Social Networks and Unraveling in Labor Markets*

Itay P. Fainmesser†

July 27, 2012

Abstract

This paper studies the phenomenon of early hiring in entry-level labor markets affected by social networks. We offer a model in which information is revealed over time. At first, workers have noisy information 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. Later on, 'hard' accurate verifiable information becomes available. 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, unraveling, entry-level labor markets, early hiring.

*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. The paper also benefited from the insightful feedback from two anonymous referees and an associate editor.

†Economics Department, Brown University, Providence RI 02906. Email: *Itay_Fainmesser@Brown.edu*

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. Nevertheless, law clerks in the US are often hired by judges as early as the fall of their first year in law school, and medical interns were in some years hired as early as two years prior to their graduation.

[41] documents several markets that exhibited a process of *unraveling* towards increasingly 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.[1] Recently, [33] documents 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 location of their training institutions, but rather that there is an inherent difference in the way hiring is conducted in different stages of 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.[2] 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.[3]

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 receive 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. This would potentially be done via workers' mentors who convey their impressions of the workers to firms in which they have contacts. At the second stage, 'hard'

[1] For direct evidence, see [16,34,35].
[2] See also [7,10,11,19,30].
[3] For previous theoretical work on unraveling see also [13,21,28,29,37,43].

verifiable (and accurate) information is revealed to the workers and can be credibly transmitted to all firms. Firms that use their connections (and hire promising candidates early) dilute the pool of high quality workers in the second stage. 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. Two-sided networks are especially suitable for describing market interactions in which the roles of the different sides are distinct (e.g. employers vs. potential employees). By studying the effect of changes to 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 connections to or deleting connections from the network; and [2] changes that correspond to changing the distribution of a fixed number of connections across workers and firms.

There are several ways to add links to a network. One way is to increase the *span* of the network - i.e., increase the number of workers and firms that have at least one connection. Another way is to increase the network's *density* - i.e., increase 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). The opposite is true for the case that the distribution of the number of connections across workers is more polarized.

We are also interested in answering 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 (either by introducing a centralized match or by modifying the market rules) leads to lesser unraveling. To this end, we offer a simple parameterization of the efficiency of a matching market 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 centralized and decentralized markets. In particular, this family of matching procedures incorporates most procedures that are studied in the market design literature (e.g., deferred acceptance, random dictator, top trading cycles, etc.).[4]

This paper is related to the literature on networks in economics. [10,11] study models of job search via personal connections and derive implications for inequality and unemployment. More broadly, there is a growing related literature on network games (e.g. [6,8]). In network games each player cares only about the actions taken by her neighbors. [18] suggests 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 simplifies the analysis significantly. This is possible due to recent graph theoretic results by [15] who study repeated games in large two-sided networks.

The large networks 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 [22] for an example used in the analysis of network formation). In particular, we provide an approximation for the number of workers hired early via the network. However, because 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

[4] We discuss the connection with known matching algorithms in section 8.1.

of unraveling. In particular, previous contributions focus on the heterogeneity of firms and candidates with respect to quality [28,29] and preferences [21], as well as on the size of the applicant pool [28]. We share with previous models the insurance element driving the unraveling process.

In the following section we motivate our analysis by reviewing evidence on patterns of unraveling in several entry-level labor markets. In sections 3 and 4 we present the model and derive 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 section 6 we derive comparative statics on unraveling with respect to the network structure and the matching procedure, and in section 7 we discuss the impact of unraveling on market outcomes and consider policy implications. Section 9 provides a discussion of important elements of the model including the robustness of our modeling assumptions, and section 10 offers concluding remarks.

2 Unraveling: prevalence and patterns

In this section we review evidence from different markets that exhibited unraveling. Much of the evidence is anecdotal. Nonetheless, it sheds light on the process of unraveling. Since we are interested in the determinants of unraveling, we also provide evidence from matching markets that did not exhibit unraveling.

Gastroenterology fellowships. Gastroenterology is a subspecialty of internal medicine. A typical gastroenterology fellow will have previously been employed for three years following medical school as a resident in internal medicine and then started a gastroenterology fellowship. The entry-level market for American gastroenterology fellowships was organized by a centralized clearinghouse from 1986 to 1996. Before, and after (until 2006), it has been conducted via a decentralized market in which appointment dates unraveled to well over a year before the beginning of employment. During those years, unraveling was wide spread. On the other hand, from 1986 until 1996, nearly all positions were filled through the match. This points to the effectiveness of a centralized clearinghouse in preventing unraveling. However, a centralized clearinghouse is not guaranteed to prevent unraveling. For example, in 1999, before the match was formally abolished, only 14 out of more than 300 positions participated in the match. [35]

suggests that simultaneous shifts in demand and supply shifted the expectations on both sides of the market and moved the system to an equilibrium with a high level of unraveling. This points to the possibility of coexistence of multiple equilibria, and suggests that coordination is important.

The special circumstances in the gastroenterology fellowships market provide additional insights with respect to mobility and wages. The insights with respect to wages are stark: in the gastroenterology fellowships market, [35] finds no effect of unraveling on wages.

Mobility corresponds to the tendency of workers to move between different geographical regions; low mobility characterizes fragmented markets. [33] finds that both before and after the years in which the centralized clearinghouse was used, gastroenterologists were less mobile and more likely to be employed by the hospital in which they were internal medicine residents than when the clearinghouse was in use. The same is true at the city and state level. We interpret the locality of early hiring to suggest that social networks are more active at early stages, when information on candidates is hard to come by. We also infer that the network structure is correlated with institutional affiliation, i.e., there are ties between gastroenterology departments and internal medicine departments at the same hospital. Finally, in the process of reinstating the gastroenterology fellowships match, market designers encountered resistance from several mid-tier gastroenterology departments that share hospitals with higher-tier internal medicine departments. This resistance suggests a belief that they are better off in an unraveling market. This also reinforces our conjecture that networks of connections play a role in the unraveling process.

Judicial clerks. Federal judicial clerkships represent an important point of entry to many of the most sought-after positions in the legal profession. Every year top students from elite law schools compete for positions with judges who can help them to land Supreme Court clerkships, plum teaching jobs, and competitive law firm positions. At the same time, federal judges depend heavily on their law clerks to aid them with their workload. Judicial clerks' wages and benefits are determined by federal law.

The unraveling in the market for judicial clerkships is a long outstanding issue and has been a source of disputes in the judicial system (see [3,4,26]). Despite the fact that early hiring is wide spread, in every year there is a substantial number of judges who do not attempt to hire early. Moreover, hiring is concentrated around a handful of top law schools (namely Chicago, Harvard, Michigan, Stanford, and Yale), and judges who are known to consistently hire early come from a subset of circuits.

Attempts to set the date of the hiring of clerks failed repeatedly, and anecdotal evidence suggests that unraveling takes a particular dynamic pattern. It is also often claimed that unraveling is triggered by judges from the 9th Circuit (California), and that the circuits on the East Coast abide more often with suggested deadlines. The reasons for this pattern 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 jumps-the-gun because it is less attractive to clerks due to lesser prestige of its positions (see also [3,4]). We interpret the claims to suggest that there are advantages to geographic proximity. This may be due to logistical considerations, or (social) communication networks between students (or their mentors) and judges. We also note that a connection between a judge and a law school may be reinforced by the fact that clerks are heavily involved in the search for their successors.

Additional markets. Every year in the US more than 20k medical students are matched with their first position. After periods of unraveling and reorganization, the market for medical residencies presently operates successfully by a centralized clearinghouse organized by the National Residency Matching Program (NRMP). In 2002, an anti-trust suit against the NRMP and numerous other defendants was brought by 16 law firms on behalf of 3 former residents seeking to represent the class of all former residents. The theory behind the suit is that a match inevitably holds down wages.[5] Nevertheless, consistent with the finding from the gastroenterology fellowships market, [34] finds that there is no difference in wages between medicine subspecialities that use a match and those that do not. The suit was dismissed on August 12, 2004 in an Opinion, Order & Judgment by Judge Paul L. Friedman.

Logistical constraints on interviewing impose a constraint on the number of early offers that firms are able to make in many labor markets. The constraint is even more explicit in the market for MBA graduates. First, one channel of unraveling is via first year summer internships, which are often followed by job offers. Second, in some schools (e.g., Harvard Business School)

[5] See [9,12,24,31] for theoretical work with different conclusions.

interviewing is restricted to school vacations and a formal "interviewing week." Firms that do not follow the rules are penalized and may be excluded from future recruiting events. Students may also be penalized.

We also note that not all entry-level labor markets are faced with a choice between a centralized clearinghouse and unraveling. The US market for junior faculty in economics is operating without a centralized clearinghouse and suffers little unraveling. One characteristic of the market is that interviews are conducted over a three-day period at a large conference, which most universities, large private companies, and job candidates attend. The conference is designed to provide coordination, reduce interviewing costs and increase the efficiency of the market. Before the conference, applications are sent and information regarding candidates is revealed to any institution to which a candidate applies. Earlier in a candidate's training, her mentors may already share information with their colleagues, and do so more often as the job market season approaches. Connections of reliable information sharing can be based on coauthoring, repeated interactions or other common history. Since high profile faculty are invited to give talks and participate in many academic events, one may suspect that the underlying network of connections is dense for students at highly ranked institutions, whose mentors are well connected.

Finally, three characteristics fit all of the labor markets described above. First, the training period for the candidates is long and involves acquiring new skills. During this time both the workers and their teachers learn about the workers' expected productivities. Much related, all of the above markets are for highly skilled workers. In such markets the number of high quality workers is often smaller than the market saturation level. This is clear in the markets for medical residents and for junior faculty in economics, in which the number of candidates is often smaller than the number of positions. However, even where this is not the case (e.g., judicial clerkships market), there is high variance in workers' abilities, and employers compete for the best workers. Third, it is not uncommon for employment offers to be open for a short duration, and it is widely acknowledged that acceptance of an employment offer is a binding commitment. For example, in the market for judicial clerks offers are sometimes open for less than 30 minutes, and there is little evidence of law students who renege on early acceptances of judges' offers.[6] Employment offers that are open for a limited duration and the acceptance

[6] See [4,36].

of which is considered binding are often called *exploding offers* in the market design literature.

3 A model of hiring and employment

In this section we present a simple model of employment and embed it in a two-stage hiring process. We defer the discussion of the assumptions imposed by the structure of the model to section 9.2.

There is a finite set of firms, $F \equiv \{1, 2, ..., n_f\}$, and a finite set of workers, $W \equiv \{1, 2, ..., n_w\}$. Each worker w can work for at most one firm and each firm f can employ at most one worker. A worker w is characterized by a productivity level $q_w \in \{L, H\}$. We assume that production depends only on the workers' productivities and normalize wages to zero. The payoff of firm f from employing worker w is captured by (1). In the next section we introduce heterogeneity in firms' payoffs.

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

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

$$u_w(f) = \upsilon_{wf}. \tag{2}$$

A firm that does not employ any worker, and an unemployed worker have a payoff of 0.

Prior to employment, there are two stages of hiring that correspond to stages in the workers' training denoted $\mathcal{S} = -1$ 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. It is assumed that prior to stage $\mathcal{S} = -1$ nature assigns each worker with productivity level $q_w = H$ or $q_w = L$ with equal probability, and preferences $\{\upsilon_{wf}\}_{f \in F}$ that are drawn independently from a distribution $\mathcal{H}$ with positive density in every point in the support $[\underline{\upsilon}, \overline{\upsilon}]$ for some $\overline{\upsilon} > 0$ and $\underline{\upsilon} \leq \overline{\upsilon}$. The realizations of q_w and $\{\upsilon_{wf}\}_{f \in F}$ are independent of each other and across workers. Workers and firms do not observe q_w and $\{\upsilon_{wf}\}_{f \in F}$ but learn about them over time as described below.

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

At $\mathcal{S} = -1$ information about own productivity and preferences is revealed to the workers and their mentors as follows: worker w observes a signal $\left(s_w, \{\upsilon_{wf}\}_{f \in F}\right)$ where $\{\upsilon_{wf}\}_{f \in F}$ are the worker's preferences and $s_w \in \{h, l\}$ is a noisy signal of the worker's productivity. The mentor of worker w observes only s_w and not the worker's preferences. If worker w is of high productivity ($q_w = H$) she receives a signal $s_w = h$ with probability $\alpha \in \left[\frac{1}{2}, 1\right]$ 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 receives the signal according to the reversed probabilities. The realizations of the signals are independent across workers.

The signal s_w consists of 'soft' information (in-class exam grades, performance as research assistant, 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 to firm f. Formally, 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 υ_{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 any worker w 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 $\mathcal{S} = 0$) 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 $\mathcal{D}(\underline{c}, \overline{c})$, independently across firms. The distribution $\mathcal{D}$ captures any unmodeled heterogeneity in the institutional flexibility of firms with respect to the timing of hiring. Finally, we assume that

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

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 at 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 $w \in W$.

3. Each worker w who receives at least one offer decides whether to accept any of the offers (at most one).

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 $\mathcal{S} = 0$.

3.1.1 Networks and information

If firm f is able to learn s_w (i.e., $f \in N_w$ 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 number of connections, are often consistent across networks and time. Thus, we assume that prior to 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).[7]

Formally, 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):

$$r_w = |N_w| \qquad (r_f = |N_f|) \,. \tag{4}$$

Let $\theta_W(r)$ ($\theta_F(r)$) be a *rational number* that captures the fraction of workers (firms) with degree r for $r = 0, 1, 2, \ldots\infty$. Given the number of workers n_w, and the degree distribution (θ_W and θ_F) there is a unique number of firms n_f that is consistent with any underlying graph G. Thus, we omit n_f and let $\mathcal{G}(n_w, \theta_W, \theta_F)$ be the set of networks consistent with $(n_w, \theta_W, \theta_F)$.[8]

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

[7] In section 5.1 we discuss further the motivation for the random process of the formation of the network.

[8] For any fixed θ_W and θ_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 [20]). All statements should be read as holding for n_w s.t. the aforementioned set is non-empty.

market (n_w and n_f); [2] her own links (N_w or N_f); and [3] the degree distribution (θ_W and θ_F). Thus, the Bayesian posterior of worker w (firm f) with degree r puts identical probability on all networks in $\mathcal{G}(n_w, \theta_W, \theta_F | r_w = r)$ ($\mathcal{G}(n_w, \theta_W, \theta_F | r_f = r)$). 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.

It is possible that some firms and workers have no connections ($r = 0$). To describe changes to the network structure that do not involve such firms and workers we use a truncated degree distribution P. Formally, let $P(r, \theta_W) = \theta_W(r) / (1 - \theta_W(0))$ be the fraction of workers with degree r as a fraction of the workers who have positive degrees. The definition extends immediately 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, in which the number of high quality workers is smaller than the market saturation level (see evidence in section 2). Recall that $\alpha \in \left[\frac{1}{2}, 1\right]$, and consider the following definition.

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

Definition 1 includes all markets in which there are not many more workers than open positions. In particular, we show later that assuming scarcity of high productivity workers guarantees that in asymptotically large markets, the number of high productivity workers who are still seeking employment at stage $\mathcal{S} = 0$ is (with probability 1) smaller than the number of firms that are looking for workers at $\mathcal{S} = 0$.

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

At $\mathcal{S} = 0$, workers graduate from their training and obtain a diploma and a track record that contain verifiable information revealing their true productivities $\{q_w\}_{w \in W}$. Thus, $\{q_w\}_{w \in W}$ can be credibly transmitted to all firms and are common knowledge. The preferences of any worker w $\left(\{\upsilon_{wf}\}_{f \in F}\right)$ are still her private information. In this environment the network is obsolete and we are in a familiar setup of one-to-one matching markets.[9]

[9] See [40] for a good introduction to matching theory.

The outcome of a matching market depends heavily on the underlying market rules (see also [38,39]). Since we are also interested in how changes in the post-graduation market affect early hiring, we consider a large class of *matching procedures* that covers both centralized and decentralized markets. 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 workers' and firms' preferences but not their identities, and that put positive probability only on *weakly stable matchings* - i.e., matchings in which 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. 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 a weakly stable matching 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 $\mathcal{S}=0$.

Lemma 1 shows that the requirement that a matching procedure is anonymous and guarantees a weakly stable matching, when applied to an asymptotically large market, 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 $\mathcal{S}=-1$, and can be varied exogenously by the choice of the particular matching procedure.

Let W_q^0 be the set of workers with productivity q who reach $\mathcal{S}=0$ unmatched and let F^0 be the set of firms that reach $\mathcal{S}=0$ unmatched. Given a network G, signal accuracy level α, and 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 $f\in F^0$. 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 *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 consistent with θ_W,θ_F and n_w. Assume that no worker who receives a low signal at $\mathcal{S}=-1$ is hired early (at $\mathcal{S}=-1$). Then,*

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

   (a) *For any worker $w\in W_L^0$, $E_{\widehat{G}(\theta_W,\theta_F,n_w),M,\alpha}[u_w|L,W_H^0,W_L^0,F^0]=0$.*

(b) *For any* $\xi > 0$

$$lim_{n_w \to \infty}\ sup_f\ \left| E_{\widehat{G}(\theta_W,\theta_F,n_w),M,\alpha}\left[\pi_f | W_H^0, W_L^0, F^0\right] - \frac{|W_H^0|}{|F^0|} \cdot \pi_H \right| < \xi.$$

Moreover, if all firms are acceptable to all workers ($\underline{v} \geq 0$)*, then for any firm* $f \in F^0$*,* $E_{\widehat{G}(\theta_W,\theta_F,n_w),M,\alpha}\left[\pi_f | W_H^0, W_L^0, F^0\right] = \frac{|W_H^0|}{|F^0|} \cdot \pi_H$.

2. *For any* $\phi \in \left[\max\left\{0, \frac{\underline{v}}{\overline{v}}\right\}, 1\right]$ *and* $\xi > 0$ *there exists an anonymous matching procedure* M *that guarantees a weakly stable matching such that*

$$lim_{n_w \to \infty}\ sup_w\ \left| E_{\widehat{G}(\theta_W,\theta_F,n_w),M,\alpha}\left[u_w | H, W_H^0, W_L^0, F^0\right] - \phi \cdot \overline{v} \right| < \xi.$$

We show later that no worker who receives a low signal in stage $\mathcal{S} = -1$ is ever hired early. 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 ϕ, and to establish that all of the market procedures that we focus on lead to (asymptotically) identical $E_{G,M,\alpha}\left[u_w | L\right]$ and $E_{G,M,\alpha}\left[\pi_f | W_H^0, W_L^0, F^0\right]$. The following definition offers a parameterization for the family of matching procedures that our analysis covers.

Definition 2 *A matching procedure* M ***is parameterized by*** $\phi_M \in [0, 1]$ *if:*
[1] M *is anonymous and guarantees a weakly 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 at* $\mathcal{S} = -1$ *is hired early (at* $\mathcal{S} = -1$*), the expected utility of a high productivity worker in the post-graduation market is asymptotically* $\phi_M \cdot \overline{v}$*. Namely,* $lim_{n_w \to \infty}\ sup_w \left| E_{G,M,\alpha}\left[u_w | H, W_H^0, W_L^0, F^0\right] - \phi_M \cdot \overline{v} \right| = 0$.

We interpret ϕ_M as a measure of the efficiency of the matching procedure. In section 8.1 we discuss further our notion of weak stability and the interpretation of ϕ_M. We also show that many of the centralized mechanisms studied in the literature generate matchings that are characterized by $\phi_M = 1$.

4 Bayesian equilibrium and ε-equilibrium

We now define the notions of Bayesian equilibrium and ε-equilibrium in our setup. 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. Let μ_w be the strategy of worker w, i.e.,

$\mu_w\left(s_w, \{\upsilon_{wf}\}_{f\in F}, N_w, \widetilde{F}_w\right)$ is the offer accepted by worker w who receives a signal s_w, has preferences that are captured by $\{\upsilon_{wf}\}_{f\in F}$, a set of firms connected to her N_w, and offers from every firm $f \in \widetilde{F}_w$ at stage $\mathcal{S} = -1$. Similarly, $\sigma_f\left(c, N_f, \widetilde{W}_h\right)$ is the probability distribution over workers to whom an early offer is made by firm f that has a cost of hiring early $c_f = c$, a set of workers connected to it N_f, and the knowledge that every worker $w \in \widetilde{W}_h \subseteq N_f$ received a signal $s_w = h$ ($\sigma_f(\cdot) = w_0$ implies that firm f does not make an offer at stage $\mathcal{S} = -1$). When it is clear from the context we let $\mu_w\left(\widetilde{F}_w\right) = \mu_w\left(s_w, \{\upsilon_{wf}\}_{f\in F}, N_w, \widetilde{F}_w\right)$ and $\sigma_f\left(\widetilde{W}_h\right) = \sigma_f\left(c_f, N_f, \widetilde{W}_h\right)$.

A family of firms' strategies that is 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 *We say that σ_f is a **label-free strategy** if $\sigma_f\left(c, N_f, \widetilde{W}_h\right)$ assigns identical probabilities to any w and w' for whom at least ONE of the following holds: [1] firm f knows that w and w' received high signals $\left(w, w' \in \widetilde{W}_h\right)$; [2] firm f knows that w and w' received low signals $\left(w, w' \in N_f \backslash \widetilde{W}_h\right)$; [3] firm f does not know what signals w and w' received $(w, w' \in W \backslash N_f)$.*

For a given network G, market procedure M, and signal accuracy α, let $\Pi_{G,M,\alpha}(f)$ be the expected payoff of firm f that employs strategy σ_f. Similarly, Let $U_{G,M,\alpha}(w)$ be the expected utility of worker w who employs strategy μ_w. We now define ε-equilibrium in our setup.

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

$$\Pi_{G(n_w,\theta_W,\theta_F),M,\alpha}\left(f|\sigma_f, c_f, N_f, \{\sigma_{f'}\}_{f'\in F\backslash\{f\}}, \{\mu_{w'}\}_{w'\in W},\right) \geq$$
$$\Pi_{G(n_w,\theta_W,\theta_F),M,\alpha}\left(f|\widetilde{\sigma}_f, c_f, N_f, \{\sigma_{f'}\}_{f'\in F\backslash\{f\}}, \{\mu_{w'}\}_{w'\in W},\right) - \varepsilon$$

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

$$U_{G(n_w,\theta_W,\theta_F),M,\alpha}\left(w|\mu_w, \{\upsilon_{wf}\}_{f\in F}, \{\mu_{w'}\}_{w'\in W\backslash\{w\}}, \{\sigma_{f'}\}_{f'\in F}\right) \geq$$
$$U_{G(n_w,\theta_W,\theta_F),M,\alpha}\left(w|\widetilde{\mu}_w, \{\upsilon_{wf}\}_{f\in F}, \{\mu_{w'}\}_{w'\in W\backslash\{w\}}, \{\sigma_{f'}\}_{f'\in F}\right) - \varepsilon$$

If $\varepsilon = 0$ the definition amounts to a Bayesian equilibrium. We now analyze firms' and workers' best response correspondences separately and show that they can be summarized using two random variables. If a firm f makes an offer at $\mathcal{S} = -1$ to a worker w who receives a signal $s_w = h$:

$$\Pi_{G,M,\alpha}\left(f|\widetilde{W}_h,\sigma_f\left(\widetilde{W}_h\right)=w \text{ for some } w\in W\right)=$$
$$Pr\left\{w \text{ accepts}\right\}\cdot\left\{Pr\left[q_w=H|N_f,\widetilde{W}_h\right]\cdot\pi_H+Pr\left[q_w=L|N_f,\widetilde{W}_h\right]\cdot\pi_L-c_f\right\}+$$
$$+Pr\left\{w \text{ rejects}\right\}\cdot E_{G,M,\alpha}\left[\pi_f\right] \tag{5}$$

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

$$\Pi_{G,M,\alpha}\left(f|\widetilde{W}_h,\sigma_f\left(\widetilde{W}_h\right)=w_0\right)=E_{G,M,\alpha}\left[\pi_f\right] \tag{6}$$

Note that $Pr\left\{w \text{ accepts}\right\}$, $Pr\left\{w \text{ rejects}\right\}$, and $E_{G,M,\alpha}\left[\pi_f\right]$ depend on $\left\{\sigma_{f'}\right\}_{f'\in F\setminus\{f\}}$ and $\left\{\mu_{w'}\right\}_{w'\in W}$. 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 who receives early job offers (at $\mathcal{S}=-1$) from a set of firms $\widetilde{F}_w\subseteq F$. If w accepts the offer of firm $f\in\widetilde{F}_w$ then

$$U_{G,M,\alpha}\left(w|\widetilde{F}_w,\mu_w\left(\cdot,\widetilde{F}_w\right)=f\right)=\upsilon_{wf} \tag{7}$$

and otherwise

$$U_{G,M,\alpha}\left(w|\widetilde{F}_w,\mu_w\left(\cdot,\widetilde{F}_w\right)=f_0\right)=Pr\left[q_w=H|s_w\right]\cdot E_{G,M,\alpha}\left[u_w|H\right]+Pr\left[q_w=L|s_w\right]\cdot E_{G,M,\alpha}\left[u_w|L\right] \tag{8}$$

Conditional on $\widetilde{W}_h$ $\left(\widetilde{F}_w\right)$ 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 $E_{G,M,\alpha}\left[\pi_f\right]$ $\left(E_{G,M,\alpha}\left[u_w|H\right]\right)$. Formally,

[1] Consider a firm f. Conditional on $\widetilde{W}_h\neq\emptyset$, firm f makes an offer at $\mathcal{S}=-1$ if and only if $c_f<\alpha\cdot\pi_H+(1-\alpha)\cdot\pi_L-E_{G(n_w,\theta_W,\theta_F),M,\alpha}\left[\pi_f\right]$. Given that c_f is drawn from $\mathcal{D}$, independently across firms, the ex-ante probability that a firm f with $\widetilde{W}_h\neq\emptyset$ makes an offer at $\mathcal{S}=-1$ is captured by

$$\sigma_{\theta_W,\theta_F,M,\alpha}\left(\cdot\right)=\mathcal{D}\left(\alpha\cdot\pi_H+(1-\alpha)\cdot\pi_L-E_{G(n_w,\theta_W,\theta_F),M,\alpha}\left[\pi_f\right]\right). \tag{9}$$

[2] Consider a worker w. Conditional on w receiving *exactly* one offer (from a firm f) at $\mathcal{S}=-1$, w accepts the offer if and only if $\upsilon_{wf}>\alpha\cdot E_{G(n_w,\theta_W,\theta_F),M,\alpha}\left[u_w|H\right]$. Given that υ_{wf} is drawn from $\mathcal{H}$, independent across firms and workers, the ex-ante probability that a worker with one early job offer accepts the offer is captured by

$$\mu_{G,M,\alpha}\left(\cdot\right)=1-\mathcal{H}\left(\alpha\cdot E_{G(n_w,\theta_W,\theta_F),M,\alpha}\left[u_w|H\right]\right) \tag{10}$$

Similarly, the ex-ante probability that a worker w who receives exactly m offers at $\mathcal{S}=-1$ accepts an offer is $1-\left(1-\mu_{G,M,\alpha}\left(\cdot\right)\right)^{m}$, and conditional on accepting an offer, w accepts the offer that maximizes $\left\{\upsilon_{wf}\right\}_{f\in\widetilde{F}_w}$.

5 Equilibrium existence and structure

In this section we show that equilibrium 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 ε-equilibria exist for arbitrary low ε. 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 $\mathcal{S}=-1$ for any given realization of workers' signals. Consider a mapping

$$\gamma_0=\psi_{G,M,\alpha}\left(\widehat{\gamma}_0\right)=\psi_{G,M,\alpha}\left(\left\{\sigma_{G,M,\alpha}\left(w|\widehat{\gamma}_0\right)\right\}_{w\in W},\left\{\mu_{G,M,\alpha}\left(f|\widehat{\gamma}_0\right)\right\}_{f\in F}\right) \tag{11}$$

that maps from the expectations of workers and firms with respect to hiring at $\mathcal{S}=-1$ to the random variable that captures the same outcome at $\mathcal{S}=-1$. Any γ_0' such that $\gamma_0'=\psi_{G,M,\alpha}\left(\gamma_0'\right)$ captures an equilibrium level of hiring at $\mathcal{S}=-1$, and any equilibrium with γ_0^* corresponds to a fixed point $\gamma_0^*=\psi_{G,M,\alpha}\left(\gamma_0^*\right)$. However, $\psi_{G,M,\alpha}\left(\gamma\right)$ is an extremely complex object that depends on the entire network structure and is stochastic even for a fixed network structure and fixed realization of workers' signals. Thus, instead of characterizing $\psi_{G,M,\alpha}$ directly, we establish that in asymptotically large networks and for any $\langle\theta_W,\theta_F,M,\alpha\rangle$, the outcome of $\psi_{G(n_w,\theta_W,\theta_F),M,\alpha}\left(\gamma_0\right)$ converges to a well behaved function with a deterministic output. We are then able to characterize ε-equilibria for arbitrarily small ε.

We now illustrate our analysis using a simple **hypothetical** exercise: consider a network in which there is 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 of $r_{f'}$ for any $f'\in N_w$) and captured by $\widetilde{P}\left(r,\theta_F\right)=\frac{P(r,\theta_F)\cdot r}{\overline{r}_f}$, where $P\left(r,\theta_F\right)$ is the fraction of firms with degree r out of the firms that have positive degrees (as defined in section 3.1.1), and $\overline{r}_f=E_P\left[r_f|r_f\geq 1\right]=\sum_{r\in\{1,2,\dots\infty\}}P\left(r,\theta_F\right)\cdot r$. Suppose further (**hypothetically**) that for every worker w, $E_{G(n_w,\theta_W,\theta_F),M,\alpha}\left[u_w|H\right]=\phi_M\cdot\overline{\upsilon}$ and that for any firm f, $E_{G(n_w,\theta_W,\theta_F),M,\alpha}\left[\pi_f\right]=\frac{\left|W_H^0\right|}{\left|F^0\right|}\cdot\pi_H$ (recall that W_H^0 is the set of workers with high productivity that reach $\mathcal{S}=0$ unmatched and F^0 is the set of firms that reach $\mathcal{S}=0$ unmatched). Now consider a worker w chosen u.a.r. from the workers who received a high signal ($s_w = h$). Then, the probability that w receives a job offer from a firm f that is chosen u.a.r. from N_w is

$$\begin{aligned}\widetilde{\tau}_{\theta_W,\theta_F,M,\alpha}(\gamma) &= \sum_{r_f=1}^{\infty}\widetilde{P}_F(r_f,\theta_F)\left[\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma)\sum_{m=0}^{r_f-1}\binom{r_f-1}{m}0.5^m 0.5^{r_f-m-1}\frac{1}{m+1}\right] \\ &= \sum_{r_f=1}^{\infty}\widetilde{P}(r_f,\theta_F)\cdot\left[\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma)\cdot\left(1-0.5^{r_f}\right)/\left(0.5\cdot r_f\right)\right]\end{aligned} \tag{12}$$

where

$$\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma)=\mathcal{D}\left(\alpha\cdot\pi_H+(1-\alpha)\cdot\pi_L-\frac{\frac{1}{2}-\alpha\cdot\gamma}{\frac{\sum_{r=0}^{\infty}\theta_W(r)\cdot r}{\sum_{r=0}^{\infty}\theta_F(r)\cdot r}-\gamma}\cdot\pi_H\right). \tag{13}$$

To see why, note that $\binom{r-1}{m}0.5^m0.5^{r-m-1}$ is the probability that there are m other $s_w = h$ workers connected to firm f conditional on $|N_f| = r$. Finally, $\frac{1}{m+1}$ is the conditional probability that f makes the offer to w. The derivation of $\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma)$ is available in the appendix.

Given that the realizations of the signals, the offers received and the acceptance of offers are independent across workers, the expected number of workers hired in stage $\mathcal{S}=-1$ is captured by

$$\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma) = \frac{1}{2}\cdot(1-\theta_W(0))\cdot\sum_{r_w=1}^{\infty}P(r_w,\theta_W)\cdot\left(1-\left[1-\widetilde{\tau}_{\theta_W,\theta_F,M,\alpha}(\gamma)+\widetilde{\tau}_{\theta_W,\theta_F,M,\alpha}(\gamma)\cdot\left(1-\widetilde{\mu}_{\theta_W,\theta_F,M,\alpha}\right)\right]^{r_w}\right) \tag{14}$$

where

$$\widetilde{\mu}_{\theta_W,\theta_F,M,\alpha}=1-\mathcal{H}\left[\alpha\cdot\phi_M\cdot\overline{v}\right]. \tag{15}$$

The calculation of $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma)$ 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 for any $\gamma\in[0,1]$, $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma)$ is deterministic and well behaved. Thus, establishing that for any $\widehat{\gamma}\in[0,1]$ and $\epsilon>0$,

$$lim_{n_w\to\infty}\ Pr\left(\left|\psi_{G(n_w,\theta_W,\theta_F),M,\alpha}(\widehat{\gamma})-\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\widehat{\gamma})\right|<\epsilon\right)=1 \tag{16}$$

would allow 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 any* $\varepsilon > 0$ *there exists* $n_w \in \mathbb{Z}^+$ *such that for every* $n'_w > n_w$ *there exists an* ε*-equilibrium with* $G(n_w, \theta_W, \theta_F), M, \alpha,$ *and* γ^* *in which* $\Pr\left(\psi_{G(n_w,\theta_W,\theta_F),M,\alpha}(\gamma^*) \in [\gamma^* - \varepsilon, \gamma^* + \varepsilon]\right) > 1 - \varepsilon.$

The set of 0-equilibria corresponds to the set of fixed points of $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma)$.

Theorem 1 *Let* θ_W, θ_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* γ *is a* 0*-equilibrium with* $\langle \theta_W, \theta_F, M, \alpha \rangle$*. Assume further that firms employ label-free strategies. Then,* $\gamma^* = \widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma^*)$ *if and only if* γ^* *is a* 0*-equilibrium with* $\langle \theta_W, \theta_F, M, \alpha \rangle$*.*

In Lemma 4 in the Appendix, we derive a limit closed form expression for $\psi_{G,M,\alpha}(\widehat{\gamma})$ without formally expressing $\psi_{G,M,\alpha}(\widehat{\gamma})$ for any finite network G. To this end, we rely on a recent graph theoretic result by [15] 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 workers hired at $\mathcal{S} = -1$ converges to the mean and that (16) 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]* θ_W, θ_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. This is not surprising given that firms' actions are strategic complements. More specifically, multiplicity is determined by the properties of $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma)$ which correspond to the properties of θ_W, θ_F, M, α, and $\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma)$. For example, if $\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma)$ 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 that hire early (at $\mathcal{S} = -1$) create expectations that they will do so in subsequent years, when new cohorts of workers graduate. The expectations that some firms hire early trigger preemptive actions by other firms, and the fraction of workers hired early increases from year to year. We also provide a definition of *greater unraveling* that is suitable for environments with multiple equilibria.

Consider an entry level labor market in which a new cohort of workers graduates every year and firms hire new workers every year. For simplicity, assume that all cohorts of workers are of the same size n_w. The hiring process in each year follows stages $\mathcal{S} = -1, 0$ that are described above. The network of connections between firms and a new cohort of workers is drawn u.a.r. from all of the possible networks with the same θ_W and θ_F. To motivate the changes to the network structure, recall that in every year different workers are searching for jobs. Given that each worker can have a different mix of mentors, the network defined by the connections of the different subsets of mentors that each worker has is different across cohorts, even if the connections of each mentor are the same. The specific assumptions that the network is selected u.a.r. and that θ_W and θ_F stay exactly the same can be relaxed.

In this dynamic environment, consider an unraveling process driven by firms that maximize their immediate payoffs within each year (workers maximize their overall utility, but each worker is in the market for one year). At year $t = 0$, some fraction γ^0 of the workers (all with $s_w = h$) are hired at $\mathcal{S} = -1$. At each year $t > 0$, each firm and worker best respond to the outcome of the previous year's hiring cycle. Let γ^t denote the fraction of workers (all with $s_w = h$) hired at stage $\mathcal{S} = -1$ in year 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 (16) and Theorem 1, the limiting dynamic process (when n_w is arbitrary large) is captured by $\gamma^t = \widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma^{t-1})$ and any rest point of the system corresponds to a static 0-equilibrium. Note that $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}$ 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 of some well studied labor markets (see [41] and references therein) and suggestive evidence from the experimental market design literature (see also [23]). Moreover, considering the process of unraveling as dynamic lends itself to a natural way of capturing the notion that some markets generate *greater unraveling* then other markets.

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

$$lim_{n_w\to\infty}\ lim_{t\to\infty}\left[\gamma^t\left(n_w, \langle\cdot\rangle^1, \gamma^0\right) - \gamma^t\left(n_w, \langle\cdot\rangle^2, \gamma^0\right)\right] \geq 0 \tag{17}$$

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

There are several systematic 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*. This distinction turns out to be important; increasing a network's density and increasing a network's span have significantly different effects on unraveling.

Network span. Adding connections by increasing the span of the network generates greater unraveling.

Proposition 1 *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* $P(r, \theta_F^1) =$

$P\left(r, \theta_F^2\right)$. *Assume that all firms employ label-free strategies. Then,* $\left\langle \theta_W^1, \theta_F^1, M, \alpha \right\rangle$ *generates greater unraveling than* $\left\langle \theta_W^2, \theta_F^2, M, \alpha \right\rangle$.

The intuition for Proposition 1 is straightforward: increasing the number of connected workers and firms increases the number of offers at $\mathcal{S} = -1$. The connection between Proposition 1 and the evidence in section 2 is also suggestive. In the markets for gastroenterology fellowships, every gastroenterology department has a connection with at least one internal medicine department (at the same hospital). In the market for judicial clerkships, judges as well as law students from top schools are well connected – essentially all judges have colleagues in top universities and are also connected via their current clerks.

Network density. Increasing the networks' density leads to *greater* unraveling if the initial network is sparse and the increase is small. On the other hand, if the initial network is dense, increasing the network's density leads to *lesser* unraveling.

For degree distributions θ_W, θ_F, let $\theta_W^\rho, \theta_F^\rho$ be 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 call ρ the *density multiplier* of θ_W, θ_F.[10]

Proposition 2 *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,*

[1] for any M *and* α, *there exists* $\overline{r} = \overline{r}(M, \alpha) \in \mathbb{Z}^+$ *such that if min* $\left\{\rho^L \cdot r | r \geq 1, \theta_F(r) > 0\right\} > \overline{r}$, *then* $\left\langle \theta_W^{\rho^L}, \theta_F^{\rho^L}, M, \alpha \right\rangle$ *generates greater unraveling than* $\left\langle \theta_W^{\rho^H}, \theta_F^{\rho^H}, M, \alpha \right\rangle$; *and*

[2] for any M *and* α *such that* $\mathcal{H}\left[\alpha \cdot \phi_M \cdot \overline{\upsilon}\right] > \frac{1}{3}$ *there exists* $\underline{r} = \underline{r}(M, \alpha) \in \mathbb{Z}^+$ *such that if max* $\left\{\rho^H \cdot r | \theta_F(r) > 0\right\} < \underline{r}$ *then* $\left\langle \theta_W^{\rho^H}, \theta_F^{\rho^H}, M, \alpha \right\rangle$ *generates greater unraveling than* $\left\langle \theta_W^{\rho^L}, \theta_F^{\rho^L}, M, \alpha \right\rangle$.

Part 1 of Proposition 2 is surprising partly because it establishes that unraveling is *not* maximized when the network is very dense or in the well studied complete market. In particular, 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.

To better understand the forces behind the non monotonicity captured by Proposition 2, it is useful to think of stage $\mathcal{S} = -1$ as divided to two steps: in the first step, the network is

[10] Our claims apply to ρ such that $\rho \cdot r \in \mathbb{Z}^+$ for every r in the support of θ_F and θ_W.

pruned by eliminating [1] all workers who receive a low signal ($s_w = l$); and [2] all firms that would not make early offers at stage $\mathcal{S} = -1$ even if they are connected to workers who receive high signals. In the second step the induced (pruned) network is analyzed – each firm that has at least one connection makes an offer to one of the workers connected to it at random.

Now suppose that we start with two networks of different initial densities. At the end of the first step the denser network remains (weakly) denser and has a (weakly) larger span. To see why, consider the networks in figure 1. If both firms plan to make early offers and both workers receive high signals, then the networks remain the same at the end of the first step and the network in figure 1b is denser. On the other hand, if only firm f_2 plans to make an early offer and only worker w_1 receives a high signal then the network induced by pruning figure 1b has a larger span than the network induced by pruning figure 1a.

In the second step, we compare across networks that may be different in span and density, and include only firms that plan to make early offers and workers who receive high signals. The effect of the difference in span is straightforward; a larger span leads to greater unraveling. The effect of differences in density in the second step goes in the opposite direction: greater density implies that less workers are hired early. To see why, suppose that at the end of the first step (the pruning step) we are left with the networks in figure 1. For simplicity, assume further that both workers will accept an early offer from either firm rather than stay unmatched until $\mathcal{S} = 0$. Then, in figure 1a both workers are hired early with probability 1, whereas in figure 1b there is a positive probability $\left(= \frac{1}{2}\right)$ that both firms make an offer to the same worker and only one worker is hired early. More generally, additional links reduce the ability of the network to act as a coordination devise for determining the worker to whom each firm makes an offer. To see the impact of (lack of) coordination, recall that in the second step each firm makes at most one early offer. Making this offer uniformly at random to a high signal worker on a complete network is equivalent to making it on a network drawn uniformly at random from those in which all firms have a single link to a high signal worker. However, a network drawn uniformly at random from those in which all firms have a single link to a high signal worker has a low expected span (many high signal workers are not connected to any firm). This implies that due to the lack of coordination, the complete network performs as if it were a network with a lower span. We therefore conclude that an increase in density in the second step is equivalent to an indirect decrease in effective span.

To aggregate these two forces and derive the non monotonic conclusion captured by Proposition 2, note that if a network is very dense, the direct effect of the pruning step on the span of the network is minimal – with high probability all of the workers who receive a high signal and all of the firms who are interested in making early offers have at least one connection in the induced network after the first step. Therefore, if two networks with high densities are compared, the induced networks at the end of the first step differ mainly in density, and the lower density network exhibits greater unraveling. On the other hand, if two networks that have low densities are compared, the induced networks at the end of the first step differ greatly in span, and the lower density network exhibits lesser unraveling.

figure 1a **figure 1b**

Figure 1

To summarize, increasing a network's density increases the probability that a worker w with $s_w = h$ and a firm f with a low enough c_f are connected. This leads to an increase in the number of offers made at $\mathcal{S} = -1$. On the other hand, increasing the network's density also increases the probability that fewer workers receive a larger portion of the offers.

6.2 The network structure – redistribution of links

We focus on redistribution of links across workers and firms that have degree of at least 1, i.e., with no effect on the span of the network (which is covered by Proposition 1). Proposition 3 predicts greater unraveling in markets in which some firms have many connections and others have few, than in markets in which firms' degrees are similar. On the other hand, lesser unraveling is predicted in markets in which some workers are much more connected than others, than in markets in which workers' degrees are similar. While only suggestive, these predictions fit the discussion of the evidence in section 2. The US market for junior faculty in economics exhibits very little unraveling. It is also often claimed that Ph.D. candidates from top universities have many connections via their high profile mentors whereas students from lower tier institutions are less connected. The discussion of the role of connectivity in triggering unraveling in the market for judicial clerkships fits the suggested patterns as well: the debate about the role of the 9th Circuit (California) in triggering unraveling highlights the claim that unraveling may be triggered because some judges are less connected than their peers to students from the highest ranked law schools.

Proposition 3 *Consider two regular environments* $\left\langle \theta_W^1, \theta_F^1, M, \alpha \right\rangle$ *and* $\left\langle \theta_W^2, \theta_F^2, M, \alpha \right\rangle$ *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* $P\left(\cdot, \theta_W^2\right)$ *is a Mean Preserving Spread (MPS) of* $P\left(\cdot, \theta_W^1\right)$ *and* $P\left(\cdot, \theta_F^1\right)$ *is a MPS of* $P\left(\cdot, \theta_F^2\right)$, *then* $\left\langle \theta_W^1, \theta_F^1, M, \alpha \right\rangle$ *generates greater unraveling than* $\left\langle \theta_W^2, \theta_F^2, M, \alpha \right\rangle$.

We begin with the intuition for the effect of the spread of workers' degrees. For any equilibrium strategies, the probability that a worker receives at least one acceptable early offer is increasing in her degree. However, adding a link to a worker with a high degree makes a smaller difference in the aforementioned probability than adding a link to a worker with a lower degree. More generally, the probability that any worker w receives at least one acceptable offer at $\mathcal{S} = -1$ is increasing and concave in r_w which implies the required result. Another way of reaching the same conclusion is by considering the role of the network in facilitating coordination between firms. Recall that firms do not know the degrees of workers connected to them. Therefore, firms cannot condition their offers on the degrees of workers, and the distribution of offers across workers is skewed towards highly connected workers. Since each worker can accept at most one offer, a skewed distribution of early offers leads to low levels of early hiring.

Now consider the effect of the spread of firms' degrees. For any equilibrium strategies, the probability that a worker w receives an early offer from a firm f (that is connected to her) is decreasing in the degree of f. However, if firm f has a high degree, one additional link does not make a large difference in the aforementioned probability, while if f has a low degree, one additional link has a larger effect. More generally, the probability that a worker w receives an offer from a firm f that is connected to her is decreasing and convex in r_f which implies the required result. Additional insight can be gained by examining the coordination provided by the network in the distribution of offers. At first glance the effect of an increased spread of firms' degrees is not conclusive; some firms become more connected and can cause more situations in which a single worker receives multiple offers. However, as the spread of firms' degrees increases, the *number* of firms who have large degrees decreases. For example, consider the networks in figure 2 and suppose that all workers receive high signals and all firms make early offers. In the network in figure 2b, if f_1 makes an offer to w_2 or to w_3, then there will be exactly one worker who receives two offers and exactly one worker who receives no offer (w_1). This occurs with probability $\frac{2}{3}$ (if f_1 makes an offer to w_1 then each worker receives exactly one offer). On the other hand, in the network in figure 2a if f_1 makes an offer to w_2, then there will be exactly one worker who receives two offers and exactly one worker who receives no offer. This occurs with probability $\frac{1}{2}$. However, even if f_1 makes an offer to w_1 (which occurs with probability $\frac{1}{2}$), there is still a positive probability that one worker receives multiple offers and another receives none: with probability $\frac{1}{2}$, f_2 makes an offer to w_3. Therefore, in the network in figure 2a with probability $\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$ only two workers receive early offers and in the network in figure 2b the equivalent probability is $\frac{2}{3}$. The reason is that in figure 2b more of the firms are coordinated, and any miscoordination must come from the actions of f_1.

Figure 2

6.3 Market rules and the accuracy of early information

In this section we study the effect of changes in the post-graduation matching procedure (captured by ϕ_M), as well as the effect of changes in the accuracy of early signals of workers productivities (i.e., changes in α).

The post-graduation matching procedure (ϕ_M). In line with the evidence reviewed in section 2 regarding the connection between the efficiency of the post-graduation market (centralized or decentralized) and the level of unraveling, the following result shows that an increase in ϕ_M leads to lesser unraveling. This connection is also recognized in empirical and experimental work in market design (e.g. [23]), providing additional motivation for the design of efficient clearinghouses.

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 interesting to note that if $\alpha \geq \frac{\underline{v}}{\overline{v}}$ then $\langle \theta_W, \theta_F, M^1, \alpha \rangle$ generates greater unraveling than $\langle \theta_W, \theta_F, M^2, \alpha \rangle$ whereas if $\alpha < \frac{\underline{v}}{\overline{v}}$ then $\langle \theta_W, \theta_F, M^1, \alpha \rangle$ and $\langle \theta_W, \theta_F, M^2, \alpha \rangle$ generate the same level of unraveling. To see why, note that if $\alpha < \frac{\underline{v}}{\overline{v}}$ workers who receive at least one offer at stage $\mathcal{S} = -1$ always accept one of the offers that they receive (independent of ϕ_M). On the other hand, if $\alpha \geq \frac{\underline{v}}{\overline{v}}$ then for some ϕ_M workers reject early offers from some firms and their rejection threshold increases in ϕ_M.

The accuracy of early signals (α)**.** If α is higher, then 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 $\mathcal{S} = -1$. These two forces result in a non monotonic relationship between α and the unraveling level in the market: moderate levels of accuracy of early signals generate greater unraveling than high or low (extreme) levels.

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.*

[1] If $\underline{v} > \frac{1}{2} \cdot \overline{v}$ *then there exist* $\underline{\alpha} \in \left(\frac{1}{2}, 1\right)$ *such that if* $\alpha^2 \leq \underline{\alpha}$ *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] There exist $\overline{\alpha} \in \left(\frac{1}{2}, 1\right)$ *such that if* $\alpha^1 \geq \overline{\alpha}$ *then* $\langle \theta_W, \theta_F, M, \alpha^1 \rangle$ *generates greater unraveling than* $\langle \theta_W, \theta_F, M, \alpha^2 \rangle$*.*

Consider a worker w who receives at least one early offer. If all firms are acceptable to a worker w, and if α is very small, then w accepts one of the offers with probability 1. Thus, a small increase in α 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 α 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 α only decreases the probability that worker w accepts, and $\widetilde{\mu}_{\theta_W, \theta_F, M, \alpha^2} \leq \widetilde{\mu}_{\theta_W, \theta_F, M, \alpha^1}$.

7 Welfare

Welfare analysis of two-sided matching markets is subtle due to the 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 *Greater unraveling leads to higher aggregate utility of connected workers* $\left(\sum_{w\in\{w\in W|r_w\geq 1\}} u_w\right)$, *and to lower aggregate payoffs of firms* $\left(\sum_{f\in F} \pi_f\right)$.

The first part captures the insurance that unraveling provides to connected workers, whereas the second part captures the hiring of low productivity workers that is caused by the early hiring and that reduces aggregate firms' profits. The anecdotal evidence described in section 2 provides similar insights into the winners and losers of unraveling. For example, the idea that unraveling makes some workers better off, and that a centralized match can hinder unraveling (Proposition 4) helps to account for the observed resistance to the centralized match by a small group of medical residents. In the same spirit, the medical residencies match as well as the gastroenterology fellowships match were coordinated efforts of firms. On the other hand, in the gastroenterology fellowships and in the judicial clerkships markets some firms (potentially well connected ones) objected to policies that may stop unraveling.

Evaluating aggregate welfare changes in the economy as a whole requires further assumptions. For example, if firms do not hire low productivity workers at $\mathcal{S} = 0$ because there is a large supply of 'medium productivity' means of production (e.g., machines), then unraveling decreases overall productivity in the market and may also decrease aggregate welfare. On the other hand, consider the case that production using low productivity workers is more socially efficient than alternative modes of production (excluding high productivity workers), and that firms do not hire low productivity workers at $\mathcal{S} = 0$ due to wage rigidity in the market. Then, unraveling increases overall productivity, and with quasi-linear utilities unraveling also increases aggregate welfare.[11]

[11] I thank an anonymous referee for suggesting the second example.

8 Policy implications

8.1 Centralized clearinghouse

For a large range of parameters the equilibrium exhibits a *tipping point* structure with multiple stable equilibria. In such environments cooperation is of first order. A centralized clearinghouse or an "interviewing conference" (see section 2) provide effective means of coordination. Moreover, Proposition 4 shows that a centralized clearinghouse may also serve to prevent unraveling if it improves the efficiency of the post-graduation market (increasing ϕ). In this section, we emphasize the latter point following two steps: first, we show that when the post-graduation market consists of a centralized clearinghouse, any of the matching procedures commonly proposed by the market design literature is guaranteed to generate a matching parameterized by $\phi = 1$. Then, we argue that decentralized marketplaces generally facilitate an efficiency level characterized by $\phi < 1$.

The following Corollary provides the connection between ϕ and a stronger notion of stability studied in the market design literature – the additional requirement beyond weak stability is that there is no firm-worker pair who prefer being matched to each other rather than to their assigned partner. For clarity, we call this more restrictive stability notion **strict stability**.

Corollary 3 *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 consistent with θ_W, θ_F and n_w. Assume that no worker who receives a low signal in stage $\mathcal{S} = -1$ is hired early. Then, for any anonymous matching procedure M that guarantees a* ***strictly stable*** *matching, and for any $\xi > 0$*

$$lim_{n_w \to \infty}\ sup_w \left| E_{\widehat{G}(\theta_W, \theta_F, n_w), M, \alpha} \left[u_w | H, W_H^0, W_L^0, F^0 \right] - \overline{\upsilon} \right| < \xi.$$

To see why, assume that a matching procedure M is anonymous and puts positive probability only on strictly stable matchings. Now assume by contradiction that the outcome is such that the expected utility of a high productivity worker is strictly below $\overline{\upsilon} - \xi$ for some $\xi > 0$. By the assumption of scarcity of high productivity workers and Lemma 2, there are firms that remain unmatched. In particular, the number of unmatched firms increases proportionally with the size of the market. Consequently, in large markets, for any worker w and $\xi > 0$, the probability that there is an unmatched firm f such that $\upsilon_{wf} > \overline{\upsilon} - \xi$ is (asymptotically) 1. This contradicts the assumption that the matching procedure puts positive probability only on strictly stable matchings.

An interesting implication of Corollary 3 is that many of the algorithms used in the design of centralized clearinghouses can be parameterized by $\phi = 1$ and are therefore generating lesser unraveling than the equivalent decentralized markets.

Corollary 4 *The Gale-Shapley worker proposing deferred acceptance algorithm [17], the random serial dictator algorithm for high productivity workers (see also [1]), the top trading cycle algorithm with initial random assignments of firms to high productivity workers (see [2]), and the so-called "Boston mechanism" with preferences submitted by high productivity workers are all anonymous and in our setup all generate stable matchings. Therefore, they can be parametrized by* $\phi = 1$.

Notably, the deferred acceptance algorithm generates strictly stable matchings in more general environments, whereas the top trading cycle algorithm, the random serial dictator algorithm, and the Boston mechanism generate stable matchings in our setup due to the firms' indifference between any two high productivity workers.

In contrast with the centralized clearinghouse environment, and following [32] we consider the notion of strict stability to be inconsistent with many decentralized markets.[12] Yet even in a decentralized environment, we find it reasonable to expect that: [1] an unmatched worker eventually finds an unmatched firm if such a firm exists (and vice versa); and [2] a worker who prefers to stay unmatched rather than being matched with a certain firm will not end up matched to that firm (and vice versa). Thus, our assumptions allow our model to incorporate decentralized environments in addition to centralized ones. For example, a decentralized market in which firms and workers meet at random and match (or not) upon their first meeting is covered by our model and can generate an expected utility that is much lower than $\overline{\upsilon}$ (but never negative) for any high productivity worker who reaches $\mathcal{S} = 0$ unmatched (equivalent to a matching procedure parameterized by lower values of ϕ).

[12] [32] studies the stability properties of matchings that are generated by decentralized matching markets and finds that in the presence of market frictions and preference uncertainty, strong assumptions on the richness of the economy have to be made in order for decentralized markets to generate stable outcomes in equilibrium.

8.2 Worker- and firm-driven unraveling, and "exploding offers"

In our model, firms make "exploding offers" (i.e. offers that expire at the end of stage $\mathcal{S} = -1$). This is a necessary condition for unraveling to occur. Otherwise, firms would be concerned with having their early offers held until the formal market at stage $\mathcal{S} = 0$, and could not benefit from early hiring. A simple-to-state policy rules out offers that are open for only a short period of time. However, legal consideration may deem this infeasible. An alternative policy is to make the acceptances of early offers not binding. Legally, this is already the case – any worker has the right to resign. However, market norms and repeated games considerations often enforce early agreements and therefore facilitate unraveling (see also [36]).

More generally, in evaluating any policy, one must keep in mind that the incentives of workers and firms may be conflicting. This raises an important question: should policies be directed at preventing early offers or early acceptances? Our model offers a distinction between worker- and firm-driven unraveling, which should be handled differently.[13] To be precise, 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 training institutions can be divided into two time segments. In early training α is small and workers would accept any job offer to insure themselves against unemployment. As a result, $\sigma_{\theta_W,\theta_F,M,\alpha}(\gamma) \in [0,1]$ whereas $\widetilde{\mu}_{\theta_W,\theta_F,M,\alpha} = 1$. This *worker-driven unraveling* can be influenced by changes to firms' incentives, as the incentives of workers to contract early are too strong to be affected. Closer to graduation, α is large, and firms try to hire any high signal worker, whereas workers decline less desirable job offers (for all $\gamma \in [0,1]$, $\sigma_{\theta_W,\theta_F,M,\alpha}(\gamma) = 1$ whereas $\widetilde{\mu}_{\theta_W,\theta_F,M,\alpha} \in [0,1]$). This *firm-driven unraveling* can be most effectively influenced by policies that affect workers' incentives.

8.3 Network structure

The paper provides comparative statics with respect to the effect of the network structure on unraveling. The design of policies that utilize these comparative statics requires a good model of network formation or a notion of network stability. In the context of information sharing networks regarding the qualities of job candidates, [14,15] propose a model of repeated games in two-sided networks and explore the structure of networks that allow for truthful information

[13] The distinction between firm- and worker-driven unraveling was suggested in [29] that focuses on the different qualities of the firms that trigger the unraveling process in each of the types of unraveling.

to flow via connections in the network. Their results suggest that the structure of the network can be affected by changes in the observability of the outcomes of interactions in the market. For example, not revealing the identity of the recommender to competing firms may impose a restriction that the two-sided network be more sparse.

The current paper also challenges some previous misconceptions. For example, Proposition 2, and to some extent also Proposition 3, show that making communication of the early signal more costly may not reduce unraveling.

9 Discussion

9.1 (In)finite economies and networks

Even ignoring the underlying topology, solving for equilibrium behavior and outcomes in matching markets under various market rules is often an intractable exercise. Partly for this reason much of the matching literature focuses on equilibria in dominant strategies. In the context of college admissions, [5] suggests in that solving for equilibrium behavior when there is a continuum of agents on one side of the market is a much more tractable problem.

The introduction of networks poses additional complications. For example, computing the expected number of workers hired early is a cumbersome operation. This is true even for highly stylized network structures, agents' strategies, and market procedures. Moreover, solving for the expected number of workers hired is not enough. In order to pin down equilibrium behavior, the entire distribution of the number of workers hired early is needed.

In this paper, we derive results for ε-equilibria (for arbitrarily small ε) in asymptotically large finite networks. Solving for ε-equilibria has the flavor of a continuum analysis because exact equilibrium is only guaranteed to emerge when the size of the economy goes to infinity. In what follows we review and evaluate the simplifications provided by the focus on large finite networks and by the 0-equilibrium solution concept.

Focusing on large networks is done for the following reasons:

[1] If a network is sufficiently large, a strong random component in the selection of the network guarantees that the network exhibits no degree correlation.

[2] Firms' costs of hiring early as well as workers' preferences and productivities are determined by random processes. In large economies, the law of large numbers kicks-in, and predetermined

smooth realization functions emerge.

[3] Combining 1. and 2., the fraction of workers hired early is a deterministic function of firms' and workers' strategies. Thus, at stage $\mathcal{S} = -1$, firms and workers ignore their private information in evaluating their expected future payoff from reaching the post-graduation market unmatched.

While 2. is standard in the economics literature, 1. relies on graph theoretic results by [15]. Notably, from a graph theoretic perspective, the requirement that the network is large can be relaxed if we are willing to restrict attention to specific families of networks. One such family of networks includes all semi-regular networks – networks in which all firms have the same degree and all workers have the same degree (but not necessarily the same degree as firms). In semi-regular networks the degree correlation is zero by definition. Relaxing 2. is trickier; even in semi-regular networks, the fraction of workers hired early is a complicated random function of firms' and workers' strategies. One way around it is by considering a large network that consists of many copies of the same small network.

The 0-equilibrium solution concept guarantees that firms do not lose more than ε (for arbitrary small ε) by choosing the ε-equilibrium action. We also know that the fractions of firms and workers who are required to make non-optimal decisions is zero in asymptotically large economies. However, the distribution of firms' early hiring costs is continuous, so we cannot rule out that there is a marginal firm which is indifferent whether it makes or does not make an offer in an ε-equilibrium. Such a firm may be assigned a non-optimal action relative to an exact equilibrium in any finite economy. To determine whether an exact equilibrium exists for finite but large networks, one needs to know something about the speed of convergence of the range of firms' early hiring costs that require a deviation from optimal action in the ε-equilibrium. This may depends on the distributions of firms' early hiring costs ($\mathcal{D}$). Deriving formally such a result requires a closed form expression for $\psi_{G(n_w,\theta_W,\theta_F),M,\alpha}(\gamma)$. Focusing on 0-equilibria provides a more tractable analysis.

In what follows we present two examples of the analysis of unraveling in simple finite economies. We use the examples to replicate two of the comparative statics of the infinite economies. Example 1 considers changes to the span of the network, whereas in example 2 we consider changes in the network's density. For simplicity, in both examples we restrict attention to environment in which the unraveling is worker-driven, i.e. workers accept early offers with probability 1.

Example 1 *The network in figure 3b has a higher span than the network in figure 3a. In both networks if worker w_1 receives a low signal, then firm f_1 will not find it profitable to make an early offer (at $\mathcal{S} = -1$). Similarly, in both networks if w_1 receives a high signal and is hired early by firm f_1 then the expected payoff of f_1 is $\alpha\pi_H + (1-\alpha)\,\pi_L - c_f$. However, the expected payoff of f_1 from not extending an early offer to a high signal w_1 depends on the network structure.*

Figure 3

In the network in figure 3a, if w_1 receives a high signal and is not hired early by firm f_1 then a simple counting exercise shows that the expected payoff of f_1 is $\left(\frac{1}{3}\alpha + \frac{1}{6}\right)\pi_H$. Therefore, f_1 makes an early offer to a high signal w_1 if and only if $\alpha\pi_H + (1-\alpha)\,\pi_L - c_f \geq \left(\frac{1}{3}\alpha + \frac{1}{6}\right)\pi_H$. Now consider the network in figure 3b and suppose that conditional on w_2 receiving a high signal, firm f_2 makes an early offer with ex-ante probability σ. Then, if w_1 receives a high signal and is not hired early by firm f_1 then the expected payoff of f_1 is $\left(\frac{1}{3}\alpha + \frac{1}{6} - \frac{1}{12}\sigma\alpha\right)\pi_H$. As a result, f_1 makes an early offer to a high signal w_1 if and only if $\alpha\pi_H + (1-\alpha)\,\pi_L - c_f \geq \left(\frac{1}{3}\alpha + \frac{1}{6} - \frac{1}{12}\sigma\alpha\right)\pi_H$. Plugging in a distribution $\mathcal{D}$ (for c_f) we are able to solve for an equilibrium strategy σ^*. More important, for any $\sigma > 0$, $\left(\frac{1}{3}\alpha + \frac{1}{6} - \frac{1}{12}\sigma\alpha\right)\pi_H < \left(\frac{1}{3}\alpha + \frac{1}{6}\right)\pi_H$, so consistent with Proposition 1 the incentives of firm f_1 to make early offers is greater in the network with the larger span.

Measuring the effect of an increase in density on unravelling in finite networks is less straightforward. In a small network, a firm that is connected to more than one worker may have incentives to act differently if one of its connected workers receives a high signal than if two of its connected workers receive high signals. In large networks this effect is eliminated. However, studying large finite network poses the difficulties reviewed above. To this end, in the following example we perform the following restricted equilibrium analysis in finite networks: holding constant the strategies of all but one firm, we evaluate the incentives of a firm to make an early offer.

Example 2 *Consider the networks in figure 4 and suppose that firms f_2 and f_3 follow the same strategy: conditional on being connected to at least one worker who receives a high early signal, make an early offer with probability σ to one of the high signal workers (if more than one such worker exists, choose a worker to make the offer to uniformly at random). Similar to the previous example, in both networks if worker w_1 receives a low signal, then firm f_1 does not make an early offer (at $\mathcal{S} = -1$). Also, in both networks, if w_1 receives a high signal and is hired early by firm f_1 then the expected payoff of f_1 is $\alpha\pi_H + (1-\alpha)\pi_L - c_f$. We now compare the expected payoff of f_1 from not extending an early offer to a high signal w_1 in both networks.*

Figure 4

Let $\pi(k_f, k_h, k_l)$ be the expected payoff of a firm that reaches unmatched a post-graduation market ($\mathcal{S} = 0$) with k_f firms, k_h workers with high early signals, and k_l workers with low early signals. Then, in the network in figure 4a, if f_1 does not hire early a high signal w_1 then f_1 has expected payoff of $\frac{1}{4}\left[\sigma^2\pi(2,1,0) + 2\sigma(1-\sigma)\pi(3,2,0) + (1-\sigma)^2\pi(4,3,0)\right] + \frac{1}{2}\left[\sigma\pi(3,1,1) + (1-\sigma)\pi(4,2,1)\right] + \frac{1}{4}\pi(4,1,2)$. The corresponding value for the network in figure 4b is $\frac{1}{4}\left[\sigma^2\left[\frac{1}{2}\pi(2,1,0) + \frac{1}{2}\pi(3,2,0)\right] + 2\sigma(1-\sigma)\pi(3,2,0) + (1-\sigma)^2\pi(4,3,0)\right] + {}$ $+\frac{1}{2}\left[(\sigma^2 + 2\sigma(1-\sigma))\pi(3,1,1) + (1-\sigma)^2\pi(4,2,1)\right] + \frac{1}{4}\pi(4,1,2)$. Therefore, the probability that f_1 makes an early offer in the network in figure 4a is higher than in the network in figure 4b if $\frac{1}{2}\sigma^2(\pi(3,2,0) - \pi(2,1,0)) + 2(\sigma^2 - \sigma)[\pi(4,2,1) - \pi(3,1,1)] > 0$. Whether the inequality holds depends on the parameters of the model.[14] For example, in environments in which f_2 and f_3 have strong incentives to make early offers (high σ) the inequality holds and the less dense network in figure 4a generates higher probability with which f_1 makes early offers. The opposite is true in environments in which f_2 and f_3 have weak incentives to make early offers (low σ).

[14] Note that $\pi(3,2,0) - \pi(2,1,0) > 0$ and $\pi(4,2,1) - \pi(3,1,1) > 0$.

9.2 Modeling decisions

Several restrictions are imposed by the structure of the model. The goal of this section is twofold: [1] relate our key modeling decisions to the corresponding activity rules in markets that motivate this paper, and [2] discuss the robustness of our results to alternative assumptions.

Our model abstracts from the wage determination process. The discussion of the role of wages in the analysis of matching markets and of unraveling is not new. Models of matching markets can be analyzed using the *assignment model* [25,42] where wages are a part of the clearing mechanism, or using the *marriage model* [17] where wages are assumed out. We note that allowing for wage heterogeneity across worker-firm pairs does not affect our analysis per se – much of the wage heterogeneity can be incorporated into preferences. The substantiative assumption is that wages do not vary with the timing of hiring. This assumption is supported by evidence from the markets motivating this paper that is reviewed in section 2. For example, wages in the market for judicial clerkships are regulated, and evidence reviewed from the gastroenterology fellowships market and from the market for medical residencies suggests that wages do not vary significantly with the timing of hiring or with the matching procedure used. Perhaps for that reason, [29] that analyzes unraveling using the assignment model in the context of college admissions, admits that the assignment model analysis *"applies with a greater force to assignment markets in which payments transfers are explicitly negotiated"*.

More important, our results are robust to the introduction of endogenous wages. To see that, normalize the minimal wage to zero. We now take any wage determination rule in stage $\mathcal{S} = 0$ as given and consider a wage determination rule for the early offers. Assume that when firms make early offers a worker who receives at least one offer bargains on the wage with the offering firms before deciding which offer to accept (if any). Assume further that the bargaining process is efficient, i.e., if there is a positive wage that is desireable both to the worker and to one of the offering firms, then the worker accepts one of her offers, and conditional on accepting an offer, the worker always accepts the offer from the most preferred firm that made her an offer. Assume further that the bargaining process is label-free (depends only on the firms' hiring costs and the worker preferences). Then, the analysis of the endogenous wages model follows directly from the analysis of the model studied in this paper with small modifications to the probability that a worker accepts an early offer (μ) and to the probability that a firm makes an early offer (σ). More specifically, the modified μ is higher than the μ in our model and so wage bargaining generates weakly higher levels of unraveling. On the other hand, wage bargaining has no direct effect on σ, which is affected only indirectly via the increase in the number of workers hired early (due to the increase in μ).

Second, in our model each firm hires at most one worker. This is realistic in several of the markets motivating the paper. For example, in the market for judicial clerks, many judges hire only one clerk per year. An alternative, more general model would also consider firms with multiple job openings. We note that as long as workers are substitutes, allowing firms to hire multiple workers does not alter our analysis of the post-graduation market (stage $\mathcal{S} = 0$). The analysis of the early hiring (stage $\mathcal{S} = -1$) with the possibility of multiple hires depends on the number of early offers that firms are able to make. If each firm can make one early offer, our analysis goes through with only minor changes. Allowing firms to make a number of offers that is not much larger than the number of openings that they have introduces an additional layer of complexity but does not change our results qualitatively.

A related assumption is that in stage $\mathcal{S} = -1$ each firm can make at most one offer. The single offer assumption is a simplification of the idea that the number of offers that each firm can make early is bounded. This is motivated by the observation that making offers and waiting for a response takes time. In particular, recall the discussion in section 2 about the labor market for MBA graduates, in which an early offer requires either a summer internship or at least an interview. The restrictions dictated by schools on the timing of interviews and the constraint posed by the need for a summer internship limit the number of early offers that a firm is able to make. Additional examples are provided in [41]. In some cases, such as in the labor market for MBA graduates, the constraint is exogenous, whereas in other examples the limiting factor is that workers hold offers until the deadline imposed by the firms.

We now consider alternatives to the assumption that each firm makes at most one early offer. The closest extension is to assume a fixed and exogenous upper limit on the number of early offers each firm can make. If this upper limit is binding, i.e., it is small enough relative to the degrees of firms in the network, then we do not expect any qualitative change to our results. A second alternative is a large (i.e., on the order of magnitude of a firm's degree) upper limit on the number of offers that a firm can make. The complete analysis under this assumption is cumbersome and we cannot rule out some changes to our results with respect to the effects of changes in the degree distribution (Propositions 2 and 3). However, the non monotonicity of Proposition 2 does not completely go away. To see that recall our discussion of the non monotonicity after Proposition 2, and consider stage $\mathcal{S} = -1$ as divided to two steps: in the first step, the network is pruned by eliminating [1] all workers who receive low signals ($s_w = l$); and [2] all firms that would not make early offers at $\mathcal{S} = -1$ even if they are connected to workers who receive high signals. In the second step, the induced (pruned) network is analyzed – each firm that has at least one connection makes an offer to one of the workers connected to it at random. Now suppose that we start with two networks of different initial densities. At the end of the first step, the denser network remains (weakly) denser and has a (weakly) larger span. In the second step we compare across networks that may differ in span and density, and that include only firms that make early offers and workers who receive high signals. The effect of the difference in span is independent of the assumption on the number of early offers that each firm can make – in the network that has a larger span more workers are hired early. On the other hand, with many early offers the effect of the difference in density is complex. Nonetheless, the following example demonstrates that increased density in the pruned network can decrease unraveling, at least for some range of network densities – reminiscent of the corresponding effect when each firm makes at most one early offer.

Example 3 *Suppose that at the end of the first step (the pruning step) we are left with the networks in figure 5. For simplicity, assume further that any worker will accept immediately an early offer from any of the firms. Then, in figure 5a all three workers are hired early with probability 1, whereas in figure 5b there is a positive probability that only two workers are hired, e.g. if f_1 makes its first offer to w_2 and f_2 makes its first offer to w_3.*

Figure 5

A third approach calls for a more general market mechanism to operate in stage $\mathcal{S} = -1$, subject to the constraint that firms can be matched early only with workers who are connected to them. The general mechanism approach poses many technical difficulties because there is no known mathematical framework to analyze the number of matchings generated by a general matching algorithm that is constrained by a network structure, even for simple random networks. For now we make the following observations. First, our result with respect to the effects of changes in the span of the network (Proposition 1), as well as our results with respect to changes in the information accuracy and in the matching procedure (Propositions 5 and 4 respectively) are not qualitatively sensitive to the matching procedure used in stage $\mathcal{S} = -1$. On the other hand, our result with respect to the effect of changes in the density of the network (Proposition 2) is sensitive to the elimination of frictions in the hiring process in stage $\mathcal{S} = -1$. In particular, for an arbitrary matching procedure at $\mathcal{S} = -1$, adding links to the network (without increasing the span) can generate more or less early hiring depending on the matching procedure used. For example, consider a matching mechanism that always picks a maximal match subject to the constraint that only connected firm and worker are matched. Then, increasing a network's density could not decrease (and generally would increase) the number of worker-firm pairs that are matched.

Finally, our analysis focuses on the role of information flows from workers to firms. We study two-sided networks of connections between workers and firms and do not consider connections between firms or between workers. This is not to say that no such connections exist. However, same type connections must have a different role – it is generally not in the best interest of a firm to inform a competitor about the quality of a worker that it intends to hire. One role of connections between firms can be to coordinate on which firm makes an offer to which of the workers. Such collusive strategies can take many forms and are in general very complex. Moreover, given that unraveling by competing firms is payoff reducing to a firm, collusive strategies may be difficult to enforce when the network of inter-firm connections is very dense. On the other hand, when the network of inter-firm connections is sparse, a firm might agree to coordinate with the firms connected to it as a part of a long term interaction. This is because the effect of the local coordination will not affect overall unraveling in a large network. We consider the topic of coordination between firms in the hiring process and its effect on unraveling to be an interesting topic for future research.[15] Connections between workers may also be of interest. If a connection between workers captures having a mentor in common, the mentor can signal different qualities to different firms, thus "coordinating" the firms to make non overlapping offers to her mentees. A mentor may be incentivized to do so, because a firm may prefer to make an early offer to a worker based on information from such a mentor – the mentee is more likely to have fewer competing offers and more likely to accept the firm's early offer. The consequences of such coordination are similar to the consequences of direct coordination across firms, yet it alleviates the incentive problems that may hinder coordination between firms.

10 Conclusion

This paper tackles 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 means 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 depends 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: [1] increasing the span of the network leads to greater unraveling; [2] increasing a network's density leads to greater unraveling if the network has low density, whereas increasing a network's density leads to lesser unraveling if the network has high density; [3] increasing the variance of workers' degrees leads to lesser unraveling; and [4] increasing the variance of firms' degrees leads to greater

[15] The consequences of coordination across firms in the context of a one-period market is considered in [27]: firms may interview similar or different subsets of workers, thus generating different two-sided networks between firms and workers in which a firm can make an offer only to a worker connected to it.

unraveling. Moreover, we show that improving the design of the post-graduation market by improving the expected quality of the matching between workers and firms leads to lesser unraveling.

11 Appendix

11.1 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 *(Due to [40]) For a set of workers W' and a set of firms F', a* ***one-to-one matching*** *is a function $\mathcal{M}: W' \cup F' \to W' \cup F' \cup \{w_0, f_0\}$ 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 $\Delta\left(\overline{\mathcal{M}}(W', F')\right)$ be the set of all probability distributions on elements of $\overline{\mathcal{M}}(W', F')$. A* ***matching procedure*** *is a function $M: \mathcal{P}(W') \times \mathcal{P}(F') \to \Delta\left(\overline{\mathcal{M}}(W', F')\right)$.*

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

Definition 10 *Let $U \equiv \left(u_{w_1}, u_{w_2}, ..., u_{w_{|W'|}}\right)$ and $U' \equiv \left(u'_{w_1}, u'_{w_2}, ..., u'_{w_{|W'|}}\right)$ 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 $U'' \equiv \left(u''_{w_1}, u''_{w_2}, ..., u''_{w_{|W'|}}\right)$ 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 $f \in F \backslash \{f', f''\}$ 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:*[16]

[16] 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. Because firms' payoff functions are identical at stage $\mathcal{S}=0$, the outcome of an anonymous matching procedure will depend only on the workers' preferences as captured by the definition.

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

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

Definition 11 *A matching procedure* ***guarantees a weakly 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).*

11.2 Derivation of $\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma)$

Consider firms' best responses as captured by (9) and suppose further (**hypothetically**) that for any firm f, $E_{G(n_w,\theta_W,\theta_F),M,\alpha}[\pi_f] = \frac{|W_H^0|}{|F^0|} \cdot \pi_H$. Then, (9) becomes

$$\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma) = \mathcal{D}\left(\alpha \cdot \pi_H + (1-\alpha) \cdot \pi_L - E\left(\frac{|W_H^0|}{|F^0|}\right) \cdot \pi_H\right). \tag{18}$$

Moreover,

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

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, (18) can be rewritten as

$$\widetilde{\sigma}_{\theta_W,\theta_F,M,\alpha}(\gamma) = \mathcal{D}\left(\alpha \cdot \pi_H + (1-\alpha) \cdot \pi_L - \frac{\frac{1}{2} - \alpha \cdot \gamma}{\frac{\sum_{r=0}^{\infty} \theta_W(r) \cdot r}{\sum_{r=0}^{\infty} \theta_F(r) \cdot r} - \gamma} \cdot \pi_H\right)$$

11.3 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* θ_W, θ_F *and* n_w. *Assume further that no worker who receives a low signal at* $\mathcal{S} = -1$ *is hired early. 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 at* $\mathcal{S} = -1$ *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 at* $\mathcal{S} = -1$. *Then, there exist* $\zeta > 0$ *such that*

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

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 at $\mathcal{S}=-1$, 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 at $\mathcal{S}=-1$ AND are revealed to be 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 at $\mathcal{S}=-1$. 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 $\mathcal{S}=-1$ (as a proportion of n_w). Finally, let $p_f^{\widehat{G}(\theta_W,\theta_F,n_w)}(\cdot)=\frac{n_f^{\widehat{G}(\theta_W,\theta_F,n_w)}(0,\alpha)}{n_f}$ be the r.v. that captures the proportion of firms that are not matched before $\mathcal{S}=0$. Then,

$$p_w^{0,\widehat{G}(\theta_W,\theta_F,n_w)}(H,0)\leq p_w^{0,n_w}(l,H)+p_w^{-1,n_w}(h)-\frac{\gamma^{\widehat{G}(\theta_W,\theta_F,n_w)}(\cdot)}{n_w} \tag{19}$$

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 $\mathcal{S}=-1$. This equals the proportion of workers who receive high signal, are of high productivity, and are not hired at $\mathcal{S}=-1$, only if all of the high signal workers who are not hired at $\mathcal{S}=-1$ are also of high productivity.

By the strong law of large numbers

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

Thus, inequality (19) implies that,

$$Pr\left(lim_{n_w\to\infty}\left(p_w^{0,\widehat{G}(\theta_W,\theta_F,n_w)}(H,0)-\left[\frac{1-\alpha}{2}+\frac{1}{2}-\frac{\gamma^{\widehat{G}(\theta_W,\theta_F,n_w)}}{n_w}\right]\right)\leq 0\right)=1 \tag{20}$$

By the assumption of scarcity of high productivity workers for every $\left\langle \widehat{G}(n_w,\theta_W,\theta_F),\alpha\right\rangle$ there exists $\eta>1$ such that $n_f>\eta\cdot\left(1-\frac{\alpha}{2}\right)\cdot n_w$. Consequently,

$$n_f\cdot p_f^{\widehat{G}(\theta_W,\theta_F,n_w)}(\cdot)=n_f-\gamma^{\widehat{G}(\theta_W,\theta_F,n_w)}>\eta\cdot\left(1-\frac{\alpha}{2}\right)\cdot n_w-\gamma^{\widehat{G}(\theta_W,\theta_F,n_w)} \tag{21}$$

Combining (20) and (21) and some algebra yields that there exists $\eta>1$ such that,

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

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 proofs of part 1-(a) and the second claim in part 1-(b) of the Lemma are immediate from the definitions above. The proof of the first claim in part 1-(b) follows from Lemma 2 – if there are more firms than high quality workers at $\mathcal{S}=0$, then independence across workers' preferences implies that in an asymptotically large market all high quality workers are matched in any weakly stable matching, and that firms have identical expected payoffs given any anonymous matching procedure. The proof for part 2 of the Lemma is as follows:

Let $\digamma\left(w, F^0, \xi, \phi\right) \equiv \left\{f \in F^0 \middle| \; |\upsilon_{wf} - \phi \cdot \overline{\upsilon}| < \xi\right\}$ and denote the empty set by $\varnothing$. For given $\phi \in \left[\max\left\{0, \frac{\underline{v}}{\overline{v}}\right\}, 1\right]$, and $\xi > 0$ consider the following algorithm:

Let $\mathcal{W} = W_H^0$ and $\mathcal{F} = F^0$
While $\mathcal{W} \neq \varnothing$ and $\mathcal{F} \neq \varnothing$
  Select $w \in \mathcal{W}$ uniformly at random
    If $\digamma\left(w, \mathcal{F}, \xi, \phi\right) \neq \varnothing$
      Pick a firm $f \in \digamma\left(w, \mathcal{F}, \xi, \phi\right)$ uniformly at random
      Match w to f
      Let $\mathcal{W} = \mathcal{W} \backslash w$ and $\mathcal{F} = \mathcal{F} \backslash f$
    Otherwise
      Pick a firm $f \in \mathcal{F}$ uniformly at random
      Match w to f
      Let $\mathcal{W} = \mathcal{W} \backslash w$ and $\mathcal{F} = \mathcal{F} \backslash f$

The algorithm matches a firm and a worker at every iteration and therefore always stops when either $\mathcal{W} \neq \varnothing$ or $\mathcal{F} \neq \varnothing$ and provides a weakly stable matching. The probability distribution over the outcomes of the algorithm is a matching procedure $M\left(W^0, F^0\right)$. 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 $\widehat{G}\left(\theta_W, \theta_F, n_w\right)$,

$$lim_{n_w \to \infty} Pr\left[\digamma\left(w, \mathcal{F}, \xi, \phi\right) \neq \varnothing\right] = 1$$

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

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

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 $\widetilde{w} \in \mathcal{W}$ 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

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

or

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

To complete the proof, fix $\zeta > 0$ and $\xi > 0$, and consider a randomly selected worker $\widetilde{w}$ and a set $\widetilde{F}(n_w)$ 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 $\widetilde{f} \in \widetilde{F}(n_w)$ such that $\left|v_{\widetilde{w}\widetilde{f}} - (\phi \cdot \overline{v})\right| < \xi$. Recall that $v_{\widetilde{w}\widetilde{f}}$ is distributed $\mathcal{H}$ with positive density in every point in the support $[\underline{v}, \overline{v}]$. Then,

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

and

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

It is only left to recall that when there is no firm $\widetilde{f} \in \widetilde{F}(n_w)$ for which $\left|v_{\widetilde{w}\widetilde{f}} - \phi \cdot \overline{v}\right| < \xi$, $v_{\widetilde{w}f}$ 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 ΔG let $\overline{b}^F(r, \Delta G)$ be the random vector of length r chosen as follows: [1] choose a network G according to ΔG, [2] choose a worker with degree r u.a.r. from all workers with degree r in G, and [3] let $\overline{b}^F(r, \Delta G)$ be the vector of the degrees of all firms in N_w ordered randomly (with equal probability given to each ordering). Let $\overline{b}^{0F}(r, \theta_F)$ be a vector of length r such that for every $i \in \{1, 2, ..., r\}$, $\overline{b}_i^{0F}$ equals r' with probability $\widetilde{P}^F(r', \theta_F)$ and such that $\left\{\overline{b}_i^{0F}\right\}_{i \in \{1,2,...,r\}}$ are determined independently of each other.

Lemma 3 *(Due to [15]) For all r and finite support θ_W, θ_F,*

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

Lemma 4 *Let θ_W, θ_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 $\psi_{G(n_w,\theta_W,\theta_F),M,\alpha}(\widehat{\gamma})$ be the r.v. that captures the fraction of workers hired at $\mathcal{S} = -1$ if all firms and workers best respond to the belief that $\gamma = \widehat{\gamma}$. Then, for every $\epsilon > 0$ and $\widehat{\gamma} \in [0,1]$*

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

Proof. Following condition (3), a worker w has a positive probability to be hired at $\mathcal{S} = -1$ 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 combining Lemma 3 and the limit result in Lemma 1-(1)-(b), and repeating the algebra at the bottom of section 11.2 we get that

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

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 $\xi > 0$

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

and therefore

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

which in turn, together with the independence of $\{v_{wf}\}_{f \in F}$ and Lemma 3 imply that as $n_w \to \infty$ 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

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

For a given graph G let $\widehat{w}_h(G)$ 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 $x_W(G) = \frac{\widehat{w}_h(G)}{n_w}$. 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

$$\sum_{r_w=1}^{\infty} P^{\widehat{W}}(r_w|G) \left(1 - \left[(1 - \tau_{G,M,\alpha}(\gamma)) + \tau_{G,M,\alpha}(\gamma) \cdot \left(1 - \mu_{G,M,\alpha}(\gamma)\right)\right]^{r_w}\right)$$

The proof is then completed by applying the strong law of large numbers twice to establish that
(1) $Pr\left(lim_{n_w \to \infty} x_W(G) = \frac{1}{2} \cdot (1 - \theta_W(0))\right) = 1$; and
(2) for any r, $Pr\left(lim_{n_w \to \infty} P^{\widehat{W}}(r|G(n_w,\theta_W,\theta_F)) = P(r,\theta_W)\right) = 1$. ■

Theorem 1 - Proof. By the convexity of the support for γ and the continuity of $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma)$ a fixed point exists. We now show that if $\gamma^* = \widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma^*)$ then γ^* 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 $n_w' > n_w$, $\Pr\left(\psi_{G(n_w,\theta_W,\theta_F),M,\alpha}(\gamma^*) \in [\gamma^* - \zeta, \gamma^* + \zeta]\right) > 1 - \zeta$.

Second, we show that for any $\varepsilon > 0$ there exists ζ such that for any $\zeta' < \zeta$, workers' and firms' best responses for γ^* satisfy the conditions for an ε-equilibrium for any $\gamma \in [\gamma^* - \zeta', \gamma^* + \zeta']$. This follows from Lemma 1, the continuity of $\lim_{n_w \to \infty} E_{G(n_w,\theta_W,\theta_F),M,\alpha}[\pi_f]$ in γ, and the independence of $\lim_{n_w \to \infty} E_{G(n_w,\theta_W,\theta_F),M,\alpha}[u_w]$ of γ. Since the payoffs of workers and firms are bounded, this completes the proof that if $\gamma^* = \widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}(\gamma^*)$ then γ^* is a 0-equilibrium with $\langle \theta_W, \theta_F, M, \alpha \rangle$.

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

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

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

Since for all $r \geq 1$, $P\left(r,\theta_W^1\right) = P\left(r,\theta_W^2\right)$ and $P\left(r,\theta_F^1\right) = P\left(r,\theta_F^2\right)$, 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 $r \geq 1$, $\widetilde{P}\left(r,\theta_F^1\right) = \widetilde{P}\left(r,\theta_F^2\right)$. Thus, for every $\gamma \in [0,1]$,

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

Then, by the definition of $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha}$ (expression 14), the difference in θ_W implies that for every $\gamma \in [0,1]$,

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

as required. ■

Proposition 2 - Proof. First, note that $\left\langle \theta_W^{\rho^L},\theta_F^{\rho^L},M,\alpha\right\rangle$ and $\left\langle \theta_W^{\rho^H},\theta_F^{\rho^H},M,\alpha\right\rangle$ are by definition regular environments, and note that for every $\langle \theta_W^{\rho},\theta_F^{\rho},M,\alpha\rangle$ and γ, $\frac{\partial\sigma_{\theta_W^{\rho},\theta_F^{\rho},M,\alpha}(\gamma)}{\partial\rho} = 0$ and $\frac{\partial\widetilde{\mu}_{\theta_W^{\rho},\theta_F^{\rho},M,\alpha}(\gamma)}{\partial\rho} = 0$. In regular environments, we can rely on Corollary 1 and prove the Proposition by showing that there exists $\overline{r},\underline{r} \in \mathbb{Z}^+$ such that if $max\left\{\rho^H\cdot r|\theta_F(r) > 0\right\} < \underline{r}$ then $\widetilde{\psi}_{\theta_W^{\rho^H},\theta_F^{\rho^H},M,\alpha}(\gamma) \geq \widetilde{\psi}_{\theta_W^{\rho^L},\theta_F^{\rho^L},M,\alpha}(\gamma)$, and if $min\left\{\rho^L\cdot r|r\geq 1,\theta_F(r) > 0\right\} > \overline{r}$, then $\widetilde{\psi}_{\theta_W^{\rho^H},\theta_F^{\rho^H},M,\alpha}(\gamma) \leq \widetilde{\psi}_{\theta_W^{\rho^L},\theta_F^{\rho^L},M,\alpha}(\gamma)$.

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

$$\begin{aligned}
\widetilde{\psi}_\rho &= \frac{1}{2}\cdot(1-\theta_W(0))\cdot\sum_{r_w=1}^{\infty}P(r_w,\theta_W)\cdot\left(1-\left[1-\widetilde{\tau}_\rho+\widetilde{\tau}_\rho\cdot(1-\widetilde{\mu})\right]^{\rho\cdot r_w}\right) &(22)\\
&= \frac{1}{2}\cdot(1-\theta_W(0)) - \frac{1}{2}\cdot(1-\theta_W(0))\cdot\sum_{r_w=1}^{\infty}P(r_w,\theta_W)\cdot\\
&\quad\cdot\left[1-x\cdot\left\{\sum_{r_f=1}^{\infty}\widetilde{P}(r_f,\theta_F)\cdot\left[\left(1-0.5^{\rho\cdot r_f}\right)/\left(0.5\cdot\rho\cdot r_f\right)\right]\right\}\right]^{\rho\cdot r_w} &(23)
\end{aligned}$$

and

$$\frac{\partial\widetilde{\psi}_\rho}{\partial\rho} = -\frac{1}{2}\cdot(1-\theta_W(0))\cdot\sum_{r_w=1}^{\infty}P(r_w,\theta_W)\cdot\frac{\partial\varphi_\rho(r_w)}{\partial\rho} \tag{24}$$

where

$$\varphi_\rho(r_w) = \left[1-x\cdot\left\{\sum_{r_f=1}^{\infty}\widetilde{P}(r_f,\theta_F)\cdot\left[\left(1-0.5^{\rho\cdot r_f}\right)/\left(0.5\cdot\rho\cdot r_f\right)\right]\right\}\right]^{\rho\cdot r_w} \tag{25}$$

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

1. $\varphi_\rho(r_w)$ and $\varphi_\rho(r_w, r_f)$ are twice differentiable.

2. If $\frac{\partial \varphi_\rho(r_w,r)}{\partial \rho} \geq 0$ $\left(\frac{\partial \varphi_\rho(r_w,r)}{\partial \rho} < 0\right)$ for every $r \in \mathbb{R}^+$ in the convex hull of $\{r_f | \theta_F(r_f) > 0\}$, then $\frac{\partial \varphi_\rho(r_w)}{\partial \rho} \geq 0$ $\left(\frac{\partial \varphi_\rho(r_w)}{\partial \rho} < 0\right)$.

3. $sign\left\{\frac{\partial \varphi_\rho(r_w,r_f)}{\partial \rho}\right\} = sign\left\{\frac{\partial \ln(\varphi_\rho(r_w,r_f))}{\partial \rho}\right\}$.

Thus to prove Proposition 2 it is sufficient to show that:

- **Step 1:** $\mathcal{H}[\alpha \cdot \phi \cdot \overline{v}] > \frac{1}{3}$ implies that $\varphi_2(r_w, 1) \leq \varphi_1(r_w, 1)$.
- **Step 2:** there exists $\overline{\rho}(M, \alpha) \in \mathbb{Z}^+$ such that for all $\rho \geq \overline{\rho}(M, \alpha)$ and any $r_w > 0$, $\frac{\partial \ln(\varphi_\rho(r_w,1))}{\partial \rho} \geq 0$.

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

$$\frac{\varphi_2(r_w, 1)}{\varphi_1(r_w, 1)} = \frac{[1 - 0.75 \cdot x]^{2 \cdot r_w}}{[1 - x]^{r_w}} = \left(\frac{[1 - 0.75 \cdot x]^2}{1 - x}\right)^{r_w} \tag{26}$$

To conclude that $\frac{\varphi_2(r_w,1)}{\varphi_1(r_w,1)} \leq 1$ as required, we note that if $\mathcal{H}[\alpha \cdot \phi \cdot \overline{v}] > \frac{1}{3}$, then $\widetilde{\mu} < \frac{2}{3}$ and $\sigma \leq 1$, so that $x < \frac{2}{3}$. Therefore,

$$\frac{\partial \frac{[1-0.75 \cdot x]^2}{1-x}}{\partial x} = \frac{-0.5 + 0.75 \cdot x}{(1 - x)^2} \cdot (1 - 0.75 \cdot x) < 0 \tag{27}$$

and

$$\left(\frac{\varphi_2(r_w, 1)}{\varphi_1(r_w, 1)} | x = 0\right) = 1 \tag{28}$$

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

$$\ln(\varphi_\rho(r_w, 1)) = \rho \cdot r_w \cdot \ln[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)] \tag{29}$$

which with some algebra amounts to

$$\frac{\partial \ln(\varphi_\rho(r_w, 1))}{\partial \rho} = r_w \cdot \ln[1 - x \cdot (1 - 0.5^{\rho \cdot r}) / (0.5 \cdot \rho \cdot r)] + \tag{30}$$

$$+ \frac{[x \cdot \rho \cdot r_w \cdot \ln(0.5) \cdot 0.5^{\rho \cdot r} \cdot r + x \cdot r_w - x \cdot r_w \cdot 0.5^{\rho \cdot r}]}{0.5 \cdot \rho \cdot r - x + x \cdot 0.5^{\rho \cdot r}} \tag{31}$$

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

$$\ln[1 - x \cdot [(1 - 0.5^\rho) / (0.5 \cdot \rho)]] + \frac{[x \cdot \rho \cdot \ln(0.5) \cdot 0.5^\rho + x - x \cdot 0.5^\rho]}{0.5 \cdot \rho - x + x \cdot 0.5^\rho} \geq 0 \tag{32}$$

In Step 2, we are interested in the sign of $\frac{\partial \ln\left(\varphi_\rho(r_w,1)\right)}{\partial \rho}$ for large ρ. For any x and for any $\rho \geq 2$, $0.5 \cdot \rho - x + x \cdot 0.5^\rho > 0$, so inequality (32) holds if and only if

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

With some additional algebra we get that inequality (32) holds if and only if

$$\frac{2^\rho\{(0.5\cdot\rho - x)\ln[1 - x\cdot(1-0.5^\rho)/(0.5\cdot\rho)] + x\}}{\rho} \geq \frac{x\cdot\ln[1-x\cdot(1-0.5^\rho)/(0.5\cdot\rho)] + x\cdot\rho\cdot\ln(0.5) - x}{\rho} \quad (34)$$

where

$$lim_{\rho\to\infty}\frac{x \cdot \ln\left[1 - x \cdot (1 - 0.5^\rho) / (0.5 \cdot \rho)\right] + x \cdot \rho \cdot \ln(0.5) - x}{\rho} = x \cdot \ln(0.5) < \infty \quad (35)$$

Consequently, a sufficient condition for inequality (32) to hold is

$$lim_{\rho\to\infty}\frac{2^\rho\left\{(0.5 \cdot \rho - x)\ln\left[1 - x \cdot \left[(1 - 0.5^\rho) / (0.5 \cdot \rho)\right]\right] + x\right\}}{\rho} = \infty \quad (36)$$

Also,

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

and

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

As a result, condition (36) is satisfied for any $x > 0$, and inequality (32) holds for any $x \geq 0$ as required. ■

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

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 $\gamma \in [0,1]$,

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

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

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

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

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

which completes the proof. ■

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

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

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

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

which completes the proof. ■

Proposition 5 - Proof. We rely on Corollary 1 and prove the Proposition by showing that: [1] if $\alpha^2\leq\frac{\underline{v}}{\phi_M\cdot\overline{v}}$ then for every $\gamma\in[0,1]$, $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha^2}\left(\gamma\right)\geq\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha^1}\left(\gamma\right)$, and [2] if $\alpha^1\geq\frac{\overline{c}+\frac{\frac{1}{2}}{\frac{\sum_{r=0}^{\infty}\theta_W(r)\cdot r}{\sum_{r=0}^{\infty}\theta_F(r)\cdot r}-\gamma}\cdot\pi_H-\pi_L}{\pi_H-\pi_L+\frac{\gamma}{\frac{\sum_{r=0}^{\infty}\theta_W(r)\cdot r}{\sum_{r=0}^{\infty}\theta_F(r)\cdot r}-\gamma}\cdot\pi_H}$ for all $\gamma\in[0,1]$, then for every $\gamma\in[0,1]$, $\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha^1}\left(\gamma\right)\geq\widetilde{\psi}_{\theta_W,\theta_F,M,\alpha^2}\left(\gamma\right)$.

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

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

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

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

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

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

12 References

[1] A. Abdulkadiroglu, T. Sonmez, Random serial dictatorship and the core from random endowments in house allocation problems, Econometrica 66 (1998), 689-701.

[2] A. Abdulkadiroglu, T. Sonmez, House allocation with existing tenants, J. Econ. Theory 88 (1999), 233–260.

[3] C. Avery, C. Jolls, R.A. Posner, A.E. Roth, The Market for Federal Judicial Law Clerks, The Univ. of Chicago Law Review 68 (2001), 793-902.

[4] C. Avery, C. Jolls, R.A. Posner, A.E. Roth, The New Market for Federal Judicial Law Clerks, The Univ. of Chicago Law Review 74 (2007), 447-486.

[5] E. Azevedo, J.D. Leshno, A Supply and Demand Framework for Two-Sided Matching Markets, working paper, 2012.

[6] C. Ballester, A. Calvó-Armengol, Y. Zenou, Who's Who in Networks. Wanted: The Key Player, Econometrica 74 (2006), 1403-1417.

[7] T.F. Bewley, Why Wages Don't Fall During a Recession, Harvard Univ. Press, Cambridge, MA, 1999.

[8] Y. Bramoullé, M. D'Amours, R. Kranton, Strategic Interaction and Networks, working paper, 2010.

[9] J. Bulow, J. Levin, Matching and Price Competition, Amer. Econ. Rev. 96 (2006), 652-668.

[10] A. Calvo-Armengol, M.O. Jackson, The Effects of Social Networks on Employment and Inequality, Amer. Econ. Rev. 94 (2004), 426-454.

[11] A. Calvo-Armengol, Y. Zenou, Job matching, social network and word-of-mouth communication, J. Urban Econ. 57 (2005), 500-522.

[12] V.P. Crawford, The Flexible-Salary Match: A Proposal to Increase the Salary Flexibility of the National Resident Matching Program, J. Econ. Behav. Organ. 66 (2008), 149-160.

[13] E. Damiano, H. Li, W. Suen, Unravelling of Dynamic Sorting, Rev. Econ. Stud. 72 (2005), 1057-1078.

[14] I.P. Fainmesser, Community Structure and Market Outcomes: a Repeated Games in Networks Approach, American Economic Journal: Microeconomics 4 (2012), 32-69.

[15] I.P. Fainmesser, D.A. Goldberg, Cooperation in Partly Observable Networked Markets, working paper, 2012.

[16] G.R. Fréchette, A.E. Roth, U.M. Ünver, Unraveling yields inefficient matchings: evidence from post-season college football bowls, RAND J. Econ. 38 (2007), 967–982.

[17] D. Gale, L. Shapley, College Admissions and the Stability of Marriage, Amer. Mathematical Monthly 69 (1962), 9-15.

[18] A. Galeotti, S. Goyal, M.O. Jackson, F. Vega-Redondo, L. Yariv, Network Games, Rev. Econ. Stud. 77 (2010), 218-244.

[19] M.S. Granovetter, Getting A Job: A Study of Contacts and Careers, Harvard Univ. Press, Cambridge, MA, 1974.

[20] C. Greenhill, B.D. McKay, X. Wang, Asymptotic enumeration of sparse 0–1 matrices with irregular row and column sums, J. Combinatorial Theory, Series A 113 (2006), 291-324.

[21] H.W. Halaburda, Unravelling in Two-Sided Matching Markets and Similarity of Preferences, Games Econ. Behav. 69 (2010), 365-393.

[22] M.O. Jackson, B.W. Rogers, Meeting Strangers and Friends of Friends: How Random are Socially Generated Networks? Amer. Econ. Rev. 97 (2007), 890-915.

[23] J.H. Kagel, A.E. Roth, The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment, Quart. J. Econ. 115 (2000), 201-235.

[24] F. Kojima, Matching and Price Competition: Comment, Amer. Econ. Rev. 97 (2007), 1027-1031.

[25] T.C. Koopmans, M. Beckmann, Assignment Problems and the Location of Economic Activities, Econometrica 25 (1957), 53-76.

[26] A. Kozinski, Confessions of a Bad Apple, The Yale Law Journal, 100 (1991), 1707-1730.

[27] R. Lee, M. Schwarz, Interviewing in Two-Sided Matching Markets, working paper, 2012.

[28] H. Li, S. Rosen, Unraveling in Matching Markets, Amer. Econ. Rev. 88 (1998), 371-87.

[29] H. Li, W. Suen, Risk Sharing, Sorting, and Early Contracting, J. Polit. Economy 108 (2000), 1058-1087.

[30] J.D. Montgomery, Social Networks and Labor-Market Outcomes: Toward an Economic Analysis, Amer. Econ. Rev. 81 (1991), 1407-18.

[31] M. Niederle, Competitive Wages in a Match with Ordered Contracts, Amer. Econ. Rev. 97 (2007), 1957-1969.

[32] M. Niederle, L. Yariv, Decentralized Matching with Aligned Preferences, working paper, 2009.

[33] M. Niederle, A.E. Roth, Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and without a Centralized Match, J. Polit. Economy 111 (2003), 1342-1352.

[34] M. Niederle, A.E. Roth, Relationship Between Wages and Presence of a Match in Medical Fellowships, Journal of the American Medical Association 290 (2003), 1153-1154.

[35] M. Niederle, A.E. Roth, The Gastroenterology Fellowship Market: Should There Be a Match? Amer. Econ. Rev. 95 (2005), 372-375.

[36] M. Niederle, A.E. Roth, Market Culture: How Rules Governing Exploding Offers Affect Market Performance, American Economic Journal: Microeconomics, 1 (2009), 199-219.

[37] M. Ostrovsky, M. Schwarz, Information Disclosure and Unraveling in Matching Markets, American Economic Journal: Microeconomics, 2 (2010), 34-63.

[38] A.E. Roth, The Economist as Engineer: Game Theory, Experimental Economics and Computation as Tools of Design Economics, Econometrica 70 (2002), 1341-1378.

[39] A.E. Roth, What have we learned from market design?, Econ. J. 118 (2008), 285-310.

[40] A.E. Roth, M. Sotomayor, Two-Sided Matching, Cambridge Univ. Press, Cambridge, UK, 1990.

[41] A.E. Roth, X. Xing, Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions, Amer. Econ. Rev. 84 (1994), 992-1044.

[42] L. Shapley, M. Shubik, The Assignment Game I: The Core, Int. J. Game Theory 1 (1971), 111-130.

[43] W. Suen, A Competitive Theory of Equilibrium and Disequilibrium Unravelling in Two-Sided Matching, RAND J. Econ. 31 (2000), 101-120.