

# Social Networks and Unraveling in Labor Markets\*

Itay P. Fainmesser<sup>†</sup>

September 30, 2010

## Abstract

This paper studies the phenomenon of early hiring in entry-level labor markets (e.g. the market for gastroenterology fellowships and the market for judicial clerks) in the presence of social networks. We offer a two-stage model in which workers in training institutions reveal information on their own ability over time. In the early stage, workers receive a noisy signal about their own ability. The early information is ‘soft’ and non-verifiable, and workers can convey the information credibly only to firms that are connected to them (potentially via their mentors). At the second stage, ‘hard’ verifiable (and accurate) information is revealed to the workers and can be credibly transmitted to all firms. We characterize the effects of changes to the network structure on the unraveling of the market towards early hiring. Moreover, we show that an efficient design of the matching procedure can prevent unraveling. (JEL: A14, D85, C78, L14)

**Keywords:** Networks, market design, unraveling, entry-level labor markets, early hiring.

## 1 Introduction

The timing of transactions is an important part of a market’s activity. In entry-level labor markets, such as the market for judicial clerks or for medical interns, hiring a worker before sufficient information is revealed on her quality can lead to inefficient placement of workers. Nevertheless, law clerks in the US are often hired by judges as early as the fall of their first

---

\*I am grateful to Susan Athey, Drew Fudenberg, Markus Mobius and Al Roth for their advice, guidance, and support. I would like to thank Philippe Aghion, Attila Ambrus, Itai Ashlagi, Eric Budish, James Burns, Ulrich Doraszelski, Hanna Halaburda, Fuhito Kojima, Robin Lee, Sendhil Mullainathan, Muriel Niederle, Laura Serban, Adam Szeidl, Alexander Westkamp, Leeat Yariv, and the participants of many seminars and conferences for helpful comments and discussion. All remaining errors are my own.

<sup>†</sup>Economics Department, Brown University, Providence RI 02906. Email: *Itay\_Fainmesser@Brown.edu*

year in law school, medical interns were in some years hired as early as two years prior to their graduation, and NBA teams recently draft high school seniors, skipping the college stage entirely.

Roth and Xing (1994), document several markets that exhibited a process of *unraveling* towards earlier and earlier contracting dates when market participants repeatedly ‘jump the gun’. Unraveling is found to affect the outcomes of markets with respect to both distribution and welfare.<sup>1</sup> Recently, Niederle and Roth (2003) document that workers’ mobility decreased during the unraveling in the gastroenterology fellowships market that was triggered by the collapse of the central match. The observation that early hiring is ‘more local’ than late hiring suggests that locality is not merely driven by the preferences of the workers to stay in the same location as their training institutions, but rather that there is an inherent difference in the way hiring is conducted in different stages in the workers’ training.

The local nature of the hiring process is not surprising. Sociologists and economists have long recognized that many workers find their jobs through friends and relatives.<sup>2</sup> It is only natural that social networks affect an inherently connection-based phenomenon such as early hiring. Nevertheless, none of the earlier models of unraveling accounted for the underlying topology of markets that motivate the study of unraveling, whether it is based on geography or on personal connections.<sup>3</sup>

In this paper, we propose a model in which some firms and workers are *connected* - e.g. via personal connections of workers’ mentors. Our model consists of two stages in which workers are in training institutions and reveal information on their own ability over time. In the early stage, workers get a noisy signal about their own ability. The early information is ‘soft’ and non-verifiable. Thus, workers can convey the information credibly only to firms that are connected to them, potentially by learning the mentors’ impressions of the workers. At the second stage, ‘hard’ verifiable (and accurate) information is revealed to the workers and can be credibly transmitted to all of the firms. Firms that use their connections (and hire promising candidates early) dilute the pool of high quality workers in the second stage.

---

<sup>1</sup>For direct evidence, see Niederle and Roth (2003 and 2005) and Fréchette, Roth, and Ünver (2007).

<sup>2</sup>See also Granovetter (1974), Montgomery (1991), Bewley (1999), Calvo-Armengol and Jackson (2004), and Calvo-Armengol and Zenou (2005).

<sup>3</sup>For previous theoretical work on unraveling see also Li and Rosen (1998), Li and Suen (2000), Suen (2000), Damiano, Li, and Suen (2005), Halaburda (2010), and Ostrovsky and Schwarz (2010).

The externality imposed on other firms triggers a process of unraveling towards more and more firms using their connections and hiring early.

We model the pattern of connections between workers and firms as a two-sided (bipartite) network. A connection links a worker with a firm to which she is able to convey private information credibly at an early stage of her training.<sup>4</sup> Two-sided networks are especially adept to describing market interactions in which the roles are well defined as in many labor markets. By studying the effect of changes in the network structure on unraveling we provide a rigorous analysis of changes in information asymmetries in the market and their impact on unraveling. We find that differences in the patterns of connections can account for differences in market outcomes, including unraveling. In our comparative statics, we focus on two types of changes to the network structure: [1] changes that correspond to adding or deleting connections from the network; and [2] changes that correspond to changing the distributions of connections across workers and firms.

There are several ways of adding links to a network. One way is by increasing the *span* of the network - i.e. by increasing the number of workers and firms that have at least one connection. Another way is by increasing the network's *density* - i.e. by increasing the number of connections of workers and firms that have at least one connection (without changing the network's span). We find that increasing the network's density has a non monotonic effect on unraveling. In particular, if a network is *sufficiently dense* then any increase in density leads to lesser unraveling. An immediate implication is that a complete market in an early stage of the workers' training does not generate more unraveling than a networked market. Increasing the span of the network always generates greater unraveling.

We further characterize the effects of redistributing connections across workers and firms. If the distribution of the number of connections across firms is more *polarized* unraveling is greater (a distribution is more polarized if there is higher density at the tails of the distribution). This is consistent with suggestive evidence from the market for judicial clerks.<sup>5</sup>

---

<sup>4</sup>Different mechanisms can allow to convey private information credibly over personal connections, e.g. teachers or other recommenders may have personal connections with the hiring entity. Such are the connections between internal medicine departments and gastroenterology departments in the same hospital that potentially allow for transmission of information about interns graduating from internal medicine internships and looking for a gastroenterology fellowship. See also Fainmesser (2010), Fainmesser and Goldberg (2010), and Karlan et. al. (2009) for a theoretical analysis of credible information transmission in a network.

<sup>5</sup>In the market for judicial clerks, attempts to set the date of the hiring of clerks failed repeatedly. The

The opposite is true for the case that the distribution of the number of connections across workers is more polarized.

We are also interested to answer the following questions: what is the scope of market design in this networked environment? Can a better design of the post-graduation market prevent unraveling? Consistent with much of the evidence from the market design literature, we show that improving the underlying mechanism for matching workers and firms in labor markets leads to lesser unraveling. To this end, we offer a simple parameterization of the efficiency of a market procedure in our setup. We prove that in large markets in which high productivity workers are scarce, our parameterization is supported by a family of matching procedures that follow from activity rules in both centralized and decentralized markets. In particular, this family of matching procedures incorporates most algorithms that are studied in the market design literature (e.g. deferred acceptance, random dictator, top trading cycles, etc.).

This paper is also related to the literature on networks in economics. Calvo-Armengol and Jackson (2004), and Calvo-Armengol and Zenou (2005) study models of job search via personal connections and derive implications to inequality and unemployment. More broadly, there is a growing related literature on network games (e.g. Bramoulle, D’Amours, and Kranton 2010, Ballester, Calvo-Armengol, and Zenou 2006). In networks games each player cares only about the actions taken by her neighbors. Galeotti et. al. (2010) suggest that in network games the analysis is simplified if players are assumed to hold incomplete knowledge of the network structure.

This simplification cannot be directly applied to our setup because a firm cares not only about the actions taken by its neighbors, but also about the aggregate outcome in the market which depends on the actions of all of the firms and workers in the market as well as on the entire network structure. Nevertheless, we show that if there are many workers and many firms and if the network is formed with a sufficiently salient random component, the assumption that workers and firms have incomplete knowledge of the network structure can

---

9th Circuit (California) is often claimed to be the one to unravel first. The reasons for this are under debate: the 9th Circuit judges point out that the East Coast judges have a geographical advantage as they are close to more top-ranked law schools, suggesting that a close connection to candidates matters, and that it is difficult to prevent unraveling in the presence of a skewed distribution of the firms’ degrees. See also Avery et. al. (2001) and (2007).

simplify the analysis significantly. This is possible due to recent graph theoretic results by Fainmesser and Goldberg (2010) who study repeated games in large two-sided networks.

The large network approach leads to an analysis that has the flavor of a *mean-field* approximation that is often assumed to approximate discrete and stochastic processes by a continuous and deterministic process (see Jackson and Rogers 2007 for an example used in the analysis of network formation). In particular, we approximate the number of workers hired early via the network. However, as we are interested in equilibrium behavior, we take a more explicit approach that provides bounds on the quality of the approximation and allows us to derive the firms' and workers' best response correspondences.

Finally, our model is different both in approach and in predictions from earlier models of unraveling. In particular, previous contributions focus on the heterogeneity of firms and candidates with respect to quality (Li and Rosen 1998, Li and Suen 2000) and preferences (Halaburda 2010), and on the size of the applicant pool (Li and Rosen 1998). We share with previous models the insurance element driving the unraveling process, and an agreement that unraveling lowers the efficiency in markets.<sup>6</sup>

In the following three sections we lay out the model and define the best response correspondences for workers and firms. In section 5, we characterize the structure of the equilibria in our model and define a notion of equilibrium stability that captures the dynamic nature of unraveling. In particular, our equilibrium analysis relies on an important approximation result that is reviewed in section 5. In sections 6 and 7 we derive comparative statics on unraveling with respect to the network structure and the matching procedure and discuss the impact of unraveling on market outcomes. Section 8 offers concluding remarks.

## 2 A simple model of employment

There is a finite set of firms,  $F \equiv \{1, 2, \dots, n_f\}$ , and a finite set of workers,  $W \equiv \{1, 2, \dots, n_w\}$ . Each worker  $w \in W$  can work for at most one firm and each firm  $f \in F$  can employ at most one worker. A worker  $w$  is characterized by a productivity level  $q_w \in \{L, H\}$ . We assume

---

<sup>6</sup>Li and Suen (2000) suggest that firms that do not unravel suffer a welfare loss. We extend this claim to all firms in an ex-ante sense.

that production depends only on the workers' productivities and normalize wages to zero (assuming a fixed wage is consistent with the markets motivating this paper and can also be substituted, with some technical burden, with milder restrictions).<sup>7,8</sup>

The payoff of firm  $f$  from employing worker  $w$  is captured by (1). In the following section we introduce heterogeneity to firms' payoffs.

$$\pi_f(q_w) = \pi(q_w) = \begin{cases} \pi_H & \text{if } q_w = H \\ \pi_L & \text{if } q_w = L \end{cases} \quad (\text{where } \pi_H > \pi_L) \quad (1)$$

Workers have idiosyncratic preferences over firms. Specifically, let worker  $w$ 's utility be:

$$u_w(f) = 1 + v_{wf} \quad (2)$$

where for every  $w$  and  $f$ ,  $v_{wf} \in [-v, v]$  for some  $v \in (0, 1)$ . A firm that does not employ any worker, and a worker that is unemployed have a payoff of 0.

### 3 The hiring process

In this section, we describe the process of hiring in the market. We start by an overview of the two stages of the workers' training during which they can be hired. Later, we review separately the details of the hiring process within each stage.

There are two stages in the workers' training,  $\mathcal{S} = -1$  and  $\mathcal{S} = 0$ . At stage  $\mathcal{S} = -1$  workers are in training institutions (i.e. law school, medical school, internship programs, etc.) and cannot yet be employed. At  $\mathcal{S} = 0$ , workers graduate from their studies and are ready to be employed.

Before stage  $\mathcal{S} = -1$  nature assigns each worker with a productivity level  $q_w = H$  or

---

<sup>7</sup>Judicial clerks' wages are determined by federal law. Medical residents' wages are limited to a small number of wage steps. In particular, in the gastroenterology fellowships market, Niederle and Roth (2003,2005) do not find any effect of unraveling on wages.

<sup>8</sup>The discussion of the role of wages in the analysis of unraveling is not new. Models of matching markets can be analyzed using the assignment model (Koopmans and Beckmann 1957, Shapley and Shubik 1971) where wages are a part of the clearing mechanism, or using the marriage model (Gale and Shapley 1962) where wages are assumed out. Li and Suen (2000), who analyze unraveling using the assignment model in the context of college admissions, admit that "*our [assignment model] analysis applies with a greater force to assignment markets in which payments transfers are explicitly negotiated*".

$q_w = L$  with equal probability, and preferences  $\{v_{wf}\}_{f \in F}$  that are drawn from a distribution  $\mathcal{H}$  with mean 0 and positive density in every point in the support  $[-v, v]$  where  $v \in (0, 1)$ . The realizations of  $q_w$  and  $\{v_{wf}\}_{f \in F}$  are independent of each other and across workers. Workers and firms do not observe  $q_w$  and  $\{v_{wf}\}_{f \in F}$  but will learn about them over time as described below.

We now describe intuitively the two stages of the training and hiring. In stage  $\mathcal{S} = -1$ , each worker  $w$  learns her own preferences  $\left(\{v_{wf}\}_{f \in F}\right)$ .<sup>9</sup> In addition, worker  $w$  and her mentors receive a noisy signal of her productivity ( $s_w$ ). The noisy signal  $s_w$  consists of ‘soft’ information (in-class exam grades, reinforcements from teachers, etc.). In particular, there is no official document or public track record that allows worker  $w$  to prove that she received a given signal. However, the worker’s mentors have pre-existing connections with a subset of the firms that allow for the credible transmission of workers’ signals to these firms. If firm  $f$  and at least one of the mentors of worker  $w$  are connected,  $f$  learns  $s_w$  accurately at stage  $\mathcal{S} = -1$ . Since mentors do not have a strategic role in our model, we say that a worker  $w$  and a firm  $f$  are connected if one of the mentors of worker  $w$  is connected with firm  $f$ . Firm  $f$  is not able to learn  $v_{wf}$  for any worker  $w$ .<sup>10</sup>

At the end of stage  $\mathcal{S} = -1$ , each firm  $f$  can make an offer to at most one worker, and a worker that receives at least one offer can accept one of the offers or reject all. If a worker  $w$  accepts an offer from a firm  $f$ , both  $w$  and  $f$  leave the market and  $w$  works for  $f$  starting at stage  $\mathcal{S} = 0$ .

In stage  $\mathcal{S} = 0$ , worker  $w$  and her mentors receive verifiable information that indicates accurately the worker’s productivity ( $q_w$ ) (transcripts, publications, etc.). Thus,  $q_w$  can be credibly transmitted to all firms. We explore a family of matching procedures governing the labor market in stage  $\mathcal{S} = 0$ . If a worker  $w$  and a firm  $f$  are matched,  $w$  works for  $f$  starting

---

<sup>9</sup>Assuming that a worker learns her preferences for a firm only upon receiving an offer from that firm does not change the analysis.

<sup>10</sup>There may be several reasons why firm  $f$  is not able to learn  $v_{wf}$  for worker  $w$ . First, a worker may strategically transmit her preferences to her mentor. Second, mentors are likely to be able to misrepresent workers’ preferences without being penalized by firms in the future. The reason for the latter is unmodelled in this paper. Informally, mentors are less likely to be held accountable by firms for misrepresenting workers’ preferences for several reasons: [1] mentors are not expected to be able to elicit workers’ preferences; [2] information on preference is rich  $\left(\{v_{wf}\}_{f \in F}\right)$  and cannot be transmitted easily; and [3] the relation between  $v_{wf}$  and outcomes that are observable by the firm (acceptance or rejection of an offer) depend on several unobserved variables, e.g. which other offers the worker received.

at stage  $\mathcal{S} = 0$ .

### 3.1 Stage $\mathcal{S} = -1$ : early hiring

We now describe formally stage  $\mathcal{S} = -1$  of the workers training, including the information available to workers, the network between firms and workers, and the process of early hiring.

At stage  $\mathcal{S} = -1$ , information about own productivity and preferences is revealed to the workers in the following way: worker  $w$  observes a signal  $(s_w, \{v_{wf}\}_{f \in F})$  where  $\{v_{wf}\}_{f \in F}$  are the worker's preferences and  $s_w \in \{h, l\}$  is a noisy signal of the worker's productivity. If worker  $w$  has productivity  $q_w = H$  she receives a signal  $s_w = h$  with probability  $\alpha \in [\frac{1}{2}, 1]$  and a signal  $s_w = l$  with probability  $(1 - \alpha)$ . For ease of notation assume that if worker  $w$  has productivity  $q_w = L$  she has the reversed probabilities over her signal  $s_w$ . The realizations of the signals are independent across workers.

For each worker  $w$  there exists a set of firms  $N_w \subset F$  that can learn  $s_w$ . Denote by  $N_f \subset W$  the set of workers such that firm  $f$  can learn  $\{s_w\}_{w \in N_f}$ . Firms cannot learn  $v_{wf}$  for any worker  $w$ .

After learning  $\{s_w\}_{w \in N_f}$ , each firm can make at most one offer. Firm  $f$  can make an offer to worker  $w$  independent of whether  $w \in N_f$  or not. Each worker  $w$  can then choose to accept one offer or none. If firm  $f$  makes an offer to worker  $w$  and worker  $w$  accepts, both commit that after graduation (at stage  $\mathcal{S} = 0$ )  $w$  will be employed by  $f$ . The commitment is binding and both  $w$  and  $f$  exit the labor market.

We assume further that firm  $f$  incurs a cost  $c_f$  for hiring at stage  $\mathcal{S} = -1$ , and let  $c_f$  be drawn from a distribution with a continuous cumulative distribution function  $\mathcal{D}(\underline{c}, \bar{c})$ , independently across firms. There are two potential (unmodelled) sources for the cost: [1] early commitment of funding that leads to a reduction in the firm's liquidity,<sup>11</sup> and [2] repugnance - in some entry level labor markets, such as the gastroenterology fellowship market and the market for judicial clerkships, firms are publicly objecting to early hiring.

---

<sup>11</sup>During the economic downturn of 2008-2010, employment contracts of freshly minted MBA's were canceled at a (potentially reputational) cost to the hiring firms. In the absence of a global financial crisis a firm that suffers from financial difficulties might incur even a greater cost because the cancellation cannot be attributed to market-wide economic recession.

Thus, hiring early might lead to loss of goodwill.<sup>12</sup>

Finally, we assume that

$$\pi_L < 0; \text{ and } -\pi_L \geq \pi_H \tag{3}$$

Given that a worker’s ex-ante probability of having high productivity is  $\frac{1}{2}$ , condition (3) guarantees that a firm  $f$  does not hire a worker  $w$  if  $f$  has no (or negative) information about the productivity of  $w$ .

To summarize, the timeline of the early labor market in stage  $\mathcal{S} = -1$  is as follows:

1. Each worker  $w$  observes a noisy signal  $s_w$ . Each firm  $f \in N_w$  learns  $s_w$ .
2. Each firm  $f$  makes an offer to at most one worker  $w \in W \cup \{w_0\}$ . We say that a firm makes an offer to worker  $w_0$  if the firm does not make any offer.
3. Each worker  $w$  who received at least one offer decides whether to accept an offer or not. Worker  $w$  can accept at most one offer.
4. If firm  $f$  makes an offer to worker  $w$  and worker  $w$  accepts, both exit the labor market and worker  $w$  is employed by firm  $f$  starting at stage  $\mathcal{S} = 0$ .

**Remark 1** *Two assumptions deserve additional discussion. First, the assumption that each firm can make at most one offer is a simplification of the idea that the number of offers that each firm can make early is limited. In many markets, this results from the hiring norms and technology prior to graduation. The market for Gastroenterology fellowships provides an example. During the period before the match, departments try to hide from each other the fact that they are making early offers.<sup>13</sup> Making many early offers might expose a department’s early hiring attempts before the department is able to hire. As a result, departments might try*

---

<sup>12</sup>The discussion against practices of early hiring and exploding offers in the judicial clerkships market were especially fierce. For a controversial defense of the common practice of early exploding offers, see Kozinski (1991).

<sup>13</sup>This is demonstrated in the following E-mail sent by Debbie Proctor, the gastroenterologist who took the lead in reorganizing the match, to the economists assisting in redesigning the Gastroenterology fellowships market:

*“I’m answering 3-4 emails per day especially on this issue. ‘I want to make sure MY competition is in the match and that they don’t cheat.’ Well, this is another way of saying that if they cheat, then I will too!...Have you ever seen this before? The distrust amongst program directors? I find it hard to believe that we are unique. Maybe this is [a] social science phenomenon?”.* See also Niederle and Roth (2009).

to make few offers that are likely to be accepted. An additional example of a different flavor is provided by the market for new MBA graduates. At Harvard Business School, periods in which students are available for interviews are dictated by the school and firms are not allowed to invite students for interview during the semester. This puts a bound on the number of interviews a firm can make at every stage. Since an interview is required prior to making an offer, this restricts the number of early offers.

A second assumption is that an early agreement between a firm and a worker is binding and that an early offer expires if it is not accepted at stage  $\mathcal{S} = -1$ . Early offers that are open only for a short period of time are often called ‘exploding offers’ in the market design literature. In many markets that unravel exploding offers are prevalent (see Niederle and Roth 2009) and their acceptance is binding. For example, in the market for judicial clerks offers are sometimes open for less than 30 minutes,<sup>14</sup> and there is little evidence of law students that renege on an early acceptance of a judge’s offer.

### 3.1.1 Networks and information

If firm  $f$  is able to learn  $s_w$  ( $f \in N_w$  and  $w \in N_f$ ) we say that  $f$  and  $w$  are *connected*. We note that the sets of firms, workers, and connections (links) induce a network. We now describe the network structure, as well as firms’ and workers’ knowledge and beliefs with respect to the network structure.

We are mainly interested in large markets. It is by now widely accepted that in large networks: [1] the underlying process of network formation has a strong stochastic element, and [2] some aggregate characteristics of the network structure, such as the distribution of the numbers of connections, are usually consistent across networks and time. Thus, we assume that before stage  $\mathcal{S} = -1$ ,  $\{N_w\}_{w \in W}$  and  $\{N_f\}_{f \in F}$  are determined by a random process that is described below. Firms and workers know the random process of network formation, but do not have complete knowledge of the network. Instead, a worker  $w$  (firm  $f$ ) observes only  $N_w$  ( $N_f$ ). Both assumptions fit well the labor markets motivating this paper

---

<sup>14</sup>Avery et. al. (2007) cite a 2005 applicant for federal judicial clerkships: "I received the offer via voicemail while I was in flight to my second interview. The judge actually left three messages. First, to make the offer. Second, to tell me that I should respond soon. Third, to rescind the offer. It was a 35 minute flight."

and can be relaxed at the cost of added technical burden.<sup>15,16</sup>

We now describe formally workers' and firms' knowledge of the network: we capture the network of connections between workers and firms with a graph  $G \equiv \langle F, W, E \rangle$ , where  $E \subset F \times W$  is the set of connections (edges) between firms and workers. The *degree* of worker  $w$  (firm  $f$ ) is the number of connections of worker  $w$  (firm  $f$ ):

$$r_w = |N_w| \quad (r_f = |N_f|)$$

Let  $\theta_W(r)$  ( $\theta_F(r)$ ) be the fraction of workers (firms) in the market with degree  $r$  for  $r = 0, 1, 2, \dots, \infty$ . Note that specifying  $n_w$ ,  $\theta_W(r)$ , and  $\theta_F(r)$  implies a unique  $n_f$  that is consistent with a graph  $G$ . Thus, we omit  $n_f$  and let  $\mathcal{G}(n_w, \theta_W, \theta_F)$  be the set of networks consistent with  $(n_w, \theta_W, \theta_F)$ .<sup>17</sup>

We assume that before stage  $\mathcal{S} = -1$  the network is chosen from  $\mathcal{G}(n_w, \theta_W, \theta_F)$  uniformly at random (u.a.r.) and that worker  $w$  (firm  $f$ ) knows: [1] the number of workers and firms in the market  $n_w$  and  $n_f$ ; [2] the set  $N_w$  ( $N_f$ ); and [3]  $\theta_W$  and  $\theta_F$ . We also note that the Bayesian posterior of worker  $w$  (firm  $f$ ) with degree  $r$  puts identical probability on all networks in  $\mathcal{G}(n_w, \theta_W, \theta_F | r_w = r)$  ( $\mathcal{G}(n_w, \theta_W, \theta_F | r_f = r)$ ). We denote by  $G(n_w, \theta_W, \theta_F | \cdot)$  a member of  $\mathcal{G}(n_w, \theta_W, \theta_F | \cdot)$  that is chosen u.a.r.

Some firms and workers might have no connections ( $r = 0$ ). To describe changes to the network structure that do not involve such buyers and sellers we use a modified degree distribution  $P$ . Formally, let  $P(r, \theta_W)$  be the fraction of workers with degree  $r$  as a fraction of the workers that have positive degrees and note that

$$P(r, \theta_W) = \frac{\theta_W(r)}{1 - \theta_W(0)}$$

---

<sup>15</sup>In section 5.1 we offer a dynamic model in which the hiring process is repeated with subsequent graduating cohorts in subsequent hiring cycles and discuss further the motivation for the random process of the formation of the network.

<sup>16</sup>The idea that workers and firms have incomplete knowledge of the network fits well the labor markets motivating this paper. For example, a firm might know several mentors, but might not know the students of these mentors and their other mentors as these change across graduating cohorts and hiring cycles. Similarly worker  $w$  might have several mentors and know which firms they are connected to, but may not know how well connected each of these firms are.

<sup>17</sup>For any fixed  $n_w$ ,  $\theta_W$ , and  $\theta_F$  there exists an infinite strictly increasing sequence of integers  $\{n_w\}$  s.t.  $\mathcal{G}(n_w, \theta_W, \theta_F) \neq \emptyset$  (see Greenhill et. al. 2006). All statements should be read as holding only for  $n_w$  s.t. the aforementioned set is non-empty.

The definition immediately extend to  $P(r, \theta_F)$ .

### 3.1.2 Scarcity of high productivity workers

Many of the labor markets that motivate this paper are markets for highly skilled workers (e.g. lawyers, physicians, psychologists). In such markets, the number of high quality workers is usually smaller than the market saturation level.

**Definition 1** We say that  $\langle \theta_W, \theta_F, \alpha \rangle$  exhibits *scarcity of high productivity workers* if for every  $\langle G, \alpha \rangle$  such that  $G \in \mathcal{G}(n_w, \theta_W, \theta_F)$  there exists  $\eta > 1$  such that,  $n_f > \eta \cdot (1 - \frac{\alpha}{2}) \cdot n_w$ .

Definition 1 includes all markets in which there are not many more workers than open positions. In particular, we show later that Definition 1 guarantees that in markets that exhibit scarcity of high productivity workers, the number of high productivity workers that are still looking for employment at stage  $\mathcal{S} = 0$  is always smaller than the number of firms that are looking for a worker at stage  $\mathcal{S} = 0$ .

## 3.2 Stage $\mathcal{S} = 0$ : graduation

We now describe formally stage  $\mathcal{S} = 0$ , which we also call the post-graduation market. We define and parameterize a family of market mechanisms on which we focus.

At stage  $\mathcal{S} = 0$  workers graduate from their training and obtain a diploma and a track record that contain verifiable information that reveals their true qualities  $\{q_w\}_{w \in W}$ . At this stage,  $\{q_w\}_{w \in W}$  are common knowledge, but the preferences of any worker  $w$  ( $\{v_{wf}\}_{f \in F}$ ) are still her private information. In this environment, the network is obsolete and we are in a familiar setup of a one-to-one matching market.<sup>18</sup>

The market design literature shows that the outcome in a matching market depends heavily on the underlying market rules.<sup>19</sup> Since we are also interested in how changes in the post-graduation market affect early hiring, we consider a class of *matching procedures* that includes many of the matching procedures studied in the market design literature.

---

<sup>18</sup>Roth and Sotomayor (1990) provide a good introduction to matching theory.

<sup>19</sup>See also Roth (2002, 2008).

Intuitively, a matching procedure is a function from sets of workers and firms to a probability distribution over a set of matchings. We focus on matching procedures that are *anonymous* - i.e. take into considerations the workers' and firms preferences but not their identities; and that put positive probability only on *stable matchings* - i.e. matchings that guarantee that no firm and worker that are matched prefer to stay unmatched and no worker and firm that would like to be matched to each other remain unmatched.<sup>20</sup> The requirement that a matching procedure be anonymous excludes matching procedures in which there is an ad-hoc reason that some firms and workers are matched at stage  $\mathcal{S} = 0$ .

The formal definition of an anonymous matching procedure that guarantees stable matchings builds on definitions from matching theory and is deferred to the Appendix. Instead, we present now the main result of this section and discuss its implications for the modeling of stage  $\mathcal{S} = 0$ .

Let  $W_q^0$  be the set of workers with productivity  $q$  that reach  $\mathcal{S} = 0$  unmatched and let  $F^0$  be the set of firms that reach  $\mathcal{S} = 0$  unmatched. Given  $G$  and  $\alpha$ , and given a matching procedure  $M$  denote by  $E_{G,M,\alpha}[u_w|q]$  the expected utility of worker  $w \in W_q^0$  and denote by  $E_{G,M,\alpha}[\pi_f]$  the expected payoff of firm  $f \in F^0$ . For given  $W_q^0$  ( $q \in \{H, L\}$ ) and  $F^0$ , denote by  $E_{G,M,\alpha}[u_w|q, W_H^0, W_L^0, F^0]$  and  $E_{G,M,\alpha}[\pi_f|W_H^0, W_L^0, F^0]$  the corresponding conditional expectations.

Lemma 1 shows that the requirement that a matching procedure is anonymous and guarantees a stable matching pins down a unique expected payoff for all firms that participate in the post-graduation market. Moreover, the expected utility of high productivity workers who reach stage  $\mathcal{S} = 0$  unmatched is asymptotically independent of the hiring at stage  $\mathcal{S} = -1$ . Therefore, the expected utility of high productivity workers can be varied exogenously by the choice of the particular matching procedure without affecting the expected payoffs of firms.

**Lemma 1** *Let  $\langle \theta_W, \theta_F, \alpha \rangle$  exhibit scarcity of high productivity workers, and let  $\widehat{G}(\theta_W, \theta_F, n_w)$  be any network that is consistent with  $\theta_W, \theta_F$  and  $n_w$ . Assume further that no worker who*

---

<sup>20</sup>In the matching literature, stability of a matching also requires that no firm-worker pair prefers to be matched to each other rather than to the partner that they are assigned with. However, in our model, this requirement is obsolete because firms have identical preferences over workers.

receives a low signal in stage  $\mathcal{S} = -1$  is hired early (in stage  $\mathcal{S} = -1$ ) and let  $\widehat{W}_H^0, \widehat{W}_L^0, \widehat{F}^0$  be any  $W_H^0, W_L^0, F^0$  that are possible under these assumptions. Then,

1. Given any anonymous matching procedure  $M$  that guarantees a stable matching:

(a) for any firm  $f \in \widehat{F}^0$ ,  $E_{\widehat{G}(\theta_W, \theta_F, n_w), M, \alpha} [\pi_f | \widehat{W}_H^0, \widehat{W}_L^0, \widehat{F}^0] = \frac{|\widehat{W}_H^0|}{|\widehat{F}^0|} \cdot \pi_H$ .

(b) for any worker  $w \in \widehat{W}_L^0$ ,  $E_{\widehat{G}(\theta_W, \theta_F, n_w), M, \alpha} [u_w | L, \widehat{W}_H^0, \widehat{W}_L^0, \widehat{F}^0] = 0$ .

2. For every  $\phi \in [0, 1]$  and for every  $\xi > 0$  there exists an anonymous matching procedure  $M$  that guarantees a stable matching such that

$$\lim_{n_w \rightarrow \infty} \sup_w \left| E_{\widehat{G}(\theta_W, \theta_F, n_w), M, \alpha} [u_w | H, \widehat{W}_H^0, \widehat{W}_L^0, \widehat{F}^0] - (1 + \phi \cdot v) \right| < \xi$$

We show later that it is always true that no worker who receives a low signal in stage  $\mathcal{S} = -1$  is hired early (in stage  $\mathcal{S} = -1$ ). Thus, Lemma 1 applies throughout our analysis. While the result is of interest on its own, we only use Lemma 1 to motivate exogenous variations in  $\phi$ , and to establish that all of the market procedures that we focus on lead to identical  $E_{G, M, \alpha} [u_w | L]$  and  $E_{G, M, \alpha} [\pi_f | W_H^0, W_L^0, F^0]$ . The following definition offers a parameterization for the family of market procedures that our analysis covers.

**Definition 2** A matching procedure  $M$  is parameterized by  $\phi_M \in \mathbb{R}$  if:

[1]  $M$  is anonymous and guarantees a stable matching, and

[2] for all  $G \in \mathcal{G}(n_w, \theta_W, \theta_F)$  and for any  $W_H^0, W_L^0$ , and  $F^0$  that are possible under the assumption that no worker who receives a low signal in stage  $\mathcal{S} = -1$  is hired (in stage  $\mathcal{S} = -1$ ), the expected utility of a high productivity worker in the post-graduation market is asymptotically  $1 + \phi_M \cdot v$ . Namely,

$$\lim_{n_w \rightarrow \infty} \sup_w \left| E_{G, M, \alpha} [u_w | H, W_H^0, W_L^0, F^0] - (1 + \phi_M \cdot v) \right| = 0$$

We interpret  $\phi_M$  as a continuous measure of the efficiency of the market rules.

## 4 Bayesian equilibrium and $\varepsilon$ -equilibrium

We now define the notions of Bayesian equilibrium and  $\varepsilon$ -equilibrium in our setup. Note that non trivial strategic decisions are made only at stage  $\mathcal{S} = -1$ : firms decide who to make offers to, and workers decide which offers to accept.

A strategy of a firm  $f$  is a mapping  $\sigma_f : \mathbb{R}^+ \times \mathcal{P}(W) \times \mathcal{P}(W) \rightarrow \Delta(W \cup \{w_0\})$  where  $\mathcal{P}(W)$  is the set of all subsets of workers and  $\Delta(W \cup \{w_0\})$  is the set of all probability distributions on  $W \cup \{w_0\}$ . Thus,  $\sigma_f(c, N_f, \widetilde{W}_h)$  is a mixed strategy of firm  $f$  with a cost of hiring early  $c_f = c$ , set of workers connected to it  $N_f$ , and that learned that every worker  $w \in \widetilde{W}_h \subseteq N_f$  received a signal  $s_w = h$  (and that every worker  $w \in N_f \setminus \widetilde{W}_h$  received a signal  $s_w = l$ ). The firm's mixed strategy is over the workers to which the firm makes an offer at stage  $\mathcal{S} = -1$  ( $\sigma_f(\cdot) = w_0$  implies that firm  $f$  does not make an offer at stage  $\mathcal{S} = -1$ ). For brevity, when it is clear from the context we let  $\sigma_f(\widetilde{W}_h) = \sigma_f(c_f, N_f, \widetilde{W}_h)$ .

A family of firms' strategies that is especially natural in our context includes strategies in which firms ignore the *names* (or *labels*) of the workers and make their offers based only on the economically meaningful attributes of the workers. Formally,

**Definition 3** Consider a firm  $f$  with  $(c, N_f, \widetilde{W}_h)$ . We say that  $\sigma_f$  is a **label-free strategy** if  $\sigma_f(c, N_f, \widetilde{W}_h)$  assigns identical probabilities to any  $w$  and  $w'$  for whom at least ONE of the following holds: [1]  $w, w' \in \widetilde{W}_h$ ; [2]  $w, w' \in N_f \setminus \widetilde{W}_h$ ; [3]  $w, w' \in W \setminus N_f$ .

A strategy for worker  $w$  is a mapping  $\mu_w : \{h, l\} \times [-v, v]^{n_f} \times \mathcal{P}(F) \times \mathcal{P}(F) \rightarrow \Delta(F \cup \{f_0\})$ . Thus,  $\mu_w(s_w, \{v_{wf}\}_{f \in F}, N_w, \widetilde{F}_w)$  is a mixed strategy of worker  $w$  that receives a signal  $s_w$ , has preferences that are captured by  $\{v_{wf}\}_{f \in F}$ , a set of firms connected to her  $N_w$ , and offers from every firm  $f \in \widetilde{F}_w$  at stage  $\mathcal{S} = -1$ . For brevity, when it is clear from the context we let  $\mu_w(\widetilde{F}_w) = \mu_w(s_w, \{v_{wf}\}_{f \in F}, N_w, \widetilde{F}_w)$ .

For given network  $G$ , market procedure  $M$ , and signal accuracy  $\alpha$ , let  $\Pi_{G,M,\alpha}(f) = \Pi_{G,M,\alpha}(f, \sigma_f, c_f, N_f, \{\sigma_{f'}\}_{f' \in F \setminus \{f\}}, \{\mu_{w'}\}_{w' \in W})$  be the expected payoff of firm  $f$  that employs strategy  $\sigma_f$ . Similarly, Let  $U_{G,M,\alpha}(w) = U_{G,M,\alpha}(w, \mu_w, \{v_{wf}\}_{f \in F}, \{\mu_{w'}\}_{w' \in W \setminus \{w\}}, \{\sigma_{f'}\}_{f' \in F})$  be the expected utility of worker  $w$  who employs strategy  $\mu_w$ . We are now ready to define equilibrium and  $\varepsilon$ -equilibrium in our setup.

**Definition 4** The vectors of strategies  $\{\sigma_f\}_{f \in F}$  and  $\{\mu_w\}_{w \in W}$  are an  $\varepsilon$ -equilibrium if for all  $\tilde{\sigma}_f \in \text{supp}(\sigma_f)$ ,  $f \in F$ :

$$\begin{aligned} & \Pi_{G(n_w, \theta_W, \theta_F), M, \alpha} \left( f, \sigma_f, c_f, N_f, \{\sigma_{f'}\}_{f' \in F \setminus \{f\}}, \{\mu_{w'}\}_{w' \in W} \right) \geq \\ & \Pi_{G(n_w, \theta_W, \theta_F), M, \alpha} \left( f, \tilde{\sigma}_f, c_f, N_f, \{\sigma_{f'}\}_{f' \in F \setminus \{f\}}, \{\mu_{w'}\}_{w' \in W} \right) - \varepsilon \end{aligned}$$

and for all  $\tilde{\mu}_w \in \text{supp}(\mu_w)$ ,  $w \in W$ :

$$\begin{aligned} & U_{G(n_w, \theta_W, \theta_F), M, \alpha} \left( w, \mu_w, \{v_{wf}\}_{f \in F}, \{\mu_{w'}\}_{w' \in W \setminus \{w\}}, \{\sigma_{f'}\}_{f' \in F} \right) \geq \\ & U_{G(n_w, \theta_W, \theta_F), M, \alpha} \left( w, \tilde{\mu}_w, \{v_{wf}\}_{f \in F}, \{\mu_{w'}\}_{w' \in W \setminus \{w\}}, \{\sigma_{f'}\}_{f' \in F} \right) - \varepsilon \end{aligned}$$

If  $\varepsilon = 0$  the definition amounts to a Bayesian equilibrium in our setup. We now analyze firms and workers best response functions separately and show that they can be summarized using two random variables.

If a firm  $f$  makes an offer at stage  $\mathcal{S} = -1$  to a worker  $w$  that receives a signal  $s_w = h$ :

$$\begin{aligned} & \Pi_{G, M, \alpha} \left( f | \widetilde{W}_h, \sigma_f \left( \widetilde{W}_h \right) = w \text{ for some } w \in W \right) = \\ & \quad Pr \{w \text{ accepts}\} \cdot \left\{ Pr \left[ q_w = H | N_f, \widetilde{W}_h \right] \cdot \pi_H + Pr \left[ q_w = L | N_f, \widetilde{W}_h \right] \cdot \pi_L - c_f \right\} + \\ & \quad + Pr \{w \text{ rejects}\} \cdot E_{G, M, \alpha} [\pi_f] \end{aligned} \tag{4}$$

and if firm  $f$  does not make an offer at stage  $\mathcal{S} = -1$ :

$$\Pi_{G, M, \alpha} \left( \sigma_f | \widetilde{W}_h, \sigma_f \left( \widetilde{W}_h \right) = w_0 \right) = E_{G, M, \alpha} [\pi_f]$$

Note that  $Pr \{w \text{ accepts}\}$ ,  $Pr \{w \text{ rejects}\}$ , and  $E_{G, M, \alpha} [\pi_f]$  depend on  $\{\sigma_{f'}\}_{f' \in F \setminus \{f\}}$  and  $\{\mu_{w'}\}_{w' \in W}$ . On the other hand,  $Pr \left[ q_w = H | N_f, \widetilde{W}_h \right]$ , and  $Pr \left[ q_w = L | N_f, \widetilde{W}_h \right]$  are independent of the strategies employed by all firms and workers.

Now consider a worker  $w$  that receives early job offers (stage  $\mathcal{S} = -1$ ) from a set of firms  $\tilde{F}_w \subseteq F$ . If  $w$  accepts the offer of firm  $f \in \tilde{F}_w$  then

$$U_{G, M, \alpha} \left( \cdot | \tilde{F}_w, \mu_w \left( \cdot, \tilde{F}_w \right) = f \right) = 1 + v_{wf}$$

and otherwise

$$U_{G, M, \alpha} \left( \cdot | \tilde{F}_w, \mu_w \left( \cdot, \tilde{F}_w \right) = f_0 \right) = Pr [q_w = H | s_w] \cdot E_{G, M, \alpha} [u_w | H] + Pr [q_w = L | s_w] \cdot E_{G, M, \alpha} [u_w | L]$$

Note that: [1] conditional on  $\widetilde{W}_h$  ( $\widetilde{F}_w$ ) the best response of firm  $f$  (worker  $w$ ) depends on the network structure and on the strategies of all other firms and workers only via  $E_{G,M,\alpha}[\pi_f]$  ( $E_{G,M,\alpha}[u_w|L]$ ); and [2] conditional on the number of firms and of high productivity workers that arrive stage  $\mathcal{S} = 0$  unmatched,  $E_{G,M,\alpha}[\pi_f]$  and  $E_{G,M,\alpha}[u_w|L]$  are independent of the network structure. Combining observations [1] and [2] we conclude that equilibrium behavior depends on the network structure only through its effect on the number of firms and of high productivity workers that arrive at stage  $\mathcal{S} = 0$  unmatched.

Formally, let  $\gamma$  be the fraction of workers hired at stage  $\mathcal{S} = -1$ . Then both of the following hold:

1. Consider a firm  $f$ . Conditional on  $\widetilde{W}_h \neq \emptyset$  and on the expectation that  $\gamma = \gamma_1$ , the probability that  $f$  makes an offer at stage  $\mathcal{S} = -1$  depends on  $c_f$  and is captured by

$$\begin{aligned} \sigma_{\theta_W, \theta_F, M, \alpha}(\gamma_1) &= \sigma_{G(\theta_W, \theta_F, n_w), M, \alpha}\left(r, \widetilde{W}_h, \gamma_1\right) = \\ &= \mathcal{D}\left(\alpha \cdot \pi_H + (1 - \alpha) \cdot \pi_L - \frac{\frac{1}{2} - \alpha \cdot \gamma_1}{1 - \gamma_1} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H\right) \end{aligned} \quad (5)$$

2. Consider a worker  $w$ . Conditional on  $\gamma = \gamma_1$  and on  $w$  receiving *exactly* one job offer at stage  $\mathcal{S} - 1$ , the probability that  $w$  accepts the offer depends on  $\{v_{wf}\}_{f \in F}$  and is captured by

$$\mu_{G, M, \alpha}(\gamma) = \mu_{G(n_w, \theta_W, \theta_F), M, \alpha}\left(r_w, \widetilde{F}_w, \gamma\right) = 1 - \mathcal{H}\left(\alpha \cdot E_{G(n_w, \theta_W, \theta_F), M, \alpha}[u_w|H, \gamma_1] - 1\right) \quad (6)$$

Similarly, a worker that received exactly  $m$  offers at stage  $\mathcal{S} = -1$  accepts an offer with probability  $1 - (1 - \mu_{G, M, \alpha}(\gamma))^m$ , and conditional on accepting an offer, worker  $w$  picks the offer that maximizes  $\{1 + v_{wf}\}_{f \in \widetilde{F}_w}$ .

Expression 5 is surprising because it establishes that in any equilibrium, for any given firm  $f$ , conditional on  $\gamma_1$  and as long as  $\widetilde{W}_h \neq \emptyset$ ,  $\sigma_{G(\theta_W, \theta_F, n_w), M, \alpha}\left(r, \widetilde{W}_h, \gamma_1\right)$  is independent of  $r$ ,  $\widetilde{W}_h$ , and  $n_w$  and depends on the network structure only via  $\theta_W, \theta_F$ . Condition 5 also captures the observation that the incentives of firms to make offers at stage  $\mathcal{S} = -1$  increase in the fraction of workers hired at stage  $\mathcal{S} = -1$ . A complete derivation of (5) and (6) is

deferred to the Appendix.

## 5 Equilibrium existence and structure

In this section we show that equilibrium in our setup corresponds to a fixed point of a mapping from the fraction of workers hired at stage  $\mathcal{S} = -1$  to itself, and find that in large networks,  $\varepsilon$ -equilibria exist for arbitrary low  $\varepsilon$ . Moreover, when firms employ label-free strategies, the set of equilibria is fully characterized as the set of fixed points of a simple function.

Let  $\widehat{\gamma}_0$  be a random variable that describes the (common and rational) expectations of workers and firms with respect to the fraction of workers hired at stage  $\mathcal{S} = -1$ . Consider a mapping

$$\gamma_0 = \psi_{G,M,\alpha}(\widehat{\gamma}_0) = \psi_{G,M,\alpha}(\sigma_{G,M,\alpha}(\widehat{\gamma}_0), \mu_{G,M,\alpha}(\widehat{\gamma}_0))$$

that maps from the expectations of workers and firms with respect to hiring at stage  $\mathcal{S} = -1$  to the random variable that captures the same outcome in stage  $\mathcal{S} = -1$ . Any  $\gamma'_0$  such that  $\gamma'_0 = \psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma'_0)$  captures an equilibrium level of hiring at stage  $\mathcal{S} = -1$ , and any equilibrium with  $\gamma_0^*$  corresponds to a fixed point  $\gamma_0^* = \psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma_0^*)$ . However, for any network  $G$ ,  $\psi_{G,M,\alpha}(\gamma)$  is complicated as it depends on the entire network structure and its outcome is stochastic even conditional on the network structure. In particular, for any network with more than a few workers and firms, a closed form expression for  $\psi_{G,M,\alpha}(\gamma)$  is prohibitively complicated.

Instead of trying to characterize  $\psi_{G,M,\alpha}(\gamma)$  directly, we establish that in (asymptotically) large networks and for every  $\langle \theta_W, \theta_F, M, \alpha \rangle$ ,  $\psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma_0)$  converges to a well behaved function with a deterministic output. As a result, we are able to characterize  $\varepsilon$ -equilibria for arbitrarily small  $\varepsilon$ .

We now illustrate our analysis using a simple exercise: suppose that there was no correlation between the degrees of firms and workers that are connected - i.e. if we choose a worker  $w \in W$  u.a.r. and then choose a firm  $f \in N_w$  u.a.r., then the probability that  $r_f = r$

is independent of  $r_w$  and captured by

$$\tilde{P}(r, \theta_F) = \frac{P(r, \theta_F) \cdot r}{\bar{r}_f}$$

where  $\bar{r}_f = E_P[r_f | r_f \geq 1] = \sum_{r \in \{1, 2, \dots, \infty\}} P(r, \theta_F) \cdot r$ . Now consider a worker  $w$  chosen u.a.r. The probability that worker  $w$  receives a job offer from a firm  $f$  that is chosen u.a.r. from  $N_w$  is

$$\begin{aligned} \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) &= \sum_{r_f=1}^{\infty} \tilde{P}_F(r_f, \theta_F) \cdot \left[ \sigma_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot \sum_{m=0, \dots, r_f-1} \binom{r_f-1}{m} 0.5^m 0.5^{r_f-m-1} \cdot \frac{1}{m+1} \right] \\ &= \sum_{r_f=1}^{\infty} \tilde{P}(r_f, \theta_F) \cdot [\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - 0.5^{r_f}) / (0.5 \cdot r_f)] \end{aligned}$$

To see why, note that  $\binom{r-1}{m} 0.5^m 0.5^{r-m-1}$  is the probability that there are  $m$  other  $s_w = h$  workers in  $N_f$  conditional on  $|N_f| = r$ . Finally,  $\frac{1}{m+1}$  is the conditional probability that  $f$  makes the offer to  $w$ . Given that the realizations of the signals, the offers received and the acceptance of offers are independent across workers, if the matching procedure  $M$  is parameterized by  $\phi_M$ , the expected number of workers hired in stage  $\mathcal{S} = -1$  is captured by

$$\begin{aligned} \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma) &= \\ \frac{1}{2} \cdot (1 - \theta_W(0)) \cdot \sum_{r_w=1}^{\infty} P(r_w, \theta_W) \cdot (1 - [1 - \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) + \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - \tilde{\mu}_{\theta_W, \theta_F, M, \alpha})]^{r_w}) \end{aligned} \tag{7}$$

where

$$\tilde{\mu}_{\theta_W, \theta_F, M, \alpha} = 1 - \mathcal{H}[\alpha \cdot (1 + \phi_M \cdot v) - 1]$$

The calculation of  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  above follows a naive counting exercise. Namely, it is equivalent to going over all of the workers, one by one, and evaluating their probabilities of receiving at least one acceptable early offer. Note that  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  is deterministic and

well behaved. Thus, establishing more generally that for any  $\gamma \in [0, 1]$  and  $\epsilon > 0$ ,

$$\lim_{n_w \rightarrow \infty} \Pr \left( \left| \psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\hat{\gamma}) - \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\hat{\gamma}) \right| < \epsilon \right) = 1 \quad (8)$$

allows us to characterize the equilibrium structure in large networks. Formally,

**Definition 5** *we say that  $\gamma^* \in [0, 1]$  is a **0-equilibrium in large networks** (or simply **0-equilibrium**) with  $\langle \theta_W, \theta_F, M, \alpha \rangle$  if for every  $\epsilon > 0$  there exists  $n_w \in \mathbb{Z}^+$  such that for every  $n'_w > n_w$  there exists an  $\epsilon$ -equilibrium with  $G(n_w, \theta_W, \theta_F), M, \alpha$  in which  $\Pr \left( \psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma^*) \in [\gamma^* - \epsilon, \gamma^* + \epsilon] \right) > 1 - \epsilon$ .*

Theorem 1 characterizes the set of 0-equilibria as the set of solutions for a fixed point problem in  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$ .

**Theorem 1** *Let  $\theta_W, \theta_F$  have finite support and  $\langle \theta_W, \theta_F, \alpha \rangle$  exhibit scarcity of high productivity workers. Consider a market procedure  $M$  that is parameterized by  $\phi_M \in [0, 1]$ . Then, there exists  $\gamma \in [0, 1]$  such that  $\gamma$  is a 0-equilibrium with  $\langle \theta_W, \theta_F, M, \alpha \rangle$ .*

*Assume further that firms employ label-free strategies. Then,  $\gamma^* = \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma^*)$  if and only if  $\gamma^*$  is a 0-equilibrium with  $\langle \theta_W, \theta_F, M, \alpha \rangle$ .*

In Lemma 4, which is deferred to the Appendix, we derive a limit closed form expression for  $\psi_{G, M, \alpha}(\hat{\gamma})$  without formally expressing  $\psi_{G, M, \alpha}(\hat{\gamma})$  for any finite network  $G$ . To this end, we rely on a recent graph theoretic result by Fainmesser and Goldberg (2010) that implies that in a network that is chosen u.a.r. conditional on a degree distribution, as the network grows, the degree correlation goes to zero. We then apply the law of large numbers to conclude that the fraction of worker hired at stage  $\mathcal{S} = -1$  converges to the mean and that (8) holds.

For the remainder of the paper, we focus on the analysis of 0-equilibria in *regular* environments in which Theorem 1 and Lemma 4 apply.

**Definition 6** *An environment  $\langle \theta_W, \theta_F, M, \alpha \rangle$  is **regular** if: [1]  $\theta_W, \theta_F$  have finite support; [2]  $\langle \theta_W, \theta_F, \alpha \rangle$  exhibit scarcity of high productivity workers; and [3]  $M$  is parameterized by some  $\phi_M \in [0, 1]$ .*

**Multiplicity.** Theorem 1 does not rule out multiplicity of 0-equilibria. In particular, multiplicity is determined by the properties of  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  which in turn correspond to the properties of  $\theta_W$ ,  $\theta_F$ ,  $M$ ,  $\alpha$ , and  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma)$ . For example, if  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma)$  is concave for every  $\gamma \in [0, 1]$ , then there are at most three equilibria, one at  $\gamma^* = 0$  and one or two additional equilibria.

## 5.1 Unraveling

In this section we define unraveling as a dynamic process in which firms and workers that hire early (at stage  $\mathcal{S} = -1$ ) trigger a response from other firms and workers, and the fraction of workers hired early increases. We also provide a definition of *greater unraveling* that is adept for environments with multiple equilibria.

Consider an entry level labor market in which new cohorts of workers graduate in different hiring cycles (e.g. years) and firms hire new workers in every hiring cycle. For simplicity, assume that all cohorts of workers are of the same size  $n_w$ . The hiring process in each hiring cycle follows stages  $\mathcal{S} = -1, 0$  that are described above. The network of connection is re-drawn in each period u.a.r. from all of the possible networks with the same  $\theta_W$  and  $\theta_F$ . To motivate the changes to the network structure, recall that the network is defined by the connections of the different subsets of mentors that each worker has. As the set of workers varies across periods, so does the network. However, as we are motivated by large markets, we follow the observation that large networks tend to maintain some aggregate characteristics constant even as the network undergo local changes. The assumptions that the network is selected u.a.r. and that  $\theta_W$  and  $\theta_F$  stay exactly the same can be relaxed.

In this dynamic environment, consider an unraveling process governed by myopic best responses. At hiring cycle  $t = 0$ , an arbitrary fraction  $\gamma^0$  of the workers (all with  $s_w = h$ ) is hired at stage  $\mathcal{S} = -1$ . At each hiring cycle  $t > 0$ , each agent best responds to the outcome of the play in hiring cycle  $t - 1$ . Let  $\gamma^t$  denote the fraction of the workers (all with  $s_w = h$ ) that are hired at stage  $\mathcal{S} = -1$  in hiring cycle  $t$ , then

$$\gamma^t = \psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma^{t-1})$$

and any rest point of the system correspond to a static Bayesian equilibrium. Following (8) and Theorem 1, the limiting dynamic process (when  $n_w$  is arbitrary large) is approximated by

$$\gamma^t = \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma^{t-1})$$

and any rest point of the system correspond to a static 0-equilibrium. Note that  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}$  has a positive slope for every  $\gamma \in [0, 1]$ . Hence, from any starting point, the convergence of the limiting dynamic process is monotone, either upwards or downwards.

A dynamic process of unraveling, in which the market participants modify their strategies based on the previous hiring cycle's outcomes, captures the dynamics in some well studied labor markets (see Roth and Xing 1994 and references therein) and suggestive evidence from the experimental market design literature (see also Kagel and Roth 2000). Moreover, considering the process of unraveling as dynamic lends itself to a natural way of capturing the notion that some markets generate *greater unraveling* than other markets.

**Definition 7** Let  $\gamma^t(n_w, \langle \theta_W, \theta_F, M, \alpha \rangle, \gamma^0) = \psi_{G(n_w, \theta_W, \theta_F), M, \alpha} \circ \psi_{G(n_w, \theta_W, \theta_F), M, \alpha}^{t-1}(\gamma^0)$ . We say that an environment  $\langle \cdot \rangle^1 = \langle \theta_W^1, \theta_F^1, M^1, \alpha^1 \rangle$ , **generates greater unraveling** than another environment  $\langle \cdot \rangle^2 = \langle \theta_W^2, \theta_F^2, M^2, \alpha^2 \rangle$ , if for any  $\gamma^0 \in [0, 1]$

$$\lim_{n_w \rightarrow \infty} \lim_{t \rightarrow \infty} [\gamma^t(n_w, \langle \cdot \rangle^1, \gamma^0) - \gamma^t(n_w, \langle \cdot \rangle^2, \gamma^0)] \geq 0 \quad (9)$$

Definition 7 captures the idea that an environment  $\langle \cdot \rangle^1$  generates greater unraveling than  $\langle \cdot \rangle^2$  if from every starting point environment  $\langle \cdot \rangle^1$  leads to a 0-equilibrium with more hiring at stage  $\mathcal{S} = -1$  than environment  $\langle \cdot \rangle^2$ . The following Corollary is implied by Lemma 4 and establishes a useful connection between  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}$  and the unraveling generated by an environment  $\langle \theta_W, \theta_F, M, \alpha \rangle$ .

**Corollary 1** Consider two regular environments  $\langle \cdot \rangle^1 = \langle \theta_W^1, \theta_F^1, M^1, \alpha^1 \rangle$  and  $\langle \cdot \rangle^2 = \langle \theta_W^2, \theta_F^2, M^2, \alpha^2 \rangle$ , and assume that all firms employ label-free strategies. Then, if  $\tilde{\psi}_{\theta_W^1, \theta_F^1, M^1, \alpha^1}(\gamma) \geq \tilde{\psi}_{\theta_W^2, \theta_F^2, M^2, \alpha^2}(\gamma)$  for every  $\gamma \in [0, 1]$ , then  $\langle \cdot \rangle^1$  generates greater unraveling than  $\langle \cdot \rangle^2$ .

## 6 Comparative statics

### 6.1 The network structure

We now investigate how changes to the network structure affect unraveling. We consider changes that correspond to addition/deletion of links or to redistribution of links.

**Addition of links.** There are several ways in which links can be added to a network. Let  $W^1 \subseteq W$  ( $F^1 \subseteq F$ ) be the set of workers (firms) that have a degree of at least 1. One way of adding links is by increasing the degrees of workers in  $W^1$  (firms in  $F^1$ ) so that the number of workers (firms) that have a degree of at least 1 does not change. We call such an addition of links an increase in the network's *density*. A different way for adding links involves changes to  $W^1$  ( $F^1$ ). In particular, one can add links that connect workers (firms) that were not connected before and had a degree of zero. We call such an addition of links an increase in the network's *span*. It turns out that increasing a network's density and increasing a network's span have significantly different effects on unraveling.

Consider degree distributions  $\theta_W, \theta_F$ . We denote by  $\theta_W^\rho, \theta_F^\rho$  the degree distributions such that for every  $r$ ,  $\theta_W^\rho(\rho \cdot r) = \theta_W(r)$ , and  $\theta_F^\rho(\rho \cdot r) = \theta_F(r)$ . We say that  $\rho$  is the *density multiplier* of  $\theta_W, \theta_F$ .<sup>21</sup> Proposition 1 implies that increasing the networks' density leads to *greater* unraveling when the initial network is sparse and the increase is small. When the initial network is dense, increasing the network's density leads to *lesser* unraveling.

**Proposition 1** *Let  $\langle \theta_W, \theta_F, M, \alpha \rangle$  be a regular environment and let all firms employ label-free strategies. Consider  $\rho^H > \rho^L \geq 1$ . Then, there exists  $\bar{r}(M, \alpha), \underline{r}(M, \alpha) \in \mathbb{Z}^+$  such that*

1. *if  $\max \{ \rho^H \cdot r \mid \theta_F(r) > 0 \} < \underline{r}$  then  $\langle \theta_W^{\rho^H}, \theta_F^{\rho^H}, M, \alpha \rangle$  generates greater unraveling than  $\langle \theta_W^{\rho^L}, \theta_F^{\rho^L}, M, \alpha \rangle$ , and*
2. *if  $\min \{ \rho^L \cdot r \mid r \geq 1, \theta_F(r) > 0 \} > \bar{r}$ , then  $\langle \theta_W^{\rho^L}, \theta_F^{\rho^L}, M, \alpha \rangle$  generates greater unraveling than  $\langle \theta_W^{\rho^H}, \theta_F^{\rho^H}, M, \alpha \rangle$ .*

---

<sup>21</sup>Our claims apply to  $\rho$  such that  $\rho \cdot r \in \mathbb{Z}^+$  for every  $r$  in the support of  $\theta_F$  and  $\theta_W$ .

Proposition 1 is surprising as it establishes that unraveling is *not* maximized when the network is very dense or in the well studied complete market. In particular, Proposition 1 highlights the role of the network in the unraveling process. In markets in which early information diffusion is not based on personal connections we would expect lower levels of unraveling than in some markets in which connections are important.

The forces that generate the non monotonicity suggested by Proposition 1 are demonstrated in figure 1.



Figure 1: Consider two markets with two firms and two workers as depicted in figure 1a and 1b. Assume that one of the firms has a very low cost of hiring early and that the other firm has a very high cost of hiring early. Assume further that only one of the workers receives a signal  $s_w = h$ . Then, in the network in figure 1b the firm with the low costs of hiring will make an early offer to the high signal worker, whereas in figure 1a there is a positive probability that the high signal worker is not connected to the firm with the low early hiring costs, and therefore no early offers will be made. Now assume that both firms have very low costs of hiring early, and that both workers receive a signal  $s_w = h$ . Then the network in figure 1a guarantees that both workers will receive early offers, whereas in the network in figure 1b there is a positive probability that both firms make an early offer to the same worker.

More generally, increasing the network's density increases the probability that a worker  $w$  with  $s_w = h$  and a firm with low enough  $c_f$  are connected. This leads to an increase in the number of offers made in stage  $\mathcal{S} = -1$ , as captured by the increase in  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  for any  $\tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma)$ . On the other hand, increasing the network's density increases the probability that fewer workers receive a larger portion of the offers, as a worker cannot accept more than one offer, this leads to a lower level of unraveling. This is captured by a decrease in  $\tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  for any  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma)$ . Since  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  is increasing in  $\tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma)$ , this implies a decrease in  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  for any  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma)$ .

In contrast, adding connections by increasing the span of the network has a monotone effect on unraveling.

**Proposition 2** Consider two regular environments  $\langle \theta_W^1, \theta_F^1, M, \alpha \rangle$  and  $\langle \theta_W^2, \theta_F^2, M, \alpha \rangle$  such that  $\theta_W^1(0) < \theta_W^2(0)$ ,  $\theta_F^1(0) < \theta_F^2(0)$ , and for all  $r \geq 1$ ,  $P(r, \theta_W^1) = P(r, \theta_W^2)$  and  $P(r, \theta_F^1) = P(r, \theta_F^2)$ . Assume that all firms employ label-free strategies. Then,  $\langle \theta_W^1, \theta_F^1, M, \alpha \rangle$  generates greater unraveling than  $\langle \theta_W^2, \theta_F^2, M, \alpha \rangle$ .

The intuition for Proposition 2 is straightforward. Increasing the number of connected workers and firms increases the number of offers at stage  $\mathcal{S} = -1$  without the adverse effect of an increase in density.

**Redistribution of links.** To capture redistribution of links in the network we use the standard notion of Mean Preserving Spread (MPS). We focus on redistribution of links among the workers (firms) that have a degree of at least 1, i.e. with no effect on the span of the network.<sup>22</sup> Proposition 3 predicts greater unraveling in markets in which some firms have many connections and others have few, relative to markets in which all firms have similar degrees. On the other hand, lesser unraveling is predicted in markets in which some workers have many connections and others have few, relative to markets in which all workers have similar degrees.

**Proposition 3** Consider two regular environments  $\langle \theta_W^1, \theta_F^1, M, \alpha \rangle$  and  $\langle \theta_W^2, \theta_F^2, M, \alpha \rangle$  such that  $\theta_W^1(0) = \theta_W^2(0)$ ,  $\theta_F^1(0) = \theta_F^2(0)$ , and assume that all firms employ label-free strategies. If  $P(\cdot, \theta_W^2)$  is a MPS of  $P(\cdot, \theta_W^1)$  and  $P(\cdot, \theta_F^1)$  is a MPS of  $P(\cdot, \theta_F^2)$ , then  $\langle \theta_W^1, \theta_F^1, M, \alpha \rangle$  generates greater unraveling than  $\langle \theta_W^2, \theta_F^2, M, \alpha \rangle$ .

If  $P(\cdot, \theta_W^2)$  is a MPS of  $P(\cdot, \theta_W^1)$  then in  $\langle \theta_W^2, \theta_F^2, M, \alpha \rangle$  there are more workers with high and low degrees and less workers with intermediate degrees compared to  $\langle \theta_W^1, \theta_F^1, M, \alpha \rangle$ . The probability that a worker receives at least one offer at stage  $\mathcal{S} = -1$  is increasing and concave in  $r_w$  and so is the probability that a worker receives at least one acceptable offer  $(1 - [1 - \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) + \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - \tilde{\mu}_{\theta_W, \theta_F, M, \alpha})]^{r_w})$ . Therefore,  $\tilde{\psi}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma) \leq \tilde{\psi}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma)$ . If  $P(\cdot, \theta_F^1)$  is a MPS of  $P(\cdot, \theta_F^2)$  then in  $\langle \theta_W^1, \theta_F^1, M, \alpha \rangle$  there are more firms with high and low degrees and less firms with intermediate degrees compared to

---

<sup>22</sup>A redistribution of links in a way that changes the span of the network is dominated by the effect of the change to the span. In particular, an increase  $W^1$  and/or  $F^1$  leads to greater unraveling.

$\langle \theta_W^2, \theta_F^2, M, \alpha \rangle$ . The probability that a firm  $f$  makes an offer to a worker  $w$  that is connected to it  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - 0.5^{r_f}) / (0.5 \cdot r_f)$  is decreasing and convex in  $r_f$ . Therefore,  $\tilde{\psi}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma) \leq \tilde{\psi}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma)$ .

The observation that greater unraveling is expected in markets in which some firms have many connections and others have few sheds light on some of the difficulties that market designers face when trying to prevent unraveling. In the market for judicial clerks, attempts to set the date of the hiring of clerks failed repeatedly. The 9th Circuit (California) is often claimed to be the one to unravel first. The reasons for this are under debate: the 9th Circuit judges point out that the East Coast judges have a geographical advantage as they are close to more top-ranked law schools. The East Coast circuits claim that the 9th Circuit unravels as it is less attractive to clerks, due to lower quality of its positions. While not resolving the debate, Proposition 3 suggests that asymmetry across firms with respect to their connections leads to greater unraveling.

## 6.2 Market rules and the accuracy of early information

In this section we study the effect of changing the post-graduation matching procedure via changes to  $\phi_M$  and the effect of changing the accuracy of early signals available to the workers at stage  $\mathcal{S} = -1$  via changes to  $\alpha$ .

**The post-graduation matching procedure's quality** ( $\phi_M$ ). The following result shows that increasing  $\phi_M$  leads to lesser unraveling, providing yet an additional motivation for the design of more statically-efficient markets.

**Proposition 4** *Consider two regular environments  $\langle \theta_W, \theta_F, M^1, \alpha \rangle$  and  $\langle \theta_W, \theta_F, M^2, \alpha \rangle$  such that  $\phi_{M^1} \leq \phi_{M^2}$ , and assume that all firms employ label-free strategies. Then  $\langle \theta_W, \theta_F, M^1, \alpha \rangle$  generates greater unraveling than  $\langle \theta_W, \theta_F, M^2, \alpha \rangle$ .*

It is also interesting to note that if  $\alpha \geq \frac{1-v}{1+v}$  then  $\langle \theta_W, \theta_F, M^1, \alpha \rangle$  generates strictly greater unraveling than  $\langle \theta_W, \theta_F, M^2, \alpha \rangle$  whereas otherwise  $\langle \theta_W, \theta_F, M^1, \alpha \rangle$  and  $\langle \theta_W, \theta_F, M^2, \alpha \rangle$  generate the same level of unraveling. To see why, note that if  $\alpha < \frac{1-v}{1+v}$  workers that receive at least one offer at stage  $\mathcal{S} = -1$  always accept one of the offers that they receive (independent

of  $\phi_M$ ), and  $\tilde{\mu}_{\theta_W, \theta_F, M^1, \alpha} = \tilde{\mu}_{\theta_W, \theta_F, M^2, \alpha} = 1$ . On the other hand, if  $\alpha \geq \frac{1-v}{1+v}$  then for some  $\phi_M$ , workers reject early offers from some firms, therefore  $\tilde{\mu}_{\theta_W, \theta_F, M^1, \alpha} \geq \tilde{\mu}_{\theta_W, \theta_F, M^2, \alpha}$ .

The connection between the matching procedure in the post-graduation market and the level of unraveling in a market were recognized in empirical and experimental work in market design. For example, Kagel and Roth (2000) examine the effect of the mechanism used in the post graduation market on the ability of the market to recover from unraveling. Kagel and Roth find that "*when the centralized mechanisms are introduced, there is only a small rollback of the unraveling that developed when the market was decentralized. But because of the congestion and competition in the market, some firms and workers who intend to make early matches find themselves unable to do so, and these participate in the centralized mechanism.*" As a result, Kagel and Roth conclude, firms and workers find out that the matching procedure improved and that they can do better by not hiring early.

**The accuracy of the early signal ( $\alpha$ ).** If  $\alpha$  is higher a worker  $w$  with  $s_w = h$  has a higher probability of being  $q_w = H$ . Thus  $w$  is more likely to reject offers at stage  $\mathcal{S} = -1$ . For the same reason, a firm  $f \in N_w$  is more likely to make an offer to  $w$  at stage  $\mathcal{S} = -1$ . These two forces result in a non monotonic relationship between  $\alpha$  and the unraveling level in the market.

**Proposition 5** Consider two regular environments  $\langle \theta_W, \theta_F, M, \alpha^1 \rangle$  and  $\langle \theta_W, \theta_F, M, \alpha^2 \rangle$  such that  $\alpha^1 \leq \alpha^2$ , and assume that all firms employ label-free strategies. Then,

1. If  $\alpha^2 \leq \frac{1-v}{1+\phi_M \cdot v}$  then  $\langle \theta_W, \theta_F, M, \alpha^2 \rangle$  generates greater unraveling than  $\langle \theta_W, \theta_F, M, \alpha^1 \rangle$ ,  
and
2. If  $\alpha^1 \geq \frac{\bar{c} + \frac{1}{2} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H - \pi_L}{\pi_H - \pi_L}$  then  $\langle \theta_W, \theta_F, M, \alpha^1 \rangle$  generates greater unraveling than  $\langle \theta_W, \theta_F, M, \alpha^2 \rangle$ .

If  $\alpha$  is very small, a worker  $w$  that receives at least one early offer accepts one of the offers with probability 1. Thus, a small increase in  $\alpha$  only increases the probability that a firm makes an early offer to a worker  $w$  with  $s_w = h$ , and for every  $\gamma \in [0, 1]$ ,  $\sigma_{\theta_W, \theta_F, M, \alpha^2}(\gamma) \geq \sigma_{\theta_W, \theta_F, M, \alpha^1}(\gamma)$ . On the other hand, if  $\alpha$  is large, a firm that is connected to at least one

worker  $w$  such that  $s_w = h$  makes an early offer with probability 1. Thus, a further increase in  $\alpha$  only decreases the probability that worker  $w$  accepts, and  $\tilde{\mu}_{\theta_W, \theta_F, M, \alpha^2} \leq \tilde{\mu}_{\theta_W, \theta_F, M, \alpha^1}$ .

**Remark 2** *Worker- and firm-driven unraveling.*<sup>23</sup>

*In many entry-level labor markets, information about workers becomes more accurate over time. Propositions 4 and 5 suggest that the time that workers spend in the training institutions can be divided into two time segments: early on,  $\alpha$  is small and workers would accept any job offer in order to insure themselves against unemployment. As a result,  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma) \in [0, 1]$  whereas  $\tilde{\mu}_{\theta_W, \theta_F, M, \alpha} = 1$ . This worker-driven unraveling can be more easily influenced by changes to firms' incentives, as the incentives of workers to contract early are too strong to be affected. Closer to graduation,  $\alpha$  is large, and the balance of power shifts; firms try to hire any high potential worker, whereas workers decline less desirable job offers (for all  $\gamma \in [0, 1]$ ,  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma) = 1$  whereas  $\tilde{\mu}_{\theta_W, \theta_F, M, \alpha} \in [0, 1]$ ). This firm-driven unraveling can be more easily influenced by policies that affect workers' incentives. Determining whether stage  $\mathcal{S} = -1$  is dominated by worker- or firm-driven unraveling depends on the market of interest.*

## 7 Unraveling and market outcomes

Welfare analysis for two-sided matching markets is subtle as there is an inherent trade-off between the gain of one agent and the loss of other agents on the same side of the market. Nevertheless, we are able to make the following observations.

**Corollary 2 :**

- (1) *Greater unraveling leads to higher aggregate utility of connected workers  $\left(\sum_{w \in \{w \in W | r_w \geq 1\}} u_w\right)$ .*
- (2) *Greater unraveling leads to lower aggregate payoffs of firms  $\left(\sum_{f \in F} \pi_f\right)$ .*

The first claim captures the insurance that unraveling provides to connected workers. The second claim points out that unraveling leads to the hiring of low productivity workers

---

<sup>23</sup>The division to firm vs. worker driven unraveling was suggested by Li and Suen (2000) who focus on the different qualities of the firms that trigger the unraveling process in each of the types of unraveling.

and reduces aggregate firms' profits. Evaluating aggregate welfare changes in the economy as a whole requires further assumptions about the outside options of workers and firms.

The idea that unraveling makes some workers better off, and that a centralized match can hinder unraveling (Proposition 4) helps to account for observed resistance to a match by groups of individual workers. For example, the match for medical interns to internships was under scrutiny when an anti-trust suit against the National Residency Matching Program (NRMP) and numerous other defendants was brought in 2002 by 16 law firms on behalf of 3 former residents seeking to represent the class of all former residents (and naming as defendants a class including all hospitals that employ residents). It was dismissed on August 12, 2004 in an Opinion, Order & Judgment by Judge Paul L. Friedman.<sup>24</sup>

## 8 Conclusion

This paper is a first attempt to tackle the phenomenon of early hiring in entry-level labor markets in the presence of social networks connecting employers and potential workers. To this end, we propose a model of local interaction in which information flows via connections in a network. While the idea that social networks are used as a mean of transferring information, and in particular information related to job search, is widely accepted in the economic literature, it has not yet been incorporated into the analysis of the timing of hiring in labor markets. Our model provides a first step in this direction.

In our model, the incentives of firms to make early offers depend on the aggregate level of early hiring which in turn depends on the entire network structure in complex ways. Thus, a firm's best response depend on the firm's beliefs with respect to the entire network structure. To overcome that we provide formal analysis of firms beliefs in large networks that are chosen at random and combine tools from graph theory, matching theory, and market design.

We find that the structure of the network affects unraveling in systematic ways. Moreover, we show that improving the design of the post-graduation market by improving the quality of the match between workers and firms leads to lesser unraveling.

---

<sup>24</sup>Much of the discussion surrounding the lawsuit revolved around the effect of the match on salaries. See also Bulow and Levin (2006), Niederle and Roth (2003), and Kojima (2007). Nevertheless, Niederle and Roth find no such effect.

## 9 Appendix

### 9.1 Matching procedures - definitions

In the absence of a meaningful network, the analysis of stage  $\mathcal{S} = 0$  lends itself to the more familiar analysis of one-to-one matching markets. Formally, let  $w_0$  and  $f_0$  be the null worker and firm.

**Definition 8** (Roth and Sotomayor 1990) *For a set of workers  $W'$  and a set of firms  $F'$ , a **one-to-one matching** is a function  $\mathcal{M} : W' \cup F' \rightarrow W' \cup F' \cup \{w_0, f_0\}$  such that  $w = \mathcal{M}(f)$  if and only if  $\mathcal{M}(w) = f$  and for all  $w \in W'$  and  $f \in F'$ :*

- either  $\mathcal{M}(w) \in F'$  or  $\mathcal{M}(w) = f_0$ , and
- either  $\mathcal{M}(f) \in W'$  or  $\mathcal{M}(f) = w_0$

Much of the matching literature focuses on fixed exogenous sets of workers and firms with perfect information. In our environment, the sets of workers and firms that reach stage  $\mathcal{S} = 0$  unmatched ( $W_H^0, W_L^0$ , and  $F^0$ ) are determined endogenously in stage  $\mathcal{S} = -1$  and we are required to define a notion of a matching procedure. Intuitively, a matching procedure captures the rules of the market which in turn determine the mapping from sets of workers and firms to a probability distribution over matchings.

Denote the set of all subsets of a set  $A$  (the power set of set  $A$ ) by  $\mathcal{P}(A)$ .

**Definition 9** *Let  $\overline{\mathcal{M}}(W', F')$  be the set of all one-to-one matchings over  $W'$  and  $F'$ , and let  $\Delta(\overline{\mathcal{M}}(W', F'))$  be the set of all probability distributions on elements of  $\overline{\mathcal{M}}(W', F')$ . A **matching procedure** is a function  $M : \mathcal{P}(W') \times \mathcal{P}(F') \rightarrow \Delta(\overline{\mathcal{M}}(W', F'))$ .*

We now define the notions of anonymous matching procedures and of matching procedures that guarantee a stable matching.

**Definition 10** *Let  $U \equiv (u_{w_1}, u_{w_2}, \dots, u_{w_{|W'|}})$  and  $U' \equiv (u'_{w_1}, u'_{w_2}, \dots, u'_{w_{|W'|}})$  be two profiles of workers' utility functions such that for some  $i \neq j$ ,  $u_{w_i} = u'_{w_j}$  and  $u_{w_j} = u'_{w_i}$  and for any  $k \notin \{i, j\}$ ,  $u_{w_k} = u'_{w_k}$ . Let  $U'' \equiv (u''_{w_1}, u''_{w_2}, \dots, u''_{w_{|W'|}})$  be a profile of workers' utility functions such that there exist two firms  $f'$  and  $f''$  such that: (1) for every  $j$ ,  $u_{w_j}(f') = u''_{w_j}(f'')$  and  $u_{w_j}(f'') = u''_{w_j}(f')$ ; and (2) for every  $f \in F \setminus \{f', f''\}$  and for every  $j$ ,  $u_{w_j}(f) = u''_{w_j}(f)$ . A matching procedure is **anonymous** if for every  $W' \subseteq W$  and  $F' \subseteq F$  and in every matching  $\mathcal{M}$  that has a positive probability given the matching procedure the following holds:<sup>25</sup>*

1. for any  $f \in F'$ ,  $Pr[\mathcal{M}(w_i) = f|U] = Pr[\mathcal{M}(w_j) = f|U']$  and for any  $k \notin \{i, j\}$ ,  $Pr[\mathcal{M}(w_k) = f|U] = Pr[\mathcal{M}(w_k) = f|U']$ .

<sup>25</sup>A matching procedure is anonymous if it depends only on the preferences of workers' and firms' and not on their labels or position in the network. As firms' payoff functions are identical at stage  $\mathcal{S} = 0$ , an anonymous matching procedure will depend only on the workers' preferences as captured by the definition.

2. for any  $w \in W$ ,  $Pr[\mathcal{M}(w) = f'|U] = Pr[\mathcal{M}(w) = f''|U'']$  and for every  $f \in F' \setminus \{f', f''\}$ ,  $Pr[\mathcal{M}(w) = f|U] = Pr[\mathcal{M}(w) = f|U'']$ .

**Definition 11** A matching procedure **guarantees a stable matching** if for every  $W'$  and  $F'$  and in every matching  $\mathcal{M}$  that has a positive probability given the matching procedure the following holds: [1] for any  $w \in W'$  and  $f \in F'$  such that  $\mathcal{M}(w) = f$ ,  $u_w(\cdot|\mathcal{M}(w) = f) \geq u_w(\cdot|\mathcal{M}(w) = f_0)$  and  $\pi_f(\cdot|\mathcal{M}(f) = w) \geq \pi_f(\cdot|\mathcal{M}(f) = w_0)$  (for any matched worker and firm, both prefer to be matched to the other than be unmatched); and [2] for any  $w \in W'$  and  $f \in F'$  such that  $\mathcal{M}(w) = f_0$  and  $\mathcal{M}(f) = w_0$ ,  $u_w(\cdot|\mathcal{M}(w) = f_0) > u_w(\cdot|\mathcal{M}(w) = f)$  or  $\pi_f(\cdot|\mathcal{M}(f) = w_0) > \pi_f(\cdot|\mathcal{M}(f) = w)$  (for any unmatched worker and firm, at least one of them prefers not to be matched to the other).

## 9.2 Best response correspondence: firms

In any equilibrium, firm  $f$  makes a job offer at stage  $\mathcal{S} = -1$  if there exists at least one worker  $w \in W$  such that

$$\Pi_{G(n_w, \theta_W, \theta_F), M, \alpha} \left( f | \sigma_f \left( \widetilde{W}_h \right) = w \right) > E_{G(n_w, \theta_W, \theta_F), M, \alpha} [\pi_f] \quad (10)$$

and does not make an offer if the inequality is reversed for all  $w \in W$ . Since by definition  $Pr\{w \text{ accepts}\} = 1 - Pr\{w \text{ rejects}\}$ , condition (10) can be restated.

$$Pr \left[ q_w = H | N_f, \widetilde{W}_h \right] \cdot \pi_H + Pr \left[ q_w = L | N_f, \widetilde{W}_h \right] \cdot \pi_L - c_f > E_{G(n_w, \theta_W, \theta_F), M, \alpha} [\pi_f] \quad (11)$$

Recall that  $W_q^0$  is the set of workers of productivity  $q$  that reach stage  $\mathcal{S} = 0$  unmatched, and  $F^0$  is the set of firms that reach stage  $\mathcal{S} = 0$  unmatched. By Lemma 1, if high productivity workers are scarce and if no worker who receives a low signal in stage  $\mathcal{S} = -1$  is hired early (in stage  $\mathcal{S} = -1$ ),  $E_{G(n_w, \theta_W, \theta_F), M, \alpha} [\pi_f] = \frac{|W_H^0|}{|F^0|} \cdot \pi_H$ . Moreover, following condition (3), condition (11) never holds for any worker  $w$  such that  $s_w = l$ , or for any worker  $w \notin N_f$  (because the firm did not receive credible information regarding  $s_w$ ). Consequently, scarcity of high productivity workers implies that firm  $f$  is better off making a job offer at stage  $\mathcal{S} = -1$  if and only if  $\widetilde{W}_h \neq \emptyset$  and

$$\alpha \cdot \pi_H + (1 - \alpha) \cdot \pi_L - c_f > E \left( \frac{|W_H^0|}{|F^0|} \right) \cdot \pi_H \quad (12)$$

Let  $\gamma$  be the fraction of workers hired at stage  $\mathcal{S} = -1$ . Conditional on  $\gamma$ ,

$$E \left( \frac{|W_H^0|}{|F^0|} | \gamma \right) = \frac{(\frac{1}{2} - \alpha \cdot \gamma) \cdot n_w}{(1 - \gamma) \cdot n_f} = \frac{\frac{1}{2} - \alpha \cdot \gamma}{1 - \gamma} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r}$$

To see why the second equality holds, note that for  $n_w, n_f, \theta_W(r), \theta_F(r)$  to be consistent with a network structure, it must hold that  $\frac{n_w}{n_f} = \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r}$ .

Therefore, condition (12) can be rewritten as

$$\alpha \cdot \pi_H + (1 - \alpha) \cdot \pi_L - c_f > \frac{\frac{1}{2} - \alpha \cdot \gamma}{1 - \gamma} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H \quad (13)$$

and the ex-ante probability that a randomly chosen firm with degree  $r$  makes an offer at stage  $\mathcal{S} = -1$ , conditional on  $\widetilde{W}_h \neq \emptyset$  and the expectation that  $\gamma = \gamma_1$ , is

$$\begin{aligned} \sigma_{\theta_W, \theta_F, M, \alpha}(\gamma_1) &= \sigma_{G(n_w, \theta_W, \theta_F), M, \alpha}(r, \widetilde{W}_h, \gamma_1) \\ &= Pr \left\{ \alpha \cdot \pi_H + (1 - \alpha) \cdot \pi_L - c_f > \frac{\frac{1}{2} - \alpha \cdot \gamma_1}{1 - \gamma_1} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H \right\} \\ &= \mathcal{D} \left( \alpha \cdot \pi_H + (1 - \alpha) \cdot \pi_L - \frac{\frac{1}{2} - \alpha \cdot \gamma_1}{1 - \gamma_1} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H \right) \end{aligned} \quad (14)$$

### 9.3 Best response correspondence: workers

In any equilibrium, worker  $w$  accepts a job offer at stage  $\mathcal{S} = -1$  if there exists at least one firm  $f \in \widetilde{F}_w$  such that

$$1 + v_{wf} > U_{G(n_w, \theta_W, \theta_F), M, \alpha} \left( \cdot | \mu_w \left( \cdot, \widetilde{F}_w \right) = f_0 \right) \quad (15)$$

and does not accept any job offer if the inequality is reversed.

Combining Lemma 1 and the observation that a worker  $w$  does not receive an early offer unless  $s_w = h$ , condition (15) can be simplified to

$$1 + v_{wf} > \alpha \cdot E_{G(n_w, \theta_W, \theta_F), M, \alpha} [u_w | H] \quad (16)$$

Pick an edge  $(f, w) \in E$  u.a.r. If all workers expect  $\gamma = \gamma_1$ , then the ex-ante probability that condition (16) holds for  $(f, w)$  is

$$\mu_{G, M, \alpha}(\gamma) = \mu_{G(n_w, \theta_W, \theta_F), M, \alpha}(r_w, \widetilde{F}_w, \gamma) \quad (17)$$

$$= Pr \left\{ 1 + v_{wf} > \alpha \cdot E_{G(n_w, \theta_W, \theta_F), M, \alpha} [u_w | H, \gamma_1] \right\} \quad (18)$$

$$= 1 - \mathcal{H} \left( \alpha \cdot E_{G(n_w, \theta_W, \theta_F), M, \alpha} [u_w | H, \gamma_1] - 1 \right) \quad (19)$$

Recall that  $\mathcal{H}$  is continuous. Thus,  $\mu_{G, M, \alpha}(\gamma)$  is the induced probability that a worker that received *exactly* one job offer at stage  $-1$  accepts the offer. Similarly, a worker that received exactly  $m$  offers at stage  $\mathcal{S} = -1$  accepts an offer with probability  $1 - (1 - \mu_{G, M, \alpha}(\gamma))^m$ .

Conditional on accepting an offer, worker  $w$  solves the following maximization problem to decide which offer to accept:

$$Max_{f \in \widetilde{F}_w} 1 + v_{wf}$$

## 9.4 Proofs

**Lemma 2** Let  $\langle \theta_W, \theta_F, \alpha \rangle$  exhibit scarcity of high productivity workers, and let  $\widehat{G}(\theta_W, \theta_F, n_w)$  be any network that is consistent with  $\theta_W, \theta_F$  and  $n_w$ . Assume further that no worker who receive a low signal in stage  $\mathcal{S} = -1$  is hired early (in stage  $\mathcal{S} = -1$ ). Let  $n_w^{\widehat{G}(\theta_W, \theta_F, n_w)}(H, 0, \alpha)$  be a random variable (r.v.) that captures the number of workers of high productivity that are not hired in stage  $\mathcal{S} = -1$  and let  $n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0, \alpha)$  be a r.v. that captures the number of firms that did not hire in stage  $\mathcal{S} = -1$ . Then, there exist  $\zeta > 0$  such that

$$Pr \left( \lim_{n_w \rightarrow \infty} \frac{n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0, \alpha) - n_w^{\widehat{G}(\theta_W, \theta_F, n_w)}(H, 0, \alpha)}{n_w} > \zeta \right) = 1$$

**Proof.** Let  $p_w^{-1, n_w}(s_w)$  be a r.v. that captures the proportion of workers that receive a signal  $s_w$  in stage  $\mathcal{S} = -1$ , and let  $p_w^{0, n_w}(s_w, H)$  be a r.v. that captures the proportion of workers that receive a signal  $s_w$  in stage  $\mathcal{S} = -1$  AND are of high productivity in stage  $\mathcal{S} = 0$ . Let  $\gamma^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot)$  be a r.v. that captures the number of workers hired in stage  $\mathcal{S} = -1$ . Let  $p_w^{0, \widehat{G}(\theta_W, \theta_F, n_w)}(H, 0)$  be a r.v. that captures the proportion of workers that are of high productivity and are not hired at stage  $\mathcal{S} = -1$  (as a proportion of  $n_w$ ). Finally, note that  $p_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot) = \frac{n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0, \alpha) - \gamma^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot)}{n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0, \alpha)}$  is a r.v. that captures the proportion of firms that are not matched before stage  $\mathcal{S} = 0$ . Then,

$$p_w^{0, \widehat{G}(\theta_W, \theta_F, n_w)}(H, 0) \leq p_w^{0, n_w}(l, H) + p_w^{-1, n_w}(h) - \frac{\gamma^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot)}{n_w} \quad (20)$$

where the inequality holds because  $p_w^{-1, n_w}(h) - \frac{\gamma^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot)}{n_w}$  is the proportion of workers who receive high signal and are not hired at stage  $\mathcal{S} = -1$ . This equals the proportion of workers that receive high signal, are of high productivity, and are not hired at stage  $\mathcal{S} = -1$  only if all of the high signal workers who are not hired at stage  $\mathcal{S} = -1$  are also of high productivity.

By the strong law of large numbers

$$Pr \left( \lim_{n_w \rightarrow \infty} p_w^{-1, n_w}(h) = \frac{1}{2} \right) = 1; \quad Pr \left( \lim_{n_w \rightarrow \infty} p_w^{0, n_w}(l, H) = \frac{1 - \alpha}{2} \right) = 1$$

Thus, inequality (20) implies that,

$$Pr \left( \lim_{n_w \rightarrow \infty} \left( p_w^{0, \widehat{G}(\theta_W, \theta_F, n_w)}(H, 0) - \left[ \frac{1 - \alpha}{2} + \frac{1}{2} - \frac{\gamma^{\widehat{G}(\theta_W, \theta_F, n_w)}}{n_w} \right] \right) \leq 0 \right) = 1 \quad (21)$$

By the assumption of scarcity of high productivity workers for every  $\langle \widehat{G}(n_w, \theta_W, \theta_F), \alpha \rangle$

there exists  $\eta > 1$  such that  $n_f > \eta \cdot \left(1 - \frac{\alpha}{2}\right) \cdot n_w$ . Consequently,

$$n_f \cdot p_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot) = n_f - \gamma^{\widehat{G}(\theta_W, \theta_F, n_w)} > \eta \cdot \left(1 - \frac{\alpha}{2}\right) \cdot n_w - \gamma^{\widehat{G}(\theta_W, \theta_F, n_w)} \quad (22)$$

Combining (21) and (22) and some algebra yields that there exists  $\eta > 1$  such that,

$$Pr \left( \lim_{i \rightarrow \infty} \frac{n_f \cdot p_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot) - n_w \cdot p_w^{0, \widehat{G}(\theta_W, \theta_F, n_w)}(H, 0)}{n_w} > (\eta - 1) \cdot \left(1 - \frac{\alpha}{2}\right) \right) = 1$$

since by definition  $n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0, \alpha) = n_f \cdot p_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(\cdot)$ ,  $n_w^{\widehat{G}(\theta_W, \theta_F, n_w)}(H, 0, \alpha) = n_w \cdot p_w^{0, \widehat{G}(\theta_W, \theta_F, n_w)}(H, 0)$ , and since  $(\eta - 1) \cdot \left(1 - \frac{\alpha}{2}\right) > 0$ , the proof is complete. ■

**Lemma 1 - Proof.** The proof of part 1 of the Lemma is immediate from the definitions above. The proof for part 2 is as follows:

Let  $F(w, F^0, \xi, \phi) \equiv \{f \in F^0 \mid |1 + v_{wf} - (1 + \phi \cdot v)| < \xi\}$  and denote the empty set by  $\emptyset$ . For given  $\phi \in [0, 1]$ , and  $\xi > 0$  consider the following algorithm:

Let  $\mathcal{W} = \widehat{W}_H^0$  and  $\mathcal{F} = \widehat{F}^0$   
While  $\mathcal{W} \neq \emptyset$  and  $\mathcal{F} \neq \emptyset$   
  Select  $w \in \mathcal{W}$  uniformly at random  
  If  $F(w, \mathcal{F}, \xi, \phi) \neq \emptyset$   
    Pick a firm  $f \in F(w, \mathcal{F}, \xi, \phi)$  uniformly at random  
    Match  $w$  to  $f$   
    Let  $\mathcal{W} = \mathcal{W} \setminus w$  and  $\mathcal{F} = \mathcal{F} \setminus f$   
  Otherwise  
    Pick a firm  $f \in \mathcal{F}$  uniformly at random  
    Match  $w$  to  $f$   
    Let  $\mathcal{W} = \mathcal{W} \setminus w$  and  $\mathcal{F} = \mathcal{F} \setminus f$

The algorithm matches a firm and a worker at every iteration and therefore always stops when either  $\mathcal{W} \neq \emptyset$  or  $\mathcal{F} \neq \emptyset$  and provides a stable matching. The probability distribution over the outcomes of the algorithm is a matching procedure  $M(W^0, F^0)$ . The anonymity of the procedure is directly implied by the randomness in the selection of the worker and the firm out of the relevant sets.

It is left to show that for every  $w \in \mathcal{W}$  that is selected by the algorithm and any  $\mathcal{F}$  that is reached by the algorithm given a network  $\widehat{G}(\theta_W, \theta_F, n_w)$ ,

$$\lim_{n_w \rightarrow \infty} Pr [F(w, \mathcal{F}, \xi, \phi) \neq \emptyset] = 1$$

Let  $n_w^{\widehat{G}(\theta_W, \theta_F, n_w)}(H, 0)$  be a r.v. that captures the number of workers of high productivity that are not hired at stage  $\mathcal{S} = -1$  and let  $n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0)$  be a r.v. that captures the number

of firms that did not hire at stage  $\mathcal{S} = -1$ . Then, by Lemma 2 there exists  $\zeta > 0$  such that

$$Pr \left( \lim_{n_w \rightarrow \infty} \frac{n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0, \alpha) - n_w^{\widehat{G}(\theta_W, \theta_F, n_w)}(H, 0, \alpha)}{n_w} > \zeta \right) = 1$$

Let  $|A|$  be the number of elements in a set  $A$ . For any  $n_w$  and at every iteration of the algorithm,  $|\mathcal{F}| \geq n_f^{\widehat{G}(\theta_W, \theta_F, n_w)}(0, \alpha) - n_w^{\widehat{G}(\theta_W, \theta_F, n_w)}(H, 0, \alpha)$ , and this holds with equality only when the last worker  $\tilde{w} \in \mathcal{W}$  is chosen by the algorithm. As a result, there exists  $\zeta > 0$  such that when a random worker  $w \in \mathcal{W}$  is chosen by the algorithm

$$Pr \left( \lim_{n_w \rightarrow \infty} \frac{|\mathcal{F}|}{n_w} > \zeta \right) = 1$$

or

$$Pr(\lim_{n_w \rightarrow \infty} |\mathcal{F}| > \zeta \cdot n_w) = 1$$

To complete the proof, fix  $\zeta > 0$  and  $\xi > 0$ , and consider a randomly selected worker  $\tilde{w}$  and a set  $\tilde{F}(n_w)$  of  $\zeta \cdot n_w$  firms that is chosen independently of the worker's preferences. Let  $B(n_w)$  be the event that there is no firm  $\tilde{f} \in \tilde{F}(n_w)$  such that  $\left| 1 + v_{\tilde{w}\tilde{f}} - (1 + \phi \cdot v) \right| < \xi$ . Recall that  $v_{\tilde{w}\tilde{f}}$  is distributed  $\mathcal{H}$  with positive density in every point in the support  $[-v, v]$ . Then,

$$Pr(B(n_w)) = (1 - [\mathcal{H}(\phi \cdot v + \xi) - \mathcal{H}(\phi \cdot v - \xi)])^{\zeta \cdot n_w}$$

and

$$\lim_{n_w \rightarrow \infty} Pr(B(n_w)) = 0$$

It is only left to recall that when there is no firm  $\tilde{f} \in \tilde{F}(n_w)$  such that  $\left| 1 + v_{\tilde{w}\tilde{f}} - (1 + \phi \cdot v) \right| < \xi$ ,  $v_{\tilde{w}\tilde{f}}$  is bounded. ■

**Definition 12** For two random variables (r.v.)  $X, Y$  with support on some countably infinite set  $\mathcal{X}$ , the total variational distance between  $X$  and  $Y$ ,  $TVD(X, Y)$ , is defined as  $\sum_{x \in \mathcal{X}} |Pr(X = x) - Pr(Y = x)|$ .

For a distribution over networks  $\Delta G$  let  $\bar{b}^F(r, \Delta G)$  be the random vector of length  $r$  chosen as follows: [1] choose a network  $G$  according to  $\Delta G$ , [2] choose a worker with degree  $r$  u.a.r. from all workers with degree  $r$  in  $G$ , and [3] let  $\bar{b}^F(r, \Delta G)$  be the vector of the degrees of all firms in  $N_w$  ordered randomly (with equal probability given to each ordering). Let  $\bar{b}^{0F}(r, \theta_F)$  be a vector of length  $r$  such that for every  $i \in \{1, 2, \dots, r\}$ ,  $\bar{b}_i^{0F}$  equals  $r'$  with probability  $\tilde{P}^F(r', \theta_F)$  and such that  $\left\{ \bar{b}_i^{0F} \right\}_{i \in \{1, 2, \dots, r\}}$  are determined independently of each other.

**Lemma 3** (Fainmesser and Goldberg 2010) For all  $r$  and finite support  $\theta_W, \theta_F$ ,

$$\lim_{n_w \rightarrow \infty} TVD \left| \bar{b}^F(r, G(n_w, \theta_W, \theta_F)), \bar{b}^{0F}(r, \theta_F) \right| = 0$$

**Lemma 4** Let  $\theta_W, \theta_F$  have finite support and  $\langle \theta_W, \theta_F, \alpha \rangle$  exhibit scarcity of high productivity workers. Consider a market procedure  $M$  that is parameterized by  $\phi_M \in [0, 1]$ . Assume that all firms employ label-free strategies. Finally, let  $\psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\hat{\gamma})$  be the r.v. that captures the fraction of workers hired at stage  $\mathcal{S} = -1$  if all firms and workers best respond to the belief that  $\gamma = \hat{\gamma}$ . Then, for every  $\epsilon > 0$  and  $\hat{\gamma} \in [0, 1]$

$$\lim_{n_w \rightarrow \infty} \Pr \left( \left| \psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\hat{\gamma}) - \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\hat{\gamma}) \right| < \epsilon \right) = 1$$

**Proof.** Following condition (3) and the discussion in sections 9.2 and 9.3, a worker  $w$  has a positive probability to be hired at stage  $\mathcal{S} = -1$  only if  $s_w = h$  and  $r_w \geq 1$ .

Consider a network  $G$ . Select worker  $w$  u.a.r. and then select a firm  $f \in N_w$  u.a.r. Let  $\tau_{G, M, \alpha}(\gamma)$  be the probability that  $f$  makes an early offer to  $w$  if  $s_w = h$ . Recall that the realization  $(q_w, s_w)$  is independent of anything else, then by Lemma 3,

$$\lim_{n_w \rightarrow \infty} \tau_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma) = \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma)$$

Note also that by Lemma 1 and Definition 2, for any  $W_H^0, W_L^0, F^0$  that are possible under the assumptions above and for any  $\xi > 0$

$$\lim_{n_w \rightarrow \infty} \sup_w \left| E_{G(n_w, \theta_W, \theta_F), M, \alpha} [u_w | H, W_H^0, W_L^0, F^0] - (1 + \phi_M \cdot v) \right| < \xi$$

and therefore

$$\lim_{n_w \rightarrow \infty} \mu_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma) = \tilde{\mu}_{\theta_W, \theta_F, M, \alpha}$$

which in turn, together with the independence of  $\{v_{wf}\}_{f \in F}$  and Lemma 3 imply that as  $n_w \rightarrow \infty$  the probability that  $w$  receives at least one early offer (at  $\mathcal{S} = -1$ ) that she would like to accept, conditional on  $s_w = h$ , converges to

$$1 - \left[ (1 - \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma)) + \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - \tilde{\mu}_{\theta_W, \theta_F, M, \alpha}) \right]^{r_w}$$

For a given graph  $G$  let  $\hat{w}_h(G)$  be a r.v. that captures the number of workers in the set  $\widehat{W}_h(G) = \{w \in W | r_w \geq 1 \text{ and } s_w = h\}$  and let  $x_W(G) = \frac{\hat{w}_h(G)}{n_w}$ . Let  $P^{\widehat{W}}(\cdot | G)$  be the degree distribution of workers in  $\widehat{W}_h(G)$ . The probability that a randomly chosen worker  $w \in \widehat{W}_h(G)$  is hired is

$$\sum_{r_w=1}^{\infty} P^{\widehat{W}}(r_w | G) \left( 1 - \left[ (1 - \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma)) + \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - \tilde{\mu}_{\theta_W, \theta_F, M, \alpha}) \right]^{r_w} \right)$$

The proof is then completed by applying the strong law of large numbers twice to establish that

- (1)  $\Pr \left( \lim_{n_w \rightarrow \infty} x_W(G) = \frac{1}{2} \cdot \theta_W \right) = 1$  and
- (2) for any  $r$ ,  $\Pr \left( \lim_{n_w \rightarrow \infty} P^{\widehat{W}}(r | G(n_w, \theta_W, \theta_F)) = P(r, \theta_W) \right) = 1$ . ■

**Theorem 1 - Proof.** By the convexity of the support for  $\gamma$  and the continuity of  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma)$  a fixed point exists.

We now show that if  $\gamma^* = \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma^*)$  then  $\gamma^*$  is a 0-equilibrium with  $\langle \theta_W, \theta_F, M, \alpha \rangle$ . The proof goes in two steps. First, Lemma 4 implies that for any  $\zeta > 0$  there exists  $n_w$  such that for any  $n'_w > n_w$ ,  $\Pr(\psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma^*) \in [\gamma^* - \zeta, \gamma^* + \zeta]) > 1 - \zeta$ .

Second, we are required to show that for any  $\varepsilon > 0$  there exists  $\zeta$  such that for any  $\zeta' < \zeta$ , workers' and firms' best responses for  $\gamma^*$  satisfy the conditions for an  $\varepsilon$ -equilibrium for any  $\gamma \in [\gamma^* - \zeta', \gamma^* + \zeta']$ . The latter follows from the continuity (and independence of  $n_w$ ) of the conditions on firms' and workers' decisions (equations 13 and 16).

Since the payoffs of workers and firms are bounded, Step 2 completes the proof that if  $\gamma^* = \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma^*)$  then  $\gamma^*$  is a 0-equilibrium with  $\langle \theta_W, \theta_F, M, \alpha \rangle$ .

We are left to show that if  $\gamma^*$  is a 0-equilibrium with  $\langle \theta_W, \theta_F, M, \alpha \rangle$  then  $\gamma^* = \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma^*)$ . Assume by contradiction that  $\gamma^* = \tilde{\psi}_{\theta_W, \theta_F, M, \alpha}(\gamma^*) + \varkappa$  for some  $\varkappa \neq 0$ . Then by Lemma 4 and following the argument from step 1 above, for any  $\zeta > 0$  there exists  $n_w$  such that for any  $n'_w > n_w$ ,  $\Pr(\psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma^*) \in [\gamma^* + \varkappa - \zeta, \gamma^* + \varkappa + \zeta]) > 1 - \zeta$ .

As a result, there exist  $\zeta_1 > 0$  and  $n_w$  such that for any  $n'_w > n_w$ ,  $\Pr(\psi_{G(n_w, \theta_W, \theta_F), M, \alpha}(\gamma^*) \in [\gamma^* - \zeta_1, \gamma^* + \zeta_1]) \leq \zeta_1 < 1 - \zeta_1$ , contradiction to  $\gamma^*$  being a 0-equilibrium. ■

**Proposition 1 - Proof.** First, note that  $\langle \theta_W^{\rho^L}, \theta_F^{\rho^L}, M, \alpha \rangle$  and  $\langle \theta_W^{\rho^H}, \theta_F^{\rho^H}, M, \alpha \rangle$  are by definition regular environments, and note that for every  $\langle \theta_W^\rho, \theta_F^\rho, M, \alpha \rangle$  and  $\gamma$ ,  $\frac{\partial \sigma_{\theta_W^\rho, \theta_F^\rho, M, \alpha}(\gamma)}{\partial \rho} = 0$  and  $\frac{\partial \tilde{\mu}_{\theta_W^\rho, \theta_F^\rho, M, \alpha}(\gamma)}{\partial \rho} = 0$ . In regular environments, we can rely on Corollary 1 and prove the Proposition by showing that there exists  $\bar{r}(M, \alpha), \underline{r}(M, \alpha) \in \mathbb{Z}^+$  such that if  $\max\{\rho^H \cdot r | \theta_F(r) > 0\} < \underline{r}$  then  $\tilde{\psi}_{\theta_W^{\rho^H}, \theta_F^{\rho^H}, M, \alpha}(\gamma) \geq \tilde{\psi}_{\theta_W^{\rho^L}, \theta_F^{\rho^L}, M, \alpha}(\gamma)$ , and if  $\min\{\rho^L \cdot r | r \geq 1, \theta_F(r) > 0\} > \bar{r}$ , then  $\tilde{\psi}_{\theta_W^{\rho^H}, \theta_F^{\rho^H}, M, \alpha}(\gamma) \leq \tilde{\psi}_{\theta_W^{\rho^L}, \theta_F^{\rho^L}, M, \alpha}(\gamma)$ .

To reduce the notation that we carry throughout the proof, fix  $\theta_W, \theta_F, M, \alpha$ , and  $\gamma$  and let  $\tilde{\psi}_\rho = \tilde{\psi}_{\theta_W^\rho, \theta_F^\rho, M, \alpha}(\gamma)$ ,  $\tilde{\tau}_\rho = \tilde{\tau}_{\theta_W^\rho, \theta_F^\rho, M, \alpha}(\gamma)$ ,  $\sigma = \sigma_{\theta_W^\rho, \theta_F^\rho, M, \alpha}(\gamma)$ ,  $\tilde{\mu} = \tilde{\mu}_{\theta_W^\rho, \theta_F^\rho, M, \alpha}$  and  $x = \sigma \cdot \tilde{\mu}$ . We can drop the  $\gamma$  argument since a claim of greater diffusion is proved by a shift in  $\tilde{\psi}_{\theta_W^\rho, \theta_F^\rho, M, \alpha}(\gamma)$  for every  $\gamma \in [0, 1]$  and since  $\sigma$  and  $\tilde{\mu}$  (and  $x$ ) are independent of  $\rho$  and therefore can be treated of as exogenous for a given  $\gamma$ . Substituting in the definitions of  $\tilde{\tau}$  and  $x$  and some algebra yields,

$$\tilde{\psi}_\rho = \frac{1}{2} \cdot (1 - \theta_W(0)) \cdot \sum_{r_w=1}^{\infty} P(r_w, \theta_W) \cdot (1 - [1 - \tilde{\tau}_\rho + \tilde{\tau}_\rho \cdot (1 - \tilde{\mu})]^{\rho \cdot r_w}) \quad (23)$$

$$= \frac{1}{2} \cdot (1 - \theta_W(0)) - \frac{1}{2} \cdot (1 - \theta_W(0)) \cdot \sum_{r_w=1}^{\infty} P(r_w, \theta_W) \cdot \left[ 1 - x \cdot \left\{ \sum_{r_f=1}^{\infty} \tilde{P}(r_f, \theta_F) \cdot [(1 - 0.5^{\rho \cdot r_f}) / (0.5 \cdot \rho \cdot r_f)] \right\}^{\rho \cdot r_w} \right] \quad (24)$$

and

$$\frac{\partial \tilde{\psi}_\rho}{\partial \rho} = -\frac{1}{2} \cdot (1 - \theta_W(0)) \cdot \sum_{r_w=1}^{\infty} P(r_w, \theta_W) \cdot \frac{\partial \varphi_\rho(r_w)}{\partial \rho} \quad (25)$$

where

$$\varphi_\rho(r_w) = \left[ 1 - x \cdot \left\{ \sum_{r_f=1}^{\infty} \tilde{P}(r_f, \theta_F) \cdot [(1 - 0.5^{\rho r_f}) / (0.5 \cdot \rho \cdot r_f)] \right\} \right]^{\rho r_w} \quad (26)$$

Let  $\varphi_\rho(r_w, r_f) = [1 - x \cdot [(1 - 0.5^{\rho r_f}) / (0.5 \cdot \rho \cdot r_f)]]^{\rho r_w}$ . We note three important facts:

1.  $\varphi_\rho(r_w)$  and  $\varphi_\rho(r_w, r_f)$  are twice differentiable.
2. If  $\frac{\partial \varphi_\rho(r_w, r)}{\partial \rho} \geq 0$  ( $\frac{\partial \varphi_\rho(r_w, r)}{\partial \rho} < 0$ ) for every  $r \in \mathbb{R}^+$  in the convex hull of  $\{r_f | \theta_F(r_f) > 0\}$ , then  $\frac{\partial \varphi_\rho(r_w)}{\partial \rho} \geq 0$  ( $\frac{\partial \varphi_\rho(r_w)}{\partial \rho} < 0$ ).
3.  $\text{sign} \left\{ \frac{\partial \varphi_\rho(r_w, r_f)}{\partial \rho} \right\} = \text{sign} \left\{ \frac{\partial \ln(\varphi_\rho(r_w, r_f))}{\partial \rho} \right\}$ .

Thus to prove Proposition 1 it is sufficient to show that:

- **Step 1:**  $\varphi_2(r_w, 1) \leq \varphi_1(r_w, 1)$ .
- **Step 2:** there exists  $\bar{\rho}(M, \alpha) \in \mathbb{Z}^+$  such that for all  $\rho \geq \bar{\rho}(M, \alpha)$  and any  $r_w \in \mathbb{Z}^+$ ,  $\frac{\partial \ln(\varphi_\rho(r_w, 1))}{\partial \rho} \geq 0$ .

The proof of Step 1 follows a direct comparison of  $\varphi_2(r_w, 1)$  and  $\varphi_1(r_w, 1)$ .

$$\frac{\varphi_2(r_w, 1)}{\varphi_1(r_w, 1)} = \frac{[1 - 0.75 \cdot x]^{2r_w}}{[1 - x]^{r_w}} = \left( \frac{[1 - 0.75 \cdot x]^2}{1 - x} \right)^{r_w} \quad (27)$$

To conclude that  $\frac{\varphi_2(r_w, 1)}{\varphi_1(r_w, 1)} \leq 1$  as required, we note that given that  $\mathcal{H}$  has median 0, for any  $\phi_M \in [0, 1]$ ,  $\tilde{\mu} \leq \frac{1}{2}$  and  $\sigma \leq 1$ , so that  $x \leq \frac{1}{2}$ . Therefore,

$$\frac{\partial \frac{[1 - 0.75 \cdot x]^2}{1 - x}}{\partial x} = \frac{-0.5 + 0.75 \cdot x}{(1 - x)^2} \cdot (1 - 0.75 \cdot x) < 0 \quad (28)$$

and

$$\left( \frac{\varphi_2(r_w, 1)}{\varphi_1(r_w, 1)} \Big|_{x=0} \right) = 1 \quad (29)$$

imply that  $\frac{\varphi_2(r_w, 1)}{\varphi_1(r_w, 1)} \leq 1$  and  $\varphi_2(r_w, 1) \leq \varphi_1(r_w, 1)$  which concludes the proof of Step 1.

We now prove Step 2. We start by taking the derivative of

$$\ln(\varphi_\rho(r_w, 1)) = \rho \cdot r_w \cdot \ln[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)] \quad (30)$$

which with some algebra amounts to

$$\frac{\partial \ln(\varphi_\rho(r_w, 1))}{\partial \rho} = r_w \cdot \ln[1 - x \cdot (1 - 0.5^{\rho \cdot r}) / (0.5 \cdot \rho \cdot r)] + \quad (31)$$

$$+ \frac{[x \cdot \rho \cdot r_w \cdot \ln(0.5) \cdot 0.5^{\rho \cdot r} \cdot r + x \cdot r_w - x \cdot r_w \cdot 0.5^{\rho \cdot r}]}{0.5 \cdot \rho \cdot r - x + x \cdot 0.5^{\rho \cdot r}} \quad (32)$$

Thus, for any  $r_w > 0$ ,  $\frac{\partial \ln(\varphi_\rho(r_w, 1))}{\partial \rho} \geq 0$  whenever

$$\ln[1 - x \cdot [(1 - 0.5^\rho) / (0.5 \cdot \rho)]] + \frac{[x \cdot \rho \cdot \ln(0.5) \cdot 0.5^\rho + x - x \cdot 0.5^\rho]}{0.5 \cdot \rho - x + x \cdot 0.5^\rho} \geq 0 \quad (33)$$

In Step 2, we are interested in the sign of  $\frac{\partial \ln(\varphi_\rho(r_w, 1))}{\partial \rho}$  for large  $\rho$ . For any  $x$  and for any  $\rho \geq 2$ ,  $0.5 \cdot \rho - x + x \cdot 0.5^\rho > 0$ , so inequality (33) holds if and only if

$$(0.5 \cdot \rho - x + x \cdot 0.5^\rho) \cdot \ln[1 - x \cdot [(1 - 0.5^\rho) / (0.5 \cdot \rho)]] + x \cdot \rho \cdot \ln(0.5) \cdot 0.5^\rho + x - x \cdot 0.5^\rho \geq 0 \quad (34)$$

With some additional algebra we get that inequality (33) holds if and only if

$$\frac{2^\rho \{(0.5 \cdot \rho - x) \ln[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)] + x\}}{\rho} \geq \frac{x \cdot \ln[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)] + x \cdot \rho \cdot \ln(0.5) - x}{\rho} \quad (35)$$

where

$$\lim_{\rho \rightarrow \infty} \frac{x \cdot \ln[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)] + x \cdot \rho \cdot \ln(0.5) - x}{\rho} = x \cdot \ln(0.5) > -\infty \quad (36)$$

Consequently, a sufficient condition for inequality (33) to hold is

$$\lim_{\rho \rightarrow \infty} \frac{2^\rho \{(0.5 \cdot \rho - x) \ln[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)] + x\}}{\rho} = \infty \quad (37)$$

Also,

$$\begin{aligned} & \frac{2^\rho \{(0.5 \cdot \rho - x) \ln[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)] + x\}}{\rho} = \\ & = 2^\rho \left\{ \left(0.5 - \frac{x}{\rho}\right) \ln[1 - x \cdot [(1 - 0.5^\rho) / (0.5 \cdot \rho)]] + \frac{x}{\rho} \right\} \\ & \geq 2^\rho \left\{ \left(0.5 - \frac{x}{\rho}\right) \ln[1 - x \cdot [(1) / (0.5 \cdot \rho)]] + \frac{x}{\rho} \right\} \\ & = 2^{\rho-1} \left\{ \left(1 - \frac{2x}{\rho}\right) \ln \left[1 - \frac{2x}{\rho}\right] + \frac{2x}{\rho} \right\} \\ & = \frac{2^{\rho-1}}{\rho^2} \left\{ \rho(\rho - 2x) \ln \left[1 - \frac{2x}{\rho}\right] + 2x\rho \right\} \end{aligned}$$

and

$$\begin{aligned} \lim_{\rho \rightarrow \infty} \rho(\rho - 2x) \ln \left[1 - \frac{2x}{\rho}\right] + 2x\rho & = 2x^2 \\ \lim_{\rho \rightarrow \infty} \frac{2^{\rho-1}}{\rho^2} & = \infty \end{aligned}$$

As a result, condition (37) is satisfied for any  $x > 0$ , and inequality (33) holds for any  $x \geq 0$  as required. ■

**Proposition 2 - Proof.** We rely on Corollary 1 and prove the Proposition by showing that for every  $\gamma \in [0, 1]$ ,  $\tilde{\psi}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) \geq \tilde{\psi}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma)$ .

Since for all  $r \geq 1$ ,  $P(r, \theta_W^1) = P(r, \theta_W^2)$  and  $P(r, \theta_F^1) = P(r, \theta_F^2)$ , we have that  $\frac{\sum_{r=0}^{\infty} \theta_F^1(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W^1(r) \cdot r} = \frac{\sum_{r=0}^{\infty} \theta_F^2(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W^2(r) \cdot r}$  and also that for all  $r \geq 1$ ,  $\tilde{P}(r, \theta_F^1) = \tilde{P}(r, \theta_F^2)$ . Thus, for every  $\gamma \in [0, 1]$ ,

$$\sigma_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) = \sigma_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma); \tilde{\tau}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) = \tilde{\tau}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma); \text{ and } \tilde{\mu}_{\theta_W^1, \theta_F^1, M, \alpha} = \tilde{\mu}_{\theta_W^2, \theta_F^2, M, \alpha}$$

Then, by the definition of  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha}$  (expression 7), the difference in  $\theta_W$  implies that for every  $\gamma \in [0, 1]$ ,

$$\tilde{\psi}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) \geq \tilde{\psi}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma)$$

as required. ■

**Proposition 3 - Proof.** We rely on Corollary 1 and prove the Proposition by showing that for every  $\gamma \in [0, 1]$ ,  $\tilde{\psi}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) \geq \tilde{\psi}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma)$ .

The proof goes in 3 short steps. First, by the definition of MPS,  $\frac{\sum_{r=0}^{\infty} \theta_F^1(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W^1(r) \cdot r} = \frac{\sum_{r=0}^{\infty} \theta_F^2(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W^2(r) \cdot r}$ . Therefore, for every  $\gamma \in [0, 1]$ ,

$$\sigma_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) = \sigma_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma) \text{ and } \tilde{\mu}_{\theta_W^1, \theta_F^1, M, \alpha} = \tilde{\mu}_{\theta_W^2, \theta_F^2, M, \alpha}$$

Second, by the convexity and negative monotonicity of  $(1 - 0.5^r) / (0.5 \cdot r)$  in  $r$ , for every  $\gamma \in [0, 1]$ ,

$$\tilde{\tau}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) \geq \tilde{\tau}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma)$$

Finally,  $(1 - [1 - \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) + \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - \tilde{\mu}_{\theta_W, \theta_F, M, \alpha})]^{r_w})$  is increasing in  $\tilde{\tau}_{\theta_W, \theta_F, M, \alpha}$  and increasing and concave in  $r_w$ . Thus, for every  $\gamma \in [0, 1]$ ,

$$\tilde{\psi}_{\theta_W^1, \theta_F^1, M, \alpha}(\gamma) \geq \tilde{\psi}_{\theta_W^2, \theta_F^2, M, \alpha}(\gamma)$$

which completes the proof. ■

**Proposition 4 - Proof.** We rely on Corollary 1 and prove the Proposition by showing that for every  $\gamma \in [0, 1]$ ,  $\tilde{\psi}_{\theta_W, \theta_F, M^1, \alpha}(\gamma) \geq \tilde{\psi}_{\theta_W, \theta_F, M^2, \alpha}(\gamma)$ .

By definition, for every  $\gamma \in [0, 1]$ ,

$$\sigma_{\theta_W, \theta_F, M^1, \alpha}(\gamma) = \sigma_{\theta_W, \theta_F, M^2, \alpha}(\gamma); \tilde{\tau}_{\theta_W, \theta_F, M^1, \alpha}(\gamma) = \tilde{\tau}_{\theta_W, \theta_F, M^2, \alpha}(\gamma); \text{ and } \tilde{\mu}_{\theta_W, \theta_F, M^1, \alpha} \geq \tilde{\mu}_{\theta_W, \theta_F, M^2, \alpha}$$

Finally,  $(1 - [1 - \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) + \tilde{\tau}_{\theta_W, \theta_F, M, \alpha}(\gamma) \cdot (1 - \tilde{\mu}_{\theta_W, \theta_F, M, \alpha})]^{r_w})$  is decreasing in  $\tilde{\mu}_{\theta_W, \theta_F, M, \alpha}$ . Thus, for every  $\gamma \in [0, 1]$ ,

$$\tilde{\psi}_{\theta_W, \theta_F, M^1, \alpha}(\gamma) \geq \tilde{\psi}_{\theta_W, \theta_F, M^2, \alpha}(\gamma)$$

which completes the proof. ■

**Proposition 5 - Proof.** We rely on Corollary 1 and prove the Proposition by showing that: [1] if  $\alpha^2 < \underline{\alpha}$  then for every  $\gamma \in [0, 1]$ ,  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha^2}(\gamma) \geq \tilde{\psi}_{\theta_W, \theta_F, M, \alpha^1}(\gamma)$ , and [2] if  $\alpha^1 > \bar{\alpha}$  then for every  $\gamma \in [0, 1]$ ,  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha^1}(\gamma) \geq \tilde{\psi}_{\theta_W, \theta_F, M, \alpha^2}(\gamma)$ .

**Part 1:** Recall that  $\tilde{\mu}_{\theta_W, \theta_F, M, \alpha} = 1 - \mathcal{H}[\alpha \cdot (1 + \phi_M \cdot v) - 1]$  and that  $\mathcal{H}$  has the support  $[-v, v]$  where  $v \in (0, 1)$ . Let  $\alpha^2 \leq \frac{1-v}{1+\phi_M \cdot v}$ , thus  $\alpha^i \cdot (1 + \phi_M \cdot v) - 1 \leq -v$  for  $i = 1, 2$ , and  $\tilde{\mu}_{\theta_W, \theta_F, M, \alpha^1} = \tilde{\mu}_{\theta_W, \theta_F, M, \alpha^2} = 1$ . Recall that by definition, for every  $\gamma \in [0, 1]$ ,

$$\sigma_{\theta_W, \theta_F, M, \alpha^2}(\gamma) \geq \sigma_{\theta_W, \theta_F, M, \alpha^1}(\gamma) \text{ and } \tilde{\tau}_{\theta_W, \theta_F, M, \alpha^2}(\gamma) \geq \tilde{\tau}_{\theta_W, \theta_F, M, \alpha^1}(\gamma)$$

implying that for every  $\gamma \in [0, 1]$ ,  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha^2}(\gamma) \geq \tilde{\psi}_{\theta_W, \theta_F, M, \alpha^1}(\gamma)$ .

**Part 2:** Recall that  $\sigma_{\theta_W, \theta_F, M, \alpha}(\gamma_1) = \mathcal{D}\left(\alpha \cdot \pi_H + (1 - \alpha) \cdot \pi_L - \frac{\frac{1}{2} - \alpha \cdot \gamma_1}{1 - \gamma_1} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H\right)$  and that  $\mathcal{D}$  has the support  $[\underline{c}, \bar{c}]$ . Let

$$\alpha^1 \geq \frac{\bar{c} + \frac{1}{2} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H - \pi_L}{\pi_H - \pi_L}$$

thus for every  $\gamma \in [0, 1]$ ,  $\alpha^i \cdot \pi_H + (1 - \alpha^i) \cdot \pi_L - \frac{\frac{1}{2} - \alpha^i \cdot \gamma_1}{1 - \gamma_1} \cdot \frac{\sum_{r=0}^{\infty} \theta_F(r) \cdot r}{\sum_{r=0}^{\infty} \theta_W(r) \cdot r} \cdot \pi_H \geq \bar{c}$  for  $i = 1, 2$ , and

$$\sigma_{\theta_W, \theta_F, M, \alpha^2}(\gamma) = \sigma_{\theta_W, \theta_F, M, \alpha^1}(\gamma) = 1 \text{ and } \tilde{\tau}_{\theta_W, \theta_F, M, \alpha^2}(\gamma) = \tilde{\tau}_{\theta_W, \theta_F, M, \alpha^1}(\gamma) = \sum_{r_f=1}^{\infty} \tilde{P}(r_f, \theta_F) \cdot \frac{(1 - 0.5^{r_f})}{0.5 \cdot r_f}$$

Recall that by definition  $\tilde{\mu}_{\theta_W, \theta_F, M, \alpha^2} \leq \tilde{\mu}_{\theta_W, \theta_F, M, \alpha^1}$  implying that for every  $\gamma \in [0, 1]$ ,  $\tilde{\psi}_{\theta_W, \theta_F, M, \alpha^2}(\gamma) \leq \tilde{\psi}_{\theta_W, \theta_F, M, \alpha^1}(\gamma)$ . ■

## 10 References

Avery, Christopher, Christine Jolls, Richard A. Posner, and Alvin. E. Roth (2001), "The Market for Federal Judicial Law Clerks," *The University of Chicago Law Review*, Vol. 68(3), pp. 793-902.

Avery, Christopher, Christine Jolls, Richard A. Posner, and Alvin. E. Roth (2007), "The New Market for Federal Judicial Law Clerks," *The University of Chicago Law Review*, Vol. 74(2), pp. 447-486.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou (2006), "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 74, 1403-1417.

Bewley, Truman F. (1999), "Why Wages Don't Fall During a Recession," Harvard University Press.

Bramoullé, Yann, Martin D'Amours, and Rachel Kranton (2010), "Strategic Interaction and Networks," working paper.

Bulow, Jeremy and Jonathan Levin (2006), "Matching and Price Competition," *American Economic Review*, 96(3), pp. 652-668.

Calvo-Armengol, Antoni, and Matthew O. Jackson (2004), "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, Vol. 94(3), pp. 426-454.

Calvo-Armengol, Antoni, and Yves Zenou (2005), "Job matching, social network and word-of-mouth communication," *Journal of Urban Economics*, Vol. 57(3), pp. 500-522.

Damiano, Ettore, Hao Li, and Wing Suen (2005), "Unravelling of Dynamic Sorting," *Review of Economic Studies*, Vol. 72, pp. 1057-1078.

Fainmesser, Itay P., and David A. Goldberg (2010), "Bilateral and Community Enforcement in a Networked Market with Simple Strategies," working paper.

Fainmesser, Itay P. (2010), "Community Structure and Market Outcomes: a Repeated Games in Networks Approach," working paper.

Fréchette, Guillaume R., Alvin E. Roth, and Utku M. Ünver (2007), "Unraveling yields inefficient matchings: evidence from post-season college football bowls," *The RAND Journal of Economics*, Vol. 38(4), pp. 967-982.

Gale, David and Lloyd Shapley (1962), "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9-15.

Galeotti, Andrea, Sanjeev Goyal, Matthew O. Jackson, Fernando Vega-Redondo, and Leeat Yariv (2010), "Network Games," *Review of Economic Studies*, Vol. 77 (1), pp. 218-244.

Granovetter, Mark S. (1974), "Getting A Job: A Study of Contacts and Careers," Harvard University Press.

Greenhill, Catherine, Brendan D. McKay, and Xiaoji Wang (2006), "Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums," *Journal of Combinatorial Theory, Series A*, 113, 291-324.

Halaburda, Hanna W. (2010), "Unravelling in Two-Sided Matching Markets and Similarity of Preferences," *Games and Economic Behavior* 69(2), pp. 365-393.

Jackson, Matthew O., and Brian W. Rogers (2007), "Meeting Strangers and Friends of Friends: How Random are Socially Generated Networks?" *American Economic Review*, 97(3), pp. 890-915.

Kagel, John H., and Alvin E. Roth (2000), "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment," *Quarterly Journal of Economics*, Vol. 115(1), pp. 201-235.

Karlan, Dean, Tanya Rosenblat, Markus M. Mobius, and Adam G. Szeidl (2009), "Trust and Social Collateral," *Quarterly Journal of Economics*, 124(3), pp. 1307-1361.

Kojima, Fuhito (2007), "Matching and Price Competition: Comment," *American Economic Review*, 97(3), pp. 1027-1031.

Koopmans, Tjalling C., and Martin Beckmann (1957), "Assignment Problems and the Location of Economic Activities," *Econometrica*, Vol. 25(1), pp. 53-76.

Kozinski Alex (1991), "Confessions of a Bad Apple," *The Yale Law Journal*, Vol. 100(6), pp. 1707-1730.

Li, Hao, and Sherwin Rosen (1998), "Unraveling in Matching Markets," *American Economic Review*, Vol. 88(3), pp. 371-87.

Li, Hao, and Wing Suen (2000), "Risk Sharing, Sorting, and Early Contracting," *Journal of Political Economy*, Vol. 108(5), pp. 1058-1087.

Montgomery, James D. (1991), "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review*, Vol. 81(5), pp. 1407-18.

Niederle, Muriel, and Alvin E. Roth (2003), "Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and without a Centralized Match." *Journal of Political Economy*, 111(6), pp. 1342-1352.

Niederle, Muriel, and Alvin E. Roth (2005), "The Gastroenterology Fellowship Market: Should There Be a Match?," *American Economic Review*, 95(2), pp. 372-375.

Niederle, Muriel, and Alvin E. Roth (2009), "Market Culture: How Rules Governing Exploding Offers Affect Market Performance," *American Economic Journal: Microeconomics*, 1(2), pp. 199-219.

Niederle, Muriel, and Alvin E. Roth (2003), "Relationship Between Wages and Presence of a Match in Medical Fellowships," *Journal of the American Medical Association*, 290(9), pp. 1153-1154.

Ostrovsky, Michael, and Michael Schwarz (2010), "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, v. 2(2), pp. 34-63.

Roth, Alvin E. (1991), "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K.," *American Economic Review*, 81(3), pp. 415-440.

Roth, Alvin E. (2002), "The Economist as Engineer: Game Theory, Experimental Economics and Computation as Tools of Design Economics," *Econometrica*, 70(4), pp. 1341-1378.

Roth, Alvin E. (2008), "What have we learned from market design?" Hahn Lecture, *Economic Journal*, Vol. 118, pp. 285-310.

Roth, Alvin E., and Merilda Sotomayor (1990), "Two-Sided Matching," Cambridge University Press.

Roth, Alvin E., and Xiaolin Xing (1994), "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," *American Economic Review*, 84(4), pp. 992-1044.

Shapley, Lloyd, and Martin Shubik (1971), "The Assignment Game I: The Core," *International Journal of Game Theory*, Vol. 1, pp. 111-130.

Suen, Wing (2000), "A Competitive Theory of Equilibrium and Disequilibrium Unraveling in Two-Sided Matching," *RAND Journal of Economics*, 31(1), pp. 101-120.