A Structural Analysis of Job Referrals and Social Networks: The Case of the Corporate Executives Market

Yiran Chen *

November 8, 2018
Latest Version Available Here

Abstract

I develop and structurally implement a labor market search model in which workers, in addition to directly receiving job offers, also receive referrals from their social contacts. In the model, referrals are generated endogenously: an external referral occurs when a friend rejects an offer he/she receives, and an internal referral occurs when a friend leaves his/her current job. I estimate the model by Generalized Method of Moments using data on the labor market history and the social connections of executives in S&P 500 firms. Using the estimated model, I find that referrals play a substantial role in the executive labor market. More than one quarter of the job transitions and raises are driven by referrals. Shutting down referrals reduces executives' welfare by an equivalence of a two to seven percentage points reduction in income. I also evaluate the impacts of the structure of the social networks by comparing the outcomes under the observed networks and alternative randomly formed networks. I find that the welfare distribution is more unequal under the random networks. I further investigate the mechanisms for these effects through the lens of two network statistics: friends' popularity and local community clustering.

Keywords: job search, referrals, social networks, managerial labor market

*Department of Economics, University of Pennsylvania, The Ronald O. Perelman Center for Political Science and Economics, 133 South 36th Street, Suite 150, Philadelphia, PA 19104. Email: yiranc@sas.upenn.edu. I am extremely grateful to Hanming Fang, Andrew Postlewaite, Petra Todd and Kenneth Burdett for their continuing guidance and support. I have benefited greatly from discussions with Gorkem Bostanci, Tomas Larroucau, Hongxun Ruan, Andrew Shephard, Holger Sieg, Takeaki Sunada, Hanna Wang, Weilong Zhang, and participants at the UPenn Empirical Micro Lunch Seminar and the Empirical Micro Workshop. All errors are my own.

1 Introduction

Job referrals are an important channel through which workers find jobs and firms fill vacancies.[1] Because of referrals, workers who are more socially interconnected may experience better labor market outcomes and advance more quickly in their careers. In this paper, I study the network effects of job referrals and investigate the impacts of social network structure on workers' labor market outcomes. I develop a job search model that incorporates referral and non-referral job offers and different kinds of social networks. The network structure is dynamic and evolves as workers move across jobs. To estimate the model, I combine three different data sources on the corporate executive labor market.

In my model, workers directly receive job offers as well as referrals from their friends. Referrals are generated endogenously: an external referral occurs when a friend rejects an offer that he/she receives, and an internal referral occurs when a friend leaves his/her current job. As a result of this referral process, the quantity and the quality of referrals depend not only on the worker's number of friends but also on the quality of these friends' jobs. Moreover, the model generates rich network effects beyond immediate friends. First, the *popularity of a friend* affects referrals. On the one hand, a popular friend means high competition for referrals, lowering a worker's probability of receiving a particular referral sent by his/her popular friend (*competition effect*). On the other hand, a popular friend benefits from his/her large set of friends' referrals, increasing the quantity and the quality of referrals he/she sends out (*ripple effect*). Second, *local clustering*, defined as the fraction of a worker's friends who are also friends with one another, also affects referrals. An advantage of high clustering is that it keeps the positive spillovers in an inner circle (*closeness effect*). A disadvantage is that it limits the positive spillovers from a distance (*isolation effect*).

My empirical analysis is based on three data sets: (1) Compustat Executive Compensation (ExecuComp), (2) BoardEx, and (3) U.S. Stock Database from the Center for Research in Security Prices (CRSP). The first data set is used to construct a panel of individuals' employment history in the executive market along with their on-the-job compensation. The second data set is used to construct three social networks representing social connections established via education, work, and other social activities. The third data set provides firms' financial variables.

My analysis first provides reduced-form evidence on job referrals in the executive labor market. First, I show that socially connected executives' compensation is positively correlated, controlling for time-invariant individual characteristics, time-specific shocks, and industry-specific shocks. This

[1]For example, Granovetter (1973) reports that 56% of the job seekers in a Boston suburb obtained their jobs through social contacts. Other surveys at different locations and times report numbers ranging from 25% to 87%. Marsden and Gorman (2001) report that 37% of U.S. firms often use referral in recruitment. Other surveys at different locations and times report numbers ranging from 36% to 88%. See Topa (2011) for a detailed summary.

result supports my model's prediction that executives with better jobs send better quality referrals, increasing their friends' compensation. Second, I show that an individual is more likely to make a career advancement when his/her executive friends leave their current jobs, controlling for time-invariant individual characteristics and time-specific shocks. This finding suggests that executives who leave their current jobs send referrals that help their friends advance in their careers.

I then estimate the structural job search model by Generalized Method of Moments (GMM). The estimation results show both the statistical and economical significance of job referrals. My model nests a model with no referrals, and a specification test rejects such a model. Simulations from my model provide a way of assessing the importance of referrals in the job market dynamics in this labor market. A decomposition shows that 27.86% of non-executive to executive transitions, 66% of executive job-to-job transitions, and 82.1% of raises are driven by referrals.

I use the estimated model to perform two counterfactual experiments. The first experiment evaluates the welfare effect of referrals by varying the probability of referrals. I find that shutting down referrals reduces executives' welfare by an equivalence of a 2-7% decrease in annual income and that increasing referral probability to one boosts executives' welfare by an equivalence of a 6-16% increase in annual income.

The second experiment examines the welfare effect of network structure by varying the network structure. Specifically, a new set of counterfactual networks are generated in which the individuals have the same number of friends as the observed networks, but the connections are formed randomly. I find that the welfare distribution is more unequal under the randomly formed networks. I further investigate the mechanisms for these effects through the lens of two network statistics: friends' popularity and local clustering. First, in terms of friends' popularity, the *competition effect* dominates the *ripple effect*. The random networks increase friends' popularity for individuals with a small number of friends and decrease friends' popularity for individuals with a large number of friends. Therefore, the increased competition for referral makes individuals with a small number of friends worse off, and the decreased competition for referral makes individuals with a large number of friends better off, which increases inequality. Second, in terms of local clustering, the *isolation effect* dominates for individuals with a small number of friends, and the *closeness effect* dominates for individuals with a large number of friends. The random networks universally decrease the local clustering. Therefore, the decreased isolation effect makes individuals with a small number of friends better off, and the decreased closeness effect makes individuals with a large number of friends worse off, which decreases inequality. Overall, the competition effect resulting from the change in friend popularity dominates, generating greater inequality under the random networks. This experiment highlights the effects of network structures beyond the number of friends.

My paper contributes to three strands of the literature. First, it contributes to studies of the impact of social connections on executive compensation (e.g., Shue, 2013; Engelberg, Gao and Parsons, 2013) by building a model that formalizes the mechanisms by which social connections impact executives' compensation and job transitions. Second, it contributes to studies of labor search models with job referrals (e.g., Montgomery, 1992; Mortensen and Vishwanath, 1994; Calvó-Armengol and Zenou, 2005; Ioannides and Soetevent, 2006; Galenianos, 2014; Arbex, O'Dea and Wiczer, 2018) by incorporating the full structure of the network (beyond the number of friends). The model presents a general framework to study a rich set of network effects of job referrals. Third, it contributes to studies of labor market dynamics on social networks (e.g., Topa, 2001; Calvó-Armengol and Jackson, 2004; Calvó-Armengol and Jackson, 2007) by modeling micro-founded workers' decisions and wage bargaining processes.

Literature Review

My paper is related to four broad strands of the literature. First, it complements the literature on labor market peer effect and job referrals in the general labor market.[2] Surveys show that referrals are frequently used by workers to find jobs (e.g., Granovetter, 1973; Pellizzari, 2010) and by firms to fill vacancies (e.g., Marsden and Gorman, 2001).[3] Additionally, empirical studies (e.g., Topa, 2001; Conley and Topa, 2007; Schmutte, 2014) show that local job referrals generate positive spatial correlation in workers' employment statuses and wage premia. More recently, Burks et al. (2015) and Brown, Setren and Topa (2016) characterize the relationships among referrals, match quality, wage trajectories, and turnover using data sets with direct information on referrals. In addition to documenting empirical patterns, some studies develop theoretical models on referrals, starting with the pioneering work by Boorman (1975). Some studies model referrals as means to reduce search friction by providing more opportunities for workers and firms to meet. Others model referrals as means to reduce information friction by providing firms with more information on worker quality or match quality (e.g., Rees, 1966; Montgomery, 1991; Arrow and Borzekowski, 2004; Dustmann et al., 2015). My paper focuses on the role referral plays in mitigating search friction, and I discuss related papers in reviewing the search literature.

Second, my paper contributes to the literature on executive compensation.[4] In the literature,

---

[2]See Ioannides and Datcher Loury (2004) and Topa (2011) for comprehensive surveys.

[3]See Holzer (1988), Corcoran, Datcher and Duncan (1980), Granovetter (1995), Gregg and Wadsworth (1996), and Addison and Portugal (2002) for additional evidence on workers' usage of referrals; see Holzer (1987), Neckerman and Kirschenman (1991), Miller and Rosenbaum (1997), Fernandez, Castilla and Moore (2000), and Brown, Setren and Topa (2016) for additional evidence on firms' usage of referrals. A more comprehensive and detailed listing can be found in Topa (2011).

[4]See Frydman and Jenter (2010) and Edmans and Gabaix (2016) for comprehensive surveys.

some researchers view the level of total compensation as a competitive outcome (e.g., Gabaix and Landier, 2008; Terviö, 2008), while others view it as rent extraction and managerial entrenchment (e.g., Bebchuk and Fried, 2004; Kuhnen and Zwiebel, 2008). My model combines both views by incorporating competition from outside offers and allowing executives to earn positive rent through bargaining. Because the focus of my study is job search and referrals, my paper abstracts from problems arising from asymmetric information such as learning about executive ability (e.g., Taylor, 2013) or contracting in the presence of hidden action and hidden information (e.g., Gayle, Golan and Miller, 2015; Gayle, Li and Miller, 2016).

Among the literature on executive compensation, my paper directly contributes to studies examining the impact of executives' social connections on their compensation. Previous studies focus mostly on documenting the empirical relationship between social connection and compensation.[5] For example, Shue (2013) uses the random assignment of MBA students to sections at the Harvard Business School to show that executive compensation is significantly more similar among graduates from the same section than among graduates from different sections, and that this effect is more than twice as strong in the year following alumni reunions. Engelberg, Gao and Parsons (2013) investigate CEOs' social connections to other executives and directors outside their firms and show that these connections increase CEO compensation, and that the increase is higher for connections with "important" people such as CEOs of big firms. My paper contributes to this literature by providing additional empirical evidence, and moreover, a structural model that formalizes the mechanisms by which social connections impact executives' compensation and job transitions.

Third, my paper contributes to the theoretical and empirical literature on labor search models.[6] There is a large body of literature on sequential random job search. For example, the wage posting model in Burdett and Mortensen (1998) is adapted and empirically implemented in Van den Berg and Ridder (1998), Bontemps, Robin and Van den Berg (1999), Bontemps, Robin and Van den Berg (2000), Meghir, Narita and Robin (2015), Shephard (2017), and Aizawa and Fang (2018). The wage bargaining model in Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006) is adapted and empirically implemented in Dey and Flinn (2005) and Bagger et al. (2014). My model is based on that of Cahuc, Postel-Vinay and Robin (2006), in which workers search both on and off the job, climb the productivity ladder, and bargain their wages with the firms.

Among the literature on labor search, my paper directly contributes to studies of referrals in

[5]There is another set of studies examining the impact of executives and directors' social connections on firm performance. See El-Khatib, Fogel and Jandik (2015), Hwang and Kim (2009), Fracassi and Tate (2012), Cai and Sevilir (2012), Larcker, So and Wang (2013), and Ruan (2017).

[6]See Rogerson, Shimer and Wright (2005) for a survey of theoretical studies and Eckstein and Van den Berg (2007) for a survey of empirical studies.

search-based frameworks, in which referral is modeled as an additional channel of job arrival. Previous studies do not accommodate rich social network structure and generate very limited network effects. In early studies such as Montgomery (1992) and Mortensen and Vishwanath (1994), there is no notion of social connections. In later studies such as Calvó-Armengol and Zenou (2005), Ioannides and Soetevent (2006), Galenianos (2014), and Arbex, O'Dea and Wiczer (2018), referrals are sent through social networks. However, they do not incorporate network structure beyond the degree distribution (i.e., the number of social connections), and they impose strong assumptions on the network structure to achieve this tractability.[7] My paper differs from this set of papers in that it does not impose assumptions on how the networks are formed, and that it uses the full structure of the network (the adjacency matrices), which generates rich network effects. I highlight the importance of the network structure beyond the number of friends by showing in my counterfactual experiment that networks with the same degree sequence but different connection patterns lead to different welfare distribution. Additionally, my model allows for a rich characterization of referrals, both internal and external. In Calvó-Armengol and Zenou (2005) and Ioannides and Soetevent (2006), there are only external referrals. In their models, jobs are identical, so employed workers send referrals whenever they receive outside job offers, while in my paper jobs are heterogeneous, and employed workers only pass along unwanted jobs. In Galenianos (2014) and Arbex, O'Dea and Wiczer (2018), there are only internal referrals. In their models, employed workers randomly send referrals identical to their own jobs, while in my paper this only occurs when a worker leaves his/her current job, a modeling choice driven by the small number of executive positions.

Finally, my paper contributes to studies on the labor market dynamics on social networks such as Topa (2001), Calvó-Armengol and Jackson (2004), Calvó-Armengol and Jackson (2007), and Fontaine (2008).[8] These studies model job transitions of socially connected workers as a Markov process on a network, with the full network structure incorporated. Topa (2001) and Calvó-Armengol and Jackson (2004) focus on employment dynamics. In their models, an employed worker randomly contacts an unemployed friend and then the friend becomes employed. Calvó-Armengol and Jackson (2007) further incorporate wage dynamics. These three papers take a statistical approach and abstract from workers' decision-making problem and firms' wage-setting problem. My paper complements their studies by formally formulating workers' decision problem and the wage bargaining process, providing micro-founded employment and wage dynamics. In Fontaine (2008),

[7]In Calvó-Armengol and Zenou (2005) and Ioannides and Soetevent (2006), a worker randomly draws new friends in every period, so only the immediate friends matter. In Galenianos (2014), a worker has a continuum of friends, so the unemployment rate of one's friends does not change in the steady state. In Arbex, O'Dea and Wiczer (2018), the random network formation model, combined with the worker's information structure, implies that a worker's degree is a sufficient statistic for his/her future value.

[8]See Jackson, Rogers and Zenou (2017) for a survey of the broad literature on social networks.

wage is bargained, but the network is assumed to be complete, while my model accommodates any network configuration.[9] To the best of my knowledge, there are not many empirical studies of job referrals with explicit network models.[10] My paper contributes to this literature by developing a tractable yet flexible model with micro foundation for empirical analysis.

The remainder of the paper is structured as follows. Section 2 presents a job search model with referrals. Section 3 describes the data and shows reduced-form evidence. Section 4 describes the estimation strategy. Section 5 presents the estimation results. Section 6 discusses the counterfactual experiments. Section 7 concludes.

2 Model

2.1 Social Networks

In this section I describe the social networks through which referrals are sent. There are n workers in the labor market, indexed by i. Among them, there are three types of social connections established via overlapping experience in school, work, and social activities. Connections are undirected, and new ones are formed over time. These three types of connections constitute three networks. In these networks, a node represents a worker and an edge represents a social connection. At time t, the networks are characterized by adjacency matrices

$$Y^{Edu,t} = \{y_{ij}^{Edu,t}\}_{1 \leq i,j \leq n}, \; Y^{Work,t} = \{y_{ij}^{Work,t}\}_{1 \leq i,j \leq n}, \; Y^{Social,t} = \{y_{ij}^{Social,t}\}_{1 \leq i,j \leq n},$$

where $y_{ij}^{Edu,t} = 1$ if workers i and j have established a school connection by time t, and $y_{ij}^{Edu,t} = 0$ otherwise; the same holds true for $y_{ij}^{Work,t}$ and $y_{ij}^{Social,t}$.

Both $Y^{Edu,t}$ and $Y^{Social,t}$ are exogenous. Part of $Y^{Work,t}$ is exogenous, and part of $Y^{Work,t}$ is endogenous as a result of transition in the executive labor market.[11] I assume that when individuals make labor market decisions, they do not consider the implied changes to the endogenous work network. This assumption is an implication of Assumption 1 (static expectation) which will be discussed later.

[9] In a complete network, every pair of distinct nodes is connected. A complete network simplifies analysis because it eliminates heterogeneity in network position.

[10] Topa (2001) is one such empirical paper. As discussed, it does not feature micro-founded worker's decision problem and wage-setting process.

[11] Endogenous connections are established between the executives in the same company. Other workplace connections are viewed as exogenous if one of the individuals is not an executive. For example, when executives in different companies serve on the same board of directors in a third company, this connection is viewed as exogenous to the model of executive labor market.

Additionally, define a simplified network that does not distinguish the different types of connections

$$Y^{All,t} = \{y_{ij}^{All,t}\}_{1\le i,j\le n},$$

where $y_{ij}^{All,t} = \max\{y_{ij}^{Edu,t}, y_{ij}^{Work,t}, y_{ij}^{Social,t}\}$, i.e., $y_{ij}^{All,t} = 1$ as long as there is some kind of social connection between i and j.

For any network $k \in \{Edu,\ Work,\ Social,\ All\}$, define worker i's friends in network k at time t as the set of nodes it is directly connected to:

$$N^{k,t}(i) = \{j \mid y_{ij}^{k,t} = 1\}. \tag{1}$$

Define worker i's degree in network k at time t as the number of i's friends:

$$d_i^{k,t} = |N^{k,t}(i)|. \tag{2}$$

Finally, define worker i's local clustering coefficient in network k at time t as the number of connections between i's friends divided by the number of possible connections between them:

$$c_i^{k,t} = \begin{cases} \frac{|\{y_{jl}^{k,t}: j,l\in N^{k,t}(i), y_{jl}^{k,t}=1\}|}{\binom{d_i^{k,t}}{2}} & \text{if } d_i^{k,t} > 1 \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

where $d_i^{k,t}$ and $N^{k,t}(i)$ are defined in (2) and (1). Local clustering coefficient is a measure of how tightly connected a local network is. It takes a value between $[0, 1]$, characterizing the fraction of worker i's friends who are also friends with one another.

To avoid verbosity, throughout the paper, when I refer to friend, degree, or local clustering coefficient without specifying a particular network, I mean those associated with the simplified network Y^{All}. Additionally, to ease notation, I omit time superscript t when it does not cause confusion.

2.2 Job Search with Referrals

In this section I describe the job search model with referrals through social connections. I follow the terminology in the standard search literature, but certain terms should be interpreted in the context of the executive labor market: specifically, "workers" refer to executives and candidates for executive jobs, "wage" refers to the value of the compensation package, and "unemployment" refers to the status of not working in an executive job. The basic settings are similar to Cahuc, Postel-Vinay and Robin (2006). A novel feature in my model is that workers can send referrals to friends after they reject an offer or leave their current jobs, which affects both job transition and wage bargaining.

I consider a labor market with a finite number of workers and firms. Workers and firms are matched randomly in a frictional labor market. Time is continuous. Workers enjoy instantaneous utility $U(x)$ from income x and discount the future at rate ρ.

Production. Workers differ in their abilities (denoted as a) as well as their positions in the social networks. Firms differ in their productivities (denoted as p). A firm is modeled as a collection of jobs with the same productivity. The marginal product for a worker-firm pair (a, p) is ap. An unemployed worker receives an income flow of ab, which he/she has to forgo upon finding a job.

Direct Job Arrivals and Job Separations. Firms and workers meet randomly. Unemployed workers receive direct job offers at Poisson rate λ_0, and employed workers at rate λ_1. The productivity p of the firm from which an offer originates is randomly distributed on $[p^{min}, p^{max}]$ according to cdf $F(p)$. Exogenous job separations occur at Poisson rate δ.

Referrals. Referrals are by-products of traditional types of labor market transitions. They always follow direct job arrivals or job separations, when workers reject offers or leave their current firms.[12] Specifically, there are three situations under which a worker sends a referral:

1. External referral following direct job arrival: if a worker rejects a direct offer, he/she may refer one of his/her friends to the firm he/she declines. I call it an external referral because it is initiated by an individual outside of the firm.

2. Internal referral following direct job arrival: if a worker leaves his/her current firm because he/she accepts a better job offer, he/she may refer one of his/her friends to the firm he/she leaves. I call it an internal referral because it is initiated by an individual who has worked at the firm.

3. Internal referral following job separation: if a worker leaves his/her firm as a result of exogenous separation, he/she may refer one of his/her friends to the firm he/she leaves.[13] This is also an internal referral.

[12] In this model, a worker cannot refer a friend to his/her current firm unless he/she himself/herself leaves. This modeling choice is driven by the small number of executive positions in each firm.

[13] Job separations can be either involuntary or voluntary. Voluntary separations such as retirement or leaving for health reasons may lead to internal referrals.

Workers are assumed to be nonstrategic in sending referrals. They send referrals with probability π_1 after direct job arrivals and with probability π_0 after job separations. Conditional on sending a referral, a recipient is sampled according to the following sequential statistical process:

1. Sample employment status: sample unemployed friends with probability ν_u and employed friends with probability $\nu_e = 1 - \nu_u$.

2. Sample a network: sample school friends, work friends, and social-activity friends with probability $(\omega^{Edu}, \omega^{Work}, \omega^{Social}) \in \Delta^2$. When not all three types of friends are present, only sample from the available types and normalize the probability accordingly.

3. Sample a friend: conditional on employment status and type of network, sample one friend randomly.

More precisely, conditional on friend $j \in N^{All}(i)$ sending a referral, the probability for worker i to receive this referral is

$$\gamma_{i\leftarrow j} = \begin{cases} \nu_u(\xi_j^{Edu\cap u} \cdot \frac{y_{ij}^{Edu}}{d_j^{Edu\cap u}} + \xi_j^{Work\cap u} \cdot \frac{y_{ij}^{Work}}{d_j^{Work\cap u}} + \xi_j^{Social\cap u} \cdot \frac{y_{ij}^{Social}}{d_j^{Social\cap u}}) & \text{if } s_i = u, \\ \nu_e(\xi_j^{Edu\cap e} \cdot \frac{y_{ij}^{Edu}}{d_j^{Edu\cap e}} + \xi_j^{Work\cap e} \cdot \frac{y_{ij}^{Work}}{d_j^{Work\cap e}} + \xi_j^{Social\cap e} \cdot \frac{y_{ij}^{Social}}{d_j^{Social\cap e}}) & \text{if } s_i = e, \end{cases} \tag{4}$$

where $\mathbf{s} = \{s_i\}_{i=1,\ldots,n}$ is a vector of employment statuses for all workers. For each network $k \in \{Edu, Work, Social\}$. $d_j^{k\cap u} = |N^k(j) \cap \{i : s_i = u\}|$ is the number of j's unemployed friends in network k, and $d_j^{k\cap e} = |N^k(j) \cap \{i : s_i = e\}|$ is the number of j's employed friends in network k. $\xi_j^{k\cap u} = \frac{\mathbb{1}(d_j^{k\cap u}>0)\omega^k}{\mathbb{1}(d_j^{Edu\cap u}>0)\omega^{Edu}+\mathbb{1}(d_j^{Work\cap u}>0)\omega^{Work}+\mathbb{1}(d_j^{Social\cap u}>0)\omega^{Social}}$ is the probability for j to send referrals to unemployed friends in network k, and $\xi_j^{k\cap e} = \frac{\mathbb{1}(d_j^{k\cap e}>0)\omega^k}{\mathbb{1}(d_j^{Edu\cap e}>0)\omega^{Edu}+\mathbb{1}(d_j^{Work\cap e}>0)\omega^{Work}+\mathbb{1}(d_j^{Social\cap e}>0)\omega^{Social}}$ is the probability for j to send referrals to employed friends in network k.

In this model, a referral is a chance for a worker to meet with a firm: an additional source of job arrival. It is tied to the firm's productivity, not to the wage offered to the worker sending the referral. All the wages are bargained between worker-firm pairs.

I make the following additional assumptions of the referral process. First, sending a referral is costless and takes no time. Second, unemployed workers do not send referrals when they reject job offers.[14] Third, referrals have no immediate "chain effect": when a worker rejects a referral, he/she no longer passes it along to his/her friends; when a worker accepts a referral and change

[14] This assumption guarantees the tractability of the model. It ensures that the "reservation productivity" for an unemployed worker is a single agent decision problem. Otherwise, if unemployed workers are allowed to send referrals after rejecting job offers, the "reservation productivities" for unemployed workers will be interdependent, resulting in a game among the workers. In terms of the empirics, the estimation results show few unemployed workers reject offers. Therefore, this assumption is relatively innocuous quantitatively.

jobs, he/she does not send an internal referral about the job he/she leaves.[15] This assumption is supported by reduced-form evidence in Section 3.3 that there is no significant co-movement in job transitions beyond immediate friends.

Wage Bargaining. Wages are bargained over by workers and firms, and the bargaining process is the same for direct offers and referrals. In the bargaining, information is complete. All participating agents observe one another's types: the firm's productivity p, and the worker's ability a as well as his/her state variable Γ which embodies information relevant to the arrivals of referrals. For expository purposes, I focus on bargaining here and defer the description of Γ to later. Firms can vary their wage offers according to the characteristics (a, Γ) of the particular worker they meet. They can also counter the offers received by their employees from competing firms. Additionally, wage contracts are long-term contracts that can be renegotiated by mutual agreement only. Finally, the bargaining outcome is such that a worker obtains his/her reservation value and a share β of the additional worker rent, where β represents the worker's bargaining power.[16]

Formally, let $V_0(a, \Gamma)$ denote the lifetime value of an unemployed worker with ability a and state variable Γ; let $V_1(a, w, p, \Gamma)$ denote that of the same worker when employed at a firm of productivity p and paid a wage w. The bargained wage between a type-(a, Γ) unemployed worker and a type-p firm, denoted as $\phi_0(a, p, \Gamma)$, satisfies

$$V_1(a, \phi_0(a, p, \Gamma), p, \Gamma) = V_0(a, \Gamma) + \beta[V_1(a, ap, p, \Gamma) - V_0(a, \Gamma)]. \tag{5}$$

In this equation, $V_0(a, \Gamma)$ is the worker's reservation value, $V_1(a, ap, p, \Gamma)$ is the maximum value the worker can hope to extract from the match with the firm, and $[V_1(a, ap, p, \Gamma) - V_0(a, \Gamma)]$ is the additional worker rent brought about by this match. Note that the worker only accepts the offer if $V_1(a, ap, p, \Gamma) \geq V_0(a, \Gamma)$. Equivalently, this can be characterized by a productivity threshold $p_0(a, \Gamma)$, defined by

$$V_1(a, ap_0(a, \Gamma), p_0(a, \Gamma), \Gamma) = V_0(a, \Gamma). \tag{6}$$

If $p \geq p_0(a, \Gamma)$, the worker accepts the offer; otherwise, he/she rejects it.

When a worker employed at firm p is contacted by an outside firm p', the incumbent firm and the

[15] This assumption reduces the computational intensity of the model. Relaxing this assumption will not substantially change the analytical property of the model. It is useful to note that this assumption is not as restrictive as it seems. Although in the short run referrals only affect immediate friends, in the long run the network effects will propagate beyond immediate friends (the ripple effect).

[16] When the worker's utility is linear, this can be interpreted as a Nash bargaining solution. In more general cases, Cahuc, Postel-Vinay and Robin (2006) show that this is the outcome of a strategic bargaining game adapted from Rubinstein (1982).

poaching firm compete for the worker. If the poaching firm is more productive than the incumbent ($p' > p$), it wins the bargain by offering a wage $\phi_1(a, p, p', \Gamma)$ such that

$$V_1(a, \phi_1(a, p, p', \Gamma), p', \Gamma) = V_1(a, ap, p, \Gamma) + \beta[V_1(a, ap', p', \Gamma) - V_1(a, ap, p, \Gamma)]. \tag{7}$$

In this case, the worker's reservation value is the maximum value he/she can extract from the incumbent firm, $V_1(a, ap, p, \Gamma)$.

If the poaching firm is less productive than the incumbent ($p' \leq p$), the incumbent retains the worker by offering a renegotiated wage $\phi_1(a, p', p, \Gamma)$ such that

$$V_1(a, \phi_1(a, p', p, \Gamma), p, \Gamma) = V_1(a, ap', p', \Gamma) + \beta[V_1(a, ap, p, \Gamma) - V_1(a, ap', p', \Gamma)]. \tag{8}$$

Note that renegotiation requires mutual agreement, so it is only triggered if $\phi_1(a, p', p, \Gamma)$ is higher than the worker's current wage w. Equivalently, this can be characterized by a productivity threshold $q(a, w, p, \Gamma)$, defined by

$$\phi_1(a, q(a, w, p, \Gamma), p, \Gamma) = w. \tag{9}$$

If the poaching firm's productivity is relatively high ($q(a, w, p, \Gamma) < p' \leq p$), the worker gets a raise in the incumbent firm; if the poaching firm's productivity is too low ($p' \leq q(a, w, p, \Gamma)$), the outside offer is discarded.

Labor Market Transitions and Referrals. With the descriptions above, I can fully characterize the worker's labor market transitions and referrals.

1. When an unemployed type-(a, Γ) worker meets a firm p, either from direct arrival or referral,
   (a) If $p \geq p_0(a, \Gamma)$, the worker accepts the offer at a bargained wage $\phi_0(a, p, \Gamma)$;
   (b) Otherwise, the worker rejects the offer.
2. When an employed type-(a, Γ) worker at firm p paid at wage w meets an outside firm p', either from direct arrival or referral,
   (a) If $p' > p$, the worker moves to the new firm with a wage $\phi_1(a, p, p', \Gamma)$.
   Additionally, if this follows a direct job arrival, the worker may refer a friend to the firm he/she leaves (internal referral).
   (b) If $q(a, w, p, \Gamma) < p' \leq p$, the worker stays at the current firm with a raise to $\phi_1(a, p', p, \Gamma)$.
   Additionally, if this follows a direct job arrival, the worker may refer a friend to the firm he/she rejects (external referral).

(c) If $p' \leq q(a, w, p, \Gamma)$, the worker keeps the current wage at the current firm. Additionally, if this follows a direct job arrival, the worker may refer a friend to the firm he/she rejects (external referral).

3. When an employed worker is exogenously separated from his/her current job, he/she may refer a friend to the firm he/she leaves (internal referral).

State Variable Γ^i. $\Gamma^i = \{\Gamma_0^i(\cdot), \Gamma_1^i(\cdot)\}$ are distributions of the productivities of worker i's friends' jobs, when i is unemployed and employed respectively.[17] More precisely, $\Gamma_0^i(p)$ and $\Gamma_1^i(p)$ are the "head count" of i's friends working at firms with productivities no greater than p, where each friend j is weighted by $\gamma_{i\leftarrow j}$, his/her probability of sending referral to i. As $\gamma_{i\leftarrow j}$, defined in (4), depends on recipient i's employment status, so does Γ^i. Specifically, Γ_0^i is calculated using $\gamma_{i\leftarrow j}$ for unemployed i, and Γ_1^i is calculated using $\gamma_{i\leftarrow j}$ for employed i:

$$\begin{aligned} \Gamma_0^i(p) &= \sum_{j \in N^{All}(i)} \gamma_{i\leftarrow j}(s_i = u) \cdot \mathbb{1}(p_j \leq p), \\ \Gamma_1^i(p) &= \sum_{j \in N^{All}(i)} \gamma_{i\leftarrow j}(s_i = e) \cdot \mathbb{1}(p_j \leq p). \end{aligned} \tag{10}$$

In the following part, I discuss two properties of Γ^i. First, $(\lambda_1, \pi_1, \delta, \pi_0, F, \Gamma^i)$ fully characterize worker i's *instantaneous* arrival rate and distribution of referrals. To illustrate this point, first consider the referrals from one particular friend j. External referral from j arrives at rate $\lambda_1 \pi_1 F(p_j) \gamma_{i\leftarrow j}$, and the associated productivity distribution is F truncated above at p_j. Internal referral from j arrives at rate $[\lambda_1 \pi_1 (1 - F(p_j)) + \delta \pi_0] \gamma_{i\leftarrow j}$, and the associated productivity is p_j. Then use Γ^i to aggregate over friends.

Second, $(\lambda_1, \pi_1, \delta, \pi_0, F, \Gamma^i)$ are not sufficient statistics to forecast the dynamics of Γ^i. The dynamics of Γ^i depend on the dynamics of friends' jobs $\{p_j\}_{j \in N^{All}(i)}$. Friends' jobs are affected by referrals they receive, i.e. $\{\Gamma^j\}_{j \in N^{All}(i)}$, and thus their friends jobs $\{\{p'_j\}_{j' \in N^{All}(j)}\}_{j \in N^{All}(i)}$. By similar argument, the interdependence goes further on the network (ripple effects). Therefore, the dynamics of Γ^i is determined by the structure of the full networks and the dynamics of all workers' jobs.[18] Even though referrals do not have immediate effect beyond direct friends, in the long run, referrals generate ripple effects that spread across the whole network.

[17] Rigorously speaking, they are not probability distributions because the total measures do not necessarily equal one.

[18] For networks with multiple components, rigorously speaking, the dynamics of Γ^i is determined by the structure of the connected component containing i and the dynamics of the workers in the component.

Worker's Information Set. First, a worker i observes the following information about his/her local networks:

$$\begin{aligned}&N^{Edu}(i), N^{Work}(i), N^{Social}(i);\\&\{d_j^{Edu\cap u}, d_j^{Edu\cap e}, d_j^{Work\cap u}, d_j^{Work\cap e}, d_j^{Social\cap u}, d_j^{Social\cap e}\}_{j\in N^{All}(i)}.\end{aligned}$$ [19]

He/she knows the identities of his/her friends and the types of their connections. Additionally, he/she knows his/her friends' degree in unemployed and employed school network, work network, and social-activity network, so he/she can calculate the level of "competition" for referrals. Second, he/she observes the productivities of his/her friends' jobs:

$$\{p_j\}_{j\in N^{All}(i)}.$$

Therefore, a worker has enough information to calculate the instantaneous Γ^i.

Workers' Forward-Looking Behavior. I make the following assumptions on workers' forward-looking behavior.

Assumption 1. Workers have static expectation of Γ.

Workers do not forecast the dynamics of Γ. When workers bargain with firms, they calculate Γ using the latest information and assume it does not change. In other words, workers ignore future changes of their networks as well as their friends' jobs, and thus ignore future changes of the arrival rate and distribution of referrals in their calculation of future values. This assumption can be interpreted as bounded rationality. Given workers' limited information, it is forbiddingly difficult to calculate Γ's law of motion for two main reasons. First, Γ is high dimensional; second, the calculation requires integration over the structures of the unobserved part of the networks and integration over non-friends' jobs.[20]

This assumption introduces discrepancy between the worker's problem of solving bargained wage and my (the researcher's) problem of studying the dynamics. It should be emphasized that this assumption only applies to the worker's problem and does not apply to my study of the dynamics. I incorporate the evolution of the networks and analyze the labor market transitions of all workers according to the model description in the previous parts.

[19] It is reasonable to assume that a worker knows only his/her local network. Previous studies (e.g., Friedkin, 1983; Krackhardt, 1987; Krackhardt, 2014; and Banerjee et al., 2017) show that people have little knowledge of their social network structure beyond immediate friends.

[20] An alternative modeling approach is to approximate Γ with a low dimensional object and approximate its law of motion by imposing further assumptions. This alternative approach accommodates dynamics at the expense of accuracy. In this paper, I choose to use the accurate characterization of Γ and forgo the dynamics.

This assumption has two implications. First, in making labor market decisions, a worker does not consider the implied change to the endogenous work network. For example, a worker will not accept an "undesirable" job for the sole purpose of becoming friends with another worker in the same firm. This implication is not unrealistic, especially in the executive labor market, because an individual can achieve similar purpose by actively building on his/her social-activity network, whose evolution is accommodated in my model. Second, in sending referrals, a worker does not consider the implied change to friends' jobs. This implication is innocuous, especially for referrals following direct job offers. To illustrate, consider a situation when worker i refers a job to friend j and j accepts. This will not increase helpful referrals from j to i because the best referral from j is as good as j's own job, which i has forgone.

Assumption 2. Workers are rationally forward looking in all other aspects.

Given the arrivals of referrals, workers have rational expectations on future job arrivals, job separations, and bargained wages.

Worker's Problem and Bargained Wage. An unemployed worker's value function is characterized by

$$\rho V_0(a,\Gamma) = U(ab) \tag{11a}$$

$$+ \lambda_0 \int_{p_0(a,\Gamma)}^{p^{max}} [V_1(a,\phi_0(a,x,\Gamma),x,\Gamma) - V_0(a,\Gamma)]dF(x) \tag{11b}$$

$$+ \delta\pi_0 \int_{p_0(a,\Gamma)}^{p^{max}} [V_1(a,\phi_0(a,y,\Gamma),y,\Gamma) - V_0(a,\Gamma)]d\Gamma_0(y) \tag{11c}$$

$$+ \lambda_1\pi_1 \int_{p_0(a,\Gamma)}^{p^{max}} \int_{p_0(a,\Gamma)}^{y} [V_1(a,\phi_0(a,x,\Gamma),x,\Gamma) - V_0(a,\Gamma)]dF(x)d\Gamma_0(y) \tag{11d}$$

$$+ \lambda_1\pi_1 \int_{p_0(a,\Gamma)}^{p^{max}} \int_{y}^{p^{max}} [V_1(a,\phi_0(a,y,\Gamma),y,\Gamma) - V_0(a,\Gamma)]dF(x)d\Gamma_0(y), \tag{11e}$$

where $\phi_0(a,p,\Gamma)$ is the bargained wage defined in (5), and $p_0(a,\Gamma)$ is the reservation productivity defined in (6). To understand expression (11), note that line (11a) represents the flow utility; line (11b) represents the expected value of receiving a direct job offer; line (11c) represents the expected value of receiving an internal referral following friends' job separations; line (11d) represents the expected value of receiving an external referral following friends' job arrivals; and finally, line (11e) represents the expected value of receiving an internal referral following friends' job arrivals.

An employed worker's value function is characterized by

$$
\begin{aligned}
\rho V_1(a,w,p,\Gamma) &= U(w) && \text{(12a)}\\
&+ \delta[V_0(a,\Gamma) - V_1(a,w,p,\Gamma)] && \text{(12b)}\\
&+ \lambda_1 \int_{q(a,w,p,\Gamma)}^{p} [V_1(a,\phi_1(a,x,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x) && \text{(12c)}\\
&+ \lambda_1 \int_{p}^{p^{max}} [V_1(a,\phi_1(a,p,x,\Gamma),x,\Gamma) - V_1(a,w,p,\Gamma)]dF(x) && \text{(12d)}\\
&+ \delta\pi_0 \int_{q(a,w,p,\Gamma)}^{p} [V_1(a,\phi_1(a,y,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]d\Gamma_1(y) && \text{(12e)}\\
&+ \delta\pi_0 \int_{p}^{p^{max}} [V_1(a,\phi_1(a,p,y,\Gamma),y,\Gamma) - V_1(a,w,p,\Gamma)]d\Gamma_1(y) && \text{(12f)}\\
&+ \lambda_1\pi_1 \int_{q(a,w,p,\Gamma)}^{p} \int_{q(a,w,p,\Gamma)}^{y} [V_1(a,\phi_1(a,x,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y) && \text{(12g)}\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{q(a,w,p,\Gamma)}^{p} [V_1(a,\phi_1(a,x,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y) && \text{(12h)}\\
&+ \lambda_1\pi_1 \int_{q(a,w,p,\Gamma)}^{p} \int_{y}^{p^{max}} [V_1(a,\phi_1(a,y,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y) && \text{(12i)}\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{p}^{y} [V_1(a,\phi_1(a,p,x,\Gamma),x,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y) && \text{(12j)}\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{y}^{p^{max}} [V_1(a,\phi_1(a,p,y,\Gamma),y,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y), && \text{(12k)}
\end{aligned}
$$

where $\phi_1(a,p,p',\Gamma)$ is the bargained wage defined in (7), and $q(a,w,p,\Gamma)$ is the productivity threshold defined in (9). To understand expression (12), note that line (12a) represents the flow utility; line (12b) represents the expected value of experiencing a job separation; line (12c) represents the expected value of a wage raise resulting from a direct job offer; line (12d) represents the expected value of a job change resulting from a direct job offer; line (12e) represents the expected value of a wage raise resulting from an internal referral following friends' job separations; line (12f) represents the expected value of a job change resulting from an internal referral following friends' job separations; lines (12g) and (12h) represent the expected value of a wage raise resulting from an external referral following friends' job arrivals for friends with productivity lower than and higher than p respectively; line (12i) represents the expected value of a wage raise resulting from an internal referral following friends' job arrivals; line (12j) represents the expected value of a job change resulting from an external referral following friends' job arrivals; and finally, line (12k) represents the expected value of a job change resulting from an internal referral following friends' job arrivals.

In Appendix A, I provide details on solving the reservation productivity $p_0(\cdot)$ and bargained wages $\phi_0(\cdot), \phi_1(\cdot)$. Specifically, I show that for CRRA utility function

$$U(x) = \begin{cases} \ln x & \text{if } \alpha = 1 \\ \frac{x^{1-\alpha}-1}{1-\alpha} & \text{if } \alpha \neq 1 \end{cases}, \tag{13}$$

the reservation productivity is given by the implicit function

$$\ln p_0(a,\Gamma) = \begin{cases} \ln b + \beta(\lambda_0 - \lambda_1)\int_{p_0(a,\Gamma)}^{p^{max}} \frac{\bar{F}(x)x^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx \\ +\delta\pi_0\beta\int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(y)-\bar{\Gamma}_1(y))y^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(y)+[\lambda_1\pi_1\beta\bar{F}(y)+\delta\pi_0\beta]\bar{\Gamma}_1(y)}dy & \text{if } \alpha = 1 \\ +\lambda_1\pi_1\beta\int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(x)-\bar{\Gamma}_1(x))\bar{F}(x)x^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx \\ \\ \frac{1}{1-\alpha}\ln\Big\{b^{1-\alpha} + (1-\alpha)\beta(\lambda_0-\lambda_1)\int_{p_0(a,\Gamma)}^{p^{max}} \frac{\bar{F}(x)x^{-\alpha}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx \\ +(1-\alpha)\delta\pi_0\beta\int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(y)-\bar{\Gamma}_1(y))y^{-\alpha}}{\rho+\delta+\lambda_1\beta\bar{F}(y)+[\lambda_1\pi_1\beta\bar{F}(y)+\delta\pi_0\beta]\bar{\Gamma}_1(y)}dy & \text{if } \alpha \neq 1 \\ +(1-\alpha)\lambda_1\pi_1\beta\int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(x)-\bar{\Gamma}_1(x))\bar{F}(x)x^{-\alpha}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx\Big\} \end{cases}, \tag{14}$$

and the bargained wages are given by

$$\ln\phi_1(a,p^L,p^H,\Gamma) = \begin{cases} \ln a + \beta\ln p^H + (1-\beta)\ln p^L \\ -\delta\pi_0(1-\beta)^2\int_{p^L}^{p^H} \frac{\bar{\Gamma}_1(y)y^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(y)+[\lambda_1\pi_1\beta\bar{F}(y)+\delta\pi_0\beta]\bar{\Gamma}_1(y)}dy \\ -\lambda_1(1-\beta)^2\Big[\int_{p^L}^{p^H} \frac{\bar{F}(x)x^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx & \text{if } \alpha = 1 \\ +\pi_1\int_{p^H}^{p^{max}}\int_{p^L}^{p^H} \frac{\bar{F}(x)x^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dxd\Gamma_1(y) \\ +\pi_1\int_{p^L}^{p^H}\int_{p^L}^{y} \frac{\bar{F}(x)x^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dxd\Gamma_1(y)\Big] \\ \\ \ln a + \frac{1}{1-\alpha}\ln\Big\{\beta(P^H)^{1-\alpha} + (1-\beta)(P^L)^{1-\alpha} \\ -(1-\alpha)\delta\pi_0(1-\beta)^2\int_{p^L}^{p^H} \frac{\bar{\Gamma}_1(y)y^{-\alpha}}{\rho+\delta+\lambda_1\beta\bar{F}(y)+[\lambda_1\pi_1\beta\bar{F}(y)+\delta\pi_0\beta]\bar{\Gamma}_1(y)}dy \\ -(1-\alpha)\lambda_1(1-\beta)^2\Big[\int_{p^L}^{p^H} \frac{\bar{F}(x)x^{-\alpha}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx & \text{if } \alpha \neq 1 \\ +\pi_1\int_{p^H}^{p^{max}}\int_{p^L}^{p^H} \frac{\bar{F}(x)x^{-\alpha}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dxd\Gamma_1(y) \\ +\pi_1\int_{p^L}^{p^H}\int_{p^L}^{y} \frac{\bar{F}(x)x^{-\alpha}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dxd\Gamma_1(y)\Big]\Big\} \end{cases}, \tag{15}$$

and

$$\phi_0(a,p,\Gamma) = \phi_1(a,p_0(a,\Gamma),p,\Gamma). \tag{16}$$

Model Property.

Proposition 1. Job transitions are independent of ability.

All the job transitions, including UE, EE, and EU, are independent of workers' abilities. First, by model assumption, direct job arrivals and job separations are independent of workers' abilities. Additionally, workers' optimal accept/reject decisions are independent of abilities: for employed workers, they always climb the productivity ladder; for unemployed workers, the reservation productivities are independent of their abilities as shown in (14). Therefore, referrals are independent of workers' abilities because they are generated as results of employed workers' job separations and accept/reject decisions after direct job offers. Finally, workers' optimal accept/reject decisions with respect to referrals are independent of abilities.

This property is mainly driven by two modeling assumptions. The first assumption is that direct job offer arrivals and job separations do not depend on workers' abilities. This assumption rules out the possibility for high-ability workers to encounter more frequent or better quality jobs, because previous empirical studies find little evidence of sorting between executive ability and firm productivity.[21] This assumption also rules out the possibility for low-ability workers to experience more frequent job separations, because in the executive market most of the separations are voluntary due to personal reasons/retirement/death and only a small fraction are forced due to low competence.[22]

The second assumption is no information asymmetry in wage bargaining. This assumption guarantees that worker ability is fully compensated in all bargained wages, and thus when contacted by the same firm, low-ability and high-ability workers have the same accept/reject incentives. It abstracts from asymmetric information problems because they are not of first order importance for the purpose of studying network effects of referrals. Moreover, in the executive market, the problem of asymmetric information is less severe because there are typically abundant records on a candidate's past performance, upon which a firm can evaluate his/her ability.

Corollary 1. The speed of career advancement depends on referrals.

Career advancement (UE and EE transitions) depends on referrals because referrals are an additional channel of job arrivals. As discussed earlier, the arrival rate and distribution of referrals are determined by Γ, the productivities of friends' jobs. Variation in Γ leads to variation in referrals

[21] For example, Terviö (2008) and Gabaix and Landier (2008) estimate assignment models between firms and CEOs. Under the assumption of positive assortative matching, their empirical results show very small dispersion in CEO ability. This suggests that, quantitatively, there is no significant sorting between worker ability and firm productivity.

[22] For example, Huson, Parrino and Starks (2001), Kaplan and Minton (2006), and Taylor (2010) use the news on the *Wall Street Journal* to categorize whether a separation is voluntary or forced. They show that on average, only 2% of the CEOs are forced to leave their job each year. Additionally, some of the forced separations are caused by personal scandals that are not directly related to the executives' competence in generating output, the ability in this model.

and ultimately to variation in the speed of career advancement.

Proposition 2. The level of compensation depends on worker ability.

Specifically, log wage is additively separable in ability as shown in (15) and (16): $\ln\phi_1(a, p^L, p^H, \Gamma) = \ln a + \ln\phi_1(1, p^L, p^H, \Gamma)$, $\ln\phi_0(a, p, \Gamma) = \ln a + \ln\phi_0(1, p, \Gamma)$. In this model, ability only comes into play in the wage bargaining process, and thus it is reflected in the level of wage.

This property is mainly driven by two parametric assumptions. First, output and unemployment income are multiplicative in worker ability. Second, workers receive their reservation value and a share of the additional rent through bargaining. These are standard assumptions in the literature, with a reasonable depiction of the reality and the convenience of tractability.

It is useful to note that the additive separability of log wage implies that the wage growth rate is independent of ability. Instead, wage growth depends on referrals because it results from competing offers.

2.3 Discussion: Networks and Referrals

My model generates a rich set of network effects, and I discuss some of them in this section.

First, a worker's number of friends is not the only thing that matters. A worker's friends act like filters in sending referrals. The distribution of referrals from one particular friend is censored by the productivity of this friend's job. In other words, the best referral one can expect from a friend is as good as the friend's own job. Therefore, the better a friend's job, the better the referral he/she sends. Referral prospects depend not only on the number of friends, but also on friends' success in their careers.

Second, friends of a friend matter. On the one hand, they "compete" for job referrals, which lowers a worker's probability of receiving a particular referral (competition effect).[23] On the other hand, they send referrals to a worker's friend, improving the friend's job. This increases the quantity and the quality of the referrals the friend sends out, which ultimately benefits the worker (ripple effect). These two mechanisms work in opposite directions, and the net effect of friends of friends is qualitatively ambiguous. Through a quantitative analysis with a counterfactual experiment in Section 6.2, I find that the competition effect is more important.

Third, mutual friends matter. More precisely, the clustering structure, i.e., whether a worker's friends are also friends with one another, affects referrals. I illustrate the effects of clustering using Figure 1 as an example. Consider worker A in these two networks.[24] According to the definition

[23]Workers do not actively compete for referrals. I use "compete" in a statistical sense.

[24]Note that in both networks, worker A has four friends, and each friend has two friends (a fifty-percent chance for A to receive a friend's referral). Therefore, for worker A, both the number of friends and the level of competition for referral are the same in the two networks. The major difference of the two networks is clustering.

in (3), worker A's local clustering coefficient $c_A = 0$ for the network on the left because none of his/her friends are connected; it is $c_A = \frac{2}{\binom{4}{2}} = \frac{1}{3}$ for the network on the right because there exist two connections among his/her friends over six possible connections. Therefore, worker A's network on the right has a higher level of clustering. An advantage of clustering is that it keeps the

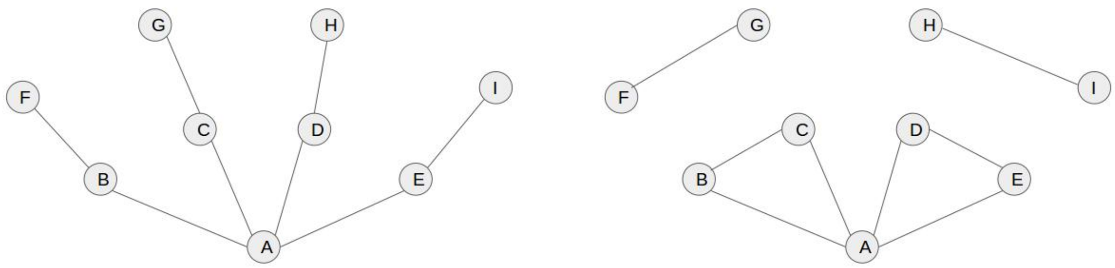

Figure 1: Low vs. High Clustering

positive spillovers in an inner circle (closeness effect). Consider the scenario when friend B sends out a referral, but not to A. In the low-clustering network, this referral goes to F, who will not be able to help A directly. In comparison, in the high-clustering network, this referral goes to C who is a friend of worker A and will be able to help A directly. In this sense, clustering helps. A disadvantage of clustering is that it limits positive spillovers from afar (isolation effect). Consider an alternative scenario when friends $\{B, C, D, E\}$ get unlucky and experience few good shocks. In the low-clustering network, worker A can still benefit in the long run from the good shocks to friends of friends $\{F, G, H, I\}$ through their referrals to $\{B, C, D, E\}$. In the high-clustering network, however, there is no such channel. In this sense, clustering hurts. These two mechanisms work in opposite directions, and the net effect of clustering is qualitatively ambiguous. Through a quantitative analysis with a counterfactual experiment in Section 6.2, I find that the relative importance of these two effects is heterogeneous.

3 Data

In this section I describe my data sets and present summary statistics and reduced-form evidence. I combine three data sets: (1) Compustat Executive Compensation (ExecuComp); (2) BoardEx; and (3) the Center for Research in Security Prices (CRSP) U.S. Stock Database.

3.1 Data Sets

Compustat Executive Compensation. Compustat Executive Compensation (hereafter, ExecuComp) provides executive compensation data collected directly from companies' annual proxy statements filed with the U.S. Securities and Exchange Commission (SEC).[25] Most companies report around 5 executives for a given year, typically the C-Suite, including the chief executive officer (CEO), the chief financial officer (CFO), and the chief operating officer (COO), etc. I use ExecuComp to construct a panel of individuals' employment history in the executive market along with their on-the-job compensation.

BoardEx. BoardEx provides network connection data among board members and senior executives in notable public and private companies collected from publicly available information. Social connections are defined by overlaps in education, private and public companies, and other social activities such as charities, clubs, business associations, and university board memberships, etc. I use BoardEx to construct three evolving networks for education, workplace, and social activities respectively. Additionally, BoardEx also provides information on the gender and nationality composition of the board of the directors.

U.S. Stock Database from the Center for Research in Security Prices. The U.S. Stock Database from the Center for Research in Security Prices (hereafter, CRSP) provides stock market data for equity securities traded on the major U.S. stock exchanges. CRSP contains information such as price, quote, market capitalization, shares outstanding, trading volume, etc. I use CRSP to construct companies' financial variables.

Sample Construction. In my estimation sample, I include individuals who have ever worked as senior executives in S&P 500 companies between (and including) 2007 and 2015 and merge the data in ExecuComp and BoardEx.[26] The resulting sample consists of a total of 4,192 individuals. In constructing employment history, I code the time when an individual is not an executive in an S&P 500 firm as "unemployment". This should not be interpreted literally; instead it should be interpreted as not being employed in the specific labor market under study. Additionally, as my model rules out downward job-to-job transition, I code a transition from an executive job in a high

[25] A proxy statement is a document containing the information the Securities and Exchange Commission (SEC) requires companies to provide to shareholders so they can make informed decisions about matters that will be brought up at an annual or special stockholder meeting. It is also known as Form DEF 14A.

[26] I use a sample period after 2006 because of a change in the reporting rule of the executives' total compensation. In 2006, the Financial Accounting Standard (FAS) 123R changed the reporting requirements of the DEF 14A form. Under this new reporting regime, the cost of all employee stock options, as well as other equity-based compensation arrangements, have to be reflected in the financial statements based on the estimated fair value of the awards.

market cap company to another executive job in a low market cap company as if the individual goes through unemployment in between the jobs. Finally, I construct three evolving social networks. For each year t, a network includes all the connections established before t. Additionally, I assume that once a connection is formed, it will never be lost. The resulting education network is fixed over time, and the work and social-activity networks monotonically grow larger over time. It is useful to note that not all work connections are established as a result of movements in the executive labor market. Examples of exogenous work connections are those formed when one or both individuals serve on the board of directors or in lower management positions. In the end, I discuss the potential issue of missing connections in the constructed networks. On the one hand, connections with individuals out of the sample are not included. On the other hand, some connections between individuals in the sample may not be recorded in BoardEx if the information is not publicly available such as attending the same church. Due to these data limitations, the constructed networks are subnetworks of the "true" networks and may potentially lead to underestimation of the effect of referrals.

3.2 Summary Statistics

Executives and Firms. Table 1 reports the summary statistics for the executives and the firms in the data. Executives in S&P 500 firms have an average annual compensation of 6.09 million dollars, with a median of 3.89 million. Most of the individuals in the sample are between the ages of 38 and 61 at the beginning of the sample period, with both the mean and the median being around 49. Moreover, 91% of them are male. S&P 500 firms in the sample period have an average monthly stock return of 0.69% and an average market capitalization of 29.47 billion dollars. The size of the board of directors ranges from 7 to 15, with a mean of 10.96. The average proportion of male directors is 85%, and the average proportion of foreign directors is 11%.

Networks. Table 2 reports the summary statistics for the social networks. For each network, I report the distributions of degree (defined in (2)) and clustering coefficient (defined in (3)) at the beginning (2007), the middle (2011), and the end (2015) of the sample period. The education network does not change over time. About half of the individuals have no school friends. Conditional on having school friends, the average number of school friends is 6.35. The work network grows moderately; the average number of work friends increases from 16.40 in 2007 to 24.76 in 2015. The social-activity network grows more rapidly; the average number of social-activity friends increases from 0.45 in 2007 to 13.57 in 2015. These new social-activity connections are formed unevenly. About half of the individuals have no social-activity connections throughout the entire sample period, whereas the top 5% added more than 60 new connections.

Variable	Mean	5%	25%	50%	75%	95%
			Panel A: Executives			
Compensation	6.09	1.15	2.31	3.89	7.27	17.68
Age in 2007	48.98	38	44	49	54	61
Male	0.91					
			Panel B: Firms			
Stock Return	0.69%	-0.62%	0.35%	0.71%	1.11%	2.21%
Market Capitalization	29,472.24	4,698.01	7,481.18	13,489.17	29,597.99	121,726.30
Board Size	10.96	7.80	9.31	10.87	12.18	14.81
Board Gender Ratio	0.85	0.72	0.81	0.85	0.89	0.97
Board Nationality Mix	0.11	0	0	0.07	0.18	0.38

Table 1: Summary Statistics: Executives and Firms

Notes: (1). Compensation is the the total annual compensation reported in the SEC filings. It is the sum of the salary, bonus, the value of stock and option awards, non-equity incentive plan compensation, change in pension value and nonqualified deferred compensation earnings, and other compensations such as perquisites and personal benefits etc. Compensation reported in this table is adjusted for inflation, in the unit of 2015 U.S. million dollar. (2). Stock return is monthly return. (3). Market capitalization is adjusted for inflation, in the unit of 2015 U.S. million dollar. (4). Board size is the number of directors on the board. (5). Board gender ratio is the proportion of male directors. (6). Board nationality mix is the proportion of foreign directors.

These networks exhibit two key features. The first feature is high level of clustering.[27] The average clustering coefficient conditional on $d > 0$ ranges from 0.40 to 0.65, meaning for an average individual more than 40% of his/her friends are also friends with one another.[28] In comparison, the average clustering coefficient for a randomly formed network is around 0.05. The second feature is sorting on degree, i.e., high-degree individuals' friends are often high-degree, and low-degree individuals' friends are often low-degree. The correlations between own degree and the average of friends' degrees are 0.85, 0.69, and 0.74 respectively for the education, work, and social-activity networks. In comparison, the correlation for a randomly formed network is around 0.20.

Labor Market Transitions. Tables 3 and 4 report the summary statistics for the labor market transitions. Each year, the average fraction of non-executives is 47.21%, and their annual transition rate from non-executive to executives (UE) is 16.09%. The average fraction of executives is 52.79%, their annual transition rate to executives in a more productive firm (EE) is 0.72%, and their annual transition rate to non-executives (EU) is 13.03%. Career advancement is defined as either a UE

[27]The definition of a connection requires two individuals to have overlapping time in an organization. Therefore, with people joining and leaving at staggered time, the resulting networks are not collections of cliques (a clique is a subgraph such that every two distinct nodes are connected; all the nodes in a clique have clustering coefficients of 1).

[28]Clustering coefficient is meaningful only for individuals with friends. Clustering coefficient for an individual with no friend is defined to be 0.

Variable	Mean	Mean Conditional on $d > 0$	5%	25%	50%	75%	95%
		Panel A: Network Y^{Edu}					
d^{Edu}	3.67	6.35	0	0	1	5	24
c^{Edu}	0.27	0.47	0	0	0	0.52	1
		Panel B: Network Y^{Work}					
$d^{Work,2007}$	16.40	17.08	1	7	11	20	52
$c^{Work,2007}$	0.63	0.65	0	0.41	0.64	0.92	1
$d^{Work,2011}$	20.76	20.97	5	9	14	25	61
$c^{Work,2011}$	0.60	0.61	0.20	0.39	0.58	0.86	1
$d^{Work,2015}$	24.76	24.76	7	11	18	31	69
$c^{Work,2015}$	0.56	0.56	0.21	0.35	0.52	0.77	1
		Panel C: Network Y^{Social}					
$d^{Social,2007}$	0.45	3.87	0	0	0	0	3
$c^{Social,2007}$	0.04	0.40	0	0	0	0	0.41
$d^{Social,2011}$	4.39	12.25	0	0	0	2	26
$c^{Social,2011}$	0.18	0.52	0	0	0	0.32	0.89
$d^{Social,2015}$	13.57	26.48	0	0	1	9	66
$c^{Social,2015}$	0.27	0.53	0	0	0	0.53	0.96
		Panel D: Network Y^{All}					
$d^{All,2007}$	20.42	20.92	2	8	15	26	60
$c^{All,2007}$	0.50	0.51	0.13	0.29	0.43	0.69	1
$d^{All,2011}$	28.41	28.57	5	12	20	37	78
$c^{All,2011}$	0.46	0.46	0.16	0.27	0.39	0.61	1
$d^{All,2015}$	41.15	41.15	8	15	27	49	130
$c^{All,2015}$	0.43	0.43	0.17	0.26	0.37	0.55	0.94

Table 2: Network Statistics

Notes: (1). d is degree, defined in Eq.(2). It is the number of an individual's friends. (2). c is clustering coefficient, defined in Eq.(3). It gives the proportion of an individual's friends who are friends with one another. (3). The time superscript for Y^{Edu} is omitted because it does not change in the sample period.

or EE transition. The annual rate of career advancement is 7.97%. The average duration of the non-executive spell is 3.97 years for the right-censored spells and 3.87 years for those ending with an executive job. The average duration of an executive job spell is 5.26 years for the right-censored spells, 3.30 years for those ending with transitions to non-executives, and 3.10 years for those ending with transitions to executive jobs in more productive firms.

	Conditional on U	Conditional on E		Unconditional
	UE Transition	EE Transition	EU Transition	Career Advancement
Annual Transition Rate	16.09%	0.72%	13.03%	7.97%

Table 3: Job Transition

Notes: (1). A UE transition is defined as the transition from non-executive to executive. (2). An EE transition is defined as the transition from one executive job to another. (3). An EU transition is defined as the transition from executive to non-executive. (4). A career advancement is defined as either a UE or EE transition.

Panel A: Unemployment Spell			
Type of Transition	Number of Spells	Mean of Spell Duration	Std. Dev. of Spell Duration
Right Censored	1,984	3.97	2.32
To Employment	2,547	3.87	2.37
Panel B: Job Spell			
Type of Transition	Number of Spells	Mean of Spell Duration	Std. Dev. of Spell Duration
Right Censored	2,208	5.26	3.08
To Unemployment	2,422	3.30	2.15
To Another Job	127	3.10	1.85

Table 4: Durations

Notes: (1). An unemployment spell is defined as the period when an individual is not an executive. (2). The unit of time is one year.

3.3 Reduced-form Evidence

In this part, I present suggestive evidence on the presence of job referrals in the executive labor market and supporting evidence on my model assumption of no immediate "chain effect" from friends of friends.

Correlation in Friends' Compensation. This analysis shows that socially connected executives' compensation is correlated, which is consistent with the implication of my job referral model. I regress an executive's compensation on the average of his/her friends' compensation and a set of covariates according to the following Spatial Auto-Regressive model (SAR):

$$\ln w_{i,t} = \rho \frac{\sum_{j \in N^{k,t}(i)} \ln w_{j,t}}{d_i^{k,t}} + (x_{i,t})^T \beta + u_i + v_t + \epsilon_{i,t}, \tag{17}$$

where $\ln w_{i,t}$ is executive i's log compensation in year t, and $\frac{\sum_{j \in N^{k,t}(i)} \ln w_{j,t}}{d_i^{k,t}}$ is the average of i's executive friends' log compensation in year t. $x_{i,t}$ is a vector of exogenous covariates including the executive's age, his/her employer's financial variables and governance variables, and industry dummy. u_i is the individual fixed-effects, and v_t is the time fixed-effects. $\epsilon_{i,t}$ is an i.i.d. error term. Coefficient ρ, sometimes referred to as the "social interaction parameter", measures the percentage change in i's compensation in response to a one percentage increase to friends' compensation.

In this regression, $\frac{\sum_{j \in N^{k,t}(i)} \ln w_{j,t}}{d_i^{k,t}}$ is correlated with the error term $\epsilon_{i,t}$ because of simultaneous equations. To address the endogeneity problem, I use instrumental variables. The instruments are the exogenous covariates of friends and friends of friends. More precisely, (17) can be written in the following compact form

$$\ln w = \rho W \ln w + x\beta + u + v + \epsilon, \tag{18}$$

and I use Wx and W^2x as instruments for $W \ln w$.

Table 5 reports the results of the Spatial Auto-Regressive regression in (17), where different columns correspond to different network specifications. Column (1) uses the simplified network, which does not distinguish between different types of social connections; column (2) uses the school network; column (3) uses the work network; and column (4) uses the social-activity network. These results illustrate that, across all specifications of networks, an executive's compensation is positively correlated with his/her friends' compensation. This correlation is driven by neither the correlation in time-invariant individual characteristics nor time- or industry-specific shocks because of the inclusion of individual, time, and industry fixed-effects. The positive correlation in socially connected executives' compensation supports my model's prediction that executives with better jobs send better quality referrals, increasing their friends' compensation.

Co-movement in Friends' Job Transitions. This analysis shows that individuals are more likely to make career advancement when their executive friends leave their current jobs, which suggests the presence of referrals. Table 6 reports the results of a Logit regression. It shows that an individual is more likely to make an upward transition (UE or EE) when there is a friend(s) who experiences job-to-job transition (EE) or job separation (EU) in the same or the previous year. This corresponds to internal referrals following direct job arrivals or job separations. The individual fixed effects control for time-invariant individual characteristics, and the year fixed effects rule out

	Dependent variable: log(Comp)			
	(1)	(2)	(3)	(4)
Mean (Friend's log(Comp))	0.3858***			
	(0.0274)			
Mean (School Friend's log(Comp))		0.2105***		
		(0.0404)		
Mean (Work Friend's log(Comp))			0.5573***	
			(0.0364)	
Mean (Social-Activity Friend's log(Comp))				0.0207*
				(0.0082)
Observations	18,481	18,481	18,481	18,481
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes

Table 5: Correlation of Socially Connected Executives' Compensation

Notes: (1). This table summarizes the results of a Spatial Auto-Regressive regression with individual, year, and industry fixed effects, where the dependent variable is the log of annual compensation. (2). Column (1) uses the simplified network, which does not distinguish between types of social connections; column (2) uses the school network; column (3) uses the work network; and column (4) uses the social-activity network. (3). Instrumental variables are used to address the endogeneity problem arising from simultaneous equations. (4). Other covariates include age, age^2, and dummy for serving on the board of directors; the board's size, gender ratio, and nationality mix; and firm's lagged log(market cap) and stock return. (5). Standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

time-specific shocks.[29] It is useful to note that given the small number of EE transitions, most of the co-movements are in the form of individuals making upward transitions following friends' downward transitions, which are unlikely to be driven by cohort effect or industry-specific shocks. To interpret the quantitative results, note that the coefficients in Logit regressions represent log odds ratios. For example, a log odds ratio of 0.6957 is equivalent to an odds ratio of 2.005, meaning the odds of career advancement for an individual whose friends change jobs is twice that of an individual whose friends do not change jobs. In terms of the probability of career advancement, it corresponds to a 4.7% increase in absolute terms according to the OLS estimates in Table B2.

	Dependent variable: $\mathbb{1}$*(Career Advancement)*$\in \{0,1\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}$ (Job Transition Among Executive Friends)	0.6957^{***}			
	(0.0934)			
... from School		0.2706^{***}		
		(0.0769)		
... from Work			0.7954^{***}	
			(0.0870)	
... from Social Activity				0.2508^{**}
				(0.0944)
Observations	19,360	19,360	19,360	19,360
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Table 6: Co-movement in Socially Connected Executives' Job Transitions

Notes: (1). This table summarizes the results of a Logit regression with year and individual fixed effects, where the dependent variable is whether an individual experiences a career advancement. (2). A career advancement is defined as either a transition from non-executive job to executive job (UE) or a transition between executive jobs with productivity increase (EE). (3). Variable $\mathbb{1}$ (Job Transition Among Executive Friends) is a dummy variable that equals one if any friend experiences a transition between executive jobs with productivity increase (EE) or a transition from executive job to non-executive job (EU) in the same year or the previous year. (4). The scope of friends varies for different network specifications. Column (1) uses the simplified network, which does not distinguish between types of social connections; column (2) uses the school network; column (3) uses the work network; and column (4) uses the social-activity network. (5). The Logit regression with individual fixed effects uses a subsample of the individuals with the same total number of career advancements: in this case, individuals with exactly one career advancement during the sample period. (6). Other covariates include age^2, age^3, whether an individual was an executive in the previous year, and the productivity of the job in the previous year if the individual was an executive. (7). Standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Lack of Co-movement Beyond Immediate Friends. This analysis shows the lack of co-movement in job transitions beyond immediate friends, which motivates my model assumption

[29] The Logit regression with individual fixed effects uses a subsample of the individuals with the same total number of career advancements: in this case, individuals with exactly one career advancement during the sample period. As a robustness check, the results of an OLS regression with year and individual fixed effects using the full sample are reported in Table B2.

ruling out immediate "chain effect" from friends of friends. Table 7 shows that an individual's probability of career advancement responds to the job-to-job transition (EE) and job separation (EU) of immediate friends, but not to that of friends of friends.[30] The results suggest that, in the short run, individuals do not significantly benefit from friends of friends.

Dependent variable: $\mathbb{1}$*(Career Advancement)*$\in \{0,1\}$	(1)	(2)
$\mathbb{1}$(Job Transition Among Executive Friends)	0.6338***	
	(0.1401)	
$\mathbb{1}$(Job Transition Among Executive Friends of Friends)	0.2637	
	(0.4536)	
Number of Job Transition Among Executive Friends		0.0595**
		(0.0216)
Number of Job Transition Among Executive Friends of Friends		0.0035
		(0.0031)
Observations	19,360	19,360
Individual FE	Yes	Yes
Year FE	Yes	Yes

Table 7: Lack of Co-movement in Job Transition Beyond Immediate Friends.

Notes: (1). This table summarizes the results of a Logit regression with year and individual fixed effects, where the dependent variable is whether an individual experiences a career advancement. (2). A career advancement is defined as either a transition from non-executive job to executive job (UE) or a transition between executive jobs with productivity increase (EE). (3). The simplified network, which does not distinguish between types of social connections, is used to define friends as well as friends of friends. (4). Friends of friends exclude immediate friends and only include those who are not immediate friends. (5). Variable $\mathbb{1}$(Job Transition Among Executive Friends (or Friends of Friends)) is a dummy variable that equals one if any friend (or friend of friends) experiences a transition between executive jobs with productivity increase (EE) or a transition from executive job to non-executive job (EU) in the same year or the previous year. (6). Variable Number of Job Transition Among Friends (of Friends of Friends) is the total number of EE or EU transitions among friends (or friends of friends) in the same year or the previous year. (7). The Logit regression with individual fixed effects uses a subsample of the individuals with the same total number of career advancements: in this case, individuals with exactly one career advancement during the sample period. (8). Other covariates include age^2, age^3, whether an individual was an executive in the previous year, and the productivity of the job in the previous year if the individual was an executive. (9). Standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

[30] I exclude immediate friends in defining friends of friends and only include those who are not immediate friends. Additionally, in this exercise, I use the simplified network Y^{All} which does not distinguish between different types of social connections because it is problematic to restrict to friends of friends in each of the networks $Y^{Edu}, Y^{Work}, Y^{Social}$. Such restriction excludes a significant number of friends of friends across different types of social connections, for example, a work friend of a school friend.

4 Estimation Strategy

In this section I present my estimation method. In my empirical application, I use the following specifications. The productivity distribution upon job arrival, F, is truncated log normal:

$$\ln p \sim \mathcal{N}(\mu_p, \sigma_p), \quad p \in [p^{min}, p^{max}]. \tag{19}$$

Following Terviö (2008) and Gabaix and Landier (2008), I use the firm's market capitalization as the empirical counterpart for productivity. Additionally, I set p^{min} and p^{max} to be the observed lowest and highest market capitalization. An executive's bargaining power is a function of his/her number of executive friends:

$$\beta = \frac{\exp(\beta_0 + \beta_1 d^{All \cap e})}{1 + \exp(\beta_0 + \beta_1 d^{All \cap e})}, \tag{20}$$

where $d_i^{All \cap e} = |\{j | j \in N^{All}(i), s_j = e\}|$ is i's number of executive friends. Executive ability a follows a log normal distribution:

$$\ln a \sim \mathcal{N}(\mu_a, \sigma_a). \tag{21}$$

Log compensation has additive i.i.d. measurement error ϵ:

$$\epsilon \sim \mathcal{N}(0, \sigma_\epsilon).\text{[31]} \tag{22}$$

The unit of time is one year, and I set the continuous time discount rate ρ to be 0.05 (an annual discount rate of 0.95).[32]

I estimate the parameters via a minimum-distance estimator, which follows Imbens and Lancaster (1994) and Petrin (2002). The objective function is based on the generalized method of moments (GMM) with three sets of moments. The first two sets of moments are in the form of log-likelihood, which I aim to maximize by requiring that the first derivatives should equal zero. Specifically, the first set of moments characterizes the likelihood of the executives' compensation immediately after career advancement; the second set of moments characterizes the likelihood of the productivities of socially isolated workers' first jobs. The third set of moments are in the form of traditional moments, characterizing executives' labor market transitions. Accordingly, parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ can be grouped into three sets, where $\theta_1 = \langle \alpha, \beta_0, \beta_1, \ln(b), \sigma_a, \sigma_\epsilon \rangle$ is mostly relevant to the first set of moments, $\theta_2 = \langle \mu_p, \sigma_p \rangle$ is most relevant to the second set, and

[31] This measurement error should be interpreted as benefits in the compensation package that are not included in the calculation of the value of total compensation. For example, perks such as traveling with a private jet.

[32] Flinn and Heckman (1982) shows that it is difficult to separately identify the discount factor from the flow unemployed income in standard search models.

$\theta_3 = \langle \lambda_0, \lambda_1, \delta, \pi_0, \pi_1, \nu_u, \omega^{Edu}, \omega^{Work} \rangle$ is most relevant to the third set.

Specifically, the targeted moments are

$$\mathcal{M}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial \ln(L_1(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta_1}} \\ \frac{\partial \ln(L_2(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta_2}} \\ \boldsymbol{m} - \mathbf{E}[\boldsymbol{m}; \boldsymbol{\theta}] \end{bmatrix}, \tag{23}$$

where L_1 is the first likelihood of executive compensation, which I discuss in more detail in Section 4.1; and L_2 is the second likelihood of firm productivity, which I discuss in more detail in Section 4.2; and $\boldsymbol{m}$ is a vector of moments for labor market transitions, which I describe in Section 4.3. Then I estimate the parameters using the following objective function:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathcal{M}(\boldsymbol{\theta})' W \mathcal{M}(\boldsymbol{\theta}), \tag{24}$$

where W is the weight matrix. I choose W to be the diagonal elements of the optimal weight matrix.[33]

4.1 Likelihood of Executive Compensation

L_1 is the likelihood of the executives' compensation immediately after career advancement. When executive i experiences a career advancement (UE or EE transition) at time t, the observed compensation in the new job is given by

$$\ln w_{it} = \ln \phi_1(1, p_{it}^L, p_{it}^H, \Gamma_{it}) + \ln a_i + \epsilon_{it}, \tag{25}$$

where p_{it}^H is the productivity of the new job, p_{it}^L is the productivity of the old job for an EE transition or the reservation productivity calculated according to (14) for a UE transition, ϕ_1 is the equilibrium wage function (15), a_i is unobserved ability, and ϵ_{it} is unobserved measurement error. Let $u_i = \ln a_i - \mu_a$, then log compensation can be rewritten as

$$\ln w_{it} = \ln \phi_1(1, p_{it}^L, p_{it}^H, \Gamma_{it}) + \mu_a + u_i + \epsilon_{it}. \tag{26}$$

This is a standard random effect model (Maddala, 1971) satisfying

$$E(u_i | p_{it}^L, p_{it}^H, \Gamma_{it}) = 0, \tag{27}$$

[33] Altonji and Segal (1996) show that the asymptotic optimal weight matrix may induce bias in finite sample estimates.

$$E(\epsilon_{it}|p^L_{it}, p^H_{it}, \Gamma_{it}) = 0, \; E(u_i \epsilon_{it}|p^L_{it}, p^H_{it}, \Gamma_{it}) = 0. \tag{28}$$

The orthogonality condition (27) holds because of the model implication that worker i's ability is uncorrelated with his/her own and his/her friends' job dynamics, which is shown in Proposition 1. Condition (28) holds because of the independence assumption on the measurement errors.

Additionally,

$$E(u_i u_j|p^L_{it}, p^H_{it}, \Gamma_{it}) = \begin{cases} \sigma_a^2 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}, \quad E(\epsilon_{it}\epsilon_{js}|p^L_{it}, p^H_{it}, \Gamma_{it}) = \begin{cases} \sigma_\epsilon^2 & \text{if } i = j \text{ and } t = s \\ 0 & \text{otherwise} \end{cases}. \tag{29}$$

The complete expression of the likelihood function is given in Appendix C.1.

4.2 Likelihood of Firm Productivity

L_2 is the likelihood of the productivities of socially isolated workers' first jobs. The isolated unemployed workers only face direct job arrivals with productivity distribution F. Upon arrival, they accept jobs above their reservation productivity p_0 as described by (14). Therefore, their accepted jobs are distributed independently according to F truncated below by p_0. The complete expression of the likelihood function is given in Appendix C.2.

4.3 Moments of Labor Market Transitions

Referrals introduce interdependence in friends' jobs and thus interdependence in labor market transitions, which is difficult to characterize by likelihood. Instead, I simulate the model and target some moments of the transitions.

Targeted Moments. I include the following standard moments: the frequencies of UE, EE, and EU transitions, and the means and the standard deviations of the employment and the unemployment spells. In addition, I include moments characterizing the co-movement in socially connected individuals' job transitions to capture the effects of referrals. These moments include, for each type of social connection, the fraction of unemployed individuals experiencing UE transition conditional on whether they have friends experiencing EU transition, the fraction of unemployed individuals experiencing UE transition conditional on whether they have friends experiencing EE transition, the fraction of employed individuals experiencing EE transition conditional on whether they have friends experiencing EU transition, and the fraction of employed individuals experiencing EE transition conditional on whether they have friends experiencing EE transition.

Simulation. In the model, direct job arrivals and job separations are independent continuous-time Poisson processes.[34] In the simulation, I preserve the continuous-time setting by simulating the waiting times.[35] After the simulation, I discretize the simulated data so that it is comparable to the discrete-time observations in the data. The simulation procedure is described in the following.

The simulation follows an iterative procedure. Each iteration k starts at time t^{k-1} with the employment/unemployment status and job productivity of all the workers being $\mathbf{s} = (s_1, ..., s_n)$ and $\mathbf{p} = (p_1, ...p_n)$. For each unemployed worker, simulate one waiting time for direct job offer $\Delta t_i^{\text{ArrivalU},k}$; for each employed worker, simulate one waiting time for direct job offer $\Delta t_i^{\text{ArrivalE},k}$ and another waiting time for job separation $\Delta t_i^{\text{Separation},k}$. The shortest waiting time among all is $\Delta t^{*,k} = \min\left\{ \{\Delta t_i^{\text{ArrivalU},k} : s_i = u\}, \{\Delta t_i^{\text{ArrivalE},k}, \Delta t_i^{\text{Separation},k} : s_i = e\} \right\}$.[36] Set $t^k = t^{k-1} + \Delta t^{*,k}$. If $t^k > T$, where T is a pre-specified length of time, the simulation is completed. Otherwise, proceed with the following. Record t^k as the time of the k-th event, along with the corresponding event type $\in$ {Arrival for U, Arrival for E, Separation} and person i^*. If the event is an arrival for an unemployed worker i^*, first sample the job's productivity, then the worker makes the decision to accept or reject (s_{i^*} and p_{i^*} change accordingly). If the event is an arrival for an employed worker i^*, first sample the job's productivity, then the worker makes the decision to accept or reject (s_{i^*} and p_{i^*} change accordingly), which leads to an internal or external referral with probability π_1. If a referral is generated, sample a recipient $j^* \in N^{All,t^k}(i^*)$, then the recipient makes the decision to accept or reject (s_{j^*} and p_{j^*} change accordingly). If the event is a job separation, i^* becomes unemployed (s_{i^*} and p_{i^*} change accordingly), which leads to an internal referral with probability π_0. If a referral is generated, sample a recipient $j^* \in N^{All,t^k}(i^*)$, then the recipient makes the decision to accept or reject (s_{j^*} and p_{j^*} change accordingly). This completes iteration k. The next iteration starts at time t^k with the new $\mathbf{s}$ and $\mathbf{p}$. Note that the memoryless property allows me to "reset the Poisson clock" in each iteration.[37] Specifically, in each iteration, I only use the shortest waiting time and do not carry the rest of the waiting times to the subsequent iterations. For the subsequent iteration, I simulate a new set of waiting times to generate the arrival time and the type of the event.

[34] The continuous-time setting rules out simultaneous events.

[35] For a Poisson process with arrival rate λ, at any given time point, the waiting time until the next arrival follows an exponential distribution with mean $\frac{1}{\lambda}$, i.e., the cdf of waiting time is $F(t) = 1 - e^{-\lambda t}$.

[36] In the implementation, I use a slightly different but statistically equivalent approach to reduce computational intensity. Instead of sampling separate arrival times for each person, I only sample three order statistics $t^{\text{ArrivalU}} = \min\{t_i^{\text{ArrivalU}} : s_i = u\}$, $t^{\text{ArrivalE}} = \min\{t_i^{\text{ArrivalE}} : s_i = e\}$, $t^{\text{Separation}} = \min\{t_i^{\text{Separation}} : s_i = e\}$. t^{ArrivalU} follows an exponential distribution with mean $\frac{1}{|\{i:s_i=u\}|\cdot\lambda_0}$; similarly, t^{ArrivalE} follows an exponential distribution with mean $\frac{1}{|\{i:s_i=e\}|\cdot\lambda_1}$, and $t^{\text{Separation}}$ follows an exponential distribution with mean $\frac{1}{|\{i:s_i=e\}|\cdot\delta}$.

[37] Memoryless refers to the property that the distribution of the waiting time (until an arrival) does not depend on how much time has elapsed already.

4.4 Identification

In this part, I provide heuristic arguments for identification of the structural parameters. The first set of parameters, $\theta_1 = \langle \alpha, \beta_0, \beta_1, \ln(b), \sigma_a, \sigma_\epsilon \rangle$, are identified from executive compensation immediately after career advancement. The utility parameter α governs the inter-temporal incentives, and it is therefore identified from the curvature of the log compensation function with respect to log productivity.[38] The bargaining parameters β_0 and β_1 are identified from the overall correlation between log compensation and log productivity, and the intuition is the following. In one extreme case of no bargaining power, an executive's initial compensation is negatively correlated with employer productivity because he/she trades lower present compensation for higher continuation value. In the other extreme case of full bargaining power, an executive's compensation is equal to the marginal product, and thus is positively correlated with employer productivity. Unemployed productivity $\ln(b)$ is identified from the within-person difference between compensation after UE transition and EE transition.[39] The dispersion parameter of ability σ_a is identified from the covariance of within-person residual terms, and the dispersion parameter of measurement error σ_ϵ is identified from the variance of the residual terms net of σ_a^2.

The second set of parameters, the location and dispersion parameters of productivity distribution, $\theta_2 = \langle \mu_p, \sigma_p \rangle$, are identified from the observed distribution of the productivities of socially isolated workers' first jobs. With the log-normal parametric assumption, they are identified from the mean and variance of the observed distribution.

The third set of parameters, $\theta_3 = \langle \lambda_0, \lambda_1, \delta, \pi_0, \pi_1, \nu_u, \omega^{Edu}, \omega^{Work} \rangle$, are identified from the labor market transitions. The job separation rate δ is identified from the duration of employment spell and the frequency of employment-to-unemployment transition. The probability of referral following job separation π_0 is identified from the co-movement of individuals' career advancement and their friends' job separations. Similarly, the probability of referral following job arrival π_1 is identified from the co-movement of individuals' career advancement and their friends' job-to-job transitions. The probability of referring an unemployed friend ν_u is identified by comparing the co-movement of individuals' unemployment-to-employment transitions and their friends' job-to-job transitions or job separations vs. the co-movement of individuals' job-to-job transitions and their friends' job-to-job transitions or job separations, given the productivity distribution. The job arrival rates λ_0 and λ_1 are identified from the frequency of unemployment-to-employment transition and job-to-job transition given the job separation rate δ, the referral parameters (π_0, π_1, ν_u), and the productivity

[38] $\frac{1}{\alpha}$ gives the inter-temporal elasticity of substitution.

[39] When $\alpha \neq 1$, $\ln(b)$ and $\mathbb{E}\ln(a)$ can be separately identified even without within-person comparison. The reason is that $\mathbb{E}\ln(a)$ enters the expression of $\ln w$ linearly, while $\ln(b)$ enters non-linearly.

distribution parameters (μ_p, σ_p). The probabilities of referring through different types of social connections, ω^{Edu} and ω^{Work}, are identified from the difference in co-movement pattern in the three different social networks.[40]

5 Estimation Results

5.1 Parameter Estimates

Table 8 reports the parameter estimate. The estimates of the bargaining parameters show that the executives generally have high bargaining power, and their bargaining power is increasing in their number of executive friends. Most of the individuals' bargaining powers range from 0.62 to 0.79.[41] The estimate of the unemployment productivity implies an annual income of 0.52 million dollars for an average-ability individual in a non-executive job. The estimated offer arrival rates for non-executives and executives are 0.130 and 0.006 respectively, implying average waiting times of 7.67 years and 165 years respectively for direct job offers. The estimated referral probabilities after job separation and job arrival are 0.228 and 0.794 respectively. The referral probability after job separation is lower than that after job arrival potentially because some separations are forced, and in these cases there is no chance for referral. The estimated probability of referring non-executive friends is 0.626, and that of referring executive friends is 0.374. Therefore, the model implication that the executives are more selective in accepting jobs, combined with the empirical estimates that they receive fewer direct offers as well as referrals, jointly contribute to the low number of job-to-job transitions observed in the data. Finally, the estimated probabilities of referral through different types of networks show that the executives are most likely to refer their work friends (with probability 0.653), then their social-activity friends (with probability 0.192), and lastly their school friends (with probability 0.155).

5.2 Statistical and Economical Significance of Referrals

It is useful to note that the estimated referral probabilities π_0 and π_1 are both statistically and economically significantly different from zero. A Wald test of the joint hypotheses $H_0 : \pi_0 = \pi_1 = 0$ rejects a nested model with no referral, $\chi^2(2) = 1171618$, p=0.[42] To further understand the economic significance of referrals, I decompose the sources of job dynamics through simulation and report the results in Table 9. The table shows that for non-executives' UE transitions, 72.15%

[40] $\omega^{Social} = 1 - \omega^{Edu} - \omega^{Work}$.

[41] 99% of the individuals have fewer than 50 *executive* friends.

[42] The critical value for a significance level of 0.001 is 13.82.

Parameter	Notation	Estimate	Std. Err.
Panel A: Parameters in $\theta_1 = \langle \alpha, \beta_0, \beta_1, \ln(b), \sigma_a, \sigma_\epsilon \rangle$			
CRRA Coefficient	α	1.393	0.13340
Constant in the Logit Function for Bargaining Power	β_0	0.500	0.38150
Slope in the Logit Function for Bargaining Power	β_1	0.017	0.00240
Log (Unemployment Productivity)	$\ln(b)$	6.866	0.07653
Dispersion Para. of Executives' Lognormal Ability Dist.	σ_a	0.355	0.00725
Dispersion Para. of Compensations' Normal Measurement Error Dist.	σ_ϵ	0.634	0.00088
Panel B: Parameters in $\theta_2 = \langle \mu_p, \sigma_p \rangle$			
Location Para. of Firms' Lognormal Productivity Dist.	μ_p	9.550	0.01772
Dispersion Para. of Firms' Lognormal Productivity Dist.	σ_p	1.240	0.00083
Panel C: Parameters in $\theta_3 = \langle \lambda_0, \lambda_1, \delta, \pi_0, \pi_1, \nu_u, \omega^{Edu}, \omega^{Work} \rangle$			
Offer Arrival Rate for the Unemployed	λ_0	0.130	0.00004
Offer Arrival Rate for the Employed	λ_1	0.006	0.00001
Job Separation Rate	δ	0.147	0.00004
Probability of Referral Following Job Separation	π_0	0.228	0.00045
Probability of Referral Following Job Arrival	π_1	0.794	0.00073
Probability of Referring an Unemployed Friend	ν_u	0.626	0.00051
Probability of Referring a School Friend	ω^{Edu}	0.155	0.00013
Probability of Referring a Work Friend	ω^{Work}	0.653	0.00014

Table 8: Parameter Estimates

result from direct job offers, 1.60% from referrals following friends' job arrivals, and 26.26% from referrals following job separations. For executives' EE transitions, the proportions are 34.00%, 3.07%, and 62.93% respectively; and for their wage raises, the proportions are 17.91%, 5.37%, and 76.73% respectively. These numbers show the following patterns. First, for all three types of transitions, referrals following friends' job separations lead to far more transitions than those following friends' job arrivals. This is due to the large number of job separations compared to the number of on-the-job offer arrivals, even though the probability of sending a referral is lower in the former case than the latter. Second, referrals contribute to higher proportions of job dynamics for executives (EE transition and wage raise) than for non-executives (UE transition). This suggests that the ratio of referral arrival rate and direct job offer arrival rate is larger for the executives than the non-executives. Third, for executives, referrals contribute to a higher fraction of wage raises than job-to-job transitions. The reason is that compared with direct offers, referred jobs are less likely to be more productive than executives' current jobs as they are no better than friends' own jobs.

Panel A: Source of Accepted Offer in UE Transition			
	Direct Offer	Referrals Following Friends' Job Arrivals	Referrals Following Friends' Job Separations
Number	1851.7	41.0	673.9
%	72.15	1.60	26.26
Panel B: Source of Accepted Offer in EE Transition			
	Direct Offer	Referrals Following Friends' Job Arrivals	Referrals Following Friends' Job Separations
Number	37.6	3.4	69.6
%	34.00	3.07	62.93
Panel C: Source of Competing Offer in Wage Raise			
	Direct Offer	Referrals Following Friends' Job Arrivals	Referrals Following Friends' Job Separations
Number	34.7	10.4	148.7
%	17.91	5.37	76.73

Table 9: Sources of Job Dynamics: Direct Offers and Referrals

Notes: The reported numbers are averages over 100 simulations.

5.3 Model Fit

Figure 2 shows the fit for executive compensation conditional on firm productivity and number of executive friends. The model is able to replicate the empirical pattern that executive compensation is increasing in both the firm productivity and the number of executive friends. Figure 3 shows the fit for the productivity distribution, in which I compare the empirical distribution of the socially isolated individuals' first jobs in the data and the parametrically estimated probability density function in the model. The model is capable of capturing the general shape of the empirical histogram. Table 10 shows the fit for the moments in job transitions. The model is able to replicate the spell durations and the frequencies of transitions in the model. Additionally, it is able to replicate the co-movement pattern that the probability of career advancement is higher when some friends leave current jobs than when no friend does. I show the fit for co-movement moments for each type of social connection separately in an additional Table D3 in Appendix D.

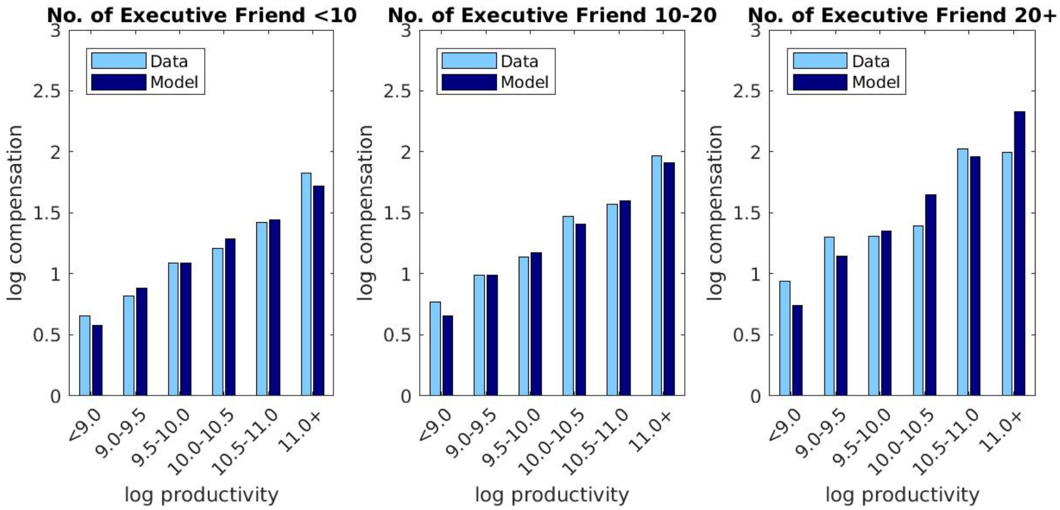

Figure 2: Model Fit for Executive Compensation

Notes: (1). The empirical measurement for firm productivity is market capitalization. (2). The unit for compensation and market capitalization are both 2015 U.S. million dollar.

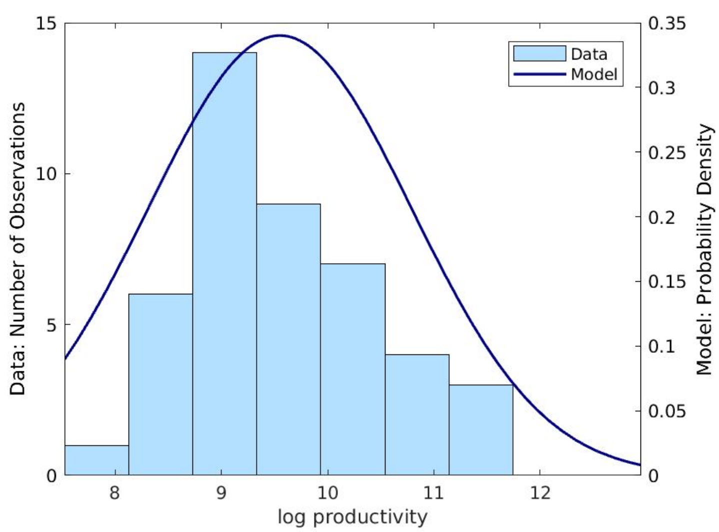

Figure 3: Model Fit for Productivity Distribution

Notes: (1). The empirical measurement for firm productivity is market capitalization, and the unit is 2015 U.S. million dollar. (2). The histogram plots the distribution of the socially isolated individuals' first jobs' productivities.

Moments	Data	Model
Mean of Employment Spell	3.9193	4.1457
Std. Deviation of Employment Spell	2.3548	2.8232
Mean of Unemployment Spell	4.4119	4.0953
Std. Deviation of Unemployment Spell	2.8509	2.8276
Frequency of UE Transition	0.6076	0.6109
Frequency of EE Transition	0.5778	0.5832
Frequency of EU Transition	0.0303	0.0296
Fraction of unemployed individuals experiencing UE transition		
... if there is no friend experiencing EU transition	0.0812	0.1344
... if there are friends experiencing EU transition	0.1717	0.1753
... if there is no friend experiencing EE transition	0.1541	0.1603
... if there are friends experiencing EE transition	0.1870	0.1926
Fraction of employed individuals experiencing EE transition		
... if there is no friend experiencing EU transition	0.0028	0.0031
... if there are friends experiencing EU transition	0.0077	0.0078
... if there is no friend experiencing EE transition	0.0054	0.0053
... if there are friends experiencing EE transition	0.0130	0.0156

Table 10: Model Fit: Moments in Job Transitions

Notes: (1). A UE transition is defined as the transition from non-executive to executive. An EE transition is defined as the transition from one executive job to another with a productivity increase. An EU transition is defined as the transition from executive to non-executive. (2). Frequency of transition is calculated as the total number of transitions during sample period divided by the number of individuals. (3). The table reports the co-movement moments (last 8 rows) for the simplified network, which does not distinguish between different types of social connections.

6 Counterfactual Experiments

In this section, I use the estimated model to perform two sets of counterfactual experiments. First, I evaluate the effect of referrals by varying the referral probabilities. Second, I examine the effect of network structure by comparing the observed networks with randomly formed networks.

6.1 Varying Referral Probabilities

In this set of counterfactual experiments, I vary the referral probabilities, π_0 and π_1. In the baseline model, these probabilities are estimated to be $\pi_0 = 0.228$ for referrals following friends' job separations and $\pi_1 = 0.794$ for referrals following friends' job arrivals. In two counterfactual experiments, I set the referral probabilities to be zero ($\pi_0 = 0,\ \pi_1 = 0$) and one ($\pi_1 = 1,\ \pi_1 = 1$) respectively.

I compare the outcomes in the counterfactual models with the baseline model, in terms of the number of referrals received and the executive's lifetime utility measured by consumption equivalence. In Table 11, I break down the individuals by their network degrees and report their outcomes. As expected, for both counterfactual experiments and both outcome measures, referrals have stronger effects for higher degree individuals. High degree individuals have higher gains from referrals partly because they receive more referrals, and partly because they can extract a higher share of the surplus in wage bargaining once they receive a referral.[43]

In the case of setting referral probabilities to zero ("no referral"), the lowest degree group loses 0.17 referrals (in a total time span of 9 years), which is equivalent to a 2.62% reduction in annual income; the highest degree group loses 0.35 referrals, which is equivalent to a 6.93% reduction in annual income.

In the other case of setting referral probabilities to one ("mandatory referral"), the changes in outcomes are more significant. For all degree groups, the change generates more than 0.5 additional referrals, which is equivalent to a 6.73% increase in annual income for the lowest degree group and a 15.96% increase for the highest degree group.

[43] Recall that bargaining power depends on the number of executive friends. This heterogeneity generates, even in the absence referrals ($\pi_0 = \pi_1 = 0$), a difference in individual welfare (consumption equivalence) across degree groups.

Degree	$d < 20$	$20 \leq d < 40$	$40 \leq d < 60$	$60 \leq d < 80$	$d >= 80$
	Number of Referrals Received				
Baseline	0.17	0.22	0.25	0.31	0.35
$\pi_0 = 0, \pi_1 = 0$	0.00	0.00	0.00	0.00	0.00
Δ	-0.17	-0.22	-0.25	-0.31	-0.35
$\pi_0 = 1, \pi_1 = 1$	0.56	0.71	0.80	0.99	1.07
Δ	0.39	0.49	0.55	0.68	0.71
	Annual Consumption Equivalence (million \$)				
Baseline	1.35	1.43	1.53	1.62	1.82
$\pi_0 = 0, \pi_1 = 0$	1.32	1.37	1.46	1.51	1.68
$\%\Delta$	-2.62%	-3.59%	-4.26%	-6.31%	-6.93%
$\pi_0 = 1, \pi_1 = 1$	1.45	1.56	1.72	1.86	2.12
$\%\Delta$	6.73%	9.38%	12.19%	15.09%	15.96%
Group Size	1,684	1,188	593	289	438
mean(Degree)	11.58	28.30	48.40	68.62	145.60

Table 11: Counterfactual Experiment: Varying Referral Probabilities

Notes: (1). All the changes are calculated with respect to the baseline model, in which referral probabilities are $\pi_0 = 0.228$ and $\pi_1 = 0.794$. (2). The reported numbers are averages over 100 simulations. (3). The time span used in the simulation is 9 years, the same as the sample length. (4). d is the degree for $Y^{All,2015}$, the simplified network in the final sample period.

6.2 Random Networks

In this counterfactual experiment, I vary the network structure. Specifically, I generate a new set of education, work, and social-activity networks such that the degree of each individual in each year is the same as the observed networks (baseline), but the connections are formed at random.[44] I find that the welfare distribution is more unequal under the random networks, and I investigate the underlying mechanism through the lens of two local network statistics.

6.2.1 Differences in the Local Network Structures

Before I present the results, it is useful to first examine the difference between the structure of the observed networks and the random networks. By design, the two sets of networks have equal degrees for all individuals at all time. The main difference lies in two other local structures: friends' popularity and local clustering.

Friends' Popularity. I measure friends' popularity by average friend degree. Define i's average friend degree in network k to be

$$\frac{\sum_{j \in N^k(i)} {d^k}_j}{|N^k(i)|}, \tag{30}$$

the average degree of i's friends'. It can be interpreted as the level of "popularity" of friends.

The patterns of the difference in average friend degree between the two sets of networks are heterogeneous with respect to own degree. As described in Section 3.2, the observed networks exhibit sorting on degree. Low-degree individuals tend to have more low-degree friends in the observed network. Therefore, they have lower average friend degree in the observed network, compared with the random network. High-degree individuals tend to have more high-degree friends in the observed network. Therefore, they have higher average friend degree in the observed network, compared with the random network. Figure 4 breaks down the individuals by their own degrees and plots the distributions of average friend degrees for the observed and the random networks respectively.

Local Clustering. I measure the level of clustering by the local clustering coefficient defined in (3). It gives the fraction of an individual's friends who are also friends with one another.

The observed networks exhibit universally a far higher level of clustering than the random networks. Figure 5 plots the distributions of local clustering coefficients for the observed and the random networks respectively. In the observed networks, 88% of the individuals have clustering

[44] I describe the algorithm for simulating random network with a given degree sequence in Appendix E.

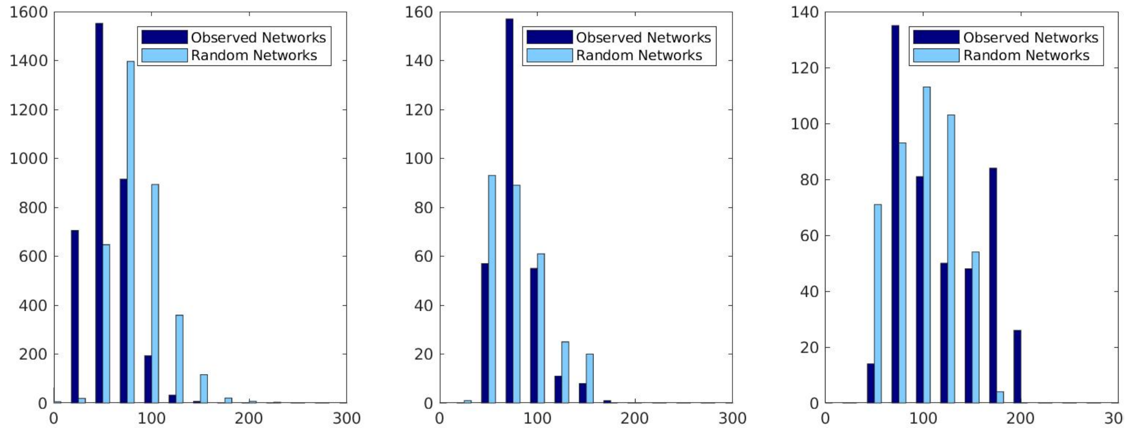

Figure 4: Distribution of Average Friend Degree for Low (Left), Medium (Middle), High (Right) Degree Individuals

Notes: (1). Average friend degree, defined in Eq.(30), gives the average degree of an individual's friends. (2). The figure on the left uses the subset of individuals with degree less than 60; the figure in the middle uses the subset of individuals with degree greater than or equal to 60 and less than 80; and the figure on the right uses the subset of individuals with degree greater than or equal to 80. (3). The degree and the average friend degree are calculated based on the simplified network in the final sample period, $Y^{All,2015}$, for the observed and the random network respectively.

coefficients above 0.2, and 50% above 0.35. In stark contrast, in the random networks, 97.50% of the individuals have clustering coefficients below 0.2, and 50% below 0.03.

6.2.2 Effects of Network Structures on Referrals

In this part, I consider a change from the observed networks to random networks with the same degree sequences and examine the resulting changes in referral patterns as well as executive welfare.

Overall Effects. Figure 6 shows that the random networks lead to greater inequality, compared with the observed networks. Under random networks, there are more individuals with low welfare and also more individuals with high welfare, which is a result of increased inequality in referrals. Table 12 breaks down the individuals by degree and reports the changes in their local network structures and labor market outcomes. The decrease in the clustering coefficient is significant for all degree groups. The average friend degree increases for the lower degree groups and decreases for the highest degree group. In terms of the labor market outcomes, low-degree individuals are worse off under the random networks, whereas high-degree individuals are better off, as shown in their welfare measured by consumption equivalence. For example, the lowest degree group receives 0.10 fewer referrals, which is equivalent to a 1.41% reduction in annual income. The highest degree

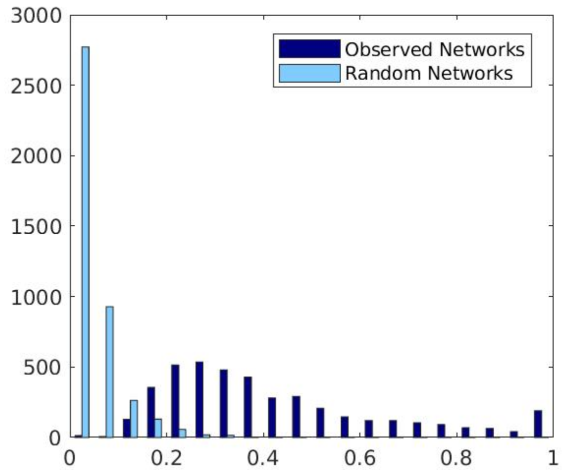

Figure 5: Distribution of Clustering Coefficient

Notes: (1). Clustering coefficient, defined in Eq.(3), gives the fraction of an individual's friends who are also friends with one another. (2). The clustering coefficients are calculated based on the simplified network in the final sample period, $Y^{All,2015}$, for the observed and the random network respectively.

group receives 0.20 more referrals, which is equivalent to a 3.52% increase in annual income.

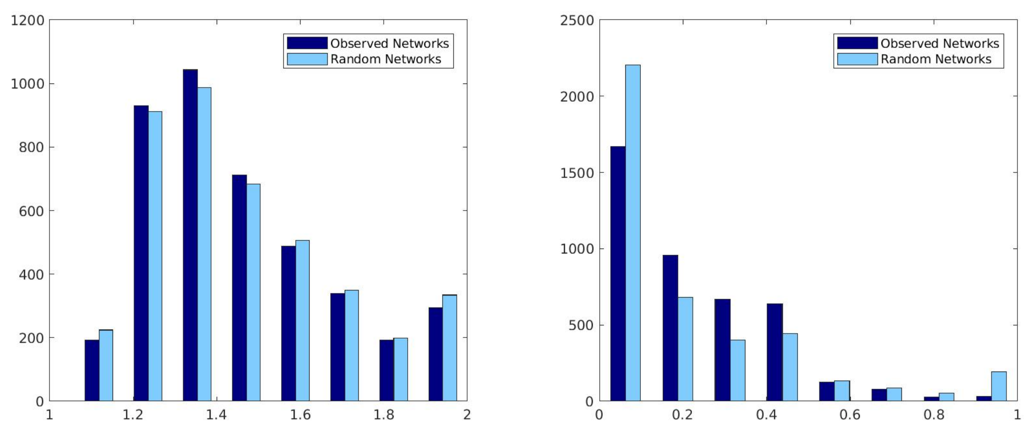

Figure 6: Distributions of Executive Welfare (Left) and Number of Referrals Received (Right)

Notes: (1). Executive welfare is measured by annual consumption equivalence in million \$. (2). In the plot of distribution of executive welfare, all individuals with consumption equivalence higher than 2 million are plotted at 2 million. (3). In the plot of distribution of the number of referrals received, all individuals with more than 1 referral are plotted at 1 referral. (4). The numbers are averages over 100 simulations. (5). The time span used in the simulation is 9 years, the same as the sample length.

Effects of Friend Popularity and Clustering by Degree Group. I investigate the mechanisms for the overall effects through the lens of two local network structures discussed in Section 6.2.1, friends' popularity and local clustering. As discussed in Section 2.3, the qualitative effects

Degree Quintile	1	2	3	4	5
	Clustering Coefficient				
Observed Networks	0.70	0.49	0.35	0.31	0.31
Random Networks	0.04	0.04	0.04	0.05	0.07
Δ	-0.65	-0.45	-0.30	-0.26	-0.23
	Average Friend Degree				
Observed Networks	35.74	48.52	58.84	68.67	98.80
Random Networks	89.91	84.40	83.23	84.05	91.14
Δ	54.17	35.88	24.40	15.39	-7.66
	Number of Referrals Received				
Observed Networks	0.16	0.18	0.21	0.25	0.33
Random Networks	0.06	0.10	0.16	0.27	0.52
Δ	-0.10	-0.08	-0.05	0.02	0.20
	Annual Consumption Equivalence (million $)				
Observed Networks	1.34	1.37	1.42	1.49	1.71
Random Networks	1.32	1.36	1.42	1.49	1.77
%Δ	-1.41%	-0.51%	0.04%	0.39%	3.52%
Group Size	719	860	898	854	861
mean(Degree)	7.32	14.23	24.26	40.94	105.89

Table 12: Counterfactual Experiment: Random Networks

Notes: (1). All the changes are calculated with respect to the observed networks. (2). The reported labor market outcomes are averages over 100 simulations. (3). The time span used in the simulation is 9 years, the same as the sample length. (4). Degree, average friend degree, and clustering coefficient are defined in (2), (30), and (3) respectively. They are calculated based on the simplified network in the final sample period, $Y^{All,2015}$, for the observed and the random network respectively.

of these network structures on job referrals are ambiguous. First, friends' popularity has two countervailing effects on the arrivals of referrals. On the one hand, a popular friend means high competition for referrals, lowering a worker's probability of receiving a particular referral sent by his/her popular friend (competition effect). On the other hand, a popular friend benefits from his/her large set of friends' referrals, increasing the quantity and the quality of referrals he/she sends out (ripple effect). Second, local clustering also has two countervailing effects on the arrivals of referrals. An advantage of high clustering is that it keeps the positive spillovers in an inner circle (closeness effect). A disadvantage is that it limits the positive spillovers from a distance (isolation effect).

I quantitatively analyze these network effects by regressing the percentage change in the number of referrals an individual receives on the change in his/her average friend degree and the change in his/her local clustering coefficient. I allow heterogeneous effects with respect to an individual's own degree by interacting the regressors with degree group dummies. Table 13 reports the results.

Dependent variable: $\Delta\%$ Number of Referrals			
(ΔAvg. Friend degree)	$\times\, \mathbb{1}(\text{degree} < 20)$	-0.0056***	(0.0008)
...	$\times\, \mathbb{1}(20 \leq \text{degree} < 40)$	-0.0081***	(0.0011)
...	$\times\, \mathbb{1}(40 \leq \text{degree} < 60)$	-0.0074***	(0.0016)
...	$\times\, \mathbb{1}(60 \leq \text{degree} < 80)$	-0.0085**	(0.0028)
...	$\times\, \mathbb{1}(80 \leq \text{degree} < 100)$	-0.0149**	(0.0047)
...	$\times\, \mathbb{1}(100 \leq \text{degree} < 150)$	-0.0095	(0.0056)
...	$\times\, \mathbb{1}(\text{degree} \geq 150)$	-0.0189*	(0.0074)
(ΔClustering coef.)	$\times\, \mathbb{1}(\text{degree} < 20)$	-0.0009	(0.1127)
...	$\times\, \mathbb{1}(20 \leq \text{degree} < 40)$	-0.6455***	(0.1783)
...	$\times\, \mathbb{1}(40 \leq \text{degree} < 60)$	-0.8782***	(0.2514)
...	$\times\, \mathbb{1}(60 \leq \text{degree} < 80)$	-0.4454	(0.4031)
...	$\times\, \mathbb{1}(80 \leq \text{degree} < 100)$	0.4937	(0.6450)
...	$\times\, \mathbb{1}(100 \leq \text{degree} < 150)$	1.5269	(0.9156)
...	$\times\, \mathbb{1}(\text{degree} \geq 150)$	4.3564***	(1.2196)
Observations		4,192	

Table 13: Effects of Friend Popularity and Clustering on Job Referral

Notes: (1). This table summarizes the results of an OLS regression, where the dependent variable is the percentage change in the number of referrals received. (2). Degree, average friend degree, and clustering coefficient are defined in (2), (30), and (3) respectively. They are calculated based on the simplified network in the final sample period, $Y^{All,2015}$, for the observed and the random network respectively. (3). Other covariates are degree, degree^2, degree^3. (4). Standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The results show that for all degree groups, the marginal impact of higher friend popularity is negative. This is driven by the competition effect from friends of friends. As the change from the observed networks to the random networks induces heterogeneous changes in friend popularity for different degree groups, the welfare implication will also be heterogeneous.

Moreover, the marginal impact of higher clustering is negative for low-degree groups, but positive for high-degree groups. The intuition for the heterogeneous effects of clustering is the following. When an individual's degree is low, the number of good shocks among his/her immediate friends is low, so it is relatively more important to be able to benefit from the spillovers from a distance. High clustering limits such chances, generating a negative impact on referrals (isolation effect dominates). When an individual's degree is high, however, the number of good shocks among his/her immediate friends is large enough, so it is relatively more important to keep these spillovers in an inner circle. High clustering provides such protections, generating a positive impact on referrals (closeness effect dominates). As the change from the observed networks to the random networks decreases the level of clustering universally, the heterogeneous effects of clustering for different degree groups generate heterogeneous welfare implication.

Qualitative Discussion. In summary, I present the change in two network statistics in Section 6.2.1, the marginal effects of these changes in Table 13, and the overall effects on referrals and thus welfare in Table 12. Many of these changes and effects are heterogeneous, so it is useful to provide a qualitative summary. Table 14 summarizes the qualitative effects of the changes in network structure on referrals through the lenses of the two local network structures discussed above.

	Low Degree	High Degree
Friend Popularity		
Δ Avg. Friend Degree	increase	decrease
Marginal Effect	negative	negative
Implied Change in Referrals	decrease	increase
Clustering		
Δ Clustering Coefficient	decrease	decrease
Marginal effect	negative	positive
Implied Change in Referrals	increase	decrease
Overall		
Change in Referrals	decrease	increase

Table 14: Qualitative Effects of a Change from the Observed Networks to the Random Networks

For low-degree individuals, their friend popularity increases, and a negative marginal effect implies a decrease in referrals. Their local clustering decreases, and a negative marginal effect implies an increase in referrals. Overall, the number of referrals–and thus welfare–decreases.

For high-degree individuals, their friend popularity decreases, and a negative marginal effect implies an increase in referrals. Their local clustering decreases, and a positive marginal effect implies a decrease in referrals. Overall, the number of referrals–and thus welfare–increases.

This counterfactual experiment highlights the importance of the network structure beyond the number of friends. Although the observed and the counterfactual networks have the same degree sequences, their difference in connection patterns leads to different welfare distribution. Specifically, with high-degree individuals gaining and low-degree individuals losing under the random networks, the distribution of referrals, and thus welfare, is more unequal. This counterintuitive finding further demonstrates the delicacy of the mechanisms in network models and the need for rigorous study.

7 Conclusion

In this paper, I study the network effects of job referrals and investigate the impact of social network structure on workers' labor market outcomes. I develop a job search model that incorporates referral and non-referral job offers and different kinds of social networks. In my model, the network structure is dynamic and evolves as workers move across jobs. Workers directly receive job offers as well as referrals from their friends. Referrals are generated endogenously: an external referral occurs when a friend rejects an offer that he/she receives, and an internal referral occurs when a friend leaves his/her current job. As a result of this referral process, the quantity and the quality of referrals depend not only on the worker's number of friends but also on the quality of these friends' jobs. Moreover, the model incorporates the full network structure beyond the number of friends, and it generates rich network effects beyond immediate friends.

My empirical analysis combines three different data sources on the corporate executive labor market. I first provide reduced-form evidence on job referrals in the executive labor market. I then estimate the structural job search model by Generalized Method of Moments (GMM). The estimation results show both the statistical and economical significance of job referrals. My model nests a model with no referrals, and a specification test rejects such a model. Simulations from my model show that more than one quarter of job transitions and raises are driven by referrals.

I use the estimated model to perform two counterfactual experiments. The first experiment evaluates the welfare effect of referrals by varying the probability of referrals. I find that shutting down referrals reduces executives' welfare by an equivalence of a 2-7% decrease in annual income and that mandatory referral boosts executives' welfare by an equivalence of a 6-16% increase in annual income.

The second experiment examines the welfare effect of network structure by varying the network structure. Specifically, a new set of counterfactual networks are generated in which the individuals have the same number of friends as the observed networks, but the connections are formed randomly. I find that the welfare distribution is more unequal under the randomly formed networks. I further investigate the mechanisms for these effects through the lens of two network statistics: friends' popularity and local clustering, and my findings are as follows. First, in terms of friends' popularity, the competition effect dominates the ripple effect. Second, in terms of local clustering, the isolation effect dominates for individuals with a small number of friends, and the closeness effect dominates for individuals with a large number of friends. Overall, the competition effect resulting from the change in friend popularity dominates, generating greater inequality under the random networks. This experiment highlights the effects of network structure beyond the number of friends.

This paper focuses on worker heterogeneity in network position. An interesting extension is to incorporate more heterogeneity such as gender into the analysis and study how referrals and social networks affect the gender pay gap and the underrepresentation of female executives. Additionally, this paper abstracts from the roles played by intermediaries such as executive search firms in soliciting referrals as well as the executives' strategic concerns in providing referrals. An interesting extension to the model is to explicitly incorporate the executive search companies and study their interactions with the executives in generating referrals. These are exciting areas for future research.

References

Addison, John T, and Pedro Portugal. 2002. "Job search methods and outcomes." *Oxford Economic Papers*, 54(3): 505–533.

Aizawa, Naoki, and Hanming Fang. 2018. "Equilibrium labor market search and health insurance reform." National Bureau of Economic Research.

Altonji, Joseph G, and Lewis M Segal. 1996. "Small-sample bias in GMM estimation of covariance structures." *Journal of Business & Economic Statistics*, 14(3): 353–366.

Arbex, Marcelo, Dennis O'Dea, and David Wiczer. 2018. "Network Search: Climbing the Job Ladder Faster." *International Economic Review*.

Arrow, Kenneth J, and Ron Borzekowski. 2004. "Limited network connections and the distribution of wages."

Bagger, Jesper, Fran Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics." *American Economic Review*, 104(6): 1551–96.

Banerjee, Abhijit V, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2017. "Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials."

Bebchuk, Lucian, and Jesse Fried. 2004. *Pay without performance.* Vol. 29, Cambridge, MA: Harvard University Press.

Blitzstein, Joseph, and Persi Diaconis. 2011. "A sequential importance sampling algorithm for generating random graphs with prescribed degrees." *Internet mathematics*, 6(4): 489–522.

Bontemps, Christian, Jean-Marc Robin, and Gerard J Van den Berg. 1999. "An empirical equilibrium job search model with search on the job and heterogeneous workers and firms." *International Economic Review*, 40(4): 1039–1074.

Bontemps, Christian, Jean-Marc Robin, and Gerard J Van den Berg. 2000. "Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation." *International Economic Review*, 41(2): 305–358.

Boorman, Scott A. 1975. "A combinatiorial optimization model for transmission of job information through contact networks." *The bell journal of economics*, 216–249.

Brown, Meta, Elizabeth Setren, and Giorgio Topa. 2016. "Do informal referrals lead to better matches? Evidence from a firms employee referral system." *Journal of Labor Economics*, 34(1): 161–209.

Burdett, Kenneth, and Dale T Mortensen. 1998. "Wage differentials, employer size, and unemployment." *International Economic Review*, 257–273.

Burks, Stephen V, Bo Cowgill, Mitchell Hoffman, and Michael Housman. 2015. "The value of hiring through employee referrals." *The Quarterly Journal of Economics*, 130(2): 805–839.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "Wage bargaining with on-the-job search: Theory and evidence." *Econometrica*, 74(2): 323–364.

Cai, Ye, and Merih Sevilir. 2012. "Board connections and M&A transactions." *Journal of Financial Economics*, 103(2): 327–349.

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

Calvó-Armengol, Antoni, and Matthew O Jackson. 2007. "Networks in labor markets: Wage and employment dynamics and inequality." *Journal of economic theory*, 132(1): 27–46.

Calvó-Armengol, Antoni, and Yves Zenou. 2005. "Job matching, social network and word-of-mouth communication." *Journal of urban economics*, 57(3): 500–522.

Conley, Timothy G, and Giorgio Topa. 2007. "Estimating dynamic local interactions models." *Journal of Econometrics*, 140(1): 282–303.

Corcoran, Mary, Linda Datcher, and Greg Duncan. 1980. "Information and influence networks in labor markets." *Five thousand American families: Patterns of economic progress*, 8(S 1): 37.

Dey, Matthew S, and Christopher J Flinn. 2005. "An equilibrium model of health insurance provision and wage determination." *Econometrica*, 73(2): 571–627.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2015. "Referral-based job search networks." *The Review of Economic Studies*, 83(2): 514–546.

Eckstein, Zvi, and Gerard J Van den Berg. 2007. "Empirical labor search: A survey." *Journal of Econometrics*, 136(2): 531–564.

Edmans, Alex, and Xavier Gabaix. 2016. "Executive compensation: A modern primer." *Journal of Economic literature*, 54(4): 1232–87.

El-Khatib, Rwan, Kathy Fogel, and Tomas Jandik. 2015. "CEO network centrality and merger performance." *Journal of Financial Economics*, 116(2): 349–382.

Engelberg, Joseph, Pengjie Gao, and Christopher A Parsons. 2013. "The Price of a CEO's Rolodex." *The Review of Financial Studies*, 26(1): 79–114.

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

Flinn, Christopher, and James Heckman. 1982. "New methods for analyzing structural models of labor force dynamics." *Journal of Econometrics*, 18(1): 115–168.

Fontaine, Francois. 2008. "Why are similar workers paid differently? The role of social networks." *Journal of Economic Dynamics and Control*, 32(12): 3960–3977.

Fracassi, Cesare, and Geoffrey Tate. 2012. "External networking and internal firm governance." *The Journal of Finance*, 67(1): 153–194.

Friedkin, Noah E. 1983. "Horizons of observability and limits of informal control in organizations." *Social Forces*, 62(1): 54–77.

Frydman, Carola, and Dirk Jenter. 2010. "CEO compensation." *Annu. Rev. Financ. Econ.*, 2(1): 75–102.

Gabaix, Xavier, and Augustin Landier. 2008. "Why has CEO pay increased so much?" *The Quarterly Journal of Economics*, 123(1): 49–100.

Galenianos, Manolis. 2014. "Hiring through referrals." *Journal of Economic Theory*, 152: 304–323.

Gayle, George-Levi, Chen Li, and Robert A Miller. 2016. "Was Sarbanes-Oxley Costly? Evidence from Optimal Contracting on CEO Compensation."

Gayle, George-Levi, Limor Golan, and Robert A Miller. 2015. "Promotion, turnover, and compensation in the executive labor market." *Econometrica*, 83(6): 2293–2369.

Granovetter, Mark. 1995. *Getting a job: A study of contacts and careers.* University of Chicago Press.

Granovetter, Mark S. 1973. "The strength of weak ties." *American journal of sociology*, 78(6): 1360–1380.

Gregg, Paul, and Jonathan Wadsworth. 1996. "How effective are state employment agencies? Jobcentre use and job matching in Britain." *Oxford Bulletin of Economics and Statistics*, 58(3): 443–467.

Holzer, Harry J. 1987. "Hiring procedures in the firm: Their economic determinants and outcomes."

Holzer, Harry J. 1988. "Search method use by unemployed youth." *Journal of labor economics*, 6(1): 1–20.

Huson, Mark R, Robert Parrino, and Laura T Starks. 2001. "Internal monitoring mechanisms and CEO turnover: A long-term perspective." *The Journal of Finance*, 56(6): 2265–2297.

Hwang, Byoung-Hyoun, and Seoyoung Kim. 2009. "It pays to have friends." *Journal of financial economics*, 93(1): 138–158.

Imbens, Guido W, and Tony Lancaster. 1994. "Combining micro and macro data in microeconometric models." *The Review of Economic Studies*, 61(4): 655–680.

Ioannides, Yannis M, and Adriaan R Soetevent. 2006. "Wages and employment in a random social network with arbitrary degree distribution." *American Economic Review*, 96(2): 270–274.

Ioannides, Yannis M, and Linda Datcher Loury. 2004. "Job information networks, neighborhood effects, and inequality." *Journal of economic literature*, 42(4): 1056–1093.

Jackson, Matthew O, Brian W Rogers, and Yves Zenou. 2017. "The economic consequences of social-network structure." *Journal of Economic Literature*, 55(1): 49–95.

Kaplan, Steven N, and Bernadette Minton. 2006. "How has CEO turnover changed? Increasingly performance sensitive boards and increasingly uneasy CEOs." National Bureau of Economic Research.

Krackhardt, David. 1987. "Cognitive social structures." *Social networks*, 9(2): 109–134.

Krackhardt, David. 2014. "A preliminary look at accuracy in egonets." In *Contemporary Perspectives on Organizational Social Networks*. 277–293. Emerald Group Publishing Limited.

Kuhnen, Camelia M, and Jeffrey Zwiebel. 2008. "Executive pay, hidden compensation and managerial entrenchment."

Larcker, David F, Eric C So, and Charles CY Wang. 2013. "Boardroom centrality and firm performance." *Journal of Accounting and Economics*, 55(2-3): 225–250.

Maddala, Gangadharrao Soundalyarao. 1971. "The use of variance components models in pooling cross section and time series data." *Econometrica: Journal of the Econometric Society*, 341–358.

Marsden, Peter V, and Elizabeth H Gorman. 2001. "Interpersonal ties, social capital, and employer staffing practices." *Social capital: Theory and research*, 105–26.

Meghir, Costas, Renata Narita, and Jean-Marc Robin. 2015. "Wages and informality in developing countries." *American Economic Review*, 105(4): 1509–46.

Miller, Shazia Rafiullah, and James E Rosenbaum. 1997. "Hiring in a Hobbesian world: Social infrastructure and employers' use of information." *Work and Occupations*, 24(4): 498–523.

Montgomery, James D. 1991. "Social networks and labor-market outcomes: Toward an economic analysis." *The American economic review*, 81(5): 1408–1418.

Montgomery, James D. 1992. "Job search and network composition: Implications of the strength-of-weak-ties hypothesis." *American Sociological Review*, 586–596.

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

Neckerman, Kathryn M, and Joleen Kirschenman. 1991. "Hiring strategies, racial bias, and inner-city workers." *Social Problems*, 38(4): 433–447.

Pellizzari, Michele. 2010. "Do friends and relatives really help in getting a good job?" *ILR Review*, 63(3): 494–510.

Petrin, Amil. 2002. "Quantifying the benefits of new products: The case of the minivan." *Journal of political Economy*, 110(4): 705–729.

Postel-Vinay, Fabien, and Jean-Marc Robin. 2002. "Equilibrium wage dispersion with worker and employer heterogeneity." *Econometrica*, 70(6): 2295–2350.

Rees, Albert. 1966. "Information networks in labor markets." *The American Economic Review*, 56(1/2): 559–566.

Rogerson, Richard, Robert Shimer, and Randall Wright. 2005. "Search-theoretic models of the labor market: A survey." *Journal of economic literature*, 43(4): 959–988.

Ruan, Hongxun. 2017. "Social Capital and Innovation: Evidence from Connected Holdings."

Rubinstein, Ariel. 1982. "Perfect equilibrium in a bargaining model." *Econometrica: Journal of the Econometric Society*, 97–109.

Schmutte, Ian M. 2014. "Job referral networks and the determination of earnings in local labor markets." *Journal of Labor Economics*, 33(1): 1–32.

Shephard, Andrew. 2017. "Equilibrium search and tax credit reform." *International Economic Review*, 58(4): 1047–1088.

Shue, Kelly. 2013. "Executive networks and firm policies: Evidence from the random assignment of MBA peers." *The Review of Financial Studies*, 26(6): 1401–1442.

Steger, Angelika, and Nicholas C Wormald. 1999. "Generating random regular graphs quickly." *Combinatorics, Probability and Computing*, 8(4): 377–396.

Taylor, Lucian A. 2010. "Why are CEOs rarely fired? Evidence from structural estimation." *The Journal of Finance*, 65(6): 2051–2087.

Taylor, Lucian A. 2013. "CEO wage dynamics: Estimates from a learning model." *Journal of Financial Economics*, 108(1): 79–98.

Terviö, Marko. 2008. "The difference that CEOs make: An assignment model approach." *The American Economic Review*, 642–668.

Topa, Giorgio. 2001. "Social interactions, local spillovers and unemployment." *The Review of Economic Studies*, 68(2): 261–295.

Topa, Giorgio. 2011. "Labor markets and referrals." In *Handbook of social economics*. Vol. 1, 1193–1221. Elsevier.

Van den Berg, Gerard J, and Geert Ridder. 1998. "An empirical equilibrium search model of the labor market." *Econometrica*, 1183–1221.

Appendix A Bargained Wage

Unemployed worker's value function is

$$\begin{aligned}\rho V_0(a,\Gamma) =& U(ab)\\ &+ \lambda_0 \int_{p_0(a,\Gamma)}^{p^{max}} [V_1(a,\phi_0(a,x,\Gamma),x,\Gamma) - V_0(a,\Gamma)]dF(x)\\ &+ \delta\pi_0 \int_{p_0(a,\Gamma)}^{p^{max}} [V_1(a,\phi_0(a,y,\Gamma),y,\Gamma) - V_0(a,\Gamma)]d\Gamma_0(y)\\ &+ \lambda_1\pi_1 \int_{p_0(a,\Gamma)}^{p^{max}} \int_{p_0(a,\Gamma)}^{y} [V_1(a,\phi_0(a,x,\Gamma),x,\Gamma) - V_0(a,\Gamma)]dF(x)d\Gamma_0(y)\\ &+ \lambda_1\pi_1 \int_{p_0(a,\Gamma)}^{p^{max}} \int_{y}^{p^{max}} [V_1(a,\phi_0(a,y,\Gamma),y,\Gamma) - V_0(a,\Gamma)]dF(x)d\Gamma_0(y),\end{aligned} \tag{31}$$

where state variable $\Gamma = \{\Gamma_0, \Gamma_1\}$, and $p_0(a,\Gamma)$ is the reservation productivity, i.e.

$$V_0(a,\Gamma) = V_1(a, ap_0(a,\Gamma), p_0(a,\Gamma), \Gamma). \tag{32}$$

Simplify the expression using the property of ϕ_0 implied from the bargaining process,

$$\begin{aligned}\rho V_0(a,\Gamma) =& U(ab)\\ &+ \lambda_0 \int_{p_0(a,\Gamma)}^{p^{max}} \beta[V_1(a,ax,x,\Gamma) - V_0(a,\Gamma)]dF(x)\\ &+ \delta\pi_0 \int_{p_0(a,\Gamma)}^{p^{max}} \beta[V_1(a,ay,y,\Gamma) - V_0(a,\Gamma)]d\Gamma_0(y)\\ &+ \lambda_1\pi_1 \int_{p_0(a,\Gamma)}^{p^{max}} \int_{p_0(a,\Gamma)}^{y} \beta[V_1(a,ax,x,\Gamma) - V_0(a,\Gamma)]dF(x)d\Gamma_0(y)\\ &+ \lambda_1\pi_1 \int_{p_0(a,\Gamma)}^{p^{max}} \int_{y}^{p^{max}} \beta[V_1(a,ay,y,\Gamma) - V_0(a,\Gamma)]dF(x)d\Gamma_0(y).\end{aligned} \tag{33}$$

Employed worker's value function is

$$
\begin{aligned}
&\rho V_1(a,w,p,\Gamma)\\
&= U(w) + \delta[V_0(a,\Gamma) - V_1(a,w,p,\Gamma)]\\
&+ \lambda_1 \int_{q(a,w,p,\Gamma)}^{p} [V_1(a,\phi_1(a,x,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)\\
&+ \lambda_1 \int_{p}^{p^{max}} [V_1(a,\phi_1(a,p,x,\Gamma),x,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)\\
&+ \delta\pi_0 \int_{q(a,w,p,\Gamma)}^{p} [V_1(a,\phi_1(a,y,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]d\Gamma_1(y)\\
&+ \delta\pi_0 \int_{p}^{p^{max}} [V_1(a,\phi_1(a,p,y,\Gamma),y,\Gamma) - V_1(a,w,p,\Gamma)]d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{q(a,w,p,\Gamma)}^{p} \int_{q(a,w,p,\Gamma)}^{y} [V_1(a,\phi_1(a,x,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{q(a,w,p,\Gamma)}^{p} \int_{y}^{p^{max}} [V_1(a,\phi_1(a,y,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{q(a,w,p,\Gamma)}^{p} [V_1(a,\phi_1(a,x,p,\Gamma),p,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{p}^{y} [V_1(a,\phi_1(a,p,x,\Gamma),x,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{y}^{p^{max}} [V_1(a,\phi_1(a,p,y,\Gamma),y,\Gamma) - V_1(a,w,p,\Gamma)]dF(x)d\Gamma_1(y),
\end{aligned}
\tag{34}
$$

where $q(a,w,p,\Gamma)$ is the cutoff productivity for wage raise, i.e.,

$$\phi_1(a,q(a,w,p,\Gamma),p,\Gamma) = w, \tag{35}$$

or equivalently,

$$V_1(a,w,p,\Gamma) = (1-\beta)V_1(a,aq(a,w,p,\Gamma),q(a,w,p,\Gamma),\Gamma) + \beta V_1(a,ap,p,\Gamma). \tag{36}$$

Simplify the expression using the property of ϕ_1 implied from the bargaining process,

$$
\begin{aligned}
&\rho V_1(a,w,p,\Gamma)\\
&= U(w) + \delta[V_0(a,\Gamma) - V_1(a,w,p,\Gamma)]\\
&+ \lambda_1 \int_{q(a,w,p,\Gamma)}^{p} \Big\{[\beta V_1(a,ap,p,\Gamma) + (1-\beta)V_1(a,ax,x,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} dF(x)\\
&+ \lambda_1 \int_{p}^{p^{max}} \Big\{[\beta V_1(a,ax,x,\Gamma) + (1-\beta)V_1(a,ap,p,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} dF(x)\\
&+ \delta\pi_0 \int_{q(a,w,p,\Gamma)}^{p} \Big\{[\beta V_1(a,ap,p,\Gamma) + (1-\beta)V_1(a,ay,y,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} d\Gamma_1(y)\\
&+ \delta\pi_0 \int_{p}^{p^{max}} \Big\{[\beta V_1(a,ay,y,\Gamma) + (1-\beta)V_1(a,ap,p,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{q(a,w,p,\Gamma)}^{p} \int_{q(a,w,p,\Gamma)}^{y} \Big\{[\beta V_1(a,ap,p,\Gamma) + (1-\beta)V_1(a,ax,x,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{q(a,w,p,\Gamma)}^{p} \int_{y}^{p^{max}} \Big\{[\beta V_1(a,ap,p,\Gamma) + (1-\beta)V_1(a,ay,y,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{q(a,w,p,\Gamma)}^{p} \Big\{[\beta V_1(a,ap,p,\Gamma) + (1-\beta)V_1(a,ax,x,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{p}^{y} \Big\{[\beta V_1(a,ax,x,\Gamma) + (1-\beta)V_1(a,ap,p,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} dF(x)d\Gamma_1(y)\\
&+ \lambda_1\pi_1 \int_{p}^{p^{max}} \int_{y}^{p^{max}} \Big\{[\beta V_1(a,ay,y,\Gamma) + (1-\beta)V_1(a,ap,p,\Gamma)] - V_1(a,w,p,\Gamma)\Big\} dF(x)d\Gamma_1(y).
\end{aligned}
\tag{37}
$$

First, define $\bar{F}(\cdot) = 1 - F(\cdot), \bar{\Gamma}_0(\cdot) = \Gamma_0(p^{\max}) - \Gamma_0(\cdot), \bar{\Gamma}_1(\cdot) = \Gamma_1(p^{\max}) - \Gamma_1(\cdot)$.

Simplify (37) by integrating by part, using property (36) and collecting terms of $V_1(a,w,p,\Gamma)$,

$$
\begin{aligned}
(\rho+\delta)V_1(a,w,p,\Gamma) =& U(w)+\delta V_0(a,\Gamma) \\
&+\lambda_1(1-\beta)\int_{q(a,w,p,\Gamma)}^{p}\bar{F}(x)\frac{dV_1(a,ax,x,\Gamma)}{dx}dx \\
&+\lambda_1\beta\int_{p}^{p^{max}}\bar{F}(x)\frac{dV_1(a,ax,x,\Gamma)}{dx}dx \\
&+\delta\pi_0(1-\beta)\int_{q(a,w,p,\Gamma)}^{p}\bar{\Gamma}_1(y)\frac{dV_1(a,ay,y,\Gamma)}{dy}dy \\
&+\delta\pi_0\beta\int_{p}^{p^{max}}\bar{\Gamma}_1(y)\frac{dV_1(a,ay,y,\Gamma)}{dy}dy \\
&+\lambda_1\pi_1(1-\beta)\int_{q(a,w,p,\Gamma)}^{p}\int_{q(a,w,p,\Gamma)}^{y}\bar{F}(x)\frac{dV_1(a,ax,x,\Gamma)}{dx}dxd\Gamma_1(y) \\
&+\lambda_1\pi_1(1-\beta)\int_{p}^{p^{max}}\int_{q(a,w,p,\Gamma)}^{p}\bar{F}(x)\frac{dV_1(a,ax,x,\Gamma)}{dx}dxd\Gamma_1(y) \\
&+\lambda_1\pi_1\beta\int_{p}^{p^{max}}\int_{p}^{y}\bar{F}(x)\frac{dV_1(a,ax,x,\Gamma)}{dx}dxd\Gamma_1(y).
\end{aligned}
\tag{38}
$$

For the special case where $w = ap$, $q(a,ap,p,\Gamma) = p$, and (38) becomes

$$
\begin{aligned}
(\rho+\delta)V_1(a,ap,p,\Gamma) =& U(ap)+\delta V_0(a,\Gamma) \\
&+\lambda_1\beta\int_{p}^{p^{max}}\bar{F}(x)\frac{dV_1(a,ax,x,\Gamma)}{dx}dx \\
&+\delta\pi_0\beta\int_{p}^{p^{max}}\bar{\Gamma}_1(y)\frac{dV_1(a,ay,y,\Gamma)}{dy}dy \\
&+\lambda_1\pi_1\beta\int_{p}^{p^{max}}\int_{p}^{y}\bar{F}(x)\frac{dV_1(a,ax,x,\Gamma)}{dx}dxd\Gamma_1(y).
\end{aligned}
\tag{39}
$$

Differentiate (39) with respect to p,

$$
\frac{dV_1(a,ap,p,\Gamma)}{dp} = \frac{aU'(ap)}{\rho+\delta+\lambda_1\beta\bar{F}(p)+[\lambda_1\pi_1\beta\bar{F}(p)+\delta\pi_0\beta]\bar{\Gamma}_1(p)}. \tag{40}
$$

To solve for bargained wage, note that $\phi_1(a,p^L,p^H,\Gamma)$ is equal to ϕ_1^* that solves

$$
V_1(a,\phi_1^*,p^H,\Gamma) = \beta V_1(a,ap^H,p^H,\Gamma) + (1-\beta)V_1(a,ap^L,p^L,\Gamma). \tag{41}
$$

Plug in expressions (38) and (39) and use the property that $q(a, \phi_1^*, p^H, \Gamma) = p^L$,

$$
\begin{aligned}
U(\phi_1(a, p^L, p^H, \Gamma)) =& \beta U(ap^H) + (1-\beta)U(ap^L) \\
& - \delta\pi_0(1-\beta)^2 \int_{p^L}^{p^H} \bar{\Gamma}_1(y) \frac{dV_1(a, ay, y, \Gamma)}{dy} dy \\
& - \lambda_1(1-\beta)^2 \Bigg\{ \int_{p^L}^{p^H} \bar{F}(x) \frac{dV_1(a, ax, x, \Gamma)}{dx} dx \\
& + \pi_1 \int_{p^H}^{p^{max}} \int_{p^L}^{p^H} \bar{F}(x) \frac{dV_1(a, ax, x, \Gamma)}{dx} dx d\Gamma_1(y) \\
& + \pi_1 \int_{p^L}^{p^H} \int_{p^L}^{y} \bar{F}(x) \frac{dV_1(a, ax, x, \Gamma)}{dx} dx d\Gamma_1(y) \Bigg\},
\end{aligned} \tag{42}
$$

where $\frac{dV_1(a,ax,x,\Gamma)}{dx}$ is given in (40).

Simplify (33) by integrating by part,

$$
\begin{aligned}
\rho V_0(a, \Gamma) =& U(ab) \\
& + \lambda_0\beta \int_{p_0(a,\Gamma)}^{p^{max}} \bar{F}(x) \frac{dV_1(a, ax, x, \Gamma)}{dx} dx + \delta\pi_0\beta \int_{p_0(a,\Gamma)}^{p^{max}} \bar{\Gamma}_0(y) \frac{dV_1(a, ay, y, \Gamma)}{dy} dy \\
& + \lambda_1\pi_1\beta \int_{p_0(a,\Gamma)}^{p^{max}} \int_{p_0(a,\Gamma)}^{y} \bar{F}(x) \frac{dV_1(a, ax, x, \Gamma)}{dx} dx d\Gamma_0(y).
\end{aligned} \tag{43}
$$

To solve for unemployed worker's reservation productivity, note that $p_0(a, \Gamma)$ is equal to p_0^* that solves

$$
V_1(a, ap_0^*, p_0^*, \Gamma) = V_0(a, \Gamma). \tag{44}
$$

Plugging in expressions (43) and (39) gives an implicit function for $p_0(a, \Gamma)$

$$
\begin{aligned}
U(ap_0(a, \Gamma)) =& U(ab) + \beta(\lambda_0 - \lambda_1) \int_{p_0(a,\Gamma)}^{p^{max}} \bar{F}(x) \frac{dV_1(a, ax, x, \Gamma)}{dx} dx \\
& + \delta\pi_0\beta \int_{p_0(a,\Gamma)}^{p^{max}} (\bar{\Gamma}_0(y) - \bar{\Gamma}_1(y)) \frac{dV_1(a, ay, y, \Gamma)}{dy} dy \\
& + \lambda_1\pi_1\beta \int_{p_0(a,\Gamma)}^{p^{max}} (\bar{\Gamma}_0(x) - \bar{\Gamma}_1(x)) \bar{F}(x) \frac{dV_1(a, ay, y, \Gamma)}{dx} dx,
\end{aligned} \tag{45}
$$

where $\frac{dV_1(a,ax,x,\Gamma)}{dx}$ is given in (40).

By the definition of $p_0(a, \Gamma)$,

$$
\phi_0(a, p, \Gamma) = \phi_1(a, p_0(a, \Gamma), p, \Gamma). \tag{46}
$$

In summary, the bargained wage is characterized by (40), (42), (45), and (46).

For CRRA utility with $\alpha \neq 1$, (42) simplifies to

$$\begin{aligned}\ln \phi_1(a, p^L, p^H, \Gamma) = \ln a + \frac{1}{1-\alpha} \ln \Bigg\{ & \beta (P^H)^{1-\alpha} + (1-\beta)(P^L)^{1-\alpha} \\ & - (1-\alpha)\delta\pi_0(1-\beta)^2 \int_{p^L}^{p^H} \frac{\bar{\Gamma}_1(y) y^{-\alpha}}{\rho + \delta + \lambda_1 \beta \bar{F}(y) + [\lambda_1 \pi_1 \beta \bar{F}(y) + \delta\pi_0 \beta] \bar{\Gamma}_1(y)} dy \\ & - (1-\alpha)\lambda_1 (1-\beta)^2 \Big[\int_{p^L}^{p^H} \frac{\bar{F}(x) x^{-\alpha}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx \\ & + \pi_1 \int_{p^H}^{p^{max}} \int_{p^L}^{p^H} \frac{\bar{F}(x) x^{-\alpha}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx d\Gamma_1(y) \\ & + \pi_1 \int_{p^L}^{p^H} \int_{p^L}^{y} \frac{\bar{F}(x) x^{-\alpha}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx d\Gamma_1(y) \Big] \Bigg\},\end{aligned} \tag{47}$$

and (45) simplifies to

$$\begin{aligned}\ln p_0(a, \Gamma) = & \frac{1}{1-\alpha} \ln \Bigg\{ b^{1-\alpha} + (1-\alpha)\beta(\lambda_0 - \lambda_1) \int_{p_0(a,\Gamma)}^{p^{max}} \frac{\bar{F}(x) x^{-\alpha}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx \\ & + (1-\alpha)\delta\pi_0\beta \int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(y) - \bar{\Gamma}_1(y)) y^{-\alpha}}{\rho + \delta + \lambda_1 \beta \bar{F}(y) + [\lambda_1 \pi_1 \beta \bar{F}(y) + \delta\pi_0 \beta] \bar{\Gamma}_1(y)} dy \\ & + (1-\alpha)\lambda_1\pi_1\beta \int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(x) - \bar{\Gamma}_1(x)) \bar{F}(x) x^{-\alpha}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx \Bigg\};\end{aligned} \tag{48}$$

for CRRA utility with $\alpha = 1$, (42) simplifies to

$$\begin{aligned}\ln \phi_1(a, p^L, p^H, \Gamma) = \ln a + \beta \ln p^H + (1-\beta) \ln p^L & \\ - \delta\pi_0(1-\beta)^2 \int_{p^L}^{p^H} & \frac{\bar{\Gamma}_1(y) y^{-1}}{\rho + \delta + \lambda_1 \beta \bar{F}(y) + [\lambda_1 \pi_1 \beta \bar{F}(y) + \delta\pi_0 \beta] \bar{\Gamma}_1(y)} dy \\ - \lambda_1(1-\beta)^2 \Big[\int_{p^L}^{p^H} & \frac{\bar{F}(x) x^{-1}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx \\ + \pi_1 \int_{p^H}^{p^{max}} \int_{p^L}^{p^H} & \frac{\bar{F}(x) x^{-1}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx d\Gamma_1(y) \\ + \pi_1 \int_{p^L}^{p^H} \int_{p^L}^{y} & \frac{\bar{F}(x) x^{-1}}{\rho + \delta + \lambda_1 \beta \bar{F}(x) + [\lambda_1 \pi_1 \beta \bar{F}(x) + \delta\pi_0 \beta] \bar{\Gamma}_1(x)} dx d\Gamma_1(y) \Big],\end{aligned} \tag{49}$$

and (45) simplifies to

$$\begin{aligned}\ln p_0(a,\Gamma) = &\ln b + \beta(\lambda_0-\lambda_1)\int_{p_0(a,\Gamma)}^{p^{max}} \frac{\bar{F}(x)x^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx \\ &+ \delta\pi_0\beta\int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(y)-\bar{\Gamma}_1(y))y^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(y)+[\lambda_1\pi_1\beta\bar{F}(y)+\delta\pi_0\beta]\bar{\Gamma}_1(y)}dy \\ &+ \lambda_1\pi_1\beta\int_{p_0(a,\Gamma)}^{p^{max}} \frac{(\bar{\Gamma}_0(x)-\bar{\Gamma}_1(x))\bar{F}(x)x^{-1}}{\rho+\delta+\lambda_1\beta\bar{F}(x)+[\lambda_1\pi_1\beta\bar{F}(x)+\delta\pi_0\beta]\bar{\Gamma}_1(x)}dx.\end{aligned} \tag{50}$$

Appendix B Additional Table for Data Description

	Career Advancement		UE Transition		EE Transition		EU Transition	
#(Transition)	#(Individual)	%	#(Individual)	%	#(Individual)	%	#(Individual)	%
0	1,772	42.27	1,836	43.80	4,066	96.99	1,941	46.30
1	2,178	51.96	2171	51.79	125	2.98	2,085	49.73
2	231	5.51	179	4.27	1	0.02	161	3.84
3	10	0.24	6	0.14	0	0	5	0.12
4	1	0.02	0	0	0	0	0	0
Total	4,192	100.00	4,192	100.00	4,192	100.00	4,192	100.00

Table B1: Tabulation of Individuals' Numbers of Transitions

Notes: (1). A UE transition is defined as the transition from non-executive to executive. (2). An EE transition is defined as the transition from one executive job to another. (3). An EU transition is defined as the transition from executive to non-executive. (4). A career advancement is defined as either a UE or EE transition.

	Dependent variable: $\mathbb{1}$*(Career Advancement)*$\in \{0,1\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}$ (Job Transition Among Executive Friends)	0.0470***			
	(0.0055)			
... from School		0.0166***		
		(0.0049)		
... from Work			0.0532***	
			(0.0050)	
... from Social Activity				0.0174**
				(0.0059)
Observations	33,536	33,536	33,536	33,536
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Table B2: Co-movement in Socially Connected Executives' Job Transitions

Notes: (1). This table summarizes the results of an OLS regression with year and individual fixed effects, where the dependent variable is whether an individual experiences a career advancement. (2). A career advancement is defined as either a transition from non-executive job to executive job (UE) or a transition between executive jobs with productivity increase (EE). (3). Variable $\mathbb{1}$ (Job Transition Among Executive Friends) is a dummy variable that equals one if any friend experiences a transition between executive jobs with productivity increase (EE) or a transition from executive job to non-executive job (EU) in the same year or the previous year. (4). The scope of friends varies for different network specifications. Column (1) uses the simplified network, which does not distinguish between types of social connections; column (2) uses the school network; column (3) uses the work network; and column (4) uses the social-activity network. (5). Other covariates include age^2, age^3, whether an individual was an executive in the previous year, and the productivity of the job in the previous year if the individual was an executive. (6). Standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix C Estimation Strategy

C.1 Likelihood L_1

The likelihood contribution of a worker i with m career advancements at time $t_1, t_2, ...t_m$ is

$$f_i(\boldsymbol{w_i};\boldsymbol{\theta}) = \frac{\exp\Big(-\frac{1}{2}(\ln\boldsymbol{w_i} - \ln\boldsymbol{\phi_i} - \boldsymbol{\mu_a})^T\boldsymbol{\Sigma_i}^{-1}(\ln\boldsymbol{w_i} - \ln\boldsymbol{\phi_i} - \boldsymbol{\mu_a})\Big)}{\sqrt{(2\pi)^m|\boldsymbol{\Sigma_i}|}}, \tag{51}$$

where

$$\ln\boldsymbol{w_i} = \begin{bmatrix} \ln w_{it_1} \\ ... \\ \ln w_{it_m} \end{bmatrix}_{m\times 1}, \quad \ln\boldsymbol{\phi_i} = \begin{bmatrix} \phi_1(1, p^L_{it_1}, p^H_{it_1}, \Gamma_{it_1}) \\ ... \\ \phi_1(1, p^L_{it_m}, p^H_{it_m}, \Gamma_{it_m}) \end{bmatrix}_{m\times 1}, \quad \boldsymbol{\mu_a} = \begin{bmatrix} \mu_a \\ ... \\ \mu_a \end{bmatrix}_{m\times 1}, \tag{52}$$

and

$$\boldsymbol{\Sigma_i} = \begin{bmatrix} \sigma_a^2 + \sigma_\epsilon^2 & ... & \sigma_a^2 \\ ... & ... & ... \\ \sigma_a^2 & ... & \sigma_a^2 + \sigma_\epsilon^2 \end{bmatrix}_{m\times m}. \tag{53}$$

For all the UE transitions, the model implies an additional constraint on the parameter space: $p^H_{it} \geq p_{0,it}(\boldsymbol{\theta})$, the observed productivity of the accepted job is no less than the calculated reservation productivity. I incorporate this constraint as a penalty term in the objective function. Instead of directly solving the constrained optimization problem

$$\begin{aligned} &\max \sum_i \ln f_i(\boldsymbol{w_i};\boldsymbol{\theta}) \\ &\text{s.t. } p^H_{it} \geq p_{0,it}(\boldsymbol{\theta}) \text{ for UE transition,} \end{aligned} \tag{54}$$

set the following objective function to solve an unconstrained optimization problem

$$\max \sum_i \ln f_i(\boldsymbol{w_i};\boldsymbol{\theta}) - \kappa\sum_i g_i(\boldsymbol{\theta}), \tag{55}$$

where κ is a small positive number, and $g_i(\boldsymbol{\theta}) = \sum_{t\in\{t_1,...,t_m\}} g_{it}(\boldsymbol{\theta})$ is a penalty function defined as

$$g_{it}(\boldsymbol{\theta}) = \begin{cases} 0 & \text{for EE transition,} \\ -\ln(p^H_{it} - p_{0,it}(\boldsymbol{\theta})) & \text{for UE transition where } p^H_{it} > p_{0,it}(\boldsymbol{\theta}), \\ +\infty & \text{for UE transition where } p^H_{it} \leq p_{0,it}(\boldsymbol{\theta}). \end{cases} \tag{56}$$

Function g is an interior penalty function, which guarantees that the solution to (55) satisfies the constraint. When κ is small enough, the solution to (55) will be close to the solution to original constrained problem (54).[45]

C.2 Likelihood L_2

The likelihood contribution of the productivity of a socially isolated worker i's first job p_i is

$$f_i(p_i;\boldsymbol{\theta}) = \frac{\phi(\frac{\ln p_i - \mu_p}{\sigma_p})}{\sigma_p(\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{\ln \tilde{p} - \mu_p}{\sigma_p}))} \cdot \frac{1}{p_i}, \tag{57}$$

where ϕ and Φ are the pdf and cdf of the standard normal distribution, $\tilde{p} = \max\{p^{\min}, p_0\}$, and p_0 is the reservation productivity for an isolated worker satisfying

$$\ln p_0 = \begin{cases} \ln b + \frac{\exp(\beta_0)}{1+\exp(\beta_0)}(\lambda_0 - \lambda_1)\displaystyle\int_{p_0}^{p^{max}} \frac{x^{-1} \cdot \frac{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{x - \mu_p}{\sigma_p})}{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{\ln p^{\min} - \mu_p}{\sigma_p})}}{\rho + \delta + \lambda_1 \frac{\exp(\beta_0)}{1+\exp(\beta_0)} \cdot \frac{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{x - \mu_p}{\sigma_p})}{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{\ln p^{\min} - \mu_p}{\sigma_p})}} dx & \text{if } \alpha = 1, \\ \\ \frac{1}{1-\alpha} \ln \left\{ b^{1-\alpha} + (1-\alpha)\frac{\exp(\beta_0)}{1+\exp(\beta_0)}(\lambda_0 - \lambda_1)\displaystyle\int_{p_0}^{p^{max}} \frac{x^{-\alpha} \cdot \frac{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{x - \mu_p}{\sigma_p})}{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{\ln p^{\min} - \mu_p}{\sigma_p})}}{\rho + \delta + \lambda_1 \frac{\exp(\beta_0)}{1+\exp(\beta_0)} \cdot \frac{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{x - \mu_p}{\sigma_p})}{\Phi(\frac{\ln p^{\max} - \mu_p}{\sigma_p}) - \Phi(\frac{\ln p^{\min} - \mu_p}{\sigma_p})}} dx \right\} & \text{if } \alpha \neq 1. \end{cases} \tag{58}$$

[45] I use $\kappa = 10^{-6}$ in the estimation.

Appendix D Additional Table for Model Fit

Moments	Data	Model
Co-movement in School Network		
Fraction of unemployed individuals experiencing UE transition		
… if there is no school friend experiencing EU transition	0.1601	0.1612
… if there are school friends experiencing EU transition	0.1631	0.1767
… if there is no school friend experiencing EE transition	0.1603	0.1636
… if there are school friends experiencing EE transition	0.1784	0.1922
Fraction of employed individuals experiencing EE transition		
… if there is no school friend experiencing EU transition	0.0065	0.0063
… if there are school friends experiencing EU transition	0.0086	0.0091
… if there is no school friend experiencing EE transition	0.0071	0.0068
… if there are school friends experiencing EE transition	0.0079	0.0128
Co-movement in Work Network		
Fraction of unemployed individuals experiencing UE transition		
… if there is no work friend experiencing EU transition	0.0805	0.1381
… if there are work friends experiencing EU transition	0.1766	0.1784
… if there is no work friend experiencing EE transition	0.1547	0.1615
… if there are work friends experiencing EE transition	0.1934	0.1920
Fraction of employed individuals experiencing EE transition		
… if there is no work friend experiencing EU transition	0.0026	0.0041
… if there are work friends experiencing EU transition	0.0080	0.0079
… if there is no work friend experiencing EE transition	0.0055	0.0056
… if there are work friends experiencing EE transition	0.0156	0.0175
Co-movement in Social-Activity Network		
Fraction of unemployed individuals experiencing UE transition		
… if there is no social-activity friend experiencing EU transition	0.1624	0.1600
… if there are social-activity friends experiencing EU transition	0.1547	0.1871
… if there is no social-activity friend experiencing EE transition	0.1607	0.1629
… if there are social-activity friends experiencing EE transition	0.1667	0.2051
Fraction of employed individuals experiencing EE transition		
… if there is no social-activity friend experiencing EU transition	0.0072	0.0062
… if there are social-activity friends experiencing EU transition	0.0071	0.0096
… if there is no social-activity friend experiencing EE transition	0.0071	0.0067
… if there are social-activity friends experiencing EE transition	0.0076	0.0100

Table D3: Model Fit: Moments in Job Transitions

Notes: A UE transition is defined as the transition from non-executive to executive. An EE transition is defined as the transition from one executive job to another with a productivity increase. An EU transition is defined as the transition from executive to non-executive.

Appendix E Simulating Random Network with A Given Degree Sequence

To simulate a random network with a given degree sequence, I follow the Steger and Wormald (1999) algorithm.[46] The algorithm starts with an empty graph and adds edges sequentially. Each time, generate a potential edge by sampling uniformly a pair of nodes that have not yet received their full allotment of edges. Add this edge if it does not induce a loop or multiple edges. Continue this process until no more permissible pairs can be found. If the algorithm gets stuck in the sense that there are unmatched nodes left over that are not allowed to be paired with each other, start over and try again.

To simulate a series of monotonically growing (over time) networks, first generate the initial network exactly as described above. For the subsequent networks, to simulate Y^t, start the algorithm with Y^{t-1} instead of an empty graph and add edges as described above.

[46] It is a variant of the pairing model (also known as the configuration model or stubs model). See Blitzstein and Diaconis (2011) for a survey of other algorithms.