History Dependence, Cohort Attachment, and Job Referrals in Networks of Close Relationships

Ayal Chen-Zion, James E. Rauch

History Dependence, Cohort Attachment, and Job Referrals in Networks of Close Relationships

Abstract

We model network formation in a firm. Agents learn about the quality of their working relationships with each other. Their good relationships become their networks. Accumulating relationships becomes increasingly costly, however. Over time agents become less open to forming relationships with others unknown to them, leading their networks to be front-loaded with agents they met near the beginning of their careers. The interaction of this dynamic with turnover yields predictions about the time pattern of history dependence in an agent's network as a function of his tenure. Mutual openness of newly arrived agents in a firm also leads to the cross-section prediction of "cohort attachment," a tendency for members of an agent's hiring cohort to be disproportionately represented in his network. When members of a network formed within a firm are subsequently split across many firms, the desire to renew their successful working relationships can lead to job referrals. Former co-workers who provide referrals will be drawn disproportionately from the referred workers' hiring cohorts at their previous employers.

Ayal Chen-Zion
Amazon / USA
chenzion@amazon.com

James E. Rauch
UC San Diego / USA
jrauch@ucsd.edu

1 Introduction

People form close social and work relationships inside organizations such as firms and schools. A distinguishing feature of close relationships is that they require significant time and energy. In this paper we argue that these relationships are also persistent. We show how this persistence of close relationships, in combination with their time-intensity, shapes the way agents form their networks of close relationships and the resulting patterns of history dependence in these networks. Specifically, 1) agents' networks tend to be front-loaded with co-workers they met near the beginning of their careers, 2) turnover within a firm erodes this front-loading and leads to a U-shaped pattern of history dependence for the networks of agents with careers of moderate length, and 3) members of an agent's hiring cohort tend to be disproportionately represented in his network. In an extension to multiple firms, we show that when members of a network formed within a firm are subsequently split across many firms, the desire to renew their successful working relationships can lead to job referrals, and former co-workers who provide referrals will be drawn disproportionately from the referred workers' hiring cohorts at their previous employers.

In our model agents learn whether they get along well or work productively together by trying to do so. If they discover they are well matched, they continue to socialize or work collaboratively in the future, given the opportunities. Denoting a focal agent by ego and designating the others in the organization as alters, we consider the set of alters with whom ego has learned he is well matched to constitute his network of close relationships. The size of ego's network, hence its value, increases monotonically with his tenure toward its steady-state value, consistent with the typical wage-tenure profile (see, e.g., Topel 1991).

Ego expands his network by trying out relationships with alters of unknown match quality and learning with which new alters match quality is good. Adding members to his network becomes increasingly costly, however, because close interaction with each one eventually interferes with close interaction with the others, given limited time and energy. Considering an ego entering a new environment, he will be most open to trying out relationships at the beginning, and less open later when his network is growing large. Agents' networks thus tend to be front-loaded with people they met near the beginning of their organizational careers. At the same time there is turnover within the organization, which tends to yield the highest shares in ego's network of alters he met most recently. We show that the interaction of these two forces leads to a U-shaped pattern of history dependence for the networks of agents with careers of moderate length. That is, a longitudinal plot

of the shares in ego's network of alters against the dates in his career at which he met them will have a U-shape for egos with careers of moderate length.

When turnover brings a new cohort of agents into the organization, they find that the agents already there are not very open to trying out new relationships, so the new agents try out relationships with each other. A pattern of network links (close relationships) forms within the organization in which within-cohort links are overrepresented. Our model thus gives rise to predictions about the cross-section pattern of network links within an organization as well as predictions regarding how individual networks evolve over time.

We will give the name "cohort attachment" to the tendency for within-cohort links to be overrepresented within an organization. This prediction of our model allows us to understand better several related ideas and results in the literature. Wagner, Pfeffer, and O'Reilly III (1984, p. 76) write, "Thus, because of the effects of free communication capacity and interest in forming relationships, persons who enter [the organization] at roughly the same time are more likely to communicate with each other than with those who entered either much earlier or later." This idea is used by Zenger and Lawrence (1989) to examine the impact of tenure similarity (equivalent to time-of-entry similarity) on subsequent communication. They find that tenure similarity strongly predicts the frequency with which engineers and engineering managers in the research division of a medium-sized U.S. electronics firm communicate outside of their project groups. Bandiera, Barankay, and Rasul (2008, Table 4) find that "same arrival date" is a strong predictor of friendship among college students working on seasonal contracts picking fruit on a UK farm, controlling for a wide range of ascriptive characteristics and potential correlates such as same living site. Chen-Zion (2016, Chapter 2) uses matched employer-employee data for Brazil to examine which co-workers an employee entrepreneur brings from a parent firm to his spinoff firm. He finds that parent firm employees hired in the same first plant and same cohort as the entrepreneur were 21 percent more likely to join him at the spinoff than other parent employees hired in the same first plant, controlling for similarity between co-worker and entrepreneur characteristics and for length of co-worker tenure overlap with the entrepreneur at the parent firm.

In our model the value generated by each pairwise match does not depend on which other matches form, and in the central case we analyze there are always enough agents of unknown match quality with whom to explore matching. Thus there are neither network externalities nor search and matching externalities, and egocentric networks form optimally from the social point of

view. Departures from optimality and corrective policies can be analyzed starting from this base. For example, network externalities could be present because the network of relationships in the organization as a whole serves the often-studied function of diffusion of information. The whole network may not be configured optimally to serve this function. Information could be bottled up within hiring cohorts, a version of the "siloing" problem (Tett 2015). In this situation, suppose we consider formation of special project work teams, the subject of a large literature (e.g., Chen 2005, Reagans, Zuckerman, and McEvily 2007). When choosing from employees with relatively low tenure, management should make sure that different hiring cohorts are represented on each team. This need not be a concern for employees with high tenure because turnover will have eroded their cohort network shares.

We can extend our model to allow members of ego's network formed within an organization who are subsequently split across many organizations to be his "contacts." The desire of contacts to renew their successful working relationships can lead to job referrals.[1] We show that this value of contacts causes agents to increase their target numbers of relationships, and provide sufficient conditions under which they form qualitatively the same longitudinal and cross-sectional patterns of network links within their organizations as in the model without contacts. Under these sufficient conditions our model predicts that former co-workers who provide referrals will be drawn disproportionately from the referred workers' hiring cohorts at their previous employers.

In much of the existing theoretical literature on job referrals, networks are given exogenously (e.g., Montogomery 1991, Calvo-Armengol and Jackson 2004). Galeotti and Merlino (2014) model investment in contacts, but unlike in our model investment in contacts is not joint with investment in intrafirm networks, because intrafirm networks are not covered. We do not contribute to explaining to what extent firms use referrals, which is the subject of Galenianos (2013, 2014). Our work is more closely related to the small but growing empirical literature that examines

---

[1]Note that workers do not have incentives to provide "negative referrals" if not forced to renew unsuccessful relationships when former co-workers with whom they are poorly matched are hired. This corresponds to the central case of our model in which agents of unknown match quality are always available to try out at the margin. The teamwork motivation for referral has recently been confirmed experimentally by Pallais and Sands (2016). In randomized controlled trials with referred job applicants, they report (p. 1796), "each referral completed one task with her referrer and one task with another randomly chosen referrer. Referred workers performed substantially better when paired with their own referrers." They also report (Online Appendix, p. 3), "referrers were more than twice as likely to want to partner again with their own referral as with someone else's referral. Similarly, referred workers were substantially more likely to want to work again with their own referrer than with someone else's referrer." On the other hand, Pallais and Sands (p. 1793) "do not find evidence that referrals exert more effort because they believe their performance will affect their relationship with their referrer or their referrer's position at the firm."

job referrals from former co-workers (Cingano and Rosolia 2012, Glitz 2017, Saygin, Weber, and Weynandt forthcoming, Eliason, Hensvik, Kramarz, and Skans 2019). These authors use administrative data from, respectively, Italy, Germany, Austria, and Sweden to identify worker overlap at previous employers. They examine outcomes for workers displaced by firm closure. A worker's former colleagues are linked to his reduced unemployment duration or greater re-employment probability (Cingano and Rosolia 2012, Glitz 2017), and his increased probability of being hired at a former colleague's plant (Saygin, Weber, and Weynandt forthcoming, Eliason, Hensvik, Kramarz, and Skans 2019). Using Brazilian data, Chen-Zion (2016, Chapter 1) extends the framework of Saygin, Weber, and Weynandt (forthcoming) and finds that the presence of a hiring-cohort former co-worker increases the probability of job acquisition at a specific hiring plant nearly three times more than the presence of a nonhiring-cohort former co-worker.

Our work is also related to, and has implications for, the peer effects literature. Both are concerned with networks formed as a result of being in the same place at the same time. The current state of the art in the peer effects literature is to examine peer groups created by random assignment (see Sacerdote 2014 for a survey). Typically random assignment occurs at the beginning of the agents' tenure in an organization. A popular example is random assignment of college freshmen to dorm rooms. The alters to which ego is randomly assigned are then found to influence a wide range of his behaviors, from binge drinking to buying a new car. The results of our model suggest that this influence would be much weaker if the random assignments occurred at the ends instead of the beginnings of organizational careers, because egos will be less open to establishing new relationships with the alters to whom they have been assigned. At the same time, persistence of close relationships suggests that it would be worth pursuing follow-up studies of the influence of randomly assigned peers.

In the next section we develop our model of network formation and history dependence for one firm, and derive results for the longitudinal and cross-sectional structure of agents' networks. Section 3 extends the model to incorporate contacts and allow for job referrals across firms. Our conclusions are in Section 4.

2 Network Formation Within One Firm

2.1 Model description and assumptions

We will consider the formation of personal networks by agents within an organization. It is convenient to call this organization a firm, but we believe that our model applies to network formation in other institutional settings as well.

A key inspiration for our model is Jovanovic (1979). In his model, one worker meets with one firm, and the pair learn about the quality of their match. Roughly speaking, if they learn that the quality of their match is good, they stay together, and if they learn that the quality of their match is bad, they separate. In our model, matches are between workers (agents) within a firm. Well matched agents become members of each others' networks (stay together), and poorly matched agents avoid each other in the future (separate). Different from Jovanovic (1979), an agent can in principle form matches with any number of other agents, up to the limit of all the agents in the firm.

We will follow the evolution of agents' networks in the firm over time $t = \{0, 1, 2, \dots\}$. These agents are risk-neutral and symmetric. In every period, they engage in pairwise work relationships or matches. Agents learn their match qualities with other agents by experience: at the end of every period, the qualities of all unknown matches formed in that period are revealed.

Assumption 1. *The firm is born at the beginning of period $t = 0$ with a continuum of agents with mass equal to N.*

The continuum assumption allows us to avoid integer problems. In this section, we will ignore agents outside the boundary of the firm.

The firm undergoes a constant, exogenous rate of worker turnover:

Assumption 2. *At the beginning of every period $t = \{1, 2, \dots\}$, a share δ of agents selected at random separates from the firm, and the departing agents are replaced by a* cohort *of mass δN, where $\delta \in [0, 1]$.*

This assumption serves two purposes. First, it creates network decay at a constant rate δ: an agent who remains with the firm finds that a share δ of the agents with whom he would be well matched disappears each period. Second, it creates a cohort structure for the firm in which every cohort except the founding ($t = 0$) cohort is of equal initial size, and cohort size declines at a constant

6

rate with tenure. When we analyze the cohort shares of agents' networks, the patterns generated by the interplay of forces in our model will stand out more clearly against this simple baseline. We could decouple hiring from network decay, allowing the firm to grow or shrink in any period, without affecting the longitudinal results in the next subsection, but we will note in subsection 2.3 that some of our cross-sectional results could be affected.

Remark. It is useful to consider what happens with extreme values of δ. If $\delta = 1$, the firm is re-created from scratch every period. Since the firm and its agents have no history, there can be no analysis of history dependence in the agents' networks. If $\delta = 0$, the firm consists of a fixed set of agents whose networks do not decay. This polar case is of some interest and will be covered in the next subsection.

Each agent maximizes the discounted sum of his per-period payoffs. The next three assumptions concern these payoffs.

Assumption 3. *Every match is one of two types determined by the surplus it yields to the matched parties in the period in which it occurs: high quality yielding y_H or low quality yielding y_L ($y_H > y_L > 0$).[2] The unconditional probability that a match is high quality is $p \in (0, 1)$.*

The match surplus can be thought of as net of any benefit derived by the firm. We assume that the firm, like its workers, prefers high to low quality matches. For example, the firm could receive a fixed share of the value generated by any match. The assumption that the firm prefers high to low quality matches is of no consequence in this section but is relevant in the next section when we discuss job referrals.

As stated above, a match is a pairwise work relationship between risk-neutral agents. We believe that work relationships are ubiquitous even where employees appear to work in isolation, as in a typical cubicle environment, for example. Employees find others with whom they work well solving non-routine problems or filling in for each other. They interact during breaks and lunch, where good relationships boost morale and reduce absenteeism. The yield y from a match should thus be interpreted flexibly. It could be money, such as a piece rate received by the agent pair for the output or service they produce, or a sum of money and utility, as when an enjoyable workplace relationship also raises productivity and thereby increases pay.

Assumption 4. *Every match is of equal value to both parties, i.e., the matched parties divide the surplus equally.*

---

[2]This follows the Moscarini (2005) simplification of Jovanovic (1979).

When the context is appropriate, it is possible to interpret this assumption as the outcome of Nash bargaining with a disagreement point of $(0, 0)$. For example, we could suppose that if the matched parties cannot agree on who deserves how much credit, they cannot turn in their project to their boss to get paid. When there are non-monetary benefits of the match, we are effectively assuming that the "technology of friendship" divides the surplus equally.

Assumptions 3 and 4 imply that, in the period in which the match occurs, each agent receives $\frac{y_H}{2}$ when the match is high quality and $\frac{y_L}{2}$ when the match is low quality. We assume that all matches contribute equally, regardless of type, to an agent's time and energy cost. Recalling that every agent is symmetric, let z_t be the total mass of matches formed by an agent in period t:

Assumption 5. *The cost to an agent of forming z_t matches is $c(z_t)$, where $c(0) = 0$, $c'(z_t) > 0$, $c''(z_t) > 0$ and $\lim_{z_t \to \infty} c'(z_t) = \infty$.*

We assume $c''(z_t) > 0$ because, as the mass of work relationships grows, the agent gets tired, has scheduling conflicts, etc.

Let us call the agent on whose decisions we are focusing ego and all other agents alters. In this subsection and the next, without loss of generality we select ego from the founding cohort. This saves on notation because period t is identical to ego's tenure. In subsection 2.3 we consider egos from later cohorts and introduce notation that allows us to distinguish ego tenure from time period.

In each period t, ego inherits from the previous period knowledge that allows him to partition alters into three sets: alters with whom he knows he is well matched, alters with whom his match quality is unknown, and alters with whom he knows he is poorly matched.

Definition. *Network and network size We call the set of alters in his firm with whom ego knows he is well matched at the end of the period his* network. *We call the mass of this set ego's* network size *and denote it by* n.

We assume that, prior to his entry to the firm, ego has never matched with any other agent in the firm and therefore has initial network size equal to zero. (We relax this assumption in Section 3.) We denote the mass of the set of alters unknown to ego in period t by u_t. The decisions that each agent needs to make in any period are how many matches z_t to form and with whom. Clearly ego prefers to match with alters within his network before trying matches with unknown alters, and prefers trying matches with unknown alters before matching with alters with whom he knows he is poorly matched. Noting that ego inherits a network of size $(1 - \delta)n_{t-1}$ from the previous period,

we can consider three cases: i) $z_t \leq (1-\delta)n_{t-1}$; ii) $(1-\delta)n_{t-1} < z_t \leq (1-\delta)n_{t-1} + u_t$; and iii) $z_t > (1-\delta)n_{t-1} + u_t$. In case i), ego's desired mass of matches is smaller than his surviving network; in case ii), ego wishes to match with his network and also try out matches with unknown alters; and in case iii), ego is willing to match with alters with whom he has not worked well because there are not enough unknown alters.

We rule out case iii) by imposing an additional condition on the cost function, derived in Appendix A and expressed in terms of the model parameters, that prevents the mass of matches ego desires to form from exceeding $(1-\delta)n_{t-1} + u_t$ in equilibrium:

Assumption A. *For* $\delta \in (0, 1 - \sqrt{1-p})$, *we assume* $c'(\delta N[1 + (1-\delta)]) \geq \frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)p\frac{(1-\delta)(1-p)}{1-\beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2}$. *For* $\delta = 0$ *or* $\delta \geq 1 - \sqrt{1-p}$), *we assume* $c'(pN) \geq \frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)p\frac{(1-\delta)(1-p)}{1-\beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2}$.

Assumption A can be thought of as an addendum to Assumption 5. Intuitively, it states that if the cost of matching rises sufficiently fast, case iii) will not occur. We will show in the next subsection that case i) also never occurs. Therefore case ii) applies: ego always matches with an unknown alter at the margin, ego matches with any alter within his network with probability one, and ego's desire to match with alters in his network is always reciprocated.

Remark. One of the most robust features of social networks is "homophily": the tendency for egos to be linked to alters who are similar to them along observable dimensions (McPherson, Smith-Lovin, and Cook 2001). It is possible to incorporate homophily into our model. Suppose we divide our continuum of agents into two types. Matches within type are identical to matches between types, except that the former (latter) are of high quality with probability $p''(p')$, where $p'' > p'$. Agents therefore match within type if they can. If case ii) prevails within type then agents never match across types, so all matches are of high quality with probability $p = p''$. The same reasoning holds for more than two types, but as the number of types increases it becomes increasingly unrealistic for there to be enough unknowns within type for case ii) to hold. In other words, we can incorporate homophily into our model by reinterpreting p as the probability that a match is of high quality within type and strengthening Assumption A sufficiently to ensure that case ii) prevails within type.[3]

---

[3]Carrell, Sacerdote, and West (2013) created treatment Air Force squadrons with concentrations of high- and low-ability students and found a negative impact on the academic performance of the low-ability students. They also found evidence that this result occurred because the low-ability students were excluded from the social networks of the high-ability students. We can see their treatment as having made it more likely that case ii) prevailed within the high-ability

2.2 Longitudinal results

In this subsection we will build toward a complete longitudinal characterization of ego's network. We will find his network size as a function of his tenure at the firm and we will find the share of his network accounted for by agents he met during each period of his tenure. To accomplish this we need to derive the time path of ego's matches with alters of unknown match quality.

We denote by x_t the mass of matches ego chooses to form in period t with alters of unknown match quality. He incurs matching costs $c(z_t) = c(x_t + (1 - \delta)n_{t-1})$. At the end of the period match qualities are revealed and surplus is divided. Ego's total per-period payoff is thus given by the sum of his payoffs from matching within his network and matching outside his network less his matching costs,

$$(1 - \delta)n_{t-1}\frac{y_H}{2} + x_t\frac{py_H + (1 - p)y_L}{2} - c(x_t + (1 - \delta)n_{t-1}).$$

His network size evolves according to

$$n_t = (1 - \delta)n_{t-1} + px_t. \tag{1}$$

We assume that if ego separates from the firm he is immediately hired by another firm, at which he accumulates a new network. (We consider the possibility of unemployment in Section 3.) Ego's value function is then given by

$$V(n_{t-1}) = \max_{x_t}\{(1 - \delta)n_{t-1}\frac{y_H}{2} + x_t\frac{py_H + (1 - p)y_L}{2} - c(x_t + (1 - \delta)n_{t-1}) + \beta[(1 - \delta)V(n_t) + \delta V(0)]\}, \tag{2}$$

where β is the constant discount factor and δ in the continuation value term is the probability that ego separates from his firm.

In Appendix B we show that the first-order condition reduces to

$$\frac{py_H + (1 - p)y_L}{2} + \beta(1 - \delta)p\frac{(1 - \delta)(1 - p)}{1 - \beta(1 - \delta)^2(1 - p)}\frac{y_H - y_L}{2} = c'(x_t + (1 - \delta)n_{t-1}) \equiv c'(z^*). \tag{3}$$

---

type. Chen-Zion (2016, Chapter 1) finds that controlling for homophily among all observable worker characteristics reduces but does not eliminate the impact of same hiring cohort on job referrals from former co-workers, which is consistent with the predictions of our model applying within type.

10

Equation (3) states that ego sets the marginal cost of a match with an unknown alter equal to the current period payoff plus the expected future payoff to having a larger network. We see from this equation that ego forms a constant total mass of matches z^* in every period. (Assumption A that rules out case iii) above also ensures the existence of a z^* that solves equation (3).) Equation (4) then yields the mass of matches with unknowns that ego forms in any period:

$$x_t = z^* - (1 - \delta)n_{t-1}. \tag{4}$$

We can see from equations (1) and (4) that the increase in network size from its inherited value $(1 - \delta)n_{t-1}$ equals a fraction p of the difference between z^* and $(1 - \delta)n_{t-1}$, so if the initial value of network size is less than z^* then it can never exceed z^*. Since network size begins at zero, we can rule out case i) above ($z_t \leq (1 - \delta)n_{t-1}$).

We can substitute equation (4) into equation (1), yielding

$$n_t = pz^* + (1 - p)(1 - \delta)n_{t-1}.$$

We can then derive the complete time paths for network size and for the mass of matches with unknowns ego forms in each period:

$$n_t = \sum_{\tau=0}^{t} (1 - p)^\tau (1 - \delta)^\tau pz^* \qquad x_t = z^* - (1 - \delta)\sum_{\tau=0}^{t-1} (1 - p)^\tau (1 - \delta)^\tau pz^*. \tag{5}$$

Note that the expression for n_t gives the value of network size at the end of the period. In particular, for $t = 0$ the expression yields $n_0 = pz^*$, but the value of network size at the beginning of period 0 is zero, which also implies $x_0 = z^*$.

As $t \to \infty$, network size and the mass of matches of unknown quality ego forms approach their steady state values:

$$n_t \to \bar{n} = \frac{p}{[\delta + p(1 - \delta)]} z^* \qquad x_t \to \bar{x} = \frac{\delta}{[\delta + p(1 - \delta)]} z^*. \tag{6}$$

We see from equations (5) and (6) that n_t increases monotonically from zero to its steady state value, which never exceeds z^*. Note that steady state network size increases with the probability of a good match and decreases with the rate of network decay. If the network does not decay ($\delta = 0$), then in the limit all matches are within network and matches with unknowns drop to zero.

The firm becomes completely static, with a fixed set of agents and fixed relationships between them.

Inspection of equation (5) establishes our first proposition:

Proposition 1. *Ego's mass of matches with unknowns x_t declines monotonically with t.*

Proposition 1 states that ego becomes monotonically less open over time to meeting alters of unknown match quality. This occurs because ego's network size increases monotonically with time whereas his optimally chosen capacity for work relationships remains unchanged.

Clearly ego's network is valuable to him, in that the same mass of work relationships without a network yields less benefit. We can compute the value of a network of size n_{t-1} explicitly by comparing $V(n_{t-1})$ to $V(0)$.[4] Since

$$V'(n_{t-1}) = \frac{(1-\delta)(1-p)}{1-\beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2}, \text{ then}$$

$$V(n_{t-1}) = \frac{(1-\delta)(1-p)}{1-\beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2}n_{t-1} + V(0).$$

It follows that

$$V(n_{t-1}) - V(0) = \frac{(1-p)(1-\delta)}{1-\beta(1-\delta)^2(1-p)}\frac{(y_H - y_L)}{2}n_{t-1}.$$

Inspection of this expression establishes:

Proposition 2. *The value of ego's network is increasing in n_{t-1}, decreasing in δ, decreasing in p, increasing in β, and increasing in $y_H - y_L$.*

Proposition 2 states that the value of ego's network is increasing in (last period's) network size, decreasing in the rate of network decay, decreasing in the probability that a match with an unknown is good, decreasing in the rate at which future payoffs are discounted, and increasing in the difference between good and bad match values. Since matching within his network guarantees ego a good match quality, his network is less valuable if the probability is high that a match with an agent of unknown match quality is good or if the difference between good and bad match values is small. Discounting reduces the value of ego's network since part of its value is that ego has to match with a smaller mass of unknowns in the future.

---

[4]It is straightforward to show that $V(0) = \frac{z^* c'(z^*) - c(z^*)}{(1-\beta)}$.

Since n_{t-1} is monotonically increasing with time, we have the following corollary to Proposition 2:

Corollary 1. *The value of ego's network is monotonically increasing with his tenure at the firm.*

We conclude this subsection with our results on history dependence. In our model an ego with tenure t looking at his network retrospectively will see that he met the alters at various times t'. We denote the mass of matches that were first formed in t' that are still in ego's network at the end of t by $n_t(t') = px_{t'}$ if $t = t'$ and $(1 - \delta)^{t-t'}px_{t'}$ if $t > t'$.

Definition. *History dependence $HD_t(t') \equiv \frac{n_t(t')}{n_t}$, the probability that a member of ego's network in t resulted from an initial meeting from a given previous period t'.*

Substituting for $x_{t'}$ in the expression for $n_t(t')$ using equation (4) yields $HD_t(t') = \frac{(1-\delta)^{t-t'}p[z^* - (1-\delta)n_{t'-1}]}{n_t}$. Note that if $\delta = 1$, $HD_t(t') = 0$ for $t > t'$: if network decay is complete, there is no history dependence.

We see that the past time period t' has two counteracting influences on our measure of history dependence. On the one hand, since $n_{t'-1}$ increases with t', $HD_t(t')$ tends to decrease with t', reflecting the "front-loading" of agents' networks caused by persistence of relationships and time constraints as discussed above. On the other hand, since $(1 - \delta)^{t-t'}$ increases with t', $HD_t(t')$ tends to increase with t', showing how a constant rate of decay of network relationships tends to establish a more conventional pattern of history dependence where more recent meetings are more influential. We also see that $HD_t(t')$ is decreasing in t, so the longer is an agent's tenure in a firm the less is the influence on his network of meetings from any particular time in the past.

The influences of t' and t on our measure of history dependence can be summarized in the following proposition:

Proposition 3. *Assume $p > \frac{\delta}{(1-\delta)}$. For $\delta > 0$, there exists a $\underline{t}' \geq 1$ such that $HD_t(t')$ is monotonically decreasing in t' for $t' < \underline{t}'$ and monotonically increasing in t' for $t' > \underline{t}'$. Moreover, there exists a $\underline{t} > \underline{t}'$ such that $HD_t(0) > HD_t(t)$ for $t < \underline{t}$ and $HD_t(0) < HD_t(t)$ for $t > \underline{t}$.*

The proof of Proposition 3, and all remaining proofs in this section, are in Appendix B.

Proposition 3 shows that when an agent's tenure in a firm is sufficiently short ($t \leq \underline{t}'$), representation of alters in his network is least influenced by his most recent meetings and most influenced by his very first meetings. With longer tenure ($t > \underline{t}'$), the share of alters resulting his most

13

recent meetings increases relative to less recent meetings, creating a U-shaped pattern of history dependence where ego's network is dominated by alters he met at the beginning and most recent periods of his organizational career. Eventually $(t > \underline{t})$, the influence of the distant past diminishes sufficiently that the most recent meetings account for the largest share of alters of any period.

Let us consider two informative special cases. It is helpful if from this point forward we denote firm age by T.

Example 1: $\delta = 0$. We have $n_t(t') = pz^*[1 - \sum_{\tau=0}^{t'-1}(1-p)^\tau p] = pz^*[1-1+(1-p)^{t'}] = pz^*(1-p)^{t'}$. That is, as we move from the past toward the present, $HD_t(t')$ decreases at rate $(1-p)$.

Example 2: $\delta > 0$, T large so that $n_T \approx \bar{n}$. In this case $\bar{x} = \frac{\delta \bar{n}}{p}$, $HD_T(T) \approx \frac{p\bar{x}}{\bar{n}} = \delta$, and $HD_T(t') \approx \delta(1-\delta)^{T-t'}$. That is, as we move t' from the present toward the past, $HD_T(t')$ decreases at rate $(1-\delta)$.

The second example shows that, for any positive rate of network decay, a conventional pattern of network history dependence will be established if a sufficient amount of time passes. For relatively short time horizons or small rates of network decay, the greatest representation in ego's network will be from alters he met early in his career at the firm.

Figure 1: History Dependence
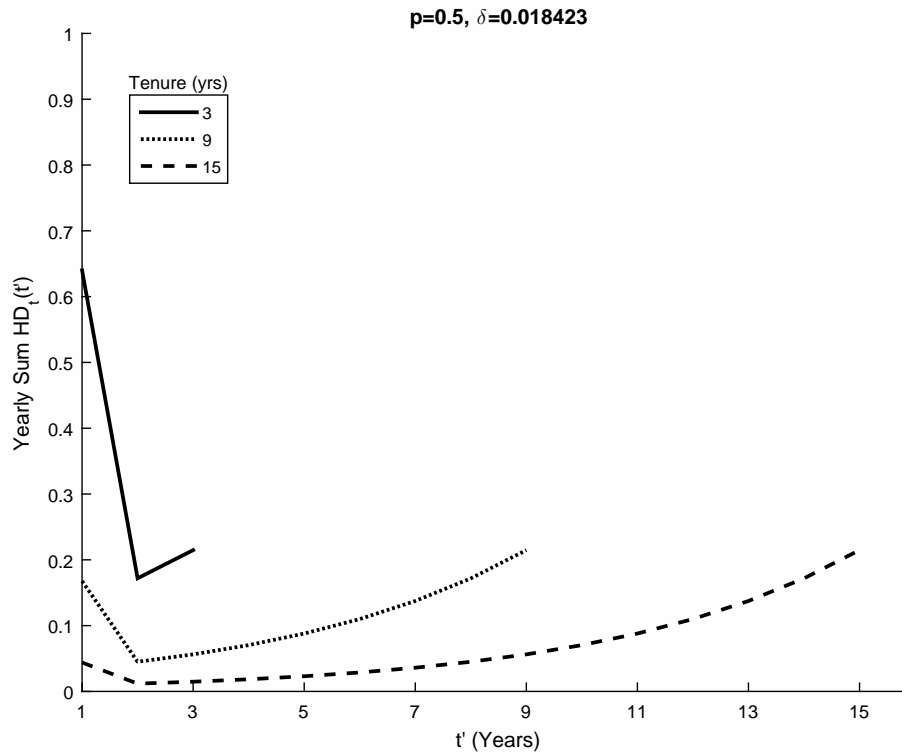

p=0.5, δ=0.018423

Figure 1 calculates the time pattern of history dependence for various lengths of ego tenure at the firm, where time is measured in years to build intuition. The calculation assumes an even bet that matches are of high quality ($p = 0.5$) and a 20 percent rate of network decay per year ($\delta \approx 0.018$, where the underlying periods are months, the typical unit used to measure tenure in administrative data). At three years of tenure, more than 60 percent of ego's network consists of alters he met in his first year at the firm. We also see that the U-shaped pattern of history dependence already appears, with greater representation of alters met during ego's third year at the firm than during his second. At nine years of tenure, representation of alters met during ego's most recent year at the firm finally surpasses representation of alters met during his first year. After fifteen years of tenure ego's pattern of history dependence is dominated by network decay.

2.3 Cross-sectional results

Absent retrospective interviews, we cannot see in what periods ego initially met with which alters. Alters are, however, horizontally differentiated by cohort of entry to the firm, which is much more easily observable. A natural analog to $HD_t(t')$ is the share of each cohort in ego's network at time t. This equals the probability, conditional on remaining in the firm at time t, that a given alter in each cohort belongs to ego's network, multiplied by cohort size and divided by ego's network size. Insofar as one can observe alters' exit from as well as entry to the firm, the conditional probability is of independent empirical interest. In this subsection we will prove two results on the conditional probability that a given alter in each cohort belongs to ego's network and two results on cohort shares of ego's network.

Let us denote cohort by c. Tenure of an agent in cohort c is given by $t - c$. For the founding cohort, tenure is given by $t - 0 = t$. However, we can no longer afford the notational convenience of using the founding cohort to represent all cohorts. We will now denote network size and mass of matches with agents whose match quality is unknown by n_t^c and x_t^c, respectively. To compute n_t^c and x_t^c, we can use equation (5) substituting $t - c$ for t. For example, $x_c^c = z^*$, and $x_c^{c-1} = z^* - (1 - \delta)pz^*$.

Let $P_t^c(c')$ be the probability that a given alter in cohort c' is in the network of a given ego in cohort c at the end of period t, conditional on his remaining with the firm. Let L_t^c be the mass of matches with unknown alters formed in period t by agents whose match quality is unknown to a given ego in cohort c. Noting that $P_t^c(c')$ equals p times the probability that a given alter in cohort

c' who remains with the firm is of known match quality to a given ego in cohort c at the end of period t, we have

$$P_t^c(c') = \begin{cases} 0 & \text{if} \quad t < \max\{c, c'\} \\ P_{t-1}^c(c') + [1 - \frac{P_{t-1}^c(c')}{p}]px_t^c\frac{x_t^{c'}}{L_t^c} & \text{if} \quad t \geq \max\{c, c'\} \end{cases} \tag{7}$$

$$L_t^c = \begin{cases} N(1-\delta)^t x_t^0 + \sum_{c'=1}^t \delta N(1-\delta)^{t-c'} x_t^{c'} & \text{if} \quad t = c \\ N(1-\delta)^t [1 - \frac{P_{t-1}^c(0)}{p}]x_t^0 + \sum_{c'=1}^t \delta N(1-\delta)^{t-c'}[1 - \frac{P_{t-1}^c(c')}{p}]x_t^{c'} & \text{if} \quad t > c \end{cases} \tag{8}$$

Equations (7) and (8) provide recursive solutions for $P_t^c(c')$ and L_t^c. We can use these solutions to compute the share $S_t^c(c')$ of any cohort c' in the network of an ego in cohort c at the end of period t:

$$S_t^c(c') = P_t^c(c')N^{c'}(1-\delta)^{t-c'}/n_t^c \tag{9}$$

where $N^{c'} = N$ if $c' = 0$ and $N^{c'} = \delta N$ if $c' > 0$.

Despite the complex, recursive formulas for $P_t^c(c')$ and $S_t^c(c')$, we are able to derive some useful analytical results. First, we show that the conditional probability that an alter from ego's own cohort is a member of his network is greater than the conditional probability that an alter from any incumbent cohort is a member of ego's network:

Proposition 4. $P_t^c(c) > P_t^c(c')$ *for all* $c' \in [0, c-1]$, *for all* $t \geq c$.

The intuition for Proposition 4 is provided by Proposition 1: alters become monotonically less open to meetings with unknowns as their tenure increases. Proposition 4 applies to incumbent cohorts. Does the intuition for Proposition 4 apply to later cohorts, since ego becomes monotonically less open to meetings with unknowns as his tenure increases? Yes, but this intuition is insufficient. It is possible that, because some alters are known to him, fewer total meetings with unknowns will be available to ego when he first meets with later cohorts than were available when he first met with his own cohort, raising $P_{c'}^c(c')$ relative to $P_c^c(c)$ for $c' > c$ and creating the possibility that $P_t^c(c') > P_t^c(c)$ for some c', t. The condition stated in Lemma 1 eliminates this possibility:

Lemma 1. *If* z^*/N *is sufficiently small,* $x_t^c/L_t^c > x_{t+b}^c/L_{t+b}^c$ *for all* $b \in [1, T-c]$, *for all* $t \geq c$.

The Lemma states that the ratio of desired to available meetings with unknowns by ego in cohort c is greater in period t than b periods later. It follows from diminishing openness to meetings

with unknowns that desired meetings decrease from period t to $t + b$, and the condition that z^*/N is sufficiently small prevents any reduction in available meetings from overturning the result. Note that if we were to weaken Assumption 2 by decoupling firm hiring from network decay (layoffs), firm shrinkage could undermine Lemma 1. This in turn would make it possible for the next two propositions (but not Propositions 1-4) to fail.

With Lemma 1 in place, we can prove a parallel to Proposition 4 for cohorts that arrive after ego's cohort:

Proposition 5. *If z^*/N is sufficiently small, $P_t^c(c) > P_t^c(c')$ for all $c' \in [c + 1, T]$, for all $t \geq c'$.*

Together, Propositions 4 and 5 demonstrate *cohort attachment*: conditional on remaining with the firm, a member of ego's own cohort is more likely to belong to his network than is a member of any other cohort.

Remark. Consider a change in perspective from the egocentric networks of the agents in the firm to the network of relationships in the firm as a whole, and measure "clustering" by the average probability that agents j and k have a relationship given that both have a relationship with i (Jackson 2008, p. 35). Since probabilities of relationships in our model are independent, as in a random graph, this clustering measure simply equals the average probability that two agents have a relationship. Considering a cohort or collection of cohorts as subnetworks of the firm network, Propositions 4 and 5 imply that clustering is greater for any one cohort than for any collection of cohorts.

Our results for cohort shares of ego's network differ from our results for conditional probabilities of belonging to ego's network because they incorporate network decay. Over time, network decay erodes the dominant position of ego's own cohort in his network and opens up space for more recent cohorts.

Proposition 6. *Consider a firm of age $T \geq 2$. If z^*/N is sufficiently small and $\delta[1+\delta p/2(1-p)] < 1/2$, then $S_t^c(c')$ reaches its maximum over cohorts $c' \in [1, T]$ at cohort c for at least the two most recent cohorts, i.e., $c \in [T - 1, T]$.*

The condition on δ in Proposition 6 is needed because otherwise the higher conditional probability that an alter in ego's network is in ego's own cohort than in the next cohort is dominated by turnover of his own cohort.[5] However, any positive rate of turnover must eventually cause the own

---

[5]It can be shown that the upper bound on δ in Proposition 6 lies between 0.472 and 0.5, depending on the value of p.

cohort network share to fall below the share of the most recent cohort.

Proposition 7. *If z^*/N is sufficiently small and $T - c$ is sufficiently large, $S_T^{c'}(T) > S_T^{c'}(c')$ for all cohorts $c' \in [1, c]$.*

The founding cohort ($c = 0$) is excluded from Propositions 6 and 7 because, under the condition on δ given in Proposition 6, it is of larger initial size than all the other cohorts.

The two propositions combined predict that the networks of more recent cohorts are dominated by agents that joined the firm at or near the same time as themselves, the networks of the oldest cohorts are dominated by the most recent cohorts, and the networks of agents with intermediate tenure are dominated by a combination of cohorts close to their own and the most recent cohorts.

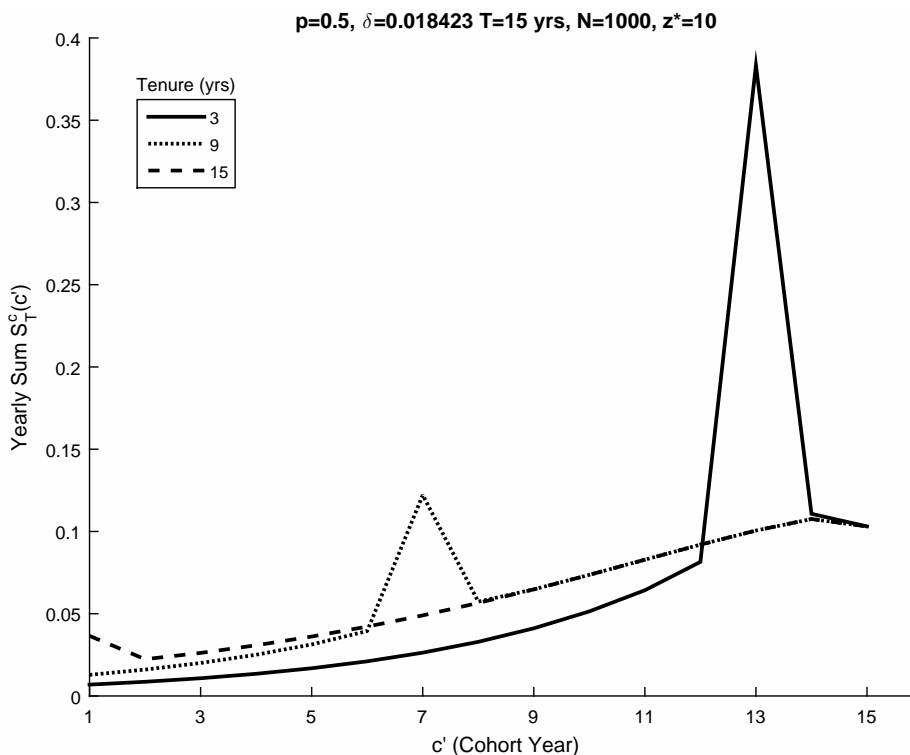Figure 2: Cohort Network Shares

Figure 2 illustrates Propositions 6 and 7. It imitates Figure 1: the parameters p and δ are the same, the firm is 15 years old, the underlying periods are months that are aggregated into years to build intuition, and the plots are for egos with 3, 9, and 15 years tenure. More specifically, the egos we plot are assigned to the seventh month of their cohort years, so that the egos with 3, 9, and 15 years tenure entered the firm in the seventh month of years 13, 7, and 1, respectively. We see

18

that the share of own cohort year in the network of the ego with 3 years tenure is four times that of any other cohort year, whereas the share of own cohort year in the network of the ego with 9 years tenure is only slightly larger than the shares of the most recent cohorts, and the most recent cohorts are clearly the largest in the network of the ego with 15 years tenure. Figure 2 also shows that by the end of their cohort years the egos' networks are essentially in the steady state given by equation (6), so that the network shares of all cohorts arriving later are virtually equal across the egos despite their different tenures. The reason for the downturn in cohort network share at T is that for the first few periods after an alter enters the firm his probability of having met ego increases at a faster rate than δ.

3   Contacts and Job Referral

Job referrals are a canonical application of network models in economics. The small but growing empirical literature cited in our Introduction specifically analyzes job referrals to ego from alters he met in previous employment.[6] This indicates the importance of exactly the kind of history-dependent network formation we emphasize in this paper. However, a job referral necessarily connects an ego outside a firm to alters inside the firm, whereas our model has focused entirely on network formation and operation within a firm (or, more broadly, within any one organization).

In this section ego will take into account the value that alters may provide for him if he separates from his firm. We will find conditions under which the qualitative results of the previous section do not change despite this added complexity, and for a special case we will determine the quantitative changes in ego's network formation program. This extension of our model allows us to predict which alters are more likely to generate job referrals for ego.

Let us extend our model to include many firms, finite in number. We assume that an agent can be employed by at most one firm in any period. An agent who does not work for any firm in a given period is unemployed in that period. Firms can form or dissolve. We assume an exogenous, constant probability of firm dissolution.

We only consider referrals of unemployed agents to firms. The pool of unemployed is filled by the exogenous separations of the previous section and by exogenous firm dissolution.[7] The pool

---

[6]There is a much broader job referral literature, which covers all types of connections between egos and alters rather than focusing on those formed through previous work at a common employer. For surveys see Ioannides and Loury (2004) and Topa (2011).

[7]The small empirical literature cited above has only considered referrals of egos who are unemployed because their

of unemployed is drained by firms replacing the workers who separated from them and by the founding of new firms. The mass of hires, respectively δN and N, is exogenous.

When alters in ego's network separate from him because they or ego leave their firm or because their firm dissolves, they become the *contacts* of ego.

Definition. *Contacts Alters with whom ego knows he is well matched but who are not in ego's current organization are the* contacts *of ego.*

A contact is different from an alter in ego's network because, in the current period, ego cannot form a match with him. Ego therefore neither derives value nor incurs costs from his contacts in the current period. We assume that ego's contacts return to unknown match quality at a constant, exogenous rate, as ego and alters "drift apart" over time following separation from their common employer. This is consistent with the results of Eliason, Hensvik, Kramarz, and Skans (2019), who find that former co-workers are an important source of job referrals "especially if they worked together relatively recently (i.e., the value of the connection seems to depreciate over time)."

In line with the referral literature, we assume an information structure such that a firm only knows of contacts that are brought to its attention by its current employees. In particular, the firm is unaware of contacts that may exist between the unemployed workers themselves. If the firm knew of such contacts, it might want to hire a "ready-made" network of unemployed workers. Under our assumed information structure, and given the assumption in the previous section that the firm prefers high to low quality matches, it is clear that the interests of the firm's employees and the firm are aligned. The employees want the firm to be aware of their contacts among the unemployed, and the firm wants to hire the unemployed workers with the greatest mass of contacts among its employees. The firm will rank the unemployed by these masses and hire until all vacancies are filled or until it exhausts all unemployed with positive masses, in which case we assume it chooses randomly among the remaining unemployed.[8]

Recall that ego's contacts are alters who were formerly in his networks. When ego meets alters,

---

previous firms dissolved. By focusing on egos whose firms have closed, these papers avoid a potential selection bias from studying egos who have been laid off from thriving firms.

[8]The two cases have different implications for the relationship between an unemployed worker's mass of contacts at a firm and the probability that he is hired. In the case where all unemployed workers with positive masses are hired, this probability takes a discrete jump when the worker's mass increases from zero to positive, then remains constant. In the case where not all unemployed workers with positive masses are hired, this probability strictly increases with the worker's mass of contacts at the firm. Saygin, Weber, and Weynandt (forthcoming) use as their proxy for referral an indicator for whether the number of former co-workers at the hiring firm is positive, which corresponds to our case where all unemployed workers with positive masses of contacts at the firm are hired.

then, he recognizes that in the future they may become contacts who help him out of unemployment through referral. We greatly simplify ego's decision problem by assuming there is only one round of referral, so that workers hired through referral cannot refer other workers in turn. In this case, an alter's contacts are of no consequence to ego: if an alter is hired because of his contacts, he cannot refer an unemployed ego by assumption. Alters in his firm therefore remain symmetric from ego's point of view within the three sets delineated in the previous section: alters with whom he knows he is well matched, alters with whom his match quality is unknown, and alters with whom he knows he is poorly matched.

We would like the results of the previous section, including cohort attachment, to continue to apply to the networks that egos form with workers who were of unknown match quality to them when they or the workers joined their firms. The key to retaining these results is retaining Proposition 1: ego becomes monotonically less open over time to meeting alters of unknown match quality.

We can see immediately that ego's behavior in the previous section is not qualitatively changed by the prospect of dissolution of his employer. From ego's point of view separation from his current firm and dissolution of his current firm are equivalent, so he can add the probabilities together, leaving the results obtained in the previous section qualitatively unchanged.

Likewise, being hired through referral does not in itself qualitatively change ego's subsequent behavior, provided his mass of contacts at the hiring firm is small (in particular, smaller than his steady state network size). It only means that his initial network size is positive rather than zero.

The two changes made to our model in this section that threaten Proposition 1 are 1) ego's network at his current employer grows when his employer hires his contacts as well as when he matches with unknowns, and 2) ego has an additional incentive to match with unknowns, because they could turn out to be useful contacts in the future. 1) becomes a problem if the firm tends to hire ego's contacts early rather than late in his tenure, because this will cause his demand for matches with unknowns to increase rather than decrease over time. 2) becomes a problem if ego's marginal value of adding contacts grows with his stock of contacts, causing his desired mass of matches z_t to increase over time.

We now re-solve ego's problem in the previous section incorporating the changes to our model in this section, under the assumptions that 1) the firm hires an exogenous mass of ego's contacts each period, and 2) ego's marginal value of adding contacts does not increase with his stock of

contacts. We show that ego's desired total mass of matches remains constant if his marginal value of adding contacts is constant, and decreases over time if his marginal value of adding contacts is decreasing. An additional condition limiting the decrease in firm hiring of ego's contacts ensures that his demand for matches with unknowns declines monotonically with time. Hence these assumptions and condition are sufficient to retain our qualitative results from the previous section for ego's network with workers who were of unknown match quality to him when he or the workers joined the firm.

If the cohort attachment result continues to apply to ego's network in any firm, it will extend to his contacts as well. That is, just as co-workers in ego's cohort are more likely to belong to his network, former co-workers who were in ego's cohort are more likely to be his contacts. If former co-workers who were in ego's hiring cohort are more likely to be his contacts, they are more likely to generate referrals for him. This is the most readily testable prediction we can add to the existing empirical literature on job referrals, with Chen-Zion (2016, Chapter 1) cited above being a first attempt in that direction. Note that our prediction does not apply to former co-workers who referred or were referred by ego, because they were not of unknown match quality to him when he or the workers joined the firm.

We introduce some new notation:

$r_t =$ mass of ego's contacts hired by his employer at the beginning of period t

$g_t =$ ego's stock of contacts at the beginning of period t.

We now have two state variables, n_t and g_t. The equations of motion are

$$n_t = (1 - \delta)n_{t-1} + r_t + px_t$$

$$g_t = (1 - \psi)g_{t-1} + \delta n_{t-1},$$

where ψ is the rate at which ego's contacts return to unknown match quality.

We must now distinguish between ego's value function when employed, $V(n_{t-1}, g_{t-1})$, and his

value function when unemployed, $V_u(g_t)$. We have[9]

$$V(n_{t-1}, g_{t-1}) = \max_{x_t}\{[(1-\delta)n_{t-1} + r_t]\frac{y_H}{2} + x_t\frac{py_H + (1-p)y_L}{2} - c(x_t + r_t + (1-\delta)n_{t-1})$$
$$+ \beta[(1-\delta)V(n_t, g_t) + \delta V_u(g_t)]\}. \tag{10}$$

We assume that $V_u(g_t)$ is concave, so that $V_u'' \leq 0$. In the case where the marginal value of adding contacts V_u' is constant, we can find a closed-form solution for x_t as we did in Section 2. In the case where V_u' is decreasing, we cannot find a closed-form solution for x_t but can show that it declines monotonically with time.[10]

The first order condition yields

$$\frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)p\frac{\partial V}{\partial n_t} = c'(x_t + r_t + (1-\delta)n_{t-1}). \tag{11}$$

In Appendix B we show that, if V_u' is constant, we obtain a constant solution for $\frac{\partial V}{\partial n}$:

$$\frac{\partial V}{\partial n} = \frac{(1-\delta)(1-p)\frac{y_H - y_L}{2} + \beta\delta^2 V_u'[1 + \frac{\beta(1-\delta)(1-\psi)}{1 - \beta(1-\delta)(1-\psi)}]}{1 - \beta(1-\delta)^2(1-p)}. \tag{12}$$

Except for the term including V_u', the right-hand side of this expression is the same as the right-hand side of the expression for $V'(n)$ in Section 2. Substituting back into the first-order condition, we obtain the equivalent of equation (3):

$$\frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)p\frac{(1-\delta)(1-p)\frac{y_H - y_L}{2} + \beta\delta^2 V_u'[1 + \frac{\beta(1-\delta)(1-\psi)}{1 - \beta(1-\delta)(1-\psi)}]}{1 - \beta(1-\delta)^2(1-p)}$$
$$= c'(x_t + r_t + (1-\delta)n_{t-1}) \equiv c'(\hat{z}^*), \tag{13}$$

where $\hat{z}^*$ is the constant mass of matches that ego forms in every period. A comparison of equations (3) and (13) shows that the left-hand side of equation (13) is strictly larger if $V_u' > 0$. We have therefore proven that $\hat{z}^* > z^*$, where z^* solves equation (3):

---

[9]To conserve on notation we treat the probability of dissolution of the firm as negligible relative to the probability δ of separation from the firm. Denoting the firm dissolution probability by $\tilde{\delta}$ and writing $\beta[(1 - (\delta + \tilde{\delta}))V(n_t, g_t) + (\delta + \tilde{\delta})V_u(g_t)]$ in equation (10) would leave our results qualitatively unchanged.

[10]The marginal value of adding contacts will be influenced by the marginal impact of contacts on the probability of finding a new job and on the value of a new job. Endogenizing the marginal value of adding contacts in general equilibrium is left to future research.

Proposition 8. *If the marginal value of adding contacts V_u' is constant and positive, ego forms a constant mass of matches in every period, and this mass is strictly greater than when contacts have no value.*

Intuitively, a constant, positive marginal value of adding contacts gives ego a constantly greater incentive to form matches relative to the model of Section 2 in which contacts have no value. Note that to maintain $x_t > 0$ we must assume the exogenous masses of referrals satisfy $r_t < \hat{z}^* - (1-\delta)n_{t-1}$: the firm does not hire such a large mass of ego's contacts in any period that ego does not want to match with unknowns. Next we add a restriction on changes to this restriction on levels of r_t.

Given Proposition 8, as ego's network grows the mass of matches he seeks with agents of unknown match quality must decline, provided his firm does not decrease its hiring of his contacts (who are of known match quality) too fast. In Appendix B we prove:

Proposition 9. *If the marginal value of adding contacts V_u' is constant, x_t decreases monotonically with time if the firm's hiring of ego's contacts satisfies*
$[1 - (1-\delta)(1-p)]r_t - r_{t+1} < (1-\delta)^{t+1}(1-p)^t p \hat{z}^*$, *where $\hat{z}^*$ is determined by equation* (13).

One way to put some structure on r_t is to note that while ego is employed at the firm his stock of prior contacts is decreasing at rate ψ. It is then reasonable to think that the firm's hiring of ego's contacts will decrease at the same rate: $r_{t+1} = r_t(1-\psi)$. In this case $\psi < (1-\delta)(1-p)$ is a sufficient condition for x_t to decrease monotonically.

We conclude by considering the case where $V_u'' < 0$. We immediately see that our constant solution above for $\partial V/\partial n$ can no longer hold since V_u' will decrease as g increases over time, causing $\partial V/\partial n$ to decrease rather than remain constant if equation (12) holds. Indeed, $\partial V/\partial n$ decreasing over time is precisely what we want to show, because we can see from the first order condition (11) that this implies the desired mass of matches $\hat{z}_t^*$ will decrease with time, which will reinforce the monotonic decrease of x_t.

Standard arguments show that V is concave, hence $\partial^2 V/\partial n^2 \leq 0$. What remains to be shown is $\partial^2 V/\partial n \partial g \leq 0$. Then, as n and g increase toward their steady state values, $\partial V/\partial n$ (weakly) decreases over time. The proof that $V_u'' < 0$ implies $\partial^2 V/\partial n \partial g \leq 0$ is given at the end of Appendix B.

As before, to maintain $x_t > 0$ we must assume that the exogenous masses of referrals satisfy $r_t < \hat{z}_t^* - (1-\delta)n_{t-1}$. A final concern is that with $\hat{z}_t^*$ decreasing rather than constant, the logic

24

we used in the model without contacts to rule out $z_t \leq (1 - \delta)n_{t-1}$ (case i) in Section 2) no longer applies. Since it is easily shown that $\hat{z}_t^*$ cannot decrease below z^*, to rule out this case it is sufficient to ensure that $(1 - \delta)n_{t-1}$ cannot increase to z^*. A low enough value of p, a high enough value of δ, or some combination will ensure this.

4   Conclusions

When chance meetings reveal compatibility, the agents involved have incentives to maintain their relationships. Accumulating relationships becomes increasingly costly, however, causing agents to become less open to chance meetings over time. The interaction of this dynamic with turnover leads to Proposition 3 in our paper, describing for an egocentric network in an organization the time pattern of history dependence as a function of ego's tenure in the organization. Mutual openness of newly arrived agents also leads to the cross-section prediction of "cohort attachment," a tendency for members of ego's hiring cohort to be disproportionately represented in his network. In an extension of our model we allowed members of ego's network who are subsequently split across many organizations to be his "contacts." The desire of contacts to renew their successful working relationships leads to job referrals. We provide sufficient conditions under which agents form qualitatively the same longitudinal and cross-sectional patterns of network links within their organizations as in the model without contacts. Under these sufficient conditions our model predicts that former co-workers who provide referrals will be drawn disproportionately from the referred workers' hiring cohorts at their previous employers.

In addition to homophily, non-mechanical theories of network formation emphasize contingency (e.g., Small and Sukhu 2016) and strategy (e.g., Jackson and Wolinsky 1996). Random matching with agents of unknown match quality places our model clearly in the former group. Our contribution shows that contingency does not imply complete unpredictability. On the contrary, the time structure induced by the persistence and costliness of relationships allows us to make a number of testable predictions. More broadly, we hope our model helps investigators to "see" endogenous networks, just as we can see networks based on ascriptive characteristics such as subcaste (Munshi and Rosenzweig 2016).

Our prediction of cohort attachment receives some tentative support from the papers cited in our Introduction, though clearly more work is needed. In the future, surveys of individuals in

organizations could map out the times at which they met the alters in their networks and thus directly test Proposition 3. Data at the firm rather than individual level could also be relevant if the agents in our model were firms instead of individuals, establishing relationships with other firms. Proposition 3 then suggests, for example, that the networks of young firms would be dominated by the clients and suppliers with which they were matched at startup, whereas the networks of firms that survive to "maturity" would be dominated by more recent clients and suppliers. As Chaney (2014) and others have shown, modeling firms as agents forming networks can be especially productive for the study of international trade. Our model could help to interpret and add structure to the history dependence in international trade found in the work of Eichengreen and Irwin (1998) and Head, Mayer, and Ries (2010). The generality of our framework should accommodate many applications.

References

BANDIERA, O., I. BARANKAY, AND I. RASUL (2008): "Social Capital in the Workplace: Evidence on its Formation and Consequences," *Labour Economics*, 15(4), 724–748.

CALVO-ARMENGOL, A., AND M. O. JACKSON (2004): "The effects of social networks on employment and inequality," *The American Economic Review*, 94(3), 426–454.

CHANEY, T. (2014): "The Network Structure of International Trade," *American Economic Review*, 104(11), 3600–3634.

CHEN, G. (2005): "Newcomer adaptation in teams: Multilevel antecedents and outcomes," *Academy of Management Journal*, 48(1), 101–116.

CHEN-ZION, A. (2016): "Networks in Labor Markets and Information," Ph.D. thesis, University of California, San Diego.

CINGANO, F., AND A. ROSOLIA (2012): "People I Know: Job Search and Social Networks," *Journal of Labor Economics*, 30(2), 291–332.

EICHENGREEN, B., AND D. A. IRWIN (1998): "The Role of History in Bilateral Trade Flows," in *The Regionalization of the World Economy*, pp. 33–62. University of Chicago Press.

ELIASON, M., L. HENSVIK, F. KRAMARZ, AND O. N. SKANS (2019): "Social Connections and the Sorting of Workers to Firms," IZA Discussion Paper 12323.

GALENIANOS, M. (2013): "Learning about match quality and the use of referrals," *Review of Economic Dynamics*, 16(4), 668–690.

——— (2014): "Hiring Through Referrals," *Journal of Economic Theory*, 152, 304–323.

GALEOTTI, A., AND L. P. MERLINO (2014): "Endogenous job contact networks," *International Economic Review*, 55(4), 1201–1226.

GLITZ, A. (2017): "Coworker Networks in the Labour Market," *Labour Economics*, 44, 218–230.

HEAD, K., T. MAYER, AND J. RIES (2010): "The Erosion of Colonial Trade Linkages after Independence," *Journal of International Economics*, 81(1), 1–14.

IOANNIDES, Y. M., AND L. D. LOURY (2004): "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42(4), 1056–1093.

JACKSON, M. (2008): *Social and Economic Networks*. Princeton Univ Press.

JACKSON, M. O., AND A. WOLINSKY (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.

JOVANOVIC, B. (1979): "Job Matching and the Theory of Turnover," *The Journal of Political Economy*, 87(5), 972.

MCPHERSON, M., L. SMITH-LOVIN, AND J. M. COOK (2001): "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, pp. 415–444.

MONTOGOMERY, J. D. (1991): "Social networks and labor-market outcomes: Toward an economic analysis," *The American Economic Review*, 81(5), 1408–1418.

MOSCARINI, G. (2005): "Job Matching and the Wage Distribution," *Econometrica*, 73(2), 481–516.

MUNSHI, K., AND M. ROSENZWEIG (2016): "Networks and misallocation: Insurance, migration, and the rural-urban wage gap," *The American Economic Review*, 106(1), 46–98.

PALLAIS, A., AND E. G. SANDS (2016): "Why the referential treatment? Evidence from field experiments on referrals," *Journal of Political Economy*, 124(6), 1793–1828.

REAGANS, R. E., E. ZUCKERMAN, AND B. MCEVILY (2007): "On firmer ground: The collaborative team as strategic research site for verifying network-based social-capital hypotheses," *The Missing Links: Formation and Decay of Economic Networks*, pp. 147–82.

SACERDOTE, B. (2014): "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?," *Annual Review of Economics*, 6(1), 253–272.

SAYGIN, P. O., A. WEBER, AND M. A. WEYNANDT (forthcoming): "Coworkers, Networks, and Job-Search Outcomes," *ILR Review*.

SMALL, M. L., AND C. SUKHU (2016): "Because they were there: Access, deliberation, and the mobilization of networks for support," *Social Networks*, 47, 73–84.

TETT, G. (2015): *The silo effect: The peril of expertise and the promise of breaking down barriers*. Simon and Schuster.

TOPA, G. (2011): "Labor markets and referrals," *Handbook of Social Economics, Amsterdam: Elsevier Science*, pp. 1193–1221.

TOPEL, R. (1991): "Specific capital, mobility, and wages: Wages rise with job seniority," *Journal of Political Economy*, 99(1), 145–176.

WAGNER, W. G., J. PFEFFER, AND C. A. O'REILLY III (1984): "Organizational Demography and Turnover in Top-management Group," *Administrative Science Quarterly*, pp. 74–92.

ZENGER, T. R., AND B. S. LAWRENCE (1989): "Organizational Demography: The Differential Effects of Age and Tenure Distributions on Technical Communication," *Academy of Management Journal*, 32(2), 353–376.

Appendices

A Conditions Ensuring Sufficient Availability of Agents of Unknown Match Quality

We derive conditions on the cost of matching that ensure that agents will not desire to form so many matches that they have to form matches they know to be of low quality. That is, our sufficient conditions ensure that in equilibrium there are more agents of unknown match quality to ego than the number of matches he desires to form outside his network.

We start with the special case $\delta = 0$, so the set of agents is fixed. We assume that enough unknowns are available to ego in every period, then derive a condition under which this assumption holds. Under this assumption, ego builds his network as described by equations (4) and (5) (with $\delta = 0$). In period 0 he meets with $x_0 = z^*$ unknowns, yielding a network of size $n_1 = pz^*$ in period 1. In period 1 he meets with $x_1 = z^* - pz^*$ unknowns, and so on. Thus in any period t one can compute the total number of matches ego has formed with unknowns by dividing n_t from equation (5) by p. As shown in equation (6), with $\delta = 0$ n_t reaches its maximum, steady state value at z^*. Hence ego can never have matched with more than z^*/p unknowns. Since all agents are initially unknown to ego, the condition $z^*/p \leq N$ is sufficient to ensure that ego never runs out of unknowns with whom to match.

Now we consider the main case $\delta > 0$, so that agents enter the firm in cohorts. We will show that the candidate sufficient condition $z^*/p \leq N$ is too weak when small δ is combined with large p, and derive an alternative sufficient condition. Recall that $P_t^c(c')$ is the probability that a given alter in cohort c' is in the network of a given ego in cohort c at the end of period t, conditional on the alter remaining with the firm. It follows that $P_t^c(c')/p$ is the probability that an ego in cohort c has met with a given alter in cohort c' by the end of period t, conditional on the alter remaining with the firm. Since $P_t^c(c')/p$ was derived by dividing the demand for matches by ego with a given unknown alter by the supply of matches by all unknown alters, we want a condition that ensures $P_t^c(c')/p \leq 1$.

Consider the polar case $p = 1$. Derivation of the sufficient condition is greatly simplified by Propositions 4 and 5, which imply $P_t^c(c)/p > P_t^c(c')/p$ for $c' \neq c$. (Proposition 5 itself is conditioned on "z^*/N small enough," but it can be shown that for $p = 1$ this condition is satisfied

if $z^*/N < 1$, which is equivalent to our weaker sufficient condition.) We therefore only need a condition sufficient for $P_t^c(c)/p \leq 1$.

When $p = 1$ every alter is of high match quality. Each agent therefore makes z^* matches with alters of high match quality in his entry period and δz^* replacement matches in every subsequent period. It follows that the number of matches with unknowns sought by every incumbent alter in the period ego enters the firm (and every subsequent period) is fixed at its minimum value δz^*. Using this information along with equations (7) and (8) allows us to compute $P_c^c(c)/p = z^*/[N(1 - \delta)\delta + \delta N] = z^*/\delta N[1 + (1 - \delta)]$. Setting $z^* = \delta N[1 + (1 - \delta)]$ therefore yields $P_c^c(c)/p = 1$. From equation (7) we see that $P_c^c(c)/p = 1$ implies $P_t^c(c)/p = 1$ for all $t > c$. We also see from equation (7) that $P_t^c(c)/p$ is monotonically increasing in z^* for all $t \geq c$. Hence $z^* \leq \delta N[1 + (1 - \delta)]$ is sufficient to ensure $P_t^c(c')/p \leq 1$ for all $t \geq c$ when $\delta > 0$ and $p = 1$. Finally, inspection of equations (7) and (8) shows that reducing p below 1 decreases $P_c^c(c)/p$ since the numerator remains unchanged and the denominator increases, after which $P_t^c(c)/p$ cannot rise above 1 for $t > c$ provided $z^* \leq \delta N[1 + (1 - \delta)]$.

We can now summarize. For $p \in (0, 1)$ and $\delta = 0$, the condition $z^* \leq pN$ is sufficient to ensure that ego never runs out of unknowns with whom to match. For $p \in (0, 1)$ and $\delta > 0$, the condition $z^* \leq pN$ may be too weak if $\delta N[1 + (1 - \delta)] < pN$ or $\delta < 1 - \sqrt{1 - p}$, in which case we use $z^* \leq \delta N[1 + (1 - \delta)]$ as our condition sufficient to ensure that ego never runs out of unknowns with whom to match.

We must translate these conditions on z^* into the primitives of our model. We assume that the cost of matching rises sufficiently fast that the desired number of matches z^* falls below the levels in the relevant inequalities:

Assumption A. *For $\delta \in (0, 1 - \sqrt{1 - p})$, we assume* $c'(\delta N[1 + (1 - \delta)]) \geq \frac{py_H + (1-p)y_L}{2} + \beta(1 - \delta)p\frac{(1-\delta)(1-p)}{1 - \beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2}$. *For $\delta = 0$ or $\delta \geq 1 - \sqrt{1 - p})$, we assume* $c'(pN) \geq \frac{py_H + (1-p)y_L}{2} + \beta(1 - \delta)p\frac{(1-\delta)(1-p)}{1 - \beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2}$.

The right-hand sides of the inequalities in Assumption A equal $c'(z^*)$ by equation (3), and therefore imply the inequalities $\delta N[1 + (1 - \delta)] \geq z^*$ and $pN \geq z^*$, respectively.

B Derivations and Proofs

Derivation of equation (3).

From equation (2) in the text, the first-order condition yields

$$\frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)V'(n_t)p = c'(x_t^* + (1-\delta)n_{t-1}).$$

Note that

$$V'(n_{t-1}) = (1-\delta)\frac{y_H}{2} - (1-\delta)c'(x_t^* + (1-\delta)n_{t-1}) + \beta(1-\delta)V'(n_t)(1-\delta)$$
$$+ \left[\frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)V'(n_t)p - c'(x_t^* + (1-\delta)n_{t-1})\right]\frac{\partial x_t^*}{\partial n_{t-1}}.$$

The coefficient on $\frac{\partial x_t^*}{\partial n_{t-1}}$ equals zero by the first order condition. We also use the first-order condition to substitute for $c'(x_t^* + (1-\delta)n_{t-1})$, obtaining

$$V'(n_{t-1}) = (1-\delta)\frac{y_H}{2} - (1-\delta)\left[\frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)pV'(n_t)\right] + \beta(1-\delta)V'(n_t)(1-\delta)$$
$$= (1-\delta)(1-p)\frac{y_H - y_L}{2} + \beta(1-\delta)^2(1-p)V'(n_t).$$

This is a linear difference equation for $V'(n_t)$, which admits a constant solution

$$V'(n_{t-1}) = V'(n_t) = \frac{(1-\delta)(1-p)}{1 - \beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2}.$$

The constant solution is the only solution that satisfies the transversality condition.[11] We can substitute it back into the first-order condition to obtain

$$\frac{py_H + (1-p)y_L}{2} + \beta(1-\delta)p\frac{(1-\delta)(1-p)}{1 - \beta(1-\delta)^2(1-p)}\frac{y_H - y_L}{2} = c'(x_t + (1-\delta)n_{t-1}) \equiv c'(z^*),$$

which is equation (3) in the text.

Proposition 3. *Assume $p > \frac{\delta}{(1-\delta)}$. For $\delta > 0$, there exists a $\underline{t}' \geq 1$ such that $HD_t(t')$ is monotonically decreasing in t' for $t' < \underline{t}'$ and monotonically increasing in t' for $t' > \underline{t}'$. Moreover, there exists a $\underline{t} > \underline{t}'$ such that $HD_t(0) > HD_t(t)$ for $t < \underline{t}$ and $HD_t(0) < HD_t(t)$ for $t > \underline{t}$.*

---

[11]We can show that $V'(n_t)$ grows at rate $[\beta(1-\delta)^2(1-p)]^{-1} > 1$ unless it is constant. But by the transversality condition, $[\beta(1-\delta)]^t V'(n_t)$ must be bounded, and since $\beta(1-\delta) \times [\beta(1-\delta)^2(1-p)]^{-1} = [(1-\delta)(1-p)]^{-1} > 1$, this is impossible. Hence the only possibility is $V'(n) = $ constant.

Proof. Substituting $n_{t'-1}$ and n_t into the definition of $HD_t(t')$:

$$HD_t(t') = \frac{(1-\delta)^{t-t'}p[z^* - (1-\delta)\sum_{\tau=0}^{t'-1}(1-p)^\tau(1-\delta)^\tau p z^*]}{\sum_{\tau=0}^{t}(1-p)^\tau(1-\delta)^\tau p z^*}$$

$$HD_t(t') = \frac{(1-\delta)^{t-t'}[1 - p(1-\delta)\sum_{\tau=0}^{t'-1}(1-p)^\tau(1-\delta)^\tau]}{\sum_{\tau=0}^{t}(1-p)^\tau(1-\delta)^\tau}$$

$$HD_t(t') = \frac{(1-\delta)^{t-t'}[1 - p(1-\delta)\frac{1-(1-p)^{t'}(1-\delta)^{t'}}{1-(1-p)(1-\delta)}]}{\frac{1-(1-p)^{t+1}(1-\delta)^{t+1}}{1-(1-p)(1-\delta)}}$$

$$HD_t(t') = \frac{(1-\delta)^{t-t'}[\delta + p(1-\delta) - p(1-\delta)[1 - (1-p)^{t'}(1-\delta)^{t'}]]}{1 - (1-p)^{t+1}(1-\delta)^{t+1}}$$

$$HD_t(t') = \underbrace{\frac{(1-\delta)^t}{1 - (1-p)^{t+1}(1-\delta)^{t+1}}}_{A}[(1-\delta)^{-t'}\delta + (1-\delta)p(1-p)^{t'}] \tag{B.1}$$

Inspection of equation (B.1) shows that $HD_t(t')$ is increasing in t' for t' sufficiently large. Straightforward computation shows that $HD_t(0) > HD_t(1)$ given $p > \frac{\delta}{(1-\delta)}$. Moreover, if we treat t' as continuous and differentiate $HD_t(t')$ twice with respect to t', we obtain $A[(1-\delta)^{-t'}\delta[\ln(1-\delta)]^2 + (1-\delta)p(1-p)^{t'}[\ln(1-p)]^2] > 0$. Thus $HD_t(t')$ is strictly convex in continuous t' and has a global minimum, and in discrete time reaches a minimum for some $\underline{t}' \geq 1$.

It follows from the first part of the proposition that $HD_t(0) > HD_t(t)$ for $t < \underline{t}'$, hence $\underline{t} > \underline{t}'$. Next, we can use equation (B.1) to show that the inequality $HD_t(0) < HD_t(t)$ reduces to $\delta + (1-\delta)p < (1-\delta)^{-t}\delta + (1-\delta)p(1-p)^t$. From the first part of the proposition the right-hand side of this inequality is monotonically increasing in t for $t > \underline{t}'$. Since, in fact, the right-hand side of this inequality increases without bound in t, the existence of a $\underline{t}$ as described in the second part of the proposition follows. $\square$

Proposition 4. $P_t^c(c) > P_t^c(c')$ *for all* $c' \in [0, c-1]$, *for all* $t \geq c$.

Proof. We will show that $P_t^c(c) - P_t^c(c') > 0$ for $c' < c$, $t \geq c$. The proof proceeds by induction. For the base case $t = c$, from equation (7) we have $P_c^c(c) - P_c^c(c') = px_c^c(x_c^c - x_c^{c'})/L_c^c > 0$, because $x_c^c > x_c^{c'}$ for $c' < c$ by equation (5). For the inductive step for period $t > c$, use equation (7) to compute $P_t^c(c) - P_t^c(c') > 0$ as

$$[P_{t-1}^c(c) - P_{t-1}^c(c')] + [1 - \frac{P_{t-1}^c(c)}{p}]px_t^c\frac{x_t^c}{L_t^c} - [1 - \frac{P_{t-1}^c(c')}{p}]px_t^c\frac{x_t^{c'}}{L_t^c}.$$

Let $k_c = \frac{P_{t-1}^c(c)}{p}$, $k_{c'} = \frac{P_{t-1}^c(c')}{p}$, $m_c = x_t^c \frac{x_t^c}{L_t^c}$ and $m_{c'} = x_t^c \frac{x_t^{c'}}{L_t^c}$. Note that $0 < k_{c'} < k_c < 1$ by the inductive hypothesis. Additionally, $0 < m_{c'} < m_c < 1$ because $x_t^c > x_t^{c'}$ for $c' < c$, $t \geq c$ by equation (5). Substituting $k_c, k_{c'}, m_c$ and $m_{c'}$ into the expression above, we have

$$p\{(k_c - k_{c'}) + [(1 - k_c)m_c - (1 - k_{c'})m_{c'}]\} = p\{(1 - m_c)(k_c - k_{c'}) + (1 - k_{c'})(m_c - m_{c'})\} > 0.$$

$\square$

Lemma 1. *If z^*/N is sufficiently small, $x_t^c/L_t^c > x_{t+b}^c/L_{t+b}^c$ for all $b \in [1, T - c]$, for all $t \geq c$.*

Proof. Rearrange the inequality as $\frac{L_{t+b}^c}{L_t^c} > \frac{x_{t+b}^c}{x_t^c}$. From equation (8), we have

$$\frac{L_{t+b}^c}{L_t^c} = \frac{N(1 - \delta)^{t+b}[1 - \frac{P_{t+b-1}^c(0)}{p}]x_{t+b}^0 + \sum_{c'=1}^{t+b} \delta N(1 - \delta)^{t+b-c'}[1 - \frac{P_{t+b-1}^c(c')}{p}]x_{t+b}^{c'}}{N(1 - \delta)^t[1 - \frac{P_{t-1}^c(0)}{p}]x_t^0 + \sum_{c'=1}^t \delta N(1 - \delta)^{t-c'}[1 - \frac{P_{t-1}^c(c')}{p}]x_t^{c'}}.$$

Note that N factors out of both the numerator and denominator of this expression. Likewise, z^* factors out of every x, hence z^* factors out of both the numerator and denominator. Changes in N or z^* therefore affect $\frac{L_{t+b}^c}{L_t^c}$ only through the terms containing P. From equation (7), we see that we can make any P arbitrarily small by shrinking z^*/N.

Letting z^*/N approach zero, we have

$$\frac{L_{t+b}^c}{L_t^c} \approx \frac{N(1 - \delta)^{t+b}x_{t+b}^0 + \sum_{c'=1}^{t+b} \delta N(1 - \delta)^{t+b-c'}x_{t+b}^{c'}}{N(1 - \delta)^t x_t^0 + \sum_{c'=1}^t \delta N(1 - \delta)^{t-c'}x_t^{c'}}$$

$$= \frac{N(1 - \delta)^{t+b}x_{t+b}^0 + \sum_{a=1}^b \delta N(1 - \delta)^{t+b-a}x_{t+b}^a + \sum_{c'=1+b}^{t+b} \delta N(1 - \delta)^{t+b-c'}x_{t+b}^{c'}}{N(1 - \delta)^t x_t^0 + \sum_{c'=1}^t \delta N(1 - \delta)^{t-c'}x_t^{c'}}.$$

We will show that this last expression is greater than $\frac{x_{t+b}^c}{x_t^c}$, from which it follows that $\frac{L_{t+b}^c}{L_t^c} > \frac{x_{t+b}^c}{x_t^c}$ for z^*/N sufficiently small.

Note that the last term in the numerator of the expression equals the last term in the denominator of the expression because $x_{t+\tau}^{c'} = x_t^{c'-\tau}$. From this fact and $\frac{x_{t+b}^c}{x_t^c} < 1$, it follows that a sufficient

condition for the expression to be greater than $\frac{x^c_{t+b}}{x^c_t}$ is

$$N(1-\delta)^{t+b}x^0_{t+b} + \sum_{a=1}^{b}\delta N(1-\delta)^{t+b-a}x^a_{t+b} - N(1-\delta)^t x^0_t \frac{x^c_{t+b}}{x^c_t} \geq 0,$$

$$\text{or}(1-\delta)^b x^0_{t+b} + \sum_{a=1}^{b}\delta(1-\delta)^{b-a}x^a_{t+b} - x^0_t \frac{x^c_{t+b}}{x^c_t} \geq 0.$$

Using the fact that $(1-\delta)^b = 1 - \sum_{a=1}^{b}\delta(1-\delta)^{b-a}$, and dividing through by x^0_t, the sufficient condition becomes

$$\frac{x^0_{t+b}}{x^0_t} + \sum_{a=1}^{b}\delta(1-\delta)^{b-a}\frac{x^a_{t+b} - x^0_{t+b}}{x^0_t} - \frac{x^c_{t+b}}{x^c_t} \geq 0.$$

Note that by example 1 of Proposition 3, the left-hand side goes to zero as δ goes to zero. For $\delta > 0$, it follows from equation (5) that $\frac{x^0_{t+b}}{x^0_t} > \frac{x^c_{t+b}}{x^c_t}$, hence the sufficient condition holds. $\qquad\square$

Proposition 5. *If z^*/N is sufficiently small, $P^c_t(c) > P^c_t(c')$ for all $c' \in [c+1, T]$, for all $t \geq c'$.*

Proof. We will show by induction that $P^c_t(c) - P^c_t(c+b) \geq \sum_{\tau=0}^{b-1}[1 - \frac{P^c_{t-\tau-1}(c)}{p}]px^c_{t-\tau}\frac{x^c_{t-\tau}}{L^c_{t-\tau}} > 0$, $b = 1, ..., T-c$ and $t = c+b, ..., T$. We first establish the base case

$$P^c_{c+b}(c) - P^c_{c+b}(c+b) \geq \sum_{\tau=0}^{b-1}[1 - \frac{P^c_{c+b-\tau-1}(c)}{p}]px^c_{c+b-\tau}\frac{x^c_{c+b-\tau}}{L^c_{c+b-\tau}}.$$

From repeated applications of equation (7), we have

$$P^c_{c+b}(c) = P^c_c(c) + \sum_{\tau=0}^{b-1}[1 - \frac{P^c_{c+b-\tau-1}(c)}{p}]px^c_{c+b-\tau}\frac{x^c_{c+b-\tau}}{L^c_{c+b-\tau}}$$

$$P^c_{c+b}(c+b) = px^c_{c+b}\frac{x^{c+b}_{c+b}}{L^c_{c+b}}.$$

Hence, we must show $P^c_c(c) - px^c_{c+b}\frac{x^{c+b}_{c+b}}{L^c_{c+b}} \geq 0$, or $px^c_c\frac{x^c_c}{L^c_c} - px^c_{c+b}\frac{x^{c+b}_{c+b}}{L^c_{c+b}} \geq 0$. Since $x^{c+b}_{c+b} = x^c_c$, this reduces to $\frac{x^c_c}{L^c_c} \geq \frac{x^c_{c+b}}{L^c_{c+b}}$ or $\frac{L^c_{c+b}}{L^c_c} \geq \frac{x^c_{c+b}}{x^c_c}$, which follows from Lemma 1.

Now consider the inductive step

$$P^c_t(c) - P^c_t(c+b) \geq \sum_{\tau=0}^{b-1}[1 - \frac{P^c_{t-\tau-1}(c)}{p}]px^c_{t-\tau}\frac{x^c_{t-\tau}}{L^c_{t-\tau}} > 0.$$

Application of equation (7), and then of the inductive hypothesis, yields

$$P_t^c(c) - P_t^c(c+b) = P_{t-1}^c(c) - P_{t-1}^c(c+b) - [1 - \frac{P_{t-1}^c(c+b)}{p}]px_t^c\frac{x_t^{c+b}}{L_t^c} + [1 - \frac{P_{t-1}^c(c)}{p}]px_t^c\frac{x_t^c}{L_t^c}$$

$$\geq \sum_{\tau=0}^{b-1}[1 - \frac{P_{t-\tau-2}^c(c)}{p}]px_{t-\tau-1}^c\frac{x_{t-\tau-1}^c}{L_{t-\tau-1}^c} - [1 - \frac{P_{t-1}^c(c+b)}{p}]px_t^c\frac{x_t^{c+b}}{L_t^c} + [1 - \frac{P_{t-1}^c(c)}{p}]px_t^c\frac{x_t^c}{L_t^c}$$

$$= \sum_{\tau=0}^{b-1}[1 - \frac{P_{t-\tau-1}^c(c)}{p}]px_{t-\tau}^c\frac{x_{t-\tau}^c}{L_{t-\tau}^c} + [1 - \frac{P_{t-b-1}^c(c)}{p}]px_{t-b}^c\frac{x_{t-b}^c}{L_{t-b}^c} - [1 - \frac{P_{t-1}^c(c+b)}{p}]px_t^c\frac{x_t^{c+b}}{L_t^c}.$$

Therefore we must prove

$$[1 - \frac{P_{t-b-1}^c(c)}{p}]px_{t-b}^c\frac{x_{t-b}^c}{L_{t-b}^c} - [1 - \frac{P_{t-1}^c(c+b)}{p}]px_t^c\frac{x_t^{c+b}}{L_t^c} \geq 0.$$

Noting that $x_{t-b}^c = x_t^{c+b}$, this reduces to

$$\frac{[1 - \frac{P_{t-b-1}^c(c)}{p}]L_t^c}{[1 - \frac{P_{t-1}^c(c+b)}{p}]L_{t-b}^c} \geq \frac{x_t^c}{x_{t-b}^c}.$$

Letting z^*/N become arbitrarily small, we can make $\frac{1 - \frac{P_{t-b-1}^c(c)}{p}}{1 - \frac{P_{t-1}^c(c+b)}{p}}$ arbitrarily close to one. Since $\frac{L_t^c}{L_{t-b}^c} > \frac{x_t^c}{x_{t-b}^c}$ by Lemma 1, the result follows. □

Proposition 6. *Consider a firm of age $T \geq 2$. If z^*/N is sufficiently small and $\delta[1+\delta p/2(1-p)] < 1/2$, then $S_t^c(c')$ reaches its maximum over cohorts $c' \in [1,T]$ for cohort c for at least the two most recent cohorts, i.e., $c \in [T-1, T]$.*

Proof. For $c = T$, the proposition follows from Proposition 4 and equation (9). For $c = T - 1$, the proposition follows from Proposition 4 and equation (9) for cohorts $c' \in [1, T-1]$. It remains to be shown that $S_T^{T-1}(T-1) > S_T^{T-1}(T)$. We have $\frac{S_T^{T-1}(T-1)}{S_T^{T-1}(T)} = \frac{P_T^{T-1}(T-1)}{P_T^{T-1}(T)}(1-\delta)$, so we need

$\frac{P_T^{T-1}(T-1)}{P_T^{T-1}(T)} > \frac{1}{1-\delta}$. Using equation (7), we have

$$\frac{P_T^{T-1}(T-1)}{P_T^{T-1}(T)} = \frac{P_{T-1}^{T-1}(T-1) + [1 - \frac{P_{T-1}^{T-1}(T-1)}{p}]px_T^{T-1}x_T^{T-1}/L_T^{T-1}}{P_T^{T-1}(T)}$$

$$= \frac{px_{T-1}^{T-1}x_{T-1}^{T-1}/L_{T-1}^{T-1}}{px_T^{T-1}x_T^T/L_T^{T-1}} + \frac{[1 - \frac{P_{T-1}^{T-1}(T-1)}{p}]px_T^{T-1}x_T^{T-1}/L_T^{T-1}}{px_T^{T-1}x_T^T/L_T^{T-1}}$$

$$= \frac{x_{T-1}^{T-1}L_T^{T-1}}{x_T^{T-1}L_{T-1}^{T-1}} + [1 - \frac{P_{T-1}^{T-1}(T-1)}{p}]\frac{x_T^{T-1}}{x_T^T}$$

$$= \frac{x_{T-1}^{T-1}L_T^{T-1}}{x_T^{T-1}L_{T-1}^{T-1}} + [1 - \frac{P_{T-1}^{T-1}(T-1)}{p}][1 - (1-\delta)p],$$

where we have used equation (5). From Lemma 1, it follows that $\frac{P_T^{T-1}(T-1)}{P_T^{T-1}(T)} \geq 1 + 1 - (1-\delta)p$ for z^*/N sufficiently small. Algebra (available on request) then shows that the condition $2 - (1-\delta)p > \frac{1}{1-\delta}$ reduces to $\delta[1 + \frac{\delta p}{2(1-p)}] < \frac{1}{2}$. $\square$

Proposition 7. *If z^*/N is sufficiently small and $T - c$ is sufficiently large, $S_T^{c'}(T) > S_T^{c'}(c')$ for all cohorts $c' \in [1, c]$.*

Proof. Using equation (9), we can reduce the inequality for cohort c, $S_T^c(T) > S_T^c(c)$, to $P_T^c(T) > P_T^c(c)(1-\delta)^{T-c}$. The right-hand side of this inequality can be made arbitrarily small for $T - c$ sufficiently large because $P_T^c(c)$ is bounded from above by p. On the left-hand side of the inequality, we use equation (7) to obtain $P_T^c(T) = px_T^c\frac{x_T^T}{L_T^c} = px_T^c\frac{z^*}{L_T^c}$. x_T^c is bounded from below by $\bar{x}$. L_T^c is bounded from above by Nz^*, the maximal number of desired meetings with unknowns by all agents in the firm. Thus $P_T^c(T)$ is bounded from below by $p\bar{x}\frac{z^*}{Nz^*} = p\bar{x}/N$.

This establishes the proposition for cohort c. Intuitively, the proposition should extend to cohorts $c' \in [1, c-1]$: the number of agents remaining in each of these cohorts is smaller than for cohort c, hence the own-cohort share of ego's network should be even smaller than for cohort c. (Note that if we were to weaken Assumption 2 by decoupling firm hiring from layoffs, it would be possible for earlier cohorts to be larger than later cohorts so we could not extend Proposition 7 to all cohorts $c' \in [1, c-1]$.) We will show that $S_T^c(T) > S_T^c(c)$ implies $S_T^{c-1}(T) > S_T^{c-1}(c-1)$ for $T - c$ sufficiently large and z^*/N sufficiently small. The proof can then be repeated to show $S_T^{c-1}(T) > S_T^{c-1}(c-1)$ implies $S_T^{c-2}(T) > S_T^{c-2}(c-2)$, ... , $S_T^2(T) > S_T^2(2)$ implies $S_T^1(T) > S_T^1(1)$.

From equation (9) we see that $S_T^c(T) > S_T^c(c)$ implies $P_T^c(T) > P_T^c(c)(1 - \delta)^{T-c}$. We will show that this implies $P_T^{c-1}(T) > P_T^{c-1}(c - 1)(1 - \delta)^{T+1-c}$, from which $S_T^{c-1}(T) > S_T^{c-1}(c - 1)$ follows from equation (9). Specifically, we will show that $P_T^{c-1}(T)$ can be made arbitrarily close to $P_T^c(T)$ and that $P_T^{c-1}(c - 1) < P_T^c(c)$ or $P_T^{c-1}(c - 1)$ can be made arbitrarily close to $P_T^c(c)$. Since $P_T^c(T)/P_T^c(c) > (1 - \delta)^{T-c}$, it follows that $P_T^{c-1}(T)/P_T^{c-1}(c - 1) > (1 - \delta)(1 - \delta)^{T-c} = (1 - \delta)^{T+1-c}$.

From equation (7) we have $P_T^c(T) = px_T^c \frac{x_T^T}{L_T^c}$ and $P_T^{c-1}(T) = px_T^{c-1} \frac{x_T^T}{L_T^{c-1}}$, where

$$L_T^c = N(1 - \delta)^T \left[1 - \frac{P_{T-1}^c(0)}{p}\right] x_T^0 + \sum_{c'=1}^T \delta N(1 - \delta)^{T-c'} \left[1 - \frac{P_{T-1}^c(c')}{p}\right] x_T^{c'}$$

and

$$L_T^{c-1} = N(1 - \delta)^T \left[1 - \frac{P_{T-1}^{c-1}(0)}{p}\right] x_T^0 + \sum_{c'=1}^T \delta N(1 - \delta)^{T-c'} \left[1 - \frac{P_{T-1}^{c-1}(c')}{p}\right] x_T^{c'}$$

by equation (8). We then have $P_T^{c-1}(T)/P_T^c(T) = (x_T^{c-1}/x_T^c)(L_T^c/L_T^{c-1})$. From equation (5) we see that x_T^{c-1}/x_T^c becomes arbitrarily close to one as $T - c$ grows large. Now consider L_T^c/L_T^{c-1}. Note that N factors out of both the numerator and denominator of this ratio. Likewise, z^* factors out of every x, hence z^* factors out of both the numerator and denominator. Changes in N or z^* therefore affect L_T^c/L_T^{c-1} only through the terms containing P. From equation (7), we see that we can make any P arbitrarily small by shrinking z^*/N. It follows from our expressions for L_T^c and L_T^{c-1} that we can make the ratio L_T^c/L_T^{c-1} arbitrarily close to one by shrinking z^*/N.

Finally, consider the ratio $P_T^{c-1}(c - 1)/P_T^c(c)$. From equation (7), we can write

$$P_T^{c-1}(c - 1) = px_{c-1}^{c-1} \frac{x_{c-1}^{c-1}}{L_{c-1}^{c-1}} + \left[1 - \frac{P_{c-1}^{c-1}(c - 1)}{p}\right] px_c^{c-1} \frac{x_c^{c-1}}{L_c^{c-1}} + \left[1 - \frac{P_c^{c-1}(c - 1)}{p}\right] px_{c+1}^{c-1} \frac{x_{c+1}^{c-1}}{L_{c+1}^{c-1}} + \dots$$
$$+ \left[1 - \frac{P_{T-1}^{c-1}(c - 1)}{p}\right] px_T^{c-1} \frac{x_T^{c-1}}{L_T^{c-1}}$$

and

$$P_T^c(c) = px_c^c \frac{x_c^c}{L_c^c} + \left[1 - \frac{P_c^c(c)}{p}\right] px_{c+1}^c \frac{x_{c+1}^c}{L_{c+1}^c} + \left[1 - \frac{P_{c+1}^c(c)}{p}\right] px_{c+2}^c \frac{x_{c+2}^c}{L_{c+2}^c} + \dots$$
$$+ \left[1 - \frac{P_{T-1}^c(c)}{p}\right] px_T^c \frac{x_T^c}{L_T^c}.$$

37

Note that each term in the expression for $P_T^c(c)$ corresponds to the term in the expression for $P_T^{c-1}(c-1)$ for the preceding period. Consider the corresponding terms $[1-\frac{P_{c+b-2}^{c-1}(c-1)}{p}]px_{c+b-1}^{c-1}\frac{x_{c+b-1}^{c-1}}{L_{c+b-1}^{c-1}}$ and $[1-\frac{P_{c+b-1}^{c}(c)}{p}]px_{c+b}^{c}\frac{x_{c+b}^{c}}{L_{c+b}^{c}}$, $b = 1,\ldots,T-c$. Note that $x_{c+b-1}^{c-1} = x_{c+b}^{c}$. Also note that $(z^*)^2/z^*N = z^*/N$ factors out of every term, so that changes in N or z^* can only affect $P_T^{c-1}(c-1)/P_T^c(c)$ through $P_{c+b-2}^{c-1}(c-1)$ and $P_{c+b-1}^{c}(c)$. As above, we can make $P_{c+b-2}^{c-1}(c-1)$ and $P_{c+b-1}^{c}(c)$ arbitrarily small by shrinking z^*/N. The difference between the numerators of the corresponding terms can therefore be made arbitrarily small.

Now consider the difference between the denominators of the corresponding terms. From equation (8), we have

$$L_{c+b-1}^{c-1} = N(1-\delta)^{c+b-1}\left[1 - \frac{P_{c+b-2}^{c-1}(0)}{p}\right]x_{c+b-1}^{0} + \sum_{c'=1}^{c+b-1}\delta N(1-\delta)^{c+b-1-c'}\left[1 - \frac{P_{c+b-2}^{c-1}(c')}{p}\right]x_{c+b-1}^{c'}$$

$$L_{c+b}^{c} = N(1-\delta)^{c+b}\left[1 - \frac{P_{c+b-1}^{c}(0)}{p}\right]x_{c+b}^{0} + \sum_{c'=1}^{c+b}\delta N(1-\delta)^{c+b-c'}\left[1 - \frac{P_{c+b-1}^{c}(c')}{p}\right]x_{c+b}^{c'}.$$

By shrinking z^*/N, L_{c+b-1}^{c-1} and L_{c+b}^{c} can be made arbitrarily close to $\tilde{L}_{c+b-1}^{c-1}$ and $\tilde{L}_{c+b}^{c}$, respectively, which are given by

$$\tilde{L}_{c+b-1}^{c-1} = N(1-\delta)^{c+b-1}x_{c+b-1}^{0} + \sum_{c'=1}^{c+b-1}\delta N(1-\delta)^{c+b-1-c'}x_{c+b-1}^{c'}$$

and

$$\tilde{L}_{c+b}^{c} = N(1-\delta)^{c+b}x_{c+b}^{0} + \sum_{c'=1}^{c+b}\delta N(1-\delta)^{c+b-c'}x_{c+b}^{c'}.$$

We have

$$\tilde{L}_{c+b-1}^{c-1} - \tilde{L}_{c+b}^{c} = N(1-\delta)^{c+b-1}[x_{c+b-1}^{0} - (1-\delta)x_{c+b}^{0}] - \delta N(1-\delta)^{c+b-1}x_{c+b}^{1}$$
$$= N(1-\delta)^{c+b-1}[x_{c+b-1}^{0} - (1-\delta)x_{c+b}^{0} - \delta x_{c+b}^{1}] > 0$$

since

$$x_{c+b-1}^{0} - (1-\delta)x_{c+b}^{0} - \delta x_{c+b}^{1} = (1-\delta)(x_{c+b-1}^{0} - x_{c+b}^{0}) + \delta(x_{c+b-1}^{0} - x_{c+b}^{1})$$
$$= (1-\delta)(x_{c+b-1}^{0} - x_{c+b}^{0}) > 0.$$

We have now shown that, for z^*/N sufficiently small,

$$\left[1 - \frac{P_{c+b-2}^{c-1}(c-1)}{p}\right] px_{c+b-1}^{c-1} \frac{x_{c+b-1}^{c-1}}{L_{c+b-1}^{c-1}} < \left[1 - \frac{P_{c+b-1}^c(c)}{p}\right] px_{c+b}^c \frac{x_{c+b}^c}{L_{c+b}^c}, \ b = 1, \ldots, T - c.$$

The same algebra that proves $\tilde{L}_{c+b-1}^{c-1} - \tilde{L}_{c+b}^c > 0$ also proves $L_{c-1}^{c-1} - L_c^c > 0$, hence $px_{c-1}^{c-1}x_{c-1}^{c-1}/L_{c-1}^{c-1} < px_c^c x_c^c/L_c^c$. Thus for z^*/N sufficiently small, all corresponding terms in the expressions for $P_T^{c-1}(c-1)$ and $P_T^c(c)$ are smaller for the former. However, the expression for $P_T^{c-1}(c-1)$ also contains the additional term $[1 - \frac{P_{T-1}^{c-1}(c-1)}{p}]px_T^{c-1}x_T^{c-1}/L_T^{c-1}$. Suppose that the presence of this term makes the ratio $P_T^{c-1}(c-1)/P_T^c(c)$ greater than one. In this case, we can make $P_T^{c-1}(c-1)/P_T^c(c)$ arbitrarily close to one by increasing $T - c$ and thereby increasing the number of terms in the expressions for $P_T^{c-1}(c-1)$ and $P_T^c(c)$. $\qquad \square$

Derivation of equation (12).

From equation (10) and the equations of motion we obtain

$$\frac{\partial V}{\partial n_{t-1}} = (1 - \delta)\frac{y_H}{2} - (1 - \delta)c'(x_t + r_t + (1 - \delta)n_{t-1}) + \beta(1 - \delta)[(1 - \delta)\frac{\partial V}{\partial n_t} + \delta\frac{\partial V}{\partial g_t}] + \beta\delta^2 V_u',$$

where V_u' is the marginal value of adding contacts. We also have

$$\frac{\partial V}{\partial g_{t-1}} = \beta(1 - \delta)(1 - \psi)\frac{\partial V}{\partial g_t} + \beta\delta(1 - \psi)V_u'. \tag{B.2}$$

It is straightforward to show that if V_u' is constant then an affine solution $V(n_{t-1}, g_{t-1}) = \frac{\partial V}{\partial n}n_{t-1} + \frac{\partial V}{\partial g}g_{t-1} + V(0,0)$ exists where $\frac{\partial V}{\partial n}$ and $\frac{\partial V}{\partial g}$ are constants (hence $\frac{\partial V}{\partial n_t} = \frac{\partial V}{\partial n_{t-1}}$ and $\frac{\partial V}{\partial g_t} = \frac{\partial V}{\partial g_{t-1}}$). Moreover, since n_t and g_t are bounded this solution satisfies the transversality condition. We see that the equations above have a recursive structure. Given a constant V_u', we can solve for a constant $\frac{\partial V}{\partial g}$. Given constants V_u' and $\frac{\partial V}{\partial g}$, we can solve for a constant $\frac{\partial V}{\partial n}$. We thereby obtain

$$\frac{\partial V}{\partial n} = \frac{(1 - \delta)(1 - p)\frac{y_H - y_L}{2} + \beta\delta^2 V_u'[1 + \frac{\beta(1-\delta)(1-\psi)}{1-\beta(1-\delta)(1-\psi)}]}{1 - \beta(1 - \delta)^2(1 - p)},$$

which is equation (12) in the text.

Proposition 9. *If the marginal value of adding contacts V_u' is constant, x_t decreases monotonically*

with time if the firm's hiring of ego's contacts satisfies

$[1-(1-\delta)(1-p)]r_t - r_{t+1} < (1-\delta)^{t+1}(1-p)^t p\hat{z}^*$, *where $\hat{z}^*$ is determined by equation* (13).

Proof. Using the fact that $\hat{z}^* = x_t + r_t + (1-\delta)n_{t-1}$, we have $x_t = \hat{z}^* - (1-\delta)n_{t-1} - r_t$ and $x_{t+1} - x_t = -(1-\delta)(n_t - n_{t-1}) - (r_{t+1} - r_t)$. It can be shown that $n_t - n_{t-1} = (1-\delta)^t(1-p)^t p\hat{z}^* + (1-p)r_t$. Substituting, we have $x_{t+1} - x_t = -(1-\delta)^{t+1}(1-p)^t p\hat{z}^* + [1-(1-\delta)(1-p)]r_t - r_{t+1}$. Hence $x_{t+1} < x_t$ if $[1-(1-\delta)(1-p)]r_t - r_{t+1} < (1-\delta)^{t+1}(1-p)^t p\hat{z}^*$. $\qquad\square$

If $V(n,g)$ is defined by equation (10), $V_u'' < 0$ *implies $\partial^2 V/\partial n\partial g \leq 0$.*

Proof. [12] Since the presence of r_t has no bearing on whether V has the desired property, we drop it for simplicity. We further streamline notation by rewriting equation (10) as

$$V(n,g) = \max_x \{a(1-\delta)n + bx - c(x + (1-\delta)n) + \beta((1-\delta)V(n',g') + \delta V_u(g'))\},$$

where $a = \frac{y_H}{2}$, $b = \frac{py_H + (1-p)y_L}{2}$, and

$$n' = (1-\delta)n + px,$$
$$g' = (1-\psi)g + \delta n.$$

Letting TV be the right-hand side of this Bellman equation, the true value function satisfies $TV = V$, that is, it is a fixed point of T. It can be shown that n and g are bounded, so that $(n,g) \in [0,\bar{n}] \times [0,\bar{g}]$. Letting X be the space of all continuous functions on $[0,\bar{n}] \times [0,\bar{g}]$, it can be shown that $T : X \to X$ is a contraction. By the contraction mapping theorem, $TV = V$ has a unique solution, and $V = \lim_{m\to\infty} T^m V_0$, where V_0 is any function (typically chosen to be 0).

Now to show that $\partial^2 V/\partial n\partial g \leq 0$ for the true value function, it suffices to show that for any (concave) V, we have

$$\frac{\partial^2 V}{\partial n\partial g} \leq 0 \implies \frac{\partial^2 TV}{\partial n\partial g} \leq 0. \tag{B.3}$$

To see this, suppose we show (B.3). Let $V_0 = 0$ and $V_m = TV_{m-1}$ for all $m \geq 1$. Since the zero function satisfies $\partial^2 V/\partial n\partial g \leq 0$, it follows from induction that $\partial^2 V_m/\partial n\partial g \leq 0$. Letting $m \to \infty$, since V_m converges to the true value function V, we obtain $\partial^2 V/\partial n\partial g \leq 0$.

---

[12]We thank, without implicating, Alexis Toda for help with this proof.

Take any concave function V with $\partial^2 V / \partial n \partial g \leq 0$, and rewrite the first-order condition (11) as

$$0 = b - c'(x + (1-\delta)n) + \beta(1-\delta)p\frac{\partial V}{\partial n}(n', g'). \tag{B.4}$$

Let $x^*(n, g)$ be the solution of this equation. We can also rewrite equation (B.2) as

$$\frac{\partial TV}{\partial g} = \beta(1-\psi)\left((1-\delta)\frac{\partial V}{\partial g}(n', g') + \delta V_u'(g')\right),$$

where $n' = px^*(n, g) + (1-\delta)n$. Differentiating this expression with respect to n, we obtain

$$\frac{\partial^2 TV}{\partial n \partial g} = \beta(1-\psi)\left((1-\delta)\left(\frac{\partial^2 V}{\partial n \partial g}(n', g')(p\frac{\partial x^*}{\partial n} + 1 - \delta) + \delta\frac{\partial^2 V}{\partial g^2}(n', g')\right) + \delta^2 V_u''(g')\right).$$

For notational simplicity let $V_{11} \equiv \partial^2 V / \partial n^2$, $V_{12} \equiv \partial^2 V / \partial n \partial g$, and $V_{22} \equiv \partial^2 V / \partial g^2$. Then the above equation becomes

$$(TV)_{12} = \beta(1-\psi)((1-\delta)(V_{12}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{22}) + \delta^2 V_u''(g')).$$

Since $\beta > 0$, $0 < \psi < 1$, $0 < \delta < 1$, and $V_u'' < 0$, to show $(TV)_{12} \leq 0$, it suffices to show

$$V_{12}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{22} \leq 0. \tag{B.5}$$

We establish that (B.5) holds in four steps. First, we show that $\partial x^*/\partial n < 0$. Differentiating the first-order condition (B.4) with respect to n, we obtain

$$0 = -c''(\partial x^*/\partial n + 1 - \delta) + \beta(1-\delta)p(V_{11}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{12}), \tag{B.6}$$

$$\text{or } (c'' - \beta(1-\delta)p^2 V_{11})\partial x^*/\partial n = -(1-\delta)(c'' - \beta(1-\delta)pV_{11} - \beta\delta V_{12}).$$

Since $c'' > 0$ (increasing marginal cost), $V_{11} < 0$ (concavity),[13] and $V_{12} \leq 0$ (assumption), it follows that $\partial x^*/\partial n < 0$.

Second, we show that $\partial x^*/\partial n + 1 - \delta < 0$. Since we already know $\partial x^*/\partial n < 0$, it follows

---

[13]Suppose $V_{11} = 0$ (or $V_{22} = 0$). The determinant condition for concave functions $V_{11}V_{22} - V_{12}^2 \geq 0$ then implies $V_{12} = 0$, and (B.5) holds.

from $0 < p < 1$ that

$$\partial x^*/\partial n + 1 - \delta < p\partial x^*/\partial n + 1 - \delta.$$

Using this inequality together with (B.6), since $V_{11} < 0$ we obtain

$$
\begin{aligned}
0 &= -c''(\partial x^*/\partial n + 1 - \delta) + \beta(1 - \delta)p(V_{11}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{12}) \\
&< -c''(\partial x^*/\partial n + 1 - \delta) + \beta(1 - \delta)p(V_{11}(\partial x^*/\partial n + 1 - \delta) + \delta V_{12}), \text{ or} \\
&(c'' - \beta(1 - \delta)pV_{11})(\partial x^*/\partial n + 1 - \delta) < \beta(1 - \delta)p\delta V_{12}.
\end{aligned}
$$

Since $c'' > 0$, $V_{11} < 0$, and $V_{12} \leq 0$, we obtain $\partial x^*/\partial n + 1 - \delta < 0$.

Third, we show that $V_{11}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{12} < 0$. Since $c'' > 0$ and $\partial x^*/\partial n + 1 - \delta < 0$ by the previous step, it follows from (B.6) that

$$\beta(1 - \delta)p(V_{11}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{12}) = c''(\partial x^*/\partial n + 1 - \delta) < 0.$$

Therefore we obtain $V_{11}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{12} < 0$.

Finally, we rewrite the left-hand side of (B.5):

$$V_{12}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{22} = \frac{V_{12}}{V_{11}}(V_{11}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{12}) + \delta\frac{V_{11}V_{22} - V_{12}^2}{V_{11}}.$$

Since $V_{11} < 0$, $V_{12} \leq 0$ (assumption), and by the previous step $V_{11}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{12} < 0$, it follows that the first term is nonpositive. By the determinant condition for concave functions $V_{11}V_{22} - V_{12}^2 \geq 0$, hence the second term is also nonpositive. Therefore $V_{12}(p\partial x^*/\partial n + 1 - \delta) + \delta V_{22}$ is the sum of nonpositive terms, and (B.5) holds. $\qquad\square$