1970).

[43] Riley, John G, "Silver Signals: Twenty-Five Years of Screening and Signalling," *Journal of Economic Literature,* 39 (2001), 432–78.

[44] Rosenbaum, James, Stefanie DeLuca, Shazla Miller, and Kevin Roy, "Pathways into work: Short- and Long-Term Effects of Personal and Institutional Ties." *Sociology of Education*, 72 (1999), 179–96.

[45] Saloner, Garth, "The Old Boys' Network as a Screening Mechanism," *Journal of Labor Economics*, 3 (1985), 255–267.

[46] Simon, Curtis, and John Warner, "Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure," *Journal of Labor Economics*, 10 (1992), 306–30.

[47] Spence, A. Michael, "Job Market Signalling," *Quarterly Journal of Economics*, 87 (1973), 355–74.

[48] Topa, Giorgio, "Labor Markets and Referrals." *In Handbook of social economics*, Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, ed. (Amsterdam: Elsevier, 2011).

[49] Watanabe, Shin, "A Comparative Study of Male Employment Relations in the United States and Japan, " *UCLA Working Paper*, 1987.

[50] Yakubovich, Valery, "Weak Ties, Information, and Influence: How Workers Find Jobs in a Local Russian Labor Market," *American Sociological Review*, 70 (2005), 408–421.

[51] Zenou, Yves, "Spatial Versus Social Mismatch," *Journal of Urban Economics*, 74 (2013), 113–132.

[52] Zenou, Yves, "A Dynamic Model of Weak and Strong Ties in the Labor Market," *Journal of Labor Economics*, 33 (2015), 891–932.

Appendix A: Proofs

Proof of Lemma 1.

By hypothesis, types (s_l, s_h) are separating in a monotone equilibrium with mapping ρ. If ρ is constant on some interval $I \subseteq (s_l, s_h)$, then some $s \in I$ can profitably deviate and mimic a slightly higher type $s + \epsilon \in I$, contradicting equilibrium separation. Thus, ρ must be strictly increasing on (s_l, s_h), which implies that $\rho(s) \in (0,1)$ for any $s \in (s_l, s_h)$. Separation requires that s must be a solution to (1) for each $s \in (s_l, s_h)$. If ρ is differentiable on (s_l, s_h), then the first order condition obtained by differentiating (1) is (DE), as required. So it suffices to prove that ρ is differentiable on (s_l, s_h). This is done in series of steps.

CLAIM 1. *$\rho(s) \neq E\left[\theta|s\right]$ for all $s \in (s_l, s_h)$.*

Suppose there exists a $\widehat{s} \in (s_l, s_h)$ such that $\rho(\widehat{s}) = E[\theta|\widehat{s}]$. Without loss of generality, one can assume that $E[\theta|\widehat{s}] \in (0,1)$, because otherwise ρ cannot be strictly increasing on (s_l, s_h). This implies that

$$\frac{\partial E\left[\left(\rho\left(\widehat{s}\right) - \theta\right)^2 |\widehat{s}\right]}{\partial \rho} = 0. \tag{5}$$

Define $g(\varepsilon)$ as the expected utility gain for a type $\widehat{s} - \varepsilon$ by deviating from $\rho\left(\widehat{s} - \varepsilon\right)$ to $\rho(\widehat{s})$. Since I am on a separating portion of the type space when $|\varepsilon|$ is small,

$$\begin{aligned} g(\varepsilon) = &\left\{\gamma E\left[\theta|\widehat{s}\right] - (1-\gamma) E\left[\left(\rho\left(\widehat{s}\right) - \theta\right)^2 |\widehat{s} - \varepsilon\right]\right\} - \\ &\left\{\gamma E\left[\theta|\widehat{s} - \varepsilon\right] - (1-\gamma) E\left[\left(\rho\left(\widehat{s} - \varepsilon\right) - \theta\right)^2 |\widehat{s} - \varepsilon\right]\right\}. \end{aligned}$$

As $E\left[\left(\rho(\widehat{s} - \varepsilon) - \theta\right)^2 |\widehat{s} - \varepsilon\right] \geq E\left[\left(E\left[\theta|\widehat{s} - \varepsilon\right] - \theta\right)^2 |\widehat{s} - \varepsilon\right]$,

$$\begin{aligned} g(\varepsilon) \geq \phi(\varepsilon) := &\left\{\gamma E\left[\theta|\widehat{s}\right] - (1-\gamma) E\left[\left(\rho\left(\widehat{s}\right) - \theta\right)^2 |\widehat{s} - \varepsilon\right]\right\} - \\ &\left\{\gamma E\left[\theta|\widehat{s} - \varepsilon\right] - (1-\gamma) E\left[\left(E\left[\theta|\widehat{s} - \varepsilon\right] - \theta\right)^2 |\widehat{s} - \varepsilon\right]\right\}. \end{aligned}$$

ϕ is differentiable in a neighbourhood of 0. Observe that $\phi(0) = 0$ and using (5), $\phi'(0) = \gamma\lambda > 0$. Hence, for small $\varepsilon > 0$, $g(\varepsilon) \geq \phi(\varepsilon) > 0$, implying that a type $\widehat{s} - \varepsilon$ strictly prefers to imitate $\widehat{s}$, contradicting equilibrium separation of $\widehat{s}$. $\square$

CLAIM 2. *$\rho(s) > E\left[\theta|s\right]$ for all $s \in (s_l, s_h)$.*

Suppose not. Since $\rho(s) \neq E\left[\theta|s\right]$ by the previous claim, there exists a $s' \in (s_l, s_h)$ such that $\rho(s') < E[\theta|s']$. By monotonicity of ρ, $\rho(s' - \varepsilon) < \rho(s')$ for all $\varepsilon > 0$. By convexity of the function, $E\left[\left(\rho\left(s' - \varepsilon\right) - \theta\right)^2 |s' - \varepsilon\right] > E\left[\left(\rho\left(s'\right) - \theta\right)^2 |s' - \varepsilon\right]$ for small enough $\varepsilon > 0$. On the other hand, $E\left[\theta|s' - \varepsilon\right] < E\left[\theta|s'\right]$ for small enough $\varepsilon > 0$. Therefore, for small enough $\varepsilon > 0$, a type $s' - \varepsilon$ strictly prefers to imitate s', contradicting equilibrium separation of s'.

□

CLAIM 3. *ρ is continuous on (s_l, s_h).*

Suppose there is a discontinuity at some $\widehat{s} \in (s_l, s_h)$. First, consider the case where $\rho(\widehat{s}) < lim_{s \downarrow \widehat{s}} \rho(s) =: \overline{\rho}$. By continuity of $E\left[(\rho-\theta)^2 |s\right]$ and monotonicity of ρ, as $\varepsilon \downarrow 0$,

$$E\left[(\rho(\widehat{s}+\varepsilon)-\theta)^2 |\widehat{s}+\varepsilon\right] - E\left[(\rho(\widehat{s})-\theta)^2 |\widehat{s}+\varepsilon\right] \to E\left[(\overline{\rho}-\theta)^2 |\widehat{s}\right] - E\left[(\rho(\widehat{s})-\theta)^2 |\widehat{s}\right] > 0,$$

where the inequality above follows from $\overline{\rho} > \rho(\widehat{s}) > E[\theta|\widehat{s}]$. On the other hand, $E[\theta|\widehat{s}+\varepsilon] - E[\theta|\widehat{s}] \to 0$ as $\varepsilon \downarrow 0$; hence, for small enough $\varepsilon > 0$, $\widehat{s}+\varepsilon$ prefers to imitate $\widehat{s}$, contradicting equilibrium separation. The argument for the other case where $\rho(\widehat{s}) > lim_{s \uparrow \widehat{s}} \rho(s)$ is similar, establishing that $\widehat{s}$ prefers to imitate $\widehat{s}-\varepsilon$ for small enough $\varepsilon > 0$. □

CLAIM 4. *ρ is differentiable on (s_l, s_h).*

Given the previous claims that ρ is continuous and does not intersect $E[\theta|s]$ on (s_l, s_h), one can replicate the arguments of Mailath (1987)—see the proof of proposition 2 in his appendix. □ □

The following lemma proves uniqueness of the separating function, and it is used subsequently in the proof of Theorem 1.

> LEMMA A1. *Fix any $m_0 \in [0,1]$. There is a unique solution to the problem of finding a $\overline{s} \in [s_{min}, s_{max}]$ and $\rho : [s_{min}, \overline{s}] \to [0,1]$ such that (i) ρ is strictly increasing and continuous on $[s_{min}, \overline{s}]$, (ii) $\rho(s_{min}) = m_0$, (iii) ρ solves (DE) on $(s_{min}, \overline{s})$, and (iv) $\rho(\overline{s}) = 1$.*
>
> *Proof of Lemma A1.*

As $2(1-\gamma)(\rho(s) - E[\theta|s]) = 0$ at $\rho(s) = E[\theta|s]$ for all s, there is no Lipschitz condition on (DE) on the relevant domain, in particular even locally if $m_0 = 0$. Thus, standard results on differential equations do not apply. Instead, I build on Mailath (1987) and proceed as follows.

STEP 1. *Local Existence and Uniqueness.*

Consider the inverse initial value problem to find $\tau(\widehat{m})$ such that

$$\tau' = g(\widehat{m}, \tau) := \frac{2(1-\gamma)(\widehat{m} - E[\theta|\tau])}{\gamma\lambda}, \qquad \tau(\widehat{m}_0) = s_{min}. \tag{6}$$

Observe that g is continuous and Lipschitz on $[0,1] \times [0,1]$. Thus, standard theorems (e.g. Coddington and Levinson (1955), Theorem 2.3, p.10) imply that there is a unique local solution $\widetilde{\tau}$ to (6) on $[\widehat{m}_0, \widehat{m}_0+\delta)$, for some $\delta > 0$; moreover, $\widetilde{\tau} \in \mathcal{C}^1\left([\widehat{m}_0, \widehat{m}_0+\delta)\right)$.[14] Inverting $\widetilde{\tau}$ gives a strictly increasing $\widetilde{\rho}$ on $[s_{min}, \widetilde{s})$ such that $\rho(s_{min}) = \widehat{m}_0$, $\widetilde{\rho}$ solves (DE) on $(s_{min}, \widetilde{s})$ and $\widetilde{\rho} \in \mathcal{C}^1\left([s_{min}, \widetilde{s})\right)$. Observe that $\widetilde{\rho}(s) > E\left[\theta|s\right]$ for all $s \in (s_{min}, \widetilde{s})$. As the inverse of any increasing local solution to (DE) is a local solution to (6) on $[0, \eta)$ for some $\eta > 0$, local uniqueness of an increasing solution to (DE) follows from the fact that $\widetilde{\tau}$ is the unique local solution to (6) above 0.

STEP 2. *$\widetilde{\rho}(s) > E\left[\theta|s\right]$ above s_{min}.*

Suppose $\widetilde{\rho}$ is a solution to (DE) on $[s_{min}, s_{min}+\delta)$, with $\widetilde{\rho} \in \mathcal{C}^1\left([s_{min}, s_{min}+\delta)\right)$ and $\widetilde{\rho}' > 0$. Let $m_\delta := lim_{s\uparrow s_{min}+\delta}\widetilde{\rho}(s)$. I claim that $m_\delta > E\left[\theta|s_{min}+\delta\right]$. This is obviously true if $\delta = 0$. So assume, for sake of contradiction, that $\delta > 0$ and $m_\delta \leq E\left[\theta|s_{min}+\delta\right]$. Since $\widetilde{\rho}(s) > E\left[\theta|s\right]$ for all $s \in (s_{min}, s_{min}+\delta)$, one must have $m_\delta = E\left[\theta|s_{min}+\delta\right]$ and $lim_{s\uparrow s_{min}+\delta}\widetilde{\rho}'(s) = \infty$. As $\widetilde{\rho} \in \mathcal{C}^1\left([s_{min}, s_{min}+\delta)\right)$, there exists a $s_1 \in (s_{min}, s_{min}+\delta)$ such that $\widetilde{\rho}'(s) > \lambda$ for all $s \in [s_1, s_{min}+\delta)$. Pick $\varepsilon > 0$ such that $\widetilde{\rho}(s_1) > E\left[\theta|s_1\right] + \varepsilon$. Then,

$$\begin{aligned} m_\delta &= \widetilde{\rho}\left(s_1\right) + lim_{s\uparrow s_{min}+\delta}\int_{s_1}^{s}\widetilde{\rho}'(y)dy \\ &> E\left[\theta|s_1\right] + \varepsilon + \int_{s_1}^{s_{min}+\delta}\lambda dy \\ &= E\left[\theta|s_{min}+\delta\right] + \varepsilon, \end{aligned}$$

which contradicts $m_\delta = E\left[\theta|s_{min}+\delta\right]$.

STEP 3. *Unique Extension of $\widetilde{\rho}$ to $[s_{min}, \overline{s})$.*

The proof is completed as follows. If $\widetilde{\rho}$ is defined on $[s_{min}, s_{min}+\delta)$ with $lim_{s\uparrow s_{min}+\delta}\widetilde{\rho}(s) < 1$, then by the previous step, $\frac{\gamma}{2(1-\gamma)(\rho(s)-E[\theta|s])}$ is continuous, Lipschitz and bounded in a neighbourhood of $(s_{min}+\delta, lim_{s\uparrow s_{min}+\delta}\widetilde{\rho}(s))$. Thus, standard theorems (e.g. Coddington and Levinson (1955), Theorem 4.1 and preceding discussion, p. 15) imply that there is a unique extension of $\widetilde{\rho}$ to $[s_{min}, s_{min}+\delta+\eta)$ for some $\eta > 0$. Moreover, this extension is strictly increasing, in $\mathcal{C}^1\left([s_{min}, s_{min}+\delta+\eta)\right)$, and by the previous step, $lim_{s\uparrow s_{min}+\delta+\eta}\widetilde{\rho}(s) > E\left[\theta|s_{min}+\delta+\eta\right]$. Now let $\overline{s} := sup\left\{s : \widetilde{\rho} \text{ can be extended to } [s_{min}, s)\right\}$. Clearly, $\overline{s} < s_{max}$. In addition, one must have $lim_{s\uparrow\overline{s}}\widetilde{\rho}(s) = 1$; otherwise, $\widetilde{\rho}$ can be extended beyond $\overline{s}$. I am done by setting $\widetilde{\rho}(\overline{s}) = 1$. $\square$

Proof of Theorem 1.

Suppose there is a separating monotone equilibrium (ρ, w). Using Lemma 1, ρ solves (DE) on (s_{min}, s_{max}). Let $m_0 := lim_{s\downarrow s_{min}}\rho(s)$. Since $\rho(s) > E\left[\theta|s\right]$ for all $s \in (s_{min}, s_{max})$,

[14] $\mathcal{C}^1([a,b))$ is the set of all real-valued functions on $[a,b)$ that have a continuous derivative at all $\widehat{m} \in (a,b)$ and in addition have a right-hand derivative at a that is continuous from the right at a.

$m_0 \geq E[\theta|s_{min}] = 0$. However, using Lemma A1, there is no solution on $[s_{min}, s_{max})$ to the initial value problem given by $\rho(s_{min}) = m_0$ and (DE) on (s_{min}, s_{max}), a contradiction. □

The following lemma is used in the proof of Theorem 2.

LEMMA A2. *There exists* $\underline{s} \in [s_{min}, \overline{s}]$ *that solves (3) and (4).*

Moreover, if $\gamma < \frac{1}{2}$, *this is true with some* $\underline{s} > s_{min}$.

Proof of Lemma A2.

Start by defining the function

$$\Delta(w,s) := \gamma w - (1-\gamma)\, E\left[(1-\theta)^2\,|s\right] - \left\{\gamma E\left[\theta|s\right] - (1-\gamma)\, E\left[(\rho^*(s)-\theta)^2\,|s\right]\right\}.$$

$\Delta(w,s)$ is the utility gain for type s from sending the highest recommendation message ($\rho = 1$) and receiving the wage w over separating itself (thus receiving wage $E[\theta|s]$) with a recommendation message $\rho^*(s)$. Clearly, Δ is continuous in both arguments, and strictly increasing in w. Observe that, in an LSHP equilibrium, the gain from pooling over separating can be no more than $\Delta(1,s)$ and will generally be strictly less. There are two conceptually distinct cases: one where $\Delta(1,s) = 0$ for some $s \leq \overline{s}$, and the other where $\Delta(1,s) > 0$ for all $s \leq \overline{s}$. Let

$$\widehat{s} := \begin{cases} s_{min} & \text{if } \Delta(1,s) > 0 \text{ for all } s \leq \overline{s} \\ sup_{s\in[s_{min},\overline{s}]}\left\{s : \Delta(1,s) = 0\right\} & \text{otherwise.} \end{cases}$$

Observe that $\Delta(1,\overline{s}) > 0$; hence, $\widehat{s} < \overline{s}$ and for all $s \in (\widehat{s}, \overline{s}]$, $\Delta(1,s) > 0$. Let $\widetilde{w}(s)$ be the wage for type s that makes him indifferent between sending the highest recommendation message over separating itself, i.e. $\Delta(\widetilde{w}(s), s) = 0$. As $\Delta(E[\theta|s], s) \leq 0 \leq \Delta(1,s)$ for all $s \in [\widehat{s}, \overline{s}]$, it follows that for any such s, in the domain $w \in [E[\theta|s], 1]$ there exists a unique solution to $\Delta(w,s) = 0$ (by intermediate value theorem). By continuity of Δ, $\widetilde{w}(s)$ is continuous in s.

Let $B(s') := E[\theta|s \in [s', s_{max}]] - \widetilde{w}(s')$ be the difference between the value that type s' actually gets from sending the highest recommendation message (and pooling with those above him) and the wage that makes him indifferent between sending the highest recommendation message over separating itself. Observe that $E[\theta|s \in [s', s_{max}]]$ is a strictly increasing and continuous function in s'. Moreover, $E[\theta|s \in [\overline{s}, s_{max}]] > E[\theta|\overline{s}] = \widetilde{w}(\overline{s})$. There are two cases to consider.

(i) If $\widehat{s} > s_{min}$, then $E[\theta|s \in [\widehat{s}, s_{max}]] < 1 = \widetilde{w}(\widehat{s})$, and so $B(\widehat{s}) < 0 < B(\overline{s})$. Thus, by intermediate value theorem, there exists a unique $\underline{s} \in (\widehat{s}, \overline{s})$ such that $B(\underline{s}) = 0$.[15] This establishes the existence of $\underline{s}$ when $\widehat{s} > s_{min}$; in such case, $\underline{s} > s_{min}$ and condition (3) is satisfied.

(ii) The logic of case (i) fails when there is no $\widehat{s} \leq \overline{s}$ such that $B(\widehat{s}) < 0$. So suppose this the case, which in turn implies that $B(s_{min}) \geq 0$. Since $B(s_{min}) = E\left[\theta | s \in [s_{min}, s_{max}]\right] - \widetilde{w}(s_{min})$, this implies that $E\left[\theta | s \in [s_{min}, s_{max}]\right] \geq \widetilde{w}(s_{min})$. Observe that $\Delta(w, s)$ is strictly increasing in w, and $\Delta\left(\widetilde{w}(s_{min}), s_{min}\right) = 0$ by definition of $\widetilde{w}$.

Thus, $\Delta(E\left[\theta | s \in [s_{min}, s_{max}]\right], s_{min}) \geq 0$, and condition (4) is satisfied.

Observe that if $\gamma < \frac{1}{2}$, then $\Delta(1, s_{min}) = 2\gamma - 1 < 0$, $\widehat{s} > s_{min}$ and case (i) applies. □

Proof of Theorem 2.

The necessity of (3) follows from the discussion in the main text preceding the theorem. Fix any $\underline{s} \geq s_{min}$ that satisfy conditions (3) and (4). I will show that there is a corresponding LSHP equilibrium. This suffices to prove the theorem because of Lemma A2.

First, I define the strategy profile. An LSHP equilibrium features a cut-off type $\underline{s}$ such that any type $s < \underline{s}$ separates by playing $\rho^*(s)$, whereas all types above $\underline{s}$ send the highest message of one $\rho(s) = 1$. The employee's strategy ρ is clear, with the only possible ambiguity being for type $\underline{s}$—assume for concreteness that they "pool up" rather than "down" (so type $\underline{s}$ does not separate). Let $M^\rho := [0, \rho^*(\underline{s})) \bigcup \{1\}$ be the set of recommendation messages used in this strategy. It is clear what the firm's response must be for any $m \in M^\rho$: $w(m) = E\left[\theta | s = \rho^{*-1}(m)\right]$ for $m \in [0, \rho^*(\underline{s}))$, and $w(1) = E\left[\theta | s \in [\underline{s}, s_{max}]\right]$. For any $m \in [\rho^*(\underline{s}), 1)$, the firm plays $w(m) = E\left[\theta | \underline{s}\right]$.

I now argue that (ρ, w) as constructed above is an LSHP equilibrium. Obviously, ρ is an LSHP strategy and w is the firm's valuation according to beliefs that are derived by Bayes rule on the equilibrium path. So it only needs to be shown that ρ is optimal for the employee.

CLAIM 1. *Assume $\underline{s} = s_{min}$. Then no type has a profitable deviation to any $\widehat{m} \in [0, 1)$.*

Using Lemma A2, $\rho(s) = 1$ for all $s \geq s_{min}$ because $\underline{s} = s_{min}$. So I prove the result by showing that if some type has a profitable deviation to some $\widehat{m} \in [0, 1)$, then so does $\underline{s} = s_{min}$. For this, it suffices to show that if $s > s_{min}$ and $\widehat{m} \in [0, 1)$, then

$$\gamma E\left[\theta | s_{min}\right] - (1-\gamma) E\left[(\widehat{m} - \theta)^2 | s\right] > \gamma E\left[\theta | s \in [s_{min}, s_{max}]\right] - (1-\gamma) E\left[(1-\theta)^2 | s\right] \tag{7}$$

$$\Downarrow$$

$$\gamma E\left[\theta | s_{min}\right] - (1-\gamma) E\left[(\widehat{m} - \theta)^2 | s_{min}\right] > \gamma E\left[\theta | s \in [s_{min}, s_{max}]\right] - (1-\gamma) E\left[(1-\theta)^2 | s_{min}\right]. \tag{8}$$

[15] Although $\widetilde{w}(s)$ is a strictly decreasing function in s, this fact is not needed for this proof.

Hence, it suffices to show that (8) > (7), or equivalently

$$(1-\gamma)\int_{\widehat{m}}^{1}\int_{s_{min}}^{s}\left(-2\frac{\partial E\left[\theta|z\right]}{\partial z}\right)dzdy<0,$$

which holds because $\frac{\partial E[\theta|z]}{\partial z}=\lambda>0$. The proof is completed by noting that $\underline{s}=s_{min}$ does not have an incentive to deviate to any $\widehat{m}\in[0,1)$ because

$$\gamma E\left[\theta|s\in[s_{min},s_{max}]\right]-(1-\gamma)\,E\left[(1-\theta)^{2}\,|s_{min}\right]\geq\gamma E\left[\theta|s_{min}\right]-(1-\gamma)\,E\left[(\widehat{m}-\theta)^{2}\,|s_{min}\right],$$

which follows from Lemma A2 and the fact that $\rho(s_{min})=0$ minimizes $E\left[(\rho-\theta)^{2}\,|s_{min}\right]$.

From the derivation above, since $\underline{s}=s_{min}$ doesn't have an incentive to deviate to any $\widehat{m}\in[0,1)$, neither does any $s>\underline{s}$. □

CLAIM 2. *Assume $\underline{s}>s_{min}$. Type $\underline{s}$ is indifferent between playing $\rho^{*}(\underline{s})$ and $\rho=1$, but strictly prefers $\rho^{*}(\underline{s})$ to any* $[0,\rho^{*}(\underline{s}))$. Any $s\neq\underline{s}$ strictly prefers $\rho(s)$ to any $m\in M^{\rho}\setminus\rho(s)$.

It is immediate from Lemma A2 that (3) holds, and so $\underline{s}$ is indifferent between playing $\rho^{*}(\underline{s})$ and $\rho=1$.

Equation (DE) implies that for all $s\in(s_{min},\underline{s})$,

$$\gamma\lambda-2\,(1-\gamma)\left(\rho^{*}\left(s\right)-E\left[\theta|s\right]\right)\frac{d\rho^{*}(s)}{ds}=0.$$

As $\frac{\partial E[\theta|z]}{\partial z}>0$, this implies that for all $\widetilde{s}<s\in(s_{min},\underline{s})$,

$$\gamma\lambda-2\,(1-\gamma)\left(\rho^{*}\left(s\right)-E\left[\theta|\widetilde{s}\right]\right)\frac{d\rho^{*}(s)}{ds}<0.$$

Therefore, by continuity, any $s\leq\underline{s}$ strictly prefers $\rho^{*}(s)$ to any $m\in\bigcup_{\widetilde{s}\leq\underline{s}}\rho^{*}(\widetilde{s})\setminus\rho^{*}\left(s\right)$.

To show that any $s<\underline{s}$ strictly prefers $\rho^{*}(s)$ to any $\rho=1$, pick any such s. From the previous arguments,

$$\gamma E\left[\theta|s\right]-(1-\gamma)\,E\left[(\rho^{*}(s)-\theta)^{2}\,|s\right]>\gamma E\left[\theta|\underline{s}\right]-(1-\gamma)\,E\left[(\rho^{*}(\underline{s})-\theta)^{2}\,|s\right],$$

and hence it suffices to show that

$$\gamma E\left[\theta|\underline{s}\right]-(1-\gamma)\,E\left[(\rho^{*}(\underline{s})-\theta)^{2}\,|s\right]\geq\gamma E\left[\theta|s\in[\underline{s},s_{max}]\right]-(1-\gamma)\,E\left[(1-\theta)^{2}\,|s\right].\quad(9)$$

As

$$\gamma E\left[\theta|s\in[\underline{s},s_{max}]\right]-(1-\gamma)\,E\left[(1-\theta)^{2}\,|\underline{s}\right]=\gamma E\left[\theta|\underline{s}\right]-(1-\gamma)\,E\left[(\rho^{*}\left(\underline{s}\right)-\theta)^{2}\,|\underline{s}\right],$$

(9) is true if

$$\gamma E\left[\theta|\underline{s}\right]-(1-\gamma)\,E\left[\left(\rho^{*}(\underline{s})-\theta\right)^{2}|s\right]-\left\{\gamma E\left[\theta|s\in\left[\underline{s},s_{max}\right]\right]-(1-\gamma)\,E\left[(1-\theta)^{2}|s\right]\right\}$$
$$\geq\gamma E\left[\theta|s\in\left[\underline{s},s_{max}\right]\right]-(1-\gamma)\,E\left[(1-\theta)^{2}|\underline{s}\right]-\left\{\gamma E\left[\theta|\underline{s}\right]-(1-\gamma)\,E\left[\left(\rho^{*}\left(\underline{s}\right)-\theta\right)^{2}|\underline{s}\right]\right\},$$

which can be rewritten as

$$(1-\gamma)\int_{\rho^{*}(\underline{s})}^{1}\int_{s}^{\underline{s}}\left(-2\frac{\partial E\left[\theta|z\right]}{\partial z}\right)dzdy<0,$$

and this holds because $\frac{\partial E[\theta|z]}{\partial z}=\lambda>0$.

The argument for types $s>\underline{s}$ strictly preferring $\rho=1$ to any $m\in\bigcup_{\widetilde{s}\leq\underline{s}}\rho^{*}(\widetilde{s})$ is analogous to above. □

CLAIM 3. *Assume $\underline{s}>s_{min}$. No type has a profitable deviation to any $\widehat{m}\in[\rho^{*}(\underline{s}),1)$.*

Recall, any $\widehat{m}\in[\rho^{*}(\underline{s}),1)$ is responded with $w(\widehat{m})=E\left[\theta|\underline{s}\right]$. Hence when constrained to such messages, any type $\widetilde{s}\leq\underline{s}$ maximizes utility by sending $\rho^{*}(\underline{s})$ (same wage and lower cost). However, by the previous claim, all $\widetilde{s}\leq\underline{s}$ weakly prefer playing $\rho^{*}(\widetilde{s})$ to any $\bigcup_{\widetilde{s}\leq\underline{s}}\rho^{*}(\widetilde{s})$. This proves the claim for any $\widetilde{s}\leq\underline{s}$.

To prove the claim for all $s>\underline{s}$, it suffices to show that if $s>\underline{s}$ has a profitable deviation to a $\widehat{m}\in[\rho^{*}(\underline{s}),1)$, then type $\underline{s}$ has a profitable deviation to $\widehat{m}$, as the latter is not possible. For this, it suffices to show that if $\widetilde{s}>\underline{s}$ and $\widehat{m}\in[\rho^{*}(\underline{s}),1)$, then

$$\gamma E\left[\theta|\underline{s}\right]-(1-\gamma)\,E\left[\left(\widehat{m}-\theta\right)^{2}|\widetilde{s}\right]>\gamma E\left[\theta|s\in\left[\underline{s},s_{max}\right]\right]-(1-\gamma)\,E\left[(1-\theta)^{2}|\widetilde{s}\right] \tag{10}$$
$$\Downarrow$$
$$\gamma E\left[\theta|\underline{s}\right]-(1-\gamma)\,E\left[\left(\widehat{m}-\theta\right)^{2}|\underline{s}\right]>\left\{\gamma E\left[\theta|s\in\left[\underline{s},s_{max}\right]\right]-(1-\gamma)\,E\left[(1-\theta)^{2}|\underline{s}\right]\right\}. \tag{11}$$

Hence, it suffices to show that (11) > (10), or equivalently

$$(1-\gamma)\int_{\widehat{m}}^{1}\int_{\underline{s}}^{\widetilde{s}}\left(-2\frac{\partial E\left[\theta|z\right]}{\partial z}\right)dzdy<0,$$

which holds because $\frac{\partial E[\theta|z]}{\partial z}=\lambda>0$.

The above claims establish that ρ is optimal for the sender, completing the proof. □□

Proof of Theorem 3.

(1) – (3). The first part of the theorem follows from differentiating (DE) with respect to

σ_ϵ^2,

$$\frac{\partial \rho'(s)}{\partial \sigma_\epsilon^2} = \underbrace{\frac{\gamma}{2\left(1-\gamma\right)\left(\rho(s) - E[\theta|s,t]\right)^2}}_{>0} \frac{\partial \lambda(t)}{\partial \sigma_\epsilon^2} \left[\left(\rho(s) - E[\theta|s,t]\right) + \lambda(t)\left(s-\mu\right)\right].$$

Observe that $\lambda(t) = \frac{\sigma_\epsilon^{-2}}{\left(\sigma_\epsilon^{-2}+\sigma_\theta^{-2}\right)}$, and so $\frac{\partial \lambda(t)}{\partial \sigma_\epsilon^2} = -\left(\frac{1}{\sigma_\theta^2}\right)\frac{1}{\left(1+\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\right)^2} < 0$. After substituting the prior mean $E\left[\theta|s,t\right] = \mu + \lambda(t)\left(s-\mu\right)$ in the expression $\left(\rho(s) - E[\theta|s,t]\right) + \lambda(t)\left(s-\mu\right)$, I get $\rho(s) - \mu$, which is larger (smaller) than zero when $\mu \leq 0$ ($\mu \geq 1$) because $\rho > 0$ ($\rho < 1$) on a separating interval. Denote $\rho^*(s,t)$ ($\rho^*(s,t')$) as the recommendation strategy when referrals environment is $t = (\sigma_\epsilon^2, \sigma_\theta^2, \mu)$ ($t = \left(\sigma_\epsilon^{2'}, \sigma_\theta^2, \mu\right)$), and the employee has more (less) precise information. Observe that the recommendation strategy only varies on the separating part, so it is indeed the separating function ρ^*. Then, by the above arguments, the recommendation strategy varies faster for the employee with more (less) precise information $\rho_s^*(s,t) > \rho_s^*(s,t')$ ($\rho_s^*(s,t) < \rho_s^*(s,t')$) when the prior mean is sufficiently low $\mu \leq 0$ (sufficiently high $\mu \geq 1$).

Note that by definition $E\left[\theta|s_{min},t\right] = 0$, which implies $s_{min}(t) = -\frac{(1-\lambda(t))}{\lambda(t)}\mu$. Since $\frac{\partial\lambda(t)}{\partial\sigma_\epsilon^2} < 0$, $s_{min}(t) \leq s_{min}(t')$ ($s_{min}(t) \geq s_{min}(t')$) when the prior mean is sufficiently low $\mu \leq 0$ (sufficiently high $\mu \geq 1$).

For $\mu \in (0,1)$, if $min\left\{\underline{s}(t), \underline{s}(t')\right\} > \mu$, then by continuity of ρ, there exists a $\widetilde{s}(t) < \mu$ such that $\rho^*\left(\widetilde{s}(t), t\right) = \mu$. Let $s_l = min\left\{\widetilde{s}(t), \widetilde{s}(t')\right\}$ and $s_r = max\left\{\widetilde{s}(t), \widetilde{s}(t')\right\}$. Then, by monotonicity of ρ, $\rho^*(s,t) < \mu$ ($\rho^*(s,t) > \mu$) and $\rho^*(s,t') < \mu$ ($\rho^*(s,t') > \mu$) when $s < s_l$ ($s > s_r$). Thus, by above arguments, $\rho_s^*(s,t) < \rho_s^*(s,t')$ when $s < s_l$, and $\rho_s^*(s,t) > \rho_s^*(s,t')$ when $s > s_r$. $\square$

Moreover, for γ sufficiently low, one can show that the threshold, below which the firm can infer signals in an LSHP equilibrium, is lower for the employee with more (less) precise information when $\mu \leq 0$ ($\mu \geq 1$). This is shown in series of steps as follows.

STEP 1. *Solve the (DE).*

It is easy to verify that the family of solutions to (DE) is given by $\rho^*(s,t) + c = \frac{-\gamma}{2(1-\gamma)} ln\left[\frac{\gamma}{2(1-\gamma)} + E\left[\theta|s,t\right] - \rho^*\left(s,t\right)\right]$, where c is a constant to be determined. Using the initial value condition $\rho^*\left(s_{min},t\right) = 0$, it follows that $c = \frac{-\gamma}{2(1-\gamma)} ln\left[\frac{\gamma}{2(1-\gamma)}\right]$. Substituting this value of c into the solution of (DE) gives

$$\rho^*(s,t) = \frac{\gamma}{2(1-\gamma)} ln\left[\frac{\frac{\gamma}{2(1-\gamma)}}{\frac{\gamma}{2(1-\gamma)} + E\left[\theta|s,t\right] - \rho^*\left(s,t\right)}\right]. \tag{12}$$

STEP 2. *Define Critical Function.*

Observe that one can rewrite the indifference condition (3) of cut-off type from theorem

2 as follows.

$$\begin{array}{c}(1-\gamma)\,\rho^*\left(\underline{s},t\right)^2-2\,(1-\gamma)\,E\left[\theta|\underline{s},t\right]\rho^*\left(\underline{s},t\right)\\+\left\{\gamma\left[E\left[\theta|s\in\left[\underline{s},s_{max}\right],t\right]-E\left[\theta|\underline{s},t\right]\right]-(1-\gamma)\left[1-2E\left[\theta|\underline{s},t\right]\right]\right\}=0\end{array}.$$

Using the quadratic formula, one can solve for the roots of above expression, $\rho^*(\underline{s},t) = E\left[\theta|\underline{s},t\right] \pm \frac{1}{2(1-\gamma)}\sqrt{(.)}$, where

$$\sqrt{(.)}=2\sqrt{(1-\gamma)}\sqrt{(1-\gamma)\left(1-E\left[\theta|\underline{s},t\right]\right)^2-\gamma\left(E\left[\theta|s\in\left[\underline{s},s_{max}\right],t\right]-E\left[\theta|\underline{s},t\right]\right)}.$$

Since $\rho^*(\underline{s},t) > E\left[\theta|\underline{s},t\right]$ on a separating interval, one can rule out the negative root. After substituting $\rho^*(\underline{s},t) = E\left[\theta|\underline{s},t\right] + \frac{1}{2(1-\gamma)}\sqrt{(.)}$ into (12), one gets $E\left[\theta|\underline{s},t\right] + \frac{\sqrt{(.)}}{2(1-\gamma)} - \frac{\gamma}{2(1-\gamma)}ln\left[\frac{1}{1-\frac{\sqrt{(.)}}{\gamma}}\right] = 0$. Denote the left-hand side of this expression as the "critical" function Ω.

STEP 3. *Apply Implicit Function Theorem on the Critical Function.*

One can then apply the implicit function theorem to get

$$\begin{aligned}\frac{\partial\underline{s}(t)}{\partial\sigma_\epsilon^2}&=-\frac{\frac{\partial\Omega}{\sigma_\epsilon^2}}{\frac{\partial\Omega}{\partial\underline{s}}}\\&=-\frac{\left\{1+\left[\frac{\left(\gamma-\sqrt{(.)}\right)-2\gamma\sqrt{(.)}}{2\sqrt{(.)}^2\left(\gamma-\sqrt{(.)}\right)}\right]\left[2\,(1-\gamma)\left(E\left[\theta|\underline{s},t\right]-1\right)+(..)\right]\right\}\frac{\partial\lambda(t)}{\partial\sigma_\epsilon^2}\left(\underline{s}(t)-\mu\right)}{\left\{1+\left[\frac{\left(\gamma-\sqrt{(.)}\right)-2\gamma\sqrt{(.)}}{2\sqrt{(.)}^2\left(\gamma-\sqrt{(.)}\right)}\right]\left[2\,(1-\gamma)\left(E\left[\theta|\underline{s},t\right]-1\right)+(...)\right]\right\}\lambda\left(t\right)},\end{aligned}$$

where $(..) = \gamma\left(1-\frac{1}{\lambda(t)}\frac{\partial E[\theta|s\in[\underline{s},s_{max}],t]}{\partial\underline{s}}\right)$, and $(...) = \gamma\left(1-\frac{1}{\frac{\partial\lambda(t)}{\partial\sigma_\epsilon^2}(\underline{s}(t)-\mu)}\frac{\partial E[\theta|s\in[\underline{s},s_{max}],t]}{\partial\sigma_\epsilon^2}\right)$. Observe that $\lambda(t) > 0$ and $\frac{\partial\lambda(t)}{\partial\sigma_\epsilon^2} < 0$. For γ sufficiently low, $(..) \approx (...)$ and so $\frac{\partial\underline{s}(t)}{\partial\sigma_\epsilon^2} \approx -\frac{\frac{\partial\lambda(t)}{\partial\sigma_\epsilon^2}(\underline{s}(t)-\mu)}{\lambda(t)}$, which is larger (smaller) than zero when $\mu < 0$ ($\mu \geq 1$) because $\underline{s}(t) \geq s_{min}(t) > 0$ ($\underline{s}(t) < s_{max}(t) \leq 1$). For $\mu = 0$, $\underline{s}(t) \geq s_{min}(t) = 0$, and $\frac{\partial\underline{s}(t)}{\partial\sigma_\epsilon^2} \geq 0$—the inequality is strict if $\underline{s}(t) > s_{min}(t)$, which is the case when $\gamma < \frac{1}{2}$. Thus, for γ sufficiently low, the threshold below which the firm can infer signals in an LSHP equilibrium $\underline{s}(t) < \underline{s}(t')$ ($\underline{s}(t) > \underline{s}(t')$) is lower for the employee with more (less) precise information when the prior mean is sufficiently low $\mu \leq 0$ (sufficiently high $\mu \geq 1$). □

Finally, as $\sigma_\theta^2 \to \infty$, $\lambda(t) = \frac{\sigma_\epsilon^{-2}}{\left(\sigma_\epsilon^{-2}+\sigma_\theta^{-2}\right)} \to 1$, and the posterior mean $E\left[\theta|s,t\right] = \mu + \lambda(t)\left(s-\mu\right) \to s$. Observe that $s_{min}(t) = 0 = s_{min}(t')$, and $s_{max}(t) = 1 = s_{max}(t')$. Since $E\left[\theta|s,t\right]$ does not depend on t, neither does the initial value condition $\rho(0,t) = 0 = \rho(0,t')$

and the (DE). Thus, $\rho(s,t)=\rho(s,t')$ for any t and $t'\neq t$.

Observe that $w(1,t)=E\left[\theta|s\in\left[\underline{s},1\right],t\right]=\int_{\underline{s}}^{1}s\frac{f(s|t)}{\int_{\underline{s}}^{1}f(s|t)ds}ds$ where $f(s|t)=\int_{\theta}f(\theta|s,t)d\theta$, and $f(\theta|s,t)=\frac{1}{\sqrt{2\pi\sigma_{\epsilon}^{2}}}e^{-\frac{1}{2\sigma_{\epsilon}^{2}}(s-\theta)^{2}}\frac{1}{\sqrt{2\pi\sigma_{\theta}^{2}}}e^{-\frac{1}{2\sigma_{\theta}^{2}}(\theta-\mu)^{2}}$ is the posterior of θ given an exo-signal s and referrals environment $t=(\sigma_{\epsilon}^{2},\sigma_{\theta}^{2},\mu)$. As $\sigma_{\theta}^{2}\to\infty$, $\frac{f(s|t)}{\int_{\underline{s}}^{1}f(s|t)ds}\to\frac{\int_{\theta}e^{-\frac{1}{2\sigma_{\epsilon}^{2}}(s-\theta)^{2}}d\theta}{\int_{\underline{s}}^{1}\left[\int_{\theta}e^{-\frac{1}{2\sigma_{\epsilon}^{2}}(s-\theta)^{2}}d\theta\right]ds}=\frac{\sqrt{2\pi\sigma_{\epsilon}^{2}}}{\int_{\underline{s}}^{1}\sqrt{2\pi\sigma_{\epsilon}^{2}}ds}$, where the equality follows from the symmetry of normal distributions $\int_{\theta}e^{-\frac{1}{2\sigma_{\epsilon}^{2}}(s-\theta)^{2}}d\theta=\int_{s}e^{-\frac{1}{2\sigma_{\epsilon}^{2}}(s-\theta)^{2}}ds$—right-hand side is the integral of the conditional distribution of exo-signal s given θ and t, which has the mean θ and variance σ_{ϵ}^{2}. Thus, $w(1,t)=\frac{1}{(1-\underline{s})}\int_{\underline{s}}^{1}s=\frac{1+\underline{s}}{2}$, and $Var\left[\theta|s\in\left[\underline{s},1\right],t\right]=\frac{1}{(1-\underline{s})}\int_{\underline{s}}^{1}\sigma_{\epsilon}^{2}ds=\sigma_{\epsilon}^{2}$. Observe that the critical function, which determines the cut-off type,

$$\Omega=\underline{s}+2\sqrt{(1-\gamma)(1-\underline{s})}\sqrt{(1-\gamma)(1-\underline{s})-\frac{\gamma}{2}}-\frac{1}{2(1-\gamma)}ln\left[\frac{1}{1-\frac{2\sqrt{(1-\gamma)(1-\underline{s})}}{\gamma}\sqrt{(1-\gamma)(1-\underline{s})-\frac{\gamma}{2}}}\right],$$

does not depend on t. Thus, $\underline{s}(t)=\underline{s}(t')$ for any t and $t'\neq t$. Since $E\left[\theta|s\in\left[\underline{s},1\right],t\right]$ does not depend on t, and $Var\left[\theta|s\in\left[\underline{s},1\right],t\right]$ is increasing with σ_{ϵ}^{2}, $\theta(s\in\left[\underline{s},1\right],t)\succ_{SOSD}\theta(s\in\left[\underline{s},1\right],t')$. □□

Proof of Proposition 1.

As before, recommendation strategies are characterized using the (DE), $\rho_{s}\left(s_{t},t\right)=\frac{\gamma}{2(1-\gamma)(\rho(s_{t},t)-E[\theta|s_{t},t])}$, where[16]

$$E[\theta|s_{t},t]=\begin{cases}s_{strong} & \text{if } t=t_{strong}\\ E\left[\theta|s_{weak},t_{weak}\right] & \text{if } t=t_{weak}\end{cases},$$

and $E\left[\theta|s_{weak},t_{weak}\right]=\int_{s_{weak}-\alpha}^{s_{weak}+\alpha}\theta\frac{f(t_{weak}|\theta,s_{weak})}{\int_{s_{weak}-\alpha}^{s_{weak}+\alpha}f(t_{weak}|\theta,s_{weak})d\theta}d\theta$.[17]

For $\theta<min\left\{\underline{s}(t_{weak})-\alpha,\underline{s}(t_{strong})\right\}$, firms can infer exo-signals of both ties.

So it is clear what a firm's response must be for any s_{t}: $w\left(\rho\left(s_{strong},t_{strong}\right),t_{strong}\right)=s_{strong}$ when ties are strong, and $w\left(\rho\left(s_{weak},t_{weak}\right),t_{weak}\right)=E\left[\theta|s_{weak},t_{weak}\right]$ when ties are weak. For all $\theta\in\left[s_{weak}-\alpha,s_{weak}+\alpha\right]$ that have given exo-signal s_{weak}, wage offers from weak ties is fixed $E\left[\theta|s_{weak},t_{weak}\right]$ but wage offers from strong ties is increasing in θ because $s_{strong}=\theta$. Consequently, if $\widetilde{\theta}$ prefers strong ties, then $\theta>\widetilde{\theta}$ also prefer strong ties. Thus, there is a threshold $\widehat{\theta}=E\left[\theta|s_{weak},t_{weak}\right]$ such that $\theta<\widehat{\theta}$ prefer weak ties and $\theta>\widehat{\theta}$ prefer

[16] Observe that there is no lambda in the numerator because $\theta\sim U\left[0-2\alpha,1+2\alpha\right]$; similarly, $\lambda=1$ when there is a lack of prior.

[17] Here the expression for $E\left[\theta|s_{weak},t_{weak}\right]$ follows from the fact that $f\left(\theta|s_{weak},t_{weak}\right)=f\left(t_{weak}|\theta,s_{weak}\right)f\left(s_{weak}|\theta\right)f\left(\theta\right)$, $f\left(s_{weak}|\theta\right)=\frac{1}{2\alpha}$, and $f\left(\theta\right)=\frac{1}{1+4\alpha}$.

strong ties.[18] Using the condition $\widehat{\theta} = E\left[\theta|s_{weak}, t_{weak}\right]$ and applying quadratic formula (ruling out the negative root), one gets that

$$\widehat{\theta} = E\left[\theta|s_{weak}, t_{weak}\right] = s_{weak} - \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right], \tag{13}$$

where $\frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right] \in [0,1]$. Next, I examine wage differential between weak and strong ties, $WG_t(\theta) = \left[w\left(t_{weak}, \theta\right) - w\left(t_{strong}, \theta\right)\right]$, where $w\left(t_{strong}, \theta\right) = \theta$ and $w\left(t_{weak}, \theta\right) = \int_{\theta-\alpha}^{\theta+\alpha} E\left[\theta|s_{weak}, t_{weak}\right] \frac{f(t_{weak}|\theta, s_{weak})}{\int_{\theta-\alpha}^{\theta+\alpha} f(t_{weak}|\theta, s_{weak}) ds_{weak}} ds_{weak}$.[19] Observe that $\int_{\theta-\alpha}^{\theta+\alpha} f\left(t_{weak}|\theta, s_{weak}\right) ds_{weak} = \int_{\theta-\alpha}^{\widehat{s_{weak}}} (1-\delta)\, ds_{weak} + \int_{\widehat{s_{weak}}}^{\theta+\alpha} ds_{weak}$, where $\widehat{s_{weak}} = \theta + \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right]$ by inverting (13). After some manipulations, one gets $w\left(t_{weak}, \theta\right) = \theta + 2\frac{\alpha}{\delta}\sqrt{1-\delta} > \theta = w\left(t_{strong}, \theta\right)$.

For $\theta \geq min\left\{\underline{s}(t_{weak}) - \alpha, \underline{s}(t_{strong})\right\}$, returns to ties depend on how much information firms can infer from weak versus strong ties.

There are two cases to consider.

(1) If $min\left\{\underline{s}(t_{weak}) - \alpha, \underline{s}(t_{strong})\right\} = \underline{s}(t_{strong})$, then $\underline{s}(t_{strong}) < \underline{s}(t_{weak})$.

For $\theta \geq \underline{s}(t_{strong})$, wage offers from strong ties, $w(1, t_{strong}) = \frac{1+\underline{s}(t_{strong})}{2}$, does not vary with θ. So for all $\theta \in [s_{weak} - \alpha, s_{weak} + \alpha]$ that have given exo-signal s_{weak}, either $w\left(\rho\left(s_{weak}, t_{weak}\right), t_{weak}\right) > w(1, t_{strong})$ and everyone prefers weak ties $f\left(t_{weak}|\theta, s_{weak}\right) = 1$, or $w\left(\rho\left(s_{weak}, t_{weak}\right), t_{weak}\right) < w(1, t_{strong})$ and everyone prefers strong ties $f\left(t_{weak}|\theta, s_{weak}\right) = (1-\delta)$. Consequently, $f\left(t_{weak}|\theta, s_{weak}\right) = f\left(t_{weak}|s_{weak}\right)$, and $w\left(\rho\left(s_{weak}, t_{weak}\right), t_{weak}\right) = s_{weak}$. Observe that $w\left(\rho\left(s_{weak}, t_{weak}\right), t_{weak}\right) = s_{weak}$ is increasing in s_{weak} and $w(1, t_{strong}) = \frac{1+\underline{s}(t_{strong})}{2}$ is fixed. Thus, there is a threshold exo-signal $s^*_{weak} = \frac{1+\underline{s}(t_{strong})}{2}$ such that applicants prefer strong ties when $s_{weak} < s^*_{weak}$ and they prefer weak ties when $s_{weak} > s^*_{weak}$. Define $\widetilde{s_{weak}} = min\left\{s^*_{weak}, \underline{s}(t_{weak})\right\}$, and consider the following four regions.

(i) For $\theta \in \left(\underline{s}(t_{strong}), \widetilde{s_{weak}} - \alpha\right)$, applicants prefer strong ties because they can only have $s_{weak} < s^*_{weak}$.

It is easy to verify that $w\left(t_{weak}, \theta\right) = \theta < \frac{1+\underline{s}(t_{strong})}{2} = w\left(t_{strong}, \theta\right)$, everyone prefers strong ties, and so $WG_t(\theta) < 0$. Set $\theta^*_l = \underline{s}(t_{strong})$.

Moreover, since $s_{weak} \in [\theta - \alpha, \theta + \alpha]$, there exists a cut-off ability θ^*_r such that $\theta > \theta^*_r$ prefer weak ties because they are more likely to have $s_{weak} > s^*_{weak}$.

(ii) For $\theta \geq \underline{s}(t_{weak}) + \alpha$, they can only have $s_{weak} > \underline{s}(t_{weak})$.

---

[18] Note that

$$E\left[\theta|s_{weak}, t_{weak}\right] = \int_{s_{weak}-\alpha}^{\widehat{\theta}} \theta \frac{1}{\int_{s_{weak}-\alpha}^{\widehat{\theta}} d\theta + \int_{\widehat{\theta}}^{s_{weak}+\alpha} (1-\delta)\, d\theta} d\theta + \int_{\widehat{\theta}}^{s_{weak}+\alpha} \theta \frac{(1-\delta)}{\int_{s_{weak}-\alpha}^{\widehat{\theta}} d\theta + \int_{\widehat{\theta}}^{s_{weak}+\alpha} (1-\delta)\, d\theta} d\theta.$$

[19] Here, the expression for $w\left(t_{weak}, \theta\right)$ follows from the fact $f\left(s_{weak}, |\theta, t_{weak}\right) = f\left(t_{weak}|\theta, s_{weak}\right) f\left(s_{weak}|\theta\right)$ and $f\left(s_{weak}|\theta\right) = \frac{1}{2\alpha}$.

Either $w(1, t_{weak}) > w(1, t_{strong})$ and everyone prefers weak ties $f(t_{weak}|\theta, s_{weak}) = 1$, or $w(1, t_{weak}) < w(1, t_{strong})$ and everyone prefers strong ties $f(t_{weak}|\theta, s_{weak}) = (1-\delta)$. Consequently, $f(t_{weak}|\theta, s_{weak}) = f(t_{weak}|s_{weak})$, and $w(1, t_{weak}) = \frac{1+\underline{s}(t_{weak})}{2}$. Observe that $w(1, t_{weak}) > w(1, t_{strong})$ because $\underline{s}(t_{weak}) > \underline{s}(t_{strong})$. Since everyone prefers weak ties, $WG_t(\theta) > 0$.

(iii) For $\theta \in \left(\widetilde{s_{weak}} - \alpha, \underline{s}(t_{weak}) - \alpha\right)$, it can be shown after some algebra that

$$w(t_{weak}, \theta) = \frac{\left\{\delta\left(\theta^2 + \alpha^2\right) + (2-\delta)\, 2\alpha\theta - \delta\widetilde{s_{weak}}\right\}}{2\left[\delta\left(\theta - \widetilde{s_{weak}}\right) + (2-\delta)\,\alpha\right]},$$

which is a convex function for $\theta \geq \widetilde{s_{weak}} - \alpha$.[20]

(iv) For $\theta \in (\underline{s}(t_{weak}) - \alpha, \underline{s}(t_{weak}) + \alpha)$, it can be shown after some algebra that

$$w(t_{weak}, \theta) = \frac{\left\{\underline{s}^2(t_{weak}) - \delta\widetilde{s_{weak}}^2 - (1-\delta)(\theta-\alpha)^2 + \left[(\theta+\alpha) - \underline{s}(t_{weak})\right]\left[1 + \underline{s}(t_{weak})\right]\right\}}{2\left[\delta\left(\theta - \widetilde{s_{weak}}\right) + (2-\delta)\,\alpha\right]},$$

which is a concave function for $\theta \leq \widetilde{s_{weak}} + \alpha$.[21]

In addition, it can be shown that $\frac{\partial w(t_{weak},\theta)}{\partial\theta} > 0$ for $\theta \in (\underline{s}(t_{weak}) - \alpha, \underline{s}(t_{weak}) + \alpha)$.[22] Note that $w(t_{weak}, \theta)$ at $\theta \in (\underline{s}(t_{strong}), \underline{s}(t_{weak}) + \alpha)$ is always higher than for any $\theta \in \left(\widetilde{s_{weak}} - \alpha, \underline{s}(t_{weak}) - \alpha\right)$.[23] This guarantees that $w(t_{weak}, \theta)$ is monotonically increasing for sufficiently high ability, implying that if $w(t_{weak}, \theta) > w(t_{strong}, \theta)$ for $\theta \in \left(\widetilde{s_{weak}} - \alpha, \underline{s}(t_{weak}) - \alpha\right)$, then it also holds for $\theta \in (\underline{s}(t_{strong}), \underline{s}(t_{weak}) + \alpha)$.

Thus, there are two possibilities.

Since $WG_t(\theta) < 0$ for $\theta \in \left(\underline{s}(t_{strong}), \widetilde{s_{weak}} - \alpha\right)$, if $\frac{\partial w(t_{weak},\theta)}{\partial\theta} < 0$ for $\theta \in \left(\widetilde{s_{weak}} - \alpha, \underline{s}(t_{weak}) - \alpha\right)$, then $WG_t(\theta) < 0$ for $\theta \in (\underline{s}(t_{strong}), \underline{s}(t_{weak}) - \alpha)$ as well. So there exists a threshold $\theta_r^* \in (\underline{s}(t_{weak}) - \alpha, \underline{s}(t_{weak}) + \alpha)$ because $\frac{\partial w(t_{weak},\theta)}{\partial\theta} > 0$ for $\theta \in (\underline{s}(t_{weak}) - \alpha, \underline{s}(t_{weak}) + \alpha)$ and $WG_t(\theta) > 0$ for $\theta \geq \underline{s}(t_{weak}) + \alpha$.

Otherwise, $\frac{\partial w(t_{weak},\theta)}{\partial\theta} > 0$ for $\theta \in (\underline{s}(t_{strong}), \underline{s}(t_{weak}) + \alpha)$, and the threshold θ_r^* is found

[20] Note that $\frac{\partial^2 w(t_{weak},\theta)}{\partial\theta^2} = 2\delta^2\left(\theta - \widetilde{s_{weak}}\right) + 2\alpha\delta(2-\delta)$ is increasing in θ and is equal to $4\alpha\delta(1-\delta) > 0$ at $\theta = \widetilde{s_{weak}} - \alpha$.

[21] Note that $\frac{\partial^2 w(t_{weak},\theta)}{\partial\theta^2} = -2(1-\delta)\left[\alpha + \delta(\theta-\alpha) + \widetilde{s_{weak}}\right] < 0$ is decreasing in θ and is equal to $-2(1-\delta)\left[\alpha + \delta\underline{s}(t_{weak}) + \widetilde{s_{weak}}\right] < 0$ at $\theta = \underline{s}(t_{weak}) + \alpha$.

[22] Note that $\frac{\partial w(t_{weak},\theta)}{\partial\theta} = 2(1-\delta)\,\alpha\,(1 - \underline{s}(t_{weak})) + 2\delta^2\widetilde{s_{weak}}\underline{s}(t_{weak}) + \left(\delta\underline{s}(t_{weak}) - \widetilde{s_{weak}}\right)^2 > 0$ at $\theta = \underline{s}(t_{weak}) + \alpha$.

[23] To see this, observe that $w(t_{weak}, \theta) = \int_{\theta-\alpha}^{\theta+\alpha} E[\theta|s_{weak}, t_{weak}] \frac{f(t_{weak}|\theta, s_{weak})}{\int_{\theta-\alpha}^{\theta+\alpha} f(t_{weak}|\theta, s_{weak})ds_{weak}} ds_{weak}$, where $E[\theta|s_{weak}, t_{weak}] = s_{weak}$ for $s_{weak} \in [\theta-\alpha, \underline{s}(t_{weak})]$ and $E[\theta|s_{weak}, t_{weak}] = \frac{1+\underline{s}(t_{weak})}{2}$ for $s_{weak} \in [\underline{s}(t_{weak}), \theta+\alpha]$ when $\theta \in (\underline{s}(t_{strong}), \underline{s}(t_{weak}) + \alpha)$, whereas $E[\theta|s_{weak}, t_{weak}] = s_{weak}$ for $s_{weak} \in [\theta-\alpha, \theta+\alpha]$ when $\theta \in \left(\widetilde{s_{weak}} - \alpha, \underline{s}(t_{weak}) - \alpha\right)$. It follows from the fact that $\frac{1+\underline{s}(t_{weak})}{2} > \underline{s}(t_{weak}) \geq s_{weak}$ for $\theta \in \left(\widetilde{s_{weak}} - \alpha, \underline{s}(t_{weak}) - \alpha\right)$.

in this larger region. $\square$

(2) If $min\left\{\underline{s}(t_{weak})-\alpha,\underline{s}(t_{strong})\right\}=\underline{s}(t_{weak})-\alpha$, then either $\underline{s}(t_{strong})<\underline{s}(t_{weak})$ and the logic of case (1) applies, or $\underline{s}(t_{weak})<\underline{s}(t_{strong})$ and the following two sub-cases needs to be considered for $\theta\geq\underline{s}(t_{weak})-\alpha$.

(i) If $s_{weak}<\underline{s}(t_{weak})$, then the same logic as for $\theta<min\left\{\underline{s}(t_{weak})-\alpha,\underline{s}(t_{strong})\right\}$ applies, implying that returns from weak ties are higher $WG_t(\theta)>0$.

(ii) If $s_{weak}\geq\underline{s}(t_{weak})$, then

$$
\begin{aligned}
w\left(1,t_{weak}\right) &= \int_{\underline{s}(t_{weak})}^{s_{max}(t_{weak})} E\left[\theta|s_{weak},t_{weak}\right]\frac{f\left(s_{weak},|t_{weak}\right)}{\int_{\underline{s}(t_{weak})}^{s_{max}(t_{weak})} f\left(s_{weak},|t_{weak}\right)ds_{weak}}ds_{weak}\\
&= \frac{1}{s_{max}(t_{weak})-\underline{s}(t_{weak})}\int_{\underline{s}(t_{weak})}^{s_{max}(t_{weak})} E\left[\theta|s_{weak},t_{weak}\right]ds_{weak},
\end{aligned}
$$

where $E\left[\theta|s_{weak},t_{weak}\right]=s_{weak}-\frac{\alpha}{\delta}\left[(2-\delta)-2\sqrt{1-\delta}\right]$ follows from (13) above. Here, the second equality follows from $f\left(s_{weak},|t_{weak}\right)=\int_{s_{weak}-\alpha}^{s_{weak}+\alpha}f\left(s_{weak}|\theta\right)f\left(\theta\right)$, $f\left(s_{weak}|\theta\right)=\frac{1}{2\alpha}$, and $f\left(\theta\right)=\frac{1}{1+4\alpha}$. Note that $s_{max}(t_{weak})=1+\frac{\alpha}{\delta}\left[(2-\delta)-2\sqrt{1-\delta}\right]$, and so it can be shown that $w\left(1,t_{weak}\right)=\frac{1+\underline{s}(t_{weak})}{2}-\frac{1}{2}\frac{\alpha}{\delta}\left[(2-\delta)-2\sqrt{1-\delta}\right]$. Thus, there is a threshold θ^* such that returns from weak ties are higher $WG_t(\theta)>0$ when $\theta<\theta^*$ and returns from strong ties are higher when $WG_t(\theta)<0$ when $\theta>\theta^*$ prefer strong ties. If $w\left(1,t_{weak}\right)<\underline{s}(t_{strong})$, then $\theta^*=w(1,t_{weak})$; otherwise, $\theta^*=\underline{s}(t_{strong})$. $\square$

And finally, one can show that $\underline{s}(t_{weak})>\underline{s}(t_{strong})$ when γ is sufficiently low by examining the critical function Ω. To see this, define $E\left[\theta|s_t,t\right]=s_t-X(\alpha)$ where $X(\alpha)=\frac{\alpha}{\delta}\left[(2-\delta)-2\sqrt{1-\delta}\right]$. Observe that $X'(\alpha)>0$, and for sufficiently low γ, $\frac{\partial\underline{s}(t)}{\partial\alpha}=-\frac{\frac{\partial\Omega}{\alpha}}{\frac{\partial\Omega}{\partial\underline{s}}}\approx-(-X'\left(\alpha\right))>0$.

I am done by noting that $\alpha=0$ when ties are strong. $\square\square$

Proof of Proposition 2.

As in proposition 1, recommendation strategies are characterized using the (DE).

For $\theta<\underline{s}(h_t)$, firms can infer exo-signals of its employees.

So it is clear what a firm's response must be for any s_h: $w\left(\rho\left(s_t,h_t\right),h_t\right)=s_t$ when social ties are used, and $w\left(h_0\right)=E\left[\theta|s_0,h_0\right]$ when the formal market channel is selected. For all $\theta\in\left[s_0-\alpha,s_0+\alpha\right]$ that have given exo-signal s_0, wage offer from the formal market channel is fixed $E\left[\theta|s_0,h_0\right]$ but wage offer from social ties is increasing in θ because $s_t=\theta$. Following similar steps as in the proof of proposition 1, there is a threshold

$$\widehat{\theta}=E\left[\theta|s_0,h_0\right]=s_0-\frac{\alpha}{\delta}\left[(2-\delta)-2\sqrt{1-\delta}\right], \tag{14}$$

such that $\theta < \widehat{\theta}$ prefer the formal market channel and $\theta > \widehat{\theta}$ prefer social ties. Next, I examine the wage differential between social tie and formal market channels, $WG_h(\theta) = [w(h_t,\theta) - w(h_0,\theta)]$, where $w(h_t,\theta) = \theta$ and $w(h_0,\theta) = \int_{\theta-\alpha}^{\theta+\alpha} E[\theta|s_0,h_0] \frac{f(h_0|\theta,s_0)}{\int_{\theta-\alpha}^{\theta+\alpha} f(h_0|\theta,s_0)ds_0} ds_0$. Observe that $\int_{\theta-\alpha}^{\theta+\alpha} f(h_0|\theta,s_0)\,ds_0 = \int_{\theta-\alpha}^{\widehat{s_0}} (1-\delta)\,ds_0 + \int_{\widehat{s_0}}^{\theta+\alpha} ds_0$, where $\widehat{s_0} = \theta + \frac{\alpha}{\delta}\left[(2-\delta) - 2\sqrt{1-\delta}\right]$ by inverting (14). After some manipulations, one gets $w(h_0,\theta) = \theta + 2\frac{\alpha}{\delta}\sqrt{1-\delta} > \theta = w(h_t,\theta)$.

For $\theta \geq \underline{s}(h_t)$, wage offers from social ties, $w(1,h_t) = \frac{1+\underline{s}(h_t)}{2}$, does not vary with θ.

So for all $\theta \in [s_0-\alpha, s_0+\alpha]$ that have given exo-signal s_0, either $w(s_0,h_0) > w(1,h_t)$ and everyone prefers the formal market channel $f(h_0|\theta,s_0) = 1$, or $w(s_0,h_0) < w(1,h_t)$ and everyone prefers social ties $f(h_0|\theta,s_0) = (1-\delta)$. Consequently, $f(h_0|\theta,s_0) = f(h_0|s_0)$, and $w(s_0,h_0) = s_0$. Observe that $w(s_0,h_0) = s_0$ is increasing in s_0 and $w(1,h_t) = \frac{1+\underline{s}(h_t)}{2}$ is fixed. Thus, there is a threshold exo-signal $\widetilde{s_0} = \frac{1+\underline{s}(t_{strong})}{2}$ such that applicants prefer social ties when $s_0 < \widetilde{s_0}$ and they prefer the formal market channel when $s_0 > \widetilde{s_0}$. Consider the following two regions.

(i) For $\theta \in (\underline{s}(h_t), \widetilde{s_0}-\alpha)$, applicants prefer social ties because they can only have $s_0 < \widetilde{s_0}$.

It is easy to verify that $w(h_0,\theta) = \theta < \frac{1+\underline{s}(h_t)}{2} = w(h_t,\theta)$, everyone prefers social ties, and so $WG_h(\theta) > 0$. Set $\theta_l^{**} = \underline{s}(h_t)$. Moreover, since $s_0 \in [\theta-\alpha, \theta+\alpha]$, there exists a cut-off ability θ_r^{**} such that $\theta > \theta_r^{**}$ prefer the formal market channel because they are more likely to have $s_0 > \widetilde{s_0}$.

(ii) For $\theta \geq \widetilde{s_0}-\alpha$, it can be shown after some algebra that $w(h_0,\theta) = \frac{\left\{\delta\left(\theta^2+\alpha^2\right)+(2-\delta)2\alpha\theta-\delta\widetilde{s_0}\right\}}{2[\delta(\theta-\widetilde{s_0})+(2-\delta)\alpha]}$, which is a convex function for $\theta \geq \widetilde{s_0}-\alpha$.[24]

For δ sufficiently low, it is easy to verify that $\frac{\partial w(h_0,\theta)}{\partial\theta} > 0$ for $\theta \geq \widetilde{s_0}-\alpha$ and $w(h_0,1) > w(h_t,1)$. Thus, $\theta_r^{**} < 1$. $\square$

[24]Note that $\frac{\partial^2 w(h_0,\theta)}{\partial\theta^2} = 2\delta^2(\theta-\widetilde{s_0}) + 2\alpha\delta(2-\delta)$ is increasing in θ and is equal to $4\alpha\delta(1-\delta) > 0$ at $\theta = \widetilde{s_0}-\alpha$.